The comparative neuroprimatology 2018 (CNP-2018) road map for research on *How the Brain Got Language*

Michael A. Arbib,[1] Francisco Aboitiz,[2] Judith M. Burkart,[3]
Michael Corballis,[4] Gino Coudé,[5] Erin Hecht,[6] Katja Liebal,[7]
Masako Myowa-Yamakoshi,[8] James Pustejovsky,[9] Shelby Putt,[10]
Federico Rossano,[1] Anne E. Russon,[11] P. Thomas Schoenemann,[10]
Uwe Seifert,[12] Katerina Semendeferi,[1] Chris Sinha,[13]
Dietrich Stout,[14] Virginia Volterra,[15] Sławomir Wacewicz[16] and
Benjamin Wilson[17]

[1]University of California at San Diego / [2]Universidad Católica de Chile
/ [3]Universität Zürich / [4]University of Auckland / [5]Institut des Sciences
Cognitives Marc Jeannerod – CNRS / [6]Georgia State University / [7]Freie
Universität Berlin / [8]Kyoto University / [9]Brandeis University / [10]Indiana
University / [11]Glendon College of York University / [12]Universität zu Köln
/ [13]Hunan University / [14]Emory University / [15]Italian National Research
Council, CNR / [16]Nicolaus Copernicus University / [17]Newcastle University

We present a new road map for research on "How the Brain Got Language" that adopts an EvoDevoSocio perspective and highlights *comparative neuroprimatology* – the comparative study of brain, behavior and communication in extant monkeys and great apes – as providing a key grounding for hypotheses on the last common ancestor of humans and monkeys (LCA-m) and chimpanzees (LCA-c) and the processes which guided the evolution LCA-m → LCA-c → protohumans → *H. sapiens.* Such research constrains and is constrained by analysis of the subsequent, primarily cultural, evolution of *H. sapiens* which yielded cultures involving the rich use of language.

An overall perspective

The present paper presents the *Comparative Neuroprimatology 2018 (CNP-2018) Road Map* based on the papers of this Special Issue of *Interaction Studies* on "How the Brain Got Language: Towards a New Road Map." The *comparative neuroprimatology* framework for study of language evolution assesses relevant data and theories concerning the brains, behaviors and communication systems of monkeys, apes and humans to raise hypotheses about LCA-m (our last common ancestor with monkeys) and LCA-c (our last common ancestor with chimpanzees and apes more generally) as a basis for investigating the biological and cultural evolution of the human *language-ready brain*.

Four assumptions are shared by the authors of this road map (though several may remain controversial in the language evolution community at large):

1. Our ancestors evolved a capability for **protolanguage** – which had an open lexicon but little if any syntax – before they developed language. Here, "protolanguage" is being used in the sense of "something intermediate between (i) the communication systems of LCA-c and (ii) language – but which is not itself a language."
2. The quest is an exercise in *EvoDevoSocio* – the view that biological evolution defines developmental systems that can both shape and be shaped by cultural evolution, the dynamic emergence of habits of social interaction. We seek to understand how biological evolution yielded brains and bodies (Evo) that could develop (Devo) in a culture that already had (proto)language (Socio) so that children could master the use of that language with the help of caregivers to support **(proto)language acquisition**. And how did these brains enable humans in interaction to support the (extended and polymorphic) **emergence of languages** followed in turn by **historical language change?**
3. The study of language evolution must include brain mechanisms, comparing human brain imaging and lesion data with data on brain mechanisms for "language-related" functions in other species to ground an understanding of what has been conserved and what has been changed in human brain evolution. However, "language-related" functions need not be communicative.
4. Shared mechanisms that support signed as well as spoken languages are crucial. Nonetheless, the importance of spoken language requires us to understand the evolution of human vocal control.

Evolution works by bricolage (tinkering). It does not produce the optimal software on the optimal hardware. Instead, it yields a cultural artifact riddled with historical contingency on a brain whose genetic code reflects selection without conscious design operating on structures (DNA, membrane and cytoplasm) that

are far removed from what we might think are the crucial design features of language. To understand the evolution of the language-ready brain we need to understand *mechanisms* and *processes* and their variation across (at least) primates. Just because processes have the same name (e.g., imitation, pantomime, cognition, theory of mind) does not mean that they are implemented by the same circuits across species or even in one brain.

To keep citations to a minimum, we place the name of each author in **bold italics** to refer the reader to their individual papers in this special issue (listed in the bibliography, which provides names of any co-authors) for detailed references.

Aspects of language to be explained

We start by listing some key properties shared by human languages.

Language is a special form of communication

Lexicon and grammar
A language provides a framework for sharing of meaning in a community by combining *words* (we use the term to include, e.g., the signs of a signed language), perhaps modifying the words in the process, to express both familiar and novel meanings and to understand (more or less) the novel utterances of others (*parity of comprehension and production*). It combines an open-ended *lexicon* with a rich *grammar* that supports a *compositional semantics*.

The endless aboutness of language
A human language is a mechanism to support sharing of meaning in a community about physical and mental worlds. Components of this ability include:

Here-and-Now: A commonly shared assumption is that the primary drive in the evolutionary path to language was the value of being able to coordinate current behavior, with joint attention supporting the sharing of perception of the current environment and plans for acting within that environment (Common Ground).

Theory of mind: The ability to talk about the mental (including emotional) states of others; this may rest on an ability, possibly shared to some extent with other species, to infer the mental states of others, and use these to predict behavior.

Displacement: Moving beyond the co-situated context, language builds on capacities for episodic memory, planning and imagination to support the ability to talk about distant events as well as about the past and the possible future, as well as counterfactuals.

Abstraction: Moving from embodied grounding to disembodied abstractions. **Language as a tool for thought versus language as a means of communication:** Communicative tools are already tools for thought. Directing the goal/intention of a linguistic expression reflectively can be seen as self-communication, the origin of the alleged "non-communicative" aspect of language.

Social structure and the motivation to converse

A prominent question at the workshop was "Why do we talk?" but a more fruitful question is "Why do we converse?" (i) In few species of nonhuman primates does communication involves a back and forth, whereas the dominant form of non-written language is conversation. (ii) The word "talk" overemphasizes the use of speech. The OED definition of *conversation* includes a quote from Boswell's *Life of Johnson* (ed. 2, 1793) that cites Johnson as saying "we had *talk* enough, but no *conversation*; there was nothing *discussed*." This stresses the importance of the endeavor to develop a shared mental understanding, and this seems to require some aspects of theory of mind.

The motivation to converse is one linkage between emotion and language. Here we can note two further aspects: An utterance may be emotionally charged by the way in which prosody, facial expression and posture are integrated with its production – compare the power of music to sway the emotions. However, language can also express emotions without the speaker or signer being emotionally engaged, with emotion serving as just another domain for "aboutness."

Action, gesture and language

Human languages are in most cases spoken languages (though their reach is extended by the emergence of writing), but the languages of the Deaf are fully-formed languages which rely on manual signs (supplemented, e.g., by facial expressions) and make no use of voice. When people do speak, their speech is complemented by cospeech gestures of the hands as well as facial expressions. The puzzle is this: Nonhuman primates exhibit very little in the way of vocal control but do exhibit dexterous manual control. Why, then, did vocal control evolve as part of the human brain's distinctive capabilities, since language could "manage without it"? And how relevant does manual action remain in understanding the brain mechanisms of language?

The notion, in any case, is that the brain mechanisms supporting language can – to a first approximation – be separated from the mechanisms that recognize words in the sensory (auditory or visual) input and generate words in the motor (vocal or manual) output – without denying that neural plasticity will differentially restructure the brain dependent on whether language is spoken or signed – just as literacy can restructure it (Dehaene et al., 2010).

Methodologies

At the cost of some redundancy, we precede our specification of the new road map by illustrating a range of methodologies relevant to it.

Neurophysiology and comparative neuroanatomy

At a gross level, neuroanatomy characterizes distinctive brain regions and the pathways connecting them. At a finer level, it may seek to distinguish the cell types of different regions and the patterns of connectivity within and across those regions. Comparative neuroanatomy can thus suggest hypotheses about the evolutionary relationship of brains of LCA-m, LCA-c and modern humans, enriched by suggestions concerning the functions of specific regions.

Aboitiz compares the anatomy of macaque and human brains in seeking to assess how changing connectivity might have supported the emergence of an auditory working memory (WM) that could provide the "phonological loop" for language. The reference point is that macaques have good visual WM but poor auditory WM.

Hecht focuses on connectivity using diffusion tensor imaging (DTI) in living brains to compare pathways engaged in visuomotor integration across macaque, chimpanzee and human, while discussing the importance of changes in mechanisms of plasticity in complementing the "innate ground plan" of each brain. (Unfortunately, new NIH policies may preclude further US studies of DTI for great apes.) *Semendeferi* increases the level of detail by staining brains of different primates to reveal changes in neuronal structure that underlie differences in the substructure of different nuclei, especially those related to emotion.

Neurophysiology then enriches the comparative database by looking at the dynamic activity of the brain while the human or animal performs specific tasks. In monkeys, we have data on the fine structure of firing of individual neurons in a circuit. In humans, we have imaging techniques that can follow fine timing with very poor spatial resolution (e.g., EEG) or coarse timing with better spatial resolution (but still in terms of millions of neurons as the unit, e.g., fMRI).

Coudé uses neurophysiology to explore detailed activity of neurons, and especially mirror neurons in the manual and orofacial regions of region F5 of macaque prefrontal cortex and, crucially, links this to neuroanatomy showing that these two regions are linked to very different subsystems of the macaque brain.

Wilson use neuroimaging to study sequence processing in the macaque brain to assess what is conserved in the human brain and what additions may have evolved to support syntax.

Behavior, social structure and communication

Both field studies and studies in the lab can provide useful information for our quest even in the absence of neural correlates. Different species differ not only in their behavior and communication but also in the social structure in which these are embedded. Exploring the relationship between social structure and forms of communication may help us better distinguish the evolution of "social support for extensive communication" from the evolution of the general form of language (lexicon, grammar, compositional semantics) on which cultural evolution has played extensive variations – without ignoring the eventual need to explore the interactions between these two evolutionary foci.

Liebal surveys gestures, calls and facial expressions in nonhuman primates to question the view that gestures are intentional whereas facial expressions and vocalizations are emotional. This may accord with a broader theoretical assessment of the linkage between emotion and intentions, and the observation that what distinguishes vervet alarm calls is not emotion (each expresses a fearful situation) but rather the difference between eagle, snake and leopard.

Burkart notes that callitrichid monkeys (e.g., marmosets) appear to exhibit particularly elaborate vocal communication, including vocal turn-taking. She explores the hypothesis that this is linked with cooperative breeding (i.e., infant care shared among group members). Among primates, this rearing system is correlated with proactive prosociality, which can be expressed as a motivation to share information. Since humans are the only cooperative breeders in primates besides callitrichids, cooperative breeding may contribute to understanding why language evolved in our species, rather than in any other primate.

Rossano uses comparative study of social manipulation, turn-taking and cooperation in apes to develop implications for the evolution of language-based interaction in humans – but note the emphasis here on social conditions for such interactions, not the particularities of language as distinct from other forms of social competition and cooperation. Importantly, he further stresses the need for longitudinal studies to explore the emergence of different gestures in apes in comparison with language development in children.

Both imitation and pantomime have figured in discussions of the relation between action and language. *Russon* analyzes imitation in orangutans and offers evidence that they have some form of *pantomime* and that its use does not involve imitation. *Myowa* compares *imitation* in chimpanzees and young children to assess their different styles of imitation, noting that chimpanzees attend to the hands of the imitatee while children also glance back and forth at facial expression. The key lesson is that the same term may be employed for an ape and a human capacity but that there may be differences that require an evolutionary explanation.

Volterra focuses on humans, revealing a *developmental* progression from actions via gestures to "words" (whether signed or spoken). Does this support the hypothesis of an *evolutionary* progression from manual action via gesture to protolanguage? The answer will require a delicate treatment of the relation between phylogeny and ontogeny.

Corballis offers a comparative perspective (not limited to primates) of episodic memory, broadly construed, to suggest an evolutionary basis for a key property of language, *displacement.*

Archeology

Archeology asks what can be learned from the remains of protohumans (australopithecines and predecessors of *sapiens* in the genus *Homo*) and early humans and their artefacts. New findings about Neandertal culture are further enriching the database. Here, the primates with whom modern humans being compared are all extinct hominins rather than extant apes or monkeys.

Schoenemann focuses in part on the sparse set of skulls of *Australopithecus* and *Homo* and the somewhat limited inference of relative size of different cortical regions from endocasts of the skulls whose indentations are indicative of gross cortical shape.

Cognitive archeology examines "cultural remains" of the daily lives of our ancestors to hypothesize the cognitive processes involved in their making and use. *Stout* and *Putt* carry this further, employing "neuro-archeology" – they teach modern humans to make stone tools of the kind found by archeologists; see what parts and connections of the brain are "exercised" by learning the ancient skill; and hypothesize that their enlargement may have been a step in brain evolution.

High-level theory

Diverse "high-level" theoretical approaches may complement attempts to generate and directly address the data of comparative neuroprimatology.

Wacewicz offers an approach more consonant with general evolutionary theory to complement the work of Burkart and Rossano by emphasizing trust, cooperation and turn-taking in language origins.

Sinha offers a general EvoDevoSocio perspective that highlights the role of biological and cultural co-evolution, with particular emphasis on the evolution of praxis, symbols and infancy.

Seifert exemplifies a broader assessment of culture-readiness by investigating what is and is not shared between music-readiness (more attuned to emotional expression?) and language-readiness (more attuned to propositional content?).

Modeling and mechanism

Pustejovsky probes the relation between action, perception and language, taking a step toward modeling the actual mechanisms that may underlie the use of language.

Arbib 1 (2018a) introduces explicit modeling of biologically plausible *neural networks*, including frontoparietal interactions in macaque brain for the control of grasping, development of mirror neurons for manual actions, and opportunistic scheduling of sequences of actions. He then suggests how macaque mechanisms could be augmented to supply a hypothetical model of the ape brain adequate to support the emergence of novel gestures through dyadic interaction.

Arbib 2 (2018b) offers a complementary style of modeling, *schema theory*, that can be applied to other primates but is especially relevant when modeling human capabilities such as visual scene understanding and language use for which data on activity at the neural level is sparse or unavailable. He models the "aboutness" of language in comprehension and production, and develops hypotheses about the evolutionary relation between manual action and language.

Genetics

Genetics lies outside the scope of the present roadmap, and thus is of high priority for its sequels. Note the distinction between finding genes that act "merely" as markers (these remains are *sapiens*, those are Neandertal – but even these may be relevant to establishing timelines) and those that can be linked to changing functionality of brain or body. One clear target is the assessment of how different forms of neural plasticity may have evolved to provide circuits with novel capacities for learning.

Road map preliminaries

Using the term "hominin" for genera that emerged after the split from the great apes (australopithecines, *Homo*), with "hominid" including the great apes as well, we base the evolutionary account on four (probably overlapping) stages. Each subsequent stage raises the question: What is new here, and how did it build on or depart from features of the previous stage?

LCA-m.   Database: Monkeys.
LCA-c.   Database: Modern great apes (thus **Russon's** suggestion that LCA-ga would be a better term).
After LCA-c.   Database: Hominin fossil record up to c. 200 Kya
Modern *Homo sapiens*. Database: Archeology since c. 200Kya, historical record and current observation

We briefly summarize key "landmarks" and "connecting roads" for these stages but do not provide references for the details. Instead, we mark items *MSH* if they are part of the mirror system hypothesis as set forth in Arbib (2012), and present the name of the ***author*** in bold italics if they have discussed this item in this special issue. Areas of disagreement help define key challenges for future research. Since the length of the paper is limited, key points are omitted, but we have aimed to provide a firm framework for future elaboration.

Since the body of actual and future research on each stage is overwhelming, the meta-challenge is to assess what it is at each stage that may be relevant to understanding "how the brain got language." For example, if we view speech as the sine qua non for language, we might focus on monkey calls as a prime dataset. If we emphasize that human languages may be signed, then ape gesture may seem equally relevant. But once one looks at manual gesture, one may return to monkeys to study manual action more generally. Similarly, one may look at modern languages in the richness of their aboutness, or one may instead focus simply on the ability to string words into sequences, and then emphasize mechanisms in the monkey brain that support sequential behavior. We espouse a comprehensive framework.

Establishing the "Stages"

What capabilities of brain, behavior and communication should define stages in our road map? We need to avoid being seduced by the metaphor of the evolutionary tree, for we now understand that extant species at one stage may evolve differentially yet continue to cross-breed – and so at each stage we may establish a suite of capabilities that may have been distributed across different species and populations. What evolutionary principles could explain how the human brain might aggregate them? Moreover, primates in human captivity may acquire capacities never seen in the wild. To what extent does that imply they have the brain mechanisms to support that capacity but do not have the capability for the cultural evolution that led to that capacity?

Modern species did not evolve from each other. Thus, one challenge is to study various extant monkeys to extract a shared core (brain, behavior, communication) to define the LCA-m baseline. But what of traits seen in some monkey species that don't meet our criteria for LCA-m and yet are shared by humans? Perhaps convergent evolution was involved. But if so, we must hypothesize where and how this property re-emerged. Examples: Vocal turn-taking in the marmoset does not seem to qualify as a property of LCA-m. One group of cebus (capuchin) monkeys exhibits tool use (using stones to crack palm nuts), other groups and other species do not. ***Burkart*** assesses how the former may provide insights into human social

structure; the latter may be relevant for placing (proto)human stone tools (**Stout** and **Putt**) in an evolutionary context.

Similarly, we need a fuller assessment of what properties of present-day great apes can plausibly be attributed to LCA-c or may offer suggestions for convergent evolution.

In search of precise terminology

Another methodological challenge is that many of the terms that appear in this field have different meanings when applied to different species. Future work must refine the terminology to the point where we can address the questions: Which definition best characterizes the version seen in one species rather than another, and how does this license the version(s) posited for LCA-m, LCA-c and later? For version X, can we establish the properties of the X-ready brain and the cultural conditions (if necessary) that support its expression? Then, when we note version X posited for one ancestral species and version Y posited for a later ancestral species, we must investigate: Is X a precursor of Y, or was it a terminological "coincidence" that X and Y are refinements of the same term? If we can establish that X is a precursor of Y, is an X-ready brain also a Y-ready brain, with the evolution being primarily cultural? Or is a Y-ready brain different from an X-ready brain, so that biological as well as cultural evolution is involved? Note that these questions apply more generally. For example, while all would agree that an *H. sapiens* language-ready brain is a reading-ready brain, some may disagree with the view (held in MSH) that a protolanguage-ready brain is already language-ready.

Here are four of the terms whose refinement is relevant to defining our road map:

Imitation:	Inspired in part by Byrne and Russon (1998), **MSH** distinguished very limited imitation (e.g., effector priming) in LCA-m, "simple" imitation in LCA-c and "complex" imitation in humans, but **Myowa** adds a new dimension to complex imitation – attention to emotional state as well as the performance of the skill.
Pantomime:	**MSH** defines a form of pantomime that builds on complex action recognition (a prerequisite of complex imitation) and posits that it evolved in the hominin line; **Russon** reports on pantomimes in orangutans and so posits pantomime as a component of LCA-c (her LCA-ga).
Turn-taking:	**Burkart** assesses turn-taking in callitrichids; **Rossano** presents three variants of turn-taking, suggesting that the one applicable to human language may not be a descendant of the callitrichid version (see also **Wacewicz** for a similar view).

Episodic Memory: ***Corballis*** offers a wide range of capabilities (e.g., navigation in rats, recall of sites where food is cached by squirrels) as examples of the great ancestral depth of episodic memory; ***Pustejovsky*** sees the ability to conceptualize events, extracting them from the embodied flow of experience, with recalling such events as the form of episodic memory underlying much of language use, as unique to humans.

Beyond the primates

Parrots and some other birds have flexible vocal production and imitation, while dogs can acquire a large receptive (not productive) vocabulary for spoken commands; none of these appear to have grammar. Much is to be learned from studying such capabilities and their neural basis, but while such studies will usefully complement work on primates (Petkov & Jarvis, 2012), these lie outside our present purview.

The CNP-2018 road map

Capabilities of LCA-m

MSH: Manual dexterity with a related mirror neuron system supporting action recognition, but no capacity for "real" imitation. Integration of the mirror system with systems "beyond the mirror," including a visual dorsal "how" pathway and ventral "planning" pathway for the reach-to-grasp system. Serial behavior including opportunistic scheduling of actions.

Coudé argues that more attention must be paid to the oro-facial mirror system. Whereas the manual system is related to parietal-premotor circuits, the oro-facial system connects with limbic structures. Exploring the linkage between these two systems could underpin efforts to chart evolution of the linkage between communication and emotions.

Arbib 2 addresses the "aboutness" of language by suggesting that a system for linking visual perception of the current environment to a plan for manual action may be the precursor for a system of semantic representation in the language-ready brain. This notion needs to be assessed in relation to models of sequential behavior that set the baseline for *Wilson's* exploration of their relevance to syntax.

LCA-m is posited to have vocal communication (an innate call repertoire) but (almost) no vocal learning and little importance for manual gesture. Where *MSH* posits (at later stages) that (manual) protosign provided the scaffolding for the evolution of flexible vocal control and learning, *Aboitiz* (without denying the importance of gesture) argues for a direct road in evolving speech, requiring more

careful attention to the auditory system and precursors to vocal control in monkeys. He is particularly concerned with precursors of the form of *working* memory in humans called the phonological loop. Resolving the debate between the "vocal control first" and "semantics first" hypotheses is a major challenge. *Corballis* challenges us to assess what form of *episodic* memory LCA-m had: was it more than the ability to form a limited cognitive map?

Capabilities of LCA-c

LCA-m properties are conserved, but further capacities become available.

MSH emphasizes *simple imitation*, attempting to use familiar manual actions to achieve recognizable goals and the use of *gesture* to communicate (but not to converse). *Arbib 1* offers a model of how some of these could be learned by *ontogenetic ritualization* without dependence on imitation. For *MSH*, learnable gestures are on the path to language whereas primate calls are not (recall the debate on whether protosign provided essential scaffolding for the evolution of speech).

The use of gestures shows that *intentional communication is already established in LCA-c*. The acquisition of human-demonstrated "symbols" by enculturated apes shows that LCA-c was *symbol-ready*, even though LCA-c "cultures" were not symbol-rich. In what sense are these symbols similar to those of humans, with their rich conceptual repertoire?

Liebal challenges us to assess vocal calls and manual gestures in monkeys and great apes as a basis for better defining the evolutionary path (changes in brain and culture) that link them, and for grounding a more careful analysis of their links to emotion and intentionality.

Semendeferi compares emotion-related structures in different great ape species, assessing what they might offer in defining the LCA-c brain in contrast with the human brain to suggest that an expanded capacity for emotional processing could be linked to language-readiness. Determining the relevant connections and assessing their role in linking emotion and communication remains a crucial challenge. Helping address this will be *Coudé's* enrichment of the macaque mirror system database, and *Hecht's* use of DTI to compare mirror neuron connectivity in macaque, chimpanzee and human. A further challenge is to relate this to comparative studies of language-related connectivity (Rilling, 2014) as differentially assessed by *Aboitiz* and *Arbib 2.*

Russon presents pantomimes observed in great apes (especially orangutans) to argue for pantomime as a capacity of LCA-c (her LCA-ga). She observes that great ape pantomime does not rest on imitation, whereas *MSH* posits (next section) that complex action recognition and imitation evolved post LCA-c and prior to pantomime. Do her data invalidate the *MSH* claim or is this rather a challenge for

terminological refinement? In either case, modeling brain mechanisms support-
ing these forms of pantomime will be crucial to assessing these social functions.

Hominins prior to Homo sapiens

MSH posits a sequence of five stages from LCA-c to language-ready *Homo sapiens*:
The first combines *complex action recognition,* the ability to attend to the subgoals
and some details of the constituent movements of an observed behavior, with *com-
plex imitation*, the ability to use such recognition to acquire new skills. Second,
pantomime emerges, based on complex action recognition (but perhaps not on
imitation), supporting the creation of novel pantomimes "on the fly" and the abil-
ity of others to recognize them. This opens up semantics beyond the limited range
offered by innate vocalizations and ape gestures. Third, frequently used panto-
mimes become conventionalized within groups to provide *protosign*. Fourth, early
protosign constructed the niche for the emergence of sophisticated *vocal learning
and control*, thus augmenting protosign with *protospeech* in an expanding spiral.
Fifth comes *protolanguage*, the capacity to recognize classes of events and link
them to "protowords," whether signed or spoken.

Clearly, each claim here offers challenges. We have already mentioned the de-
bates over speech and pantomime. One may add the *Corballis-Pustejovsky* debate
over when event perception became developed enough to support protolanguage,
let alone the ability to converse about past and imagined events. Did the latter oc-
cur with early protolanguage, or did it await the emergence of language?

But even were the above sequence correct, serious problems remain. Here are
a few:

Timeline: When did these substages occur? In australopithecines? Did *H.
                    habilis* or *H. erectus* or yet uncharted forms of early *Homo* see the
                    emergence of key innovations? *Schoenemann* assesses the data
                    from endocasts but these provide weak constraints since we lack
                    insights into what it really takes for a brain to support any of the
                    above capabilities.

Social structure: What social structure was necessary for the success of these inno-
                    vations? *Rossano* offers a comparative view of different patterns of
                    social interaction and communication that apes may exhibit. A ma-
                    jor challenge then is to assess what combination of these were rel-
                    evant to the evolution of language, and how they contributed to the
                    "platform of trust" that *Wacewicz* sees as necessary for the success
                    of (proto)language and the ability of children to acquire it – without
                    denying the capacity of humans (shared with chimpanzees) to steal,
                    and to violate that trust in diverse ways (Byrne & Whiten, 1988).

Culture, more generally: Brains do not fossilize and the evidence from endocasts is limited. We have no record of language before the invention of writing a few thousand years ago. But we do have a profusion of stone tools and other artefacts. **Stout** and **Putt** combine instructing modern humans in Oldowan versus Acheulean stone knapping with brain imaging to hypothesize what might have changed in the parieto-frontal system to support these technologies. **Stout** assesses the pedagogy involved to calibrate forms of imitation and assess the level of (proto)language that might have been needed to support training in the relevant skills. For each form of culture, we must assess to what extent its evolution depended on the biological or cultural evolution of (proto)language, and to what extent it contributed to it. **Seifert** explores a possible relationship between the evolution of the language-ready brain and music-ready brain and raises questions as to what may be shared (could prosody be part of the overlap?) and what is distinctive. Here, again, we face the issue of what makes a brain "ready" for a domain of culture, and how cultural evolution may have exploited those resources.

Post-biological evolution in Homo sapiens

MSH holds that early *Homo sapiens* had protolanguages (diverse "protowords" with little or no grammar) in vocal and manual modalities, but not languages (with a grammar to support compositional semantics) – and that it was cultural evolution that underlay the transition via increasingly complex protolanguages to languages which in turn increased in complexity (there is no sharp boundary) along with increasing complexity of social structure. The ability to form protowords yielded to the ability to freely extend the lexicon and develop diverse constructions to support on-line production and comprehension of utterances which (in a possibly context-dependent way) convey new meanings.

Arbib 2 discusses the challenge of assaying the relative plausibility of the Bickertonian version of protolanguage (with the transition to language adding "merge" to a set of words) and the MSH version (with the transition both fractionating protowords to yield constructions and words and building from there). Dubreuil and Henshilwood (2013) alert us to the challenge of placing the transition in the hominin timeline, but our road map holds that the transition was gradual, with no clear break between complex protolanguages and simple languages.

Wilson compares monkey and human brains to suggest how sequence processing in LCA-m might have survived as the core of syntax as assayed by the learning of artificial grammars. But how do artificial grammars relate to meaningful

conversation? Neurolinguists have no pre-eminent theory of grammar whose operation in the brain they agree to study. The only candidate offered in this special issue is Template Construction Grammar (*Arbib 2*), but there is no reason to expect it to survive as more than a crude approximation.

Volterra offers insights into the aboutness of language by investigating the progression from actions to gestures to words in the young child and the emergence of cross-modal (gesture-word) combinations, although more investigation is required about the further development of grammar. Here we return to the crucial *DevoSocio* challenge of understanding the evolution of brains that not only support the human child's ability to learn a language but also the caregiver's ability to assist the process. *Stout's* notion of technological pedagogy, linking acquisition of technical skill with imitation and instruction, may prove helpful.

Sinha presents three "spheres" that together provide the setting for modern humans: the sphere of infancy and childhood, including learning and teaching; the technosphere of praxis and its products; and the semiosphere of communication and its mediating signs. The preceding pages offer pieces of the road map relevant to these spheres, and sets the grand challenge of not only providing each of them with a testable evolutionary scenario rooted in (computational) comparative neuroprimatology but also exploring their mutual dependencies during their evolutionary progressions, both biological and cultural.

Surveying artifacts from the last 100,000 years, one can seek to assess the cognitive capacities required for constructing shelters, for burial practices and for cave art – and then debate whether language was necessary for the development and transmittal of these cultural practices, or whether protolanguage or "mere" imitation would have sufficed. A thoughtful cautionary note is provided by Dubreuil and Henshilwood (2013) who survey a range of archeological evidence to conclude (p. 257) that

> Language readiness results from a combination of several neurocognitive mechanisms, often independent of one another. The absence of one of these mechanisms may not have prevented the evolution of language, but may have led to the evolution of impoverished forms of language. The most likely scenario, in our view, is that the brain was almost language-ready significantly before *Homo sapiens* and that the cultural evolution of languages was well underway when the first *sapiens* evolved. This is not to say, however, that *Homo erectus* and *Homo heidelbergensis* were speaking languages totally akin to ours. Limitations in perspective-taking and mind-reading abilities might have prevented some features of modern human languages from evolving, such as metalinguistic awareness, irony, and potentially some complex syntactical structures.

Envoi

The current road map cannot do justice to the richness of research in the diverse disciplines that it touches upon. Somewhat humblingly, one may note that Jon Kaas has recently published the second edition of *Evolution of Nervous Systems* (Kaas, 2017) in four volumes, with Volume 3, The Nervous Systems of Non-Human Primates, and Volume 4, The Evolution of The Human Brain: Apes and Other Ancestors, providing but a small part of the treasure trove to be exploited in building on the sample provided here. Meanwhile, we invite readers to explore the selected treasures in the 21 preceding papers in this special issue. Each concludes with a section "Towards a New Road Map." Their totality offers far more detail than the Road Map presented here – the one towards which the others are pointing – but the present paper offers a more integrated view than the others can provide. It is our hope that the CNP-2018 Road Map will not be the last, and we would welcome suggestions on how it might be enriched in future editions. Those sent to Michael Arbib (arbib@usc.edu) may, after editing and with your permission, be posted on ResearchGate as part of his Project at https://www.researchgate.net/project/Evolution-of-the-language-ready-brain.

Acknowledgements

References

Aboitiz, F. (2018). Voice, gesture and working memory in the emergence of speech. *Interaction Studies*, 19(1–2), 70–85.

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. (2018a). Computational Challenges of evolving the language-ready brain: 1. From Manual Action to Protosign. *Interaction Studies*, 19(1–2), 7–21.

Arbib, M. A. (2018b). Computational Challenges of evolving the language-ready brain: 2. Building towards neurolinguistics. *Interaction Studies*, 19(1–2), 22–37.

Burkart, J. M., Guerreiro Martins, E. M., Miss, F., & Zürcher, Y. (2018). From sharing food to sharing information. Cooperative breeding and language evolution. *Interaction Studies*, 19(1–2), 136–150.

Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: a hierarchical approach. *Behav Brain Sci*, 21(5), 667–684; discussion 684–721. https://doi.org/10.1017/S0140525X98001745

Byrne, R. W., & Whiten, A. (1988). *Machiavellian Intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans* (1 ed.). Oxford: Claredon Press.

Corballis, M. C. (2018). Mental travels and the cognitive basis of language. *Interaction Studies*, 19(1–2), 353–370.

Coudé, G., & Ferrari, P. F. (2018). Reflections on the differential organization of mirror neuron systems for hand and mouth and their role in the evolution of communication in primates. *Interaction Studies*, 19(1–2), 38–53.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., … Cohen, L. (2010). How Learning to Read Changes the Cortical Networks for Vision and Language. *Science*, 330(6009), 1359–1364. https://doi.org/10.1126/science.1194140

Dubreuil, B., & Henshilwood, C. S. (2013). Archeology and the language-ready brain. *Language and Cognition*, 5(2–3), 251–260. https://doi.org/10.1515/langcog-2013-0018

Hecht, E. E. (2018). Plasticity, innateness, and the path to language in the primate brain: Comparing macaque, chimpanzee and human circuitry for visuomotor integration. *Interaction Studies*, 19(1–2), 54–69.

Kaas, J. (Ed.) (2017). *Evolution of Nervous Systems (Second Edition; in 4 volumes): Elsevier*.

Liebal, K., & Oña, L. (2018). Mind the gap – moving beyond the dichotomy between intentional gestures and emotional facial and vocal signals of nonhuman primates. *Interaction Studies*, 19(1–2), 121–135.

Myowa, M. (2018). The Evolutionary Roots of Human Imitation, Action Understanding and Symbols. *Interaction Studies*, 19(1–2), 183–199.

Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, 4, 12. https://doi.org/10.3389/fnevo.2012.00012

Pustejovsky, J. (2018). From Actions to Events: Communicating through Language and Gesture. *Interaction Studies*, 19(1–2), 289–317.

Putt, S., & Wijeakumar, S. (2018). Tracing the evolutionary trajectory of verbal working memory with neuro-archaeology. *Interaction Studies*, 19(1–2), 272–288.

Rilling, J. K. (2014). Comparative primate neurobiology and the evolution of brain language systems. *Current Opinion in Neurobiology*, 28, 10–14. https://doi.org/10.1016/j.conb.2014.04.002

Rossano, F. (2018). Social manipulation, turn-taking and cooperation in apes: Implications for the evolution of language-based interaction in humans. *Interaction Studies*, 19(1–2), 151–166.

Russon, A. (2018). Pantomime and imitation in great apes: Implications for reconstructing the evolution of language. *Interaction Studies*, 19(1–2), 200–215.

Schoenemann, P. T. (2018). The evolution of enhanced conceptual complexity and of Broca's area: Language preadaptations. *Interaction Studies*, 19(1–2), 336–351.

Seifert, U. (2018). Relating the evolution of Music-Readiness and Language-Readiness within the context of comparative neuroprimatology. *Interaction Studies*, 19(1–2), 86–101.

Semendeferi, K. (2018). Why do we want to talk? Evolution of neural substrates of emotion and social cognition. *Interaction Studies*, 19(1–2), 102–120.

The comparative neuroprimatology 2018 (CNP-2018) road map for research · 337

Sinha, C. (2018). Praxis, symbol and language: developmental, ecological and linguistic issues. *Interaction Studies*, 19(1–2), 239–255.

Stout, D. (2018). Archaeology and the evolutionary neuroscience of language: the technological pedagogy hypothesis. *Interaction Studies*, 19(1–2), 256–271.

Volterra, V., Capirci, O., Rinaldi, P., & Sparaci, L. (2018). From action to spoken and signed language through gesture: some basic developmental issues for a discussion on the evolution of the human language-ready brain. *Interaction Studies*, 19(1–2), 216-238.

Wacewicz, S., & Żywiczyński, P. (2018). Language origins: The platform of trust, cooperation, and turn-taking. *Interaction Studies*, 19(1–2), 167–182.

Wilson, B., & Petkov, C. I. (2018). From evolutionarily conserved frontal regions for sequence processing to human innovations for syntax. *Interaction Studies*, 19(1–2), 318-335.

Address for correspondence

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

arbib@usc.edu