

# UC San Diego

## UC San Diego Previously Published Works

### Title

Computational challenges of evolving the language-ready brain

### Permalink

<https://escholarship.org/uc/item/9t88r1t6>

### Author

Arbib, Michael A

### Publication Date

2020-08-11

### DOI

10.1075/bct.112.03arb

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

## Computational challenges of evolving the language-ready brain

### 1. From manual action to protosign

Michael A. Arbib

University of California at San Diego

Computational modeling of the macaque brain grounds hypotheses on the brain of LCA-m (the last common ancestor of monkey and human). Elaborations thereof provide a brain model for LCA-c (c for chimpanzee). The Mirror System Hypothesis charts further steps via imitation and pantomime to protosign and protolanguage on the path to a "language-ready brain" in *Homo sapiens*, with the path to speech being indirect. The material poses new challenges for both experimentation and modeling.

**Keywords:** action pattern reorganization, computational comparative neuroprimatology, imitation, language evolution, mirror system hypothesis, mirror systems, modeling cerebellum, speech evolution

#### 1. The Mirror System Hypothesis (MSH) introduced

To have language is, in part, to be able to exploit an open lexicon and a powerful grammar to communicate and comprehend new meanings of increasing complexity. Since humans can learn a signed language as readily as a spoken language if raised in the appropriate milieu, I stress that "having language" is not synonymous with "having speech." This paper espouses an EvoDevoSocio approach to language evolution, positing that what evolved (*Evo*) was a language-*ready* brain – not a brain with an innate mechanism encoding a universal grammar (Arbib, 2007) but rather one enabling a child to acquire language (*Devo*), but only if raised in a milieu in which language is already present, something which, it is claimed, required tens of millennia of cultural evolution after the emergence of *Homo sapiens* (*Socio*). Biology and culture shape each other, with the structure and function of an adult's brain reflecting the social and physical interactions in which that person has engaged throughout a lifetime; and those interactions may in turn shape the culture.

To succeed, language must have the *parity property* that the meaning of an utterance intended by the speaker or signer will be understood (though not always) by the recipient. The F5 premotor brain region in macaque has mirror neurons which appear to support a parity property for manual action – specific neurons fire both when the monkey performs specific actions and when it observes similar actions performed by another. Several data suggested the relevance of mirror neurons to language:

- Macaque F5 is homologous to Area 44 in human Broca’s area, a crucial component of the human brain’s language system.
- Human brain imaging showed activation for both grasping and observation of grasping in or near Broca’s area
- Broca’s area plays a similar role for spoken and signed languages (Poizner et al., 1987).

Such observations motivated *The Mirror System Hypothesis (MSH)* (Arbib & Rizzolatti, 1997): the evolutionary basis for *language parity* is provided by the mirror system for grasping, rooting speech in communication based on production and recognition of manual gestures. In more detail, we hypothesize that evolutionary elaboration of mechanisms for execution, recognition and imitation of manual skills in time supported manual gesture, ad hoc pantomime and, eventually, proto-sign to lay the scaffolding for protospeech and the emergence of the language-ready brain. For MSH, the capacity for vocal learning and control that distinguishes humans from other primates still plays a role, but it is secondary (but see Section 7).

Much subsequent work has been devoted to “evolving” MSH (Arbib, 2016 for a recent overview). MSH hypothesizes mechanisms for behavior and communication in LCA-m and LCA-c (last common ancestor of humans with macaques and chimpanzees, respectively). These ground (and can be modified by) hypotheses on the path from LCA-m via LCA-c to *Homo sapiens*. Since data on LCA-m and LCA-c are virtually non-existent, we resort to *comparative (neuro) primatology* – comparing brain, behavior and communication across extant species (macaques and other monkeys, chimpanzees and other apes, and modern humans) and their precursors as revealed by archeological data.

## 2. Introducing *computational comparative neuroprimatology*

This article has “Computational Challenges” in the title because too many experimentalists and field workers fail to think through the processing challenges of the behaviors or brain scans they observe. Mathematical or computational models probe in more detail a methodology that will advance the field by asking for each study what is the overall behavior of interest and then seeking to develop a

hypothesis (call it H1) on the interacting processes (e.g., neural networks or schemas/schema instances) and the information flow between them that might support the behavior. H1 may be consistent with the data under review, but if not – or important new data emerge that H1 does not explain – then one must develop H2 with schemas and interactions that better match overall performance or brain activity ... and/or (crucial for us) are more conformable with evolutionary hypotheses based on, e.g., comparative neuroprimatology. And thus the experiment (or field work) – modeling cycle continues.

We advocate a *computational* comparative neuroprimatology. We need to understand what computations occur within each brain region and how the interactions between brain regions orchestrate them to yield behavior, and how interactions between two or more agents underpin language acquisition and cultural change. However, whereas for monkeys we have many single-cell recordings that constrain models of biologically structured neural networks, no such data exist for apes or humans. This has led to the following strategy:

1. For monkeys, build detailed models of interacting neural networks and use computer simulation to explore the relation between overt behavior and neural activity (Section 3).
2. Compare behavior in macaques and apes to chart key changes in the behavioral repertoire, and posit changes in an LCA-m model that can explain the extended behavior, offering this as a hypothesis for LCA-c (Section 4).
3. For humans, we may repeat the process to offer models of neural circuitry averaged over the activity of various simulated circuits to make predictions to be tested against ERP and fMRI (Arbib et al., 2000; Barrès et al., 2013). Alternatively, modeling may start from the human data (which may also include lesion data, or data on neurological disorders) and seek to build models directly either at the level of neural networks or interacting schemas (see, e.g., Arbib, 2016; Arbib, 2018; and Cooper, 2016).

This paper focuses on “monkey-based” brain modeling on the evolutionary path to protolanguage. A different style is required to trace the path from protolanguage to neurolinguistics (Part 2, Arbib, 2018). Meanwhile, many studies ignore the brain to conduct simulations of how various aspects of language might emerge in a computational milieu (Cangelosi & Parisi, 2002). We have only the beginnings of an integrated framework for EvoDevoSocio modeling that links brain, behavior and communication.

### 3. Setting a baseline for LCA-m

In this section, we introduce three models of the macaque brain, while leaving details to the original papers. Extended review of alternative models would be valuable *especially if we could evaluate their relevance to our evolutionary investigation*, but is outside the scope of this short paper.

#### 3.1 The FARS (Fagg-Arbib-Rizzolatti-Sakata) model

The FARS Model (Fagg & Arbib, 1998) explains how the brain may use visual information to guide the hand in grasping an object. It was based in part on macaque neurophysiology on neural correlates of “motor schemas for manual actions” in premotor area F5 (e.g., Rizzolatti et al., 1988) and “grasp affordances” (e.g., Taira et al., 1990) – visual cues as to graspability – in parietal area AIP (anterior intraparietal sulcus). Visual input travels by two pathways:

1. A dorsal (“how”) path via AIP extracts information on affordances to yield parameters for detailed motor control of each action
2. A ventral (“what”) path wherein object recognition provides input that the prefrontal cortex can combine with working memory to plan a sequence of actions while the dorsal path routes the appropriate affordances to the motor cortex to control the current action.

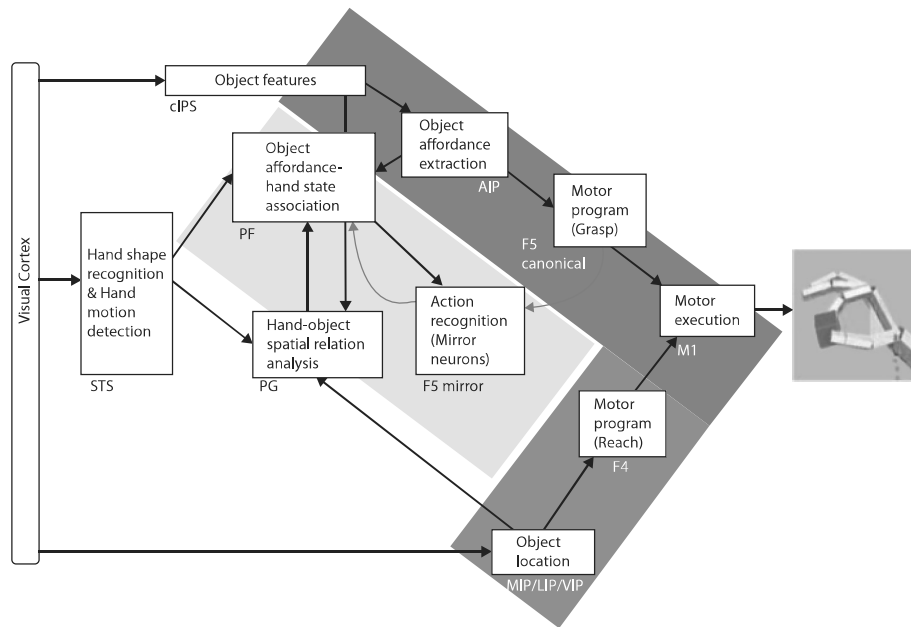
Notably, the data at that time focused on conditions where the monkey had only one affordance available for the given trial; the model further addressed having to make decisions between multiple affordances. This dorsal-ventral distinction plays a crucial role in charting the path to language processing in Part 2.

#### 3.2 Modeling mirror systems in action recognition

The Mirror Neuron System model (MNS; Oztop & Arbib, 2002) offers a *Devo* view of mirror neurons. Rather than positing an innate repertoire, it suggests how *mirror neurons for manual actions might emerge during observation of one’s own actions*. In Figure 1, the external diagonals correspond to the dorsal path of the FARS model for converting an affordance into a grasp and a complementary path for controlling the arm to bring the hand to the desired position. Since we emphasize learning, we distinguish “potential” mirror neurons (before learning) from actual mirror neurons (after their properties are defined by the learning process). These receive both (i) Efferent copy of the code for some of grasps and (ii) input from circuitry that monitors the trajectory of the hand in a reference frame centered on the chosen affordance. The efferent copy acts as a “training signal” for the neurons

it activates – the learning process strengthens synapses that encode trajectories like those for the current grasp. As learning progresses, the synaptic drive from (ii) will eventually be enough to activate the emerging mirror neurons relevant to that grasp even if input (i) is absent. Since observation of another individual’s action may evoke the same affordance-centered input pattern (ii) as for self-execution, these neurons thus become mirror neurons. MNS demonstrated how, as learning progresses, recognition of the grasp may occur earlier and earlier in the trajectory – though such anticipation will be a function of how precisely the trajectory is represented in the brain, which in turn is a function of attention as well as neural encoding.

A crucial aspect of the model, then, was to suggest that mirror neurons may have evolved first to monitor self-actions (see the ACQ model next) – matching intended action to observed trajectory – with their role in the observation of others (which is most emphasized in the literature) being an exaptation of this capability. Our 2002 hypothesis, that F5 mirror neurons of the macaque are sensitive to the sight of the monkey’s own hand during object grasping, was confirmed by Maranesi et al. (2015).



**Figure 1.** The MNS Model of Learning in the Mirror Neuron System. Note that, whatever properties mirror neurons have, they only have by dint of being part of a larger system. See text for details

### 3.3 Flexible action patterns and their rapid reorganization

Alstermark et al. (1981) demonstrated lesions of axons leaving the spinal cord that impair grasping but not reaching in cats. He taught cats to reach into a glass tube projecting horizontally from the wall and grasp a piece of food, which the cat then brings to its mouth. After just a few trials, a lesioned cat would not try to grasp but would simply bat the food from the tube and then grasp it from the floor with its jaws. How could this new skill emerge so quickly? The Augmented Competitive Queueing Model (ACQ; Bonaiuto & Arbib, 2010) explains such phenomena (and we argue that it applies to the monkey and LCA-m, even though it was inspired by cat data) by *having the mirror system monitor self-actions*, as emphasized in the previous section.

Recall that mirror neurons can be activated both by efference copy of a motor command or by observing a hand-to-object trajectory associated with the grasp. The key to ACQ is that when an *intended action* is unsuccessful, it may appear similar to an unintended action – and then the mirror neurons for the *apparent action* can serve a “what did I just do?” function. Thus, when the lesioned cat tries to grasp the food and inadvertently knocks it out of the tube, the mirror system can recognize that it looks like a “batting” action already in the cat’s repertoire. ACQ makes two evaluations of for each action:

- Desirability* depends on the current task or goal. Each time the action is performed, a measure of “expected reinforcement” is updated. This will be positive if the action leads “soon enough” to achievement of the goal, but will be greater the shorter the time required to reach that goal.
- Executability* depends on the availability of affordances (can the action be carried out now?) and the probability of the action’s success.

At each time step, the priority of available actions is set by combining executability and desirability – the highest priority action will then be executed (or, since failure is possible, its execution will be attempted). Each time an action is performed *successfully*, its desirability is updated while executability may be left as is or increased. However, when the action is unsuccessful, executability of the *intended* action is reduced while desirability of the *apparent* action is adjusted.

This explains the rapid change of behavior in Alstermark’s lesioned cat. Since the grasp keeps failing, its executability is decreased (but its desirability is unchanged) whereas the desirability of batting increases each time it is used. Consequently, in only a few trials the priority of batting comes to exceed that of grasping, and the cat has a new plan of behavior implicit in altered desirability and executability of its actions.

The model assumes that cats have mirror neurons for brachio-manual actions. This has yet to be tested. However, the suggestion is again that mirror neurons arose first for monitoring of self-actions and that this functionality is widespread.

#### 4. An LCA-c innovation built on LCA-m mechanisms

MSH does not claim that having mirror neurons for hand movements suffices for language. Rather, it claims that the ability to recognize dexterous manual actions provided a stepping stone for LCA-c to develop novel communicative gestures, and that further steps were needed en route to the language-ready brain. Consider ape gesture as a stand-in for LCA-c. Where some have argued that all ape gestures are simply extracted from an innate repertoire (Hobaiter & Byrne, 2011), others have focused on specific gestures observed in ape populations and suggested a role for social learning. Tomasello & Call (1997) proposed *ontogenetic ritualization* (OR) as a means whereby (some) ape gestures could emerge:

- i. A performs praxic behavior X and individual B consistently reacts by doing Y
- ii. Subsequently, B anticipates A's overall performance of X by starting to perform Y before A completes X.
- iii. Eventually, A anticipates B's anticipation, producing a ritualized form  $X^R$  of X to elicit Y

Halina et al. (2013) offer examples. Hobaiter & Byrne suggest that if a gesture is in frequent use in an ape group it must be innate, and that OR can only generate idiosyncratic gestures. However, the model below suggests that if a dyadic behavior is common for whatever reason, then one may expect its ritualization to be common, too.

Arbib et al. (2014) developed a model based on FARS and MNS. It introduces *dyadic brain modeling* – we simultaneously model the brains of two interacting apes. The architecture of each brain is the same but the initial states are different and thus the learning differs in the brains.

The emergence of beckoning provides an example. The child's distal goal is getting mother to hug him. In the initial episodes, the child's motivation is to be hugged while the mother's motivation is elsewhere until the child tugs on her arm and pulls her closer. She recognizes this and responds with a hug. Over subsequent episodes, two different mechanisms come into play. Thanks to the ability of the *mother's* MNS to learn to recognize an action earlier and earlier in its trajectory, she comes to recognize the child's request, and thus respond, before its completion. Similarly, the *child's* MNS allows the child to recognize earlier that his request is being granted, and to terminate his trajectory accordingly. The model shows how the child makes the transition from intending the full request and stopping



it, to simply making the initial prefix of that request as the intended action. This transition is from a *transitive* action (whose goal is set by the affordances of an object – or mother’s arm) to an *intransitive* action. Why do apes but not monkeys make this transition? Arbib et al. hypothesize that the ape can make greater use of proprioceptive information in setting a goal than the monkey can, so that an intermediate position of the arm can be recognized as a desired end state.

But if this verbal explanation is convincing, why specify the details to the point where the model can be simulated on a computer? One answer is that once implemented, we can vary parameters within the model on different simulation runs to establish parameter ranges in which the model does or does not yield the behavior that the verbal discussion rendered so plausible. Similarly, we can see the effects of adding or removing pathways between regions of the model. Hence, if we find that some settings yield behavior more akin to those of monkeys while others yield behavior more typical of apes, then we have a prediction in precise form for what might be a critical evolutionary change in brain organization.

Alas, empirical tools to establish parameter values are lacking. However, Diffusion Tensor Imaging (DTI) by Hecht et al. (2012) does speak to differences in MNS connectivity in macaque, chimpanzee and human. Among the trends seen are changes in the ventral vs. dorsal visual routes to frontal cortex which may support increased processing of visual movement details, and changes in connectivity between the parietal mirror region and inferotemporal cortex which may better support social learning of object-related actions. Of course, such connectivity data must be complemented by data on the neural circuits that are connected.

## 5. Varieties of imitation

Imitation comes in many forms (Byrne & Russon, 1998). The general definition of imitation relevant here is “the ability to use observation of others achieving a desirable goal to develop a means of achieving that goal based on the method exhibited by the performer.” Byrne (2003) argued that apes acquire new skills through “imitation as behaviour parsing,” in which the observer comes to recognize that a few subgoals are key to successful performance – but then acquires the action to get from one subgoal to the next through a lengthy process of trial and error. This is *simple imitation* that, MSH claims, was present in LCA-c, but not LCA-m. MSH then claims that our ancestors post LCA-c acquired a crucial blend of *complex action recognition* as well as *imitation* – this still “parses” the behavior but adds attention to the motion as well as the goal of subactions, with the consequent ability to achieve a first approximation to that motion without trial and error. A repertoire of “tweaks” may be used help adjust a motion, but with trial and error still available to hone a moderately successful skill with repeated practice.

MSH posits that, like various monkey species, LCA-m had an innate vocal call system, but lacked the ability – posited for LCA-c – to acquire novel manual gestures. Section 4 showed how LCA-c brains might support OR. Strikingly, though, OR does not involve imitation. Evidence remains sparse for transmission of gesture through imitation, but I predict that evidence will eventually be found.

Space does not permit an extended treatment of imitation models. Instead, a few observations. A common mistake is to think that having mirror neurons (or action recognition more broadly) is enough to be able to imitate. This is not so (Oztop et al., 2006; Oztop et al., 2013). Here are two issues:

1. If the action required to achieve a subgoal is already in the observer’s repertoire and can be recognized as such *and if the recognition can be used to guide action*, then imitation may proceed quite swiftly. Otherwise some means (trial-and-error or not) must be found to acquire that action. Here we see the need for a “reverse MNS”: Recognizing that the action being performed by another reaches a desirable goal, learn features that enable you to recognize it, and (if feasible) use those features to aid you in adding the action to your own repertoire.
2. Being able to acquire a single action does not guarantee one can develop a “program” and working memory that can link various actions to the subgoals and keep track of what subgoals have been achieved in the current behavior. Here a “reverse ACQ” might be relevant, coupling systems in cerebral cortex and basal ganglia.

Thus, human evolution may have complemented improved skills in mastering novel actions with an increased capacity to master hierarchical plans of increasing complexity.

Note the utility of complex imitation for language *learning* and complex action recognition for language *use* once this can be applied to words-as-articulatory actions – a late exaptation of a system that evolved (MSH claims) to support praxis.

Animal behaviors may be highly complex – for example, novel spatial arrangements of prey, predator and barriers as well as motivational state can yield an endless variety of trajectories in frogs (Cobas & Arbib, 1992). However, this seems qualitatively different from the on-line flexibility of conversation, where each utterance may (even though it often does not) express novel meanings.

## 6. From imitation to pantomime

We have seen that MSH claims that, post-LCA-c, intertwining changes in brain and body and in social interaction yielded a capacity for *complex action recognition and imitation* (CAR&IM), driven primarily by adaptive pressure for increased

efficacy of transfer of manual skills. But it claims that this paved the way for a new form of communication, *ad hoc pantomime*. Rejecting a previous definition (Arbib, 2012, pp. 218–219), I suggest the following:

A social group “has” pantomime if it has both the brain capability *and the social conventions* such that dyads (X,Y) of the community can *freely* engage in the following sort of exchange:

- X performs an intransitive action P that resembles an action B which might occur within a context C to achieve goal G – and does so with the intention that observer Y will “get the message” concerning some aspect of C or G;
- Y recognizes that A does indeed resemble B and, knowing that action B might occur within a context C’ with goal G’, infers that the message is some aspect of C or G.

Crucially, MSH requires that these pantomimes can be freely invented to bridge gaps in communication when previously available means fail and hypothesizes that such pantomime appeared post-CAR&IM (and thus post-LCA-c, but see Russon, 2018). The catch is that pantomime may be unsuccessful, or succeed only after much further effort. This, MSH claims, provided the selective pressure to yield social and biological evolution that yielded *protosign* in which pantomimes are ritualized in a community to provide low-energy gestures with reduced ambiguity. Indeed, variants may arise to distinguish key interpretations – e.g., a pantomime for “bird flying” might differentiate into protosigns for “bird” and for “flying.”

Is a pantomime-ready brain also a protosign-ready brain? Perhaps not. There are brain lesions in users of modern sign languages that impair language use while leaving intact the capacity for pantomime (Corina et al., 1992; Marshall et al., 2004) – suggesting that ad hoc use of pantomime is neurally different from access to a symbol within a (proto)sign system.

MSH then argues that even a limited use of protosign for communication creates an adaptive pressure for the emergence of a capacity to use gestures in other modalities, and that *protospeech* emerges through invasion of the vocal apparatus by collaterals from the protosign system. (See Arbib, 2012 for some details.) Capabilities for protosign and protospeech then emerge in an expanding spiral (Arbib, 2005): the path to speech is indirect. (Or is it? See Section 7 and Aboitiz, 2018.)

The final MSH claim is that once the capacities for complex imitation and protolanguage were in place in early *Homo sapiens*, the emergence of language – an open lexicon, a grammar supporting a rich compositional semantics, and a phonology – were all primarily the result of social innovation and dissemination (Arbib, 2012, Chapter 10). Nonetheless, Baldwinian evolution may have “tweaked” the biological substrate for acquiring or using these language features, such as

increasing control over vocal articulators and increasing capacity for symbolic working memory. Brain mechanisms for the production and comprehension of language utterances will be a central concern of Part 2 (Arbib, 2018).

## 7. Is the path to speech indirect?

### 7.1 Some macaque premotor neurons may control vocalization

Whereas the involvement of medial cortex in monkeys in the conditioning of innate calls is widely accepted (Jürgens, 2002), Coudé et al. (2011) found neurons in the ventral premotor cortex that activate during the conditioned vocalizations they studied. Fogassi et al. (2013) suggest that these neurons constitute a primitive neural substrate of a cortical center for the voluntary control of vocalization. But does this promote the hypothesis of a direct route from LCA-m vocalization to speech in which non-homologous regions are implicated, or demonstrate a restricted path on which evolution post-LCA-c could enable protosign mechanisms to scaffold protospeech? Note that orofacial control does not imply vocal control.

In support of the latter view: (a) nonhuman primates can master novel manual skills but not novel vocal skills; and (b) pantomime offers relatively direct access to a wider range of meanings than does sound symbolism, thus offering a clearer path to an open semantics via protosign than directly through protospeech.

### 7.2 Case study: The role of the cerebellum in prism adaptation

Rather than offer a model directly bearing on the path (direct or indirect) to speech, I want to turn to a model of the role of the cerebellum in prism adaptation (Arbib et al., 1995). The point made here is that computational neuroscience can contribute interpretive tools of general utility, and that lessons learned from modeling in one domain may in due course illuminate another. Our challenge was to develop a model of the role of cerebellum and related brain structures that could explain the data of Martin et al. (1996) on adaptation of throwing to a target while wearing prisms that shift the visual input laterally. The data contained two surprises.

- a. In many cases, someone who had adapted to wearing the prisms during repeated throws underarm showed little or no adaptation when, with prisms still on, she started throwing overarm. We explained this by linking different *microcomplexes* – each a patch of cerebellar cortex linked to a patch of cerebellar nuclei and circuitry in cerebral cortex – to different types of throwing. We showed that the degree of overlap between microcomplexes for underarm and overarm throws could explain the degree of transfer between prism adaptation for the different throws.

- b. After hundreds of blocks of trials, each involving adaptation and readaptation to the prisms, the (very dedicated) subject eventually reached a stage at which no adaptation was required when the prisms were donned or doffed. The basic model rested on the fact that hand areas in cerebellum and cortex are richly endowed with fibers encoding eye position, whereas there is no reason for evolution to have favored fibers encoding *prism on/off*. Thus, to complete our model, we hypothesized that a neutral mix of fibers from cerebral cortex was available to the relevant cerebellar microcomplexes, and thus a very sparse subset could convey features that might correlate with prism on/off even though neither evolution nor experience had previously selected for them. The model worked as follows: Because many fibers encoded eye position, learning could rapidly adjust enough synapses to adaptively change cerebellar modulation of the arm-throw circuitry. However, because the prism on/off-related fibers were so sparse, the chance of their being modified adaptively was very small, and thus the number of trials for their adaptation to become effective was very large.

The suggestion, then, is that the drive from ventral premotor cortex to vocal control discovered by Coudé et al. is akin to the prism on/off fibers of the model – evidence of sparse random connections rather than an evolved capability. Nonetheless, for creatures for whom greater vocal control became adaptive, this random group could become the target of Baldwinian evolution to foster the emergence of protospeech on the scaffolding of protosign. (This suggestion offers a new challenge for MSH modeling.)

## 8. Towards a new road map

The main focus of this paper has been to show how the analysis of macaque brain models can deepen the understanding of the path from LCA-m via LCA-c to *H. sapiens* as hypothesized by an “old” road map, that of MSH (Arbib, 2012; 2016). Nonetheless, it has touched on several issues beyond MSH. Here is a slightly augmented list:

1. Further neurophysiology is required to assess the prevalence of mirror neurons for different classes of actions in different species.
2. MSH is based on the hypothesis that macaque F5 is homologous to Broca’s area. However, Belmalih et al. (2009) offer a more subtle parcellation of relevant brain areas in macaque; while Ferrari et al. (2017) provide new data distinguishing mirror neuron networks in the manual and orofacial pathways (and see Coudé & Ferrari, 2018), linked to sensorimotor and limbic regions, respectively. This points to further exploration of the overlap between manual and orofacial networks “both within and beyond the mirror” to underpin

analysis of (1a) the motivation to communicate and (1b) the linkage between vocal and manual communication.

3. The brains of macaques and humans support diverse systems for working memory and for sequence learning and recall. More care is required to tease apart these subclasses before one can carefully elaborate their evolutionary relationship.
4. It would be useful to provide comparative models of modulation of innate calls in monkeys, ontogenetic ritualization and possible adaptation of an innate gestural repertoire in apes, and phoneme acquisition in human children.
5. As we further chart the evolution of the language-ready brain, we must seek to understand evolution of linked cerebro-cerebellar systems and bring other regions such as basal ganglia and hippocampus into play.

### Acknowledgements

This research was supported in part by NSF Grant No. BCS-1343544 “INSPIRE Track 1: Action, Vision and Language, and their Brain Mechanisms in Evolutionary Relationship,” (M.A. Arbib, Principal Investigator). The paper was prepared for a workshop funded by the same grant. My thanks to the reviewers for their constructive comments.

### References

- Aboitiz, F. (2018). Voice, gesture and working memory in the emergence of speech. *Interaction Studies*, 19(1–2), 70–85.
- Alstermark, B., Lundberg, A., Norrsell, U. & Sybirska, E. (1981). Integration in descending motor pathways controlling the forelimb in the cat: 9. Differential behavioural defects after spinal cord lesions interrupting defined pathways from higher centres to motoneurons. *Experimental Brain Research*, 42, 299–318. <https://doi.org/10.1007/BF00237496>
- Arbib, M. A. (2005). Interweaving Protosign and Protospeech: Further Developments Beyond the Mirror. *Interaction Studies: Social Behavior and Communication in Biological and Artificial Systems*, 6, 145–171. <https://doi.org/10.1075/is.6.2.02arb>
- Arbib, M. A. (2007). How New Languages Emerge (Review of D. Lightfoot, 2006, *How New Languages Emerge*, Cambridge University Press). *Linguist List*, 18–432, Thu Feb 08 2007, <http://linguistlist.org/issues/17/17-1250.html>.
- Arbib, M. A. (2012). *How the Brain Got Language: The Mirror System Hypothesis*. New York & Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199896684.001.0001>
- Arbib, M. A. (2016). Towards a Computational Comparative Neuroprimatology: Framing the Language-Ready Brain. *Physics of Life Reviews*, 16, 1–54.
- Arbib, M. A. (2018). Computational Challenges of evolving the language-ready brain: 2. Building towards neurolinguistics. *Interaction Studies*, 19(1–2), 22–37.
- Arbib, M. A., Billard, A., Jacoboni, M. & Oztop, E. (2000). Synthetic brain imaging: grasping, mirror neurons and imitation. *Neural Networks*, 13(8–9), 975–997. [https://doi.org/10.1016/S0893-6080\(00\)00070-8](https://doi.org/10.1016/S0893-6080(00)00070-8)

- Arbib, M. A., Ganesha, V. & Gasser, B. (2014). Dyadic Brain Modeling, Ontogenetic Ritualization of Gesture in Apes, and the Contributions of Primate Mirror Neuron Systems. *Phil Trans Roy Soc B*, 369 (1644), 20130414. <https://doi.org/10.1098/rstb.2013.0414>
- Arbib, M. A. & Rizzolatti, G. (1997). Neural expectations: a possible evolutionary path from manual skills to language. *Communication and Cognition*, 29, 393–424.
- Arbib, M. A., Schweighofer, N. & Thach, W. T. (1995) Modeling the Cerebellum: From Adaptation to Coordination in Glencross, D. J. & Piek, J. P. (eds), *Motor Control and Sensory-Motor Integration: Issues and Directions*. Amsterdam: North-Holland Elsevier Science, 11–36. [https://doi.org/10.1016/S0166-4115\(06\)80005-1](https://doi.org/10.1016/S0166-4115(06)80005-1)
- Barrès, V., Simons, A. & Arbib, M. A. (2013). Synthetic event-related potentials: A computational bridge between neurolinguistic models and experiments. *Neural Networks*, 37, 66–92. <https://doi.org/10.1016/j.neunet.2012.09.021>
- Belmali, A., Borra, E., Contini, M., Gerbella, M., Rozzi, S. & Luppino, G. (2009). Multimodal architectonic subdivision of the rostral part (area F5) of the macaque ventral premotor cortex. *The Journal of Comparative Neurology*, 512(2), 183–217. <https://doi.org/10.1002/cne.21892>
- Bonaiuto, J. J. & Arbib, M. A. (2010). Extending the mirror neuron system model, II: what did I just do? A new role for mirror neurons. *Biological cybernetics*, 102(4), 341–59. <https://doi.org/10.1007/s00422-010-0371-0>
- Byrne, R. W. (2003). Imitation as behaviour parsing. *Philos Trans R Soc Lond B Biol Sci*, 358(1431), 529–36. <https://doi.org/10.1098/rstb.2002.1219>
- Byrne, R. W. & Russon, A. E. (1998). Learning by imitation: a hierarchical approach. *Behav Brain Sci*, 21(5), 667–84; discussion 684–721. <https://doi.org/10.1017/S0140525X98001745>
- Cangelosi, A. & Parisi, D. (eds) (2002). *Simulating the Evolution of Language*. London: Springer. <https://doi.org/10.1007/978-1-4471-0663-0>
- Cobas, A. & Arbib, M. (1992). Prey-catching and predator-avoidance in frog and toad: defining the schemas. *J Theor Biol*, 157(3), 271–304. [https://doi.org/10.1016/S0022-5193\(05\)80612-5](https://doi.org/10.1016/S0022-5193(05)80612-5)
- Cooper, R. P. (2016) Schema Theory and Neuropsychology in Arbib, M. A. (ed), *From Neuron to Cognition via Computational Neuroscience*. Cambridge, MA: The MIT Press.
- Corina, D. P., Poizner, H., Bellugi, U., Feinberg, T., Dowd, D. & O’Grady-Batch, L. (1992). Dissociation between linguistic and nonlinguistic gestural systems: a case for compositionality. *Brain and Language*, 43(3), 414–47. [https://doi.org/10.1016/0093-934X\(92\)90110-Z](https://doi.org/10.1016/0093-934X(92)90110-Z)
- Coudé, G. & Ferrari, P. F. (2018). Reflections on the organization of the cortical motor system and its role in the evolution of communication in primates. *Interaction Studies*, 19(1–2), 38–53.
- Coudé, G., Ferrari, P. F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., Monti, F., Rozzi, S. & Fogassi, L. (2011). Neurons Controlling Voluntary Vocalization in the Macaque Ventral Premotor Cortex. *PLoS One*, 6(11), e26822. <https://doi.org/10.1371/journal.pone.0026822>
- Fagg, A. H. & Arbib, M. A. (1998). Modeling parietal-premotor interactions in primate control of grasping. *Neural Netw*, 11(7–8), 1277–1303. [https://doi.org/10.1016/S0893-6080\(98\)00047-1](https://doi.org/10.1016/S0893-6080(98)00047-1)
- Ferrari, P. F., Gerbella, M., Coudé, G. & Rozzi, S. (2017). Two different mirror neuron networks: The sensorimotor (hand) and limbic (face) pathways. *Neuroscience*.
- Fogassi, L., Coudé, G. & Ferrari, P. F. (2013). The extended features of mirror neurons and the voluntary control of vocalization in the pathway to language. *Language and Cognition*, 5, 145–155. <https://doi.org/10.1515/langcog-2013-0009>
- Halina, M., Rossano, F. & Tomasello, M. (2013). The ontogenetic ritualization of bonobo gestures. *Animal cognition*, 16(4), 653–666. <https://doi.org/10.1007/s10071-013-0601-7>

- Hecht, E. E., Gutman, D. A., Preuss, T. M., Sanchez, M. M., Parr, L. A. & Rilling, J. K. (2012). Process Versus Product in Social Learning: Comparative Diffusion Tensor Imaging of Neural Systems for Action Execution–Observation Matching in Macaques, Chimpanzees, and Humans. *Cerebral Cortex*, 23(5), 1014–24.
- Hobaiter, C. & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal cognition*, 14, 745–767. <https://doi.org/10.1007/s10071-011-0409-2>
- Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, 26(2), 235–258. [https://doi.org/10.1016/S0149-7634\(01\)00068-9](https://doi.org/10.1016/S0149-7634(01)00068-9)
- Maranesi, M., Livi, A. & Bonini, L. (2015). Processing of Own Hand Visual Feedback during Object Grasping in Ventral Premotor Mirror Neurons. *The Journal of Neuroscience*, 35(34), 11824–11829. <https://doi.org/10.1523/JNEUROSCI.0301-15.2015>
- Marshall, J., Atkinson, J., Smulovitch, E., Thacker, A. & Woll, B. (2004). Aphasia in a user of British Sign Language: Dissociation between sign and gesture. *Cognitive Neuropsychology*, 21, 537–554. <https://doi.org/10.1080/02643290342000249>
- Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J. & Thach, W. T. (1996). Throwing while looking through prisms. II. Specificity and storage of multiple gaze-throw calibrations. *Brain*, 119, (Pt 4), 1199–211. <https://doi.org/10.1093/brain/119.4.1199>
- Oztop, E. & Arbib, M. A. (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biol Cybern*, 87(2), 116–40. <https://doi.org/10.1007/s00422-002-0318-1>
- Oztop, E., Kawato, M. & Arbib, M. (2006). Mirror neurons and imitation: a computationally guided review. *Neural Netw*, 19(3), 254–71. <https://doi.org/10.1016/j.neunet.2006.02.002>
- Oztop, E., Kawato, M. & Arbib, M. A. (2013). Mirror neurons: Functions, mechanisms and models. *Neuroscience Letters*, 540, 43–55. <https://doi.org/10.1016/j.neulet.2012.10.005>
- Poizner, H., Klima, E. S. & Bellugi, U. (1987). *What the hands reveal about the brain*. Cambridge, MA: MIT Press.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G. & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Exp. Brain Res.*, 71, 491–507. <https://doi.org/10.1007/BF00248742>
- Russon, A. (2018). Pantomime and imitation in great apes: Implications for reconstructing the evolution of language. *Interaction studies*, 19(1–2), 200–215.
- Taira, M., Mine, S., Georgopoulos, A. P., Murata, A. & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research*, 83, 29–36. <https://doi.org/10.1007/BF00232190>
- Tomasello, M. & Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.

#### *Author's address*

Michael A. Arbib  
University of California at San Diego  
Psychology Department  
La Jolla  
CA 92093  
USA  
arbib@usc.edu