

# Language Neuromechanics: The Human Biological-Language Evolution

Dingyu Chung

Utica, MI, USA

Email: dy\_chung@yahoo.com

**How to cite this paper:** Chung, D.Y. (2018) Language Neuromechanics: The Human Biological-Language Evolution. *Journal of Behavioral and Brain Science*, 8, 447-472. <https://doi.org/10.4236/jbbs.2018.88028>

**Received:** July 4, 2018

**Accepted:** August 5, 2018

**Published:** August 8, 2018

Copyright © 2018 by author and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The paper proposes that the understanding of human language evolution requires the comprehensive understanding of language in terms of language types, formations, and learnings and the comprehensive understanding of human biological evolution in terms of the emergences of various hominin species with various language capacities. This paper proposes language neuromechanics and the human biological-language evolution. Language is derived from bodily movement. Language neuromechanics combines neuroscience to study language brain and biomechanics to study language movement. Language neuromechanics consists of language type, language formation, and language learning. Language types for advanced animals include gestural language verse vocal language, instinctive language verse controllable language, and symbolic language verse iconic language. Language formation involves the developments of the different types of languages from different bodily movements phylogenetically and ontogenetically. Language learning involves the learning of controllable language to adapt to communicative environment through language brain regions and language genes. This paper proposes a gradual and step-by-step human language evolution from the language of great apes to the human language through the human biological evolution which chronologically and geographically consists of early hominins, early Homos, middle Homos, and late Homos with different language capacities. For hominins, vocal language and gestural language were evolved together. In conclusion, combining neuroscience and biomechanics, language neuromechanics provides the comprehensive understanding of language. The combination of language neuromechanics and the human biological-language evolution provides the clear evolutionary path from great apes' articulate gestural language without articulate speech to human articulate gestural language and articulate speech.

## Keywords

Language Neuromechanics, Human Biological Evolution, Human

---

Language Evolution, Vocal Language, Gestural Language, Instinctive Language, Controllable Language, Symbolic Language, Iconic Language, Language Brain, Language Genes, Great Apes, Hominins

---

## 1. Introduction

In the last 45 years, there has been an extensive study in human language evolution, but human language evolution still remains as a mystery [1]. For an example, the clear evolutionary path from great apes' articulate gestural language without articulate speech to human articulate gestural language and articulate speech is still unknown. The paper proposes that the understanding of human language evolution requires the comprehensive understandings of language in terms of language types, formations, and learnings and the comprehensive understanding of human biological evolution in terms of the emergences of various hominin species with various language capacities.

Language is a communication type. Communication transmits signals and coordinates actions among communicators. Language is defined as a communication medium capable of expressing the entire communicative needs of an animal society with a similar view and background [2]. Language coordinates actions among different members of an animal society with a similar view and background. Language is originally derived from bodily movements. Language neuromechanics combines neuroscience to study language brain and biomechanics to study language movement. Neuromechanics is the field of study that combines neuroscience and biomechanics in an effort to understand movement and its relationship with the brain [3].

This paper proposes that language neuromechanics consists of language type, language formation, and language learning [4]. Language types for advanced animals include gestural language verse vocal language, instinctive language verse controllable language, and symbolic language verse iconic language. Language formation involves the developments of the different types of languages from different bodily movements phylogenetically and ontogenetically. Language learning involves the learning of controllable language to adapt to communicative environment through language brain regions and language genes.

The human language was evolved from great apes' languages that use mostly articulate gestural language in addition to basically instinctive inarticulate vocal language. On the other hand, human language consists of articulate gestural language and articulate vocal language (speech) sharply different from great apes' languages. This paper proposes a gradual and step-by-step human language evolution from the language of great apes to the human language through the human biological evolution which chronologically and geographically consists of early hominins, early Homos, middle Homos, and late Homos with different language capacities. In terms of language capability, great apes have intermediate gestural language and primitive vocal language. Early hominins and

early Homos have intermediate gestural language and intermediate vocal language. Middle Homos has advanced gestural language and intermediate vocal language. Late Homos have advanced gestural language and advanced vocal language. For hominins, vocal language and gestural language were evolved together as suggested by Adam Kendon [5]. The paper will discuss language neuromechanics in Section 2 and the human biological-language evolution in Section 3.

## 2. Language Neuromechanics

In this paper, language neuromechanics involves language type, language formation, and language learning. Language types for advanced animals include gestural language verse vocal language, instinctive language verse controllable language, and symbolic language verse iconic language. Language formation involves the developments of the different types of languages from different bodily movements phylogenetically and ontogenetically. Language learning involves the learning of controllable language to adapt to communicative environment through language brain regions and language genes.

### 2.1. Language Type

Language is derived from bodily movements through synaptic connections. Bodily movements produce visual actions perceived by vision and auditory actions perceived by hearing. Visual action and auditory action eventually generate gestural language and vocal language, respectively. Vocal language is acoustic, and there is a vowel/consonant distinction, and gestural language is formed by constellations of the human body [4]. Most primates have vocal language in terms of a repertoire of vocal calls, but only human and great apes regularly communicate with gestural language. Human and great apes communicate regularly with both vocal language and gestural language, while other primates communicate primarily with vocal language.

Language can be also divided into involuntary instinctive language and controllable ritualized language. Involuntary instinctive language appears at birth or shortly after birth. Instinctive language is the same for all members within a species without the need of learning. At different situations, different instinctive vocal languages (cry, scream, and laughter etc.) and instinctive gestural languages (instinctive facial expressions) appear. Advanced animals have instinctive language at birth and controllable ritualized language upon learning. The controllable ritualized language among the members within a society is exactly the same (ritualization) days after days. The change in the same language occurs very slowly. As a result, the four types of languages are involuntary instinctive gestural language, controllable ritualized gestural language, involuntary instinctive vocal language, and controllable ritualized vocal language.

Primates are born with instinctive vocalized language including crying, laughing, alarming, screaming, and grunting sounds as the instinctive languages

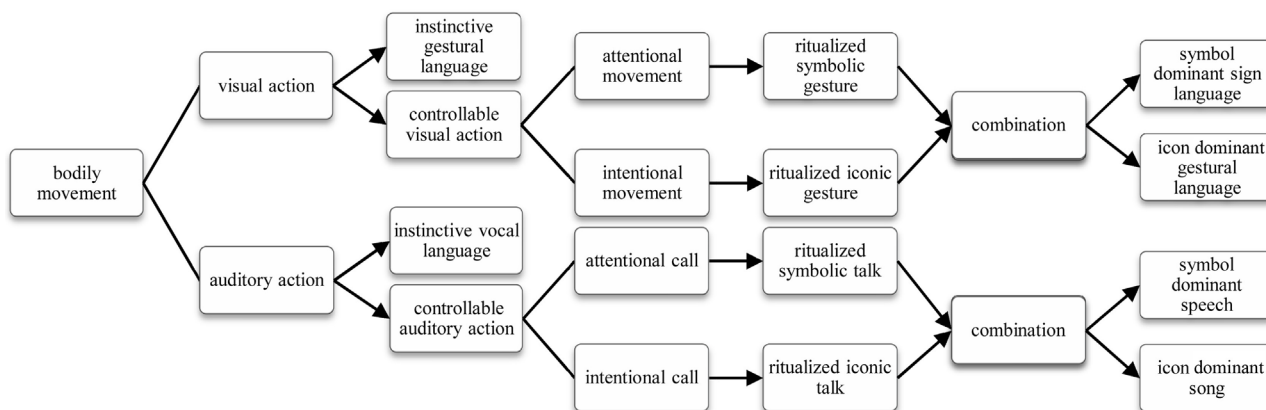
for infant primates. Infant emotional sound production appears to convey generalized meanings, e.g. pleasure and displeasure. Instinctive language is the very basic communication for infant and basic survival, such as the crying associated with social separation. The source of such instinctive language is the subcortex [6] [7]. The subcortex is located below the cerebral cortex, and consists of brainstem (reticular formation, pons, and medulla), midbrain (tectum and tegmentum), and forebrain (basal ganglia, limbic system, thalamus, and hypothalamus) below cerebrum. An infant whose cerebral hemispheres were largely absent and the basal ganglia, and cerebellum and brainstem were present could still cry [8]. When the anterior cingulate gyrus in the limbic system of an anesthetized macaque monkey was stimulated electrically, the macaque monkey made low-pitched guttural sound [9]. The electric stimulation of cerebral cortex does not cause automatic vocalization for instinctive language.

Upon maturation, the neocortex produces controllable ritualized language and controls instinct language. The neocortex is the largest part (90%) of the cerebral cortex, and is involved in higher-order brain functions such as sensory perception, cognition, generation of motor commands, spatial reasoning, and language. The neocortex is the newest part of the cerebral cortex to evolve. For humans, the two controllable language systems are (a) a learned language reception/understanding system (the speech comprehension of learned language) including a core Wernicke's area involved in word recognition and a fringe or peripheral area involved in learned language associations, and (b) a learned language production system (the motor portions of learned language) involved in speaking and grammar in a core Broca's area, some other frontal cortical areas, and subcortical areas [10]. Both the controllable language production and controllable language understanding areas are interconnected through the insula. Damage to the Broca's area in the left hemisphere results in expressive aphasia while damage to Wernicke's area in the left hemisphere results in receptive aphasia.

Furthermore, controllable language can be also divided into symbolic language and iconic language [11]. Symbolic language is based on symbols which by themselves have no meaning and specific purpose. Symbol-referent (meaning of symbol) has arbitrary relationship. Iconic language is based on icons which by themselves represent certain meaning and specific purpose. Icon-referent (meaning of icon) has deliberate relationship. As a result, the four controllable languages are ritualized symbolic gesture, ritualized iconic gesture, ritualized symbolic talk, and ritualized iconic talk. Symbolic languages and iconic languages can be combined into symbol dominant sign language, icon dominant gestural language, symbol dominant speech, and icon dominant song. Consequently, there are six language types including two instinctive language types and four controllable language types.

## 2.2. Language Formation

The language formation of the six languages is described in **Figure 1**.



**Figure 1.** Language formation.

For language, the relevant actions derived from bodily movements through synaptic connections are visual actions perceived by vision and auditory actions perceived by hearing. Some actions constitute involuntary instinctive language, and some actions constitute controllable actions which initially do not involve in language. The degree of controllable actions is determined by the plasticity of actions controlled by the nervous system. Instinctive language needs not to be learned. All members in the same species have the same expressions in instinctive language. Different expressions of instinctive language appear automatically at different proper situations. Instinctive gestural language includes instinctive facial expressions, such as happy face, sad face, surprise fact, and disgusted face. Instinctive vocal language includes cry, scream, and laugh.

Controllable actions can be attentional actions or intentional actions. Attentional actions gain attentions without meaning and specific purpose. Intentional actions have meanings and intentions. Attentional movements from visual actions include random movements of parts of body to gain attentions without meaning and specific purpose. Intentional movements have meanings and specific purpose such as giving, receiving, carrying, eating, reach, walking, running, grasp, and jump. Attentional vocal actions include random calls to gain attention without meaning and specific purpose. Intentional vocal actions include vocal imitations of known meaningful sounds (including instinctive vocal language) and vocal imitations of bodily movements.

Controllable actions can be developed into controllable language by communicators including signalers and recipients [12]. The developmental process from controllable actions into controllable languages is ontogenetic ritualization [13] [14]. Ontogenetic ritualization involves signalers and recipients mutually shaping each other's behavior during the course of repeated interactions. Ontogenetic ritualization is not instinctive (phylogenetic). Attentional actions are developed into ritualized symbols (attention-getters) to gain attention, and intentional actions are developed into ritualized icons (intention signals) to show intention [12].

Controllable attentional movements are developed into ritualized symbolic

gestures through ontogenetic ritualization. The ritualized symbolic gestures for great ape signaler include specific poking and specific throwing stuff (tactile gestures) at recipient to get the attention of intended recipient [12]. Such specific poking and throwing-stuff are the symbols to symbolize referents which are signaler and recipient. The associations of symbol and referent are arbitrary. As a result, symbolic gesture has arbitrary symbol-referent relationship. To be different from random attentional movement, ritualized symbolic gesture between signaler and recipient requires precise movement, recipient, and persistence [13].

Controllable intentional movements are developed into ritualized iconic gestures through ontogenetic ritualization by partially imitating actual intentional actions. The iconic gestures for ape signaler include extending arm toward recipient to ask recipient to give and touching the side of recipient to ask recipient to move. Extending arm toward recipient and touching the side of recipient are the icons to imitate partially actual receiving from recipient and actual pushing recipient away, respectively. The association of icon and referent (begging and moving aside) is deliberate. As a result, iconic gesture has deliberate icon-referent relationship. The iconic gestures imitate only partially the intentional actions, so the arm of the signaler does not need to reach within the distance that allows the recipient to actually give, and the touching on the side of recipient does not actually move the recipient by force. As a result, such iconic gestures are motorically ineffective. To be different from intentional movement, ritualized iconic gesture between signaler and recipient requires motorical inefficiency, recipient, and persistence [13].

The vocal language of nonhuman primates is dominated by involuntary instinctive vocal language. The controllable vocal language of nonhuman primates is very limited particularly in imitating and learning vocal language precisely. According to Fitch, de Boer, Mathur, and Ghazanfar, monkey vocal tracts are speech-ready [15]. The obstacle of monkey to controllable speech is the brain incapability of vocal learning instead of monkey anatomy. According to Tomasello, monkeys and apes have very limited controllable vocal language. In fact, all individuals of the same species share the same basic instinctive vocal language, and monkeys raised by another species do not adopt the calls of their foster species [16]. On the other hand, primates have the full volitional control of the upper limb [16]. Mirror neuron to imitate gestures from other primates helps the learning in gestural language [17]. There is no mirror neuron for vocal language. The iconic gestures are the most natural and intuitive way of expressing conceptualised contents and relations [18]. Great apes can learn human sign language precisely quite well, but cannot learn human vocal language [19]. In terms of the capability in controllable languages for all animals, great apes have primitive controllable vocal language and intermediate controllable gestural language. Generally, great apes use gestural language for less urgent actions than vocal language, which are more involved in vital functions like defense, aggression, reproduction, discovering food and avoiding potential dangers [20]. According

to Goodall, the production of sounds in the absence of the appropriate emotional state derived from the subcortex seems to be an almost impossible task for a chimpanzee [21].

As a result, the clear examples of controllable vocal languages can be found in birdsongs and cetaceans' songs, instead of nonhuman primates' vocal languages [22]. Birds use controllable and learned birdsongs to communicate as courtship songs signaling courtship [23] and as territorial songs establishing territorial boundaries [24]. Territorial birdsongs are ritualized symbolic talks between signalers and recipients who are the outsiders as potential invaders. The signalers warn the recipients (outsiders) to stay away from the signalers' territories. The territorial birdsongs are the symbols symbolizing signalers and recipients. Territorial birdsongs are arbitrary, so symbol-referent has arbitrary relationship. The birds of the same species have different territorial birdsongs in different locations. To be different from random attentional bird call, ritualized symbolic territorial birdsong requires precision, recipient, and persistence.

The equally important purpose of territorial birdsong is to direct disorientated birds to come back to home territory and home group. Highly mobile flying birds are disorientation-prone, so it is necessary for birds to have territorial birdsongs as a vocal territory locator to mark the location of home territory and home group. Highly mobile cetaceans, such as whales, with poor vision are also disorientation-prone, so they also have a vocal territory locator as group whale song to direct disorientated whale to come back to the home group [25]. Another controllable vocal locator in songbirds and cetaceans is vocal mother locator as learned contact call to direct disorientated children to come back to their mothers [26] [27]. Furthermore, bottlenose dolphins have the vocal locators as learned contact calls for the members in a social group [27]. Children learn contact calls from their mothers [28]. Controllable symbolic vocal language starts with vocal locators such as territory locator as territorial song and mother locator as learned contact call. Both vocal locators are arbitrary. Relatively slow moving nonhuman primates with good vision are not disorientation-prone, so controllable vocal territory locators and controllable vocal mother locators are not necessary. There are many other locators, such as odor locator, earth magnetic field locator, landmark locator, and GPS.

Courtship birdsongs are iconic talks to show vocally the physical-mental fitness of signalers to recipients. Recipient birds choose mates from courtship birdsongs. Courtship birdsong is a vocal show. In iconic talks, icon-referent has deliberate relationship. To be different from intentional bird call, iconic courtship birdsong requires the maximum quality, recipient, and persistence. The maximum quality is shown in the maximum numbers, vitality, and symmetry of courtship birdsongs that songbird singers can sing. Songbirds sing differently with and without recipients. Courtship birdsong-referent has deliberate relationship, instead of arbitrary relationship. Cetaceans such as whales also have diverse mating songs as vocal shows like courtship songs in songbirds [29]. For

songbirds and cetaceans, controllable symbolic vocal languages are mostly for arbitrary vocal locators, while controllable iconic vocal languages are mostly for deliberate sexual selection. In terms of the capability of controllable vocal language, songbirds and cetaceans have intermediate vocal language comparing to advanced vocal language for humans and primitive controllable vocal language for nonhuman primates. The animals with intermediate vocal language level include songbirds, cetaceans, pinnipeds, elephants, and bats [30] [31] [32]. They have vocal production learning (VPL) which is the ability to learn to modify vocal outputs in response to auditory feedback [33].

In summary, controllable languages are derived from controllable actions consisting of controllable attentional visual action, controllable intentional visual action, controllable attentional auditory action, and controllable intentional auditory action. For controllable attentional visual action, the basic controllable attentional gestural language is gestural attention getter (poking and throwing stuff) to develop into symbolic gestural language. For controllable intentional visual action, the basic controllable intentional gestural language is gestural intention signal (extending arms for begging) to develop into iconic gestural language. For controllable attentional auditory action, the basic controllable attentional vocal language is vocal locator (territorial birdsong and mother contact call) to develop into symbolic vocal language. For controllable intentional auditory action, the basic controllable intentional vocal language is vocal show (courtship birdsong and mating song) to develop into iconic vocal language as shown in **Table 1**.

The combination of ritualized symbolic gesture and ritualized iconic gesture results in two different gestural languages based on the dominance of symbolic gesture or iconic gesture. The symbol dominant gesture language is symbol dominant sign language such as human sign language. Most words in sign language have arbitrary and precise gestures. Few words have iconic origins. The icon dominant gestural language is the natural gestural language of great apes. Most words are iconic. The combination of ritualized symbolic talk and ritualized iconic talk results in symbol dominant speech and icon dominant song. The words in symbol dominant speech are arbitrary and precise. The iconic

**Table 1.** Controllable language types.

controllable action	controllable basic language	examples	controllable language type
attentional visual action	gestural attention getter	poking, throwing stuff	symbolic gestural language
intentional visual action	gestural intention signal	extending arms for begging	iconic gestural language
attentional auditory action	vocal locator	territorial song, mother contact call	symbolic vocal language
intentional auditory action	vocal show	courtship song, mating song	iconic vocal language



aspect of speech is shown in prosody [34] which are not words, and are expressed by intonation, tone, stress, and rhythm. Prosody reflects various features of speech in terms of the emotional state, form of the utterance (statement, question, or command), and the implied intentions such as irony, sarcasm, emphasis, contrast, and focus etc. The icon dominant song reflects strongly the emotional state of song even without the understanding of the lyrics in a song. The examples and the functions of the six language types are listed in **Table 2**.

### 2.3. Language Learning

The two controllable language learning regions in the brain are (a) linguistic cognitive learning region to process sensory language information including a core Wernicke's area involved in word recognition and a fringe or peripheral area involved in learned language associations, and (b) linguistic motor learning region to process vocal linguistic movement and gestural linguistic movement for the production of language in a core Broca's area, some other frontal cortical areas, and subcortical areas [10]. On the other hand, the three controllable linguistic genes are FOXP2, CNTNAP2, and FOXP1. Both FOXP2 and FOXP1 act to regulate the expression of other genes, determining when and where they are switched on or off [35]. CNTNAP2 is regulated by FOXP2. CNTNAP2 encodes a transmembrane protein (Caspr2) that facilitates clustering of proteins in specific regions of myelinated axons and at synapses [36]. The three linguistic genes involve in synaptic plasticity, which is the ability of connections between neurons (synapses) to change and adapt to experience over time. Synaptic plasticity is necessary for learning and memory. The three linguistic genes (FOXP2, FOXP1, and CNTNAP2) are expressed across many brain regions. In terms of language learning from the mapping the distribution of language related genes FOXP1,

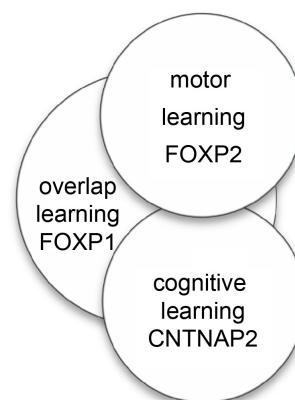
**Table 2.** The six language types.

language type	example	function
instinctive gestural language	instinctive facial expressions	for very basic gestural communication and infant's gestural communication
symbol domination controllable gestural language	human sign language	for basic gestural attention getter and complicate gestural communication
icon domination controllable gestural language	chimpanzee's gestural language	for basic gestural intention signal and simple gestural communication
instinctive vocal language	cry, scream, and laughter	for very basic vocal communication and infant's vocal communication
symbol domination controllable speech	human speech and songbird's territorial song	for basic vocal attention getter and complicate vocal communication
icon domination controllable song	human song and songbird's courtship song	for intentional vocal communication

FOXP2, and CNTNAP2 in the brains of vocal learning bat species [32], FOXP2 is the gene mostly for linguistic motor learning, CNTNAP2 gene is the gene mostly for linguistic cognitive learning, and FOXP1 is the gene for mostly linguistic overlap learning to overlap linguistic motor learning and linguistic cognitive learning. The overlapping between linguistic motor learning from FOXP2 and linguistic cognitive learning from CNTNAP2 is at minimum. FOXP2, FOXP1, and CNTNAP2 constitute the three-functional language learning. The distribution of the expressions for the three-functional language learning is shown in **Figure 2**.

FOXP2 mutations (defects) cause speech disorders shown in the people with the FOXP2 mutations [37]. Speech disorder is a linguistic motor disorder. Mutations in CNTNAP2 can produce speech problems, autistic phenotypes, intellectual disability, and epilepsy [38]. Autistic phenotypes, intellectual disability, and epilepsy are cognitive disorders. The mice without CNTNAP2 had disordered neuronal phenotypes (altered neuronal firing and seizures), improved motor coordination, reduced social interactions, and increased repetitive behavior [39]. The FOXP1 gene is closely related to FOXP2. Unlike individuals with FOXP2 mutations, the individuals with FOXP1 mutations display autism spectrum disorder, mild to moderate intellectual disabilities and motor impairments in addition to speech problem [40]. FOXP1 mutations show the overlapping of both motor and cognition disorders. In the mapping of the distribution of linguistic genes in the bats [32], the expressions of CNTNAP2 often showed an inverse pattern to the expressions of FOXP2. The expressions of FOXP1 located in more brain regions than the expressions of FOXP2.

The human FOXP2 protein has two amino acid substitutions different from the FOXP2 protein from chimpanzees [41]. This human FOXP2 is the leading genetic candidate for human speech and language proficiency [42]. The introduction of the amino acid changes for the human FOXP2 into murine FOXP2 profoundly enhances motor procedural learning in terms of striatal neuroplasticity for the mice [43] [44]. The human FOXP2 evolution likely led to the



**Figure 2.** The distribution of the expressions for the three-functional language learning.

enhancement of vocal motor procedural learning that contributed to adapting the human brain for speech and language acquisition. *Neanderthals* have similar FOXP2 gene as *Homo sapiens* [45]. As a result, *Neanderthals* likely had similar capability for speech and language.

### 3. The Human Biological-Language Evolution

The most important factor in the human biological evolution is habitat. Different habitats produced different hominins. The fossil record and the studies of human and ape DNA indicate that humans shared a common ancestor with chimpanzees and bonobos around 7.5 to 5.6 million years ago. The common ancestors lived in the forest habitat. Since then, the climate slowly cooled, resulting in different habitats at different times and regions by the expansion of woodlands and savannas and the shrinkage of forests in Africa. The five different groups including great apes, early hominins, early Homos, middle Homos, and late Homos lived at the same times or at different times in five different habitats including forest, mixed forest-woodland, mixed woodland-savanna, savanna, and fluctuating savanna, respectively. Great apes, such as chimpanzee and bonobo, live in the forest habitat. Early hominins, such as *Ardipithecus ramidus* and *Australopithecus*, lived in the mixed forest-woodland habitat. Early Homos, such as *Homo habilis* and *Homo naledi*, lived in the mixed woodland-savanna habitat. Middle Homos, such as *Homo erectus*, lived in the savanna habitat originally. Late Homos, such as *Homo sapiens* and *Neanderthals*, lived in the fluctuating savanna originally.

A most important factor in the language evolution is evolutionary pressure [46]. Under evolutionary pressure, an animal has to adopt a new trait, such a new capacity for language, to increase its reproductive success. For an example, relatively slow moving nonhuman primates with good vision have no evolutionary pressure to adopt controllable vocal territory locator, but disorientation-prone fast moving songbird or cetacean under the evolutionary pressure of disorientation had to adopt controllable vocal territory locator (territorial song) to increase its reproductive success. Two other most important factors in the language evolution are linguistic cognitive learning and linguistic motor learning. Linguistic cognitive learning processes information. The complexity of information processing is proportional to the neocortex size, so linguistic cognitive learning is proportional to the brain size. The size of the energy-fat hungry brain is limited by the foods which have to be high calorie and high fat for the large energy-fat hungry brain. For humans, sweet (high calorie) and fatty foods are attractive foods. The complexity of language in terms of information processing is proportional the brain size until the brain reaches the maximum size. Linguistic motor learning is proportional to the synaptic plasticity for linguistic motor learning. The change in linguistic genes to enhance the synaptic plasticity for linguistic motor learning improves the motor production of language.

As described in the previous section, great apes in terms of the capability of

controllable language have intermediate controllable language and primitive controllable vocal language comparing to primitive controllable gestural language and intermediate controllable vocal language. The evolution of human language, therefore, is to reach advanced controllable gestural language and advanced controllable vocal language as in **Table 3**.

### 3.1. From Ape to Early Hominins

Early hominins were evolved in the mixed habitat of dense forest and open woodland [47], where Ardi (*Ardipithecus ramidus*) [48] (4.4 million years ago) lived. The mixed habitat allowed increasingly amount of food from bushes and low branches, which could be seen and reached from the ground in open woodland. According to the observation [49] in Africa, chimpanzees today move on two legs most often when feeding on the ground from bushes and low branches. When food resources are scarce or unpredictable, chimpanzees use upright locomotion to improve food carrying efficiency. The same occurred among the early hominins which adopted bipedalism as the way to move on the ground. Other great apes evolved with different ways other than bipedalism to survive. For orangutan, the feet are much more useful to climb trees in dense rainforest than to walk on the ground, so orangutan did not develop bipedalism for walking. Gorillas, chimpanzees and bonobos did not develop bipedalism, because they needed fast and steady quadrupedal knuckle walking with the knuckle hands to escape from predators and for the large foraging ranges on the ground.

**Table 3.** The capabilities in controllable languages and the evolution of human language.

species	controllable vocal language	controllable gestural language	critical evolutionary change	language evolutionary change
songbirds and cetaceans	intermediate	primitive		
monkeys	primitive	primitive		
great apes	primitive	intermediate		
early hominins and early Homos	intermediate	intermediate	disorientation-prone from the separation of the two groups by division of labor	vocal territory locator to reduce disorientation
middle Homos	intermediate	advanced	hunting meat to increase the brain size for complex information processing	complex gestural language to express complex information processing
late Homos	advanced	advanced	new FOXP2 for vocal language procedural learning to improve vocal language	the emergency of advanced vocal language

The bipedalism of the early hominins evolved before the quadrupedal knuckle walking [50].

However, in Ardi's primitive foot, the fully opposable big toe and the absence of longitudinal arch could not provide a push needed for efficient bipedal walking and running. Its feet were still adapted for grasping trees [51] rather than walking for long distances and running fast on the ground. The movement handicap of bipedalism on the ground was serious for very young, very old, and pregnant females. To the early hominins in the mixed habitats, the forest area with many tall trees was the safe home forest area where very young, very old, and pregnant females stayed, and where they could escape quickly to the safety in tall trees, and the open woodland area with few tall trees was the unsafe exploration area for the exploration to find extra foods that could not be found in the safe home forest area. The two free hands from bipedalism allowed the early hominins to carry a large quantity of food home from the exploration as proposed by C. Owen Lovejoy [52] and to carry simple defensive weapons such as sticks and stones to defense against large predators in the unsafe area. Consequently, the bipedalism and the mixed habitat divided the early hominins into the home forest group who stayed in the safe forest home area and the exploration woodland group who explored in the unsafe exploration woodland area during daytime and return home at night. The result is division of labor for early hominins to form the forest group and the woodland group due to bipedalism as Equation (1).

$$\begin{array}{l} \text{quadrupedalism} \xrightarrow{\text{change from dense forest to mixed forest-woodland habitat}} \text{bipedalism} \\ \xrightarrow{\text{protection of small children and pregnant females with handicapped feet for running}} \\ \text{early hominins with existential division of labor} \\ = \text{forest group} + \text{woodland group} \end{array} \quad (1)$$

Existential division of labor is one of the criteria of eusociality [53] which is the highest level of organization of animal sociality in certain insects, crustaceans, and mammals. Ants, bees, and termites are eusocial animals. Human is a species of eusocial ape [54] [55]. Such division of labor is existential division of labor without which hominine society would have not been able to exist with the severe handicap of bipedalism, in the same way that the bee society would have not been able to exist without the existential division of labor between queen bees who cannot find foods and worker bees who cannot reproduce. Under division of labor, hominins became interdependent specialists. Existential division of labor produced the highly cooperative hominine society. Being cooperative, early hominins lost the large sharp canine teeth for continuous internal aggression and fighting that took place in great apes. The two important traits that distinguish early hominins from great apes are bipedalism and small canine teeth. For early hominins, division of labor did not require the large brain. Similar to other apes, Ardi's skull encased a small brain—300 to 350 cc. As a result, other than bipedalism and small canine teeth, the physical features of early hominins were ape-like.

During the evolution from great apes to early hominins, bipedal hominins in the mixed forest-woodland habitat were divided into the forest group and the woodland group for division of labor. It was imperative for the two groups to find each other. To find each other, the two groups initially used instinctive vocal call which was the same for all early hominins. The instinctive call could not distinguish one social group from another social group in the same general area. The members of a social group could get completely lost, or could get into conflicts with other social groups. As a result, unlike nonhuman primates, early hominins became disorientation-prone. The controllable vocal territory locator as territorial song had definitely evolutionary advantage to distinguish one social group from another social group. The early hominins with the capability of vocal language learning to develop the controllable vocal territory locator survived much better than the early hominins without such capability. Gradually, all early hominins had the controllable vocal territory locator. With the controllable vocal territory locator, the early hominins developed other kinds of controllable locators, resulting in the evolution into intermediate controllable vocal language as in Equation (2).

$$\begin{array}{l}
 \text{intermediate controllable gestural language} \\
 + \text{primitive controllable vocal language for great apes} \\
 \xrightarrow{\text{the separation of two groups required controllable vocal territory locator}} \\
 \text{intermediate controllable gestural language} \\
 + \text{intermediate controllable vocal language for early homonins-homos}
 \end{array} \quad (2)$$

The further evolution in vocal language learning allowed the development of the controllable vocal mother locator as contact call for children to locate their mothers. Children learn contact calls from their mothers. The controllable vocal mother locator improved the survivability of young children to gain evolutionary competitiveness. According to Oren Poliva [28], the auditory cortex communicates with the frontal lobe via the middle temporal gyrus (auditory ventral stream; AVS) or the inferior parietal lobule (auditory dorsal stream; ADS). Whereas the AVS is ascribed only with sound recognition, the ADS is ascribed with sound localization, voice detection, prosodic perception/production, lip-speech integration, phoneme discrimination, articulation, repetition, phonological long-term memory and working memory. According to Poliva, the role of the ADS in vocal control enabled early hominins to name objects using monosyllabic calls, and allowed children to learn their parents' calls by imitating their lip movements. As a result, ADS is the location in the brain for various vocal locators.

The further evolution in vocal language learning allowed early hominins to have controllable mating songs as controllable vocal iconic language, while gibbons have the instinctive mating songs which do not involve vocal learning. The controllable mating songs accelerated the evolution in controllable iconic vocal language in terms of sexual selection.

### 3.2. From Early Hominins to Early Homos

Arid climate that intensified in around 2.8 million years ago transformed the mixed forest-woodland habitat into the mixed woodland-savanna habitat [56]. The need to walk well in the mixed woodland-savanna was much stronger than in the mixed forest-woodland habitat, so the feet were evolved to much better walking feet than climbing feet. Such evolution of feet changed early hominins into early Homos, such as *Homo habilis* that still retained some body features for climbing trees in woodland. The early Homos had the division of labor with the woodland group and the savanna group as Equation (3).

$$\begin{aligned}
 & \text{early hominins with existential division of labor} \\
 & = \text{forest group} + \text{woodland group} \\
 & \xrightarrow{\text{woodland-savanna habitat}} \text{early Homos with existential division of labor} \\
 & = \text{woodland group} + \text{savanna group}
 \end{aligned} \tag{3}$$

Most early Homos became extinct quite early. One early Homo, *Homo naledi* survived at least until between 236,000 and 335,000 years ago [57]. Like *Homo habilis*, *Homo naledi* still retained some body features for climbing trees. The most important foods in savanna were plant roots and seeds.

The capability of early homos' controllable vocal language reached the same level as songbirds and cetaceans which have intermediate controllable vocal language. In songbirds, there is no substantial increase in the overall brain size despite a substantial increase in the song pre-motor region of the brain [58]. Early homos also did not have substantial increase in the overall brain size despite of substantial expansion of a language-production Broca's area [59]. The brains of early homos were small.

### 3.3. From Early Homos to Middle Homos

Arid climate that intensified further about 2 million years ago transformed the mixed woodland-savanna habitat into the savanna habitat, resulting in the emergency of middle Homos such as *Homo erectus*. Due to the loss of the foods obtained from woodland, the middle Homos were forced to hunt animals in savanna, so the middle Homos turned the division of labor into the gatherer group to gather plant foods and the hunter group to hunt animals as Equation (4).

$$\begin{aligned}
 & \text{early Homos with existential division of labor} \\
 & = \text{woodland group} + \text{savanna group} \\
 & \xrightarrow{\text{savanna habitat}} \text{middle Homos with existential division of labor} \\
 & = \text{gatherer group} + \text{hunter group}
 \end{aligned} \tag{4}$$

The gatherer group continued to gather mostly plant roots and seeds from savanna. The hunter group hunted animals. The feet were evolved to the feet for walking and running only, and were not good in climbing trees. The early Homos were evolved into the middle Homos. The intake of highly nutritious meat allowed the energy-fat hungry brain to expand [60]. The human brain is 2 percent of the body's weight but uses 20 percent of the oxygen supply and gets 20

percent of the blood flow. The calorie content of meat is high. The brain is a very fatty organ, and meat is a much better source of the necessary fats than plant foods. The stone tools were able to cut and grind the foods into digestible forms [61]. For Homos, savanna with many predators was a difficult place to survive. As a result, under the evolutionary pressure of difficult savanna and the nourishment of meat, the brain expanded to increase intelligence for advanced stone tools and advanced gestural language to survive in savanna [62] [63]. The brain size increased rapidly with the sizes between 750 and 1225 cc. The body size also increased. *Homo erectus* at least in the larger specimens had double the brain size of *Homo habilis*, and the body size was much closer to modern human body size than *Homo habilis*. Some early Homos and some middle Homos coexisted at the same time [57] in different locations for long time, because different locations in Africa had different habitats due to the differences in climate.

The gestural language inherited from great apes was good enough for early hominins and early Homos. With intermediate controllable gestural language and intermediate controllable vocal language, early hominins and early Homos had better communication than nonhuman primates. The limitation to any significant improvement in gestural language is the brain size. The significant expansion of the energy-fat hungry brain through the intake of highly nutritious meat allowed significant improvement in gestural language. The enlarged brain due the intake of highly nutritious meat allowed the advancement of intermediate gestural language to advanced gestural language. Advanced gestural language is a complex gestural language which needs syntax and grammar to provide clear and complex gestural expressions. The development of syntax and grammar is ontogenetic ritualization among communicators during the course of repeated interactions. Syntax and grammar are not instinctive (phylogenetic). There is no universal syntax and grammar [64]. The large brain allowed the evolution from intermediate controllable gestural language into advanced controllable gestural language for middle Homos Equation (5).

$$\begin{array}{l}
 \text{intermediate controllable gestural language} \\
 + \text{intermediate controllable vocal language in early hominines and early Homos} \\
 \hline
 \text{hunting meat in savanna habitat to allow the increase in the brain size to improve gestural language} \rightarrow \\
 \text{advanced controllable gestural language} \\
 + \text{intermediate controllable vocal language for middle Homos}
 \end{array}
 \tag{5}$$

The expansion of the brain accommodated advanced gestural language and increasingly advanced stone tools. With advanced controllable gestural language and intermediate controllable vocal language as well as advanced stone tools, *Homo erectus* was adaptive enough to migrate out of Africa, adapt to the environments outside of Africa, and survived for a very long time between about 1.89 million and 143,000 years ago [65].

### 3.4. From Middle Homos to Late Homos

The period between between 800,000 and 200,000 years ago is the period of



strongest climate fluctuation worldwide [66]. The savanna habitat became the unstable habitat in certain regions such as East Africa. The most successful Homos in East Africa that survived such unstable habitat had new FOXP2 gene that enhanced vocal language procedural learning, resulting in advanced vocal language [45] [67]. The middle Homos was evolved into the late Homos as in Equation (6).

middle Homos with existential division of labor = gatherer group + hunter group  
 $\xrightarrow{\text{new FOXP2 in fluctuating savanna habitat}}$   
 late Homos with advanced vocal language and with existential division of labor  
 = gatherer group + hunter group

(6)

The late Homos include *Homo sapiens* and *Neanderthals*.

The whole human evolution from great apes with quadrupedalism is as Equation (7)

quadrupedalism  $\xrightarrow{\text{change from dense forest to mixed forest-woodland habitat}}$  bipedalism  
 $\xrightarrow{\text{protection of small children and pregnant females with handicapped feet for running}}$   
 early hominins with existential division of labor = forest group + woodland group  
 $\xrightarrow{\text{woodland-savanna habitat}}$  early Homos with existential division of labor  
 = woodland group + savanna group  
 $\xrightarrow{\text{savanna habitat}}$  middle Homos with existential division of labor  
 = gatherer group + hunter group  $\xrightarrow{\text{new FOXP2 in fluctuating savanna habitat}}$   
 late Homos with advanced vocal language and with existential division of labor  
 = gatherer group + hunter group

(7)

To increase their reproductive success under the evolutionary pressure of adverse fluctuating savanna, Homos adopted a new trait that was advanced vocal language to complement advanced gestural language. Each language has its strength and weakness. For examples, while vocal language enables communication in complete darkness, gestural language on the other hand allows silent communication that does not attract the attention of predators and prey animals. Certain tasks require the usages of hands to prevent the usage of gestural language at the same time, while certain tasks require the usages of mouth to prevent the usage of vocal language at the same time, so the usage of both vocal language and gestural language has a definite advantage. The advancement of vocal language was derived from the new FOXP2 gene which improves vocal language motor procedural learning. The coexistence of advanced gestural language and advanced vocal language in late Homos overcame the adverse environment, and had definitely an evolutionary competitive advantage. As the population of late Homos increased, large inter-group group gathering became common. In a large inter-group group gathering, everyone's speech could be heard easily, while everyone's gesture could not be seen easily. Advanced vocal

language that improved the communication in large inter-group group gathering enhanced the inter-group cooperation. Gradually, advanced vocal language became the dominant language. The brain size increased to accommodate advanced vocal language and inter-group cooperation through advanced vocal language until the brain size reached the maximum. Middle Homos evolved into the late Homos as in Equation (8).

$$\begin{array}{l}
 \text{advanced intermediate controllable gestural language} \\
 + \text{intermediate controllable vocal language in middle Homos} \\
 \xrightarrow{\text{new FOXP2 in fluctuating savanna habitat to improve vocal language procedural learning}} \\
 \text{advanced controllable gestural language} \\
 + \text{advanced controllable vocal language in late Homos}
 \end{array} \quad (8)$$

The whole evolution of human language is as Equation (9).

$$\begin{array}{l}
 \text{intermediate controllable gestural language} \\
 + \text{primitive controllable vocal language in great apes} \\
 \xrightarrow{\text{bipedalism in mixed forest-woodland habitat to cause the separation of two groups requiring vocal territory locator}} \\
 \text{intermediate controllable gestural language} \\
 + \text{intermediate controllable vocal language in early homonines and early Homos} \\
 \xrightarrow{\text{hunting meat in savanna habitat to allow the increase in the brain size to improve gestural language}} \\
 \text{advanced intermediate controllable gestural language} \\
 + \text{intermediate controllable vocal language in middle Homos} \\
 \xrightarrow{\text{new FOXP2 in fluctuating savanna habitat to improve vocal language procedural learning}} \\
 \text{advanced controllable gestural language} \\
 + \text{advanced controllable vocal language in late Homos}
 \end{array} \quad (9)$$

### 3.5. Controllable Languages in Different Levels

Intermediate controllable gestural language used by great apes, early hominins, and early Homos have few symbols (gestural attention getters) and many icons (gestural intention signals), and is too simple to have the need for syntax and grammar. Great apes only learn to use signs symbolically under the very special conditions of close and continuous relationship with humans [19], but the use of symbolic signs does not occur naturally [5]. Advanced gestural language used by middle and late Homos has many symbols, few icons, and iconic language as prosody. Advanced gestural sign language is complex enough to have simple syntax and simple grammar. Intermediate vocal language used by songbirds, cetaceans, and early hominins, early Homos, and middle Homos has vocal locator in symbolic language and vocal show in iconic language, and is too simple to have the need for syntax and grammar. Advanced vocal language used by late Homos is dominated by symbolic language, while vocal iconic language is used as prosody. Advanced vocal language is complex enough to have complex syntax and complex grammar. Different levels of controllable languages different ways of expressions as in **Table 4**.

**Table 4.** Controllable languages in different levels.

<b>controllable language</b>	<b>animals</b>	<b>language components</b>	<b>syntax-grammar</b>
intermediate gestural language	great apes, early hominins, early Homos	few symbols (gestural attention getters) + many icons (gestural intention signals)	no
advanced gestural language	middle and late Homos	many symbols + few icons + iconic language as prosody	yes
intermediate vocal language	songbirds, cetaceans, early hominins, early Homos, middle Homos	locator as symbolic language + vocal show as iconic language	no
advanced vocal language	late Homos	symbolic language + iconic language as prosody	yes

#### 4. Summary and Conclusion

In summary, this paper proposes that the understanding of human language evolution requires the comprehensive understandings of language in terms of language types, formations, and learnings and the comprehensive understanding of human biological evolution in terms of the emergences of various hominin species with various language capacities. This paper proposes language neuromechanics and the human biological-language evolution. Language is derived from bodily movement. Language neuromechanics combines neuroscience to study language brain and biomechanics to study language movement. Language neuromechanics consists of language type, language formation, and language learning. Language types for advanced animals include gestural language verse vocal language, instinctive language verse controllable language, and symbolic language verse iconic language. Language formation involves the developments of the different types of languages from different bodily movements phylogenetically and ontogenetically. For language, bodily movements result in visual actions perceived by vision and auditory action perceived by hearing. Some actions become phylogenetically inarticulate instinctive gestural language and vocal language without the need of learning, while some actions become controllable actions consisting of attentional action and intentional action. Attentional action, such as random action, gains attention without meaning and specific purpose. Intentional action, such as purposeful action, has meaning and specific purpose. In the development of language from action, attentional action ontogenetically turns into symbolic language with arbitrary meaning-referent, while intentional action ontogenetically turns into iconic language with deliberate meaning-referent. The combination of symbolic language and iconic language turns into symbol dominant language and icon dominant language. As a result, the six different language types are instinctive gestural language, instinctive vocal language, symbol dominant sign language, icon dominant gestural language,

symbol dominant speech, and icon dominant song. Language learning involves the learning of controllable language to adapt to communicative environment through the language brain regions and the language genes. The language brain regions include linguistic cognitive learning region in a core Wernicke's area and linguistic motor learning region in a core Broca's area. On the other hand, the three controllable linguistic genes are FOXP2 for linguistic motor learning, CNTNAP2 for linguistic cognitive learning, and FOXP1 for linguistic overlap learning to overlap linguistic motor learning and linguistic cognitive learning.

The human language was evolved from great apes' languages that use mostly controllable gestural language in addition to basically instinctive inarticulate vocal language. On the other hand, human language consists of articulate gestural (sign) language and articulate speech sharply different from great apes' languages. This paper proposes a gradual and step-by-step human language evolution from the language of great apes to the human language through the human biological evolution. The most important factor in the human biological evolution is habitat. Different habitats produced different hominins. A most important factor in the human language evolution is evolutionary pressure. Under evolutionary pressure, an animal has to adopt a new trait, such a new capacity for language, to increase its reproductive success. The further advancement of language is limited by the brain size for linguistic cognitive learning. The size of energy-fat hungry brain is limited by the foods which have to be high energy and high fat for the large brain. The complexity of language is proportional the brain size until the brain reaches the maximum size. The large brain allows complex cognition which requires complex expression in advanced language. The most difficult motor learning is complex and fast vocal motor learning. FOXP2 is for motor learning such as motor procedural learning. The improvement from new FOXP2 to increase synaptic plasticity causes the advancement in vocal language,

Great apes, such as chimpanzee and bonobo, live in forest habitat. They are good at tree climbing, but not in bipedal walking. Their brains are small. They have intermediate gestural language and primitive vocal language. Around 7 million years ago, the climate slowly cooled, resulting in the expansion of woodlands and savannas and the shrinkage of forests in Africa to produce various habitats. Early hominins, such as *Ardipithecus ramidus* and *Australopithecus*, lived in mixed forest-woodland habitat. They were good at tree climbing and bipedal walking, but were inadequate in bipedal running. The handicap in bipedal running forced early hominins in the mixed forest-woodland habitat to have the existential division of labor with the forest group for pregnant females, young children, and old hominins and the woodland group for young hominins. Their brains were small. They had intermediate gestural language and intermediate vocal language. The advancement of vocal language from the primitive language to the intermediate language was derived from the adoption of controllable territorial song as controllable vocal locator for the two groups (the forest group and the woodland group) to find each other under the evolutionary pressure of

disorientation. Early Homos, such as *Homo habilis* and *Homo naledi*, lived in the mixed woodland-savanna habitat. They were good at tree-climbing and bipedal walking, and could run better than early hominins. They had intermediate gestural language and intermediate vocal language. Middle Homos, such as *Homo erectus*, lived in savanna habitat originally. Without the foods from woodland, they had division of labor with the gatherer group to gather plant foods and the hunter group to hunt animals for meat. They were good at bipedal walking and running, but not tree climbing. They had medium size brains. The brain size increase was due to the intake of nutritious meat to supply energy and fat for the expanding energy-fat hungry brain. They had advanced gestural language and intermediate vocal language. The advancement of gestural language was derived from the increase in the brain size. The period between 800,000 and 200,000 years ago is the period of strongest climate fluctuation worldwide, resulting in adverse habitat that endangered the existence of Homos. To increase reproductive success under such evolutionary pressure, they had to adopt a new trait that was advanced vocal language to complement advanced gestural language. The advancement of vocal language was derived from the new FOXP2 gene which improves vocal language motor procedural learning. The result was the emergency of late Homos, such as *Homo sapiens* and *Neanderthals*. Their brains were large to accommodate advanced vocal language and inter-group cooperation through vocal language until the brain size reached the maximum. The summary is listed in **Table 5**.

In conclusion, combining neuroscience and biomechanics, language neuromechanics provides the comprehensive understanding of language. The combination of language neuromechanics and the human biological-language evolution provides the clear evolutionary path from great apes' articulate gestural

**Table 5.** The combined human biological and language evolutions.

species	examples	habitat	tree climbing	bipedal walking/running	division of labor	brain size cc	gestural language	vocal language
great apes	<i>chimpanzee, bonobo</i>	forest	yes	no/no	none	small 275 - 500 cc	intermediate	primitive
early hominins	<i>Ardipithecus ramidus, Australopithecus</i>	mixed forest-woodland	yes	yes/no	forest group-woodland group	small 300 - 550 cc	intermediate	intermediate
early Homos	<i>Homo habilis, Homo naledi</i>	mixed woodland-savanna	yes	yes/no	woodland group-savanna group	small 450 - 650 cc	intermediate	intermediate
middle Homos	<i>Homo erectus</i>	savanna	no	yes/yes	gatherer group-hunter group	medium 750 - 1225 cc	advanced	intermediate
late Homos	<i>Homo sapiens, Neanderthals</i>	fluctuating savanna	no	yes/yes	gatherer group-hunter group	large human average 1350 cc	advanced	advanced

language without articulate speech to human articulate gestural language and articulate speech. Future research involves the investigation of theory of mind (unique to humans) in language neuromechanics and the human biological-language evolution.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- [1] Hauser, M., *et al.* (2014) The Mystery of Language Evolution. *Frontiers in Psychology*, **5**, 401. <https://doi.org/10.3389/fpsyg.2014.00401>
- [2] Hockett, C. (1960) The Origin of Speech. *Scientific American*, **203**, 88-96. <https://doi.org/10.1038/scientificamerican0960-88>
- [3] Nishikawa, K., *et al.* (2007) Neuromechanics: An Integrative Approach for Understanding Motor Control. *Integrative and Comparative Biology*, **47**, 16-54. <https://doi.org/10.1093/icb/icm024>
- [4] Hammarstro, H. (2016) Linguistic Diversity and Language Evolution. *Journal of Language Evolution*, **1**, 19-29. <https://doi.org/10.1093/jole/lzw002>
- [5] Kendon, A. (2017) Reflections on the “Gesture-First” Hypothesis of Language Origins. *Psychonomic Bulletin & Review*, **24**, 163-170. <https://doi.org/10.3758/s13423-016-1117-3>
- [6] Middlemis-Brown, J., Johnson, E. and Blumberg, M. (2005) Separable Brainstem and Forebrain Contributions to Ultrasonic Vocalizations in Infant Rats. *Behavioral Neuroscience*, **119**, 1111-1117. <https://doi.org/10.1037/0735-7044.119.4.1111>
- [7] Newman, J. (2007) Neural Circuits Underlying Crying and Cry Responding in Mammals. *Behavioural Brain Research*, **182**, 155-165. <https://doi.org/10.1016/j.bbr.2007.02.011>
- [8] Barnett, A., Bazelon, M. and Zapella, M. (1966) Visual and Auditory Function in an Hydranencephalic Infant. *Brain Research*, **2**, 351-360. [https://doi.org/10.1016/0006-8993\(66\)90004-7](https://doi.org/10.1016/0006-8993(66)90004-7)
- [9] Smith, W. (1945) The Functional Significance of the Rostral Cingular Cortex as Revealed by Its Responses to Electrical Excitation. *Journal of Neurophysiology*, **8**, 241-255. <https://doi.org/10.1152/jn.1945.8.4.241>
- [10] Ardila, A., Bernal, B. and Rosselli, M. (2016) How Localized Are Language Brain Areas? A Review of Brodmann Areas Involvement in Oral Language. *Archives of Clinical Neuropsychology*, **31**, 112-122. <https://doi.org/10.1093/arclin/acv081>
- [11] Tomasello, M., Gust, D. and Frost, G. (1989) A Longitudinal Investigation of Gestural Communication in Young Chimpanzees. *Primates*, **30**, 35-50. <https://doi.org/10.1007/BF02381209>
- [12] Liebal, K. and Call, J. (2012) The Origins of Non-Human Primates’ Manual Gestures. *Philosophical Transactions of the Royal Society of London B: Biological sciences*, **367**, 118-128. <https://doi.org/10.1098/rstb.2011.0044>
- [13] Tomasello, M., George, B., Kruger, A., Farrar, J. and Evans, E. (1985) The Development of Gestural Communication in Young Chimpanzees. *Journal of Human Evolution*, **14**, 175-186. [https://doi.org/10.1016/S0047-2484\(85\)80005-1](https://doi.org/10.1016/S0047-2484(85)80005-1)
- [14] Plooi, F. (1978) Some Basic Traits of Language in Wild Chimpanzees? In: Lock, A.,

- Ed., *Action, Gesture and Symbol: The Emergence of Language*, Academic Press, London, UK, 111-131.
- [15] Fitch, W.T., de Boer, B., Mathur, N. and Ghazanfar, A.A. (2016) Monkey Vocal Tracts Are Speech-Ready. *Science Advances*, **2**, e1600723. <https://doi.org/10.1126/sciadv.1600723>
- [16] Tomasello, M. (2008) *Origins of Human Communication*. The MIT Press. Cambridge, MA.
- [17] Arbib, M. (2012) *How the Brain Got Language: The Mirror Neuron Hypothesis*. Oxford University Press, Oxford, UK. <https://doi.org/10.1093/acprof:osobl/9780199896684.001.0001>
- [18] Armstrong, D. and Wilcox, S. (2007) *The Gestural Origin of Language*. Oxford University Press, New York. <https://doi.org/10.1093/acprof:oso/9780195163483.001.0001>
- [19] Wallman, J. (1992) *Aping Language*. Cambridge University Press, Cambridge, UK, 10-28. <https://doi.org/10.1017/CBO9780511611858>
- [20] Tomasello, M. and Zuberbühler, K. (2002) Primate Vocal and Gestural Communication. In: Bekoff, M., Allen, C. and Burghardt, G.M., Eds., *The Cognitive Animal. Empirical and Theoretical Perspectives on Animal Cognition*, Bradford Books/MIT, Cambridge, MA, 293-299.
- [21] Goodall, J. (1986) *The Chimpanzees of Gombe: Patterns of Behavior*. Harvard University Press, Cambridge, MA.
- [22] Berwick, R., et al. (2012) A Bird's Eye View of Human Language Evolution. *Frontiers in Evolutionary Neuroscience*, **4**, 1-25. <https://doi.org/10.3389/fnevo.2012.00005>
- [23] Catchpole, C. and Slater, P. (2008) *Bird Song: Biological Themes and Variations*. Cambridge University Press, Cambridge, UK. <https://doi.org/10.1017/CBO9780511754791>
- [24] Beecher, M. and Brenowitz, E. (2005) Functional Aspects of Song Learning in Songbirds. *Trends in Ecology and Evolution*, **20**, 143-149. <https://doi.org/10.1016/j.tree.2005.01.004>
- [25] Andre, M. and Kamminga, C. (2000) Rhythmic Dimension in the Echolocation click Trains of Sperm Whales: A Possible Function of Identification and Communication. *Journal of Marine Biological Association of the United Kingdom*, **80**, 163-169. <https://doi.org/10.1017/S002531549900168X>
- [26] Wright, T. and Wilkinson, G. (2001) Population Genetic Structure and Vocal Dialects in an Amazon Parrot. *Philosophical Transactions of the Royal Society of London B*, **268**, 609-616. <https://doi.org/10.1098/rspb.2000.1403>
- [27] King, S. and Janik, V. (2013) Bottlenose Dolphins Can Use Learned Vocal Labels to Address Each Other. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 13216-13221. <https://doi.org/10.1073/pnas.1304459110>
- [28] Poliva, O. (2016) From Mimicry to Language: A Neuroanatomically Based Evolutionary Model of the Emergence of Vocal Language. *Frontiers in Neuroscience*, **10**, 307. <https://doi.org/10.3389/fnins.2016.00307>
- [29] Stafford, K., Lydersen, C., Wiig, Ø. and Kovacs, K. (2018) Extreme Diversity in the Songs of Spitsbergen's Bowhead Whales. *Biology Letters*, **14**, Article ID: 20180056. <https://doi.org/10.1098/rsbl.2018.0056>
- [30] Petkov, C. and Jarvis, E. (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>
- [31] Knornschild, M. (2014). Vocal Production Learning in Bats. *Current Opinion in Neurobiology*, **28**, 80-85. <https://doi.org/10.1016/j.conb.2014.06.014>
- [32] Rodenas-Cuadrado, P., et al. (2018) Mapping the Distribution of Language Related Genes *FoxP1*, *FoxP2*, and *CntnaP2* in the Brains of Vocal Learning Bat Species. *Journal of Comparative Neurology*, **526**, 1235-1266. <https://doi.org/10.1002/cne.24385>
- [33] Janik, V. and Slater, P. (2000) The Different Roles of Social Learning in Vocal Communication. *Animal Behaviour*, **60**, 1-11. <https://doi.org/10.1006/anbe.2000.1410>
- [34] Brown, S. (2017) A Joint Prosodic Origin of Language and Music. *Frontiers in Psychology*, **8**, 1894. <https://doi.org/10.3389/fpsyg.2017.01894>
- [35] Li, S., Weidenfeld, J. and Morrisey, E. (2004) Transcriptional and DNA Binding Activity of the *Foxp1/2/4* Family Is Modulated by Heterotypic and Homotypic Protein Interactions. *Molecular and Cellular Biology*, **24**, 809-822. <https://doi.org/10.1128/MCB.24.2.809-822.2004>
- [36] Rodenas-Cuadrado, P., Ho, J. and Vernes, S. (2014) Shining a Light on *CNTNAP2*: Complex Functions to Complex Disorders. *European Journal of Human Genetics*, **22**, 171-178. <https://doi.org/10.1038/ejhg.2013.100>
- [37] Fisher, S. and Scharff, C. (2009) *FOXP2* as a Molecular Window into Speech and Language. *Trends in Genetics*, **25**, 166-177. <https://doi.org/10.1016/j.tig.2009.03.002>
- [38] Rodenas-Cuadrado, P., et al. (2016) Characterisation of *CASPR2* Deficiency Disorder—A Syndrome Involving Autism, Epilepsy and Language Impairment. *BMC Medical Genetics*, **17**, 8. <https://doi.org/10.1186/s12881-016-0272-8>
- [39] Penagarikano, O., et al. (2011) Absence of *CNTNAP2* Leads to Epilepsy, Neuronal Migration Abnormalities, and Core Autism-Related Deficits. *Cell*, **147**, 235-246. <https://doi.org/10.1016/j.cell.2011.08.040>
- [40] Sollis, E., et al. (2016) Identification and Functional Characterization of de Novo *FOXP1* Variants Provides Novel Insights into the Etiology of Neurodevelopmental Disorder. *Human Molecular Genetics*, **25**, 546-557. <https://doi.org/10.1093/hmg/ddv495>
- [41] Enard, W., et al. (2002) Molecular Evolution of *FOXP2*, a Gene Involved in Speech and Language. *Nature*, **418**, 869-872. <https://doi.org/10.1038/nature01025>
- [42] Konopka, G., et al. (2009) Human-Specific Transcriptional Regulation of CNS Development Genes by *FOXP2*. *Nature*, **462**, 213-217. <https://doi.org/10.1038/nature08549>
- [43] Schreiweis, C., et al. (2014) Humanized *Foxp2* Accelerates Learning by Enhancing Transitions from Declarative to Procedural Performance. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 14253-14258. <https://doi.org/10.1073/pnas.1414542111>
- [44] Enard, W., et al. (2009) A Humanized Version of *Foxp2* Affects Cortico-Basal Ganglia Circuits in Mice. *Cell*, **137**, 961-971. <https://doi.org/10.1016/j.cell.2009.03.041>
- [45] Krause, J., et al. (2007) The Derived *FOXP2* Variant of Modern Humans Was Shared with Neandertals. *Current Biology*, **17**, 1908-1912. <https://doi.org/10.1016/j.cub.2007.10.008>
- [46] Corballis, M. (2009) The Evolution of Language. *Annals of the New York Academy of Sciences*, **1156**, 19-43. <https://doi.org/10.1111/j.1749-6632.2009.04423.x>
- [47] Pickford, M. and Senut, B. (2001) The Geological and Faunal Context of Late Mi-



- ocene Hominin Remains from Lukeino, Kenya. *Comptes Rendus Académie de la Terres et des Planètes*, **332**, 145-152.
- [48] White, T., *et al.* (2009) *Ardipithecus ramidus* and the Paleobiology of Early Hominins. *Science*, **326**, 64-86. <https://doi.org/10.1126/science.1175802>
- [49] Carvalho, S., *et al.* (2012) Chimpanzee Carrying Behaviour and the Origins of Human Bipedality. *Current Biology*, **22**, R180-R181. <https://doi.org/10.1016/j.cub.2012.01.052>
- [50] Gibbons, A. (2009) A New Kind of Ancestor: *Ardipithecus* Unveiled. *Science*, **326**, 36-40. <https://doi.org/10.1126/science.326.5949.36>
- [51] Lovejoy, C., *et al.* ((2009) The Great Divides: *Ardipithecus ramidus* Reveals the Postcrania of Our Last Common Ancestors with African Apes. *Science*, **326**, 73-106. <https://doi.org/10.1126/science.1175833>
- [52] Lovejoy, C. (2009) Reexamining Human Origins in Light of *Ardipithecus ramidus*. *Science*, **326**, 74-74e8.
- [53] Crespi, B.J. and Douglas, Y. (1995) The Definition of Eusociality. *Behavior Ecology*, **6**, 109-115. <https://doi.org/10.1093/beheco/6.1.109>
- [54] Wilson, E. (2012) *The Social Conquest of the Earth*. WW Norton & Company, New York.
- [55] Chung, D. (2016) The Basic Principles of Kin Sociality and Eusociality: Human Evolution. *Natural Science*, **8**, 8-19. <https://doi.org/10.4236/ns.2016.81002>
- [56] DiMaggio, E., *et al.* (2015) Late Pliocene Fossiliferous Sedimentary Record and the Environmental Context of Early *Homo* from Afar, Ethiopia. *Science*, **347**, 1355-1359. <https://doi.org/10.1126/science.aaa1415>
- [57] Dirks, P., *et al.* (2017) The Age of *Homo naledi* and Associated Sediments in the Rising Star Cave, South Africa. *eLife*, **6**, e24231. <https://doi.org/10.7554/eLife.24231>
- [58] Garamszegi, L., *et al.* (2005) Sexually Size Dimorphic Brains and Song Complexity in Passerine Birds. *Behavioral Ecology*, **16**, 335-345. <https://doi.org/10.1093/beheco/arh167>
- [59] Tobias, P. (1987) The Brain of *Homo Habilis*. A New Level Of Organization in Cerebral Evolution. *Journal of Human Evolution*, **16**, 741-761. [https://doi.org/10.1016/0047-2484\(87\)90022-4](https://doi.org/10.1016/0047-2484(87)90022-4)
- [60] Aiello, L. and Dean, C. (1990) *An Introduction to Human Evolutionary Anatomy*. Academic Press, Cambridge, MA.
- [61] Zink, K. and Lieberman, D. (2016) Impact of Meat and Lower Palaeolithic Food Processing Techniques on Chewing in Humans. *Nature*, **531**, 500-503. <https://doi.org/10.1038/nature16990>
- [62] Hewes, G. (1973) Primate Communication and the Gestural Origin of Language. *Current Anthropology*, **14**, 5-24. <https://doi.org/10.1086/201401>
- [63] Cataldo, D., Migliano, A. and Vinicius L. (2018) Speech, Stone Tool-Making and the Evolution of Language. *PLoS ONE*, **13**, e0191071. <https://doi.org/10.1371/journal.pone.0191071>
- [64] Hinzen, W. (2012) The Philosophical Significance of Universal Grammar. *Language Sciences*, **34**, 635-649. <https://doi.org/10.1016/j.langsci.2012.03.005>
- [65] Indriati, E., *et al.* (2011) The Age of the 20 Meter Solo River Terrace, Java, Indonesia and the Survival of *Homo erectus* in Asia. *PLoS ONE*, **6**, e21562. <https://doi.org/10.1371/journal.pone.0021562>
- [66] Potts, R. (2012) Evolution and Environmental Change in Early Human Prehistory.

*Annual Review of Anthropology*, **41**, 151-167.

<https://doi.org/10.1146/annurev-anthro-092611-145754>

- [67] Coop, G., Bullaughey, K., Luca, F. and Przeworski, M. (2008) The Timing of Selection at the Human *FOXP2* Gene. *Molecular Biology and Evolution*, **25**, 1257-1259.

<https://doi.org/10.1093/molbev/msn091>