The evolution of language and thought

Philip Lieberman

Brown University, Cognitive, Linguistic and Psychological Sciences, Providence RI 02912, USA
e-mail: Philip.Lieberman@gmail.com

Summary – *Language primarily evolved as a vocal medium that transmits the attributes of human culture and the necessities of daily communication. Human language has a long, complex evolutionary history. Language also serves as an instrument of thought since it has become evident that in the course of this process neural circuits that initially evolved to regulate motor control, motor responses to external events, and ultimately talking were recycled to serve tasks such as working memory, cognitive flexibility linguistic tasks such as comprehending distinctions in meaning conveyed by syntax. This precludes the human brain possessing an organ devoted exclusively to language, such as the Faculty of Language proposed by Chomsky (1972, 2012). In essence like Fodor's (1983) modular model, a restatement of archaic phrenological theories (Spurzheim, 1815). The subcortical basal ganglia can be traced back to early anurans. Although our knowledge of the neural circuits of the human brain is at a very early stage and incomplete, the findings of independent studies over the past 40 years, discussed here, have identified circuits linking the basal ganglia with various areas of prefrontal cortex, posterior cortical regions and other subcortical structures. These circuits are active in linguistic tasks such as lexical access, comprehending distinctions in meaning conferred by syntax and the range of higher cognitive tasks involving executive control and play a critical role in conferring cognitive flexibility. The cingulate cortex which appeared in Therapsids, transitional mammal-like reptiles who lived in age of the dinosaurs, most likely enhanced mother-infant interaction, contributing to success in the Darwinian (1859) "Struggle for Existence" – the survival of progeny. They continue to fill that role in present-day mammals as well as being involved in controlling laryngeal phonation during speech and directing attention (Newman & MacLean, 1983; Cummings, 1993). The cerebellum and hippocampus, archaic structures, play role in cognition. Natural selection acting on genetic and epigenetic events in the last 500,000 years enhanced human cognitive and linguistic capabilities. It is clear that human language did not suddenly come into being 70,000 to 100,000 years as Noam Chomsky (Bolhuis et al., 2014) and others claim. The archeological record and analyses of fossil and genetic evidence shows that Neanderthals, who diverged from the human line at least 500,000 years ago possessed some form of language. Nor did the human population suddenly acquire the capability to relate two seemingly unrelated concepts by means of associative learning 100,000 years ago, re-coined "Merge" by Chomsky and his adherents, Merge supposedly is the key to syntax but associative learning, one of the cognitive processes by which children learn languages and the myriad complexities of their cultures, is a capability present in dogs and virtually all animals.*

Keywords – *Evolution, Language, Cognition, Neural circuits, Natural selection, Genetics, Noam Chomsky.*

Introduction

Humans are the only living species that possesses language, but its evolution does not appear to involve any singular evolutionary mechanism that is in any sense unique to human. The evolution of human language hinges on natural selection acting on heritable biological variation.

To Charles Darwin, biological variation was manifest:

"No one supposes that all the individuals of the same species are cast in the same mould" (1859, p. 45).

In his time, the nature and biological mechanisms that resulted in heritable variations and their transmission were unclear. However, though epigenetic changes arising from environmental effects, such as disease susceptibility of descendants of 19th century famines in Sweden (Pembry *et al.*, 2006) can affect an entire population, it remains unclear if these effects are inherited over four or more generations. Selective sweeps in which Natural Selection acts to enhance the survival of progeny of individuals remains the only attested mechanism for speciation and lesser evolutionary change. As Charles Darwin pointed out in 1859, Natural Selection will drive speciation in different ecosystems to different ends, but Natural Selection and epigenetic events which Darwin also noted but did not emphasize are the only means for the "transmutation of species."

In Darwin's words:

"…any variation, however slight and from whatever cause proceeding, if it be in any degree profitable for an individual of any species, in its infinitely complex relations to other organic beings and to external nature, will tend to the preservation of that individual, and will generally be inherited by its offspring. The offspring, also, will thus have a better chance of surviving…." (p. 61) .

It is clear that adding a source of food will increase the chance of survival to the members of a group. For example, the introduction of potatoes in societies as diverse as 19th century Norway and the Khumbu region of Nepal resulted in doubling the population. Adult lactose tolerance likewise evolved from selective sweeps on genes in cultures in which sheep, cattle or goats were already domesticated. Natural selection favored the descendants of individuals who had a mutation on one gene and the gene frequency distribution of the population shifted over generations (Tishkoff *et al.,* 2007). The shelves in supermarkets of lactose-free foodstuffs reflect the fact that the ancestors of many customers lived in settings in which milk products did not form part of an adult diet as well, as chance mutations that both confer and delete adult lactose tolerance. However, the primary "group" effect derives from selective sweeps in ecosystems where milk products were available. Some selective sweeps occurred rapidly. The survivors of the European Black Plague clearly possessed a degree of genetically transmitted immunity to the disease.

Recycling

Although Darwin stressed the gradual nature of natural selection knew that he had to account for abrupt transitions can occur, such as that from aquatic to terrestrial life. His solution was the recycling – an organ initially adapted for one function, could take on a new role. His example was the swimbladder:

"…the fact that an organ originally constructed for one purpose…may be converted into one for a wholly different purpose…." (p. 190).

Swimbladders, internal elastic sacks within fish had evolved so as to allow fish to adjust the size of their bodies to match the mass of water displaced at a given depth, allowing them to hover. Air extracted from the water by the fishes gills is transferred into the fishes's two swimbladders. "Primitive" species, such as sharks, that lack swimbladders have to be constantly in motion, else they would sink to the sea floor. By chance mutations, lungs evolved from swimbladders in which air could be extracted from the atmosphere in lung-fish stranded in dry river banks and then transferred into the fishes' bloodstreams.

Current studies confirm that recycling is a general process, accounting for phenomena as diverse as diurnal vision developing from visual receptors initially adapted for high-intensity light (Anderson & Finlay, 2014) or the cerebellum and basal ganglia which initially appear to have been adapted for motor tasks functioning in cognitive tasks, including language (e.g., Marsden & Obeso, 1994; Lieberman, 2000, 2006, 2013; Monchi *et al.,* 2001, 2006, 2007; Simard *et al.*, 2011).

The evolution of language

The focus of linguistic speculation on the evolution of language over the past 70 years has been on syntax. The dominant position, derived

from Noam Chomsky, is that the primary role of language is an instrument of thought, communication playing at most an incidental role in its evolution. Moreover, according to Chomsky and his associates, children do not learn their native languages through imitation or associative learning. Instead, every child that has been born during the past 100,000 years or so, has "preloaded" knowledge in her/his "Faculty of Language." Innate, knowledge that enables him or her to activate the detailed syntax of all human languages that existed, or will ever come to be. Arithmetic also is innate. Biological variation supposedly does not mark these innate biological capacity, nor did Natural Selection acting on variation play a role in the evolution of these genetically transmitted capacities (Chomsky, 1972, 2012; Bolhuis *et al.*, 2014).

Chomsky for decades proposed an innate "Universal Grammar" (UG) (e. g., Chomsky, 1988) that specified numerous genetically transmitted "principles and parameters" that when activated determined a language's syntax. However, genetic variation characterizes all living organisms – including humans. Therefore, UG is implausible owing to genetic variation. In some children, the principles and parameters necessary to acquire their native language would be missing, though other languages could be acquired (Lieberman, 2006, 2013). The "Strong Minimalist Thesis" described in Chomsky (2012) and Bolhuis *et al.* (2014) attempted to avoid the complication of variation and instead claims that a single innately transmitted mental operation, "Merge", suddenly came into being 50,000 to 100,000 years ago (the dates vary in Chomsky 2012 and Bolhuis *et al.*, 2015).

Chomsky's Merge operation yields what is supposed to be the defining characteristic of every human language that existed, exists, or may come into being - complex sentences with embedded clauses. Merge which accounts for the acquisition of language by children and its sudden evolution "takes exactly two (syntactic) elements and puts them together" (Bolhuis *et al.*, 2014). However, Merge entails having the same comprehensive innate knowledge base as UG

because different languages put different things together. A child exposed to English will hear two-word, definite article-noun sequences such as the cow. Chinese does not use definite articles so children exposed solely to Chinese would not hear these two-word sequences. If languages were not learned using the cognitive processes by which we learn how to use chopsticks or forks, the article-noun constructions and thousands of language-specific aspects of syntax would have had to have been innately specified 50,000, 70,000 or 100,000 years ago to account for the diversity that marks human languages. Moreover, what would happen in bilingual Chinese-English homes. Would cognitive linguistic deficits occur or does every child's brain have a set of Merge constraints for every language that existed in the last 100,000 years? And since languages continually change is there some supernatural agency that predicted the form of all languages that we may construct until our species becomes extinct? Both UG and the store of knowledge necessary for Merge are in the realm of fantasy.

Apart from this problem, many, perhaps most animals possess the ability to learn to relate two events. Pavlov demonstrated this in the 1890s in his classic experiments in which dogs were trained to relate the sound of a bell with meat. Dogs, without any instruction, learn to associate spoken words with objects or actions. The dog in the Kaminski household learned the names of hundreds of children's toys without any instruction in one trial – in other words, immediately Kaminski *et al.* (2004). Other dogs have demonstrated similar abilities. Mollusks are capable of learning to relate two stimuli, though slowly (Carew *et al.*, 1981). The claim for Merge being a unique human capability that enabled language 100,000 years ago, reflects hermetic disregard of both controlled studies over the past century, and common knowledge of anyone familiar with a dog, cat, horse and most other animals.

"In the beginning was the word" (John 1:1)

Chomsky's views on the abrupt creation of language tale shares similarities with the tenets of "Creation Theory" that takes the Bible as literal

truth in that it rejects the Darwinian theory of evolution. However the Bible correctly recognizes the central role of speech in language, a view supported by research that engineered devices that work – the ubiquitous speech recognition systems in us on computers and smartphones.

It became evident in the 1960s that speech is "special" – a species-specific feature that makes complex human language possible. It also has been possible to derive insights on the evolution of language and cognition through study of the anatomy and physiology of speech production and perception. The research group directed by Alvin Liberman and Franklin S. Cooper at Haskins Laboratories was attempting to make a text-to-speech system that would "read out" printed text to blind persons. The Haskin's group had to depart from linguistic theory which holds that phonemes (similar to the letters of the alphabet) can be freely reordered to form different words as though they were movable type - "beads on a string."

Liberman and his colleagues in their seminal 1967 paper, "The perception of the speech code," showed that phonemes did not exist in the stream of speech. The then new technology of tape recording allowed them to attempt the acoustic speech signal into phonemes. If the speech signal actually consisted of a string of phonemes then a tape recording of someone uttering the word *too* should have had a segment of tape that contained the phoneme [t], followed by a segment of tape that contained the vowel [u]. (The symbol [u] represents the word *too's* vowel in standard phonetic notation.) However, when the ostensible [t] segment excised from the word *too,* was placed before a tape segment excised from the word *tea* that hypothetically contained its [i] vowel, the result was incomprehensible.

Similar results held for all "stop" consonants (in English the phonemes b, p, d, t, k) and other vowels. It also was impossible to extract "pure" vowels from tape recordings of even bisyllabic syllables because formant frequency transitions carried on the ostensible segments of "pure" vowels determined whether a listener heard a *da, pa, gu, etc.* Analysis of the articulatory gestures involved

in speaking showed that words were produced as "encoded" entities. One articulatory gesture merged into adjacent, sometimes removed 'segments." For example, the position assumed by a speaker's lips while for the word *too* are protruded and narrowed in anticipation of the vowel [u] at the start of the word. In contrast, when anyone utters the word tea their lips are retracted at the very beginning of the word for the vowel [i]. Human speakers when they talk aim at producing words, not individual phonemes. We pull out the acoustic patterns that convey words from the stream of speech by means of a complex perceptual process. This seemingly over-complex procedure overcomes the limit imposed by the fusion-frequency of our auditory system. Rather transmitting a stream of phonemes that will heard as an incomprehensible buzz, we transmit words at a slower rate that then can be decomposed into abstract entities – phonemes.

Linguists still hold to the theory that permutable phonemes, "beads on a string", are one of the defining features of "articulated" human language. However, engineers who have to make things that work have largely abandoned the search for invariant segments of the speech signal that might correspond to phonemes. The speech recognition algorithms in common use on telephone systems, computers and smart-phones identify words. The incoming acoustic signal usually is matched against word-level templates that take into account differences in vocal tract length and dialect. Alphabetic notation is not the only writing system. The Chinese languages have for thousands of years transcribed words as entities.

The source-filter theory of speech production

The high transmission rate of speech derives from "formant frequency encoding." In order to understand how this bears on the evolution of language it's necessary to briefly discuss speech physiology. Pipe organs essentially are devices that generate formant frequencies that specify musical notes. The world's oldest existing playable pipe organ in Sion, Switzerland, was built in 1435; earlier organs probably existed. Some knowledge of speech physiology, therefore, was

known in mediaeval Europe. A bellows provided the necessary airflow which generate a "wide-band" acoustic source. Energy over a broad range of frequencies. When a person presses a "key" on the organ's keyboard a valve opens a pipe that has a particular length and is open at both ends or closed at one end. The pipe acts as an acoustic filter, allowing maximum energy to pass through it at frequencies that are determined by the length and shape of the pipe, producing a musical note. The colors that you see through a pair of sunglasses represents a similar process. Sunlight has energy distributed throughout the viable frequency range of electromagnetic energy. The tint of the sunglass lenses reduces energy at specific frequencies, the range of frequencies that pass through the sunglasses determines whether everything that you see is blueish, reddish, greyed down, etc.

Figure 1 is a sketch of the anatomy involved in speech production. The lungs provide the source of energy for speech production. As Darwin pointed out, the lungs of mammals and other terrestrial species are recycled swim bladders. Lungs retain the elastic property of swimbladders. During quiet inspiration, the diaphragm and intercostal and abdominal muscles expand the lungs. The elastic recoil of the lungs then provides the force that expels air during expiration. The duration of inspiration and expiration is almost equal. Since the elastic lungs act in a manner analogous to a rubber balloon, the alveolar (lung) air pressure is at a maximum at the start of each expiration and linearly falls as the lungs deflate. The alveolar air pressure during expiration thus starts at a high level and falls as the volume of the elastic lung sacks falls. The alveolar pressure of the outgoing flow of air impinges on the vocal cords of the larynx.

In contrast, the pattern of activity during speech and singing is quite different. The diaphragm is immobilized and the duration of expiration is keyed to the length of the sentence that a speaker produces. Alveolar air pressure is maintained at an almost uniform level until the end of expiration. This entails a speaker adjusting intercostal and abdominal muscles to "hold back"
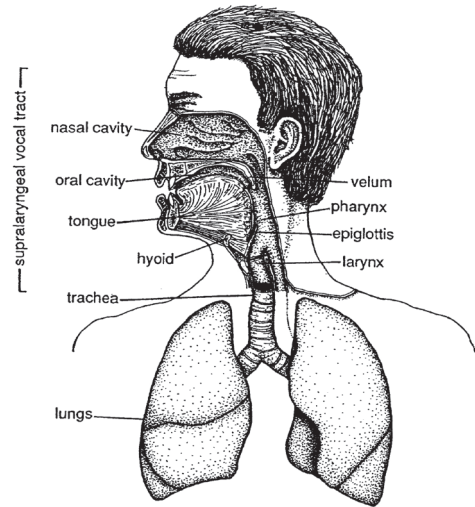
Fig. 1 -The lungs, larynx and supralaryngeal vocal tract

against the force generated by the elastic recoil force, which is high at the start of the expiration and gradually falls. The intercostal and abdominal muscles contain muscle "spindles" that monitor the force that they produce (Bouhuys, 1974). The diaphragm contains few spindles, which may account for its not being active when people talk or sing. Moreover, a speaker must anticipate the length of the sentence that he or she intends to produce, generally taking in more air before the start of a long sentence.

Since lung volume is higher before the start of a long sentence, holdback maneuvers of the intercostal and abdominal muscles must be planned and executed to take the higher elastic recoil force at the start of a sentence into account achieving a relatively level alveolar air pressure during the sentence (Lieberman, 1967). The rate at which the vocal cords open and close, the fundamental frequency of phonation (Fo), depends on the tension of the muscles that make up the vocal cords, which are complex structures made up of muscles, other soft tissue and cartilages, and other tissue, and the alveolar air pressure. The initial high alveolar air pressure would either blow apart the vocal cords, or result in a high Fo and that would always fall throughout the sentence. Instead, for

most languages the typical pattern is for a fairly level Fo, except for momentary controlled peaks that signal emphasis and a sharp decline at the sentence's end Lieberman (1967).

The meaningful calls of many species are differentiated by Fo contours and variations (e.g. Cheyney & Sayfarth, 1990) that also convey emotional information in humans cultures. But the fundamental frequency of phonation also differentiates words in tone languages such as the Chinese languages (Tseng 1981). However, the primary role of the larynx is as a source of energy for vowels and other "voiced" sounds. During phonation energy is present at Fo and integral multiples of Fo. The supralaryngeal vocal tract (SVT) acts as though it was a malleable organ pipe. Formant frequencies, which are the frequencies at which maximum acoustic energy could pass through the SVT, are the major determinants along with durational cues of the encoded phonemes that constitute words. As the tongue, lips, and the position of the larynx (which can move up or down to a limited degree) move, the formant frequencies continually change.

The absolute values of the formant frequencies for any phoneme depend on the length of a speaker's SVT. The average formant frequencies of the vowel [i], for example, are 270, 2,300, and 3,000 Hz for the adult males studied by Peterson & Barney (1952), whereas the formant frequencies of the vowel /u/ were 300, 870, and 2,240 Hz. Since the length of the SVT varies from person to person and during the years of childhood and adolescence for the same individual, the absolute values of the formant frequency pattern vary. For example, the formant frequencies of an [i] would be 1.5 times higher for a child whose SVT length was 11.3 cm long than for an adult whose SVT was 17 cm long. However, both formant frequency patterns would be perceived as examples of an [i] to human listeners owing to a process of perceptual normalization in which listeners internally estimate the length of a speaker's SVT (Nearey, 1978). Listeners can estimate SVT length after hearing a short stretch of speech or deducing SVT length after hearing a known phrase or word

such as person saying *hello.* Any successful automatic procedure would have to take the effects of differing SVT lengths into account.

Nearey's comprehensive 1978 study showed that the vowel [i] (of the word see) was an optimal signal for immediate SVT normalization explaining the results of one of the earliest attempts at machine-implemented speech recognition. Peterson & Barney in 1952 had found that words containing [i] (the vowel of see) and to a lesser extent [u] were identified with fewest errors in a project aimed at developing a voice-implemented telephone dialing system. The system would have to work for men, women and speakers who had different dialects of English. Words having the form [hVd], such as heed and had, produced by ten different speakers were presented in quasi-random order to a panel of listeners; the listeners had to immediately adjust

for different speakers' voices and identify each word. Out of 10,000 trials, listeners misidentified [i] two times, [u] six times. Words that had other vowels were misidentified hundreds of times. Hillenbrand *et al.* (1995) reported virtually identical results in a study which eliminated

dialect variation and made use of more precise computer-implemented formant measuring technology that had not existed in 1952.

Hence perceiving linguistic information conveyed by speech perception entails listeners taking these factors into account. Many species estimate the size of conspecifics by taking account of the absolute value of the formant frequencies of their cries. Vocal tract length in these species correlates with body size. Hence longer SVTs produce lower formant frequencies, providing an acoustic marker of body size (Fitch, 2000). This may account for dogs being able to recognize words (Kaminski *et al.*, 2004).

The evolution of human speech anatomy
    Charles Darwin noted:

> *"The strange fact that every particle of food and drink which we swallow has to pass over the orifice of the trachea, with some risk of falling into the lungs…." (1859, p. 191)*

Choking on food remains the fourth leading cause of accidental death in the United States (http://www.nsc.org/library/report_injury_usa.htm). This problem, unique to humans arises because of "recycling" of the human tongue to enhance the process of speech perception. The shape and position in the neck of the human tongue has been modified so as to enable us to produce the vowels [i] and [u]. The larynx which rests atop the trachea is anchored to the root of the tongue. At birth in human infants most of the tongue is positioned in the mouth and its shape is flat as is the case for other mammals. When ingesting liquids, an infant's larynx can be raised forming a sealed passage from the nose to allow breathing while the liquid flows around the larynx. There essentially are two separate pathways – one for breathing devoted to breathing. This explains why human infants and cats of all ages can lap up milk uninterrupted by breathing.

A complex developmental process that spans the first eight years of life takes place in humans to arrive at adult-like morphology (Negus, 1949; Crelin, 1968; Lieberman D.E. & McCarthy, 1999; Lieberman, 2006; Lieberman D.E., 2011). During this time the skull is restructured, shortening the relative length of the oral cavity. The tongue is reshaped towards an almost circular posterior contour and descends down into the pharynx; neck length increases. Negus (1949) thought that the unique shape of the human tongue facilitated speech communication, compensating for increasing the risk of choking. That supposition has been confirmed by computer-modeling studies that calculate the range of formant frequencies that could be produced by human and nonhuman SVTs (Lieberman *et al.,* 1969, 1972; Lieberman & Crelin 1971; Carre *et al.,* 1995; De Boer, 2010).

The initial Lieberman *et al.* (1969) study calculated the formant frequency patterns of the vowels that a rhesus macaque's tongue and SVT could produce. The range of tongue shapes was estimated by taking into account constraints on tongue deformation, which were subsequently confirmed (Takemoto, 2001, 2008). The monkey's tongue was positioned as far as possible to produce the SVT configurations used by adult human speakers to yield the "point" vowels [i], [u], and [a]. These vowels delimit the range of vowels used in human languages (Greenberg, 1963).

The human tongue's oral and pharyngeal proportions and shape explain why only adult humans can produce the vowels [i], [u], and [a] and why these vowels contribute to the robustness of human vocal communication. Half the tongue (SVTH) is positioned in the oral cavity, and half SVTV is positioned in the pharynx. SVTH and SVTV meet at an approximate right angle, owing to the tongue's posterior circular shape. The extrinsic muscles of the tongue, muscles anchored in bone, can move the almost undeformed tongue to create abrupt midpoint ten-to-one discontinuities in the SVT's cross-sectional area. Stevens's (1972) parallel research explained why the unique human tongue contributed to the robustness of human vocal communication. Stevens showed that only the species-specific human SVT can produce the ten-to-one midpoint area function discontinuities, abrupt changes in the cross-sectional area of the vocal tract, that are absolutely necessary to produce the vowels [i], [u], and [a], which he termed "quantal." Stevens employed both computer modeling using Henke's (1966) algorithm and physical models (wooden tubes that could be shifted to change the position of the 10:1 changes in SVT cross-sectional area). Quantal vowels are perceptually salient owing to the convergence of two formant frequencies which yield spectral peaks. Their formant frequency patterns do not shift when tongue position varies about one centimeter about the midpoint. Speakers thus can be sloppy and produce the "same" vowel. Nearey (1978) subsequently showed that the vowel [i] is an optimal signal for determining the length of a speaker's vocal tract - a necessary step in the complex process of recovering the linguistic content from the acoustic signals that convey speech. Whereas the identical formant frequency pattern can represent two different vowels for speakers who have different SVT lengths, no such overlap occurs for tokens of [i].

Independent computer-modeling studies carried out by Lieberman & Crelin (1971), Lieberman *et al.* (1972) and De Boer (2010) have reached similar conclusions and pointed to the evolution of the species-specific human SVT non-human primate and newborn human SVTs could not produced quantal vowels. Carre and his colleagues used a technique that directed the computer model to produce a vocal tract that could produce the full range of formant frequencies of human speech by modifying a nonhuman SVT. The system "grew" a pharynx equal in length to its oral segment. Carre *et al.* (1995) concluded that in order to produce the full phonetic range of human speech, "a vocal tract must have independently controllable oral and pharyngeal cavities nearly equal in length."

Speech capabilities of extinct hominins

The first comprehensive anatomy of the human newborn was published by Edmund S. Crelin in 1969. Crelin examined the adult male Neanderthal specimen from La Chapelle-aux-Saints and concluded that the Neanderthal SVT was similar to that of a large human newborn on the basis of the total pattern of morphological similarities between their basicrania. Lieberman & Crelin (1971) then computer modelled the reconstructed La Chapelle 1 SVT to determine the range of vowel formant frequency patterns that Neanderthals could have produced. Cineradiographic data on newborn infant cry (Truby *et al.*, 1965) guided the jaw, tongue, lip, and laryngeal maneuvers to derive the range of possible vowels. As Lieberman & Crelin (1971, p. 211) noted, "*When we were in doubt as, for example, with respect to the range of variation in the area of the larynx, we used data derived from adult Man that would enhance the phonetic ability of the Neanderthal vocal tract …*". The Neanderthal phonetic range did not include the point vowels [i], [u] and [a].

Lieberman & McCarthy (2007, 2014) used a different, more conservative, procedure to reconstruct the SVT's of extinct hominins. A series of acrobatic maneuvers of the hyoid bone and larynx is necessary to enable humans to swallow solid food. This places constraints on the possible location of the larynx in the necks of extinct hominins. The hyoid bone must be positioned at or below the mandible's inferior border; otherwise, the suprahyoid muscles could not raise the hyoid during swallowing. Similarly, the hyoid bone cannot lie below the pectoral girdle or else the infrahyoid muscles would not act to lower the hyoid. Moreover, the movements of the hyoid during swallowing would be blocked by the sternum and clavicles. Taking these constraints into account, the lowest possible position for the hyoid-larynx complex in the Neanderthal La Chapelle 1 and La Ferrasie fossils and the Skhul 5 fossil, for which vertebrae were preserved which allowing neck length to be estimated is for the vocal cords to be at cervical vertebra C6, placing the cricoid cartilage of the larynx at C7/T1. This would result in the Neanderthal hominins having SVTh/SVTv ratios of about 1.3, which when computer modelled are outside Stevens (1972) quantal range owing to their long SVThs. Skhul V would also have a 1.3 SVTh/SVTv ratio, owing to the fossil's short neck.

The two Neanderthals have a long palate, long nasopharynx, and short neck. Skhul 5 has a moderately long palate, short nasopharynx, and a short neck. A functionally modern human-like SVT, with equal long horizontal and vertical segments, did not arise until sometime after the appearance of *H. sapiens* about 200,00 ago, as the result of three processes: (1) shortening of the nasopharynx, (2) shortening of the face, and (3) slight elongation of the neck. Neanderthals and other archaic hominins had a long palate and nasopharynx, a configuration which is associated with a long SVTH. A long SVTH cannot be paired with a long SVTV, unless Neanderthals had necks that were much longer than those of fully modern *H. sapiens*. All available evidence suggests that their necks were short, even if they do fall within the range of variation of modern *H. sapiens*.

If Skhul V is to be considered anatomically modern *H. sapiens*, the final two steps in the above schedule occurred after the origin of *Homo sapiens*. However, in light of the biological cost of the

modern human SVT and tongue - an increased risk choking to death - it is apparent that speech and the neural mechanisms implicated in speech production were present in earlier hominins. The evolution of the human tongue can be regarded as an example of "recycling" – adapting an organ suited for swallowing to also enhance the process by which vocal tract length is normalized during speech perception. It, moreover, points to speech and language being present in earlier hominin species. There otherwise would have been no selective advantage for reshaping the human tongue to enhance the robustness of vocal communication (Lieberman, 1997; Lieberman & McCarthy, 2007, 2014).

Comparative Studies

Comparative studies of species of the linguistic abilities of living apes establish some limits on what aspects of language may be unique human attributes. Humans and chimpanzees share a common ancestor that lived about six million years ago (The Chimpanzee Sequencing and Analysis Consortium, 2005). Thus, although living apes also since evolved, any aspect of language that can be mastered by present-day apes most likely was present in the early stages of hominin evolution. If exposed to a language-using environment early in life, chimpanzees can acquire active vocabularies of about 150 words, communicating their needs and observations to humans and to each other (e.g., Gardner & Gardner 1969; Savage-Rumbaugh *et al.,* 1985). Chimpanzee Loulis in infancy acquired some American Sign Language (ASL) proficiency when he could only observe and participate in ASL communication with other ASL-using chimpanzees. These comparative studies of ape communication show that apes can learn words, master some aspects of ASL morphology, comprehend distinctions in meaning conveyed by simple syntax, and transmit this knowledge to another generation of chimpanzees. It thus is improbable that any "protolanguage" lacking words ever existed. Although present-day chimpanzees undoubtedly have

diverged from the common ancestor of humans and chimpanzees, any capability present in a chimpanzee surely was present in our common ancestor and lter extinct hominins.

As noted earlier domesticated dogs also can comprehend spoken words. Some dogs can

learn in one trial to reference the meaning of hundreds of spoken words with specific objects (Kaminski *et al.,* 2004). However, no nonhuman species can talk, though apes have the anatomical ability to produce all but quantal speech sounds. Apes instead can use manual sign language and other manual system to signify words, lending plausibility to the idea that manual gestures played a significant role in the early stages of language evolution (Hewes, 1973). However, the laryngeal specializations for vocal communication at the expense of respiratory are efficiency are the norm in social animals, including humans (Negus, 1949) again making it unlikely that manual gestures were ever the sole phonetic medium of language.

The neural bases of human language and cognition

Many current studies and textbooks still cling to the traditional neurophysiologic model in which Broca's frontal and Wernicke's posterior cortical areas linked by a cortical pathway constitute a system devoted to language and language alone. However, the model which has its roots in early 19th century phrenology, is wrong. Phrenologists proposed that areas of the neocortex - the outermost layer of the human brain - were the "seats" or "faculties" that each were the basis for an aspect of behavior such as religious devotion, language, mathematics and so on (Spurzheim, 1815). Since technical limits precluded direct inspection of the cortex, the area of the skull above the location of the hypothetical seat hypothetically correlated with a person's linguistic or mathematical ability or piety.

When the theory was tested, it was found wanting. Homicidal individuals were found who had large faculties of piety and trust. But

phrenology was resurrected when Paul Broca (1861) had published his study of a stroke victim whose speech was limited to a single distorted syllable. Broca whose postmortem study of the patient's brain was limited to the surface of the cortex linked one damaged cortical area to the patient's language deficits. Thus, Broca's area was born - a cortical area devoted to language that continues to be taken as a given in current studies adhering to Noam Chomsky's views on language, such as Berwick *et al.* (2013) or Fedorenko *et al.* (2011). Fedorenko and her colleagues go further, claiming to have found innate micro-sites in the left inferior gyrus of the cortex (the traditional site of Broca's area) devoted solely to comprehending distinctions in meaning conveyed by syntax.

Curiously, when the brain of Broca's first patient was examined using a high resolution MRI (the brain had been preserved in alcoholfor more than a century), the cortical damage was in the ventrolateral prefrontal cortex, not the left inferior gyrus, the traditional site of Broca's area (Dronkers *et al.*, 2007). However, the precise location of Broca's area is irrelevant because it is not critical to language. Postmortem examinations first showed that language and speech were disrupted only when subcortical brain damage was present. Current brain imaging techniques such as computerized tomography (CT) and magnetic resonance imaging (MRI) allowed the brains of thousands of patients who suffer aphasia, permanent loss of language, to be examined. Patients recover after brain damage limited to cortex. Conversely, aphasia results when subcortical structures are damaged. Alexander *et al.* (1987), for example, documented the speech production deficits of patients who had suffered strokes that damaged the basal ganglia and neural pathways connecting to it, but spared the cortex altogether. Stuss and Benson in their comprehensive 1986 book concluded that aphasia never occurs without subcortical damage. No attested case of language loss has ever occurred from a lesion limited to any cortical region, including Broca's area. Researchers holding different views on language share this view in light of the fact that both ventrolateral prefrontal cortex and Broca's area form parts of basal ganglia circuits implicated in regulating speech and language (Lieberman, 2000, 2002, 2006, 2013; Ullman, 2004).

Recycling has played a major role in the evolution of the human brain (Anderson *et al.,* 2014; Finlay & Uchiyama, 2015). The neural bases of complex motor acts, such as walking or talking, are circuits that link local operations performed in different neural structures. It is difficult to find neural structures that are strictly domain-specific. Primary visual cortex, V1, which is involved in the initial stages of visual perception also enables mice to learn to identify images and sequences of images (Gavornik & Bear, 2014; Sam *et al.*, 2015). In humans fMRI imaging shows that V1 is activated when subjects recall images (Kosslyn, *et al.*, 1999). Neural structures often perform local operations that constitute elements of different circuits that regulate seemingly unrelated aspects of behavior (e.g., Marsden & Obeso, 1994). Cars feature similar functional architecture.

If your car won't start, the repair manual will not instruct you to locate the center of starting. The manual instead points out a set of linked structures that each performs a local operation. For, example, the battery provides electrical power to the starter motor but it also powers the car's lights, radio, computer, etc., through circuits linking it to these devices. Moreover, the battery is not in itself the sole device dedicated to electrical power. The generator and voltage regulator form part of the electrical power system, and in a Toyota Prius hybrid powered car, the braking system charges the battery.

Neural Circuits

The present international initiative on mapping out the neural circuits of the human brain reflects the fact that our knowledge base is in its infancy. Invasive tracer studies of the brains of monkeys and other mammals first mapped out a class of circuits linking areas of the cortex with the basal ganglia that also exist in human beings. Retrograde tracer studies entail injecting

a chemical or virus that will propagate back down the neural pathway that transmits the electrochemical signals controlling the muscle. The animal is then sacrificed after a few days allowing the tracer to move down the circuit. The animal's brain is then sectioned and stained allowing the circuits to be microscopically discerned. Other tracers can be injected into neural structures to mark "ascending" neural pathways. Neural circuits were discovered that connected areas of prefrontal cortex through the basal ganglia and other subcortical structures to temporal and parietal cortex, thalamus, cerebellum and hippocampus (e.g., Alexander *et al.*, 1986). Noninvasive diffusion tensor imaging (DTI) has mapped similar circuits in humans and nonhuman primates (Lehéricy *et al.*, 2004).

Evidence from studies of neurodegenerative diseases such as Parkinson disease that degrade basal ganglia operations first suggested that nonhuman primates and humans had similar cortical-basal ganglia circuits. In Parkinson disease (PD), depletion of the neurotransmitter dopamine degrades the local operations of the basal ganglia (Jellinger, 1990). Patients have difficulty in sequencing the submovements that are necessary to carry out internally directed motor acts such as walking. A common clinical observation is that PD patients who have difficulty walking will do better when they are asked to mimic someone walking. External models allow them to function better. Similar problems occur when PD patients execute manual acts (Harrington & Halland, 1991) and speech motor control deteriorates (Lieberman *et al.*, 1992; Lieberman, 2006).

Cognitive inflexibility and difficulties performing cognitive acts that require planning or selecting criteria also occur in PD (e.g., Lange *et al.*, 1992). As Alexander *et al.* (1986), Cummings (1993), and other studies note, prefrontal cortical areas project to the basal ganglia, thus, accounting for cognitive deficits associated with insult to the basal ganglia component of cortical-basal ganglia circuits. "Subcortical dementia," profound diminution of cognitive flexibility can occur in Parkinson disease (PD). Patients are unable to change the direction of a thought process or action (Flowers & Robertson 1985). Other circuits involving basal ganglia, orbofrontal cortex or anterior cingulate cortex can be compromised affecting inhibition, attention and laryngeal control as Cummings (1993) shows.

Marsden & Obeso (1994) came to the conclusion that recycling had augmented basal ganglia function after reviewing the effects of surgical lesions and dopamine replacement therapy aimed at mitigating the problems associated with PD. They concluded that the basal ganglia were a sequencing engine that could link submovements - motor acts stored in motor cortex to carry out an internally guided motor act such as walking. This fits the traditional view that PD affected motor acts, but Marsden & Obeso pointed out that when circumstances dictated, the basal ganglia could change a course of action. Focal brain damage limited to the basal ganglia results in similar speech production and cognitive deficits. Bilateral lesions to the caudate nucleus and putamen of the basal ganglia in the subject studied by Pickett *et al.* (1998) resulted in severe deficits in sequencing the laryngeal, lingual, and lung motor activity necessary to produce articulate speech. The subject also had profound difficulty comprehending distinctions in meaning conveyed by syntax that are comprehended by 6-year-old children. When the subject sorted cards in the "Odd-Man-Out" test, which Flowers & Robertson (1985) devised to test PD patients' cognitive flexibility, she was incapable of changing the sorting criterion.

Neuroimaging studies confirm the role of these circuits in motor control and cognition. Monchi *et al.* (2001, 2006, 2007) used fMRI to monitor oxygen levels in the prefrontal cortex, the basal ganglia, and other subcortical structures. Depleted oxygen levels in the ventrolateral prefrontal cortex, the caudate nucleus, and the thalamus confirmed that this cortical-striatal circuit was activated when planning criterion-sorting shifts. A different cortical-striatal circuit that included posterior prefrontal cortex and the putamen was observed during the execution of a sorting criterion set shift. Dorsolateral prefrontal

cortex was involved whenever subjects made any decision as they performed card sorts, apparently monitoring whether their responses were consistent with the chosen criterion. Other fMRI studies have replicated these findings and show that the caudate nucleus uses this information when a novel action needs to be planned (Monchi *et al.,* 2006, 2007). Similar activation patterns were apparent when subjects were sorting words instead of images and had to match words on the basis of semantic similarity, phonetic similarity of the start of the syllable, or rhyme (Simard *et al.*, 2011). The neural circuits involved thus are not domain-specific, not operating solely on visual criteria. Studies ranging from electrophysiological recordings of neuronal activity in the basal ganglia of mice and other animals as they learn tasks (Graybiel 1995; Mirenowicz & Schultz, 1996; Jin & Costa, 2010) to studies of PD patients (Lang *et al.*, 1992; Monchi *et al.,* 2007) and birds (Brainard & Doupe, 2000) also show that the basal ganglia play a critical role in associative learning and in planning and executing motor acts including speech in humans and songs in songbirds.

Basal ganglia activity in these cortical-basal ganglia circuits clearly is not domain-specific, i.e., limited to language. Motor activity, associative learning and cognitive flexibility manifested in both linguistic and visual tasks involve local basal ganglia operations, reflecting the mark of evolution as cortical-basal ganglia circuits were adapted to "new" tasks.

Ventrolateral prefrontal cortex connected to posterior regions by circuits involving the subcortical basal ganglia is active during virtually all tasks that involve selecting and retrieving information according to specific criteria (Duncan & Owen, 2000). Dorsolateral prefrontal cortex is active while monitoring motor or cognitive events during a task while taking into account earlier events stored in working memory (Badre & Wagner, 2006; Monchi *et al.*, 2001, 2006, 2007; Postle, 2006; Wang *et al.*, 2005). These cognitive tasks range from retrieving information and holding it in short-term "working memory" to changing the direction of a thought process.

The antiquity of neural circuit components

The intonation of speech involves neural structures that have a deep evolutionary history. Intonation reflects laryngeal activity, and the fundamental frequency of phonation (Fo) and amplitude of speech play a central role in vocal communication, signaling sentence boundaries and other aspects of syntax (c.f., Armstrong & Ward, 1926; Lieberman, 1967). Controlled Fo contours differentiate words in "tone" languages such as Mandarin Chinese. The neural circuits and anatomy involved in controlling Fo can be traced back to therapsids, mammal-like reptiles who during the Triassic, Jurassic, and early Cretaceous eras. The initial role of these structures appears to be mother-infant communication.

Studies of human mother-infant communication reveal a special vocal mode or register, motherese, by which parents address infants in speech that has a high fundamental frequency of phonation and extreme Fo variation (Fernald *et al.*, 1989). Mammalian infants (including human infants) also signal for attention by means of a forceful isolation cry that has a high Fo and amplitude (Truby *et al.*, 1965) - the cries that can keep parents awake for months. Comparative studies suggest that therapsids employed similar anatomical specializations and neural structures to produce isolation cries. All mammals possess a paleocortex, which includes the anterior cingulate cortex (ACC). The findings of studies of the behavioral effects of damage to the ACC and the neural circuits that connect it to other parts of the human brain show that the anterior cingulate cortex plays a role in controlling Fo and directing attention to virtually anything that one wishes to do. While the soft tissue of the brains of therapsids has not survived, the inference that these mammal-like reptiles had an ACC rests on the fact that they possess three middle ear bones found in all present-day mammals. The initial function of the hinge bones of the reptilian jaw was to open the jaw wide. In the course of evolution, the hinge bones took on a dual role, functioning as "organs" of hearing. This transition from mammal-like reptiles to mammals involved the former jaw bones migrating into the middle

ear, where they serve as a mechanical amplifier that enhances auditory acuity. All mammals have both an ACC and these middle ear bones, so that middle ear bones are regarded as an index for the presence of the ACC in mammal-like therapsids.

The Darwinian struggle for existence transcends aggressive acts. Any aspect of behavior that increases reproductive success confers a selective advantage. Vocal communication that enhances mother-infant contact clearly contributes to reproductive success. Anterior cingular cortex-basal ganglia neural circuits have been recycled and are involved in both attention and laryngeal control. Lesion studies on mice show that mouse mothers do not pay attention to their infants when neural circuits to the ACC are disrupted (MacLean & Newman, 1988; Newman, 1985). Virtually every neuroimaging study ever published shows ACC activity when subjects are asked to perform any task, both directed and involuntary.

Genetic and epigenetic events that made us human

The human brain is three times as large as a chimpanzee's and there has been a steady advance in brain size over the course of hominin evolution (Herculano-Houzel, 2000). Increases in brain-size have been associated with enhanced cognitive ability in all species since Lartet's (1868) study documenting increases in brain size over time in carnivores and their prey. Many neuroscientists view the brain as a general purpose "computer," though its operations bear little resemblance to those of digital computers (Finlay & Uchiyama, 2015). Thus this general increase in computational capacity is significant and many genes have been proposed as candidates for increased human brain size. However, it is difficult to assess the significance of larger brains per-se since no correlation between absolute brain size and any aspect of cognition is apparent in living, normal human populations. Moreover, average brain size in extinct Neanderthal hominins was somewhat greater than in present-day humans.

Current studies on the effects of genetic mutations and epigenetic factors in the last 500,000 years suggests a different path towards understanding the evolution of the neural bases of human behavior that takes into account the findings of neurophysiology. Hebb in 1949 pointed out the role of synaptic malleability in cognition. Hundreds of independent studies have since pointed out that associative learning, information storage and virtually all aspects of cognition entail modification of synaptic "weights" between neurons – the basic computing elements of the brain. Electro- physiologic studies monitoring activity in the brain, for example, show that as animals learn to execute a task, it is encoded in changes in the degree to which synapses transfer information to neurons (e.g. Graybiel, 1995, 1997; Sanes *et al.*, 1999).

Advances in genotyping have revealed startling results. When the genes of fruit flies and humans are compared, few differences are evident. A complex picture is emerging that transcends the view that genes in themselves specify the distinctions that differentiate species. The morphological and behavioral differences between humans and fruit flies appear to result from the action of transcriptional factors and epigenetic events.

Transcription factors are genes that govern the transcription of information stored in DNA into a different form, single-stranded mRNA, which is later translated into functional proteins that make up the building blocks of the body. In effect, transcriptional factors are "master genes" that influence complex biology involved in how the genes ultimately form the body and brain. It also has become apparent that many DNA sequences do not code genes. These "epigenetic" factors play a major role both in the ontogenetic development of all organisms and their evolution. During the development of an organism epigenetic "enhancers" and "silencers" respectively activate and turn off genes (Reilly *et al.,* 2015). "Evo-Devo" studies have opened an area of enquiry that explains, for example, why we are not large fruit flies though we share most of our genes with these insects – different epigenetic

events shape the development of fruit flies and humans. The development of the human tongue and vocal tract noted above most likely involves genetic and epigenetic programmed events. If these epigenetic events could be identified, a better understanding of the evolution of the human speech producing anatomy could be attained. However, as is the case for most species-specific human characteristics that information is lacking.

In some few instances, such as the role of the FOXP2 transcriptional factor, insights on what makes us human have been attained. The full account is still a work in progress, but it is clear that a series of mutations on FOXP2 in the last 500,000 years and epigenetic events that are being explored acted on the basal ganglia in the circuits discussed above that regulate motor control, cognition and language – as well attention, inhibition and mood. The role of the FoxP2 transcription factor gradually became evident in a study of the members of an extended family in London who had severe deficits in speech production, sentence comprehension, and cognitive ability (Vargha-Khadem *et al*., 1995, 1998; Watkins *et al*., 2002). Anomalies in basal ganglia were noted (Watkins *et al*., 2002). The afflicted individuals possessed only one copy of the human version of the FOXP2 transcription factor, instead of the normal two (Fisher *et al*., 1998).

The Foxp2 gene is one of many transcription factors that exist in all mammals, birds, and other creatures. The mouse version of Foxp2 (the lowercase spelling indicates that it is not the human version) controls the embryonic development of the lungs, intestinal system, heart, and other muscles as well as the spinal column in mice (Shu *et al*., 2001). Humans are separated from mice by 75 million years of evolution (Mouse Genome Sequencing Consortium, 2002). However, the human version FOXP2 of the transcriptional factor and mouse version Foxp2 both act on similar neural structures in circuits involving the basal ganglia that regulate motor control and cognition - the thalamus, the caudate nucleus and putamen (components of the basal ganglia) , cerebellum and other subcortical structures (Lai *et al*., 2003). These neural structures are all intricately interconnected. The cerebellum is involved in both motor coordination and cognition. The cortical plate (layer 6), the input level of the cortex, is also affected by the FOXP2 mutation.

FOXP2 is one of the few genes that has been shown to differ from its chimpanzee version (The Chimpanzee Sequencing and Analysis Consortium 2005). A "human" version has evolved in which two amino acid occurred changes in the FOXP2[human] protein in six or seven million years, compared to one amino acid substitution between chimpanzees and mice over the previous 70 million years. This form of FOXP2 also occurs in Neanderthals and Denisovans - a group related to Neanderthals. An epigenetic change unique to humans occurred (intron 8) in the DNA sequence near the FOXP2 gene when anatomically modern humans appeared (Maricic *et al*., 2013). The date of the "selective sweep" involving the human form of FOXP2, approximately 260,000 ago, was first established by Enard and colleagues (2002).

As noted above, selective sweeps occur when a gene confers a significant advantage in the Darwinian "struggle for existence." In most instances the behavioral and neural consequences of a genetic difference between chimpanzees and humans is unclear. However, in this instance, the behavioral deficits of a FOXP2 anomaly in the KE family showed that it is played a role in the attributes of speech, language, and cognition that distinguish humans from chimpanzees and other living species.

Mouse "knock-in" studies have demonstrated that the human version of FOXP2 enhances information transfer and associative learning in the basal ganglia. In light of basal ganglia activity in both associative learning and motor control. When FOXP2[human] - the human form of the transcriptional factor was knocked into mouse pups (Enard *et al*., 2009; Lieberman, 2009; Reimers-Kipping *et al*., 2011) their vocal calls were somewhat different than the calls of mouse pups that had the normal "wild" version of Foxp2. The significant finding was that knocked-in FOXP2[human] increased synaptic

plasticity in basal ganglia neurons and increased connectivity in the basal ganglia, thalamus, and layer VI of the cortex. Significantly, the human version of FOXP2 increased synaptic plasticity in medium spiny neurons in the basal ganglia and in the substantia nigra, another structure of cortical-basal circuits. During associative learning information stored in the synaptic weights of medium-spiny neurons play a critical role in achieving goals or avoiding an aversive outcome (Bar-Gad & Bergman, 2001; Joshua *et al.,* 2008; Assad & Eskander, 2011). In short, these neurons allow animals to learn to perform complex linked sequences. Hebb's 1949 hypothesis for synaptic weights being the neuronal process by which associative learning takes place has since been validated in hundreds of independent studies. In humans, similar associative processes manifested in at the neuronal level in changes in synaptic weights allow us to learn grammatical "rules," as well as the "rules" that govern how we interact with other people, drive a car, use a fork or chopsticks, species, and learn to cope with the ever-changing conditions of life. Other "highly accelerated regions" (HARs) of the human genome also appear to be implicated in neural development (Konopka *et al.*, 2009), but we are only at the threshold of understanding the biological processes that crafted the human brain.

Cognitive flexibility and the archaeological record

Cognitive flexibility is arguably the most salient characteristic that distinguishes human behavior from that of other species. Humans continually invent new devices, create new fashions and modes of behavior and patterns of thought. We are not bound to genetically fixed responses when circumstances change as is ever the case. The Neanderthal tool-kit never approached the sophistication or variety of the stone tools employed by humans (Riel-Salvatore, 2010; Shea, 2011). However, that reflects relative distinctions rather than an absolute divide. Shea (2011) and Lieberman (2013) note the

pace of innovation that marks human enterprise and is consistent with humans prevailing and Neanderthal extinction about 40,000 years ago (Pinhasi *et al.*, 2011). However, Neanderthals and humans mated in Eurasia before that occurred when humans migrated out of Africa (Green *et al.*, 2010). Evidence for the symbolic use of marine shells and pigments by Neanderthals supports that view (Zilhao *et al.,* 2010). Neanderthal burials with grave goods may have occurred (Trinkaus, 1983), though that remains in dispute. However, Skhul V was buried with grave goods about 100,000 years ago. Intensive art "industries" that clearly involved cultural transmission through the medium of language appear to be a unique human trait. A complex paint-making technology was in place in Africa 100,000 years ago, suggesting earlier uses of pigments for ornamentation and perhaps paintings (Henshilwood *et al.,* 2011).

Language is the medium that enabled humans to transmit the aggregations of complex cultural innovations and traditions. It's modality at first was solely vocal. Exceedingly complex social orders such those of the Peruvian Inca civilization were transmitted orally as late as the 16th century (Kolata, 2013). Illiteracy still is common in much of the world. Formal sign languages were invented in the 18th century.

Concluding comments

The principles and constructs introduced by Charles Darwin in 1859 in *The origin of species* continue to inform evolutionary biology. Theories such as that proposed by Noam Chomsky (2012) which minimize or deny the role of Natural Selection in evolution bear a striking resemblance to Chambers's (1844) *Vestiges of the natural history of creation*. Chambers adopted the 18th century position that God had created a master plan that directed the evolution of animals towards its ultimate goal – humans. Chomsky, though he does not invoke God, takes a similar position – some force directs evolution. Chomsky also adopts Chambers's phrenological model of the brain.

Although Chomsky does not as phrenologists such as Spurzheim (1815) locate a hypothetical "Faculty of Language" beneath a bump on a person's skull, the Chomskian "Faculty" is devoted to language and language alone and has no affinities to neural structures or circuits that may regulate other aspects of behavior in humans and other species. Since Natural Selection has no role in the Chomskian model, selective sweeps cannot account for how some putative biological event that put Merge into one individual's brain, spread throughout the human species 100,000 years ago. Genesis is at least consistent.

No evolutionary processes that are specific to humans shaped language and no brain mechanisms appear to be specific to speech or language. A full appraisal of the biological bases of human language remains in the distant future. However, some of the neural circuits that confer the ability to master and execute the complex motor commands that underlie speech and other aspects of behavior are becoming evident. These neural circuits involve structures that also play a part in "mental" aspects of language such as associating words with their meanings and syntax. Similar neural circuits involving the same cortical and subcortical structures are implicated in a range of "higher" cognitive acts. Though many of these neural structures are "recycled" – being present in archaic species far removed from humans, they have taken on new functions and have been modified by Natural Selection acting on genetic and epigenetic events, some occurring in the last 200,000 years or so and specific to humans.

References

Alexander G. E., DeLong F. H. & Strick P. L. 1986. Parallel organization of segregated circuits linking basal ganglia and cortex. *Ann. Rev. Neurosci.,* 9: 357-381.

Alexander M. P., Naeser M. A. & Palumbo C. L. 1987. Correlations of subcortical CT lesion sites and aphasia profiles. *Brain,* 110: 961-991.

Anderson M. L. & Finlay B. L. 2014. Allocating structure to function: the strong links between neuroplasticity and natural selection. *Front. Hum. Neurosci.,* 7: 918, doi:10.3389/fnhum.2013.00918.

Armstrong L. E. & Ward I. C. 1926. *Handbook of English intonation*. Teubner, Leipzig.

Assad W. F. & Eskander E. N. 2011. Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. *J. Neurosci.,* 31: 17772-17787.

Badre D. & Wagner A. D. 2006. Computational and neurobiological mechanisms underlying cognitive flexibility. *Proc. Natl, Acad. Sci. USA,* 103: 7186-7190.

Bar-Gad I. & Bergman H. 2001. Stepping out of the box: information processing in the neural networks of the basal ganglia. *Curr. Opin. Neurobiol.,* 11: 689-695.

Berwick R. C., Friederici A. D., Chomsky N. & Bolhuis, J. J. 2013. Evolution, brain, and the nature of language, *Trends Cogn. Sci.,* 17: 89-98

Bolhuis J.J., Tattersaall I., Chomsky N. & Berwick R. C. 2014. How could language have evolved? *PLoS Biol.,* 12: e1001934, doi: 10.1371/journal.pbio.1001934 PMID: 25157536

Bouhuys A. 1974. *Breathing*. Grune and Stratton, New York.

Brainard M. S. & Doupe A. J. 2000. Interruption of a basal-ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature,* 404: 762-766.

Broca P. 1861. Nouvelle observation d'aphémie produite par une lésion de la moitié postérieure des deuxième et troisième circonvolutions frontales. *Bull. Soc. Anat.,* 6: 398-407.

Carre R., Lindblom B. & MacNeilage P. 1995 Acoustic factors in the evolution of the human vocal tract. *CR Acad. Sci. Par.,* t320: 471-476.

Chambers R. 1844. *Vestiges of the natural history of creation.* John Churchill, London.

Cheyney D.L & Seyfarth R.M. 1990. *How monkeys see the world: inside the mind of another species.* University of Chicago Press, Chicago.

Chomsky N. 1972. *Language and mind*. Harcourt Brace Jovanovich, New York.

Chomsky N. 2012. *The science of language*. Cambridge University Press, Cambridge.

Crelin E. S. 1969. *Anatomy of the newborn: an atlas*. Lea and Febiger, Philadelphia.

Cummings J.L. 1993. Frontal-subcortical circuits and human behavior. *Arch. Neurol.,* 50: 873-880.

Darwin C. 1859/1964. *On the origin of species.* Harvard University Press, Cambridge, MA.

De Boer B. 2010. Modeling vocal anatomy's significant effect on speech. *J. Evol. Psychol.,* 8: 351-366.

Dronkers N. F., Plaisant O., Iba-Zizen M. T & Cabanis E.A. 2007. Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130: 1432-1441

Duncan J. & Owen A. M. 2000 Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.,* 10: 475-483.

Enard W., Przeworski M., Fisher S. E., Lai C.S.L., Wiebe V., Kitano T., Monaco A.P. & Pääbo S. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature,* 41: 869-872.

Enard W., Gehre S., Hammerschmidt K., Hölter S.B, Blass T., Somel M., Brückner M.K., Schreiweis C., Winter C., Sohr R. *et al.* 2009. A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. *Cell,* 137: 961-971.

Federenko E., Behr M. K. & Kanwisher N. 2011. Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl. Accad. Sci. USA*, org/cgi/doi/10.1073/pnas.1112937108

Fernald A. T., Taeschner J., Dunn J. J., Papousek M., de Boysson-Bardies B. & Fukui, I. 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J. Child Lang.*, 16: 477-501.

Finlay B. L. & Uchiyama R. 2015. Developmental mechanisms channeling cortical evolution. *Trends Neurosci.,* 38: 69-76.

Flowers K.A, & Robertson C. 1985. The effects of Parkinson's disease on the ability to maintain a mental set. *J. Neurol. Neurosurg. Psych.,* 48: 517-529.

Gardner R.A. & Gardner B.T. 1969. Teaching sign language to a chimpanzee. *Science,* 165:664-672.

Gavornik J. P. & Bear M. F. 2014. Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nature Neuroscience,*17: 163-170.

Graybiel A.M. 1995. Building action repertoires: memory and learning functions of the basal ganglia. *Curr. Opin. Neurobiol.,* 5: 733-741.

Graybiel A. M. 1997. The basal ganglia and cognitive pattern generators. *Schizophrenia Bulletin* 23: 459–469.

Green R. E., Krause J., Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Patterson N., Li H., Zhai W., Hsi-Yang Fritz M. *et al.* 2010. A draft sequence of the Neandertal genome. *Science*, 328: 710-722.

Greenberg J. 1963. *Universals of language*. MIT Press, Cambridge, MA

Harrington D.L & Haaland K.Y. 1991. Sequencing in Parkinson's disease: abnormalities in programming and controlling movement. *Brain,* 114: 99-115.

Hebb D.O. 1949. *The organization of behavior: a neurophysiological theory*. Wiley, New York

Henke W. 1966. *Dynamic articulatory model of speech production using computer simulation.* Ph.D. dissertation, MIT.

Henshilwood C., d'Errico F., van Niekerk K.L., Coquinot Y., Jacobs Z., Lauritzen S.E., Menu M. & García-Moreno R. 2011. A 100,000 Year Old Ochre Processing Workshop at the Blombos Cave, South Africa. *Science*, 334: 219–221.

Herculano-Houzel S. 2009 The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neuroscience*, 3: 1-11.

Hewes G.W. 1973. Primate communication and the gestural origin of language. *Curr. Anthropol.* 14: 5-24.

Hillenbrand J, Getty L.A, Clark M.J & Wheeler K. 1995. Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.*, 97: 3099-3111.

Jellinger K. 1990. New developments in the pathology of Parkinson's disease. In Streifler M.B. Korczyn A.D, Melamed E. & Youdim M.B.H. (eds): *Advances in neurology, vol 53, Parkinson's disease: anatomy, pathology, and therapy*, pp. 1-16. Raven, New York.

Jin X & Costa R.M. 2010. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, 466:457-462

Joshua M., Adler A., Mitelman R.,Vaadia, E. & Bergman H. 2008. Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *J. Neurosci.*, 28: 11673-11684.

Kaminski J, Call J. & Fischer J. 2004. Word learning in a domestic dog: evidence for "fast mapping". *Science*, 304: 1682-1683.

Kolata A. L. 2013. *Ancient Inca*. Cambridge University Press, New York.

Konopka G., Bomar J. M, Winden K., Coppola G., JonssonZ.O., Gao F., Peng S., Preuss T.M., Wohlschlegel J.A. & Geschwind D.H. 2009. Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature*, 462: 213-217.

Kosslyn S. M., Pascual-Leone A., Felician O., Camposano S., Keenan J.P., Thompson W.L., Ganis G., Sukel K.E. & Alpert N.M. 1999. The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science*, 284: 167–170.

Lai C. S., Gerrelli D., Monaco A.P., Fisher S. E. & Copp A.J. 2001. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain*, 126: 2455-2462.

Lange K.W., Robbins T.W., Marsden C.D., James M., Owen A.M. & Paul G. M.1992. L-dopa withdrawal in Parkinson's disease selectively impairs cognitive performance in tests sensitive to frontal lobe dysfunction. *Psychopharmacology*, 107: 394-404.

Lartet E. 1868. De quelques cas de progression organique vérifiables dans la succession des temps, géologiques sur des mammifères de même famille et de même genre. *C. R. Acad. Sci.,* 66: 1119-1122.

Lehéricy S., Ducros M., Van de Moortele P. F., Francois C., Thivard L., Poupon C., Swindale N., Ugurbil K. & Kim D-S. 2004. Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann. Neurol.*, 55: 522-527.

Liberman A.M., Cooper F. S., Shankweiler D. P. & Studdert-Kennedy M. 1967. Perception of the speech code. *Psychol. Rev.,* 74: 431-461.

Lieberman P. 1967. *Intonation, perception and language*. MIT Press, Cambridge, MA.

Lieberman D.E & McCarthy R.C.1999. The ontogeny of cranial base angulation in humans and Chimpanzees and its implications for reconstructing pharyngeal dimensions. *J. Hum. Evol.,* 36: 487-517.

Lieberman P. 2000. *Human language and our reptilian brain: the subcortical bases of speech, syntax, and thought*. Harvard University Press, Cambridge, MA.

Lieberman P. 2002. On the nature and evolution of the neural bases of human language. *Yearb. Phys. Anthropol.*, 45: 36-62.

Lieberman P. 2006. *Toward an evolutionary biology of language*. Harvard University Press, Cambridge, MA.

Lieberman D.E. 2011. *The evolution of the human head*. Harvard University Press, Cambridge, MA.

Lieberman P. 2013. *The unpredictable species: what makes humans unique*. Princeton University Press, Princeton.

Lieberman P. & Crelin E.S. 1971. On the speech of Neanderthal man. *Ling. Inq.,* 2: 203-222.

Lieberman P. & McCarthy R.C. 2007. Tracking the evolution of language and speech. *Expedition*, 49: 15-20.

Lieberman P. & McCarthy R.C. 2014. *The evolution of speech and language. Handbook of Paleoarchaeology.* Springer-Verlag, Berlin, Heidelberg,

Lieberman P., Klatt D.H. & Wilson W.H. 1969. Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, 164:1185-1187

Lieberman P., Crelin E.S. & Klatt D.H. 1972 Phonetic ability and related anatomy of the newborn, adult human, Neanderthal man, and the Chimpanzee. *Am. Anthropol.,* 74: 287-307.

Lieberman P., Kako E.T, Friedman J., Tajchman G., Feldman L.S. & Jiminez E.B. 1992. Speech

production, syntax comprehension, and cognitive deficitsin Parkinson's disease. *Brain Lang.,* 43: 169-189.

Lieberman P. & McCarthy R. C. 2014. The Evolution of Speech and Language. In *Handbook of paleoanthropology,* pp. 1-41. Springer, Berlin, Heidelberg. http://www.springerreference.com/index/chapterdbid/363355 3

MacLean P.D. & Newman J.D. 1988. Role of midline frontolimbic cortex in the production of the isolation call of Squirrel monkeys. *Brain Res.,* 450: 111-123.

Maricic T., Günther V., Georgiev O., Gehre, S., Ćurlin, M., Schreiweis, C., Naumann R., Burbano H.A., Meyer M., Lalueza-Fox C. *et al.* 2013. A recent evolutionary change affects a regulatory element in the human FOXP2 gene, *Mol. Biol. Evol.*, 25: 1257-1259.

Marsden C.D. & Obeso J. A. 1994. The functions of the basal ganglia and the paradox of sterotaxic surgery in Parkinson's disease. *Brain,* 117: 877-897.

Mirenowicz J. & Schultz W. 1996. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature,* 379: 449-451.

Monchi O., Petrides M., Petre V., Worsley K. & Dagher A. 2001. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J. Neurosci.,* 21: 7739-7741.

Monchi O., Petrides M., Worsley K. J. & Doyon J. 2006. Functional role of the basal ganglia in the planning and execution of actions. *Ann. Neurol.* 59: 257-264.

Monchi O., Petrides M., Mejia-Constain B., Strafella A.P. 2007. Cortical activity in Parkinson disease during executive processing depends on striatal involvement. *Brain,* 130: 233-244.

Mouse Genome Sequencing Consortium. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature,* 420: 520-562.

Nearey T. 1978. *Phonetic features for vowels*. Indiana University Linguistics Club, Bloomington.

Negus V. E. 1949. *The comparative anatomy and physiology of the larynx*. Hafner, New York.

Pembrey M. E., Bygren L. O., Kaati G., Edvinsson S., Northstone K., Sjöström M., Golding J. & ALSPAC Study Team. 2006. Sex-specific, male-line transgenerational responses in humans. *Eur. J. Hum. Genet.,* 14: 159-66.

Peterson G.E. & Barney H.L. 1952. Control methods used in a study of the vowels. *J. Acoust. Soc. Am.,* 24: 175-184.

Pickett E.R., Kuniholm E., Protopapas A*.,* Friedman J. & Lieberman P. 1998. Selective speech motor, syntax and cognitive deficits associated with bilateral damage to the putamen and the head of the caudate nucleus: a case study. *Neuropsychology,* 36: 173-188.

Pinhasi R., Higham T.F.G., Golovanova L.V. & Doronichev V. B. 2011. Revised age of late Neanderthal occupation and the end of the Middle Paleolithic in the northern Caucasus. *Proc. Natl. Acad. Sci. USA*, 108: 21, 8611–8616.

Postle B.R. 2006. Working memory as an emergent property of the mind and brain. *Neuroscience,* 139: 23-38.

Reilly S. K., Yin J., Ayoub A. E., Emera D., Jing Leng J., Cotney J., Sarro R., Rakic P. & Noonan J. P. 2015. Evolutionary changes in promoter and enhancer activity during human corticogenesis, *Science.* 347: 1155-1159.

Riel-Salvatore J. 2010. A niche construction perspective on the Middle-Upper Paleolithic transition in Italy. *Journal of Archeological Method and Theory,* 17: 323-355.

Reimers-Kipping S., Hevers S., Pääbo S. & Enard W. 2011. Humanized Foxp2 specifically affects cortico-basal ganglia circuits. *Neuroscience*, 175:75-84.

Sam F., Cooke S. F, Komorowski R. W., Kaplan E. S., Gavornik J. P. & Bear1 M. F. 2015. Visual recognition memory, manifested as long-term habituation, requires synaptic plasticity in V1. *Nat. Neurosci.*, 18: 262-273.

Sanes J. N., Donoghue J. P., Thangaraj V., Edelman R.R. & Warach S. 1999. Shared neural substrates controlling hand movements in human motor cortex. *Science,* 268: 1775-1777.

Savage-Rumbaugh S., Rumbaugh D. & McDonald K. 1985. Language learning in two species of Apes. *Neurosci. Biobehav. Rev.*, 9: 653-665.

Shea J. J. 2011. Homo sapiens is as Homo sapiens was: behavioral variability versus "Behavioral Modernity" in Paleolithic Archaeology. *Curr. Anthropol.,* 52: 1-34.

Shu W., Yang H., Zhang L., Lu M.M. & Morrisey E. E. 2001. Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lungs and act as transcriptional repressors. *J. Biol. Chem.*, 276: 27488-27497.

Simard F., Joanette Y., Petrides M., Jubault T., Madjar C. & Monchi O. 2011. Fronto-striatal contributions to lexical set-shifting. *Cereb. Cortex,* 21:1084-1093.

Spurzheim J. K. 1815. *The physiognomical system of Drs. Gall and Spurzheim*. Baldwin Cradock and Joy, London.

Stevens K.N. 1972 Quantal nature of speech. In David E.E. Jr & Denes P.B. (eds): *Human communication: a unified view*. McGraw Hill, New York.

Stuss D. T. & Benson D. F. 1986. *The frontal lobes*. Raven, New York.

Takemoto H. 2001. Morphological analyses of the human tongue musculature for three-dimensional modeling. *J. Speech Hear Res.,* 44: 95-107.

Takemoto H. 2008. Morphological analyses and 3D modeling of the tongue musculature of the Chimpanzee (*Pan troglodytes*). *Am. J. Primatol.,* 70: 966-975.

The Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the Chimpanzee genome and comparison with the human genome. *Nature*, 437: 69-87.

Tishkoff S. A., Reed F., Ranciaro A.,Voight B.F., Babbitt C.C., Silverman J.S., Powell K., Mortensen H.M., Hirbo J.B., Osman M. *et al.* 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.*, 39: 31-40.

Trinkaus E. 1983. *The Shanidar Neandertals*. Academic, New York.

Truby H. L., Bosma J.F. & Lind J. 1965. *Newborn infant cry*. Almquist and Wiksell, Uppsala

Tseng C-Y. 1981. *An acoustic study of tones in Mandarin*. PhD dissertation, Brown University

Ullman M.T. 2004. Contributions of memory circuits to language: the declarative/procedural model. *Cognition,* 92: 311-270.

Vargha-Khadem F., Watkins K., Alcock K., Fletcher P. & Passingham R. 1995. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc. Natl. Acad. Sci. USA*, 92: 930-933.

Vargha-Khadem F., Watkins K.E., Price C.J., Ashburner J., Alcock K.J., Connelly A., Frackowiak R.S.J., Friston K.J., Pembrey M.E., Mishkin M. *et al.* 1998. Neural basis of an inherited speech and language disorder. *Proc. Natl. Acad. Sci. USA,* 95: 12695-12700.

Wang J., Rao H., Wetmore G. S., Furlan P.M., Korczykowski M., Dinges D. F. & Detre J.A. 2005. Perfusion functional MRI reveals cerebral blood flow pattern under psychological stress. *Proc. Natl. Acad. Sci. USA,* 102: 17804-17809.

Watkins K. E., Vargha-Khadem F., Ashburner J., Passingham R.E., Connelly A., Friston K.J., Frackowiak R.S.J, Mishkin M. & Gadianet D. G.2002. MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain,* 125: 465-478.

Zihao J, Angelucci D. E., Badal-Garcia E., d'Errico F., Daniel F., Dayet L., Douka K., Higham T.F., Martínez-Sánchez M.J., Montes-Bernárdez R. *et al.* 2010. Symbolic use of marine shells and mineral pigments by Iberian Neanderthals, *Proc. Natl. Acad. Sci. USA*, 107: 1023-1028.