PROCEEDINGS OF THE SEVENTH INTERNATIONAL CONGRESS OF PHONETIC SCIENCES

Held at the University of Montreal and McGill University
22-28 August 1971

Edited by

ANDRÉ RIGAULT

Professor in Linguistics at McGill University

and

RENÉ CHARBONNEAU

Professor in Phonetics at the University of Montreal

1972

MOUTON

THE HAGUE · PARIS

ON THE EVOLUTION OF HUMAN LANGUAGE

PHILIP LIEBERMAN

Human language is one of the defining characteristics that differentiate modern man from all other animals. The traditional view concerning the uniqueness of human linguistic ability is that it is based on man's mental processes (Lenneberg 1967). In other words, the 'uniqueness' of human language is supposed to be entirely due to the properties of the human brain. The particular sounds that are employed in human language are therefore often viewed as an arbitrary, fortuitously determined set of cipher-like elements. Any other set of sounds or gestures supposedly would be just as useful at the communicative, i.e., the phonetic level of human language.

The results of recent research have, however, challenged this view. The 'motor theory' of speech perception that has been developed over the past fifteen years, in essence, states that speech signals are perceived in terms of the constraints that are imposed by the human vocal apparatus (Liberman, Shankweiler and Studdert-Kennedy 1967). Other recent research, which I will attempt to summarize in this paper, indicates that the anatomic basis of human speech production is itself species-specific. This research is the product of a collaborative effort involving many skills. Edmund S. Crelin of the Yale University School of Medicine, Dennis H. Klatt of the Massachusetts Institute of Technology, Peter Wolff of Harvard University, and my colleagues at the University of Connecticut and Haskins Laboratories, have all been involved at one time or another. Our research indicates that the anatomic basis of human speech production is the result of a long evolutionary process in which the Darwinian process of Natural Selection acted to retain mutations that would enhance rapid communication through the medium of speech. The neural processes that are involved in the perception of speech and the unique species-specific aspects of the human supra-laryngeal vocal tract furthermore appear to be interrelated in a positive way.

1. VOCAL TRACT RECONSTRUCTION

The most direct approach to this topic is to start with our most recent experimental technique, the reconstruction and functional modelling of the speech-producing

anatomy of extinct fossil hominids. We have been able to reconstruct the evolution of the human supra-laryngeal vocal tract by making use of the methods of comparative anatomy and skeletal similarities that exist between extinct fossil hominids, and living primates (Lieberman and Crelin 1971). In Figure 1 inferior views of the base of the
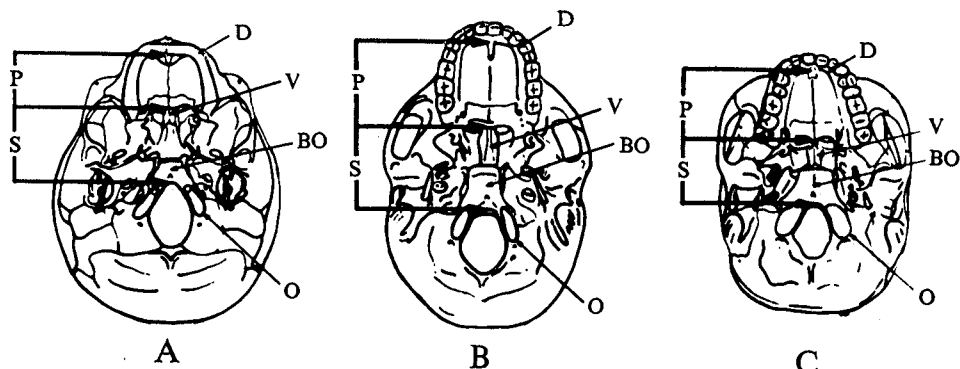

Fig. 1. Inferior views of base of skull of Newborn (A), Neanderthal (B), and adult Man (C). D — Dental Arch, P — Palate, S — Distance Between Palate and Foramen Magnum, V — Vomer Bone, BO — Basilar Part of Occipital, O — Occipital Condyle (after Lieberman and Crelin 1971).

skull are shown for newborn modern man, a reconstruction of the fossil La Chapelle-aux-Saints Neanderthal man and an adult modern man. The detailed morphology of the base of the skull and mandible, which is similar in newborn modern man and Neanderthal man, forms the basis for the Neanderthal reconstruction. Some of the skull features that are similar in newborn modern man and Neanderthal man, but different from adult modern man are as follows: (A) a generally flattened-out base, (B) lack of a chin, (C) the body of the mandible is 60 to 100 percent longer than the ramus, (D) the posterior border of the mandibular ramus is markedly slanted away from the vertical plane, (E) a more horizontal inclination of the mandibular foramen leading to the mandibular canal, (F) the pterygoid process of the sphenoid bone is relatively short and its lateral lamina is more inclined away from the vertical plane, (G) the styloid process is more inclined away from the vertical plane, (H) the dental arch of the maxilla is U-shaped instead of V-shaped, (I) 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, (J) the roof of the nasopharynx is a relatively shallow elongated arch, (K) the vomer bone is relatively short in its vertical height and its posterior border is inclined away from the vertical plane, (L) the vomer bone is relatively far removed from the junction of the sphenoid bone and the basilar side part of the occipital bone, (M) the occipital condyles are relatively small and elongated. These similarities are in accord with other skeletal features typical of Neanderthal fossils (Vlček 1970), which may be seen in the course of the ontogenetic development of modern man. This, parenthetically, does not mean that Neanderthal

man was a direct ancestral form of modern man, since Neanderthal fossils exhibit specializations like brow ridges that never occur in the ontogenetic development of modern man. Modern man, furthermore, deviates quite drastically from Neanderthal man in the course of normal maturation from the newborn state.

In Figure 2 lateral views of the skull, vertebral column, and larynx of newborn and adult modern man and Neanderthal man are presented. The significance of the afore-mentioned skeletal features with regard to the supralaryngeal vocal tract can be seen in the high position of the larynx in newborn and in Neanderthal.
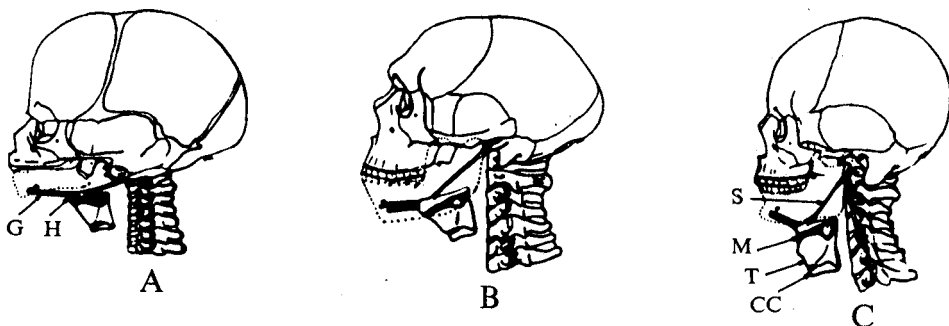

Fig. 2.   Skull, vertebral column and larynx of Newborn (A), and adult Man (C), and reconstruction of Neanderthal (B). G — Geniohyoid Muscle, H — Hyoid Bone, S — Stylohyoid Ligament, M — Thyrohyoid Membrane, T — Thyroid Cartilage, CC — Cricoid Cartilage. Note that the inclination of the styloid process away from the vertical plane in Newborn and Neanderthal results in a corresponding inclination in the stylohyoid ligament. The intersection of the stylohyoid ligament and geniohyoid muscle with the hyoid bone of the larynx occurs at a higher position in Newborn and Neanderthal. The high position of the larynx in the Neanderthal reconstruction follows, in part, from this intersection. (After Lieberman and Crelin 1971).

In Figure 3 the supralaryngeal air passages of newborn and adult man and the Neanderthal reconstruction are diagrammed so that they appear equal in size. Although the nasal and oral cavities of Neanderthal are actually larger than those
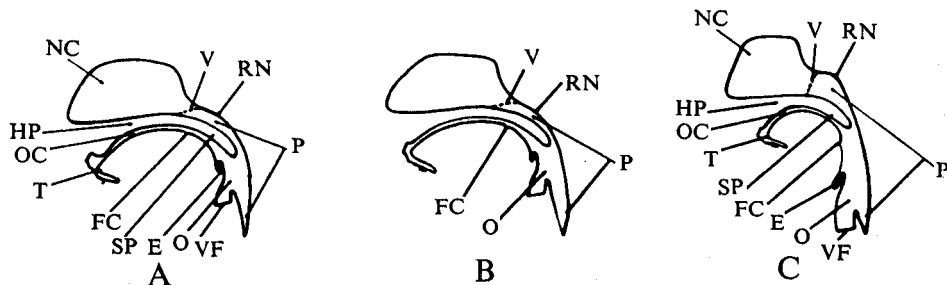

Fig. 3.   Supralaryngeal air passages of Newborn (A), Neanderthal reconstruction (B), and adult Man (C). NC — Nasal Cavity, V — Vomer Bone, RN — Roof of Nasopharynx, P — Pharynx, HP — Hard Palate, SP — Soft Palate, OC — Oral Cavity, T — Tip of Tongue, FC — Foramen Cecum of Tongue, E — Epiglottis, O — Opening of Larynx into Pharynx, VF — Level of Vocal Folds (after Lieberman and Crelin 1971).

of adult modern man, they are quite similar in shape to those of the newborn. The long 'flattened-out' base of the skull in newborn and Neanderthal is a concomitant skeletal correlate of a supralaryngeal vocal tract in which the entrance to the pharynx lies behind the entrance to the larynx. In the ontogenetic development of adult modern man the opening of the larynx into the pharynx shifts to a low position. In this shift the epiglottis becomes widely separated from the soft palate. The posterior part of the tongue, between the foramen cecum and the epiglottis, shifts from a horizontal resting position within the oral cavity to a vertical resting position, to form the anterior wall of the oral part of the pharynx (Figure 3C). In this shift the epiglottis becomes widely separated from the soft palate.

The uniqueness of the adult human supralaryngeal vocal tract rests in the fact that the pharynx and oral cavities are almost equal in length and are at right angles. No other animal has this 'bent' supralaryngeal vocal tract in which the cross-sectional areas of the oral and pharyngeal cavities can be independently modified. The human vocal tract can, in effect, function as a 'two-tube' acoustic filter. In Figure 4 we have diagrammed the 'bent' human supralaryngeal vocal tract in the production of the 'extreme', 'point' vowels /i/, /a/, and /u/. Note that the midpoint area function
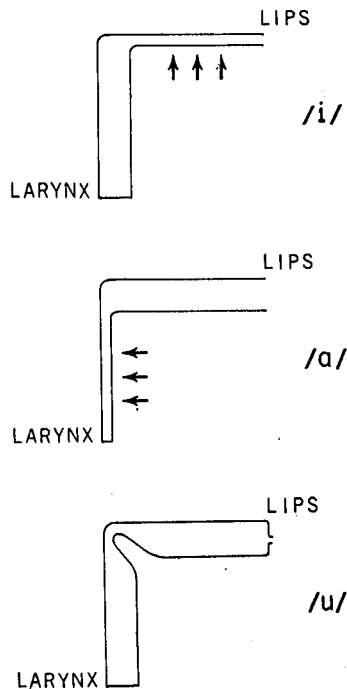

Fig. 4. Schematic diagram of the 'bent' human supralaryngeal vocal tract (note that abrupt and extreme discontinuities in cross-sectional area can occur at the midpoint).

changes are both extreme and abrupt. Abrupt discontinuities can be formed at the midpoint 'bend'. In Figure 5 the non-human 'straight' vocal tract which is typical of all living nonhuman primates (Lieberman 1968, Lieberman, Klatt and Wilson 1969 and Lieberman, Crelin and Klatt 1972), newborn humans (Lieberman, Harris,
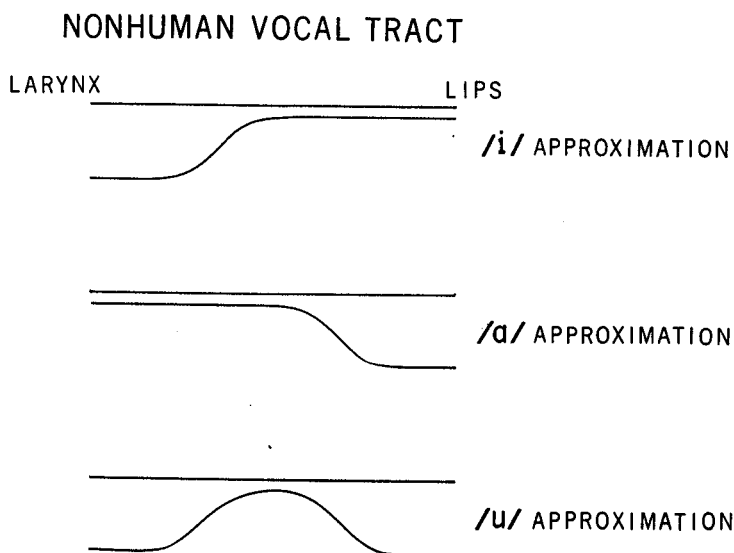
NONHUMAN VOCAL TRACT

Fig. 5.   Schematic diagram of the straight, 'single tube' nonhuman vocal tract (note that abrupt midpoint constrictions cannot be formed).

Wolff and Russell 1968) and Neanderthal man, is diagrammed as it approximates these vowels. All area function adjustments have to take place in the oral cavity in the nonhuman supralaryngeal vocal tract. Although midpoint constrictions obviously can be formed in the nonhuman vocal tract, they cannot be both extreme and abrupt. The elastic properties of the tongue prevent it from forming abrupt discontinuities at the midpoint of the oral cavity.

2. VOCAL TRACT MODELLING

Human speech is essentially the product of a source, the larynx for vowels, and a supralaryngeal vocal tract transfer function. The supralaryngeal vocal tract, in effect filters the source (Chiba and Kajiyama 1958, Fant 1960). The activity of the larynx determines the fundamental frequency of the vowel, whereas its formant frequencies are the resonant modes of the supralaryngeal vocal tract. The formant frequencies are determined by the area function of the supralaryngeal vocal tract. Man uses his articulators (the tongue, lips, mandible, pharyngeal constrictors, etc.) to modify dynamically in time the formant frequency patterns that the supralaryngeal vocal

tract imposes on the speech signal. The phonetic inventory of a language is therefore limited by (1) the number of source function modifications that a speaker is capable of controlling during speech communication, and (2) the number of formant frequency patterns available by changing the supralaryngeal area function through the dynamic manipulation of the articulators. We thus can assess the contribution of the supra-laryngeal vocal tract to the phonetic abilities of a hominid, independent of the source characteristics. A computer-implemented model of a supra-laryngeal vocal tract (Henke 1966) can be used to determine the possible contribution of the vocal tract to the phonetic repertoire. We can conveniently begin to determine whether a non-human supra-laryngeal vocal tract can produce the range of sounds that occur in human language by exploring its vowel producing ability. Consonantal vocal tract configurations can also be modelled. 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.

In Figure 6 we have presented area functions of the supra-laryngeal vocal tract of Neanderthal man that were modelled on the computer. These area functions were directed towards best approximating the human vowels /i/, /a/, and /u/. Our computer modelling (Lieberman and Crelin 1971) was guided by the results of X-ray motion pictures of speech production, swallowing and respiration in adult humans (Haskins Laboratories 1962; Perkell 1969) and in newborns (Truby, Bosma and Lind 1965). This knowledge, plus the known comparative anatomy of the living primates, allowed a fairly 'conservative' simulation of the vowel producing ability of classic Neanderthal man. We perhaps allowed a greater vowel producing range for Neanderthal man since we consistently generated area functions that were more human-like than ape-like whenever we were in doubt. Despite these compensations the Neanderthal vocal tract cannot produce /i/, /a/, or /u/.

In Figure 7 the formant frequency patterns calculated by the computer program for the numbered area functions of Figure 6 are plotted. The labelled loops are derived from the Peterson and Barney (1952) analysis of the vowels of American English of 76 adult men, women and children. Each loop encloses the data points that accounted for 90 percent of the samples in each vowel category. We have compared the formant frequencies of the simulated Neanderthal vocal tract with this comparatively large sample of human speakers since it shows that the speech deficiencies of the Neander-thal vocal tract are different in kind from the differences that characterize human speakers. Since all human speakers can inherently produce all the vowels of American English, we have established that the Neanderthal phonetic repertoire is inherently limited. In some instances we generated area functions that would be humanlike, even though we felt that we were forcing the articulatory limits of the reconstructed Neanderthal vocal tract (e.g., area functions 3, 9, and 13). However, even with these articulatory gymnastics the Neanderthal vocal tract could not produce the vowel range of American-English.
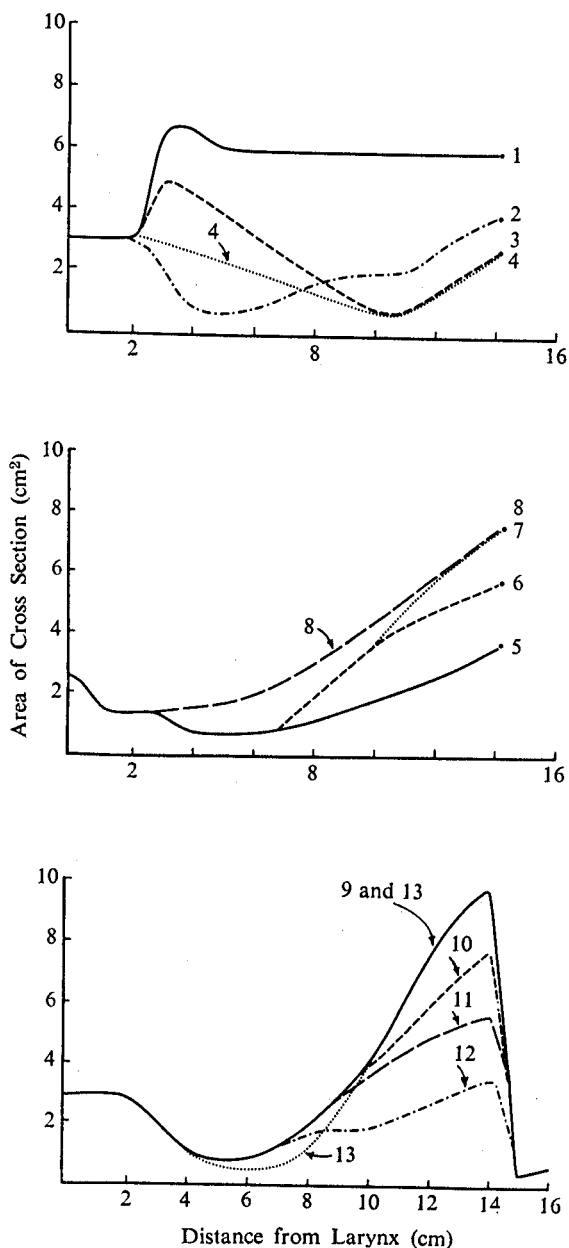
Fig. 6. Area Functions of the supralaryngeal vocal tract of Neanderthal reconstruction modelled on computer. The area function from 0 to 2 cm is derived from Fant (1960) and represents the distance from the vocal folds to the opening of the larynx into the pharynx. Curve 1 is the unperturbed tract. Curves 2, 3, and 4 represent functions directed towards a 'best match' to the human vowel /i/. Curves 5-8 are functions directed towards a 'best match' to /a/, while curves 9-13 are directed towards /u/. (after Lieberman and Crelin 1971).
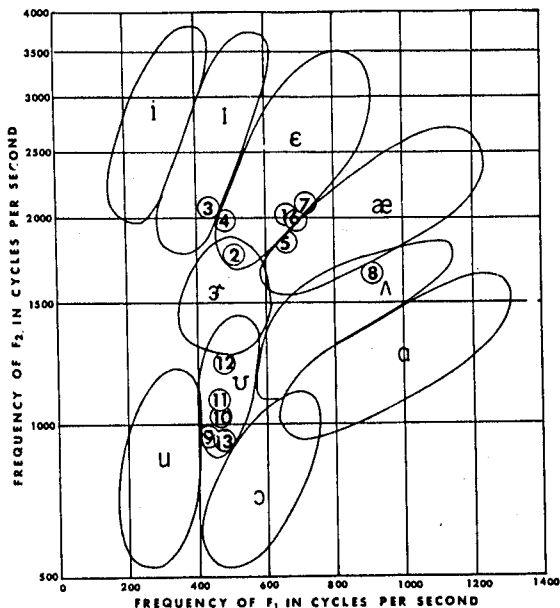
Fig. 7. Formant frequencies calculated by computer program for Neanderthal reconstruction (the numbers refer to area functions in Figure 6) (after Lieberman and Crelin 1971).

3. FUNCTIONAL PHONETIC LIMITATIONS

There are some special considerations that follow from the absence of the vowels /i/, /a/, and /u/ from the Neanderthal phonetic repertoire. Phonetic analyses have shown that these "point" vowels are the limiting articulations of a vowel triangle that is almost language universal (Troubetzkoy 1939). The special nature of /i/, /a/, and /u/ can be argued from theoretical grounds as well. Employing simplified and idealized area functions (similar to those sketched in Figure 4), 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 /i/, /a/, and /u/ have another unique acoustical property. They are the only vowels in which an acoustic pattern can be related to a unique vocal tract area function. Other 'central' vowels 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 /i/, /a/, and /u/ can calculate the size of the supra-laryngeal 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 tract can, for example, by increasing 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 /i/, /a/, and /u/ since the required discontinuities and constrictions in the supra-laryngeal vocal tract area functions produce acoustic patterns that are beyond the range of compensatory maneuvers.

4. SPEECH PERCEPTION AND SPEECH ANATOMY

We noted, at the start of this paper, that results of perceptual research have demonstrated that human listeners perceive speech in terms of the constraints imposed by the speech producing apparatus. This mode of perception, which has been termed the 'Speech' or 'Motor Theory' mode of perception, makes the rapid rate of information transfer of human speech possible (Liberman 1970). Human listeners can perceive as many as 30 phonetic segments per second in normal speech. This information rate far exceeds the temporal resolving power of the human auditory system. It is, for example, impossible to even count simple pulses at rates of 20 pulses per second. The pulses merge into a continuous tone. 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 'motor theory' of speech perception, in essence, explicates this process. The presence of vowels like /i/, /a/, and /u/ appears to be one of the anatomic factors that makes this encoding process possible.

In Figure 8 we have reproduced two simplified spectrographic patterns that will, when converted to sound, produce approximations to the syllables /di/ and /du/
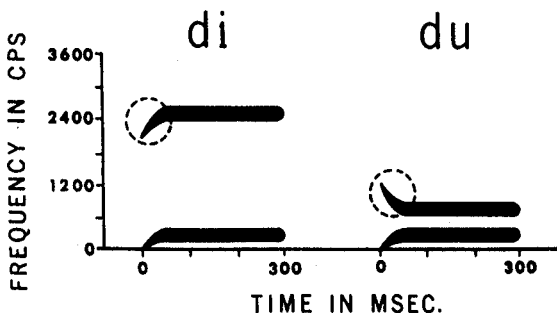

Fig. 8.   Simplified spectrographic patterns sufficient to produce the syllables /di/ and /du/ (the circles enclose the second formant frequency transitions) (after Liberman 1970).

(Liberman 1970). The dark bands on these patterns represent the first and second formant frequencies of the supra-laryngeal 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 msecs., are called 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, 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 supra-laryngeal vocal tract. The formant patterns that define the syllable /di/ in Figure 8 thus reflect the changing resonant pattern of the supra-laryngeal 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 production of speech. The listener in this process, which has been termed the "motor theory of speech perception" (Liberman, Shank, weiler and Studdert-Kennedy 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 speaker's 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 supra-laryngeal 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 syllables' 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. 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 acoustic calibration signals in human speech.

The absence of a human-like pharyngeal region in apes, newborn man and Neanderthal man is quite reasonable. The only function that the human supra-laryngeal vocal tract is better adapted to is speech production, in particular the production of vowels like /a/, /i/, and /u/. The human supra-laryngeal vocal tract is otherwise less well adapted for the primary vegetative functions of respiration, chewing, and swallowing (Lieberman, Crelin and Klatt 1972, Crelin, Lieberman and Klatt, forthcoming). This suggests that the evolution of the human vocal tract which allows vowels like /a/, /i/, and /u/ to be produced, and the universal 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.

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.

It is imperative to note that classic Neanderthal man, as typified by fossils whose skull bases are similar to the La Chapelle-aux-Saints, La Ferrasie, La Quina, Pech-

de-l'Azé and Monte Circeo fossil hominids (as well as many others), probably does not represent the mainstream of human evolution. Although Neanderthal man and modern man probably had a common ancestor, Neanderthal represents a divergent group (Boule and Vallois 1957; Vlček 1970; Lieberman and Crelin 1971). In Figure 9 we have photographed a casting of a reconstruction of the fossil Steinheim
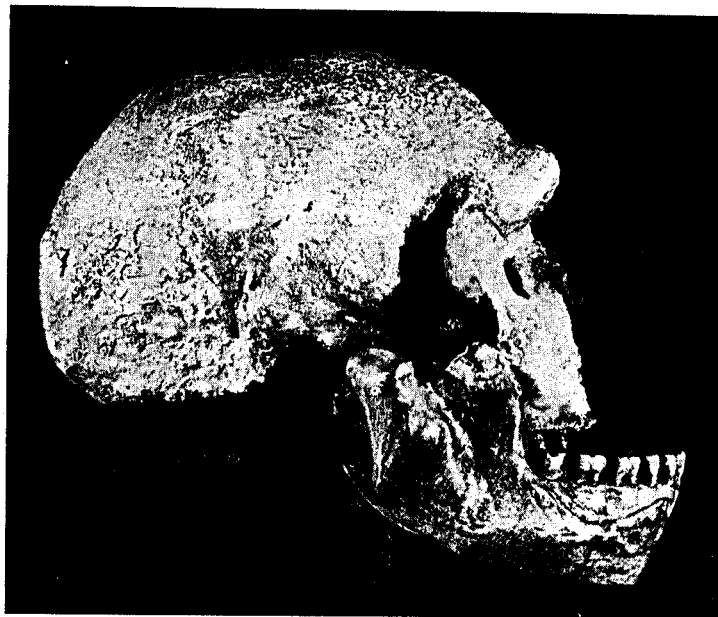

Fig. 9. Reconstructed Steinheim clavarium with Neanderthaloid mandible. Note that the Neanderthal mandible is too large. (After Crelin, Lieberman and Klatt forthcoming.)

calvarium with the mandible of the La Chapelle-aux-Saints fossil. The mandible of the Steinheim fossil hominid never was found. Note that the La Chapelle-aux-Saints mandible is too long. In Figure 10 the Steinheim fossil has been fitted with a mandible from a normal adult human, which best 'fits' the Steinheim fossil. We are in the process of reconstructing the supra-laryngeal vocal tract of the Steinheim fossil (Crelin, Lieberman and Klatt, forthcoming). It is quite likely that this fossil, which is approximately 300,000 years old, had the vocal tract anatomy that is necessary for human speech. The evolution of the anatomical basis for human speech thus would not appear to be the result of abrupt, recent change in the morphology of the skull and soft tissue of the vocal tract. We have noted a number of fossil forms that appear to represent intermediate stages in the evolution of the vocal tract. Recent fossil discoveries indicate that the evolution of the human vocal tract may have started at least 2.6 million years ago. It, therefore, is not surprising to find that the neural aspects of speech perception are matched to the anatomical aspects of speech production. Nor should we be surprised to note that 'naturalness' constraints relate the
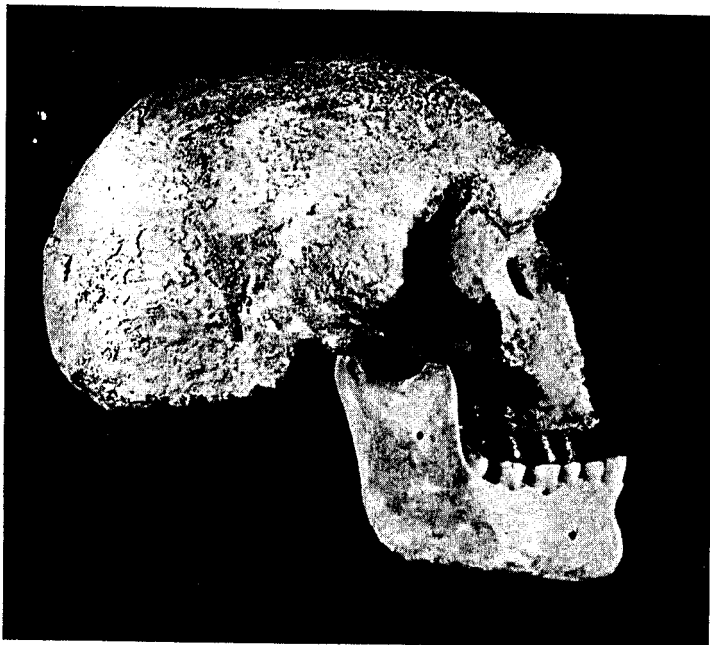
Fig. 10.   Reconstructed Steinheim clavarium with a modern human mandible. This represents the best "fit". (after Crelin, Lieberman and Klatt forthcoming).

phonetic and phonologic levels of grammar (Jakobson, Fant and Halle 1952, Postal 1968, Chomsky and Halle 1969).

Sir Arthur Keith many years ago speculated on the antiquity of man. We now know that hominid evolution can be traced back at least three million years. The evolution of phonetic ability appears to have been an integral part of this evolutionary process. It may have its origins at the very beginnings of hominid evolution.

University of Connecticut
Storrs, Connecticut, and
Haskins Laboratories
New Haven, Connecticut

REFERENCES

Boule, M. and H.V. Vallois
   1957   *Fossil Men* (New York, Dryden Press).
Chiba, T. and M. Kajiyama
   1958   *The Vowel, Its Nature and Structure* (Tokyo, Phonetic Society of Japan).
Chomsky, N. and M. Halle
   1968   *The Sound Pattern of English* (New York, Harper).
Crelin, E.S., P. Lieberman, and Dennis Klatt
   forthcoming   "Speech Abilities of the Steinheim, Skhul V, and Rhodesian Fossil Hominids".

Darwin, C.
  1971   "Ear Differences in the Recall of Fricatives and Vowels", *Quarterly Journal of Experimental Psychology* 23.
Fant, G.
  1960   *Acoustic Theory of Speech Production* (The Hague, Mouton).
Haskins Laboratories
  1962   *X-Ray Motion Pictures of Speech* (Haskins Laboratories, 305 E. 43 St., New York City).
Henke, W.L.
  1966   "Dynamic Articulatory Model of Speech Production Using Computer Simulation", unpublished Doctoral dissertation (Massachusetts Institute of Technology).
Jakobson, R., C.G.M. Fant and M. Halle
  1952   *Preliminaries to Speech Analysis* (Cambridge, M.I.T. Press).
Ladefoged, P. and D.E. Broadbent
  1957   "Information Conveyed by Vowel", *Journal of the Acoustic Society of America* 29:98-104.
Lenneberg, 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.
  1968   "Primate Vocalizations and Human Linguistic Ability", *Journal of the Acoustic Society of America* 44:1574-1584.
Lieberman, P. and E.S. Crelin
  1971   "On the Speech of Neanderthal man", *Linguistic Inquiry* 2:203-222.
Lieberman, P., E.S. Crelin, and D.H. Klatt
  1972   "Phonetic Ability and Related Anatomy of the Newborn and Adult Human, Neanderthal Man and the Chimpanzee", *American Anthropologist* 74, 3.
Lieberman, P., D. H. Klatt and W.A. Wilson
  1969   "Vocal Tract Limitations of the Vocal Repertoires of Rhesus Monkey and Other Non-Human Primates", *Science* 164:1185-1187.
Lieberman, P., K.S. Harris, P. Wolff, and L.H. Russell
  1968   "Newborn Infant Cry and Nonhuman Primate Vocalizations", *Haskins Laboratories Status Report* 17/18 (Haskins Laboratories, New York City).
Perkell, J.S.
  1969   *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study* (Cambridge, Mass., 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* 42:175-184.
Postal, P.M.
  1968   *Aspects of Phonological Theory* (New York, Harper and Row).
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   *Principes de Phonologie*, trans. 1949 (Paris, J. Cantineau, translator, Klincksieck).
Truby, H.M., J.F. Bosma, and J. Lind
  1965   *Newborn Infant Cry* (Uppsala, Almqvist and Wiksells).
Vlček, E.
  1970   "Étude comparative onto-phylogénétique de l'enfant du Pech-de-L'Azé par rapport à

d'autres enfants Neandertaliens", in Ferembach *et al.*, *L'Enfant du Pech-de-L'Azé* (*Mémoire* 33, *Archives de l'Institut de Paléontologie Humaine*) (Paris, Masson et Cie).

DISCUSSION

ALLEN (Chapel Hill, N.C.)

I wish to thank Professor Lieberman for his fine presentation of a fascinating subject. Could you please clarify one matter for me? The comparison of the skulls of human newborn, Neanderthal, and human adult was difficult for me to follow, because of the great detail present on your slides. Are such comparisons carried out by well-established and validated anthropological methods, or are they intuitive in nature?

LIEBERMAN

The procedures followed by D₁. Crelin, my collaborator, and myself are consistent with the methods of comparative anatomy. Our result is furthermore consistent with earlier reconstruction by Sir Arthur Keith (noted in Negus's *Comparative Anatomy and Physiology of the Larynx*, London, 1947) and inferences by Debrul and Coon (cf. Lieberman and Crelin 1971). The recent study of Neanderthal ontogenetic development (l'Enfant du Pech-de-l'Azé) also offers an independent replication of our conclusions.

CATFORD (Ann Arbor, Mich.)

Dr. Lieberman's contribution was as fascinating as his earlier published work on this theme. However, I would comment that even if it is proven that Neanderthal man could produce only a narrow range of central vowels, of which I am a little sceptical, he might still be capable of quite 'normal' 'modern-like' speech. We have the evidence of six N.W. Caucasian languages which have only one, two or three basically central-type vowel phonemes. Even though these languages can also have phonetic extreme vowels, as allophones dependent on surrounding semi-vowels, etc. nevertheless long sketches of perfectly intelligible speech occur in which the only vowels are of /ə/-type.

LIEBERMAN

Neanderthal man could not produce these languages precisely because he could not produce the extreme vowels that Dr. Catford notes are present in these languages.

Modern human speakers, in all likelihood, make use of these extreme vowels to ascertain the size of the vowel tracts of individual speakers. This information is essential for the speech 'decoding' that is the basis of the rapid rate of information transfer of human speech. Neanderthal man, though he could produce part of the human phonetic repertoire, would be incapable of speaking any human language.

PADDOCK (Wolfville, N.S.)

I agree that the ability to produce a number of distinctive sounds at a rapid rate is a prerequisite for language. However, there are other prerequisites. One of these is that of having something to say. For example, such an apparently primitive linguistic process as the naming of physical objects involves a great deal of abstraction and generalization. Any evidence that some non-human species possess such concept-forming ability would seem to strengthen your case. Their lack of language could then be more rightly attributed to the limitations of their articulatory abilities.

LIEBERMAN

Professor Paddock's comment is quite relevant since recent experiments with chimpanzees indicate that the logical abilities of these animals clearly exceed their communicative vocal abilities. The experiments reported by the Gardeners in *Science* show that sign language can be used to establish communication with a chimpanzee. Premack's recent work, also reported in *Science*, shows that plastic symbols can be used to communicate with a chimpanzee. The chimpanzees in these experiments may not exhibit all the logical abilities of human language, but they evidence concept-forming abilities that far exceed their vocal abilities. The primary factor that prevents these animals from communicating by means of human language would appear to rest in the limitations of their articulatory and perceptual abilities.

JONES (Buffalo, N.Y.)

Neanderthal man need not have had oral language in order to produce his artifacts. In many American Indian cultures, boys learned how to hunt, to forage, and even to tan hides — extremely hard, complicated tasks — with no verbal instruction whatever.

Neanderthal's calendars and burials, unlike his weapons, probably do imply communication of some sort. But even human communication is normally kinesic as well as linguistic. Neanderthal might have communicated with little or no verbal language, especially since his anatomy allowed only such a slow, inefficient form.

LIEBERMAN

I quite agree with this comment. It is quite likely that Neanderthal language was structured in terms of gestures as well as 'simple' vocal signals. Hewes at the University of Colorado has advanced a gestural theory. It is also likely that some 'slow' oral signals were present in Neanderthal language. Since Natural Selection always operates in slow states some oral signals will have had to be present for the relation of mutations that enhanced phonetic ability.

LAFON (Besançon)

La théorie élégante de M. Lieberman est une construction de l'esprit fort bien présentée, malheureusement en contradiction avec les données des sciences anthropologique, physiologique, anatomique. J'en note quelques exemples.

1. Le crâne présenté est celui d'un homme primitif? La découverte de crânes plus anciens et de structure plus humaine semble montrer qu'il y a probablement des lignées parallèles éteintes dont fait partie l'homme présenté par M. Lieberman; celui-là ne serait donc pas l'ancêtre de l'homme.

2. Les théories de Darwin sur l'origine simiesque de l'homme en ligne directe et progressive telle que le laisse penser le conférencier ne sont depuis longtemps plus admises.

3. Comparer ontogenèse et philogenèse est également une démarche abandonnée depuis longtemps tant elle amène à de fausses théories. Le petit de l'homme ne refait pas les démarches de l'espèce, la génétique le nie de façon formelle.

4. L'os est le tissu le plus malléable de l'organisme; il subit mais ne détermine pas. Ainsi le maxillaire inférieur a un angle tributaire de l'état dentaire, l'angle est ouvert chez les nourrissons et aussi chez le vieillard.

5. Le fœtus, les premiers mois, est identique, sur le plan morphologique, à celui du porc; que n'en tirerait-on pas sur le plan théorique en exploitant cette identité de forme!

6. Admettre une théorie mécaniste de la parole (et non du langage comme le dit le conférencier), c'est donner à l'organe le pouvoir de déterminer la fonction. Cette démarche est totalement fausse en physiologie: la fonction provoque le développement ou l'évolution de l'organe, et non l'inverse.

7. La fonction phonatoire n'a pas d'unité anatomique ni organique; elle suppose une liaison fonctionnelle de l'appareil respiratoire et de l'appareil digestif. Son usage et sa formation ne peuvent être reliés à l'organe, c'est une intercorrélation complexe neuro-physiologique qui fait la phonation.

CONCLUSION: que reste-t-il de la théorie présentée qui cependant représente un important travail expérimental?

LIEBERMAN

Professor Lafon makes a number of comments which call for correction. His comments on the evolutionary status of Neanderthal man are peculiar since I have noted that we also believe that Neanderthal man is not a direct ancestor of modern man. Neanderthal man's linguistic deficiencies may perhaps account for his being extinct. We have shown that he did not have the anatomical structure that is necessary for human speech. Our preliminary reconstructions indicate that other fossil hominids may be 'closer' to modern man insofar as they have humanlike speech producing anatomies.

Professor Lafon's comment regarding Darwin's views on a 'single' line for human evolution is incorrect. I suggest that he reread the *On the Origin of Species* for Darwin's views on evolution and natural selection.

Professor Lafon's views on the relationship between ontogeny and phylogeny are not relevant. There are other serious inadequacies in Professor Lafon's views on ontogenetic development. The fact that the human foetus shares similar traits with

pigs, of course, does not mean that man is descended from a pig — but it is consistent with the fact that modern man and pigs are both mammals and that both probably have a distant common ancestral form. Neanderthal man and modern man likewise probably had a common ancestor though they diverged. If Professor Lafon will read our previous papers as well as the recent monograph *L'Enfant du Pech-de-L'Azé* he will note that ontogenetic development provides useful insights into human evolution.

Professor Lafon's comment on changes in the morphology of the bones of the skull merely reflects the fact that extreme changes must take place as the human newborn develops towards the adult state.

Professor Lafon's final comment regarding the supposed irrelevance of anatomical structures for speech merely reflects Professor Lafon's background. Our research, which has involved acoustic analysis, anatomical examination and the modelling of supra-laryngeal vocal tracts, has demonstrated that the human supralaryngeal vocal tract is species-specific and that it has evolved in a manner that enhances speech production at the expense of vegetative functions like respiration. Human speech production is simply not an overlaid function that makes use of an apparatus that has evolved for respiration and swallowing.