Phonetic Ability and Related Anatomy of the Newborn and Adult Human, Neanderthal Man, and the Chimpanzee

PHILIP LIEBERMAN
*University of Connecticut
and Haskins Laboratories*

EDMUND S. CRELIN
Yale University School of Medicine

DENNIS H. KLATT
*Massachusetts Institute of Technology
and Research Laboratory of Electronics*

The sounds of human speech make human language a rapid medium of communication through a process of speech "encoding." The presence of sounds like the vowels [a], [i], and [u] makes this process possible. The supralaryngeal vocal tracts of newborn Homo sapiens and chimpanzee are similar and resemble the reconstructed vocal tract of the fossil La Chapelle-aux-Saints Neanderthal man. Vocal tract area functions that were directed toward making best possible approximations to the human vowels [a], [i], and [u], as well as certain consonantal configurations, were modeled by means of a computer program. The lack of these vowels in the phonetic repertories of these creatures, who lack a supralaryngeal pharyngeal region like that of adult Homo sapiens, may be concomitant with the absence of speech encoding and a consequently linguistic ability inferior to modern man.

HUMAN LANGUAGE is one of the defining characteristics of modern man. Although the evolution of human language has been the subject of hundreds of books and essays, not much is presently known. In recent years the primary focus has been directed toward the nature of the mental ability that may underly the syntactic and semantic aspects of human language.[1] This follows from a rather common opinion concerning language, i.e., that its phonetic aspect is trivial and indeed finally irrelevant to the serious study of human language and its evolution. Simpson (1966:473), for example, reviewing attempts to trace the evolution of language, notes that,

Audible signals capable of expressing language do not require any particular phonetic apparatus, but only the ability to produce sound, any sound at all. Almost all mammals and a great number of other animals can do that. Moreover, a number of animals, not only birds but also some mammals, can produce sounds recognizably similar to those of human language, and yet their jaws and palates are radically nonhuman.

Simpson essentially sets forth two premises. First, that any arbitrary set of sounds would serve as a phonetic base for human language. Second, that many animals also can produce the sounds that, in fact, occur in human language. If Simpson's premises were true, there would be little point in attempting to trace the evolution of human linguistic ability by studying either the comparative phonetic abilities of modern man and other living animals, or in attempting to reconstruct the phonetic abilities of extinct fossil hominids from their skeletal remains. Neither premise, however, is true. The results of research on the perception of human language depends on the existence of the particular sounds of human speech. No other sounds will do. The results of recent research on the anatomic basis of human speech have likewise demonstrated that no living animal, other than modern man, has the vocal mechanism that is necessary to produce the sounds of human speech.

We have discussed some of the anatomical factors that prevent living non-human pri-

mates and newborn humans from producing the range of sounds that characterize human speech (Lieberman 1968, 1969; Lieberman et al. 1969, 1972).[2] We have also been able to reconstruct the vocal apparatus of "classic" Neanderthal man (Lieberman and Crelin 1971). Our present paper has two objectives. We shall compare the anatomy and speech producing ability of the vocal mechanism of adult modern man with that of the adult chimpanzee, newborn modern man, and the reconstructed vocal mechanism of adult "classic" Neanderthal man. We will then discuss the speech perceiving and general linguistic abilities of the chimpanzee and Neanderthal man in the light of their sound making abilities. We shall, in this regard, consider some recent theoretical and experimental studies that relate the production and the perception of speech.

ACOUSTIC THEORY OF SPEECH PRODUCTION

The acoustic theory of speech production (Chiba and Kajiyama 1958; Fant 1960) relates the vocal mechanism to the acoustic signal. Human speech essentially involves the generation of sound by the mechanism of vocal cord vibration and/or air turbulence, and the acoustic shaping of these sound sources by the resonances of the supralaryngeal vocal tract. The shape of the human supralaryngeal vocal tract continually changes during the production of speech. These changes in the supralaryngeal vocal tract change its resonant properties. A useful mechanical analog to the aspect of speech production that is of concern to this discussion is a pipe organ. The musical function of each pipe is determined by its length and shape. (The pipes have different lengths and may be open at one end or closed at both ends.) The pipes are all excited by the same source. The resonant modes of each pipe determine the note's acoustic character. In human speech the phonetic properties that differentiate vowels like [i] and [a] from each other are determined by the resonant modes of the supralaryngeal vocal tract. The

frequencies at which resonances occur are called "formant" frequencies.

The acoustic theory of speech production which we have briefly outlined thus relates an acoustic signal to a supralaryngeal vocal tract configuration and a source. It is therefore possible to determine some of the constraints of an animal's phonetic range if the range of supralaryngeal vocal tract variation is known. The phonetic repertoire of an animal can obviously be expanded if different sources are used with similar supralaryngeal vocal tract configurations. We can, however, isolate the constraints that the range of supralaryngeal vocal tract - variation will impose on the phonetic repertoire.

VOCAL TRACT ANATOMY

The anatomic specializations of modern man that are necessary for human speech are evident when we compare the supralaryngeal vocal tract of adult man with creatures who lack human speech. We will start with a brief account of the skeletal similarities between Neanderthal man and newborn modern man[3] and adult chimpanzee that make it possible to reconstruct the supralaryngeal vocal tract of Neanderthal man.

In Figures 1-4, lateral views of the skulls of newborn man, adult chimpanzee, the La Chapelle-aux-Saints Neanderthal man, and adult modern man are presented. The skulls have all been drawn to appear nearly equal in size. Skull features of the newborn, the chimpanzee, and Neanderthal man that are similar to each other, but different from that of adult modern man, are as follows: (a) a generally flattened-out base, (b) lack of mastoid processes (very small in Neanderthal), (c) lack of a chin (occasionally present in the newborn), (d) the body of the mandible is much longer than the ramus (about sixty to one hundred percent longer), (e) the posterior border of the mandibular ramus is markedly slanted away from the vertical plane, (f) a more horizontal inclination of the mandibular foramen leading to the mandibular canal, (g) the pterygoid process of the sphenoid bone is relatively short and its later-
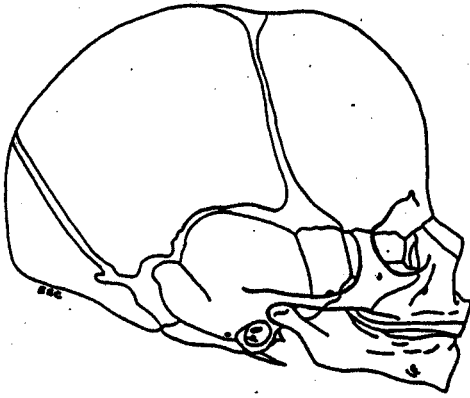
Figure 1. Skull of a human newborn.

al lamina is more inclined away from the vertical plane, (h) the styloid process is more inclined away from the vertical plane, (i) the dental arch of the maxilla is U-shaped instead of V-shaped, (j) the basilar part of the occipital bone between the foramen magnum and the sphenoid bone is only slightly inclined away from the horizontal toward the vertical plane, (k) the roof of the nasopharynx is a relatively shallow elongated arch, (l) the vomer bone is relatively short in its vertical height and its posterior border is inclined away from the vertical plane, (m) the vomer bone is relatively far removed from the junction of the sphenoid bone and the basilar part of the occipital bone, (n) the occipital condyles are relatively small and elongated.
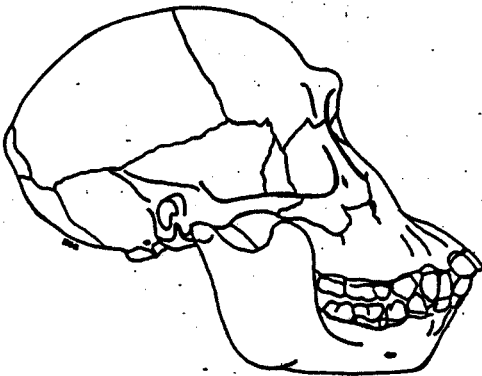
The chimpanzee differs from the newborn and adult modern man and Neanderthal man insofar as its mandible has a "simian shelf," i.e., internal buttressing of the anterior portion of mandible. The simian shelf inhibits the formation of a large air cavity behind the teeth. In adult man, a large cavity behind the teeth can be formed by pulling the tongue back in the mouth.

The significance of these skeletal features can be seen when the supralaryngeal vocal tracts that correspond to these skulls are ex-
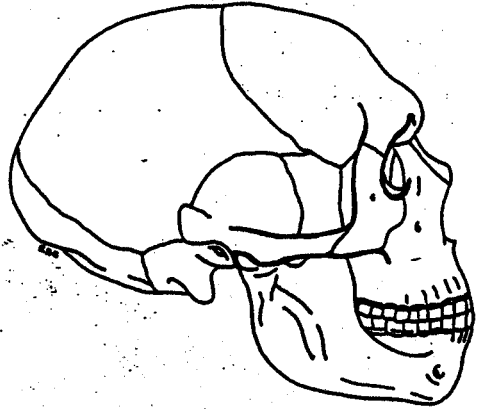

Figure 3. Skull of the La Chapelle-aux-Saints fossil Neanderthal man.

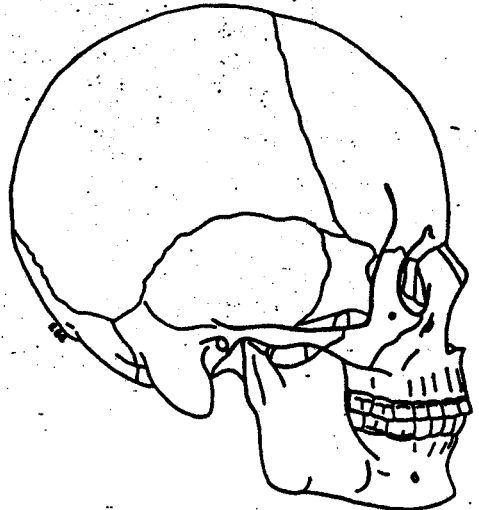Figure 2. Skull of an adult chimpanzee.

Figure 4. Skull of an adult man.

amined. The chimpanzee specimen used in this study was the head and neck of a young adult male sectioned in the midsagittal plane (Figure 5). The human newborn and adult specimens were those described by Lieberman and Crelin (1971) which included a number of heads divided in the midsagittal plane. Silicone-rubber casts were made of the air passages, including the nasal cavity, of the chimpanzee, newborn and adult man. This was done by filling each side of the split air passages separately in the sectioned heads and necks to insure perfect filling of the cavities. The casts from each side of a head and neck were then fused together to make a complete cast of the air passages. The cast of the Neanderthal air passages was made from the reconstructed nasal, oral, pharyngeal,

and laryngeal cavities of the La Chapelle-aux-Saints fossil (Lieberman and Crelin 1971). All four casts are shown in the photograph in Figure 6.

Even though the cast of the newborn air passages is much smaller than those of chimpanzee and adult modern man and Neanderthal man, it is apparent that the casts of newborn and chimpanzee are quite similar. When outlines of the air passages from all four are made nearly equal in size (Figure 7), one can more readily recognize what the basic differences and similarities are. (1) Newborn human, the chimpanzee, and Neanderthal man all have their tongue at rest completely within the oral cavity, whereas in adult man the posterior third of the tongue is in a vertical position forming the anterior

Figure 5. Left half of the head and neck of a young adult male chimpanzee sectioned in the midsaggital plane.

wall of the supralaryngeal pharyngeal cavity. The foramen cecum of the tongue is thus located far more anteriorly, in the oral cavity, in chimpanzee and newborn than it is in adult man. (2) In the newborn chimpanzee and Neanderthal the soft palate and epiglottis can be approximated, whereas they are widely separated in adult man and cannot approximate. (3) There is practically no supralaryngeal portion of the pharynx present in the direct airway out from the larynx when the soft palate shuts off the nasal cavity in the chimpanzee, Neanderthal, and newborn man. In adult man half of the supralaryngeal vocal tract is formed by the pharyngeal cavity. This difference between the chimpanzee, Neanderthal, and newborn—and adult man, is a consequence of the opening of the larynx into the pharynx, which is immediately behind the oral cavity in the chimpanzee, Neanderthal, and newborn. In adult man, this opening occurs farther down in the pharynx. Note that the supralaryngeal pharynx in adult man serves both as a pathway for the ingestion of food and liquids and as an airway to the larynx. In chimpanzee, Neanderthal, and newborn man the section of the pharynx that is behind the oral cavity is reserved for deglutition. The high epiglottis can, moreover, close the oral cavity to retain solids and liquids and allow unhampered respiration through the nose. (4) The level of the vocal folds (cords) at rest in the chimpanzee is at the upper border of the fourth cervical vertebra, whereas in adult man it is between the fifth and sixth in a relatively longer neck. The position of the
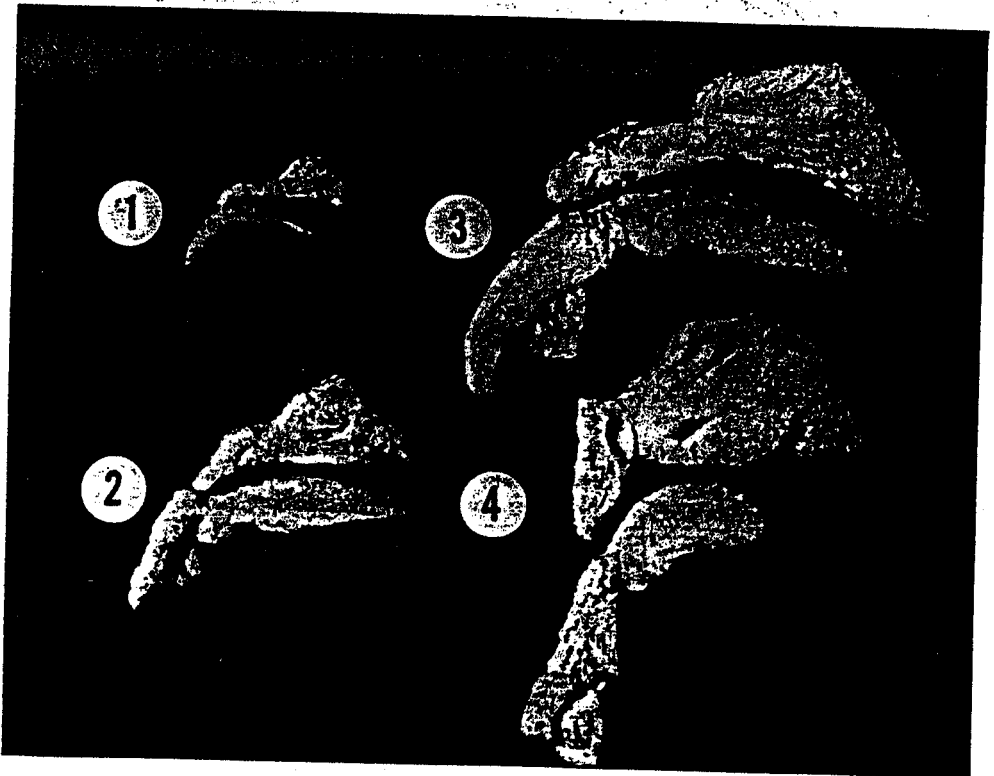

Figure 6. Casts of the nasal, oral, pharyngeal, and laryngeal cavities of (1) newborn human, (2) adult chimpanzee, (3) Neanderthal reconstruction, and (4) adult human.
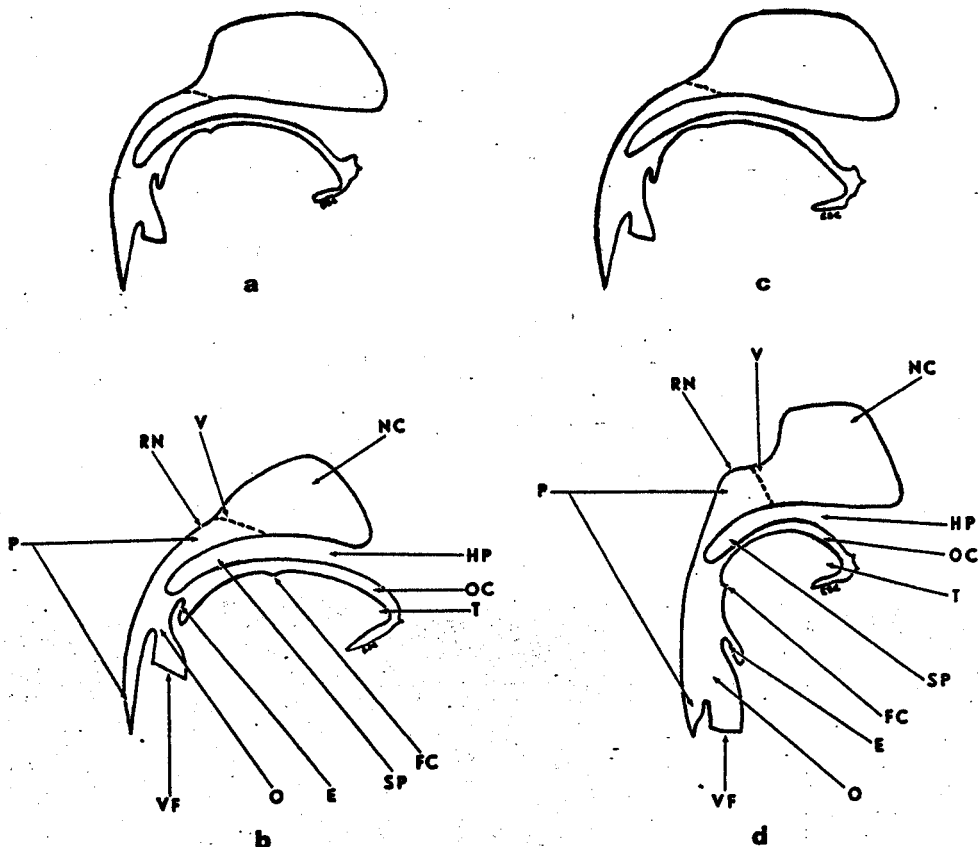
Figure 7. Diagrams of the air passages of (a) newborn human, (b) adult chimpanzee, (c) Neanderthal man, and (d) adult human. The anatomical details that are keyed on the chimpanzee and adult man are as follows: P-Pharynx, RN-Roof of Nasopharynx, V-Vomer Bone, NC-Nasal Cavity, HP-Hard Palate, OC-Oral Cavity, T-Tongue, FC-Foramen Cecum, SP-Soft Palate, E-Epiglottis, O-Opening of Larynx into Pharynx, VF-Level of Vocal Folds.

hyoid bone is high in the chimpanzee, Neanderthal and newborn. This is concomitant with the high position of the larynx.

SUPRALARYNGEAL VOCAL TRACT CONSTRAINTS ON PHONETIC REPERTOIRES

We have noted that human speech production involves a source of sound and a supralaryngeal vocal tract that acts as an acoustic "filter" or modulator. Man uses his articulators (the tongue, lips, mandible, velum, pharyngeal constrictors, etc.) to modify dynamically in time the resonant structure that the supralaryngeal vocal tract imposes on the acoustic sound pressure radiated at the speaker's lips and nares.

The phonetic inventory of a human language is therefore limited at the articulatory level by (1) the number of acoustically distinct sound sources that man is capable of controlling during speech communication, and (2) the number of distinct resonant patterns available through positioning of the articulators and dynamic manipulation of the articulators. In most human languages, a phonetic analysis reveals a phonemic inventory on the order of twenty to forty distinct sound types (Troubetzkoy 1939;

Jakobson et al. 1952). Most of the segment proliferations are achieved through the varied use of the articulators. For example, in English there are at least ten vowels that differ primarily in the articulatory configuration of the supralaryngeal vocal tract, and concomitantly in the resonant, i.e., the formant structure of the acoustic output (Peterson and Barney 1952).

There is a direct relationship between the articulatory configuration of the supralaryngeal vocal tract and the formant structure (Fant 1960). The relationship depends exclusively on the area function, i.e., the cross-sectional area of the vocal tract as a function of the distance from the vocal cords to the lips. The availability of digital computers makes it possible to determine the range of formant frequency patterns that a supralaryngeal vocal tract can produce. If the supralaryngeal vocal tract area function is systematically manipulated in accord with the muscular and anatomical constraints of the head and neck, a computer can be programmed to compute the formant frequencies that correspond to the total range of supralaryngeal vocal tract variation (Henke 1966). In other words, a computer-implemented model of a supralaryngeal vocal tract can be used to determine the possible contribution of the vocal tract to the phonetic repertoire. We can begin conveniently to determine whether a non-human supralaryngeal vocal tract can produce the range of sounds that occurs in human language by exploring its vowel producing ability. Consonantal vocal tract configurations can also be modeled. It is, however, reasonable to start with vowels since the production of consonants may also involve rapid coordinated articulatory maneuvers, and we can only speculate on the presence of this ability in fossil hominids.

THE VOWEL TRIANGLE

Articulatory and acoustic analyses have shown that the three vowels [i], [a], and [u] are the limiting articulations of a vowel triangle that is language universal (Troubetz-koy 1939). The body of the tongue is high and fronted to form a constricted oral cavity in [i] whereas it is low to form a large oral cavity in [a] and [u]. Figure 8 shows a mid-saggital outline of the vocal tract for the vowels [i], [a], and [u], as well as the cross-sectional areas of the vocal tract (Fant 1960) and the frequency domain transfer functions for these vowels (Gold and Rabiner 1968). The tongue body forms a large pharyngeal cavity in [i] and [u] and a constricted pharyngeal cavity in [a]. If the tongue body moves to form any greater constrictions, turbulent friction noise is generated at the vocal tract constriction and the articulation produces a consonant, not a vowel. Other English vowels are produced by means of supralaryngeal vocal tract configurations within the articulatory triangle[4] defined by [i], [a], and [u].

The universality and special nature of [i], [a], and [u] can be argued from theoretical grounds as well. Employing the simplified and idealized area functions shown in Figure 9, Stevens (1969) has shown that these articulatory configurations (1) are acoustically stable for small changes in articulation and therefore require less precision in articulatory control than similar adjacent articulations, and (2) contain a prominent acoustic feature, i.e., two formants that are in close proximity to form a distinct energy concentration.

The vowels [a], [i], and [u] have another unique property. They are the only vowels in which an acoustic pattern can be related to a unique vocal tract area function (Lindblom and Sundberg 1969; Stevens 1969). Other vowels like [e], [I], [U], etc. can be produced by means of several alternate area functions (Stevens and House 1955). A human listener, when he hears a syllable that contains a token of [a], [i], or [u] can calculate the size of the supralaryngeal vocal tract that was used to produce the syllable. The listener, in other words, can tell whether a speaker with a large or small vocal tract is speaking. This is not possible for other vowels since a speaker with a small vocal tract can, for example, by increasing
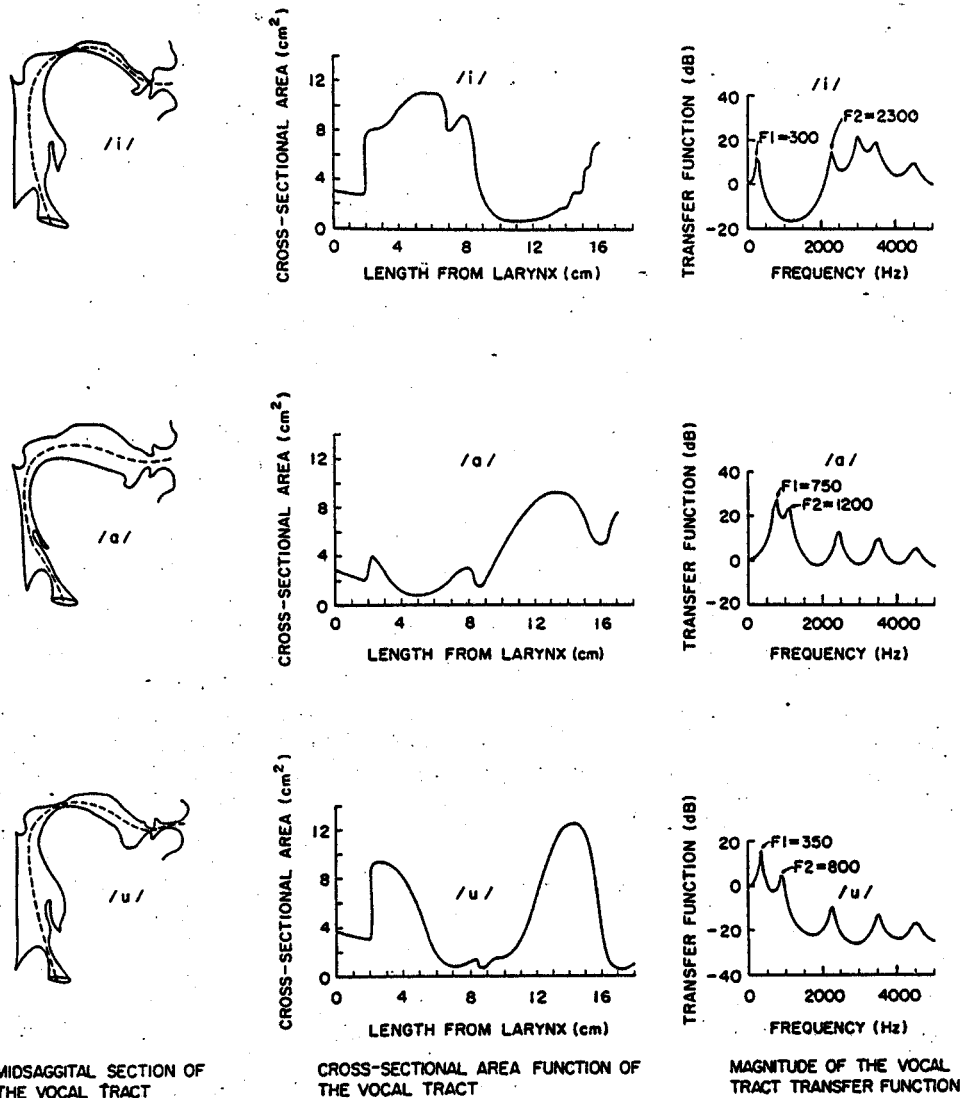
Figure 8. Illustrations of approximate (a) midsaggital sections, (b) cross-sectional area functions, and (c) acoustic transfer functions of the vocal tract for the vowels [i], [a], and [u].

the degree of lip rounding, produce a token of [U] that would be consistent with a larger vocal tract with less lip rounding. These uncertainties do not exist for [a], [i], and [u] since the required discontinuities in the supralaryngeal vocal tract area functions (Figure 8) produce acoustic patterns that are beyond the range of compensatory maneu-

vers. The degree of lip rounding for the [u] in Figure 8 is, for example, so extreme that it is impossible to constrict the lip opening any more and still produce a vowel.[5] The vowels [a], [i], and [u] are therefore different in kind from the remaining "central" vowels. These "vocal-tract size calibrating" properties of [a], [i], and [u] have a crucial
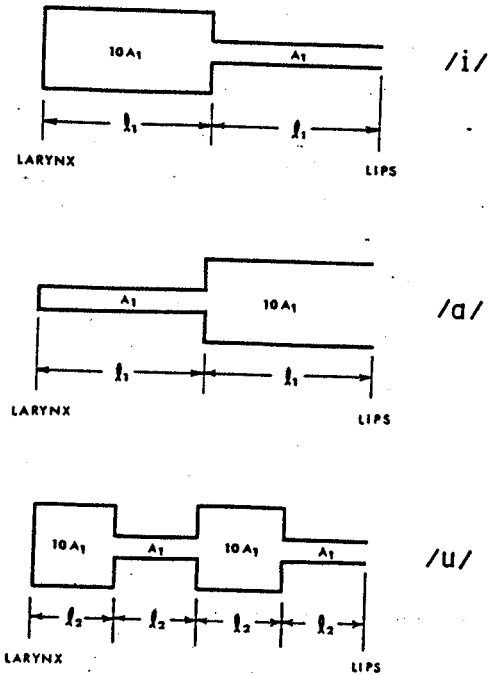
Figure 9. Stylized supralaryngeal vocal tract area functions that characterize the human vowels [a], [i], and [u].

role in the perception of speech and we will have more to say on this matter.

We can conclude from these considerations that the vowel space reserved for human language is delimited by the vowels [a], [i], and [u]. A study of the theoretical limitations on vowels produced by another related species can therefore proceed by determining the largest vowel triangle that its articulatory system is capable of generating.

THE VOWEL TRIANGLE IN CHIMPANZEE AND NEWBORN MAN AND NEANDERTHAL MAN

Some general observations are in order before detailed consideration of the vowel producing capabilities of the chimpanzee, human newborn and Neanderthal man. The idealized area functions of Stevens (Figure 9) require a relatively large ratio of the areas of the large and small section. In addition, they require rather abrupt boundaries between sections. These configurations can be approximated in adult man at the junction of the pharyngeal and oral cavities where the styloglossus muscle can be effective in pulling the body of the tongue upwards and backwards in the direction of the nasopharynx (Sobotta-Figge 1965; Perkell 1969; Lieberman 1970). The cross-sectional area of the oral and pharyngeal cavities can be independently manipulated in adult man (refer to Figure 8) while a midpoint constriction is maintained. The supralaryngeal vocal tract of adult man thus can, in effect, function as a "two" tube system. The lack of a supralaryngeal pharyngeal region prevents the chimpanzee, human newborn and Neanderthal from employing these mechanisms. They can only attempt to distort the tongue body in the oral cavity to obtain changes in cross-sectional areas. The intrinsic musculature of the tongue severely limits the range of deformations that the tongue body can be expected to employ. The chimpanzee, human newborn, and Neanderthal man, in effect, have "single tube" resonant systems.

The chimpanzee and human newborn heads are both smaller than adult man. This imposes a further difficulty since it makes it difficult to form the large cavities that are found in the vowels of man. Therefore, comparable cavity area ratios would require the use of smaller constrictions than adult man, but this would violate the requirement of non-turbulent flow in the constricted part of the vocal tract for vowels.

THE CHIMPANZEE VOWEL TRIANGLE

The vowel [a] could be articulated by a chimpanzee if he were to open his mandible sufficiently to obtain a flared area function. Taking into account the constraints mentioned above, an area function for a chimpanzee [a] has been estimated and plotted in Figure 10. Formant frequencies corresponding to the area function have been computed by means of an algorithm described by Henke (1966) and are tabu-

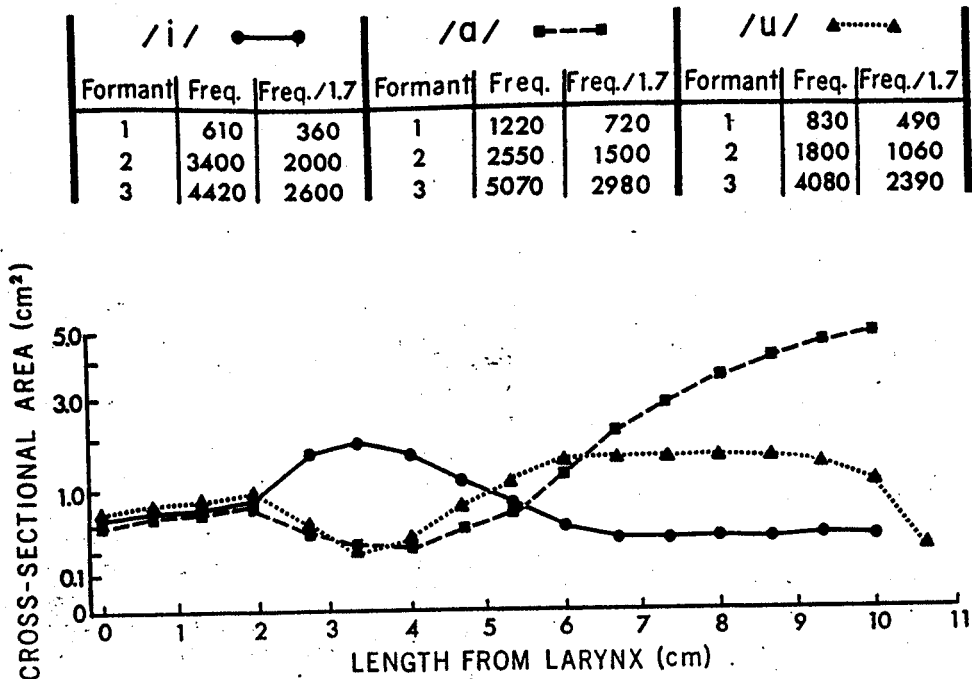| /i/ ●——●			/a/ ■–-–■			/u/ ▲·······▲		
Formant	Freq.	Freq./1.7	Formant	Freq.	Freq./1.7	Formant	Freq.	Freq./1.7
1	610	360	1	1220	720	1	830	490
2	3400	2000	2	2550	1500	2	1800	1060
3	4420	2600	3	5070	2980	3	4080	2390

Figure 10. Chimpanzee supralaryngeal vocal tract area functions modeled on computer. These functions were the "best" approximations that could be produced, given the anatomic limitations of the chimpanzee, to the human vowels [i], [a], and [u]. The formant frequencies calculated by the computer program for each vowel are tabulated and scaled to the average dimensions of the adult human vocal tract.

lated in the figure. The area of the vocal tract was specified at 0.5 cm intervals using this algorithm, which was implemented on a digital computer. When the two lowest formants are scaled down in frequency by a factor proportional to the ratio of a chimpanzee vocal tract length of 10 cm to the mean vocal tract length of 17 cm of adult man, then the chimpanzee formants can be compared directly with comparable data in adult man. This is done on a plot of first formant frequency versus second formant frequency in Figure 11 where the data point for this is denoted by the circled number "1." We see that the chimpanzee formant patterns for this vowel configuration do not fall within the range of [a] data for man, but rather lie inside the vowel triangle in the [ʌ] region. The normative data for modern man with which the chimpanzee vowel is compared is derived from a sample of

seventy-six adult men, adult women, and children (Peterson and Barney 1952). The labelled loops enclose the data points that accounted for ninety percent of the samples in each vowel category. The children in the Peterson and Barney study were sufficiently old that they all had vocal tracts that conformed to that typical of adult morphology (Lieberman et al. 1972; Crelin and Lieberman, unpublished data).

The vowel [i] could be best approximated by a chimpanzee by pulling the body of the tongue forward with the mandible lowered slightly. The cross-sectional area of the back cavity will not be large, but may aproach the area function estimated in Figure 10. This area function results in formant locations that are tabulated in Figure 10 and scaled and plotted in Figure 11 (data point "2"). The formants do not fall within the [i] region in adult man but
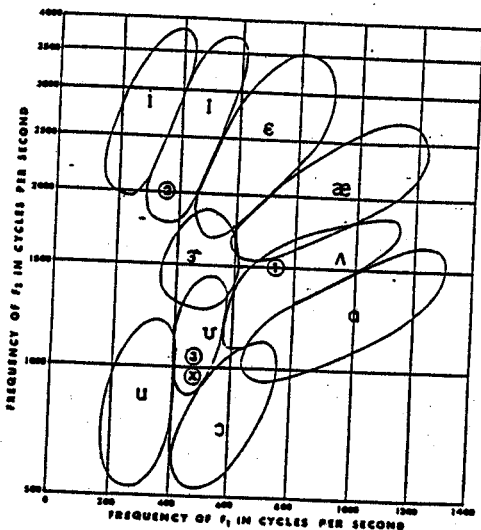
Figure 11. Plot of formant frequencies for chimpanzee vowels of Figure 9, data points (1), (2), and (3), scaled to correspond to the size of the adult human vocal tract. Data point (X) represents an additional point for human newborn. The closed loops enclose ninety percent of the data points derived from a sample of seventy-six adult men, women, and children producing American-English vowels (Peterson and Barney 1952). Note that the chimpanzee and newborn vocal tracts cannot produce the vowels [i], [u], and [a].

rather inside the vowel triangle in the [I] region.

The vowel [u] is virtually impossible for the chimpanzee to articulate. A large front cavity requires the mandible to be lowered because the simian shelf prevents the tongue body motion found in man. However, the required lip rounding is incompatible with a lowered mandible. An approximation to a chimpanzee [u] area function is estimated in Figure 10. Again, the formant locations of this area function are computed, scaled, and plotted in Figure 11 (data point "3"). They indicate that the comparable English vowel is [U] and not [u].

The discussion of the vowel triangle has not considered the effects of the chimpanzee pharynx which acts as a relatively short side-branch resonator. The pharyngeal sec-

tion may be essentially closed in a back vowel such as [a], but it probably plays an important role in [i]. The presence of a side-branch resonator has the effect of modifying formant locations and also the effect of introducing antiresonances into the vocal tract transfer function. We estimate that the lowest frequency antiresonance for [i] of a slightly flared 6 cm pharyngeal section is about 2000 Hz.[6]

NEWBORN HUMAN

The supralaryngeal vocal tract of the human newborn does not differ substantially from the chimpanzee's (Figures 1-3). The absence of a simian shelf in the mandible, however, allows the formation of a larger front cavity in the production of vowels that approximate the adult human [u]. In Figure 11 the formant locations of this area function, which resemble that of Figure 10 for the chimpanzee [u] approximation with a larger front cavity, are computed, scaled, and plotted as data point "X." The resulting vowel sound is comparable to the English vowel [U] not [u], but it is a closer acoustic approximation to [u]. The acoustic output of the newborn vocal tract does not otherwise differ substantially from the chimpanzee vocal tract. Perceptual and acoustic studies of the vocalizations of human newborn (Irwin 1957; Lieberman et al. 1972) show that all, and only, the vowels that can be produced are indeed produced.

NEANDERTHAL MAN

The vowel producing abilities of the reconstructed supralaryngeal vocal tract of the La Chapelle-aux-Saints Neanderthal fossil are presented in Figure 12. The formant frequencies of the Neanderthal supralaryngeal vocal tract configurations that best approximated the human vowels [a], [i], and [u] were computed, scaled, and plotted with respect to adult modern man (Lieberman and Crelin 1971). Note that the Neanderthal vowels which are each labeled "N" do not fall in the human ranges for [a], [i], or [u].
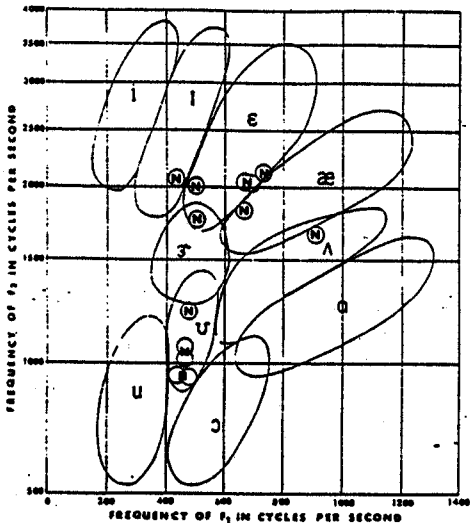
Figure 12. Plot of first and second formant frequencies for "extreme" vowels, data points (N), of reconstructed Neanderthal vocal tract (Lieberman and Crelin 1971).

The Neanderthal vocal tract was given the benefit of all possible doubts in the computer modeling. The maximum range of laryngeal cavity variation typical of modern man (Fant 1960) was, for example, used in a manner that would enhance the phonetic ability of the Neanderthal vocal tract. Articulatory maneuvers that would be somewhat acrobatic in modern man were also used to enhance Neanderthal phonetic ability. Our computer modeling was guided by the results of X-ray motion pictures of speech production, vocalization, swallowing, and respiration in adult man (Perkell 1969; Haskins Laboratories 1962) and in newborn (Truby et al. 1965). This knowledge, plus the known comparative anatomy of the living primates, allowed a fairly "conservative" simulation of the vowel producing ability of this fossil specimen who is typical of the range of "classic" Neanderthal man.[7] We perhaps allowed a greater vowel producing range for Neanderthal man since we consistently generated area functions that were more humanlike than apelike whenever we were in doubt. Despite these compensations,

the Neanderthal vocal tract cannot produce [a], [i], or [u]. The absence of these vowels from the vowel systems of chimpanzee, newborn human, and Neanderthal man in Figures 11 and 12 thus is an indirect way of showing that the vocal tracts of these creatures cannot form the abrupt area functions that are necessary for these vowels. Our modeling of the newborn vocal tract served as a control procedure since we were able to produce the vowels that newborn humans actually produce. We produced, however, a greater vowel range than has been observed in the acoustic analysis of chimpanzee vocalizations (Lieberman 1968). We will return to this point later in our discussion since it may reflect the absence of required neural mechanisms in the nonhuman primates.

SPEECH PRODUCTION AND SPEECH PERCEPTION

Supralaryngeal vocal tract area functions that approximated typical consonantal configurations for adult man (Fant 1960; Perkell 1969) were also modeled on the digital computer (Lieberman and Crelin 1971). Chimpanzee, newborn human, and Neanderthal man all appeared to have anatomical mechanisms that would allow the production of both labial and dental consonants like [b], [p], [t], [e], etc., if other muscular and neural factors were present.

It is obvious that some of these factors are not present in newborn human since neither labial nor dental consonants occur in the utterances of newborn infants (Irwin 1957). It is possible that the non-occurrence of these consonants is a consequence of a general inability to produce rapid articulatory maneuvers. The situation is more complex in chimpanzee where a discrepancy again exists between the constraints that the supralaryngeal vocal tract imposes on the phonetic repertoire and actual performance. Chimpanzees do not appear to produce dental consonants although they have the anatomical "machinery" that would permit them to do so. Observations of captive chimpanzees have not, for example, revealed patterns of vocal communication that utilize

contrasts between labial and dental consonants (Lieberman 1968). It is unlikely that the failure to observe dental consonants in chimpanzee vocalizations is due to a limited data sample since attempts to train chimpanzees to mimic human speech have not succeeded in teaching them to produce dental consonants. At least one chimpanzee has been taught to produce labial consonants like [p] and [m] (Hayes 1952), so the absence of dental consonants cannot be ascribed to a general inability to produce rapid articulatory maneuvers.

Our computer modeling of the chimpanzee vocal tract shows that these animals have the anatomic ability that would allow them to produce a number of vowels that in human speech are "phonemic" elements, i.e., sound contrasts that convey linguistically meaningful information. Chimpanzees, however, do not appear to make use of these vowel possibilities. Instead, they appear to make maximum use of the "neutral" uniform cross-section supralaryngeal vocal tract shape (Jakobson et al. 1952; Lieberman 1968) with source variations. Chimpanzees, for example, will make calls that are different insofar as the glottal excitation is weak, breathy, has a high fundamental frequency,[8] etc.

The absence of sounds that are anatomically possible may perhaps reflect perceptual limitations. In other words, chimpanzees may not use dental consonants in contrast with labial consonants because they cannot perceptually differentiate these sounds. Differences in vowel quality as between [I] and [e], for example, may also be irrelevant for chimpanzees. The absence of the vowels [a], [i], and [u] from the chimpanzee's phonetic abilities is consistent with this hypothesis, which has wider implications concerning the general phonetic and linguistic abilities of the living non-human primates and hominid fossils like Neanderthal man.

SPEECH AND LANGUAGE

Linguists have, as we noted earlier, tended to ignore the phonetic level of language and speech production. The prevailing assumption is that the interesting action is at the syntactic and semantic levels, and that just about any sequence of arbitrary sounds would do for the transfer of linguistic information. Some linguists might, for example, point out that even simple binary codes, such as Morse code, can be used to transmit linguistic information. Neanderthal man, in this view, therefore would need only one sound contrast to communicate. After all, modern man can communicate by this means; why not Neanderthal man? The answer to this question is quite simple. Human speech is a special mode of communication that allows modern man to communicate at least ten times faster than any other known method. Sounds other than speech cannot be made to convey language well.[9] That knowledge comes from fifty-five years of trying to make nonspeech sounds for use in reading machines for the blind, that is, devices that scan the print and convert it into meaningful sounds. In spite of the most diligent efforts in connection with the development of these machines, no nonspeech acoustic alphabet has yet been contrived that can be made to work more than one-tenth as well as speech (Liberman et al. 1967). Nor has any better degree of success attended efforts toward the use of visual displays in the development of "hearing" machines for the deaf (Koenig et al. 1946).

The problem is quite clear when one considers the rate at which information is transferred in human speech. Human listeners can perceive as many as twenty-five to thirty phonetic segments per second in normal speech. This segment rate far exceeds the resolving power of the human auditory system. It is, for example, impossible to even count simple pulses at rates of twenty pulses per second. The pulses simply merge into a continuous tone. Communication by means of Morse code would be possible, but it would be very slow. Human speech achieves its high information rate by means of an "encoding" process that is structured in terms of the anatomic and articulatory constraints of speech production. The presence

of vowels like [a], [i], and [u] appears to be one of the anatomic factors that makes this encoding process possible.

SPEECH ENCODING AND THE "MOTOR THEORY" OF SPEECH PERCEPTION

In human speech a high rate of information transfer is achieved by "encoding" phonetic segments into syllable-sized units. The phonetic representation of a syllable like [du] essentially states that two independent elements are being transmitted. The syllable [du] can be segmented at the phonetic level into two segments, [d] and [u], which can independently combine with other phonetic segments to form syllables like [di] or [gu]. Phonetic segments like [d], [g], [u], and [i] are also independent at the articulatory level insofar as these phonetic elements can each be specified in terms of an articulatory configuration. The phonetic element [u] thus involves a particular vocal tract configuration which approximates that in Figure 8. The phonetic element [d] likewise involves a particular vocal tract configuration in which the tongue blade momentarily occludes the oral cavity. It is possible to effect a segmentation of the syllable [du] at the articulatory level. If an X-ray motion picture of a speaker producing the syllable [du] were viewed it would, for example, be possible to see the articulatory gesture that produces the [d] in the syllable [du]. It is not, however, possible to segment the acoustic correlates of [d] from the speech signal.

In Figure 13 we have reproduced two simplified spectrographic patterns that will, when converted to sound, produce approximations to the syllables [di] and [du] (Lieberman 1970).[10] The dark bands on these patterns represent the first and second formant frequencies of the supralaryngeal vocal tract as functions of time. Note that the formants rapidly move through a range of frequencies at the left of each pattern. These rapid movements, which occur in about 50 msec, are called formant transi-

tions. The transition in the second formant, which is encircled, conveys the acoustic information that human listeners interpret as a token of a [d] *in the syllables [di] and [du]*. It is, however, impossible to isolate the acoustic pattern of [d] in these syllables. If tape recordings of these two syllables are "sliced" with the electronic equivalent of a pair of scissors (Lieberman 1963), it is impossible to find a segment that contains only [d]. There is no way to cut the tape so as to obtain a piece that will produce [d] without also producing the next vowel or some reduced approximation to it.

Note that the encircled transitions are different for the two syllables. If these encircled transitions are isolated, listeners report that they hear either an upgoing or a falling frequency modulation. In context, with the acoustic correlates of the entire syllable, these transitions cause listeners to hear an "identical" sounding [d] in both syllables. How does a human listener effect this perceptual response?

We have noted the formant frequency patterns of speech reflect the resonances of the supralaryngeal vocal tract. The formant patterns that define the syllable [di] in Figure 13 thus reflect the changing resonant pattern of the supralaryngeal vocal tract as the speaker moves his articulators from the occlusion of the tongue tip against the palate that is involved in the production of [d] to the vocal tract configuration of the [i]. A different acoustic pattern defines the [d] in the syllable [du]. The resonances of the vocal tract are similar as the speaker forms the initial occlusion of the [d] in both syllables; however, the resonances of the vocal tract are quite different for the final configurations of the vocal tract for [i] and [u]. The formant patterns that convey the [d] in both syllables are thus quite different since they involve transitions from the same starting point to different end points. Human listeners "hear" an identical initial [d] segment in both of these signals because they "decode" the acoustic pattern in terms of the articulatory gestures and the anatomical apparatus that is involved in the produc-
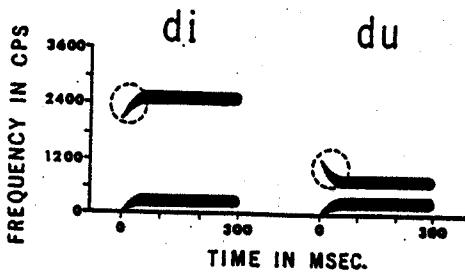
Figure 13. Simplified spectrographic patterns sufficient to produce the syllables [di] and [du]. The circles enclose the second formant frequency transitions (After Lieberman 1970).

tion of speech. The listener in this process, which has been termed the "motor theory of speech perception" (Lieberman et al. 1967), operates in terms of the acoustic pattern of the entire syllable. The acoustic cues for the individual "phonetic segments" are fused into a syllabic pattern. The high rate of information transfer of human speech is thus due to the transmission of acoustic information in syllable-sized units. The phonetic elements of each syllable are "encoded" into a single acoustic pattern which is then "decoded" by the listener to yield the phonetic representation.

In order for the process of "motor theory perception" to work the listener must be able to determine the absolute size of the speaker's vocal tract. Similar articulatory gestures will have different acoustic correlates in different sized vocal tracts. The frequency of the first formant of [a], for example, varies from 730 to 1030 Hz in the data of Peterson and Barney (1952) for adult men and children. The frequencies of the resonances that occur for various consonants likewise are a function of the size of the speakers' vocal tract. The resonant pattern that is the correlate of the consonant [g] for a speaker with a large vocal tract may overlap with the resonant pattern of the consonant [d] for a speaker with a small vocal tract (Rand 1971). The listener therefore must be able to deduce the size of the speaker's vocal tract before he can assign an

acoustic signal to the correct consonantal or vocalic class.

There are a number of ways in which a human listener can infer the size of a speaker's supralaryngeal vocal tract. He can, for example, note the fundamental frequency of phonation. Children, who have smaller vocal tracts, usually have higher fundamental frequencies than adult men or adult women. Adult men, however, have disproportionately lower fundamental frequencies than adult women (Peterson and Barney 1952), so fundamental frequency is not an infallible cue to vocal tract size. Perceptual experiments (Ladefoged and Broadbent 1957) have shown that human listeners can make use of the formant frequency range of a short passage of speech to arrive at an estimate of the size of a speaker's vocal tract. Recent experiments, however, show that human listeners do not have to defer their "motor theory" decoding of speech until they hear a two or three second interval of speech. Instead, they use the vocalic information encoded in a syllable to decode the syllable (Darwin 1971; Rand 1971). This may appear to be paradoxical, but it is not. The listener makes use of the formant frequencies and fundamental frequency of the syllable's vowel to assess the size of the vocal tract that produced the syllable. We have noted throughout this paper that the vowels [a], [i], and [u] have a unique acoustical property. The formant frequency pattern for these vowels can always be related to a unique vocal tract size and shape.[11] A listener, when he hears one of these vowels, can thus instantly determine the size of the speaker's vocal tract. The vowels [a], [i], and [u] (and the glides [y] and [w]) thereby serve as primary acoustic calibration signals in human speech.

The anatomical impossibility for the chimpanzee to produce these vowels is thus consistent with the absence of meaningful changes in vowel quality in the vocal communications of these animals. Chimpanzees probably cannot perceive these differences in vowel quality because they cannot "decode" specific vowels and consonants in

terms of the articulatory gestures that speakers use to produce these signals. A chimpanzee on hearing a particular formant frequency pattern would, for example, not be able to tell whether it was produced by a large chimpanzee who was using an [I]-like vocal tract configuration or a smaller chimpanzee who was using an [e]-like vocal tract configuration.[12] Chimpanzees simply may not have the neural mechanism that is used in modern man to decode speech signals in terms of the underlying articulatory maneuvers. The absence of a humanlike pharyngeal region in chimpanzee is thus quite reasonable. The only function that the human supralaryngeal vocal tract is better adapted to is speech production, in particular the production of vowels like [a], [i], and [u]. The adult human supralaryngeal vocal tract is otherwise less well adapted for the primary vegetative functions of swallowing and respiration (Negus 1949). It is quite easy for food to be caught in the adult human pharynx and block the entrance to the larynx with fatal consequences, whereas the high position of the laryngeal opening in chimpanzees and other non-human primates would allow them to breathe with food lodged in their pharynx. The efficiency of the respiratory apparatus is reduced considerably in adult human because the angulation of the airway (Figures 6 and 7), resulting from the low position of the larynx, appreciably lessens the volume of air which could pass through a straight tube of equal cross-section. The high position of the larynx in newborn human, chimpanzee, and Neanderthal man is efficient for respiration. As Kirchner (1970:12) notes, "the larynx of the newborn infant is, from the standpoint of position, a more efficient respiratory organ than its adult counterpart."

This suggests that the evolution of the human vocal tract which allows vowels like [a], [i], and [u] to be produced, and the widespread occurrence of these vowels in human languages reflect a parallel development of the neural and anatomic abilities that are necessary for language. This parallel development would be consistent with the evolution

of other human abilities. The ability to use tools depends, for example, on both upright posture and an opposable thumb, and neural ability. As Darwin (1859:194) noted, the theory of evolution through natural selection "can act only by taking advantage of slight successive variations; she can never take a leap, but must advance by the shortest and slowest steps." We can think of a process in which mutations that enhanced vocal communication were retained. The presence of enhanced mental ability would enhance the probability of the retention through natural selection of an anatomical mutation that enhanced the phonetic repertoire and the rate of communication. The presence of enhanced anatomical phonetic ability would, in turn, increase the probability of the retention of mutations that enhanced the neural abilities that are involved in speech encoding, decoding, syntax, etc. Positive feedback would, no doubt, result from this "circular" process. We would expect to find fossil forms like the La Chapelle-aux-Saints Neanderthal man who lacked a well developed vocal mechanism but who undoubtedly must have had a "language." The remains of Neanderthal culture all point to the presence of linguistic ability.[13]

Neanderthal man lacked the vocal tract that is necessary to produce the human "vocal tract size-calibrating" vowels [a], [i], and [u]. This suggests that the speech of Neanderthal man did not make use of syllabic encoding. While communication is obviously possible without syllabic encoding, studies of alternate methods of communication in modern man show, as we noted before, that the rate at which information can be transferred is about one-tenth that of normal human speech. The principle of encoding extends throughout the grammar of human languages. The process wherein a deep phrase marker with many elementary S's is collapsed into a derived surface structure may be viewed as an encoding process that is similar to the encoding that occurs between the phonetic level and speech (Liberman 1970). A transformational

grammar (Chomsky 1957, 1965) may be viewed as a mechanism that encodes strings of semantic units into a surface structure. The derived surface string can be readily transmitted by a speaker and perceived and stored in short time span memory by a listener. There is no other reason why adult humans do not speak in short sentences like, *I saw the boy. The boy is fat. The boy fell down.* instead of the "encoded" sentence, *I saw the fat boy who fell down.* The "encoded" sentence can be transmitted more rapidly, and it transmits the unitary reference of the single *boy* within the single breath-group (Lieberman 1967). It thus is likely that Neanderthal man's linguistic abilities were at best suited to communication at slow rates and at worst markedly inferior at the syntactic and semantic levels to modern man's linguistic ability. Neanderthal man's language is an intermediate stage in the evolution of language. It may well have employed gestural communication as well as vocal signals (Hewes 1971).

Human linguistic ability thus must be viewed as the result of a long evolutionary process that involved changes in anatomical structure through a process of mutation and natural selection which enhanced speech communication.[14] Modern man's linguistic ability is necessarily tied to his phonetic ability. Rapid information transfer through the medium of human speech must be viewed as a central property of human linguistic ability. It makes human language and human thought possible.[15]

NOTES

[1] Hewes (1971) has compiled a comprehensive annotated bibliography on the evolution of language. With the exception of studies like Hockett (1960) and Hockett and Altmann (1968), most of the emphasis has been placed on the cognitive aspects of language.

[2] These results are consistent with the fact that it has never been possible to train a non-human primate to talk. Kellogg (1968) reviews a number of recent attempts at raising chimpanzees as though they were children. It is interesting to note that similar attempts date back to at least the eighteenth century (La Mettrie 1747). The "speech" of "talking birds" is not similar to human speech at the acoustic or anatomic levels (Greenewalt 1967). A parrot's imitation of human speech is similar to a human's imitation of a siren. The signal is accepted as a mimicry. It has different acoustic properties than the siren's signal, and it is produced by a different apparatus.

[3] The similarity between human newborn and the adult Neanderthal fossil conforms to the view that modern man and Neanderthal man had a common ancestor. Darwin in *On the Origin of Species* (1859:449) clearly states the premise that we are following in making this inference; he states: "In two groups of animals, however much they may at present differ from each other in structure and in habits, if they pass through the same or similar embryonic stages, we may feel assured that they have both descended from the same or nearly similar parents, and are therefore in that degree closely related." The adult Neanderthal skull has certain specialized features, like a supraorbital torus, that are not present in newborn modern man nor in adult modern man. This indicates that Neanderthal man is probably not directly related to modern man. He is, as Boule (1911-13) recognized, probably an early offshoot from the mainstream of hominids that evolved into modern man. The skulls of present day newborn apes are quite similar to the human newborn (Schultz 1968). This would indicate an early common ancestral form for both present day apes and man. It does not show that modern man has evolved by retaining infantile characteristics. Adult modern man, in his own way, deviates as much from his newborn state (Crelin 1969; Lieberman and Crelin 1971) as adult living apes do from their newborn form. Physical anthropologists and anatomists have noted, over the years, that measurements of particular aspects of Neanderthal skulls fall within the range of variation that may be found in modern man (Patte 1955). This finding is not surprising since all adult modern men develop from the newborn morphology which has many similarities to that of adult "classic" Neanderthal man. The course of human maturation is not even and some individuals fail to develop "normally." In extreme pathologic conditions like Down's Syndrome the individual may, in fact, retain many aspects of the newborn morphology, especially those of the skull. Benda (1969) notes that Down's Syndrome may be characterized, in part, as a developmental prob-

lem. We have examined a number of subjects afflicted with Down's Syndrome who cannot produce "articulate" speech (Lieberman and Crelin, unpublished data). Some of these subjects may lack the mental ability that is necessary to control their vocal apparatus, but some of them appear to have vocal tracts that resemble the normal newborn vocal tract. They, in effect, have Neanderthaloid vocal tracts and they cannot produce human speech. The bases of their skulls and their mandibles generally resemble those of a Neanderthal. It is therefore not surprising that Virchow (1872) believed that the original Neanderthal skull, which was found in 1856, was either a pathologic specimen or the skull of an imbecile. It is also evident that different population groups of modern man have somewhat different skeletal features. In some population groups a particular skeletal feature will fall within the range characteristic of classic Neanderthal man. Laughlin (1963), for example, notes that the breadth of the ramus of the mandible in Eskimos and Aleuts can exceed the breadth of this feature in Neanderthal man. The length of the body of the mandible is also somewhat longer for Aleuts and Eskimos than is the case for other modern human skulls. The length of the body of the mandible can be about twenty percent greater than the ramus in an adult male Aleut skull. This value is, however, much smaller than is the case for either Neanderthal man or newborn human where the length of the body of the mandible is sixty to one hundred percent greater than the ramus (as measured on a lateral projection to the midline of the mandible). The total ensemble of skeletal features of the base of the skull for Aleuts and Eskimos is, moreover, consistent with the "angulation" of the vocal tract of adult modern man.

[4] It can be argued that [o] forms a fourth position on a vowel "quadrangle," but this modification will not affect our arguments in any essential way.

[5] If the size of the constriction becomes too small, turbulent noise will be generated at the constriction and the sound will no longer be a vowel.

[6] This may have a perceptual effect similar to that of nasality as transfer function zeros appear in adult human speech in nasalized vowels.

[7] We have noted (Lieberman and Crelin 1971) that a number of fossils, that differ slightly in other ways, all have a "flattened-out" skull base and other anatomical features that indicate the absence of a supra-

laryngeal vocal tract like adult modern man's. There is, in other words, a class of "Neanderthaloid" fossils who lack the ability to produce the full range of human speech.

[8] Meaningful chimpanzee calls can be "seen" in context in the recent sound motion pictures taken by P. Marler at the Gombe Stream Reserve chimpanzee project of J. Goodall (1965).

[9] I am essentially paraphrasing the discussion presented by A. M. Lieberman (1970) with regard to the linguistic status of human speech and the process of speech encoding. Lieberman's logic is clear, correct, and succinct.

[10] It can be argued that the primary acoustic cue to the identity of [d] is a brief high frequency burst of fricative noise. However, adult listeners will respond correctly to the acoustic signals defined in Figure 13 even though this cue is missing.

[11] The exact size and shape of the vocal tract can be theoretically calculated from the formant frequency pattern of these vowels if all of the theoretically infinite number of formant frequencies are known. If one, however, assumes that the formant structure of an unknown vowel is similar to [i], [u], or [a] and is produced by a cavity shape shown in Figure 9, then the two lowest formants give a good estimate of vocal tract length and size. The "quantal" nature of the speech signal discussed by Stevens (1969) makes an "exact" knowledge of vocal tract size unnecessary for speech decoding.

[12] The Ladefoged and Broadbent (1957) vowel perception study is very pertinent in this regard since it shows that human listeners also cannot tell whether the acoustic signal that is a token of a "central" vowel is an [U], an [I], or an [ae] in the absence of information that tells them the size of the speaker's vocal tract. The listeners in this experiment said that the same acoustic signal "was" the word "bit," "bat," or "but" when prior acoustic context led them to believe that the speaker had a large, medium, or small supralaryngeal vocal tract. Note that a champanzee's response to simple human verbal requests does not demonstrate that the chimpanzee can "decode" human speech. The chimpanzee may be responding to acoustic factors that are not primary linguistic units, e.g., the prosodic features that relate to the emotionally determined "tone" of the speaker's voice. Psychoacoustic experiments designed to show whether non-human primates can "decode" speech have so far

yielded negative results. It is indeed almost impossible to get non-human primates to respond to auditory signals whereas they readily respond to visual signals (Kellogg 1968; Hewes 1971).

[13] Note that the prior existence of a form of language is a necessary condition for the retention, through the process of natural selection, of mutations like the human pharyngeal region that enhance the rate of communication but are detrimental with regard to deglutition and respiration.

[14] We, therefore, see the evolution of language as a process that is ultimately based on mechanisms that exist in the more "primitive" fossil hominids. It is probable that the living apes, as well as other animals, still have similar mechanisms. Studies of animal communication therefore are relevant to the study of human linguistic ability. We obviously do not agree with the theory that bases modern man's linguistic abilities on "unique" mechanisms that require "discontinuities" in evolution (Lenneburg 1967).

[15] We would like to thank Dr. E. L. Simons of Yale University for the Chimpanzee specimen, Dr. P. F. Marler of Rockefeller University for making tape recordings and spectrograms of chimpanzee utterances available, and Drs. A. M. Liberman and C. Darwin of Haskins Laboratories and Dr. W. S. Laughlin of the University of Connecticut for their many useful comments. This work was supported in part by U. S. Public Health Service Grants AM-09499, HD-01994, DE-01774, and NB-04332-8.

REFERENCES CITED

Benda, C. E.
   1969 Down's Syndrome, Mongolism and Its Management. New York: Grune and Stratton.
Boule, M.
   1911-13 L'Homme Fossile de la Chapelle-aux-Saints. Annales de Paleontologie 6:109; 7:21, 85; 8:1.
Chiba, T., and M. Kajiyama
   1958 The Vowel, Its Nature and Structure. Tokyo: The Phonetic Society of Japan.
Chomsky, N.
   1957 Syntactic Structure. The Hague: Mouton.
   1965 Aspects of the Theory of Syntax. Cambridge: M.I.T. Press.
Crelin, E.S.
   1969 Anatomy of the Newborn: An Atlas. Philadelphia: Lea and Febiger.
Darwin, C.
   1859 On the Origin of Species. New York: Atheneum.
Darwin, C.
   1971 Ear Differences in the Recall of Fricatives and Vowels. Quarterly Journal of Experimental Psychology 23:386-392.
Fant, C. G. M.
   1960 Acoustic Theory of Speech Production. The Hague: Mouton.
Gold, B., and L. R. Rabiner
   1968 Analysis of Digital and Analog Formant Synthesizers. IEEE-Transactions on Audio Electroacoustics, AU-16:81-94.
Goodall, J.
   1965 Chimpanzees of the Gombe Stream Reserve. In Primate Behavior. I. DeVore, Ed. New York: Holt, Rinehart & Winston.
Greenewalt, C. A.
   1967 Bird Song: Acoustics and Physiology. Washington, D.C.: Smithsonian Institution.
Haskins Laboratories
   1962 X-Ray Motion Pictures of Speech. New York: Haskins Laboratories.
Hayes, C.
   1952 The Ape in Our House. New York: Harper and Brothers.
Henke, W. L.
   1966 Dynamic Articulatory Model of Speech Production Using Computer Simulation. Unpublished doctoral dissertation, M.I.T.
Hewes, G. W.
   1971 Language Origins: A Bibliography. Boulder: University of Colorado, Department of Anthropology.
Hockett, C. F.
   1960 The Origin of Speech. Scientific American 203:89-96.
Hockett, C. F., and S. A. Altmann
   1968 A Note on Design Features. In Animal Communication. T. A. Sebeok, Ed. Bloomington: Indiana University Press.
Irwin, O. C.
   1957 Speech Development in Childhood. In Manual of Phonetics. L. Kaiser, Ed. Amsterdam: North Holland Publishing Co. pp. 403-425.
Jakobson, R., C. G. M. Fant, and M. Halle
   1952 Preliminaries to Speech Analysis. Cambridge: M.I.T. Press.
Kellogg, W. N.
   1968 Communication and Language in the Home-raised Chimpanzee. Science 162:423-427.

Kirchner, J. A.
1970 Pressman and Kelemen's Physiology of the Larynx, revised edition. Rochester, Minnesota: American Academy of Ophthalmology and Otolaryngology.

Koenig, W., H. K. Dunn, and L. Y. Lacy
1946 The Sound Spectrograph. Journal of the Acoustical Society of America 18:19-49.

Ladefoged, P., and D. E. Broadbent
1957 Information Conveyed by Vowels. Journal of the Acoustical Society of America 29:98-104.

Laughlin, W. S.
1963 Eskimos and Aleuts: Their Origins and Evolution. Science 142:633-645.

La Mettrie, J. O.
1747 de. L'Homme-machine, critical edition, 1960. A. Vartanian, Ed. Princeton: Princeton University Press.

Lenneburg, E. H.
1967 Biological Foundations of Language. New York: Wiley.

Liberman, A. M.
1970 The Grammars of Speech and Language. Cognitive Psychology 1:301-323.

Liberman, A. M., D. P. Shankweiler, and M. Studdert-Kennedy
1967 Perception of the Speech Code. Psychological Review 74:431-461.

Lieberman, P.
1963 Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech. Language and Speech 6:172-187.
1967 Intonation, Perception and Language. Cambridge: M.I.T. Press.
1968 Primate Vocalizations and Human Linguistic Ability. Journal of the Acoustical Society of America 44:1574-1584.
1969 On the Acoustic Analysis of Primate Vocalizations. Behavioral Research Methods and Instrumentation 1:169-174.
1970 *Review of* Perkell's (1968) Physiology of Speech Production. Language Sciences 13:25-28.

Lieberman, P., K. S. Harris, P. Wolff, and L. H. Russell
1972 Newborn Infant Cry and Nonhuman Primate Vocalizations. Journal of Speech and Hearing Research 14:718-727.

Lieberman, P., D. H. Klatt, and W. A. Wilson
1969 Vocal Tract Limitations of the Vocal Repertoires of Rhesus Monkeys and Other Non-human Primates. Science 164:1185-1187.

Lieberman, P., and E. S. Crelin
1971 On the Speech of Neanderthal Man. Linguistic Inquiry 2(2):203-222.

Lindblom, B., and J. Sundberg.
1969 A Quantitative Model of Vowel Production and the Distinctive Features of Swedish Vowels. Speech Transmission Laboratory Report 1. Stockholm, Sweden: Royal Institute of Technology.

Negus, V. E.
1949 The Comparative Anatomy and Physiology of the Larynx. New York: Hafner.

Patte, E.
1955 Les Neanderthaliens, Anatomie, Physiologie, Comparisons. Paris: Masson et cie.

Perkell, J. S.
1969 Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study. Cambridge: M.I.T. Press.

Peterson, G. E., and H. L. Barney
1952 Control Methods Used in a Study of the Vowels. Journal of the Acoustical Society of America 24:175-184.

Rand, T. C.
1971 Vocal Tract Size Normalization in the Perception of Stop Consonants. Status Report 25/26. New Haven: Haskins Laboratories.

Schultz, A. H.
1968 The Recent Hominoid Primates. *In* Perspectives on Human Evolution. S. L. Washburn and Phyllis Jay, Eds. New York: Holt, Rinehart & Winston.

Simpson, G. G.
1966 The Biological Nature of Man. Science 152:472-478.

Sobotta, J., and F. H. J. Figge
1965 Atlas of Human Anatomy, Volume II. New York: Hafner.

Stevens, K. N.
1969 The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data. *In* Human Communication: A Unified View. E. E. David, Jr., and P. B. Denes, Eds. 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 27:484-493.

Troubetzkoy, N. S.
1939 Principles de Phonologie. Translated 1949, J. Cantineau. Paris: Klincksieck.

Truby, H. M., J. F. Bosman, and J. Lind
   1965 Newborn ·Infant Cry. Uppsala,
       Sweden: Almqvist and Wiksells.

Virchow, R.
   1872 Untersuchung des Neanderthal-
       Schädels. Zeitschrift für Ethnographie
       4:157-165.