The Evolution of Human Speech

Its Anatomical and Neural Bases

by Philip Lieberman

Human speech involves species-specific anatomy deriving from the descent of the tongue into the pharynx. The human tongue's shape and position yields the 1:1 oral-to-pharyngeal proportions of the supralaryngeal vocal tract. Speech also requires a brain that can "reiterate"—freely reorder a finite set of motor gestures to form a potentially infinite number of words and sentences. The end points of the evolutionary process are clear. The chimpanzee lacks a supralaryngeal vocal tract capable of producing the "quantal" sounds which facilitate both speech production and perception and a brain that can reiterate the phonetic contrasts apparent in its fixed vocalizations. The traditional Broca-Wernicke brain-language theory is incorrect; neural circuits linking regions of the cortex with the basal ganglia and other subcortical structures regulate motor control, including speech production, as well as cognitive processes including syntax. The dating of the FOXP2 gene, which governs the embryonic development of these subcortical structures, provides an insight on the evolution of speech and language. The starting points for human speech and language were perhaps walking and running. However, fully human speech anatomy first appears in the fossil record in the Upper Paleolithic (about 50,000 years ago) and is absent in both Neanderthals and earlier humans.

Although the focus of current linguistic research is syntax, speech is the derived feature of language, absent in even closely related living species. Speech allows us to transmit information faster than would otherwise be possible vocally. It also keeps words active in the neural computational space or verbal working memory in which the meaning of a sentence is discerned (Baddeley 1986). The neural substrate that regulates speech production appears to play a part in syntactic operations and other cognitive processes. Therefore, any account of the evolution of human language must account for the specialized anatomy and neural mechanisms that make speech possible.

I will briefly review the anatomy and physiology of speech, focusing on the species-specific anatomy of the human supralaryngeal vocal tract. I will then discuss reconstructing the supralaryngeal vocal traits of fossil hominids, taking account of recent studies of human ontogenetic development and the constraints imposed by swallowing. The findings of these studies provide a quantitative basis for inferring the speech-producing anatomy of Neanderthals and other fossil hominids.

I will then discuss the neural substrate that regulates speech production. Current findings refute the traditional theory localizing the neural bases of human language to Broca's and Wernicke's areas. These areas of the cortex play a role in speech and language, but they work in concert with other neural structures in circuits that link activity in these and other cortical areas to the basal ganglia and other subcortical structures. Evidence from neurophysiologic and behavioral studies of humans and other species show that the basal ganglia give human speech its "reiterative" quality, allowing humans to reorder a finite number of learned motor acts to form an almost unbounded store of words. Chomsky's most recent candidate for the productive capacity of syntax is a narrow faculty of language that is specific to humans and to syntax (Hauser, Chomsky, and Fitch 2002). The proposal here is that cortical-striatal-cortical neural circuits regulate syntax as well as speech production, yielding the productive qualities of syntax. Similar neural circuits grant cognitive flexibility and make possible seemingly unrelated human capacities such as composing music or dancing. I will endeavor to show that the evolutionary root of these human qualities is motor control. In this I claim no original insights; the credit goes to Karl Lashly (1951), who proposed that neural mechanisms originally adapted for motor control are the basis for syntax and human creative behavior. The isolation and dating of the human form of the FOXP2 gene, which governs the embryonic development of the subcortical structures that support

Philip Lieberman is the Fred M. Seed Professor of Cognitive and Linguistic Sciences and Professor of Anthropology at Brown University (Providence, RI 02912-1978, U.S.A. [philip_lieberman@brown.edu]). This paper was submitted 7 XII 04 and accepted 14 VI 06.

these neural circuits, provide insights on the evolution of human speech, language, and cognition.

The Anatomy and Physiology of Speech

The vocal signals of all terrestrial mammals are generated by filtering a source of acoustic energy through an airway through which maximum energy passes at frequencies termed "formants" (Fant 1960). For phonated sounds the source is a quasi-periodic series of puffs of air generated by rapidly opening and closing the vocal folds or cords of the larynx. The average fundamental frequency of phonation (F0) is perceived as the pitch of person's voice. In many languages, words are differentiated by changes in F0 over the course of a syllable, but vowel quality is largely conveyed by formant-frequency patterns enhanced by distinctions in duration (Hellwag 1781; Chiba and Kajiyama 1941; Fant 1960).

In humans, the airway above the larynx—the supralaryngeal vocal tract—continually changes its shape as we talk, producing a time-varying formant-frequency pattern. Aperiodic noise generated at a constriction in the vocal tract can also serve as a source of acoustic energy; the sound transcribed by the phonetic symbol [h] in English is essentially a vowel with a noise source generated by air moving through a fixed laryngeal opening.

Formant Frequencies

In short, the larynx provides the source of acoustic energy for vowels and other phonated speech sounds; the supralaryngeal vocal tract acts as an acoustic filter that determines the phonetic quality of the sounds. A given vocal-tract shape will let more acoustic energy through at a set of particular formant frequencies and local energy maxima occurring in inharmonic combinations. The lowest formant frequency is identified by the notation F1, the next highest as F2, the third as F3. For example, the vowels [i] and [u] of the words "see" and "sue" can be produced with identical F0's—different formant frequencies specify these vowels. As we talk, we change the vocal tract's shape and the resulting formant-frequency pattern. The relationship between formants, the laryngeal source, and speech signals may be clearer if one thinks about how sunglasses work. The difference between a pair of sunglasses that makes everything look blue and one that makes everything look pink is the balance of light energy frequencies that passes through the glasses. The tinted glass achieves these effects by attenuating—reducing the amount of light energy throughout a range of frequencies. The combination of frequencies that are least attenuated determines the color. The same source of light, sunlight, will provide a blue or pink world when filtered by different sunglasses. Formant frequencies may be thought of as the acoustic frequencies that the vocal tract allows to pass through it with minimum attenuation.

The Supralaryngeal Vocal Tract

The range of area functions and the overall length of the supralaryngeal vocal tract determine the formant frequencies that it can generate. In the eighteenth and nineteenth centuries tubes were used to model the vocal tract. The tubes acted as acoustic filters and reeds as sources of acoustic energy. Computer-implemented models are now used to determine the formant frequencies that particular vocal-tract shapes can produce (e.g., Henke 1966; Stevens 1972; Baer et al. 1991; Story, Titze, and Hoffman 1996). The adult-like human supralaryngeal vocal tract has a tongue with an almost circular sagittal (midline) contour forming two segments, a horizontal oral cavity (SVTh) and a vertical pharyngeal cavity (SVTv) of almost equal length (1:1 proportions) positioned at a right angle (fig. 1). Movements of the undistorted tongue in the space defined by the oral cavity and the pharynx can produce the abrupt midpoint 10:1 area-function discontinuities on which the format-frequency patterns of the quantal vowels [i], [u], and [a] depend. Computer modeling shows that the supralaryngeal vocal tracts of living nonhuman primates, whose tongues are almost entirely within their mouths, cannot produce these necessary discontinuities (Lieberman, Klatt, and Wilson 1969; Lieberman, Crelin, and Klatt 1972). Acoustic analyses of the vocalizations of nonhuman primates (e.g. Lieberman 1968; Fitch 1997, 2000*b*; Rendall et al. n.d.) are consistent with modeling studies. Mon-

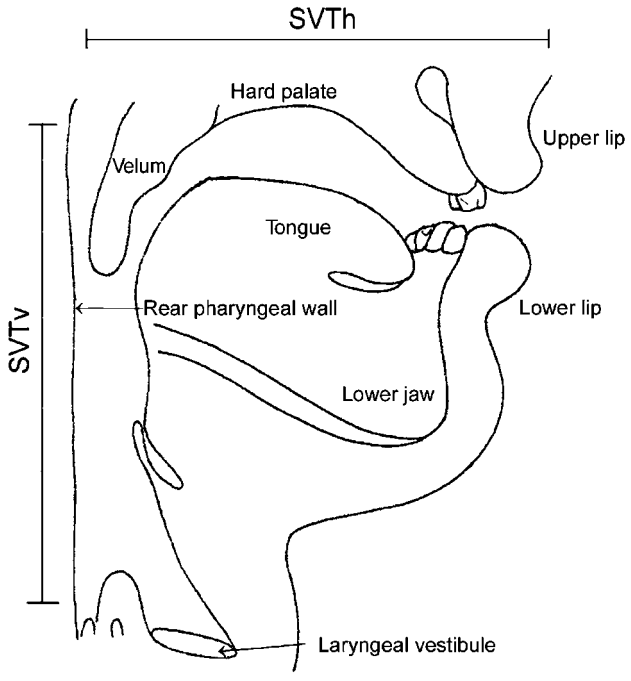

Figure 1. The adult human supralaryngeal vocal tract, showing the almost circular posterior contour of the tongue. The SVTh portion and the SVTv portion are almost equal in length. There is a natural discontinuity formed by the intersection of SVTh and SVTv that permits abrupt changes in the cross-sectional area of the human supralaryngeal vocal tract at its midpoint.

keys and apes produce schwa-like vowels (the vowel of the word "bub") because their tongues are positioned almost entirely in their mouths. One monkey species can produce two-formant frequency patterns that approximate a human [a] (Riede et al. 2005), but these vocalizations lack the third formant that would result from an [a]-like supralaryngeal vocal tract. They appear to be generated by the laryngeal air sacs' acting as resonators—a derived Diana monkey anatomical feature that has little relevance to the capabilities of ape and human vocal tracts that lack laryngeal air sacs (Lieberman 2006*b*).[1]

The Role of Quantal Vowels

Speech communication would be possible without quantal vowels. Indeed, there would have been no selective advantage for retaining whatever mutations led to the evolution of the human supralaryngeal vocal tract unless some form of speech had already been part of hominid culture. The term "quantal" was coined by Stevens (1972) to characterize speech sounds with perceptually salient acoustic properties that can be produced with a certain degree of articulatory sloppiness. The task of speech production is simplified when it is possible to produce a stable acoustic signal without having to execute exceedingly precise articulatory maneuvers. The task of speech perception is also more robust if the resulting acoustic signals are maximally distinct. These criteria are captured by Stevens's (1972) "quantal factor." The quantal factor can perhaps be illustrated by means of the following analogy: Suppose that the owner of a trendy restaurant wants to have his waiters transmit diners' orders with acoustic signals. Should he employ waiters equipped with violins or sets of handbells? If he wants to minimize the chance of errors, he will opt for handbells, each of which produces a distinct acoustic signal without requiring precise manual gestures.

Stevens demonstrated that the quantal vowels [i], [u], and [a] have perceptually salient acoustic correlates that can be produced while minimizing the need for precise motor control. Perceptual salience results from the convergence of two formant frequencies, yielding spectral peaks (Fant 1960) (fig. 2). For [i] the second and third formants, F2 and F3, converge at a high frequency; for [a], F2 and F1 converge at the midpoint of the frequency spectrum; for [u], F1 and F2 converge at a low frequency. Using quantal vowels would be similar to communicating with flags that have brilliant saturated colors. Other vowels, whose formants do not converge, produce formant patterns analogous to flags differentiated by pastel colors. Stevens demonstrated that if an abrupt area-function discontinuity occurs at the midpoint of the supralaryngeal vocal tract, the tongue can move back and forth as much as 1 cm without appreciably changing the formant frequencies. The exact position of the speaker's tongue with respect to the midpoint constriction for [i] does not have to be precise. Radiographic studies that track tongue movements confirm Stevens's theory (Beckman et al. 1995).

Carre, Lindblom, and MacNeilage (1995), using a different procedure, reached similar conclusions. Their computer model of the supralaryngeal vocal tract "grew" a vertical pharyngeal portion (SVTv) that was equal in length to its horizontal oral cavity (SVTh) when directed at producing the full range of human vowels delimited by [i], [u], and [a]. Radiographic and MRI studies show that the tongue body has a circular mid-sagittal posterior contour and is almost undeformed when we produce vowels. Producing an [i] involves moving the tongue upward and forward and producing an [a] moving the tongue back and down (Russell 1928; Chiba and Kajiyama 1941; Ladefoged et al. 1972; Nearey 1978; Baer et al. 1991; Story, Titze, and Hoffman 1996; Hiiemae et al. 2002). The human tongue and those of virtually all mammals are hydrostats (Stone and Lundberg 1996). Although muscular, the tongue cannot be squeezed into a smaller volume as we produce different vowels. The intrinsic muscles of the tongue are sometimes bunched up when speakers produce an [u] or an [i] (Fujimura and Kakita 1979), but the shape of the tongue is usually a segment of a circular arc when vowels are produced.

Vocal Tract Normalization

The vowel [i] also facilitates estimating the length of a speaker's supralaryngeal vocal tract. Longer ones yield lower formant frequencies than shorter ones for the same speech sound. Therefore, the absolute values of the formant frequencies of the same sound produced by different persons vary (Peterson and Barney 1952; Hillenbrand et al. 1995). A perceptual "normalizing" process that takes account of vocal-tract length is a critical step in speech perception. The role of vocal tract normalization in speech became evident in Peterson and Barney's (1952) study of vowel formant frequencies and vowel perception. Figure 3 shows the Peterson and Barney plot of the vowel formant frequencies of 76 adult male, adult female, and adolescent male and female speakers.

These frequencies were measured from spectrograms of each speaker's reading of a list of English words. The spoken words were identified without previously having listened to a long segment of speech produced by each particular speaker; listeners were presented with all of the words produced by ten speakers in random order. A vowel symbol that falls into a loop marked with the same phonetic symbol signifies a token that was heard as the intended vowel. The loops enclose the vowel tokens that made up 90% of the vowels that the speakers intended to convey. The loops overlap even though they do not include 10% of the stimuli that fell into a nearby vowel class. The data, for example, show that many speakers' [e] vowels had the same formant frequencies as other speakers' [ɪ]s (the vowels of the words "bet" and "bit"). The general findings of the Peterson and Barney study have been repli-

1. This rules out the possibility that nonhuman primate airways such as those of Diana monkeys could produce quantal vowels even if the degree of posterior pharyngeal expansion claimed by Riede et al. (2005) resulted in a 10:1 area function discontinuity because of its location (cf. Lieberman 2006*b*)

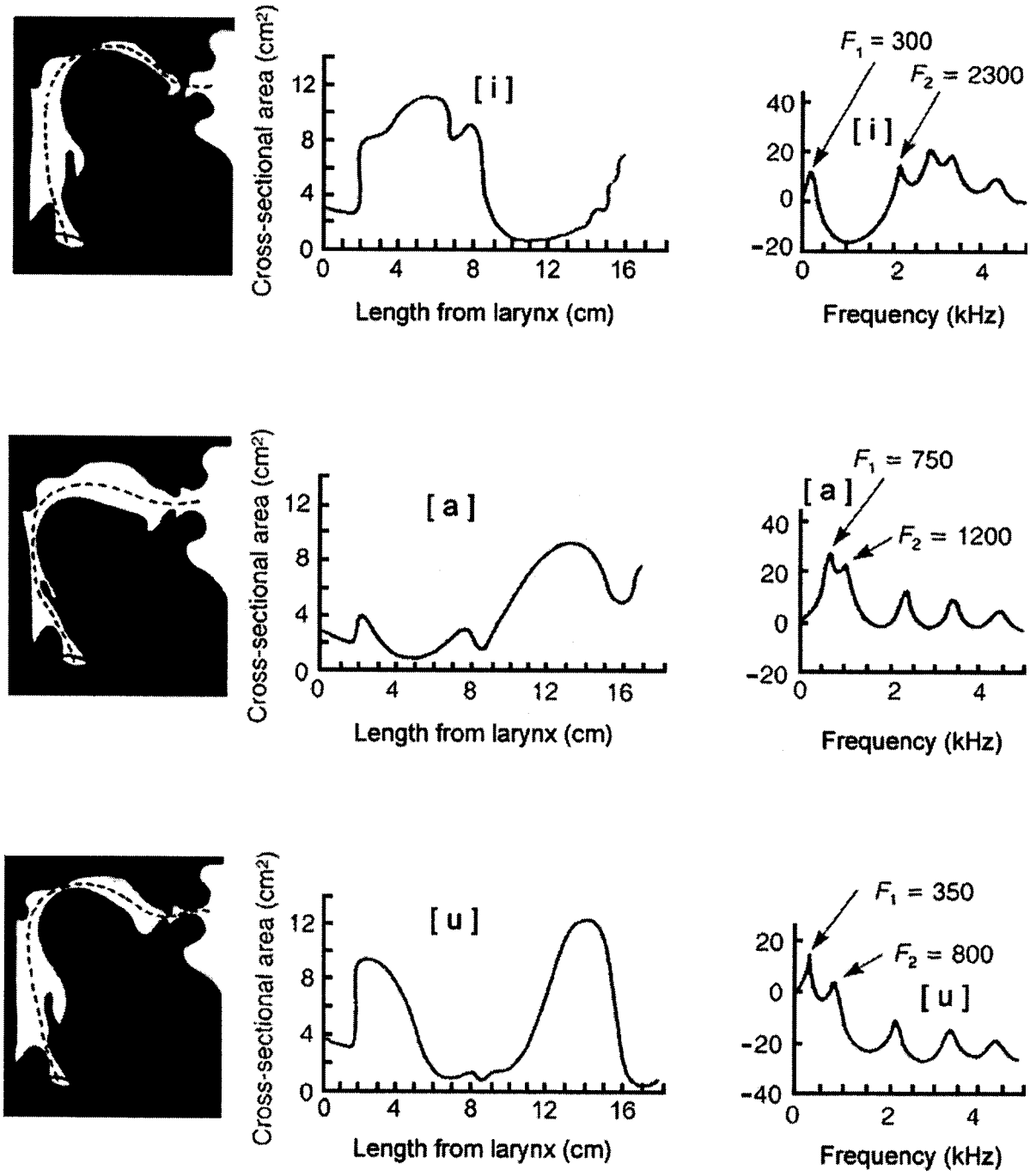

Figure 2. Midsagittal views of an adult human supralaryngeal vocal tract for the quantal vowels [i], [a] and [u] and the resulting formant-frequency patterns, showing the peaks in the frequency spectrum that follow from the convergence of two formant frequencies. The 10:1 discontinuity at the midpoint of the vocal tract allows speakers to be imprecise and still generate vowels that have spectral peaks.

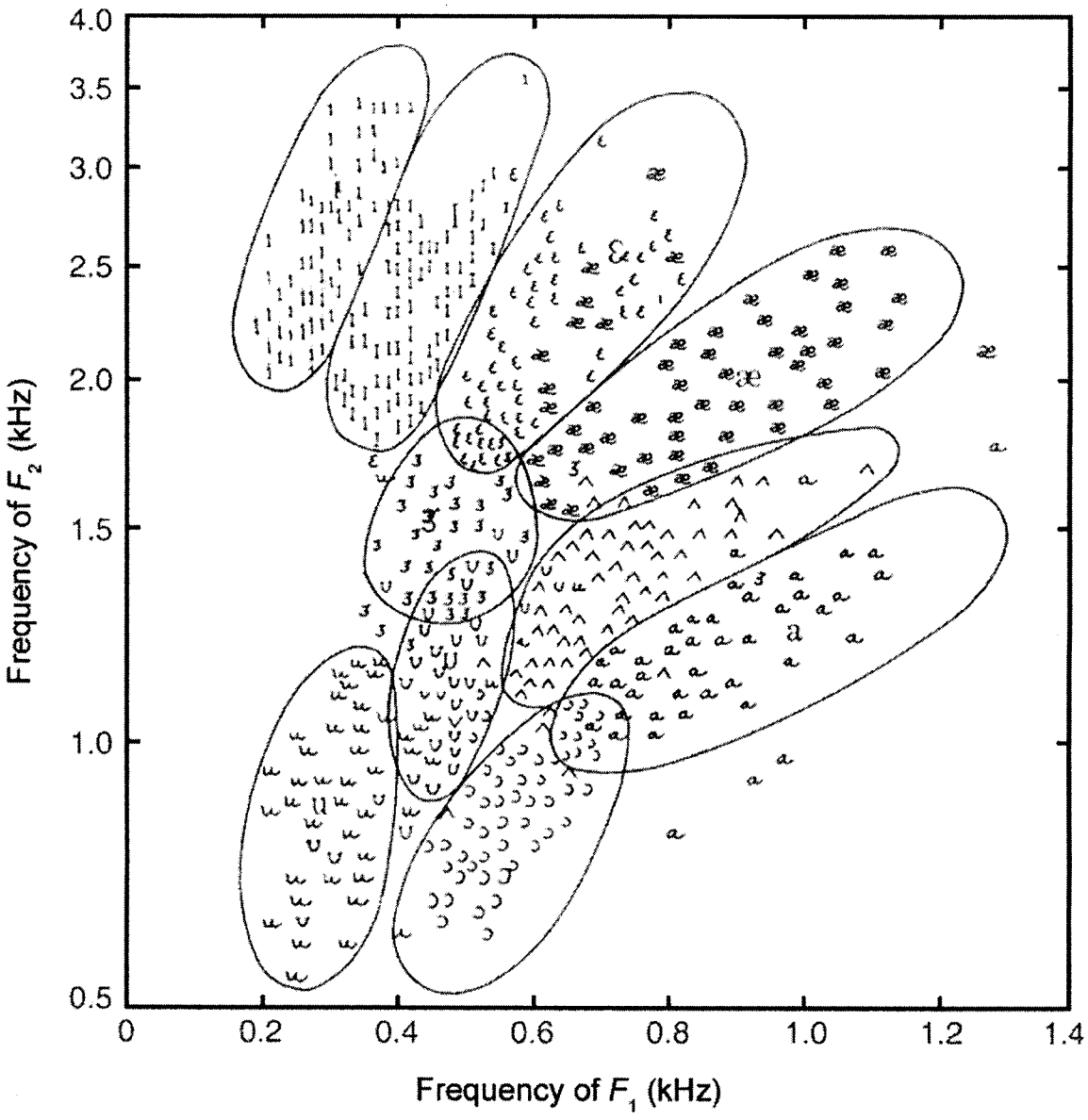

Figure 3. First and second formant frequencies derived by Peterson and Barney (1952) for the vowels produced by 76 speakers. The frequency of F2 is plotted with respect to the ordinate for each vowel token, the frequency of F1 with respect to the abscissa.

cated, using computer-implemented formant analysis, by Hillenbrand et al. (1995).

Human listeners can use different means to estimate a supralaryngeal vocal tract's length. Ladefoged and Broadbent (1957), for example, showed that the same tape-recorded word was perceived as "bit," "bat," or "but" depending on the average formant frequencies of a preceding phrase. But we generally do not need to hear a person talking before we identify a word; there are immediate normalization cues in the speech signal. There were only two errors in 10,000 trials in identifying [i]'s in the Peterson and Barney study; [u]'s had six identification errors, whereas [e] and [ɪ] were confused hundreds of times. Nearey (1978) reasoned that the formant structure of [i] might make it possible for listeners to use it as an anchor point for vocal tract normalization.

Nearey predicted that a token of a formant-frequency pattern in the [i] range would always be heard as an [i] produced by a supralaryngeal vocal tract of a particular length. If this were so, a listener would immediately "know" the length of a speaker's vocal tract and would correctly associate formant-

frequency patterns with the vowel that the speaker intended to convey. In a controlled experiment, listeners first heard a calibrating [i] followed by a synthesized formant-frequency pattern that could correspond to any vowel produced by either a short or a long vocal tract, followed by the same calibrating [i] (Nearey 1978, 98–149). Nearey used two different calibrating [i]'s, one produced by an adult male's long vocal tract and one produced by an adolescent's shorter one. Juxtaposed with the calibrating [i]'s were vowels having formant patterns that ranged over almost the total possible range of vowels for adult speakers and adolescents. Listeners heard isolated sequences that had the form [i]-V-[i], where the [i]s were either long- or short-vocal-tract [i]s and V the test stimuli. The listeners were told to identify each intermediate vowel V and to rate the naturalness of that vowel for each [i]-V-[i] sequence that they heard. The four categories of naturalness judgment ranged from "OK" to "very bad."

The listeners' responses showed that they were normalizing vocal tracts using the single token of an [i], changing their identification of the identical formant-frequency pattern when they heard it between long- or short-vocal-tract [i]s. The listeners' naturalness responses demonstrated that they interpreted these synthesized speech stimuli using a mental procedure that "knew" the range of formant frequencies that could be produced by the calibrating [i]'s vocal-tract length. For example, formant-frequency patterns that could be produced by a short vocal tract were judged to be natural when they were embedded with an [i] produced by a short vocal tract but judged to be nonspeech stimuli when they were embedded with an [i] from a long vocal tract that inherently could not produce such high formant frequencies. The V vowels clearly were perceived in a speech mode, using neural processing that took account of the speech-producing capabilities of the human supralaryngeal vocal tract.

Other speech sounds can be used for vocal tract normalization (cf. Lieberman 2006*a* for relevant studies), but the vowel [i] is an optimal calibrating sound. Its usefulness for vocal-tract length estimation follows from its unique formant pattern (high-frequency converging F2 and F3) and constraints on the vocal tract maneuvers that can be used to produce it. Whereas alternate gestures can be used to generate the formant-frequency patterns of virtually all other vowels, the tongue position and lip openings that generate an [i] are constrained (Stevens and House 1955; Nearey 1978). Speakers can protrude and constrict their lips to create the effect of having a longer vocal tract for most other vowels. Different tongue positions can be used for these vowels; speaker FSC in Nearey's (1978) study, for example, kept his tongue in almost the same position for almost all of his high F2 vowels, except for [i]. Alternate lip and larynx gestures generated his vowels' formant patterns. Fewer possibilities can generate the formant-frequency patterns for an [i]. The tongue must be placed forward and upward to the point where turbulent noise is sometimes generated in the constricted oral passage required (Fant 1960). The vowel [i] is an "honest" signal that specifies the speaker's actual vocal-tract length, and it is one of the speech sounds that a nonhuman vocal tract cannot produce.

The neural mechanisms for perceiving formant frequencies and deriving vocal-tract length appear to have a long evolutionary history. Other species appear to use formant frequencies to estimate the size of a conspecific. Fitch (1997) used a simple metric obtained by the subtracting the frequency of F1 from F3 to estimate a monkey's vocal-tract length, which is highly correlated with its body weight and length. This metric works for other species as well (Fitch 2000*a*), but only because these nonhuman-animal vocalizations are similar to the neutral schwa vowel of English, in which F3 is approximately equal to 5 (F1) (see the compilation in Riede et al. 2005). If the same metric were applied to human speech it would yield different estimated vocal-tract lengths for the same speaker depending on the vowel analyzed because the formant patterns produced by humans diverge from the schwa vowel.

Speech as the Default Mode for Language

Vocal communication frees a speaker's hands, can occur in darkness, and does not require looking at the individuals who are signaling. In addition and less obvious, speech allows the transmission of phonetic distinctions at rates of up to 20 to 30 segments per second, whereas other auditory signals merge into a continuous buzz at rates exceeding 15 items per second. It achieves this rapid transmission rate because it is an "encoded" signal in which information is transmitted at the slower syllable rate and then "decoded" into phonetic segments (Liberman et al. 1967). For example, the formant-frequency patterns that convey the phonemes of the word "cat" (approximated by the letters of the alphabet) are melded together into one syllable. As the tongue moves from the syllable-initial consonant, a formant-frequency pattern is produced that transitions into that of the vowel and then to that of the final consonant. Human speakers plan ahead. As one begins to say the word "too," one's lips "round" (protrude and narrow), anticipating the rounded [u] vowel. One's lips are not rounded at the start of the word "tea" because the following vowel is not rounded. The encoding differs somewhat from language to language (Lubker and Gay 1982) and is acquired without conscious effort by children.

Speech must have been present in hominid species that lacked supralaryngeal vocal tracts capable of producing quantal vowels because the shape of the human vocal tract increases the risk of choking to death on food lodged in the larynx. Palmer and his colleagues (1992, 187), reviewing studies of swallowing, note that, in contrast to nonhuman mammals, "normal humans are at risk for inadvertently inhaling food particles both before and after swallowing. Indeed, obstruction of the airway by inhaled food is a significant cause of morbidity and mortality in otherwise healthy individuals." Death resulting from a blocked larynx is often attributed to

other causes, but tens of thousands of incidents of fatal choking have occurred (Feinberg and Ekberg 1990). About 500,000 Americans suffer from swallowing disorders (dysphagia), and deaths from choking are the fourth-largest cause of accidental deaths in the United States (http://www.nsc.org/library/report_injury_usa.htm). There would have been no reason for retaining the mutations that resulted in the human vocal tract unless speech was already in place in hominids ancestral to humans.

Tracing the Evolution of the Human Vocal Tract

In attempting to reconstruct the soft tissue of the supralaryngeal vocal tract of a fossil when all that remains is bones, much attention has been given to the position of the larynx, which, as we will see, can rule out hypothetical vocal tracts. However, studies of the ontogenetic development of the human vocal tract reveal other factors:

1. The skeletal structure that supports the roof of the mouth rotates toward the back of the skull, effectively shortening the mouth and SVTh, during the first two years of life; the human face is flat compared with those of prognathous present-day apes and early hominids such as the australopithecines (D. Lieberman, Ross, and Ravosa 2000).

2. The human tongue gradually descends into the pharynx, changing its shape from relatively long and flat to posteriorly rounded. This yields the 1:1 SVTh/SVTv proportions seen in figure 1. This unique human developmental process is not complete until age six to eight years (D. Lieberman and McCarthy 1999). As the human tongue descends, it carries the larynx down with it.

3. The human neck gradually lengthens (Mahajan and Bharucha 1994). Neck length is critical in that a larynx positioned below the neck at the level of the sternum (collarbone) would make it impossible to swallow (Palmer et al. 1992; D. Lieberman et al. 2001).

As is the case in nonhuman primates throughout life, the tongue is positioned almost entirely in the mouth in human neonates. In the course of human ontogenetic development, the tongue moves down into the pharynx, carrying the larynx down with it. This process was first described by Victor Negus (1949, 25–26) who thought that it reflected the recession of the jaws:

> There is no prognathous snout . . . . The [human] tongue however retains the size it had in Apes and more primitive types of Man, and in consequence it is curved, occupying a position partly in the mouth and partly in the pharynx. As the larynx is closely approximated to its hinder end, there is of necessity descent in the neck; briefly stated, the tongue has pushed the larynx to a low position, opposite the fourth, fifth, and sixth cervical vertebrae.

Negus's inferences were correct insofar as extensive facial retraction occurs only in humans, but it has since become clear that the process entails more than the recession of the jaws, which occurs in the first two years of life. These findings were almost 30 years in the future when, in 1971, Edmund Crelin and I attempted to reconstruct the supralaryngeal vocal tract of the Neanderthal fossil from La Chapelle-aux-Saints (Boule 1911–13). We compared the skeletal features of the skull and mandible that support the soft tissues of the vocal tract in human newborns and in the Neanderthal fossil and noted a number of similarities between them. In addition to basicranial flexure, which became the focus of many subsequent studies, there were similarities in skeletal features supporting the muscles that move the tongue such as the pterygoid process of the sphenoid bone, the total length of the basicranium, and the distance between the end of the palate and the foramen magnum (into which the spinal column is inserted). On the basis of these findings, a range of vocal-tract area functions similar to those of newborns in the cineradiographic study of Truby, Bosma, and Lind (1965) was modeled using Henke's (1966) computer-implemented algorithm, which established the relationships between vocal-tract shapes and formant frequencies, and the computed formant-frequency vowel patterns were compared with those measured by Peterson and Barney (1952). Speech was possible because most vowel and consonant formant-frequency patterns could be produced, but the formant-frequency patterns that convey the quantal vowels of human speech could not be produced. The reconstructed Neanderthal's tongue rested for the most part in the oral cavity, and this precluded its producing the abrupt 10:1 area-function vocal-tract midpoint discontinuities required.

A number of studies subsequent to the Lieberman and Crelin (1971) paper attempted to determine the probable vocal tract of fossil hominids by establishing correlations between the cranial-base angle and the vocal tract in living nonhuman primates and then making inferences based on this angle in a fossil. A fossil that had a shallow cranial base similar to that seen in living apes and human newborns presumably had a similar vocal tract, while a fossil having a flexed adult human basicranial angle would have had a human vocal tract. Similarities between the embryonic and early stages of development have been used since Darwin (1964 [1859]) to make inferences concerning evolution. Therefore, George (1978) studied the Denver series of cephalometric X-rays, which tracked the development of basicranial skeletal features and the soft tissue of the vocal tract in subjects from age three months to adulthood (Maresh 1948; McCammon 1952) and correlated basicranial flexure with the occurrence of vowels that to her ears sounded like quantal vowels such as [i]. An acute adult-like cranial-base angle occurs at two years, when children appeared to produce quantal vowels. Since Stevens (1972) had shown that a vocal tract with adult proportions is necessary to produce these sounds, the conclusion was that the cranial-base angle was an index of vocal-tract proportions.

However, subsequent acoustic analyses showed that two-year-old children do not produce the formant-frequency patterns that specify quantal vowels. Buhr (1980) measured chil-

dren's vowel formant frequencies in the first years of life and found that they did not conform to those of adult speech. For example, the formant frequencies of a 64-week-old infant's vowels heard as [i] were actually those of [ɪ] (the vowel of "bit"). The difference in vowel quality was not apparent, however, even to trained phoneticians, when listening to these utterances (e.g., Irwin 1948). Patricia Kuhl and her colleagues (1992) solved the mystery by showing that when we listen to speech a "perceptual magnet" pulls an ill-formed formant-frequency pattern toward the ideal exemplar for the language that a person is exposed to in the early months of life. In effect, our speech perception system cleans up sloppy signals. The absence of computer-implemented digital image analysis technology in the 1970s precluded accurate measurements of tongue position by George; the perceptual-magnet phenomenon documented by Kuhl and her colleagues was not apparent until almost two decades later. In short, cranial-base flexure in itself cannot be used to predict whether a fossil had an adult human supralaryngeal vocal tract.

At the time, however, a close relationship between vocal-tract development and cranial-base angle was accepted by our and other research groups. Studies followed that linked the cranial-base angle and the length of the basicranium (which indicates oral-cavity length) with the vocal tracts of living nonhuman primates and fossil hominids (Laitman, Heimbuch, and Crelin 1978, 1979; Laitman and Heimbuch 1982). Their conclusion was that Neanderthals and earlier fossil hominids did not have human vocal tracts. The studies of Boe and his colleagues (Boe, Maeda, and Heim 1999; Boe et al. 2002) reached an opposite conclusion. Reconstructions of the vocal tracts of fossils based on cranial-base angles are problematic. When Daniel Lieberman and McCarthy (1999) reexamined the Denver series they found that the tongue and larynx continued to descend after cranial flexure stabilized and that SVTh and SVTv did not achieve their adult 1:1 proportions until age five to six years. Fitch and Giedd (1999), using MRIs, reached the same conclusion.

The low position of the human larynx is a reflex of the human tongue's reshaping and moving down into the pharynx. It is closely coupled to tongue displacement (Negus 1949; Bosma 1975; D. Lieberman and McCarthy 1999; Nishimura et al. 2003). As the tongue descends into the pharynx, it carries the larynx down with it. The descent of the tongue into the pharynx, its posterior circular shape, and the right-angle bend at its midpoint enable the human vocal tract to produce the major midpoint area-function discontinuities necessary for quantal vowels. Thus, despite the focus on the larynx in many studies on the evolution of speech, the descent and change in shape of the tongue are the key factors in the development and evolution of the human vocal tract (Lieberman 1984, 276–80).

Studies of species whose tongues are positioned in their mouths (e.g., Fitch 1997, 2000*a*) show that their vocalizations are limited to the schwa vowel. Nonhuman vocal-tract phonetic limitations characterize the deer vocalizations studied by Fitch and Reby (2001). Although the deer have low larynges, their tongues remain anchored in their long mouths. This is also the case for lions, whose larynges transiently descend as they roar; an elastic membrane links the larynx to a tongue anchored in the mouth (Weisengruber et al. 2002). The larynges of young chimpanzees descend somewhat through elongation of the distance between the hyoid bone and the larynx, but their tongues do not descend (Nishimura et al. 2003). In short, in itself a low larynx is not an indicator of potential phonetic ability. Claims such as Fitch's (2000*b*) that the human vocal tract evolved to produce lower formant frequencies by laryngeal descent (providing a false vocal impression of a larger body) cannot account for the evolution of the species-specific human vocal tract, which involves the descent of the tongue into the pharynx.

The biological mechanisms that regulate the descent and reshaping of the human tongue are unknown, and tongue position and shape cannot be inferred from the basicranial angle. Boe and his colleagues (Boe, Maeda, and Heim 1999; Boe et al. 2002) nonetheless base their Neanderthal reconstruction on the cranial-base angle of the La Chapelle-aux-Saints fossil as reconstructed by Heim (1989). The basicranial flexure of Heim's Neanderthal skull reconstruction is within the human range, but that does not signify an adult human vocal tract. Although the studies of D. Lieberman and McCarthy (1999) and Fitch and Giedd (1999) are cited, Boe and his colleagues ignore their findings and fit a vocal tract with the adult human proportions noted by Honda and Tiede (1998) to the fossil.

The relationships between skulls, jaws, and soft tissue noted by Honda and Tiede (1998) hold for adult humans; they do not apply to young children, human neonates, apes, or monkeys. Genetic evidence (Krings et al. 1997; Ovchinnikov et al. 2000) shows that Neanderthals diverged from humans about 500,000 years ago, and their skeletal morphology differs from that of modern humans (Howells 1976, 1989; D. Lieberman 1995). Adult human vocal-tract morphology therefore cannot arbitrarily be bestowed on them. Nonetheless, Boe and his colleagues model the vocal-tract shapes that adult human speakers use to produce vowels. Not surprisingly, these configurations produce the full range of human vowels. They also model a putative human infant vocal tract that does not resemble any newborn vocal tract documented by Negus (1949), Truby, Bosma, and Lind (1965), Bosma (1975), Laitman and Crelin (1976), or anyone else. Its SVTh/SVTv ratio is close to that of the five-to-six-year-old children documented by Lieberman and McCarthy (1999) and Fitch and Giedd (1999). Similar flaws mark other studies that have proposed human vocal tracts for Neanderthals (see Lieberman (1984, 2000, 2006*c* for reviews).

Dating the Modern Human Supralaryngeal Vocal Tract

A vocal tract that can produce the full range of human speech must have 1:1 SVTh/SVTv proportions. If SVTh is long, as

is the case for Neanderthals, then SVTv must also be long. But the anatomy involved in speech (tongue, hyoid bone, and larynx) has a more "primitive," basic function—eating. The hyoid, which supports the larynx, moves upward and forward about 13 mm, opening the esophagus and placing the larynx in a position in which food will not fall into it while swallowing (Ishida, Palmer, and Hiiemae 2002). A larynx in the neck can execute these maneuvers, but if the cricoid cartilage of the larynx were placed in the chest the sternum bone would make them impossible. The movements that are involved in swallowing are similar in humans and apes (Palmer et al. 1992). No human or ape descended from our common ancestor has a larynx in its chest, because it would not be able to eat.

We can determine whether Neanderthals and other fossil hominids could have had 1:1 SVTh/SVTv proportions by examining their basicrania, which provide a measure of SVTh, and their cervical vertebrae, which provide a measure of the length of their necks. McCarthy et al. (n.d.) determined these metrics for a sample of 62 specimens of *Pan troglodytes*, the WT 15000 fossil *Homo ergaster*, 3 Neanderthal fossils, and 82 specimens of *H. sapiens*, including the Middle Paleolithic Skhul V fossil, 8 Upper Paleolithic fossils, and 73 contemporary humans from seven different populations. The data show that Neanderthal necks were too short to accommodate human vocal tracts. McCarthy and his colleagues arrive at a Neanderthal neck length estimate of 120 mm in contrast to the 134–127-mm averages for two modern human samples; the short neck and long Neandertal SVTh would place the cricoid cartilage behind the sternum, permitting human speech but precluding eating. (A similar conclusion was reached by Lieberman [1984, 290–96].)

Surprisingly, a similar constraint rules out a human vocal tract in the Middle Pleistocene fossil Skhul V (McCowan and Keith 1939), which has often been thought to be fully modern. McCarthy and his colleagues estimate the cervical spine length of Skhul V to be 109 mm, at the bottom of the adult modern human range. Skhul V's SVTh is relatively long, and therefore its short neck precludes its having a fully human vocal tract with 1:1 SVTh/SVTv proportions. Fully modern speech anatomy is not evident in the fossil record until the Upper Paleolithic, about 50,000 years ago.

The Neural Substrate

It is clear that human speech entails having neural capabilities that are absent in closely related living species. Although a chimpanzee's vocal tract would suffice to establish vocal language, it cannot talk, despite the fact that acoustic analyses (e.g., Lieberman 1968) reveal "bound" formant-frequency patterns in chimpanzee calls similar to those that convey different words in human speech. These sounds could be used to differentiate words if the chimpanzees could voluntarily reorder the motor commands used to generate them. Chimpanzees could establish "protospeech," producing everything save quantal sounds, if they were able to freely reiterate—to reorder and recombine the motor commands underlying speech. Chimpanzees calls in the state of nature appear to be stereotyped and fixed (Goodall 1986). The neural circuits that confer the reiterative abilities necessary for human speech appear to be absent in chimpanzees and other nonhuman primates.

The reiterative quality of these human neural circuits extends to other aspects of behavior, including syntax. The studies that support this claim also show that the traditional Broca-Wernicke "language organ" theory is wrong. Cortical-striatal-cortical neural circuits that include the basal ganglia appear to regulate motor control, syntax, and cognition. The subcortical basal ganglia constitute a "sequencing engine" that can reiterate motor commands stored as motor pattern generators in other parts of the brain. The basal ganglia through different anatomically segregated neural circuits also reiterate cognitive pattern generators conferring cognitive flexibility and take part in associative learning. The evolutionary significance of the regulatory FOXP2 gene, which has erroneously been identified as a "language gene," rests in the fact that it governs the embryonic development of the basal ganglia and other subcortical elements of these neural circuits (see Lieberman (2000, 2002, 2006*a* and the studies noted below).

Complex brains contain many distinct neuroanatomical structures that in normal circumstances process particular tactile, visual, or auditory stimuli, while other structures and cortical regions perform local operations that regulate aspects of motor control or hold information in the short-term (working) memory, (e.g., Marsden and Obeso 1994; Mirenowicz and Schultz 1996; Monchi et al. 2001; Polit and Bizzi 1978; Sanes et al. 1995). However, an isolated structure or cortical area usually does not by itself regulate a complex behavior. Individual neural structures generally contain many anatomically segregated groups, "populations," of neurons that carry out a particular local operation. The local processes do not constitute an observable behavior. The neuronal population that carries out a local process "projects" to anatomically distinct neuronal populations in other regions of the brain. The series of linked neuronal populations form a neural circuit. The circuit constitutes the brain basis of an observable aspect of behavior—walking, talking, striking the keys of a computer's keyboard, and so on. Moreover, within a given neural structure, distinct anatomically segregated neuronal populations may occur that project to neurons in different brain structures, forming multiple circuits each of which regulates some other behavior.

As Dobzhansky (1973) put it, "Nothing in biology makes sense except in the light of evolution." Neural structures that were initially adapted to control one function took on new tasks. Seen in this light, local motor sequencing operations in the subcortical basal ganglia appear to be precursors for similar operations in cognitive domains. The basal ganglia can alter a motor act when circumstances dictate by switching

from one motor pattern generator to a more appropriate one, and during a thought process they can switch from one cognitive pattern generator to another (Graybiel 1997). For example, within the putamen, a subcortical basal ganglia structure, anatomically segregated populations of neurons form part of a system for sequencing the motor submovements that together constitute an overt movement of a monkey's hand, a rat's grooming sequence, or a person's walking or speaking (Aldridge et al. 1993; Cunnington et al. 1995; Lieberman 2000; Marsden and Obeso 1994). The putamen in itself is not the "seat" of these motor acts; it connects the submovement pattern generators to areas of motor cortex. Anatomically segregated neuronal populations in the putamen project through other subcortical structures to cortical areas implicated in higher cognition, comprehending the meaning of a sentence, attention, and reward-based learning (e.g., Alexander, DeLong, and Strick 1986; Alexander and Crutcher 1990; Cummings 1993; Graybiel 1995, 1997; Kimura, Aosaki, and Graybiel 1993; Marsden and Obeso 1994; Middleton and Strick 1994).

Experiments-in-Nature and the Broca-Wernicke Model

The study of the neural bases of human language began with experiments-in-nature that produced aphasia (permanent loss of linguistic abilities) after parts of the brain were destroyed by accidents, strokes, or other pathologies. Experiments-in-nature are still germane to the brain-language question, particularly when their findings are integrated with tracer, imaging, and electrophysiological studies. Paul Broca's (1861) observations arguably rank with the most influential such experiments. However, the interpretation of brain-behavior relationships presented here is quite different from Broca's.

Broca's patient, "Tan," had had a series of strokes. The strokes had caused extensive brain damage including but not limited to one part of the brain, the third frontal convolution (an anterior area of the cortex). Tan had limited speech ability and uttered only the syllable "tan." Broca, perhaps influenced by earlier phrenological theories (Spurzheim 1815), concluded that damage to this cortical region, which includes Broca's area, was the basis of the patient's speech deficit. If one's model of the brain is that discrete localized regions regulate observable complex behavior, it follows that destroying a region should disrupt a particular aspect of behavior. Overlooked was the fact that Tan also had extensive subcortical damage and nonlinguistic motor impairments. Wernicke (1967 [1874]) found that patients who had suffered damage in the posterior left hemisphere had difficulty comprehending speech. Again, he localized receptive linguistic ability to this neocortical area. Since language involves both comprehending and producing speech or alternate phonetic systems such as writing or sign language, Lichtheim (1885) proposed a cortical pathway linking Broca's and Wernicke's areas. According to this model, spoken language is perceived in Wernicke's area, a posterior temporal region associated with auditory perception. A cortical pathway then transmits information to Broca's region, which is adjacent to cortical areas implicated in motor control.

Although the Broca-Wernicke model has the virtue of simplicity, it is at best incomplete. The behavioral deficits of Broca's aphasia are not limited to speaking; they involve difficulty comprehending distinctions in meaning conveyed by syntax and word-finding difficulties (Blumstein 1995). Patients also suffer from cognitive deficits. Kurt Goldstein (1948) characterized Broca's aphasia as "loss of the abstract capacity" and noted an inability to adapt to changing circumstances. Contemporary clinical evidence shows that permanent loss of language does not occur absent subcortical damage, even when Broca's or Wernicke's area has been destroyed. For example, although magnetic resonance imaging (MRI) showed almost complete destruction of Wernicke's area in a 60-year-old patient, he made a full recovery; no subcortical damage was apparent (Lieberman 2000, 101–2). Moreover, damage to subcortical structures, sparing cortex, can produce aphasic syndromes. Computer-aided tomography scans and MRI now provide information on the nature and extent of brain damage that produces permanent language loss. Aphasia does not occur unless subcortical damage is present (Stuss and Benson 1986; Dronkers et al. 1992; D'Esposito and Alexander 1995).

Other studies show that subcortical damage that leaves Broca's area intact can result in Broca-like deficits in speech production and language (e.g., Naeser et al. 1982; Benson and Geschwind 1985; Alexander, Naeser, and Palumbo 1987). Alexander and his colleagues (1987), for example, noted the subcortical locus of aphasias, reviewing 19 cases of aphasia that resulted solely from subcortical lesions. The language deficits ranged from fairly mild impairment in a patient's ability to recall words to global aphasia, in which a patient produced very limited speech. In general, the severest language deficits occurred in patients who had suffered the most extensive subcortical brain damage and damage to the internal capsule (the nerve fibers that project to the cortex). Subsequent studies rule out damage to the internal capsule as causing aphasia. Deliberate surgical lesions of the internal capsule aimed at mitigating obsessive-compulsive behavior do not induce aphasia (Greenberg, Murphy, and Rasmussen 2000). Damage to the basal ganglia from strokes in the medial cerebral artery which passes through them may be the locus of Broca's aphasia. As D'Esposito and Alexander (1995, 41) conclude, "that a purely cortical lesion—even a macroscopic one—can produce Broca's or Wernicke's [aphasia] has never been demonstrated."

Cortical-Striatal-Cortical Circuits

The basal ganglia are subcortical structures located deep within the brain. They can be traced back to anurans similar to present-day frogs (Marin, Smeets, and Gonzalez 1998). The striatal component of the basal ganglia includes the caudate nucleus and the lentiform nucleus. The lentiform nucleus, in

turn, consists of the putamen and the globus pallidus. The putamen receives sensory inputs from most parts of the brain. The globus pallidus is an output structure receiving inputs from the putamen and the caudate nucleus. The caudate nucleus, the putamen, and the globus pallidus form a system with close connections to the substantia nigra, the thalamus, other subcortical structures, and the cortex. The thalamus, in turn, is connected to different cortical areas. The connections with the cortex are complex (Alexander, DeLong, and Strick 1986; Parent 1986; Alexander and Crutcher 1990; DeLong 1993; Marsden and Obeso 1994; Middleton and Strick 1994).

Seemingly unrelated disruptions in behavior such as obsessive-compulsive disorder (Greenberg, Murphy, and Rasmussen 2000), schizophrenia (Graybiel 1997), and Parkinson's disease (Jellinger 1990) derive from the disruption of neural circuits linking cortical areas with the basal ganglia. Behavioral changes usually attributed to frontal-lobe cortical dysfunction can be observed in patients with damage to basal ganglia (e.g., Cummings and Benson 1984; Flowers and Robertson 1985; Alexander, DeLong, and Strick 1986; Lange et al. 1992; DeLong 1993).

Cummings (1993) identifies five parallel basal ganglia circuits involved in motor control, cognition, attention, and other aspects of behavior. The circuit (probably circuits) projecting to the dorsolateral region of the prefrontal cortex is associated with cognitive behavior. Tracer studies confirm these circuits. These studies entail injecting substances into living animals that attach themselves to the outputs of neurons projecting to other neurons forming neural circuits. Postmortem sectioning, staining, and microscopic examination then reveal the neural pathways. Tracer studies of monkey brains confirm that the striatal basal ganglia (the caudate nucleus and the putamen) support circuits that project to cortical areas associated with motor control and cognition (Alexander, DeLong, and Strick 1986; Middleton and Strick 1994; Graybiel et al. 1994; Graybiel 1995, 1997). Noninvasive diffusion tensor imaging techniques, based on MRI technology, show similar neural circuits in humans (Lehericy et al. 2004).

Parkinson's disease damages the basal ganglia, mostly sparing the cortex (Jellinger 1990). The primary deficits of Parkinson's disease are motoric—tremors, rigidity, and movement disruptions. Speech production deficits similar to those occurring in Broca's aphasia also occur. Patients have difficulty sequencing the lip, tongue, and laryngeal maneuvers necessary to differentiate "stop" consonants. Stop consonants are produced by momentarily obstructing the supralaryngeal vocal tract with the lips (for [b] and [p]) or the tongue (for [d], [t], [g], and [k]). The lips or the tongue open the vocal tract, producing a momentary burst, an abrupt pulselike acoustic signal. The larynx must then produce phonation keyed to the burst. Phonation must occur within 20 msec. from the burst for the English voiced stops [b], [d], and [g] (the initial consonants of the words "bad," "dab," and "god"). Phonation must be delayed, usually for at least 60 msec., for the English unvoiced stops [p], [t], and [k] (the initial consonants of "pad," "tab," and "cod"). Lisker and Abramson (1964) called this phonetic distinction, which entails controlling the sequence of gestures between tongue or lips and the muscles of the larynx, "voice-onset time."

Similar voice-onset-time distinctions differentiate the stop consonants of all human languages analyzed to date. (Many languages also differentiate words by means of stops in which voicing starts before the burst.) Acoustic analyses show that a breakdown in the regulation of voice-onset time is the most symptomatic speech deficit of Broca's aphasia (Blumstein et al. 1980; Baum et al. 1990) and Parkinson's disease (Lieberman et al. 1992). In contrast, formant-frequency patterns that reflect vocal-tract maneuvers are generally preserved in both Broca's aphasia and Parkinson's disease (Blumstein 1994; Lieberman 2000).

As is the case for Broca's aphasics (Blumstein 1995), Parkinson's disease patients (Illes et al. 1988) can have difficulty producing sentences that have complex syntax. They also have difficulty comprehending sentences that have moderately complex syntax as well as long sentences that tax the brain's computational resources (e.g., Lieberman et al. 1992; Natsopoulos et al. 1993; Grossman et al. 1991, 1993; Lieberman 2000; Hochstadt 2004). As the disease progresses, dementia that differs in kind from Alzheimer's occurs (Cummings and Benson 1984). Afflicted patients retain semantic and real-world knowledge but have difficulty forming or changing cognitive sets (Flowers and Robertson 1985; Cools et al. 2001). These seemingly unrelated deficits derive from the local operations performed by the basal ganglia in the cortical-striatal-cortical circuits that regulate these aspects of behavior.

Basal Ganglia Operations

The basal ganglia operations characterized by Graybiel (1995, 1997, 1998) involve both motor pattern generators and cognitive pattern generators. In the era before medication with Levadopa was used to treat Parkinson's disease, thousands of operations were performed. The effects were reviewed in a seminal paper by Marsden and Obeso (1994), who noted (p. 889) that the basal ganglia had two different motor control functions:

> First, their normal routine activity may promote automatic execution of routine movement by facilitating the desired cortically driven movements and suppressing unwanted muscular activity. Secondly, they may be called into play to interrupt or alter such ongoing action in novel circumstances. . . . They respond to unusual circumstances to reorder the cortical control of movement.

Marsden and Obeso concluded (p. 893):

> Perhaps the basal ganglia are an elaborate machine, within the overall frontal lobe distributed system, that allows routine thought and action, but which responds to new cir-

> cumstances to allow a change in direction of ideas and movement. Loss of basal ganglia contribution, such as in Parkinson's disease, thus would lead to inflexibility of mental and motor response. . . .

Brain-imaging studies of human subjects confirm this inference. The event-related functional-MRI study of Monchi et al. (2001) shows the role of basal ganglia in the shifting of cognitive sets. Brain activity was monitored in neurologically intact subjects in a version of the Wisconsin Card Sorting Test (WCST), which evaluates a person's ability to form and shift cognitive criteria. Subjects had to sort cards by matching the images on them to the colors, shapes, or number of images on "match" cards. As predicted, neural circuits involving the prefrontal cortex and the basal ganglia were activated throughout the test. Bilateral activation was observed in the prefrontal cortex, the basal ganglia, and the thalamus. Dorsolateral prefrontal cortical areas were active at the points where the subjects had to relate the current match to earlier events stored in working memory. A cortical-striatal-cortical circuit involving a different cortical area (mid-ventrolateral prefrontal cortex), the caudate nucleus, the putamen, and the thalamus was active when subjects had to shift to a different matching criterion. Increased activity occurred in the putamen during these cognitive shifts. The behavioral study of Scott and his colleagues (2002) complements these findings. A comprehensive set of cognitive tests that assess frontal-lobe functions such as planning and tests of memory was administered to Parkinson's disease patients who had undergone neurosurgery that produced precise bilateral lesions of the globus pallidus. The sole deficits occurred on the WCST, in which the subjects were unable to shift the matching criterion as the test progressed.

Stowe et al. (2004) used PET imaging of neurologically intact subjects in a sentence-comprehension study that involved a form of set shifting. The basal-ganglia-to-dorsolateral-prefrontal-cortex circuit was active when subjects had to change their interpretation of an ambiguous sentence, confirming that basal ganglia cognitive set shifting also manifests itself in language. Other neuroimaging studies show basal ganglia as well as cortical activity during sentence-comprehension and word-retrieval tasks (Klein et al. 1994; Kotz et al. 2003; Rissman, Eliassen, and Blumstein 2003).

The focus on subcortical structures here in no way implies that the cortex is irrelevant. The imaging studies noted above and many other studies show that Broca's area is active when a person listens to speech, when a person recalls a word, and in comprehending the meaning of a sentence or identifying its emotional content. Areas in both hemispheres of the cortex are active in these tasks, including the right-hemisphere homologues of Broca's and Wernicke's areas and prefrontal areas that are not traditionally associated with language (Just et al. 1996). The absence of basal ganglia activity in other imaging studies may reflect region-of-interest procedures that did not look for subcortical activity during linguistic tasks.

Electrophysiologic studies that monitor brain activity in monkeys by means of exceedingly fine microelectrode probes show that the basal ganglia perform similar functions (reviewed in Graybiel 1995, 1997, 1998) in monkeys as well as in other mammals. When the basal ganglia of rats are destroyed, they are able to execute the individual submovements that when linked together would constitute a grooming sequence (Berridge and Whitshaw 1992), but they cannot perform the complete sequence. The rodents' basal ganglia neurons show firing patterns that sequentially inhibit and release submovements to the motor cortex, thereby stringing them into a grooming sequence (Aldridge et al. 1993).

It is generally not possible to compare the behavior of human subjects before and after an insult to the brain, nor is it ethically justifiable to test theories by placing subjects in a situation that might harm their brains. However, the climbing of Mount Everest provides a unique, ethically sound situation in which the effects of basal ganglia dysfunction on motor control, language, cognition, and other aspects of behavior can be determined. The cognitive abilities of individual subjects can be assessed before and after hypoxic (oxygen-deficit) insult to their brains, allowing the assessment of subtle as well as profound impairment. Metabolically active neural structures such as the basal ganglia are particularly sensitive to hypoxia (Inoue et al. 1992; Burke et al. 1994). Independent studies show that the globus pallidus is extremely sensitive to hypoxic damage (Laplane et al. 1984, 1989; Strub 1989). MRI imaging confirms bilateral lesions localized to the globus pallidus after exposure to altitude; the lesions produce subcortical dementia and aphasia (Jeong et al. 2002; Chie et al. 2004).

A series of experiments (Lieberman et al. 1994, 2005) shows that speech production deficits similar to these of Parkinson's disease occur as climbers ascend to higher altitudes. Voice-onset sequencing is impaired, and speech slows as the length of vowels increases. Cognitive tests such as the WCST administered at successively higher altitudes slow that set-shifting performance declines. Sentence comprehension also slows down, and error rates increase. In extreme cases, hypoxic climbers exhibiting profound speech and set-shifting errors fail to adapt their behavior to changing life-threatening events. Shifts in personality similar to those reported by Cummings (1993) for damage to cortical-striatal-cortical circuits also occur.

Motor Control and Reiterative Ability

Many linguists (e.g., Jackendoff 1994; Chomsky 1995) still hold to the view that human language bears little relation to the communication or thought of any other animal. Chomsky, whose focus has been on syntax for many years, has consistently argued that human syntactic ability involves some unique feature whose scope is restricted to language. As noted earlier, the most recent candidate (Hauser, Chomsky, and Fitch 2002) is a narrow faculty of language that allows recursion. Chomsky's (1957) initial generative syntactic theory proposed that the relative clause in the sentence "I saw the

boy who was wearing a sweater" was the end product of a process in which a hypothetical underlying sentence, "The boy was wearing a sweater," had been inserted into the frame of the carrier sentence "I saw the boy." Subsequent hypothetical transformational rules of the generative grammar then rewrote the resulting string of words to yield the sentence—"I saw the boy who was wearing a sweater"—that would actually be heard or read. Traditional grammars would straightforwardly characterize the actual, observable sentence as containing a relative clause.

In Chomsky's (1995) current minimalist grammar the syntactic rule "merge" recursively inserts sentences and other syntactic units into the framework of a carrier sentence; the minimalist syntactic rule "move" then rewrites the resulting string of words to yield the sentence that one actually hears or reads. The reiterative function of the basal ganglia includes reordering and replicating cognitive pattern generators (Graybiel 1997). The cognitive pattern generator that elicits the relative clause "who was wearing a sweater" would simply be inserted into the frame of the carrier sentence. In short, the basal ganglia sequencing engine can form a potentially infinite number of different sentences by reordering, recombining, and modifying a finite set of words using a finite set of linguistic rules. Reiteration can account for the sentences that we actually hear—inserting a relative clause, a prepositional clause, and so on, into a carrier phrase. In principle, the linguistic process is no different from inserting the dance instruction "allemande right" into a square dance or yet another variation into a rondo. Reiteration can also account for the formal phonologic operations used by linguists to describe word-level phonologic processes such as the formation of regular English plural nouns by adding the sounds coded by suffix "s" (e.g., "book" versus "books").

Chomsky and his colleagues are correct in proposing processes that can generate a potentially infinite number of sentences or words from a finite set of words and rules. However, as we have seen, the ability to reorder and recombine a finite set of elements to form an infinite set of actions is a key feature of speech motor control, other motor acts, and aspects of nonlinguistic cognition such as changing the direction of one's course of action or changing the criterion by which one categorizes objects. Many linguists may argue that language is quite different from motor control. Forming different grammatical sentences entails more than inserting a phrase or word or changing word order. The semantic-syntactic constraints on the words in any dictionary, including that in the brain, must be taken into account. Different verbs, for example, have particular constraints (the linguistic term generally used is "argument structure"). For example, the ungrammatical sentence "I wished Ann" violates a constraint because the verb "wish" cannot refer to an object, whereas "I kissed Ann" is acceptable. Motor control entails similar and, indeed, more complex constraints. As the basal ganglia release and inhibit successive pattern generators, these constraints come into play. Walking, for example, involves a sequence of submovements. Heel strike, one component of walking, can be executed only after the motor pattern generator that swings the lower leg forward, and the pattern generator that locks the legs in place when one is standing still cannot be followed with heel strike. Running, which appears to have shaped human evolution (Bramble and D. Lieberman 2004), requires exceedingly rapid and precise control of a different set of pattern generators. If walking or running seems too simple, consider the set of sequential motor commands involved in baseball, playing the violin, or dancing the tango. In short, motor pattern generators have argument structures.

The FOXP2 Gene

The embryonic development of neural structures that regulate motor control, other aspects of cognition, and emotional regulation, as well as the development of lung tissue and other structures, is governed by the FOXP2 gene. While it is undoubtedly not the only regulatory gene involved in the evolution of human language, studies of it provide some insights on the evolutionary history of the human brain. The discovery of FOXP2 results from a sustained study of a large extended family marked by a genetic anomaly. A syndrome—a suite of speech and orofacial movement disorders and cognitive and linguistic deficits—occurs in afflicted members of the KE family (Vargha-Khadem et al. 1995, 1998; Lai et al. 2001; Watkins et al. 2002). Afflicted individuals are unable to protrude their tongues while closing their lips. They have difficulty repeating two-word sequences. They have significantly lower scores on standardized intelligence tests than their nonafflicted siblings. Some have higher nonverbal IQ scores than unaffected members of the family, suggesting to some investigators that FOXP2 does not affect intelligence. However, as the different nonverbal IQs for the nonaffected members of the KE family show, intelligence is derived from the interaction of many neural systems and life experiences. It is impossible to know what the nonverbal IQs of an affected individual would have been absent the genetic anomaly, but the low mean nonverbal IQ of the affected members (86, with a range of 71–11) versus a mean of 104 (with a range of 84–119) for unaffected family members suggests that FOXP2 anomalies are responsible for generally lower intelligence.

MRI studies of affected family members shows that the caudate nucleus is abnormally small bilaterally, while the putamen, the globus pallidus, the angular gyrus, the cingulated cortex, and Broca's area are abnormal unilaterally (Vargha-Khadem et al. 1995, 1998). Watkins et al. (2002) note that reduced caudate nucleus volume is "significantly correlated with family members' performance on a test of oral praxis, non-word repetition, and the coding subtest of the Wechsler Intelligence Scale." Functional-MRI studies that compare afflicted members of the KE family with both their unaffected siblings and age-matched controls show that underactivation occurs in the putamen, Broca's area, and its right homologue (Liegeois et al. 2003), which is what would be expected in

neural circuits connecting the striatum and Broca's area (Lehericy et al. 2004). The pattern of neural anomalies and behavioral deficits is similar to that seen in individuals afflicted with Parkinson's disease, hypoxia, and lesions in basal ganglia.

This constellation of neural anomalies and behavioral deficits results from a dominant point mutation mapped to chromosome 7q31 in the FOXP2 gene (Fisher et al. 1998; Lai et al. 2001). Lai and her colleagues determined the neural expression of FOXP2 during early brain development in humans and of foxp2 in mice (Lai et al. 2003)—mammalian end points separated by 75 million years of evolution (Mouse Genome Sequencing Consortium 2002). The gene encodes a protein that regulates the expression of other genes during embryogenesis. Mutations to other similar genes have been implicated in a number of developmental disorders. In the case of family KE, the mutation changes an amino acid, apparently leading to protein dysfunction. The similar areas of expression that indicate where the gene is active in both the human and the mouse brain include structures in the cortical-striatal-cortical circuits that regulate motor control and cognition—the thalamus, the caudate nucleus, and the putamen as well as the inferior olives and the cerebellum. These structures are all intricately interconnected. Independent evidence shows that foxp2 in other mammals is expressed in the putamen as well as the caudate (Takahashi et al. 2003).

The FOXP2 gene provides a means to date the evolution of the human brain and the emergence of fully human speech capabilities. Despite the high degree of similarity, there are important distinctions between the mouse, chimpanzee, and human versions. The mouse and human versions are separated by three mutations, the chimpanzee and human versions by two. Enard et al. (2002), using the techniques of molecular genetics, estimate that the human form appeared fairly recently, sometime in the last 100,000 years—in the time frame (Stringer 1998) associated with the emergence of anatomically modern *H. sapiens*.

Walking, Running, and the Antiquity of Speech

One point concerning the evolution of human speech deserves more emphasis—its antiquity. The Lieberman and Crelin (1971) Neanderthal study is often cited to support claims that speech evolved abruptly at a recent date. Boe et al. (Boe, Maeda, and Heims 1999; Boe et al. 2002) claim that we concluded that Neanderthals were a "speechless species." However, this was not our conclusion. What we wrote was that Neanderthals represent "an intermediate stage in the evolution of language. This indicates that the evolution of language was gradual, that it was not an abrupt phenomenon. The reason that human linguistic ability appears to be so distinct and unique is that the intermediate stages in its evolution are represented by extinct species" (Lieberman and Crelin 1971, 221). Some form of speech must have been in place in the archaic hominids ancestral to both humans and Neanderthals. There would have been no selective advantage for retention of the mutations that yielded the species-specific human supralaryngeal vocal tract at the cost of increased morbidity from choking unless speech was already present. The question is when.

The basal ganglia dysfunction that is the proximate cause of Parkinson's disease impairs walking; patients have difficulty executing the internally guided sequential movements involved. Running is impossible. The Hoehn and Yahr (1967) diagnostic scale for Parkinson's disease is a measure of upright balance and locomotion. As we have seen, the subcortical structures whose expression is regulated by FOXP2—the basal ganglia and the cerebellum—play a critical role in motor control, motor learning, and cognition. Learning to execute a motor sequence involves activity in these subcortical structures as well as the prefrontal cortex (e.g., Kimura, Aosaki, and Graybiel 1993; Thatch 1996). Selection for walking, starting from the base apparent in present-day chimpanzees, which can walk for limited periods, was perhaps the starting point for the evolution of human speech, language, and cognition. The evolution of the genus *Homo* was marked by adaptations for endurance running (Bramble and D. Lieberman 2004), which places still further demands on the basal ganglia sequencing engine. Lacking more data, we can only speculate that a neural substrate permitting voluntary speech motor control was in place in early *H. erectus*. Further selection for speech production may have resulted in the human form of FOXP2 and the motor, cognitive, and linguistic abilities of contemporary humans. Developmental-neurophysiologic studies comparing the development of walking and speech may move this proposal beyond speculation.

Putting Speech Anatomy and the Brain Together

The findings discussed here concerning the evolution of human speech anatomy and the human brain point to the same conclusion. The evolution of speech was driven by Darwinian natural selection, the opportunistic use of existing structures adapted for another purpose, and mutations on regulatory genes that had far-reaching consequences. Contemporary human speech and cognitive capabilities, including enhanced syntactic and lexical abilities, are species-specific properties of *H. sapiens* derived from anatomy and neural mechanisms that appear to have coevolved. The FOXP2 gene is clearly implicated in the formation of neural circuits that regulate human cognitive and motor capacities. Natural selection acting on the mutations that yielded its human form would have enabled rapid, encoded speech, in turn enhancing the selective value of the mutations that shaped the modern human vocal tract. These events, which led to the emergence of fully modern speech, language, and cognition, appear to have occurred sometime in the period between 90,000 and 50,000 BP, the

time frame between fossils like Skhul V and fully modern humans who were capable of talking and acting (Klein 1999) as we do.

Acknowledgments

My recent research, noted in this paper, has been supported by NASA under grant NCC9–58 with the National Space Biomedical Research Institute.

Comments

Shirley Fecteau and Hugo Théoret
Center for Non-Invasive Brain Stimulation, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Ave, Boston, MA 02215, U.S.A. (sfecteau@bidmc.harvard.edu)/ Psychology, University of Montreal, CP 6128, Succ. Centre-Ville, Montréal, Qc H3C 3J7, Canada (hugo.theoret@umontreal. ca). 30 VIII 06

While we agree with Lieberman that the neural network underlying language processing is more complex than previously assumed, he may have underestimated the contribution of Broca's area. In light of recent brain imaging data suggesting important nonlinguistic functions relevant to language development in the inferior frontal gyrus, revisiting the role of Broca's area in language would strengthen Lieberman's model of speech evolution.

He mentions in the section "A Laboratory Called Mount Everest" that it is difficult to compare human behavior before and after a brain lesion. Noninvasive brain stimulation such as repetitive transcranial magnetic stimulation (rTMS) allows the study of the behavioral consequences of transient, externally induced brain disruption in healthy individuals. Such "virtual neurology" (Rafal 2001) establishes a causal relationship between brain and behavior, obviating most of the confounds inherent to natural lesions, such as neural reorganization following insult, size of lesion, and lesion site (Robertson, Théoret, and Pascual-Leone 2003). With regard to the question at hand, disruption approaches appear to contradict Lieberman's assertion that "aphasia does not occur unless subcortical damage is present." Indeed, it has been repeatedly shown that disruption of the left inferior frontal gyrus by rTMS is associated with speech arrest (e.g., Pascual-Leone, Gates, and Dhuna 1991; Stewart et al. 2001), suggesting that a healthy Broca's area is necessary for proper speech. The recovery of a lost behavior following brain damage does not mean that the lesioned area is not causally related to a given behavior. Indeed, behavior is never the result of a lesion but rather the consequence of the way the rest of the brain is capable of sustaining function following a brain lesion, and brain plasticity plays a major role in behavioral recovery (Pascual-Leone et al. 2005). In the case of language, overactivation of the right-hemisphere homologue of Broca's area has been reported in aphasia patients (Belin et al. 1996; Rosen et al. 2000; Martin et al. 2005), and it is assumed that some of these activations contribute to speech recovery. From this it follows that one should be very careful when drawing conclusions from lesion data, given that plastic events subtenting recovery of function can be very complex. For example, some of the novel right-hemisphere activations that follow left-hemisphere lesions of Broca's area may actually be *detrimental* to functional recovery from aphasia (Naeser et al. 2005).

Most relevant to our own understanding of Broca's area is Lieberman's mention that "the behavioral deficits of Broca's aphasia are not limited to speaking." Broca's area is the human homologue of monkey area F5, where mirror neurons (cells that discharge both when a monkey performs an action and when it sees someone else perform the same action) were first discovered. It is widely assumed that Broca's area is the main component of the human mirror neuron system (Rizzolatti and Craighero 2004). Importantly for theories of language, this system appears to be modality-independent, responding to both visual and auditory depiction of the same action. The existence of an auditory mirror neuron system suggests that some brain regions are involved in the production and perception of speech, in accordance with the motor theory of speech (Liberman et al. 1967). For example, hearing speech activates the motor cortex representation of the tongue and lip muscles that would be used if the heard phonemes were actually spoken (Fadiga et al. 2002; Watkins, Strafella, and Paus 2003). Importantly, these modulations of activity in the motor cortex are correlated with cerebral blood flow in Broca's area (Watkins and Paus 2004).

Finally, recognizing the expanded role of Broca's area in action perception/execution may also be relevant to our understanding of the *evolution* of language. Reading aloud increases excitability of the motor cortex representation of the dominant hand (Tokimura et al. 1996), and rTMS over Broca's area significantly impairs finger imitation (Heiser et al. 2003). Following on this, it has been proposed that changes in the primate brain (perhaps within area F5) may have supported the use of hands for communication and its evolution towards language in humans (Rizzolatti and Arbib 1998; Arbib 2005). In that sense, Broca's area and its mirror neuron system would have played a crucial role in the evolution of human speech from gestural communication in monkeys.

Despite these reservations, Lieberman is right in suggesting an increased role for motor processes in speech and language. In light of increasing data establishing functional correspondence between perception and execution of actions and speech, we believe that Broca's area is an ideal candidate for reconciling divergent views of speech evolution.

Ricardo R. García and Francisco Aboitiz
Depto. de Psiquiatría, Escuela de Medicina, Pontificia Universidad Católica de Chile, Marcoleta 387 2°piso, Santiago, Chile (rgarcia@med.puc.cl; faboitiz@puc.cl). 16 VIII 06

After extensively reviewing the evolutionary and neurobiological foundations of the development of the human supralaryngeal vocal tract, Lieberman focuses on the neural basis of human speech. Fundamentally, he proposes three main hypotheses: (1) that the neural substrate that regulates speech may be involved in syntactic operations and other cognitive processes, and thus cortical-striatal-cortical neural circuits could regulate syntax and speech production; (2) that the basal ganglia may have conferred the reiterative quality of human speech; and (3) that the neural mechanisms adapted for motor control are the basis for syntax and human creative behavior and that the isolation of the FOXP2 gene, which is linked to the development of subcortical structures supporting these neural circuits, may provide insights into the evolution of human language and speech.

His emphasis on motor systems, the basal ganglia, and FOXP2 is consistent with recent reports that link this gene with the evolution of articulate speech (Vargha-Khadem et al. 2005). For Lieberman, the basal ganglia include a cognitive central pattern generator that permits the shifting of behavioral strategies according to context and thus generates a "recursive" or "generative" motor output as in syntax. Coordinated interaction between the basal ganglia and the prefrontal cortex would subserve the development of complex functions such as syntax and speech. However, human syntax is largely known as a lateralized cortical function located in the left hemisphere. In this conceptual frame, it is possible to ask whether, if the basal ganglia are so important for language, there should not be some lateralization of this structure. Stowe et al. (2004) report a right basal ganglia involvement in a syntactic task during sentence comprehension, but further studies using neuroanatomical and imaging technology are needed to elucidate this critical point.

As Lieberman mentions, the human phonatory system differs from the ape and monkey vocalizing systems in the degree of cortical control; thus it is quite possible that the enhancement of cortical control was an important event in the development of speech and language. In this context, we (Aboitiz and García 1997; Aboitiz et al. 2006) have emphasized the existence of a cortical sensorimotor auditory-vocal circuit which was probably present in the monkey and which served as the precursor for the cortical language circuits in the human brain (i.e., Broca's and Wernicke's areas). This proposal has been confirmed by recent neuroimaging studies in the monkey (Gil-da-Costa et al. 2006). Thus the development of language is probably the result of a complex adaptive process involving adaptations at both the cortical and the subcortical level. One main contribution of the cortical circuits was probably related to the expansion of short-term working or active memory circuits, allowing the differentiation of a phonological loop that permitted mental storage and the rehearsal of complex phonological sequences. With the expansion of these short-term memory capacities, based on associative interactions (Fuster 1995), these phonological sequences were able to acquire relatively complex meanings and eventually more complex syntactic forms, which require important working memory capacities and cortical activation (Friederici 2004). Lieberman (2000, 81) proposes that a verbal working memory, instantiated in the human brain as a dynamic distributed network, would allow the comprehension of the meaning of sentences considering semantic and syntactic information coded in words and pragmatic factors.

Finally, another issue is how brain size comes into the picture. *Homo erectus* had a small brain, and there is evidence that our brains acquired their modern dimensions long before there are any direct signs of complex symbolic thinking. This implies that cultural evolution must have had an important impact on language evolution.

Ann MacLarnon
Centre for Research in Evolutionary Anthropology, School of Human and Life Sciences, Roehampton University, Holybourne Avenue, London SW15 4JD, UK (a.maclarnon@roehampton.ac.uk). 15 IX 06

Reconstructing the evolutionary history of human speech—identifying the constituent human capabilities and establishing their evolutionary paths—is a complex process, and very few factors are universally agreed upon. Lieberman identifies what he believes are unique anatomical and neurological human features necessary for fully modern speech and presents evidence against others' claims. He goes on to discuss the evolution of these features, many of which are not directly detectable from hard tissues. For example, he concludes that the neurological basis for the complex motor control required for human speech was largely present from early *Homo,* alongside adaptations for prolonged bipedal walking and running which similarly require rapid and precise control of motor patterns. He suggests that at some point between this time and the appearance of the Neanderthals, which had well-developed speech capabilities, this neurological basis was activated for use in speech. The possible time range is broad and does not conflict with the evolutionary schemas of most others. It also clarifies Lieberman's position that Neanderthals would have been able to produce most, though not quite all, of the vowels and consonants in the modern human array.

Lieberman goes on to deduce from other features that earlier modern humans, such as Skhul, still did not have fully developed speech capabilities and that these evolved only very recently, less than 50,000 years ago. The feature which is crucial to this deduction is the human ability to produce quantal vowels. Lieberman argues that the ability to produce these

sounds, [i], [u], and [a], is uniquely human and dependent on the marked descent of the posterior part of the human tongue down the pharynx. This results in a 1:1 ratio between the lengths of the horizontal and vertical portions of the tongue and the consequent ability to produce an abrupt discontinuity in the vocal tract at the point of inflection, which is the common property of the quantal vowels.

Reconstruction of the supralaryngeal vocal tract through human evolution and the shape and distribution of the tongue, perhaps the most important speech articulator, has been a fraught area of the paleoanthropological literature. Lieberman clarifies the development of his own thinking, and, in particular, on the basis of developmental evidence, he rules out the use of basicranial form to infer supralaryngeal vocal tract shape—the descent of the larynx and tongue. He also explains that laryngeal descent does not necessarily involve the descent of the posterior tongue as some writers have assumed. The evolution of the modern human tongue position must therefore be inferred otherwise, and Lieberman and colleagues favour using basicranial measures to estimate the length of the horizontal portion of the tongue (SVTh) and using the length of the cervical vertebrae to estimate the vertical portion (SVTv). This results in the conclusion that Skhul, with its relatively prognathous jaws, could not, given the length of its neck, have had an SVTh:SVTv ratio of 1:1. Therefore this feature of later modern humans must have evolved after the time of earlier modern humans, and along with it the ability to produce quantal vowels. Following Lieberman's argument, this leads to the multiple evolutionary origin of an important and possibly unique feature of fully modern human speech. I believe it is more probable that the argument is at some point flawed.

The combined reconstruction of the two parts of the supralaryngeal vocal tract is certainly not easy, and Skhul is not so markedly prognathous compared with later modern humans. There is certainly room here for further debate. In addition, as Lieberman states, recent work on the vocalizations of some non-human primates questions the uniqueness of formant production for human speech sounds. Even if the formant production of Diana monkeys reported by Riede et al. (2005) turns out to be very different from that of modern humans, making use of laryngeal air sacs, for example, further work and the inclusion of additional species are needed to clarify this aspect of human differentness.

On the basis of the evidence presented, I believe it is premature to claim that fully modern human speech evolved only as recently as 50,000 years ago and that the clarity of communication afforded by quantal vowels was not available until that time. This would suggest that our understanding of the relation between speech, language, and cultural development in later human evolution needs to change. At present I believe the more parsimonious approach is to await further work on the formant production of other primates and to be more cautious about the interpretation of reconstructions of the supralaryngeal vocal tract.

Robin Melrose
School of Language and Area Studies, University of Portsmouth, King Henry First St., Portsmouth, Hants. P01 2DZ, UK (robin.melrose@port.ac.uk). 30 VIII 06

Lieberman has made a valuable contribution to the debate on language evolution by claiming that human speech in its present form cannot have developed until some 50,000 to 90,000 years ago with the emergence of the human form of the FOXP2 gene. The reasons for this claim are of great interest to someone who operates within the framework of systemic functional linguistics (see Halliday 2004; Halliday and Matthiessen 1999; Melrose 2005). Although language in some form existed prior to this—Deacon (1997) believes that the ability to communicate symbolically (rather than indexically, as our primate ancestors did) may date back over 2 million years—it seems that it was only with the evolution of anatomically modern basal ganglia (particularly the caudate nucleus) and basal ganglia–prefrontal circuits that rapid encoded speech became possible.

Although rapid phonological encoding of speech is the main benefit conferred by anatomically modern basal ganglia, Lieberman also sees another benefit. Hauser, Chomsky, and Fitch (2002) claim that the faculty of language in the narrow sense includes only the computational mechanisms for *recursion* (reiteration), providing the capacity to generate an infinite range of expressions from a finite set of elements. And this precisely, for Lieberman, is one of the roles of the basal ganglia: as he says, the basal ganglia sequencing engine can form a potentially infinite number of different sentences by reordering, recombining, and modifying a finite set of words using a finite set of linguistic "rules."

From the perspective of systemic functional linguistics, recursion is only one aspect of language. Whenever we make an utterance, we are drawing on meanings from four *metafunctions*: the experiential (the resources for encoding our experiences of the external world and the inner world of consciousness), the interpersonal (the resources for intruding our judgements, evaluations, attitudes, and comments into the ongoing speech event), the textual (the resources for making language relevant to the context, both nonverbal and verbal), and the logical (the resources for encoding our perception of relations between things and events—resources of recursion such as coordination, subordination, and embedding).

The basis for Lieberman's claim about the development of anatomically modern basal ganglia seems to be the numerous studies carried out on the KE family, half of whose members are affected by a severe language disorder resulting from a genetic mutation. As Watkins, Dronkers, and Vargha-Khadem (2002) demonstrate, the speech of the affected members of the KE family is virtually unintelligible, since they appear to have severe problems with "sequential articulation of phonological units" (p. 461) and a deficit in the use of morpho-

syntax (though Watkins, Dronkers, and Vargha-Khadem believe that this may be again a phonological problem).

The neural basis of this language disorder is explored by Vargha-Khadem et al. (1998) and Watkins et al. (2002). Although structural abnormalities in the caudate nucleus may be the key to the KE family's speech disorder, there are other abnormalities. Most notably, the left SMA (BA 6) is underactive and has less gray matter (as do the left and right sensorimotor cortex [BA 3/4]): studies of patients with aphasia have shown that lesions in the inferior frontal gyri and motor cortex (BA 6, 44, 4), the primary somatosensory cortex (BA 43, 3, 1, 2), and the caudate are associated with deficits in nonlinguistic action comprehension (see Saygın et al. 2004), implying that BA 6 plays a significant role in encoding our experience of the external world, at least as far as actions are concerned. In addition, the left anterior cingulate is underactive and the preSMA/cingulate cortex has less gray matter: the anterior cingulate is associated with theory of mind (see Gallagher and Frith 2003), without which it would be difficult to make our utterances relevant to the context, and with the inhibition of positive attitudes (see Wood et al. 2005), which is essential for ensuring that the feelings we intrude into the ongoing speech event are appropriate ones.

While it is not possible to say whether the mutation in the FOXP2 gene conferred any benefits other than the capacity for rapid speech—that is, speech as we know it today—the other deficits found in the KE family, to say nothing of the data showing that certain areas are *overactive* or have *more* gray matter, are suggestive. What is certain is that the issues raised in Lieberman's article will be debated for some time to come.

Tobias Riede
National Center for Voice and Speech, 1101 13th St., Denver, CO 80204-5319, U.S.A. (tobiasriede@web.de). 11 IX 06

A thorough understanding of the evolution of speech requires an interdisciplinary approach and the analysis of comparative data (Jackendorff and Pinker 2005). Lieberman presents reviews of two important areas in the evolution of human speech (vocal tract anatomy and the neural basis of speech production), both of which invite critical comments.

The first comment refers to his interpretation of the role of the tongue during vocalization in animals. Contrary to what Lieberman points out, the animal tongue is not inert but likely to play a significant role when animals vocalize (Fitch 2000*a*; Riede et al. 2000*b*, 2006; Beckers, Nelson, and Suthers 2004). For example, lateral cineradiographic images of two cotton-top tamarin monkeys (*Saguinus oedipus*) and other species during vocal behavior suggest significant caudocranial movement of tongue and larynx (Fitch 2000*c*). Interpretations remain somewhat preliminary because of small sample size, limited time resolution, and a missing second imaging plane. Second, investigations of the tongue's motor neurons suggest that there is increased cortical control over tongue movement from nonprimate mammals to New World primates to Old World primates (Alipour, Chen, and Jürgens 1997, 2002). It is also relevant that sophisticated tongue movement has been observed in nonhumans during swallowing (Thexton and McGarrick 1988, 1989; Kobara-Mates et al. 1995; Hiiemae, Hayenga, and Reese 1995, or in learned motor tasks (Murray et al. 1991). Finally, changes in the acoustic signal are directly related to cross-sectional area changes in the vocal tract; whether the resulting constrictions are achieved by tongue movement or by other maneuvers such as laryngeal movement or a change of head position is less important.

Another basic claim of Lieberman is that nonhuman primate vocal tracts are uniform, in contrast to the human one, which is best modeled as a series of different-sized tubes. Our recent studies on Diana monkeys have brought new insights. These monkeys are able to produce two acoustically different alarm calls for two of their predators, eagles and leopards (Zuberbühler 2000*a*, *b*, *c*, *d*; Riede and Zuberbühler 2003*a*, *b*; Riede et al. 2005, 2006). The main call difference is in the formant characteristics. Eagle alarm calls have formant features resembling an /a/ vowel; leopard alarm calls have formant transitions from the /a/ to the /o/ vowel (Riede and Zuberbühler 2003*b*), comparable to the English expression "Ouch!" Riede et al. (2005) conducted an anatomical investigation of these monkeys' vocal tract, and the simplest model that fitted the observed anatomy proved to be a vocal tract consisting of three tubes of different diameters.

Lieberman is also incorrect in some of his statements concerning laryngeal air sacs and their acoustic role. Many nonhuman primates possess laryngeal air sacs, including (in contrast to Lieberman's account) all four great apes (Negus 1949; Schneider 1964). There is no evidence that this structure has an effect on formant position (Riede et al. 2006). One experimental study of DeBrazza's monkey (*Cercopithecus neglectus*) suggests that air sacs have a direct impact on amplitude in the higher frequency range but not on the actual position of the formants (Gautier 1971).

Vocal learning, a crucial feature of speech production in humans, is discussed only briefly by Lieberman. The gross anatomical connections of monkey and human brains are virtually identical (Deacon 1989), but there are clear functional differences (Jürgens 2002). For instance, a bilateral lesion in the inferior motor cortex completely eliminates the capacity to speak in humans but has no effect on vocal communication in monkeys (Jürgens 1999*a*, *b*). Because speech consists of learned vocal patterns, monkey calls and human nonverbal utterances most likely represent genetically programmed vocal patterns (Brokelman and Schilling 1984; Owren et al. 1992). The human motor cortex is needed for the production of learned but not innate motor patterns; patients unable to speak are sometimes still able to produce nonverbal utterances (Sem-Jacobsen and Torkildsen 1960)

such as laughing, crying, whimpering, or moaning (Owren and Bachorowski 2001, 2003; Owren 2003). Lesions in other parts of the brain, however, produce identical effects on vocal behavior in monkeys and humans. Loss of function in the periaqueductal gray of the midbrain, for instance, causes mutism in monkeys (Jürgens and Pratt 1979) and humans (Esposito et al. 1999). The idea emerging in the field of basal ganglia research is that cortico–basal ganglia circuits promote learning of motor patterns through trial-and-error learning guided by a reward-sensitive mechanism (Doya and Sejnowski 1995). In songbirds, which share the vocal-learning feature with humans, the anterior forebrain pathway, akin to a cortico–basal ganglia loop, is necessary for song learning in young birds, which copy a template of a tutor song during a critical period (Doupe and Kuhl 1999). The song is highly variable, but with practice and feedback it becomes stereotyped. The anterior forebrain pathway regulates song variability in a real-time modus (Kao, Doupe, and Brainard 2005; Olveczky, Andalman, and Fee 2005 much as basal ganglia regulate variability in mammalian (nonvocalization-related) motor patterns (Barnes et al. 2005). Furthermore, songbirds accurately recognize acoustic patterns defined by a recursive, self-embedding, context-free grammar (Gentner et al. 2006), apparently sharing the capacity for syntactic recursion that has been assumed to be the computational core of a uniquely human language faculty. Also, the transcription factor foxp2 is found in all mammals (Enard et al. 2002) and birds (Haester et al. 2004). Studies in songbirds show that during times of song plasticity foxp2 is up-regulated in a striatal region essential for song learning, suggesting that foxp2 plays important roles both in the development of neural circuits and in postnatal behavior. Teramitsu et al. (2004) describe this expression as similar in human brains.

Ian Tattersall
Division of Anthropology, American Museum of Natural History, New York, NY 10024, U.S.A. (iant@amnh.org). 2 VIII 06

Lieberman has produced an illuminating survey of the anatomical and neurological backgrounds for human speech and language, and he argues very plausibly that there is an intimate motor connection between these two functions. He also notes that the modern human vocal tract is unique not simply in the descended larynx of the adult but in the backward rotation early in life of the bony support structures of the superior vocal tract. This effectively retracts the mouth, shortening it. In combination, these two processes produce a tongue that is divided more or less equally between the mouth and the necessarily lengthened neck, and it is this feature that permits the production of the formant frequencies that are an essential attribute of speech. Laryngeal position in itself, on which so much attention has been focused over the past quarter-century, is only part of the story, and, indeed, Lieberman seems to argue that the low-lying modern human larynx is simply a passive consequence of the restructured tongue.

Applying this perspective to the fossil record, Lieberman shows convincingly that the apparatus essential to the production of articulate speech is seen only in anatomically modern humans (a group from which he wisely excludes hominids of the kind found at Skhūl). Yet he also argues that precursor hominid species must have had the ability to produce speech because only this capacity could possibly have countervailed against the grave selective disadvantage of a descended larynx: the ability to choke to death. This is, however, only very arguably the case, and Lieberman himself is at pains to point out that new structures often—one might even argue, always—arise independently of functions for which they are later co-opted. Quite simply, we have no idea what the reasons were for the acquisition of the reconfigured vocal tract, but we know that it must have been in place *before* it could be exploited for speech production.

The pattern this observation predicts appears to be borne out by the fossil and archaeological records. Anatomically modern human beings were around in Africa and even in the Levant long before there is any evidence that they were *behaving* in a "modern" way. Stone tools, however sophisticated, provide at best highly equivocal suggestions of "modernity," and clear demonstration of modern symbolic behavior patterns can be made only when ancient hominids are associated with demonstrably symbolic objects. Language is the ultimate symbolic behavior, and the modern human anatomy that permits its expression had long been established by the time that we find any convincing archaeological evidence for symbolic activity among hominids. It is thus permissible to infer that the first anatomically modern humans were not yet the symbolic creatures that their descendents were to become and that the leap to modern symbolic consciousness was achieved via a cultural stimulus (in a creature already exapted to make that leap) rather than by a biological innovation. From this perspective, it is hardly surprising that on morphological grounds Lieberman is able to reject—convincingly—the notion that Neanderthals possessed speech abilities, for only under highly unusual and transitory circumstances is there any association of Neanderthal fossils with symbolic objects.

One might therefore be surprised to find Lieberman concluding that language had its roots deep in the hominid past, possibly in conjunction with the demonstrably ancient adoption of upright locomotion. Whether facultative or obligate he does not make clear, for to do this is to blur the distinction—on which he insists—between language and precursor forms of communication (if that is indeed what language is all about). The ability to produce language and speech certainly builds upon a long and accretionary evolutionary history, among hominids and their predecessors, that involved increasingly complex forms of vocal communication; but language itself is, as far as we know, an entirely new and unanticipated form of expression, reflecting the operation of cognitive processes that are qualitatively different from any

we can observe among nonhuman organisms. Language is almost certainly a truly emergent quality, built upon what went before but entirely unpredicted by it. So while language is certainly the product of a long evolutionary past, it is equally certainly entirely novel, and, indeed, it may have been the rather recent invention of language itself that provided the cultural stimulus that declenched human symbolic thought.

Reply

The traditional model for the neural bases of speech and language posits linguistic functions localized in Broca's and Wernicke's cortical areas. The theory derives from studies of damage to the brain that result in aphasia—permament loss of various aspects of language. However, current studies show that aphasia does not occur in the absence of subcortical damage. Moreover, subcortical damage can result in aphasia without any cortical damage. The studies that I briefly reviewed, including hypoxic insult to the brain in climbers ascending Mount Everest and studies of Parkinson's disease, show that basal ganglia dysfunction yields linguistic deficits similar to those traditionally attributed to damage to Broca's area. Fecteau and Théoret claim that a "healthy Broca's area is necessary for proper speech," but the clinical record clearly shows that persons with major damage to Broca's area recover in the absence of subcortical damage. The reason seems to rest in the plasticity of cortical areas. As Fecteau and Théoret themselves note, behavior can be restored when a cortical area is destroyed because "the rest of the brain is capable of sustaining function following a brain lesion, and brain plasticity plays a major role in behavioral recovery." Studies such as Pascual-Leone et al. (2005) and Sanes et al. (1995) show that cortex is redundant and extremely malleable. In contrast, the subcortical brain structures that support cortical-striatal-cortical circuits regulating motor control, language, emotion, and other aspects of behavior (Cummings 1993) are few and do not appear to be capable of restructuring.

The fact that repetitive transcranial magnetic stimulation (TMS) of the left inferior frontal gyrus of the brain (the cortical region in which Broca's area is located) interferes with speech does not show that Broca's area is essential for speech. TMS also affects the subcortical structures that support the entire neural circuit to the left inferior gyrus. Studies of TMS of frontal cortical regions consistently show that it affects the basal ganglia structures that support the neuronal populations projecting to these cortical areas. For example, TMS of prefrontal cortex results in dopamine release in the caudate nucleus (Strafella et al. 2001), and TMS applied to the supplementary motor area increases dopamine release in the putamen (Strafella et al. 2005). I concur with the view that Broca's area is involved in aspects of behavior that transcend language. The mirror neurons identified in the monkey brain by Rizzolatti and his colleagues respond to acts such as tearing a piece of paper and the sound of the paper tearing. These responses must be learned associations and have little to do with communication.

I am in total agreement with García and Aboitiz—both cortical and subcortical regions of the brain are essential components of the neural bases of human cognitive and motor ability. The point that I hoped to convey was that the role of subcortical brain structures has been overlooked; new insights on the evolution of speech and cognition can be gained by exploring the physiology and evolution of the neural circuits linking cortical and subcortical regions of the brain. However, syntax is not localized in the left hemisphere as García and Aboitiz claim. Current brain-imaging studies do not support that view. For example, the fMRI study of Kotz et al. (2003) and other brain-imaging studies discussed in my paper and in Lieberman (2000, 2006*b*) show bilateral activity during tasks in which subjects have to comprehend distinctions in meaning conveyed by syntax.

MacLarnon appears not to have noticed the new approach presented in my paper to determining the speech capabilities of archaic hominids including Neanderthals. The data in McCarty et al. (n.d.) show that the necks of Neanderthals and all other fossils that predate 50,000 years before the present were too short to support a vocal tract in which the tongue had the proportions necessary to produce the full range of human speech. That assessment is based on "hard" skeletal evidence—the vertebrae of the neck. Neanderthals and Middle Pleistocene fossil hominids (including Skhul V, which has often been considered anatomically modern) would not have been able to eat if they had had a long SVTv. The laryngeal maneuvers necessary to swallow would have been blocked by the sternum bone.

Melrose provides further evidence that neural circuits rather than localized cortical regions are the brain bases for regulating complex behaviors. The syndromes that he mentions fit into the general framework of the premise that these circuits include the basal ganglia, other subcortical structures, and cortex. Our ongoing study of children with verbal apraxia (difficulties in sequencing speech motor commands) shows a set of motor and cognitive deficits similar to those of the KE family, which derive from an anomalous FOXP2 gene.

FOXP2 is a regulatory gene active in the embryonic development of the basal ganglia and other subcortical structures that play a key role in the neural circuits that regulate speech motor control and human cognition. The basal ganglia also participate in associative learning (e.g., Kimura, Aosaki, and Graybiel 1993). The thrust of Riede's comments on vocal learning and the avian foxp2 version of this gene is not clear. Riede correctly notes that foxp2 levels in the basal ganglia of birds appear to be involved in associative learning, a topic that I have discussed elsewhere (Lieberman 2006*b*, 173–74, 226–27). However, the fact that avian foxp2 is implicated in learning is not an argument against the probable role of human FOXP2 in the evolution of human speech and language.

Riede seems to overlook two key findings of the FOXP2 studies: (1) the human version is two mutations removed from that of chimpanzees and three from the mouse foxp2 gene (Lai et al. 2003), and (2) the time frame for the evolution of the human version is the last 100,000 years (Enard et al. 2002). The obvious behavioral differences between humans and other species in these domains may, in part, derive from these genetic distinctions. In short, as I have noted (Lieberman 2006*b*, 218–17), although FOXP2 is surely not the only gene involved in the formation of the human brain, it provides a means by which we may approach the evolution of the neural bases of human motor and cognitive ability.

To recapitulate my earlier critique (Lieberman 2006*b*) of Riede's claim that Diana monkeys produce the human vowel [a] by means of a vocal-tract shape similar to that used by humans, the monkey's tongues could not have produced the shapes that Riede modeled without surgical intervention because their tongues are positioned in their mouths and cannot produce extreme and abrupt area function discontinuities. The key is being able to produce order-of-magnitude changes in the cross-sectional area of the vocal tract at its midpoint, a feat that is impossible with a nonhuman supralaryngeal vocal tract. Moreover, if the Diana monkey were actually producing the sounds in question using the hypothetical human vocal-tract shape modeled by Riede, a third formant frequency would be present. Inspection of Riede's sound spectrograms shows a first and second formant but no trace of a third. We can thus dismiss the claim that the Diana monkey vocalizations were produced by the human-like vocal tract shape modeled by Riede. Laryngeal air sacs (which act as resonators tuned to specific frequencies) are the likely source for these sounds. In short, anatomical constraints and the principles of physical acoustics (see Chiba and Kajiyama 1941; Fant 1960; Stevens 1972) rule out the possibility of these monkeys' producing the full range of human speech.

Tattersall points out that new structures may arise independently of the functions that they now serve. Thus the restructuring of the human skull which shortens the mouth, reducing the length of SVTh, and the descent and reshaping of human tongue that yields a vocal tract in which SVTh/SVTv = 1, may initially have been chance events that had nothing to do with speech. However, the peculiar human supralaryngeal vocal tract, which increases the risk of choking to death on food lodged in the larynx, would not have been retained if it had not conferred some selective advantage. The apparent behavior that increases biological fitness is speech communication. But being able to use the fully human speech-producing anatomy that we can discern at 50,000 years before present entails having the neural capability of rapidly producing and reiterating complex sequential articulatory maneuvers. That ability is lacking in other living primates. It probably reached its present state when the human form of the FOXP2 gene evolved, but the starting point may have occurred much earlier. The role of the basal ganglia in walking and running that is evident in modern humans suggests that the process had an early origin.

The question of when language appeared hinges on one's definition of language. I have argued that while many of the components of fully human linguistic ability have a long evolutionary past, fully human speech and other language capabilities are not apparent until 50,000 years ago. As I have stressed, the anatomy that allows us to produce the full range of human speech would not have been useful without a brain that can reiterate the complex articulatory gestures that underlie speech. The evidence that is briefly reviewed in my paper and in greater detail in Lieberman (2000, 2006*b*) shows that the neural bases of speech motor control, cognition, and language involve the same structures of the human brain. There are no apparent disjoint neural motor control, syntax, or cognition "modules." The neural bases of human motor control, cognition, and language are intertwined. Therefore, it is probable that fully human syntactic and cognitive abilities were also present 50,000 years ago.

I am partially in agreement with Tattersall regarding the late appearance of fully human language. The time frames for the evolution of the human form of FOXP2 and that of speech anatomy are consistent with his view that fully human language appeared after the appearance of hominids who resembled us in many respects. However, these archaic hominids did not have vocal tracts that could produce fully human speech. They also may have lacked fully modern human brains capable of freely reiterating speech motor commands, syntactic processes, and cognitive acts. I do not think that language provided the cultural stimulus that triggered human symbolic thought. Language and other symbolic behaviors appear to derive from the evolution of a complex interdependent neural substrate—one that was not present until 50,000 or so years ago.

—Philip Lieberman

References Cited

Aboitiz, F., and R. García. 1997. The evolutionary origin of the language areas in the human brain: A neuroanatomical perspective. *Brain Research Reviews* 25:381–96. [RRG, FA]

Aboitiz, F., R. R. Garcia, C. Bosman, and E. Brunetti. 2006. Cortical memory mechanisms and language origins. *Brain and Language* 98:40–56. [RRG, FA]

Aldridge, J. W., K. C. Berridge, M. Herman, and L. Zimmer. 1993. Neuronal coding of serial order: Syntax of grooming in the neostratum. *Psychological Science* 4:391–93.

Alexander, G. E., and M. D. Crutcher. 1990. Functional architecture of basal ganglia circuits: Neural substitutes of parallel processing. *Trends in Neuroscience* 13:266–71.

Alexander, G. E., M. R. DeLong, and P. L. Strick. 1986. Parallel organization of segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience* 9:357–81.

Alexander, M. P., M. A. Naeser, and C. L. Palumbo. 1987.

Correlations of subcortical CT lesion sites and aphasia profiles. *Brain* 110:961–91.

Alipour, M., Y. Chen, and U. Jürgens. 1997. Anterograde projections of the cortical tongue area of the tree shrew (*Tupaia belangeri*). *Journal of Brain Research* 38:405–23. [TR]

———. 2002. Anterograde projections of the motorcortical tongue area of the saddle-back tamarin (*Saguinus fuscicollis*). *Brain, Behavior, and Evolution* 60:101–16. [TR]

Arbib, M. A. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavior and Brain Sciences* 28:105–67. [SF, HT]

Baddeley, A. D. 1986. *Working memory.* Oxford: Clarendon Press.

Baer, T., J. C. Gore, L. C. Gracco, and P. W. Nie. 1991. Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *Journal of the Acoustical Society of America* 90:799–828.

Barnes, T., Y. Kubota, D. Hu, D. Z. Jin, and A. M. Graybiel. 2005. Activity of striatal neurons reflect dynamic encoding and recoding of procedural memories. *Nature* 437:1158–61. [TR]

Baum, S. R., S. E. Blumstein, M. A. Naeser, and C. L. Palumbo. 1990. Temporal dimensions of consonant and vowel production: An acoustic and CT scan analysis of aphasic speech. *Brain and Language* 39:33–56.

Beckers, G. J. L., B. S. Nelson, and R. A. Suthers. 2004. Vocal-tract filtering by lingual articulation in a parrot. *Current Biology* 14:1592–97. [TR]

Beckman, M. E., T-P. Jung, S-H. Lee, K. De Jong, A. K. Krishnamurthy, S. C. Ahalt, K. B. Cohen, and M. J. Collins. 1995. Variability in the production of quantal vowels revisited. *Journal of the Acoustical Society of America* 97: 471–89.

Belin, P., P. Van Eeckhout, M. Zilbovicius, P. Remy, C. Francois, S. Guillaume, F. Chain, G. Rancurel, and Y. Samson. 1996. Recovery from nonfluent aphasia after melodic intonation therapy: PET study. *Neurology* 47:1504–11. [SF, HT]

Benson, D. F., and N. Geschwind. 1985. Aphasia and related disorders: A clinical approach. In *Principles of behavioral neurology*, ed. M. M. Mesulam, 193–228. Philadelphia: F. A. Davis.

Berridge, K. C., and I. Q. Whitshaw. 1992. Cortex, striatum, and cerebellum: Control of serial order in a grooming sequence. *Experimental Brain Research* 90:275–90.

Blumstein, S. E. 1994. The neurobiology of the sound structure of language. In *The cognitive neurosciences,* ed. M. S. Gazzaniga. Cambridge: MIT Press.

———. 1995. The neurobiology of language. In *Speech, language, and communication,* 339–70. San Diego: Academic Press.

Blumstein, S. E., W. E. Cooper, H. Goodglass, S. Statlender, and J. Gottlieb. 1980. Production deficits in aphasia: A voice-onset time analysis. *Brain and Language* 9:153–70.

Boe, L-J., J-L. Heim, K. Honda, and S. Maeda. 2002. The potential Neanderthal vowel space was as large as that of modern humans. *Journal of Phonetics* 30:465–84.

Boe, L-J., S. Maeda, and J-L. Heim. 1999. Neanderthal man was not morphologically handicapped for speech. *Evolution of Communication* 3:49–77.

Bosma, J. F. 1975. Anatomic and physiologic development of the speech apparatus. In *Human communication and its disorders,* ed. D. B. Towers, 469–81. New York: Raven.

Boule, M. 1911–13. L'homme fossile de la Chapelle-aux-Saints. *Annales de Paléontologie* 6:109; 7:21, 85; 8:1.

Bramble, D. M., and D. E. Lieberman. 2004. Endurance running and the evolution of *Homo. Nature* 432:345–52.

Broca, P. 1861. Rémarques sur le siege de la faculté de la parole articulée, suiviés d'une observation d'aphemie (perte de parole). *Bulletin de la Société d'Anatomie* (Paris) 36: 330–57.

Brokelman, W. Y., and D. Schilling. 1984. Inheritance of stereotyped gibbon calls. *Nature* 312:634–36. [TR]

Buhr, R. D. 1980. The emergence of vowels in an infant. *Journal of Speech and Hearing Research* 23:75–94.

Burke, R. E., S. O. Franklin, and C. E. Inturrisi. 1994. Acute persistent suppression of preproenkephaline mRNA expression in the striatum following developmental hypoxic-ischemic injury. *Journal of Neurochemistry* 62:1878–86.

Carre, R., B. Lindblom, and P. MacNeillage. 1995. Acoustic factors in the evolution of the human vocal tract. *Comptes Rendus de l'Académie des Sciences, Paris,* 320, series 2b: 471–76.

Chiba, T., and J. Kajiyama. 1941. *The vowel: Its nature and structure.* Tokyo: Tokyo-Kaisekan.

Chie, U., Y. Inoue, M. Kimura, E. Kirino, S. Nagaoka, M. Abe, T. Nagata, and H. Arai. 2004. Irreversible subcortical dementia following high altitude illness. *High Altitude Medicine and Biology* 5:77–81.

Chomsky, N. 1957. *Syntactic structure.* The Hague: Mouton.

———. 1995. *The minimalist program.* Cambridge: MIT Press.

Cools, R., R. A. Barker, G. J. Sahakian, and T. W. Robbins. 2001. Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain* 124:2503–12.

Cummings, J. L. 1993. Frontal-subcortical circuits and human behavior. *Archives of Neurology* 50:873–80.

Cummings, J. L., and D. F. Benson. 1984. Subcortical dementia: Review of an emerging concept. *Archives of Neurology* 41:874–79.

Cunnington, R., R. Iansek, J. L. Bradshaw, and J. G. Phillips. 1995. Movement-related potentials in Parkinson's disease: Presence and predictability of temporal and spatial cues. *Brain* 118:935–50.

Darwin, C. 1964 (1859). *On the origin of species.* Facsimile ed. Cambridge: Harvard University Press.

Deacon, T. 1989. The neural circuitry underlying primate calls and human language. *Human Evolution* 4:367–401. [TR]

———. 1997. *The symbolic species: The co-evolution of language and the human brain.* London: Penguin. [RM]

DeLong, M. R. 1993. Overview of basal ganglia function. In *Role of the cerebellum and basal ganglia in voluntary movement,* ed. N. Mano, I. Hamada, and M. R. DeLong. Amsterdam: Elsevier.

D'Esposito, M., and M. P. Alexander. 1995. Subcortical aphasia: Distinct profiles following left putaminal hemorrhage. *Neurology* 45:38–41.

Dobzhansky, T. 1973. Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* 35: 125–29.

Doupe, A. J., and P. K. Kuhl. 1999. Bird song and human speech: Common themes and mechanisms. *Annual Review of Neuroscience* 22:567–631. [TR]

Doya, K., and T. Sejnowski. 1995. A novel reinforcement model of birdsong vocalization learning. In *Advances in neural information processing systems*, vol. 7, ed. G. Tesauro, D. S. Touretzky, and T. K. Leen, 101–8. Cambridge: MIT Press. [TR]

Dronkers, N. F., J. K. Shapiro, B. Redfern, and R. T. Knight. 1992. The role of Broca's area in Broca's aphasia. *Journal of Clinical and Experimental Neuropsychology* 14: session 8, Lang Aphasia.

Enard, W., M. Prezeworski, S. E. Fisher, C. S. Lai, V. Wiebe, T. Katano, A. P. Monaco, and S. Paabo. 2002. Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 41:869–72.

Esposito, A., G. Demeurisse, B. Alberti, and F. Fabbro. 1999. Complete mutism after midbrain periaqueductal gray lesion. *Neuroreport* 10:681–85. [TR]

Fadiga, L., L. Craighero, G. Buccino, and G. Rizzolatti. 2002. Speech listening specifically modulates the excitability of tongue muscles: TMS study. *European Journal of Neuroscience* 15:399–402. [SF, HT]

Fant, G. 1960. *Acoustic theory of speech production.* The Hague: Mouton.

Feinberg, M. J., and O. Ekberg. 1990. Deglutition after near-fatal choking episode: Radiologic evaluation. *Radiology* 176: 637–40.

Fisher, S. E., F. Vargha-Khadem, K. E. Watkins, A. P. Monaco, and M. E. Pembrey. 1998. Localization of a gene implicated in a severe speech and language disorder. *Nature Genetics* 18:168–70.

Fitch, W. T. 1997. Vocal tract length and formant frequency dispersion correlate with body size in macaque monkeys. *Journal of the Acoustical Society of America* 102:1213–22.

———. 2000*a.* Skull dimensions in relation to body size in nonhuman mammals: The causal bases for acoustic allometry. *Zoology* 103:40–58.

———. 2000*b.* The evolution of speech: A comparative view. *Trends in Cognitive Science* 4:258–67.

———. 2000*c.* The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals. *Phonetica* 57:205–18. [TR]

Fitch, W. T., and J. Giedd. 1999. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America* 106:1511–22.

Fitch, W. T., and D. Reby. 2001. The descended larynx is not uniquely human. *Proceedings of Royal Society London* B 268: 1669–75.

Flowers, K. A., and C. Robertson. 1985. The effects of Parkinson's disease on the ability to maintain a mental set. *Journal of Neurology, Neurosurgery, and Psychiatry* 48: 517–29.

Friederici, A. D. 2004. The neural basis of syntactic processes. In *The cognitive neurosciences* 3, ed. M. S. Gazzaniga. Cambridge: MIT Press. [RRG, FA]

Fujimura, O., and Y. Kakita. 1979. Remarks on quantitative description of lingual articulation. In *Frontiers of speech communication research* ed. B. Lindbolm and S. Ohman, 17–24. London: Academic Press.

Fuster, J. M. 1995. *Memory in the cerebral cortex.* Cambridge: MIT Press. [RRG, FA]

Gallagher, H. L., and C. D. Frith. 2003. Functional imaging of "theory of mind." *Trends in Cognitive Sciences* 7:77–83. [RM]

Gautier, J. P. 1971. Étude morphologique et fonctionnelle des annexes extra-laryngées des Cercopithecinae: Liaison avec les cris d'espacement. *Biologia Gabonica* 7:229–67. [TR]

Gentner, T. Q., K. M. Fenn. D. Margoliash, and H. C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440:1204–7. [TR]

George, S. L. 1978. A longitudinal and cross-sectional analysis of the growth of the postnatal cranial base angle. *American Journal of Physical Anthropology* 49:171–78.

Gil-da-Costa, R., A. Martin, M. A. Lopes, M. Munoz, J. B. Fritz, and A. R. Braun. 2006. Species-specific calls activate homologs of Broca's and Wernicke's areas in the macaque. *Nature Neuroscience* 8:1064–70. [RRG, FA]

Goldstein, K. 1948. *Language and language disturbances.* New York: Grune and Stratton.

Goodall, J. 1986. *The chimpanzees of Gombe: Patterns of behavior.* Cambridge: Harvard University Press.

Graybiel, A. M. 1995. Building action repertoires: Memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology* 5:733–41.

———. 1997. The basal ganglia and cognitive pattern generators. *Schizophrenia Bulletin* 23:459–69.

———. 1998. The basal ganglia and chunking of action repertoires. *Neurobiology, Memory, Learning* 70:119–36.

Graybiel, A. M., T. Aosaki, A. W. Flaherty, and M. Kimura. 1994. The basal ganglia and adaptive motor control. *Science* 265:1826–31.

Greenberg, B. D., D. L. Murphy, and S. A. Rasmussen. 2000. Neuroanatomically based approaches to obsessive-compulsive disorder: Neurosurgery and transcranial magnetic stimulation. *Psychiatric Clinics of North America* 23:671–85.

Grossman, M. G., S. Carvell, S. Gollomp, M. B. Stern, M. Reivich, D. Morrison, A. Alavi, and H. L. Hurtig. 1993. Cognitive and physiological substrates of impaired sen-

tences processing in Parkinson's disease. *Journal of Cognitive Neuroscience* 5:480–98.

Grossman, M. G., S. Carvell, S. Gollomp, M. B. Stern, G. Vernon, and H. I. Hurtig. 1991. Sentence comprehension and praxis deficits in Parkinson's disease. *Neurology* 41: 1620–28.

Haesler, S., K. Wada, A. Nshdejan, E. E. Morrisey, T. Lints, E. J. Jarvis, and C. Scharff. 2004. FoxP2 expression in avian vocal learners and non-learners. *Journal of Neuroscience* 24: 3164–75. [TR]

Halliday, Michael Alexander Kirkwood. 2004. *An introduction to functional grammar.* 3d ed. revised by Christian M. I. M. Matthiessen. London: Edward Arnold. [RM]

Halliday, M. A. K., and C. M. I. M. Matthiessen. 1999. *Construing experience through meaning: A language-based approach to cognition.* London: Continuum. [RM]

Hauser, M. D., N. Chomsky, and W. T. Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298:1569–79.

Heim, J-L. 1989. La nouvelle réconstitution du crâne néanderthalien de la Chapelle-aux-Saints: Méthode et résultants. *Bulletin et Mémoires de la Société d'Anthropologie de Paris,* n.s., 1:95–118.

Heiser, M., M. Iacoboni, F. Maeda, J. Marcus, and J. C. Mazziotta. 2003. The essential role of Broca's area in imitation. *European Journal of Neuroscience* 17:1123–28. [SF, HT]

Hellwag, C. 1781. De formatione loquelae. Ph.D. diss., University of Tubingen.

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

Hiiemae, K. M., S. M. Hayenga, and A. Reese. 1995. Patterns of tongue and jaw movement in a cinefluorographic study of feeding in the macaque. *Archives of Oral Biology* 40: 229–46. [TR]

Hiiemae, K. M., J. B. Palmer, S. W. Medicis, J. Hegener, B. S. Jackson, and D. E. Lieberman. 2002. Hyoid and tongue movements in speaking and eating. *Archives of Oral Biology* 47:11–27.

Hillenbrand, J. L., A. Getty, M. J. Clark, and K. Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97:3099–3111.

Hochstadt, J. 2004. The nature and causes of sentence comprehension deficits in Parkinson's disease: Insights from eye tracking during sentence picture matching. Ph.D. diss., Brown University.

Hoehn, M. M., and M. D. Yahr. 1967. Parkinsonism: Onset, progression, and mortality. *Neurology* 17:427–42.

Honda, K., and M. K. Tiede. 1998. An MRI study on the relationship between oral cavity shape and larynx position. *Proceedings of the 5th International Conference on Spoken Language Processing* 2:437–40.

Howells, W. W. 1976. Neanderthal man: Facts and figures. In *Proceedings of the Ninth International Congress of Anthropological and Ethnological Sciences, Chicago 1973.* The Hague: Mouton.

———. 1989. *Skull shapes and the map: Craniometric analyses in the dispersion of modern* Homo. Papers of the Peabody Museum of Archaeology and Ethnology, Harvard University, 79.

Illes, J., E. J. Metter, W. R. Hanson, and S. Iritani. 1988. Language production in Parkinson's disease: Acoustic and linguistic considerations. *Brain and Language* 33:146–60.

Inoue, T., H. Kato, T. Araki, and K. Kogure. 1992. Emphasised selective vulnerability after repeated nonlethal cerebral ischemic insults in rats. *Stroke* 23:739–45.

Irwin, O. C. 1948. Infant speech: Development of vowel sounds. *Journal of Speech and Hearing Disorders* 13:31–34.

Ishida, R., J. B. Palmer, and K. M. Hiiemae. 2002. Hyoid motion during swallowing: Factors affecting forward and upward displacement. *Dysphagia* 17:262–72.

Jackendoff, R. 1994. *Patterns in the mind: Language and human nature.* New York: Basic Books.

Jackendorff, R., and S. Pinker. 2005. The nature of the language faculty and its implications for evolution of language (reply to Fitch, Hauser, and Chomsky). *Cognition* 97: 211–25. [TR]

Jellinger, K. 1990. New developments in the pathology of Parkinson's disease. In *Advances in neurology,* vol. 53, *Parkinson's disease: Anatomy, pathology, and therapy*, ed. M. B. Streifler, A. D. Korezyn, J. Melamed, M. B. H. Youdim, 1–15. New York: Raven Press.

Jeong, J. H., J. C. Knon, J. H. Chin, S. J. Yoon, and D. L. Na. 2002. Globus pallidus lesions associated with high mountain climbing. *Journal of Korean Medical Science* 17:861–63.

Jürgens, U. 1999*a.* Language evolution. In *Encyclopedia of neuroscience*, 2d edition, ed. G. Adelman and B. H. Smith, 1026–27. Amsterdam: Elsevier. [TR]

———. 1999*b.* Primate communication: Signaling, vocalization. In *Encyclopedia of neuroscience*, 2d edition, ed. G. Adelman and B. H. Smith, 1694–97. Amsterdam: Elsevier. [TR]

———. 2002. Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews* 26:235–58. [TR]

Jürgens, U., and R. Pratt. 1979. Role of periaqueductal gray in vocal expression of emotion. *Brain Research* 167:367–78. [TR]

Just, M. A., P. A. Carpenter, T. A. Keller, W. F. M. Eddy, and K. R. Thulborn. 1996. Brain activation modulated by sentence comprehension. *Science* 274:114–16.

Kao, M. H., A. J. Doupe, and M. S. Brainard. 2005. Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature* 433:638–43. [TR]

Kimura, M., T. Aosaki, and A. Graybiel. 1993. Role of basal ganglia in the acquisition and initiation of learned movement. In *Role of the cerebellum and basal ganglia in voluntary movements,* ed. N. Nano, I. Hamada, and M. R. DeLong, 83–87. Amsterdam: Elsevier.

Klein, D., R. J. Zatorre, B. Milner, E. Meyer, and A. C. Evans. 1994. Left putaminal activation when speaking a second language: Evidence from PET. *NeuroReport* 5:2295–97.

Klein, R. G. 1999. *The human career.* 2d ed. Chicago: University of Chicago Press.

Kobara-Mates, M., J. A. Logeman, C. Larson, and P. Kahrilas. 1995. Physiology of oropharyngeal swallow in the cat: A videofluoroscopic and electromyographic study. *American Journal of Physiology* 268:G232–41. [TR]

Kotz, S. A., M. Meyer, K. Alter, M. Besson, D. Y. Von Cramon, and A. Frederici. 2003 On the lateralization of emotional prosody: An fMRI investigation. *Brain and Language* 96: 366–76.

Krings, M., A. Stone, R. W. Schmitz, H. Krainitzki, M. Stoneking, and S. Paabo. 1997. Neanderthal DNA sequences and the origin of modern humans. *Cell* 90:19–30.

Kuhl, P. K., K. A. Williams, F. Lacerda, K. N. Stevens, and B. Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255:606–8.

Ladefoged, P. J., and D. E. Broadbent. 1957. Information conveyed by vowels. *Journal of the Acosutical Society of America* 29:98–104.

Ladefoged, P., J. De Clerk, M. Lindau, and G. Papcun. 1972. An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics* 22:48–76.

Lai, S. J., S. E. Fisher, J. A. Hurst, F. Vargha-Khadem, and A. P. Monaco. 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519–23.

Lai, C. S., D. Gerrelli, A. P. Monaco, S. E. Fisher, and A. J. Copp. 2003. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain* 126:2455–62.

Laitman, J. T., and E. S. Crelin. 1976. Postnatal development of the basicranium and vocal tract region in man. In *Symposium on development of the basicranium,* ed. J. Bosma, 206–19. Washington, D.C.: U.S. Government Printing Office.

Laitman, J. T., and R. C. Heimbuch. 1982. The basicranium of Plio-Pleistocene hominids as an indicator of their upper respiratory systems. *American Journal of Physical Anthropology* 59:323–44.

Laitman, J. T., R. C. Heimbuch, and E. S. Crelin. 1978. Developmental changes in a basicranial line and its relationship to the upper respiratory system in living primates. *American Journal of Anatomy* 152:467–82.

———. 1979. The basicranium of fossil hominids as an indicator of their upper respiratory systems. *American Journal of Physical Anthropology* 51:15–34.

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

Laplane, D., M. Baulac, and D. Widlocher. 1984. Pure psychic akinesia with bilateral lesions of basal ganglia. *Journal of Neurology, Neurosurgery, and Psychiatry* 47:377–85.

Laplane, D., M. Levasseur, B. Pillon, R. Dubois, M. Baulac, S. Tran Dinh, G. Sette, F. Danze, and J. C. Baron. 1989. Obsessive-compulsive and other behavioral changes with bilateral basal ganglia lesions. *Brain* 112:699–725.

Lashley, K. S. 1951. The problem of serial order in behavior. In *Cerebral mechanisms in behavior,* ed. L. A. Jefress, 112–46. New York: Wiley.

Lehericy, S., M. Ducros, P-E. van de Moortele, C. François, L. Thivard, C. Poopon, N. Swindale, K. Ugurbil, and D.-S. Kim. 2004. Diffusion tensor tracking shows distinct corticostriatal circuits in humans. *Annals of Neurology* 55: 522–29.

Lichtheim, L. 1885. On aphasia. *Brain* 7:433–84.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74:431–61.

Lieberman, D. E. 1995. Testing hypotheses about recent human evolution from skulls. *Current Anthropology* 36: 159–98.

Lieberman, D. E., and R. C. McCarthy. 1999. The ontogeny of cranial base angulation in humans and chimpanzees and its implications for reconstructing pharyngeal dimensions. *Journal of Human Evolution* 36:487–517.

Lieberman, D. E., R. C. McCarthy, K. M. Hiiemae, and J. B. Palmer. 2001. Ontogeny of postnatal hyoid and laryngeal descent: Implications for deglutition and vocalization. *Archives of Oral Biology* 46:117–28.

Lieberman, D. E., C. F. Ross, and M. J. Ravosa. 2000. The primate cranial base: Ontogeny, function, and integration. *Yearbook of Physical Anthropology* 43:117–69.

Lieberman, P. 1968. Primate vocalizations and human linguistic ability. *Journal of the Acoustical Society of America* 44:1157–64.

———. 1984. *The biology and evolution of language.* Cambridge: Harvard University Press.

———. 2000. *Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought.* Cambridge: Harvard University Press.

———. 2002. On the nature and evolution of the neural bases of human language. *Yearbook of Physical Anthropology* 45: 36–62.

———. 2006*a*.*Toward an evolutionary biology of language.* Cambridge: Harvard University Press.

———. 2006*b*. Limits on tongue deformation: Diana monkey vocalizations and the impossible vocal tract shapes proposed by Riede et al. (2005). *Journal of Human Evolution* 50:219–21.

———. 2006*c*. Current views on Neanderthal speech capabilities: A reply to Boe et al. (2002). *Journal of Phonetics* 34.

Lieberman, P., and E. S. Crelin. 1971. On the speech of Neanderthal man. *Linguistic Inquiry* 2:203–22.

Lieberman, P., E. S. Crelin, and D. H. Klatt. 1972. Phonetic ability and related anatomy of the newborn, adult human, Neanderthal man, and the chimpanzee. *American Anthropologist* 74:287–307.

Lieberman, P., E. T. Kako, J. Friedman, G. Tajchman, L. S.

Feldman, and E. B. Jiminez. 1992. Speech production, syntax comprehension, and cognitive deficits in Parkinson's disease. *Brain and Language* 43:169–89.

Lieberman, P., B. G. Kanki, A. Protopapas, E. Reed, and J. W. Youngs. 1994. Cognitive defects at altitude. *Nature* 372: 325.

Lieberman, P., D. H. Klatt, and W. H. Wilson. 1969. Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science* 164:1185–87.

Lieberman, P., A. Morey, J. Hochstadt, M. Larson, and S. Mather. 2005. Mount Everest: A space-analog for speech monitoring of cognitive deficits and stress. *Aviation, Space, and Environmental Medicine* 76:198–207.

Liegeois, F., T. Baldeweg, A. Connelly, D. G. Gadian, M. Mishkin, and F. Vargha-Khadem. 2003. Language fMRI abnormalities associated with FOXP2 gene mutation. *Nature Neuroscience* 6:1230–37.

Lisker, L., and A. S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20:384–442.

Lubker, J., and T. Gay. 1982. Anticipatory labial coarticulation: Experimental, biological, and linguistic variables. *Journal of the Acoustical Society of America* 71:437–38.

McCammon, R. 1952. *Human growth and development.* Springfield: Thomas.

McCarthy, R. C., D. S. Strait, F. Yates, and P. Lieberman. n.d. The origin of human speech. MS.

McCowan, T. D., and A. Keith. 1939. *The fossil remains from the Levalloiso-Mousterian.* Vol. 2. *The Stone Age of Mount Carmel.* New York: Clarendon Press.

Mahajan, P. V., and B. A. Bharucha. 1994. Evaluation of short neck: Percentiles and linear correlations with height and sitting height. *Indian Pediatrics* 31:1193–1203.

Maresch, M. M. 1948. Growth of the heart related to bodily growth during childhood and adolescence. *Pediatrics* 2: 382–402.

Marin, O., W. J. Smeets, and A. Gonzalez. 1998. Evolution of the basal ganglia in tetrapods: A new perspective based on recent studies in amphibians. *Trends in Neuroscience* 21: 487–94.

Marsden, C. D., and J. A. Obeso. 1994. The functions of the basal ganglia and the paradox of stereotaxic surgery in Parkinson's disease. *Brain* 117:877–97.

Martin, P. I., M. A. Naeser, K. W. Doron, A. Bogdan, E. H. Baker, J. Kurland, P. Renshaw, and D. Yurgelun-Todd. 2005. Overt naming in aphasia studied with a functional MRI hemodynamic delay design. *Neuroimage* 28:194–204. [SF, HT]

Melrose, R. 2005. How a neurological account of language can be reconciled with a linguist's account of language: The case of systemic-functional linguistics. *Journal of Neurolinguistics* 18:401–21. [RM]

Middleton, F. A., and P. L. Strick. 1994. Anatomical evidence for cerebellar and basal ganglia involvement in higher cognition. *Science* 266:458–61.

Mirenowicz, J. and W. Schultz. 1996. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* 379:449–51.

Monchi, O., P. Petrides, V. Petre, K. Worsley, and A. Dagher. 2001. Wisconsin Card Sorting Revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience* 21:7733–41.

Mouse Genome Sequencing Consortium. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–62.

Murray, G. M., L. D. Lin, E. M. Moustafa, and B. J. Sessle. 1991. Effects of reversible inactivation by cooling of the primate face motor cortex on the performance of a trained tongue-protrusion task and a trained biting task. *Journal of Neurophysiology* 65:511–30. [TR]

Naeser, M. A., M. P. Alexander, N. Helms-Estabrooks, H. L. Levine, S. A. Laughlin, and N. Geschwind. 1982. Aphasia with predominantly subcortical lesion sites: Description of three capsular/putaminal aphasia syndromes. *Archives of Neurology* 39:2–14.

Naeser, M. A., P. I. Martin, M. Nicholas, E. H. Baker, H. Seekins, M. Kobayashi, H. Theoret, F. Fregni, J. Maria-Tormos, J. Kurland, K. W. Doron, and A. Pascual-Leone. 2005. Improved picture naming in chronic aphasia after TMS to part of right Broca's area: An open-protocol study. *Brain and Language* 93:95–105.

Natsopoulos, D., G. Grouios, S. Bostantzopoulou, G. Mentenopoulos, Z. Katsarou, and J. Logothetis. 1993. Algorithmic and heuristic strategies in comprehension of complement clauses by patients with Parkinson's disease. *Neuropsychologia* 31:951–64.

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

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

Nishimura, T., A. Mikami, J. Suzuki, and T. Matsuzawa. 2003. Descent of the larynx in chimpanzee infants. *Proceedings of the National Academy of Sciences, U.S.A.* 100:6930–33.

Olveczky, B. P., A. S. Andalman, and M. S. Fee. 2005. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *Public Library of Science Biology* 3:e153. [TR]

Ovchinnikov, I. V., A. Gotherstrom, G. P. Romanova, V. M. Kharitonov, K. Liden, and W. Goodwin. 2000. Molecular analysis of Neanderthal DNA from the northern Caucasus. *Nature* 404:490–93.

Owren, M. J. 2003. Vocal production and perception in nonhuman primates provide clues about early hominids and speech evolution. *ATR Symposium HIS Series* 1:1–19. [TR]

Owren, M. J., and J.-A. Bachorowski. 2001. The evolution of emotional expression: A "selfish-gene" account of smiling and laughter in early hominids and humans. In *Emotion: Current issues and future directions*, ed. T. J. Mayne and G. A. Bonanno, 152–91. New York: Guilford. [TR]

———. 2003. Reconsidering the evolution of nonlinguistic

communication: The case of laughter. *Journal of Nonverbal Behavior* 27:183–200. [TR]

Owren, M. J., J. A. Dieter, R. M. Seyfarth, and D. L. Cheney. 1992. Food calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species. *Behaviour* 120:218–31. [TR]

Palmer, J. B., N. J. Rudin, G. Lara, and A. W. Crompton. 1992. Coordination of mastication and swallowing. *Dysphagia* 7:187–200.

Parent, A. 1986. *Comparative neurobiology of the basal ganglia.* New York: Wiley.

Pascual-Leone, A., A. Amedi, F. Fregni, and I. B. Merabet. 2005. The plastic human brain cortex. *Annual Review of Neurosciences* 28:377–401. [SF, HT]

Pascual-Leone, A., J. R. Gates, and A. Dhuna. 1991. Induction of speech arrest and counting errors with rapid-rate transcranial magnetic stimulation. *Neurology* 41:697–702. [SF, HT]

Peterson, G. E., and H. L. Barney. 1952. Control methods used in a study of a vowels. *Journal of the Acoustical Society of America* 24:175–84.

Polit, A. and E. Bizzi, 1978. Processes controlling arm movements in monkeys. *Science* 201:1235–37.

Rafal, R. 2001. Virtual neurology. *Nature Neuroscience* 4: 862–64. [SF, HT]

Rendall, D., S. Kollias, C. Ney, and P. Loyd. n.d. Pitch (F0) and formant profiles of human and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *Journal of the Acoustical Society of America.* In press.

Riede, T., G. J. L. Beckers, W. Blevins, and R. A. Suthers. 2004. Inflation of the esophagus and vocal tract filtering in ring doves. *Journal of Experimental Biology* 207:4025–36. [TR]

Riede, T., E. Bronson, H. Hatzikirou, and K. Zuberbühler. 2005. Vocal production in a non-human primate: Morphological data and a model. *Journal of Human Evolution* 48:85–96.

———. 2006. Multiple discontinuities in nonhuman vocal tracts: A response to Lieberman (2006). *Journal of Human Evolution* 50:222–25. [TR]

Riede, T., R. A. Suthers, N. Fletcher, and W. Blevins. 2006. Songbirds tune their vocal tract to the fundamental frequency of their song. *Proceedings of the National Academy of Sciences, U.S.A.* 103:5543–48. [TR]

Riede, T., and K. Zuberbühler. 2003*a.* Pulse register phonation in Diana monkey alarm calls. *Journal of the Acoustical Society of America* 113:2919–26. [TR]

———. 2003*b.* The relationship between acoustic structure and semantic information in Diana monkey alarm vocalization. *Journal of the Acoustical Society of America* 114: 1132–42. [TR]

Rissman, J., J. C. Eliassen, and S. E. Blumstein. 2003. An event-related fMRI study of implicit semantic priming. *Journal of Cognitive Neuroscience* 151:1160–75.

Rizzolatti, G., and M. A. Arbib. 1998. Language with our grasp. *Trends in Neuroscience* 21:188–94. [SF, HT]

Rizzolatti, G., and L. Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27:169–92. [SF, HT]

Robertson, E. M., H. Théoret, and A. Pascual-Leone. 2003. Studies in cognition: The problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience* 15:948–60. [SF, HT]

Rosen, H. J., S. E. Petersen, M. R. Linenweber, A. Z. Snyder, D. A. White, L. Chapman, A. W. Dromerick, J. A. Fiez, and M. D. Corbetta. 2000. Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology* 55:1883–94. [SF, HT]

Russell, G. O. 1928. *The vowel.* Columbus: Ohio State University Press.

Sanes, J. N., J. P. Donoghue, V. Thangaraj, R. R. Edelman, and S. Warach. 1995. Shared neural substrates controlling hand movements in human motor cortex. *Science* 268: 1775–77.

Saygın, Ayse Pınar, Stephen M. Wilson, Nina F. Dronkers, and Elisabeth Bates. 2004. Action comprehension in aphasia: Linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia* 42:1788–1804. [RM]

Schneider, R. 1964. Der Larynx der Säugetiere. In *Handbuch der Zoologie* 5(7), ed. J. G. Helmcke, H. Lengerken, D. Starck, and H. Wermuth, 1–128. Berlin: Walter de Gruyter. [TR]

Scott, R. B., J. Harrison, C. Boulton, J. Wilson, R. Gregory, S. Parkin, P. G. Bain, C. Joint, J. Stein, and T. Z. Aziz. 2002. Global attentional-executive sequelae following surgical lesions to globus pallidus interna. *Brain* 125:562–74.

Sem-Jacobsen. C., and A. Torkildsen. 1960. Depth recording and electrical stimulation in the human brain. In *Electrical studies on the unanaesthetized brain*, ed. E. R. Ramey and D. S. O'Doherty, 275–87. New York: Hober. [TR]

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

Stevens, K. N. 1972. Quantal nature of speech. In *Human communication: A unified view,* ed. E. E. David Jr. and P. B. Denes, 51–66. New York: McGraw-Hill.

Stevens, K. N., and A. S. House. 1955. Development of a quantitative description of vowel articulation. *Journal of the Acoustical Society of America* 27:484–93.

Stewart, L., V. Walsh, U. Frith, and J. C. Rothwell. 2001. TMS produces two dissociable types of speech disruption. *Neuroimage* 13:472–78. [SF, HT]

Stone, M., and A. Lundberg. 1996. Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America* 99:3728–36.

Story, B. H., I. R. Titze, and E. A. Hoffman. 1996. Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America* 100:537–54.

Stowe, L. A., A. M-J. Paans, A. A. Wijers, and F. Awarts. 2004. Activation of "motor" and other non-language structures

during sentence comprehension. *Brain and Language* 89: 290–99.

Strafella, A. P., J. H. Ko, J. Grant, M. Fraraccio, and O. Monchi. 2005. Corticostriatal functional interactions in Parkinson's disease: An rTMS/[11C]raciopride PET study. *European Journal of Neuroscience* 22:2946–52.

Strafella, A. P., T. Paus, J. Barret, and A Dagher. 2001. Repetitive transcranial magnetic stimulation of the human prefrontal cortex induces dopamine release in the caudate nucleus. *Journal of Neuroscience* 21:RC157.

Stringer, C. B. 1998. Chronological and biogeographic perspectives on later human evolution. In *Neanderthals and modern humans in western Asia,* ed. T. Akazawa, K. Akoi, and O. Bar-Yosef, 29–38. New York: Plenum.

Strub, R. L. 1989. Frontal lobe syndrome in a patient with bilateral globus pallidus lesions. *Archives of Neurology* 46: 1024–27.

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

Takahashi, K., F. C. Liu, K. Hirokawa, and H. Takahashi. 2003. Expression of FOXP2, a gene involved in speech and language in the developing and adult striatum. *Journal of Neuroscience Research* 73:62–72.

Teramitsu, I., L. C. Kudo, S. E. London, D. H. Geschwind, and S. A. White. 2004. Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *Journal of Neuroscience* 24:3152–63. [TR]

Thatch, W. T. 1996. On the specific role of the cerebellum in motor learning and cognition: Clues from PET activation and lesion studies in man. *Behavioral and Brain Sciences* 19:411–31.

Thexton, A. J., and J. D. McGarrick. 1988. Tongue movement of the cat during lapping. *Archives of Oral Biology* 33: 331–39. [TR]

———. 1989. Tongue movement in the cat during the intake of solid food. *Archives of Oral Biology* 34:2391–48. [TR]

Tokimura, H., Y. Tokimura, A. Oliviero, T. Asakura, and J. C. Rothwell. 1996. Speech-induced changes in corticospinal excitability. *Annals of Neurology* 40:628–34. [SF, HT]

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

Vargha-Khadem, F., D. G. Gadian, A. Copp, and M. Mishkin. 2005. FOXP2 and the neuroanatomy of speech and language. *Nature Review of Neuroscience* 2:131–38. [RRG, FA]

Vargha-Khadem, F., K. Watkins, R. Passingham, and P. Fletcher. 1995. Cognitive and praxic deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences, U.S.A.* 95:2695–700.

Vargha-Khadem, F., K. E. Watkins, C. J. Price, J. Ashbruner, K. J. Alcock, A. Connelly, R. S. Frackowiak, K. J. Friston, M. E. Pembrey, M. Mishkin, D. G. Gadian, and R. E. Passingham. 1998. Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences, U.S.A.* 95:2695–2700.

Watkins, K. E., N. F. Dronkers, and F. Vargha-Khadem. 2002. Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. *Brain* 125: 452–64. [RM]

Watkins, K., and T. Paus. 2004. Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience* 16:978–87. [SF, HT]

Watkins, K. E., A. P. Strafella, and T. Paus. 2003. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–94. [SF, HT]

Watkins, K. E.,. F. Vargha-Khadem, J. Ashburner, R. E. Passingham, A. Connelly, K. J. Friston, R. S. J. Frackoiwiak, M. Mishkin, and D. G. Gadian. 2002. MRI analysis of an inherited speech and language disorder: Structural brain abnormalities. *Brain* 125:465–78.

Weisengruber, G. E., G. Forstenpointner, G. Peters, A. Kubber-Heiss, and W. T. Fitch. 2002. Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis silvestris* f. *catus*). *Journal of Anatomy* 201:195–201.

Wernicke, C. 1967 (1874). The aphasic symptom complex: A psychological study on a neurological basis. In *Proceedings of the Boston Colloquium for the Philosophy of Science*, vol. 4, ed. R. S. Cohen and M. W. Wartofsky. Dordrecht: Reidel.

Wood, J. N., S. G. Romero, K. M. Knutson, and J. Grafman. 2005. Representation of attitudinal knowledge: Role of prefrontal cortex, amygdala, and parahippocampal gyrus. *Neuropsychologia* 43:249–59. [RM]

Zuberbühler, K. 2000*a.* Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition* 76:195–207. [TR]

———. 2000*b.* Referential labeling in Diana monkeys. *Animal Behavior* 59:917–27. [TR]

———. 2000*c.* Interspecies semantic communication in two forest primates. *Proceedings in Biological Science* 267: 713–18. [TR]

———. 2000*d.* Causal knowledge of predators' behaviour in wild Diana monkeys. *Animal Behavior* 59:209–20. [TR]