Available online at www.sciencedirect.com

ScienceDirect

Physics of Life Reviews 26–27 (2018) 184–197

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Reply to comment

Extending research on language foundations and evolution Reply to comments on "Rethinking foundations of language from a multidisciplinary perspective"

Tao Gong [a,b,*], Lan Shuai [b], Yicheng Wu [c,**]

[a] *Center for Linguistics & Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China*
[b] *Educational Testing Service, Princeton, NJ, United States of America*
[c] *Department of linguistics and translation, School of International Studies, Zhejiang University, Hangzhou, China*

Received 16 August 2018; accepted 4 September 2018
Available online 10 September 2018
Communicated by L. Perlovsky

It is always inspiring to digest these thought-provoking commentaries on our target article. Scholars from subfields of linguistics, as well as other disciplines spanning from anthropology, experimental psychology, evolutionary biology and computational neuroprimatology have made a variety of valuable comments, many of which have extended the scope of the target article, enriched the fields that the target article pays less attention to, and brought forth useful guidance for current and future multidisciplinary investigations on language foundations and evolution. With the accumulation of direct/indirect empirical evidence, application of multiple approaches from different disciplines, and spread, acceptance, and adoption of evolutionary perspectives, more and more researchers, especially those in disciplines other than linguistics, have acknowledged that any theoretical or empirical investigations on language foundations must pay attention to relevant issues concerning the evolutions of humans and language, such as how fundamental linguistic components or language learning/processing abilities come into being or derive from other abilities in humans, what the evolutionary trajectories of those components or abilities appear to be, and what the critical factors or minimal conditions that triggered those components or abilities are.

In this reply, we aim to: (a) clarify the challenges on our positions in the target article; (b) evaluate some popular yet erroneous views on language foundations and evolution from different perspectives; and (c) summarize a number of profitable extensions in this line of research. It is utterly impossible in this short reply for us to address all issues raised in the commentaries, some of which actually deserve careful reconsiderations and call for follow-up investigations by experts in specific fields.

DOI of original article: https://doi.org/10.1016/j.plrev.2018.04.004.

DOIs of comments: https://doi.org/10.1016/j.plrev.2018.06.001, https://doi.org/10.1016/j.plrev.2018.09.002, https://doi.org/10.1016/j.plrev.2018.07.006, https://doi.org/10.1016/j.plrev.2018.07.004, https://doi.org/10.1016/j.plrev.2018.06.015, https://doi.org/10.1016/j.plrev.2018.06.016, https://doi.org/10.1016/j.plrev.2018.06.009, https://doi.org/10.1016/j.plrev.2018.06.006, https://doi.org/10.1016/j.plrev.2018.07.005, https://doi.org/10.1016/j.plrev.2018.06.007, https://doi.org/10.1016/j.plrev.2018.07.003, https://doi.org/10.1016/j.plrev.2018.06.002.

* Corresponding author at: Educational Testing Service, Princeton, NJ, United States of America.

** Corresponding author.

E-mail addresses: tgong@ets.org (T. Gong), wuyicheng@zju.edu.cn (Y. Wu).

https://doi.org/10.1016/j.plrev.2018.09.001

1. Perspectives on language foundations and language evolution

Most commentators did not explicitly separate the two terms, language foundations and language evolution, and some even stressed language evolution more than foundations. Many commentaries and recent research mentioned therein have revealed that the candidates of language foundations lie not only inside humans, including genomic and anatomic settings (e.g., [1,2]), neural regions in human brain (e.g., [1,3,4]), and capacities in human cognition (e.g., [3,5]), but also outside humans in the socio-cultural or ecological environments of language (e.g., [6–12]).

Comrie [1], for example, mentions Dediu and Ladd's work [13] showing a high correlation between the allelic proportions of two brain size genes in human populations and the distribution of tone languages in those populations. This finding exemplifies that certain genes could serve as the genetic predispositions to linguistic features like tone, which is in line with our position in the target article that biological predispositions for language are not necessarily language-specific, but serve important roles beyond language like manipulating the brain size, as in this case. Similar to the FOXP2 gene, these genes could be one of the many genetic settings for language foundations and evolution. Nonetheless, discrepancy exists among biolinguists on whether observed correlations between genetic settings and linguistic features are really causal or just accidental [14–16].

Corballis stresses the correlative roles of the mirror neurons [17] in monkeys (production and perception of manual grasping) and in humans (production and perception of vocal signals), and proposes, along with others [18,19], the "gestural origin" hypothesis of language [20]. These theories echo, as well, our position that neural predispositions for language initially could be not for language processing. In addition, Corballis reiterates that the cognitive ability of mental time travel, a human-unique ability, is a primary driver for language evolution [20]. From an evolutionary perspective, he proposes an evolutionary scenario, as an alternative to the generative view [21], on how the defining feature of language gradually evolved along with expansion of memory and information, increase in brain size, development of manufacture, and increasingly complex social structures.

Q. Wang in his commentary [22] mentions a presumably speech-specific biological predisposition, uvula. However, according to the World Atlas of Language Structures (WALS) [22], the consonant systems of 468 out of 566 documented languages contain no uvular consonants at all [23], but this does not prevent native speakers of those languages from acquiring or using their mother tongues or many other languages. In addition, signed languages are as fully linguistic as speech. Even though signed language users might unconsciously move their speech-related muscles while signing, the (dis)appearance of uvula does not affect the acquisition and use of signed languages. This logic also applies to Q. Wang's arguments in attempt to ascribe the permanent descent of the larynx in humans as a biological predisposition to language (speech obviously). We treat the descent of the larynx as one of the examples of pursuing language-specific biological predispositions. Our position is that such attempts should not be limited to speech-related anatomical structures, even though we agree with the famous aphorism that "absence of evidence is not evidence of absence".

In rapport with the evolutionary perspective, viewing language as a complex adaptive system [24,25] or a "human-driven" complex adaptive system as in Liu's commentary [11], as well as deploying network based, connected views on a language system, its neural/genetic bases and socio-cultural/ecological environments can provide more comprehensive understanding of language foundations and evolution.

A recent neuroimaging study, for example, demonstrates that certain classic language regions show little or no response to nonlinguistic functions (e.g., arithmetic, working memory, cognitive control, or music), many of which tend to activate other regions of the brain [26]. First, without referring to some baselines (e.g., default networks [27]), direct comparisons between the activations during linguistic and nonlinguistic tasks cannot explicitly tell whether some of the brain regions responsible for nonlinguistic functions are (de)activated during linguistic tasks. Second, linguistic functions consist of a number of "low-level" functions, which, if a few of them are shared to a certain extent by some nonlinguistic tasks, there could be overlapping activations between linguistic and nonlinguistic tasks. This has been shown in many studies using more complex nonlinguistic tasks, in which certain classic language regions may play a regulating role in both types of tasks (e.g., [28–30]). Third, linguistic and complex nonlinguistic tasks could be fulfilled by a series of similar or distinct functions, some of which may escape from being captured by the neuroimaging devices due to technical limitations in the temporal and/or spatial domain. Although recent studies have begun to examine collectively the activations of various brain areas in different linguistic tasks involving semantics and/or syntax processing (e.g., [31,32]), lacking detailed comparative evidence between the "high-level", linguistic tasks and the nonlinguistic tasks requiring similar "low-level" operations (e.g., statistical or instance-based learning) could lead to a generative view that certain brain regions (or networks of several regions) serve as the foundations for language-specific functions [32]. In fact, rather than viewing language functions as "high-level" and nonlinguistic ones as "low-level", language functions may consist of many general activities. This integrative perspective could be more appropriate and useful for research seeking neural foundations of language.

Although many commentators acknowledge that language results from both biological and cultural evolution, Mufwene throws doubts on the term of "cultural transmission". Derived from biological evolution, cultural transmission can be viewed as the process by which information is passed between individuals via social learning mechanisms such as imitation or teaching [33]. In many simulation studies of language evolution (e.g., [34]) and experimental studies of language-like communication systems (e.g., [35]), cultural transmission refers to the process of language (or other communicative signaling system) adaption in a community via various kinds of communication among individuals of the same or different generations [36]. Language is acquired and modified during these transmissions, and language evolution is embedded in a socio-cultural environment, in which individuals collectively determine how to shape language during communications. The transmission channel (various forms of linguistic communications between individuals of the same or different generations), transmitted materials (expressions and meanings decoded therein, both of which provide the building blocks for linguistic knowledge), and outcome (misunderstanding is tolerated to a certain extent, and the knowledge developed based on exchanged materials could be slightly different among individuals) of cultural transmission are all distinct from those of biological transmission. Neglect of these critical differences could lead to some wrong impressions, such as language learning being largely faithful, always vertical from parents to offspring, or abilities relevant for language learning are innate.

In addition, Mufwene rightly clarifies that the oft-claimed faster speed of cultural evolution than biological evolution holds only for biological species and cultural phenomena compared. Evolution occurs at both the micro and macro scales, the pace of biological or cultural evolution is subject to the topic under evolution. In evolutionary linguistics, there are distinct timescales to record different aspects of language evolution, such as the micro-history or ontogenetic timescale to address language acquisition, meso-history or glossogenetic timescale to study language change, and macro-history or phylogenetic timescale to talk about language origin [37,38].

Arbib [4] questions the intertwined biological and cultural evolution by claiming that: biocultural evolution produced a language-ready brain before the emergence of language; and the human genome remained stabilized for a language-ready brain at a stage in which early humans had protolanguage but not language, whereas the transition from small lexicons to large lexicons, subtle grammars, and compositional semantics all rest on cultural evolution. These views implicitly assume faster cultural evolution than biological evolution, and are consistent with the perspectives of making a distinction between the evolution of the function of language and linguistic change in modern languages, associating long-term language evolution with physiological and cognitive evolutions through genetic modifications in humans (thus leading to a "language-ready" brain), and ascribing short-term language change to socio-cultural pressures while assuming that human individuals and communities do not differ in physiology, cognition, or genome.

However, there exist conflicting views on the definition and possible forms (gestural, holistic, or lexical) of protolanguage [18–21,38–42], so it seems unreasonable just drawing an explicit line somewhere at the continuous evolution of language and humans and claiming that anything at the biological side would not change after exceeding that line seem.

Additionally, if processing more complex linguistic components is also an integrative process concerning multiple general abilities, biological foundations of language, in terms of the physiological, cognitive, neural, or genetic settings, would continue to evolve along with the complexification of language. Some simulation studies as mentioned in the target article [43,44] have already shown that along with the transition from a holistic protolanguage to a simple language consisting of basic lexical items and word order, domain-general abilities critical for communicative success such as memory or joint attention could also transform themselves from domain-general abilities into language-specific properties. These abilities serve the purpose of improving communicative efficiency, their evolution would continue, despite the forms of evolved language (the evolution of those abilities may stop until some of their roles in communication could be more efficiently achieved by evolved language). The language-specificity degrees of those abilities may become stabilized, similar to what Arbib claims, but such relative equilibrium is an outcome of continuous evolution, and it can be broken and re-achieved (at the same or different levels) when changes occur on the language side due to other factors. Although these simple simulations are based upon a closed language system, what are demonstrated therein can certainly go beyond these simulations to inspire a coevolutionary hypothesis on language and relevant abilities, which serves as an alternative to the "watershed" view that a language-ready brain must be a prerequisite for language evolution [18,19]. Coevolution as such could result in the "something" unique in human biology that enables human children, but not other species, to learn language.

Furthermore, as discussed in other commentaries addressing the socio-cultural, ecological, or demographic effects on language evolution (see Section 3), individual differences at the genetic, physiological, and cognitive levels could also cast important influence on language change and diversity. Therefore, attempts to pursue causes for language change and diversity should not be limited to socio-cultural or ecological factors.

2. Misleading views in generative linguistics on language foundations and evolution

As has been pointed out by Mufwene, a lot of generative linguists still ignore the physical and social aspects of spoken or signed languages, or simply treat language as a consequence of thought enhancing (e.g., [21]). This perspective has led to some misleading views notoriously propagated among generative linguists:

(a) A single biological mutation produced the domain-specific Universal Grammar (UG), which is determined by a single (or a small set of) gene in the genome, or a language-specific acquisition device in the brain;
(b) Language is primarily an internal thinking tool, poor at and unable to serve an important role in communication;
(c) Externalization of language in usage is of little interest, since it is the "performance", not the "competence" of language.

View (a) obviously neglects the fundamental knowledge in genetics and evolutionary biology; any complex or integrated behavior like language cannot be determined by a small number of genetic settings or a set of unique skills, just as illustrated by the history of work on the FOXP2 gene [45]. According to Coupé [10], this view probably could be due to the autonomous conception of language as a closed, complex system at the cognitive and functional levels [46], and therefore, saltationist accounts of evolution and language-specific capacities in the brain appear to be a reasonable answer to the question of language origin. In addition, apparently language-specific phenomena such as Merge or Recursion are actually pervasive in many human activities including action sequences, technology, music, metacognition, theory of mind, and so on [40,47,48], and therefore, language must be a product of an all-purpose mind instead of domain-specific UG [8]. In this sense, although we agree with Q. Wang and Chu that constructing falsifiable hypotheses like the Recursion-Only Hypothesis (ROH) [49] can provide insightful guidance for experimental and comparative studies of language foundations and evolution, discussion of the ROH only in the generative grammar framework, within language structures, or based purely on biological evolution is doubtless limited. For example, a recent study has shown that recursive combination has adaptability in diversifiability of production and products, thus emerging as a driving force for material culture in human communities [50].

View (b) ignores the fact that humans are social animals and socio-cultural communications are not only inevitable but also critical for the survival of both individuals and the whole group. Humans cannot do "mind-reading" without linguistic or nonlinguistic communications. As discussed in the target article and other commentaries [7–12], in excess of biological foundations, socio-cultural/ecological environments constitute another important foundation for language; individual differences presuppose variations in idiolects, and communicative pressures and socio-cultural, ecological, and demographical factors affect communal languages, both of which drive the cultural evolution of language, intertwined with the biological evolution of language. Without the arena of use, language evolution would become a groundless "castle in the air". In addition, Mufwene also underscores the functionalist views that the ecological pressure cast by social life causes thoughts and feelings to be shared more explicitly than nonlinguistic communications and language evolves for explicit and high-fidelity communication.

By isolating language from communication, Q. Wang refuses to accept that language use or socio-cultural factors cast any constraints on language, but ascribes linguistic diversity to a limited set of underlying genes and mechanisms. However, in terms of genetics, it is nearly impossible to define linguistic phenotype, let alone its genotype, and a complex adaptive phenomenon like language cannot be reduced to correspondences between a specific cognitive repertoire and single or a few genes or gene sequences [51]. Recent empirical explorations (see [6] for a number of examples) have also shown that many recorded cases of linguistic diversity are highly correlated with socio-cultural and demographical factors and cannot be interpreted by a set of simple principles or parameters across languages.

Q. Wang contends that language communications are less efficient in delivering some messages than nonlinguistic communications, but ignores the power of language in exchanging other, more sophisticated messages. The same mistake is also committed in his argument that language is a perfect tool for thought. Mufwene gives a simple example to refute this view. In fact, evolution never has the intent to select a globally "perfect" product [52]. As for language, various types of structures could emerge to serve the purpose of expressing some types of meaning in certain socio-cultural or ecological conditions, sometimes at the cost of redundancy or low efficiency of dealing with other types of meaning in other conditions. All these are supposed to boost linguistic diversity.

Finally, view (c), in a sense, blocks linguistics from other scientific disciplines built upon the assumption that latent variables (e.g., linguistic competence, which is usually not directly observable) can be inferred from observable ones (e.g., linguistic performance during learning or use). In generative linguistics, discussion of linguistic competence without referring to observable performances is devastating. For example, linguists holding this view might pay no, or less, attention to authentic language materials or normal individuals' reflections on linguistic materials; instead, they often talk about "linguistic instinct" or self-constructed examples. This separates generative linguistics from corpus linguistics (which documents concrete universality and diversity of linguistic features in the world's languages), linguistic typology (which focuses on real examples having similar or distinct structures in different languages), pragmatics (which examines nonlinguistic or contextual influences on language communications), and psychology (which investigates human reflections on normal or manipulated linguistic materials). Authentic language materials, reflections of the general public on these materials, and other factors during daily conversations may go beyond the imagination of generative linguists and the self-centered comprehension (for example, see the human comprehension experiment [53] and the corpus-based investigation [54] on how native speakers tend to perceive a particular construct in Mandarin Chinese, which are markedly different from what is proposed by generative linguists [55]).

Q. Wang uses two self-constructed examples (sentences 1 and 2 in [2]) to argue that multiple (three) levels of embedding can occur in actual speech, and more often in written texts, and that coordinate structures favor hierarchical interpretation instead of a linear one. Here, he treats speech and writing equally, but the latter came into being much later compared to the former. In addition, understanding written sentences allows almost unlimited reanalysis till a possible comprehension is at hand. "Possible" here means that comprehending written sentences is usually self-centered, unlike normal communications that are often mutual. This could be the reason why generative linguists, based on self-comprehension, may end up with different ways of comprehension from the general public (e.g., [55]). By contrast, if language serves the primary role of communication, restrictions from both speaker and "listener" (to receive either speech or signs) can cast their influence on the construction and perception of exchanged materials. In this context, more than two levels of center-embedding structures (according to the formalist views [56,57], two and more levels of embedding is truly recursive) would occur much rarely in actual speech, because there lack such situations in reality (unless arbitrarily set), and these structures not only require heavy memory load in production and perception, but also easily cause misunderstanding, thus reducing the efficiency of communication. Memory limitation in processing more than two levels of center-embedding structure also reveals that language-specific (if any) and domain-general abilities interact in language use. In addition, although Q. Wang mentions Yang's experiment [58] to support the existence of complex recursion in natural language syntax, the sentences produced by 3–10 year-old children in that experiment only had one or at most two levels of center-embedded structures. Nonetheless, such experimental studies are highly encouraging and can help to figure out at which level language production and comprehension in normal humans (adults or children) lies, before presumptuously jumping to introspective discussions on comprehension of much more complex structures.

The "self-centered" routine in generative linguists' discussions of issues regarding language structure or evolution makes their idealization of "an ideal speaker in a completely homogeneous speech community" sound normal, legitimizes their ignorance of the socio-cultural environment in which language is learned, used, and changed, and plays a role of "firewall" blocking generative linguists themselves from seeking alternative solutions other than what they think purely from linguistic perspectives. For example, per Q. Wang's example sentences (3) and (4), a simple statistical approach is sufficient for children to regulate the relevant words to form grammatical questions, without resorting to an innate knowledge of structure dependence [59] (also see Section 4 for discussions on recent simulation studies challenging this approach). Such an approach has been deployed by infants to detect words from continuous stimuli [60]. In other words, the general assumptions of the Poverty of Stimulus (POS) argument really needs to reconsider the statistical richness of language input to children [61]. In addition, defending the validity of the POS argument should also pay attention to real language materials, general human reflections, and alternative domain-general approaches, instead of linguists' pure introspection or within the generative grammar framework [62].

Contrary to generative linguists, functional, quantitative linguists or linguistic typologists highly value authentic language materials and hence base their research on solid empirical data. For example, Liu [11] stresses the vital role of authentic language materials in reconstructing language foundations and tracing language evolution, and calls for more attentions to real language materials and relationships between language system and humans. Along with the existing large-scale language corpora of sound systems [63], basic vocabulary [64], syntactic structures and demographical information [22], Ansaldo [12] highlights language documentation and description (LDD), which helps recognize the intrinsic value of linguistic diversity by describing and archiving world languages.

Some generative linguists pay attention to or even accept the findings in neuroimaging studies about biological predispositions of language, and use them as supportive evidence for the existence of UG or human-unique and language-specific competence (e.g., [26,32]). In fact, what are captured in many neuroimaging studies are indirect reflections of brain activities (e.g., change in blood oxygen level) towards linguistic materials or other tasks, far from behavioral responses captured in behavioral psychological experiments. According to view (c), if "performance" is of little interest, why bother paying attention to such apparently more indirect evidence? Or, is the performance-competence distinction itself invalid (or should it not be held) from the very beginning?

The above clarifications of the misleading views in some commentaries could not only provide better understanding of language foundations, evolution, and relevant issues, but also benefit cultivating healthy attitudes towards multidisciplinary investigations on topics concerning language foundations and evolution.

3. Influence of socio-cultural, ecological, and demographic factors on language evolution

In contrast to generative linguists' autonomous conception of language system at the cognitive and functional levels [46], which disregards potential influence from external, nonlinguistic factors, commentators from linguistic fields other than generative linguistics or other disciplines view language as an embodied and situated activity performed by speakers, which creates an interface with the physical, biological, cognitive, and sociocultural world inhabited by these speakers [10]. Various components of language system, such as phonology, semantics, morphology, syntax, and pragmatics are all possibly under the influence of a large range of factors, and any weak effects can result in significant impacts on language. These scholars adopt a holistic perspective that integrates physiological, cognitive, genetic, environmental, socio-cultural and communicational factors so as to study long-term language evolution and short-term linguistic change.

Kallens et al. [6], for example, emphasize individual variation and linguistic diversity, two important properties of the cultural evolution of linguistic structure. A review of several studies shows that: (a) linguistic diversity exhibits statistical trends at different levels (morphological, lexical, or phonological) of language structures [65]; (b) language users in a speech community vary substantially in their language skills dealing with different linguistic components [66]; (c) the demographical factor such as the population size of a language is correlated distinctively with the lexical and morphological complexity of that language [67]; and (d) ecological factors affect the tendency of object naming [68]. Variations across and within language communities serve as the cultural evolutionary forces for language.

Zhang [9] calls for investigations on correlations between the cultural evolution of language and demographic factors. He points out that: (a) spread and change of language are commonly dependent on demographic activities; (b) population structures are associated with the culture and prestige of language, especially in a multilingual area; and (c) language contact in adjacent regions also helps shape language structures. Demographic factors serve as the rational answers to language change and diversity. Some simulation studies of language competition have already incorporated demographic factors like population size [69]. With the availability of many large-scale language databases, demographic information of different language communities also has become available (e.g., [70]), which allows detailed investigations into the correlations between demographic factors and diversity (as well as universality) of world languages.

Roberts [7] advocates examining critical social, economic and ecological conditions that could motivate the emergence of a symbolic system like language and relevant linguistic abilities in humans, based on the approach of virtual experimental archaeology (implemented through either experimental semiotics or simulation).

Coupé [10] suggests a distinction between the selective pressures induced by language learning and language use and the roles played by the socio-cultural environment, and calls for investigations into how transformations and variations in past social complexity contributed to the complexification of language forms. On this topic, paleodemographic data and ethnographic proxies could provide useful clues on the social structures of Paleolithic humans (e.g., [71]).

4. Critical period for language learning

WSY Wang in his commentary [72] reviews and discusses the research on critical periods for language learning. This line of research was also mentioned by Comrie, and was not dwelt on in the target article. After reviewing this line of research dating back to Lenneberg and followed by Penfield [73] and Kuhl [74], WSY Wang suggests that the process of learning a language is a process of learning a large set of skills, respectively, or collectively responsible for various linguistic components including phonetics, syntax, and pragmatics. Due to different characteristics of those skills, there could be multiple critical periods for different skills, each of which has its own features and maturational schedules and some may overlap. For example, the critical period for phonology could start as early as the third trimester of fetal life [75], that for morphosyntax may open in early infancy and close in the late teens [76], and that for lexicon opens around one year old [77,78].

WSY Wang points out that the critical period could only be seen clearly in the motoric production of speech, but it does not mean there is a single critical period for language as whole. In addition, he stresses that production of speech requires rapid, precise and coordinated movements of numerous muscles. In this sense, speech is an exaptation (a new function emerged from old bodily parts, and the sensori-motoric aspects of spoken language are surely domain general. This is in stark contrast with Q. Wang's opinion that there exist language-specific biological predisposition to human speech. In fact, WSY Wang suggests that the same critical period for production of speech could play similarly important roles in learning other sensori-motoric skills such as striking sequences of piano keys in coordination with pressing the foot pedal, tossing a tennis ball to an exact height for the outstretched tennis racket to smash it. In addition, he suggests that the cognitive components of language, such as words or grammar other than motoric ones, could be acquired quite late in life. All in all, WSY Wang strengthens our position on the nature of biological predispositions to language by stressing that there are essentially no language-specific biological predispositions at all.

5. Multidisciplinary research on language foundations and evolution

Many commentators share their views on the multidisciplinary research on language foundations and evolution. For example, Liu [11] points out that "multidisciplinary" does not mean a blindly made hodgepodge, but a deliberate introduction of various approaches from other disciplines to solve linguistic problems [79,80]. Roberts suggest that multidisciplinary work should be supported by concrete suggestions for how to proceed. Comrie advocates cooperation of biological and social perspectives; those heavily or exclusively working on one side of the divide should at least be aware of what is taking place on the other. Ansaldo suggests that without understanding of biology and evolution one cannot study language development, and highlights that LDD research is actually a hybrid of linguistics, anthropology, media and computer science. Coupé points out that a combination of large-scale analyses of modern languages, computer simulations, and statistics helps detect and interpret how physiological, cognitive, genetic, environmental, sociocultural and communicational factors get reinforced through time and ultimately weigh on language.

In rapport with these inspiring views, we would like to further stress that multidisciplinary researchers need to:

(a) carefully develop research questions or hypotheses on topics concerning language foundations and evolution;
(b) better understand the scope, advantages, difficulties, and limitations of the multidisciplinary approach to be adopted; and
(c) accurately apply the approach to reliable linguistic materials or materials that can reasonably reflect the target linguistic properties.

Among the three, aspect (a) seems fundamental but still deserves some thoughts. For example, Mufwene points out that many differences among the existing hypotheses on language foundations and evolution are due to disagreements on what language is, or whether UG is invoked to account for its phylogenetic emergence in Homo sapiens and/or ontogenetic emergence in children [38]. He emphasizes that in multidisciplinary research, hypotheses formulated in one discipline must be checked not only against alternatives in the same discipline but also against findings in other disciplines. In addition, Roberts suggests using the Causal Hypothesis in Evolutionary Linguistics Database

(CHIELD) [81] to dissect hypotheses into individual causal steps, differentiate causality among them, and track the amount of empirical support for connections between hypotheses. This is necessary to clarify numerous hypotheses on language foundations and evolution. Moreover, Liu states that investigations of language foundations should be based upon the characteristics of different languages and should not separate language from language users.

Aspect (b) is critical for selecting approaches from various disciplines to evaluate the proposed hypotheses on language foundations and evolution. Many commentators, based on their experiences, have shared their views on some of the popular approaches in multidisciplinary research of language evolution. For example, Roberts suggests that the value of experimental semiotics lies not in realistic recreations of specific steps of language emergence, but in systematically describing and exploring theories of how a particular linguistic ability evolved as a response to certain social, economic or ecological factors. Liu lists a number of approaches that can be used to study various questions about language evolution, including the complex network approach (to study the structure and typology of languages [82]), statistical and cognitive approaches (to explore the universal patterns of linearization of languages and the constraints of human cognition on language structure [83]), and simulation (to evaluate human's self-adaptation mechanism in processing long sentences [84]).

In our early reviews [79,85–88], we have summarized the assumptions, advantages, difficulties, and inevitable limitations of popular multidisciplinary approaches adopted in evolutionary linguistics research. For example, anthropological and archaeological research on language foundations and evolution relies primarily on two assumptions: human-unique behaviors could be determined by particular physiological or anatomical structures of humans; and nonlinguistic evidence could indicate cognitive changes relevant to language evolution. This line of research is often conducted by anatomical comparison between modern humans and fossils of extinct hominins to seek the presence/absence of certain bony conformations associated with speech, or examining archaeological records to approximate cognitive and social complexity of extinct hominins. Such research can inform human migration histories, timespans of the (re)appearance of language in humans, and relations among biological foundations, individual linguistic behaviors, and cultural activities.

However, fossil-based reconstruction of skull endocasts and bony formation usually offer less evidence of soft tissues and inner structures, or operation of some organs in a living body, just as do the discussions on the descent of the larynx in humans in the target article and some commentaries [2]. In addition, physiological apparatus does not necessarily imply the correspondent cognitive capacity, which often results in a fragile basis for research aiming to ascribe particular linguistic behavior(s) to specific anatomical structure(s).

By means of field observation (e.g., [89]), lab experiment (e.g., [90]), and enculturalization (e.g., [91]), many animal behavioral studies on language foundations and evolution have been conducted based on a comparative framework [38,92]: contrasting human behaviors with other species' culturally varied behaviors in order to get a sense of what is the likeliest range of behavioral or cognitive options that early hominins could have taken (e.g., [93]) or whether there exist distinct behaviors/abilities between humans and other species.

However, gaps exist between language-like expressions acquired and used by animals and real languages used by humans. For example, birdsongs or animal calls lack rich semantics compared to language. There are also differences between animals' behaviors and their inner representations of those behaviors. Accordingly, single experiments provide insufficient evidence to know exactly whether and how animals distinguish the acclaimed center-embedding structures from others, thus leading to discrepancy, as shown in the target article and some commentaries. A recent attempt [94] has conducted a series of experiments to demonstrate that monkeys, previously shown unable to grasp language-like hierarchical structures [95], can be trained to generalize the learned grammar to novel sequences, even longer ones, and produce hierarchical sequences following a two-level embedding structure. Discussions on the ROH and other similar topics between humans and nonhuman animals would continue if we kept interpreting animal behaviors purely from a human angle. In addition, as pointed out in the target article, research in this line should not be limited to recursion and relevant hypotheses. Imitation or statistical learning is also important for humans to acquire and develop different language structures, and degree-differences could exist in these abilities between humans and other animals. Furthermore, linking language with its socio-cultural and ecological environment, studying different social abilities between humans and other animals can also shed important light on why only our species is gifted with language.

Genetics research is based on the assumption that genotype (genetic constitution of a cell, organism, or individual) determines phenotype (observable characteristics of an organism). It is often conducted via large scale comparison, family aggregation, and twin studies, with the view to identifying particular genes that cause different performances [96].

However, lacking sufficient affected individuals restrains us from pinpointing the responsible genetic loci. In addition, correlations between the human genotype and cognitive capacities are usually indirect, irreducible to simple correspondences between a single gene, or gene sequences and a specific cognitive repertoire [51]. Therefore, studying the roles of particular genes in language development needs to consider findings in other disciplines such as psychology, physiology, and cognitive sciences [45].

Neuroimaging studies are often designed based on a violation framework (comparing neural responses in processing normal language instances with phonologically, semantically, or syntactically incongruent instances, or comparing neural responses in linguistic tasks with nonlinguistic tasks). In this way, temporal components corresponding to various levels of language processing can be discovered using electroencephalograms (EEG) (e.g., [97,98]), and brain regions responsible for both language and nonlinguistic processing or coordinating different types of language processing can be located using functional Magnetic Resonance Imaging (fMRI) or functional Near-Infrared Spectroscopy (fNIRS) (e.g., [26–32]).

However, EEG has a low spatial resolution and insensitivity to non-time-locked events, fMRI has a poor temporal resolution, fNIRS can only trace surface activations of the brain, and both fMRI and fNIRS make use of indirect indicators (change in the blood oxygen level and ratio between oxygenated and deoxygenated blood cells) to detect possible brain activities. More importantly, it is rather difficult to design neuroimaging experiments that can efficiently prune irrelevant responses, highlight responses of specific behaviors, temporally lock and spatially locate linguistic and nonlinguistic activities, and systematically compare the relevant activations. These difficulties often result in confusing results and debate on whether or not certain brain regions or networks are selectively engaged in complex functions like language.

Experimental studies based on human and/or nonhuman animal (e.g., monkeys or chimpanzees) participants (e.g., experimental semiotics studies involving or not involving a language-like communication system, or artificial language learning experiments) are suitable to examine the effects of individual cognitive, socio-cultural, and ecological factors on language learning and evolution, as well as identifying the minimal conditions and motivations to trigger a language-like communication system. As pointed out by some commentators, such experiments can inform us of the prerequisites for a language-like communication system to emerge and the characteristics that such system could possess.

However, the primary difficulty of such approach also lies in experimental design that can efficiently exclude individuals' prior knowledge and reasonably replicate the environment or social setting of language evolution. Most existing experiments use human participants as cognitively sophisticated black boxes. Other approaches such as computational modeling can help overcome this limitation by swapping human participants with carefully designed artificial agents.

Computational modeling, as operational hypotheses or theories expressed in computer programs, have recently been accepted by many researchers in evolutionary linguistics. These models can explore rich conditions and incorporate various components relevant for language evolution. Some conditions resemble the real history of language evolution, and others allow evaluation of causal factors for language evolution and interactions of linguistic and nonlinguistic components during language learning and use. In addition, models help validate incorporated theories, by illustrating how the explanatory mechanisms or theoretical constructs interact and by comparing model predictions with empirical data. Furthermore, modeling can recapitulate millions of years of evolution within a small amount of time, or magnify instant processes of language processing into a series of traceable, intermediate steps, which allow investigating language evolution and processing at various timescales.

Despite the abovementioned advantages and complementary roles to other approaches, computational modeling inevitably involves simplification, specification, and arbitrary decisions. So far, none of the existing models are within the reach of the complexity of human language or language processing. In addition, simulations only illustrate what could have happened, not what must have happened, and they cannot verify the built-in assumptions [99]. Furthermore, comparison with the relevant empirical data might sometimes be impossible, especially for models of language origin.

Arbib raised a series of criticisms on the simulations of socio-cultural evolution of language, such as: (a) such models skim over the question of how biocultural evolution could have yielded creatures that had the stated abilities and mechanisms; and (b) such simulations restrict the range of inputs and outputs, thus avoiding the challenge of how creatures in natural environments might evolve to attend to the stated variables. These criticisms are in line with the general limitations of modeling, such as the simplification in simulated language, specification by not covering all aspects of evolution, and arbitrary decisions in built-in assumptions to explore their effects but unable to verify the validity of those assumptions. Recent simulations have begun to incorporate both cultural and biological evolution to seek their collective and separate roles in shaping not only language system but also human learning abilities [43, 44]. Other computational approaches, such as the comparative (neuro)primatology work by Arbib [100,101], could complement the available simulations on socio-cultural evolution of language, yet those studies also face similar limitations of modeling studies in general (e.g., focusing too much on individual cognition but ignoring socio-cultural influences). Furthermore, research on socio-cognition and pragmatic inference in non-human primates could provide useful clues on understanding the precursors to language and the degree differences of these precursors between humans and other animals [102].

The above reiteration of the assumptions, advantages, difficulties, and limitations of different disciplines could help researchers understand studies based on different disciplines and guide selection of appropriate approaches in specific research of language evolution and foundations.

The last but not the least, aspect (c) is important as well. Misusing an adopted approach and/or misinterpreting the results obtained would lead to wrong conclusions. For instance, a series of modeling studies [103] based on Reali & Christiansen's simple recurrent network [61] have been conducted to demonstrate that the transitional probability based statistical approach is actually insufficient to resolve the auxiliary reversion question; once many verbs and auxiliaries start to share the same word forms, the simple n-gram model fails to clarify the correct or incorrect displacement of relevant words in grammatical questions. These results have been interpreted as evidence that the auxiliary inversion problem cannot be resolved by a simple statistical approach, and therefore, there must exist some innate knowledge for children to correctly resolve this problem [62].

However, due to involving homophonic word forms, simple recurrent models could not build up lexical categories to correctly distinguish homophonic verbs and auxiliaries. In other words, using simple recurrent models in this situation is a misuse of the approach. In addition, due to inability to correctly distinguish lexical categories, simple recurrent models can by no means tell which ways of displacement of relevant words match the transitional probabilities, thus failing to judge the grammaticality of sentences. This does not mean that the simple recurrent models or the transitional probability based statistical approach cannot resolve the auxiliary reversion question, because the foundation of this approach (the appropriate lexical categories) is missing in this situation. Therefore, the conclusion summarized, based on those studies, appears to be a misinterpretation of the simulation results. Nonetheless, both the authors of the original paper and the follow-up studies agree that a complete model of language acquisition cannot be developed on the basis of this distributional cue (transitional probabilities) alone; additional sources of information (semantic, phonological, prosodic) are needed for young learners to infer different aspects of the structure of the target language to be learned. Moreover, this reflects another limitation of modeling studies: different models are usually designed to address specific hypotheses or theory at a certain level, not easily scaled-up or directly adoptable to address other similar problems. In this sense, how to reasonably unify the findings of different models working on different aspects of language evolution in order to formulate a more complete theory or hypothesis about language foundations and evolution is a difficult but very important task faced by many modelers.

There could be other studies that occasionally misuse multidisciplinary approaches and/or misinterpret the obtained results. In order to correctly digest the multidisciplinary research and use appropriate findings to evaluate hypotheses or theories of language foundations and evolution, scholars who are less familiar with different approaches should not only pay attention to the results of various multidisciplinary studies but also understand the adopted approaches, at least to a certain extent.

6. Other fields and attitudes towards research of language foundations and evolution

Unlike language learning and use, WSY Wang points out that language degradation in terms of gradual/sudden breakdown of language during ageing [104] has been largely ignored in modern academic research, even though various forms of linguistic breakdown have been observed by Lenneberg and other 19th century scientists [105]. Research in this line is important to understand language foundations, and urgently needed for effectively intervening, the latter of which serves as a fruitful application of research on language foundations and evolution, especially in societies with increasingly many elderly people such as China. Decline of language is usually accompanied by decline of cognitive abilities, but we are not clear how the two aspects are correlated with each other and whether and how intervening in one aspect could influence the development of the other (e.g., being bilingual could reduce cognitive decline [106]). In line with the other directions of the research on language foundations and evolution, the perspectives of complex adaptive system and the network, connectionist view are necessary in answering the abovementioned questions, and we foresee that multidisciplinary approaches from neuroscience, genetics, and simulation would benefit the future research in this direction.

Apart from discussing recent findings in multidisciplinary studies of language foundations and evolution, Ansaldo suggests renovating future linguistics training by incorporating three courses: (a) LDD, which focuses on tracing linguistic diversity evident in empirical data; (b) evolutionary biology, which makes linguistics scientifically embedded, rather than merely metaphorically borrowing terminology at a superficial level, or even misinterpreting or misusing those terms; and (c) social sciences, which link languages with users' ways of thinking about the world, their specific cultures and histories. This is consistent with our multidisciplinary perspective on studies of language foundations and evolution [79]. In addition, we suggest that LDD should not be limited to description and documentation. Based on other approaches such as simulation or statistical analyses, we are able to interpret and predict the dynamics of language contact and change through synchronic and diachronic documentation of real world languages (e.g., [69,107,108]). Furthermore, we suggest using a mixed curriculum of humanities, social, and natural sciences to cultivate prospective students in evolutionary linguistics.

Apart from the abovementioned disciplines, Liang and colleagues [109] suggest that interpreting an oral language code translational procedure consisting of concurrent comprehension of a spoken message in a source language and reformulation of a message into a target language, could serve as a mirror for language foundations. Interpreting and language foundations share a number of common features: interpreting is a process in which cognitive processing and language use interact; attention-sharing and working memory are the bases for interpreting; interpreting training coevolves with domain-general cognitive functions; and different types of interpreting reflect integrated effects of socio-cultural environments and domain-general abilities on language production. Therefore, materials of various interpreting types could offer quantitative evidence for language production in different occasions of usage.

7. Conclusions

This reply addresses a variety of comments raised by the commentators. We first reiterate our positions in the target article that: (1) biological foundations of language could not be language-specific; (2) socio-cultural transmissions lay another equally important foundation for language; and (3) biological and cultural evolution are intertwined. Then, we argue against some of the challenges on our positions, and three misleading views on language foundations and evolution, including: (1) language is determined by domain-specific UG, a single (or a small set of) gene(s), or a language-specific acquisition device in the brain; (2) language is primarily an internal thinking tool, not for communication; and (3) externalization of language in use is of little interest. After that, we advocate the research studying the influence of socio-cultural, ecological, and demographical factors on language foundations and evolution, and the studies on multiple critical periods for acquiring different domain-general skills to learn and use various components of language. Later on, we stress three critical steps of conducting multidisciplinary research on language foundations and evolution, including: (1) carefully setting up the research questions; (2) comprehensively understanding the adopted multidisciplinary approach; and (3) properly applying the approach and interpreting the findings. We also summarize the major advantages and critical limitations of widely-adopted multidisciplinary approaches in evolutionary linguistics, and clarify some relevant terms, misuses of multidisciplinary approaches, and misinterpretations of obtained results. Finally, we list a number of fruitful extensions pointed out by some commentators on future research on language foundations and evolution, including language degradation during ageing, language documentation and description, and oral interpretation. We hope that this reply, albeit short, can provide useful guidance to open-minded linguists and keen scholars from other disciplines to conduct promising research exploring language foundations and evolution.

Acknowledgements

This work is supported in part by the National Social Science Foundation of China Project (No. 18BYY177), Zhejiang University, and the MOE Project of the Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies. We thank Keelan Evanini from Educational Testing Service, Morten H. Christiansen from Cornell University and P. Thomas Schoenemann from Indiana University for comments.

References

[1] Comrie B. The biological and the social in the origin and development of language. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by Tao Gong, Lan Shuai, Yicheng Wu. Phys Life Rev 2018;26–27:142–3. https://doi.org/10.1016/j.plrev.2018.06.001 [in this issue].

[2] Wang Q. Rethinking foundations of "Rethinking foundations of language from a multidisciplinary perspective" by T. Gong et al. Phys Life Rev 2018;26–27:167–72. https://doi.org/10.1016/j.plrev.2018.07.004 [in this issue].

[3] Corballis MC. Space, time, and language. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by Tao Gong, Lan Shuai and Yicheng Wu. Phys Life Rev 2018;26–27:144–5. https://doi.org/10.1016/j.plrev.2018.06.009 [in this issue].

[4] Arbib M. Biology matters. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by Tao Gong, Lan Shuai, Yicheng Wu. Phys Life Rev 2018;26–27:176–8. http://dx.doi.org/10.1016/j.plrev.2018.07.005 [in this issue].

[5] Chu H. Foundations of language and the recursion-only hypothesis. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by Tao Gong, Lan Shuai and Yicheng Wu. Phys Life Rev 2018;26–27:160–3. http://dx.doi.org/10.1016/j.plrev.2018.07.002 [in this issue].

[6] Kellens PC, Dale R, Christiansen MH. Linguistic diversity and individual variation: comment of "Rethinking foundations of language from a multidisciplinary perspective" by T. Gong et al. Phys Life Rev 2018;26–27:164–6. https://doi.org/10.1016/j.plrev.2018.07.003 [in this issue].

[7] Roberts S. What are the social, economic and ecological conditions for the evolution of complex communication systems? Comment on "Rethinking foundations of language from a multidisciplinary perspective" by T. Gong et al. Phys Life Rev 2018;26–27:152–4. https://doi.org/10.1016/j.plrev.2018.06.016 [in this issue].

[8] Mufwene SS. "Socio-cultural transmission" in language evolution? Comment on "Rethinking foundations of language from a multidisciplinary perspective" by Tao Gong, Lan Shuai and Yicheng Wu. Phys Life Rev 2018;26–27:155–7. https://doi.org/10.1016/j.plrev.2018.06.015 [in this issue].

[9] Zhang M. Diversity of language structure is shaped by demographic activities. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by T. Gong et al. Phys Life Rev 2018;26–27:146–8. http://dx.doi.org/10.1016/j.plrev.2018.06.007 [in this issue].

[10] Coupé C. Toward a unified view on language evolution and linguistic change. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by T. Gong et al. Phys Life Rev 2018;26–27:173–5. https://doi.org/10.1016/j.plrev.2018.07.006 [in this issue].

[11] Liu H. Language as a human-driven complex adaptive system. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by T. Gong et al. Phys Life Rev 2018;26–27:149–51. https://doi.org/10.1016/j.plrev.2018.06.006 [in this issue].

[12] Ansaldo U. The evolution of linguistic theory and the new 'core'. Phys Life Rev 2018;26–27:158–9. http://dx.doi.org/10.1016/j.plrev.2018.06.008 [in this issue].

[13] Dediu D, Ladd DR. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. Proc Natl Acad Sci USA 2007;104(26):10944–9.

[14] Ladd DR, Dediu D, Kinsella AR. Languages and genes: reflections on biolinguistics and the nature-nurture question. Biolinguistics 2008;2(2–3):114–26.

[15] Bowles JW. Some questions about determining causal inference and criteria for evidence. Response to Ladd, Dediu & Kinsella (2008). Biolinguistics 2008;2(2–3):247–55.

[16] Ladd DR, Dediu D, Kinsella AR. Reply to Bowles (2008). Biolinguistics 2008;2(2–3):256–9.

[17] Arbib MA. How the brain got language: the mirror system hypothesis. Oxford: Oxford University Press; 2012.

[18] Tomasello M. Origins of human communication. Cambridge, MA: The MIT Press; 2008.

[19] Corballis M. From hand to mouth: the origins of language. Princeton, NJ: Princeton University Press; 2002.

[20] Corballis M. Mental time travel: a case for evolutionary continuity. Trends Cogn Sci 2013;17(1):5–6.

[21] Berwick RC, Chomsky N. Why only us? Language and evolution. Cambridge, MA: The MIT Press; 2017.

[22] Dryer MS, Haspelmath M, editors. The world atlas of language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology; 2013. Available online at http://wals.info.

[23] Maddieson I. Uvular consonants. In: Dryer MS, Haspelmath M, editors. The world atlas of language structures online. Leipzig: Max Planck Institute for Evolutionary Anthropology; 2013. Chapter 6. Available online at http://wals.info/feature/6.

[24] Beckner C, Blythe R, Bybee J, Christiansen MH, Croft W, Ellis NC, et al. Language is a complex adaptive system: position paper. Lang Learn 2009;59(Suppl. 1):1–26.

[25] Steels L. Language as a complex adaptive system. In: Schoenauer M, Deb K, Rudolph G, Yao X, Lutton E, editors. Proceedings of the 6th international conference on parallel problem solving from nature. Berlin: Springer-Verlag; 2000. p. 17–26.

[26] Fedorenko E, Behr MK, Kanwisher N. Functional specificity for high-level linguistic processing in the human brain. Proc Natl Acad Sci USA 2017;108(39):16428–33.

[27] Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: anatomy, function, and relevance to disease. Ann NY Acad Sci 2008;1124(1):1–38.

[28] Monchi O, Petrides P, Petre V, Worsley K, Dagher A. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. J Neurosci 2001;21(19):7733–41.

[29] Fadiga L, Craighero L, D'Ausilio A. Broca's area in language, action, and music. Ann NY Acad Sci 2009;1169(1):448–58.

[30] Grodzinsky Y, Santi A. The battle for Broca's region. Trends Cogn Sci 2008;12:474–80.
[31] Schell M, Zaccarella E, Friederici AD. Differential cortical contribution of syntax and semantics: an FMRI study on two-word phrasal processing. Cortex 2017;96:105–20.
[32] Friederici AD, Chomsky N, Berwick RC, Moro A, Bolhuis JJ. Language, mind and brain. Nat Hum Behav 2017;1:713–22.
[33] Mesoudi A, Whiten A. The multiple roles of cultural transmission experiments in understanding human cultural evolution. Philos Trans R Soc B 2008;363:3489–501.
[34] Gong T. Exploring the roles of horizontal, vertical, and oblique transmissions in language evolution. Adapt Behav 2010;18(3–4):356–76.
[35] Scott-Phillips SC, Kirby S. Language evolution in the laboratory. Trends Cogn Sci 2010;14(9):411–7.
[36] Christiansen MH, Kirby S. Language evolution: consensus and controversies. Trends Cogn Sci 2003;7(7):300–7.
[37] Wang WSY. Explorations in language evolution. Osmania Papers in Linguistics, 8. Reprinted in: Wang WSY, editor. Explorations in language. Taipei, Taiwan; Seattle, WA: Pyramid Press, 1982. p. 105–131.
[38] Fitch WT. The evolution of language. Cambridge: Cambridge University Press; 2010.
[39] Bickerton D. Adam's tongue: how humans made language, how language made humans. New York: Hill and Wang; 2009.
[40] Jackendoff R. Foundations of language: brain, meaning, grammar, evolution. Oxford: Oxford University Press; 2002.
[41] Wray A. Formulaic language and the lexicon. New York: Cambridge University Press; 2002.
[42] Mithen S. The singing Neanderthals: the origins of music, language, mind and body. London: Weidenfeld & Nicolson; 2005.
[43] Gong T, Shuai L. Modelling the coevolution of joint attention and language. Proc - Royal Soc, Biol Sci 2012;279(1747):4643–51.
[44] Gong T, Shuai L. Modeling coevolution of language and memory capacity during language origin. PLoS ONE 2015;10(11):e0142281.
[45] Fisher SE, Scharff C. FOXP2 as a molecular window into speech and language. Trends Genet 2009;25(4):66–177.
[46] Mufwene SS, Coupé C, Pellegrino F. Complexity in language: a multifaceted phenomenon. In: Mufwene SS, Coupé C, Pellegrino F, editors. Complexity in language: developmental and evolutionary perspectives. Cambridge: Cambridge University Press; 2017. p. 1–29.
[47] Corballis MC. The recursive mind: the origins of human language, thought, and civilization. Princeton: Princeton University Press; 2011.
[48] Sperber D. Metarepresentations in an evolutionary perspective. In: Sperber D, editor. Metarepresentations: an interdisciplinary perspective. Oxford: Oxford University Press; 2000. p. 117–37.
[49] Hauser MD, Chomsky N, Fitch WT. The faculty of language: what is it, who has it, and how did it evolve? Science 2002;298:1569–79.
[50] Toya G, Hashimoto T. Recursive combination has adaptability in diversifiability of production and material culture. Front Psychol 2018. https://doi.org/10.3389/fpsyg.2018.01512 [in press].
[51] Clark A, Lappin S. Linguistic nativism and the poverty of the stimulus. Malden, MA: Wiley–Blackwell; 2011.
[52] Jacob F. Evolution and tinkering. Science 1977;196:1161–6.
[53] Shuai L, Gong T, Wu Y. Who is who? Interpretation of multiple occurrence of the Chinese reflexive: evidence from real-time sentence processing. PLoS ONE 2013;8(9):e73226.
[54] Yang X, Wu Y. Whether or not multiple occurrences of *ziji* can take distinct antecedents: a deictic perspective. J Pragmat 2015;87:142–55.
[55] Huang J, Li A, Li Y. The syntax of Chinese. Cambridge: Cambridge University Press; 2009.
[56] Martins M, Fitch W. Investigating recursion within a domain-specific framework. In: Lowenthal F, Lefebvre L, editors. Language and recursion. Berlin: Springer; 2014. p. 15–26.
[57] Roeper T. The acquisition of recursion: how formalism articulates the child's path. Biolinguistics 2011;5:57–86.
[58] Yang C. The recursion-only hypothesis and its evidence – an empirical study based on children's center-embedded ability development. Forn Lang Teach Res 2014;46(6):817–29.
[59] Chomsky N. Aspects of the theory of syntax. Cambridge, MA: MIT Press; 1965.
[60] Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science 1996;274:1926–8.
[61] Reali F, Christiansen M. Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. Cogn Sci 2005;29:1007–28.
[62] Berwick R, Pietroski P, Yankama B, Chomsky N. Poverty of the stimulus revisited. Cogn Sci 2011;35:1–36.
[63] Becker-Kristal R. Acoustic typology of vowel inventories and dispersion theory: insights from a large cross-linguistic corpus. Los Angeles, CA: University of California; 2010.
[64] Greenhill SJ, Blust R, Gray RD. The Austronesian basic vocabulary database: from bioinformatics to lexomics. Evol Bioinform 2008;4:271–83.
[65] Christiansen MH, Chater N. Creating language: integrating evolution, acquisition, and processing. Cambridge, MA: The MIT Press; 2016.
[66] Kidd E, Donnelly S, Christiansen MH. Individual differences in language acquisition and processing. Trends Cogn Sci 2018;22(2):154–69.
[67] Reali F, Chater N, Christiansen MH. Simpler grammar, larger vocabulary: how population size affects language. Proc R Soc B, Biol Sci 2018;285(1871):20172586.
[68] Regier T, Carstensen A, Kemp C. Languages support efficient communication about the environment: words for snow revisited. PLoS ONE 2016;11(4):e0151138.
[69] Zhang M, Gong T. Principles of parametric estimation in modeling language competition. Proc Natl Acad Sci USA 2013;110(24):9698–703.
[70] Nichols J, Witzlack-Makarevich A, Bickel B. The AUTOTYP genealogy and geography database: 2013 release. Zurich: University of Zurich; 2013.
[71] Hassan FA. Demographic archaeology. New York: Academic Press; 1981.
[72] Wang WSY. Critical periods for language. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by T. Gong et al. Phys Life Rev 2018;26–27:179–83. http://dx.doi.org/10.1016/j.plrev.2018.09.002 [in this issue].
[73] Penfield W, Roberts L. Speech and brain mechanisms. Princeton, NJ: Princeton University Press; 1959.
[74] Kuhl PK, Conboy BT, Coffey-Corina S, Padden D, Rivera-Gaxiola M, Nelson T. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). Phil Trans R Soc B 2008;363:979–1000.

[75] Keven N, Akins K. Neonatal imitation in context: sensory-motor development in the perinatal period. Behav Brain Sci 2017;40:e381.
[76] Hartshorne JK, Tenenbaum JB, Pinker S. A critical period for second language acquisition: evidence from 2/3 million English speakers. Cognition 2018;177:263–77.
[77] Li P, Zhao X, MacWhinney B. Dynamic self-organization and early lexical development in children. Cogn Sci 2007;31:581–612.
[78] McMurray B. Defusing the childhood vocabulary explosion. Science 2007;317:631.
[79] Gong T, Shuai L, Comrie B. Evolutionary linguistics: theory of language in an interdisciplinary space. Lang Sci 2014;41:243–53.
[80] Liang J, Liu H. Interdisciplinary studies of linguistics: language universals, human cognition and big-data analysis. J Zhejiang Univ (Hum Soc Sci) 2016:108–18.
[81] Roberts S. CHIELD: causal hypotheses in evolutionary linguistics database. In: Cuskley C, Flaherty M, Little H, McCrohon L, Ravignani A, Verhoef T, editors. The evolution of language: proceedings of the 12th international conference (EVOLANGXII). NCU Press; 2018. Available online at http://evolang.org/torun/proceedings/papertemplate.html?p=7.
[82] Cong J, Liu H. Approaching human language with complex networks. Phys Life Rev 2014;11(4):598–618.
[83] Liu H, Xu C, Liang J. Dependency distance: a new perspective on syntactic patterns in natural languages. Phys Life Rev 2017:171–93.
[84] Lu Q, Xu C, Liu H. Can chunking reduce syntactic complexity of natural languages? Complexity 2016;21(S1):33–41.
[85] Gong T, Shuai L. Computer simulation as a scientific approach in evolutionary linguistics. Lang Sci 2013;40:12–23.
[86] Gong T, Shuai L, Wu Y. Multidisciplinary approaches in evolutionary linguistics. Lang Sci 2013;37:1–13.
[87] Gong T, Shuai L, Zhang M. Modeling language evolution: examples and predictions. Phys Life Rev 2014;11(2):280–302.
[88] Gong T, Shuai L, Zhang M. Key issues for the prosperity of modelling research of language evolution. Phys Life Rev 2014;11:324–8.
[89] Clay Z, Zuberbühler K. Food-associated calling sequences in bonobos. Anim Behav 2009;77(6):1387–96.
[90] Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC. Recursive syntactic pattern learning by songbirds. Nature 2006;440:1204–7.
[91] Savage-Rumbaugh S, Shanker SG, Taylor TJ. Apes, language, and the human mind. Oxford: Oxford University Press; 1998.
[92] Oller DK, Griebel U, editors. Evolution of communication systems: a comparative approach. Cambridge, MA: The MIT Press; 2004.
[93] Hurford J. Origins of meaning. Oxford: Oxford University Press; 2007.
[94] Jiang X, Long T, Cao W, Li J, Dehaene S, Wang L. Production of supra-regular spatial sequences by macaque monkeys. Curr Biol 2018:1–29.
[95] Fitch WT, Hauser MD. Computational constraints on syntactic processing in a nonhuman primate. Science 2004;303(5656):377–80.
[96] Stromswold K. Genetics and the evolution of language: what genetic studies reveal about the evolution of language. In: Larson RK, Déprez V, Yamakido H, editors. The evolution of human language: biolinguistic perspective. Cambridge: Cambridge University Press; 2009. p. 176–90.
[97] Shuai L, Gong T. Temporal relation between top-down and bottom-up processing in lexical tone perception. Front Behav Neurosci 2014;8. article 97.
[98] Hong T, Shuai L, Frost SJ, Landi N, Pugh KR, Shu H. Cortical responses to Chinese phonemes in preschoolers predict their literacy skills at school age. Dev Neuropsychol 2018;43(4):356–69.
[99] Jaeger H, Steels L, Baronchelli A, Briscoe T, Christiansen MH, Griffiths TL, et al. What can mathematical, computational and robotic models tell us about the origins of syntax? In: Bickerton D, Szathmáry E, editors. Biological foundations and origins of syntax. Cambridge, MA: The MIT Press; 2009. p. 385–410.
[100] Arbib MA. Computational Challenges of evolving the language-ready brain, 1: from manual action to protosign. Interact Stud 2018;19(1–2):7–21.
[101] Arbib MA. Computational Challenges of evolving the language-ready brain, 2: building towards neurolinguistics. Interact Stud 2018;19(1–2):22–37.
[102] Seyfarth RM, Cheney DL. Precursors to language: social cognition and pragmatic inference in primates. Psychon Bull Rev 2017;24(1):79–84.
[103] Kam XC, Stoyneshka I, Tornyova L, Fodor JD, Sakas WG. Bigrams and the richness of the stimulus. Cogn Sci 2008;32:771–87.
[104] Wang WSY. Language and the brain in the subset years. In: Huang C-R, Jing-Schmidt Z, Meisterernst B, editors. The Routledge handbook of Chinese applied linguistics. London: Routledge; 2018 [in press].
[105] Lenneberg EH. Understanding language without ability to speak. J Abnorm Soc Psychol 1962;65(6):419–25.
[106] Anderson JAE, Grundy JG, De Frutos J, Barker RM, Grady C, Bialystok E. Effects of bilingualism on white matter integrity in older adults. NeuroImage 2018;167:143–50.
[107] Chirkova K, Gong T. Simulating vowel chain shift in Xumi. Lingua 2014;152:65–80.
[108] Chirkova K, Gong T. Modeling change in contact settings: a case study of phonological convergence. Lang Dyn Change 2018 [in press].
[109] Liang J, Lv Q, Liu Y. Interpreting as a mirror for language foundations. Comment on "Rethinking foundations of language from a multidisciplinary perspective" by Tao Gong, Lan Shuai, Yicheng Wu. Phys Life Rev 2018;26–27:139–41. http://dx.doi.org/10.1016/j.plrev.2018.06.002 [in this issue].