

METHODS & DESIGNS

On the acoustic analysis of primate vocalizations

PHILIP LIEBERMAN, DEPARTMENTS OF LINGUISTICS AND ELECTRICAL ENGINEERING, UNIVERSITY OF CONNECTICUT, Storrs, Connecticut 06268, and HASKINS LABORATORIES, 305 East 43rd Street, New York, 10017

The acoustic theory of speech production relates articulatory maneuvers to acoustic attributes of speech. Some procedures are discussed that make use of this theory to analyze the vocalizations of nonhuman primates. These procedures make use of sound spectrograms, oscillograms, and computer-implemented analogs of the primate vocal apparatus as well as anatomical measurements. The use of these techniques in recent studies of nonhuman primate vocalizations is reviewed. These studies show that nonhuman primates lack the anatomical apparatus that is necessary for the production of the full range of human speech. Some unresolved questions concerning the structure of nonhuman primate utterances are discussed with regard to human linguistic ability.

The object of this paper is to review some of the analytical methods that are appropriate for the study of the phonetics of primate utterances. We shall discuss the acoustic theory of speech production and some of the known differences between the nonhuman primates and man. We shall also discuss some possible similarities in the acoustic communications of man and the nonhuman primates that should be studied in more detail.

Although research on the acoustic and articulatory bases of speech communication has a long history, in the past 30 years a quantitative acoustic theory of speech production has been developed (Chiba & Kajiyama, 1958; Fant, 1960). This theory allows us to relate the acoustic speech signal to the articulatory maneuvers that humans use when they speak. The acoustic theory of speech production also permits us to evaluate the acoustic significance of articulatory maneuvers and anatomical structures.

PROCEDURES INVOLVED IN ACOUSTIC ANALYSIS

We will use the term "acoustic analysis" in a rather loose sense since we will actually describe some of the techniques that have been used in two recent studies (Lieberman, 1968; Lieberman, Klatt, & Wilson, 1969). These studies used anatomical and computer-modeling procedures to investigate the vocal repertoires of nonhuman primates.

Obviously, one must have adequate tape recording facilities in order to analyze the utterances of any animals. We used Sony Type TC 800 tape recorders with both Sony Type F85 and General Radio Type 1560 P5 microphones at a tape speed of 7.5 in./sec; the response was 6 dB down at 16 kHz. The system was flat to 12 kHz. The tape recorder and microphone power supplies were battery-powered, which made recordings in zoos practical. The acoustical analysis involved the use of a sound spectrograph (Voiceprint), an oscilloscope (Honeywell Visicorder), and a medium-size digital computer (Digital Equipment Corporation PDP-9).

The upper limit on the frequency response of our recording system appeared to be adequate for the primate vocalizations recorded in these studies. If one were interested in the smaller primates, e.g., squirrel monkeys (*Saimiri sciureus*) whose vocalizations appear to involve higher frequency components, the

upper limit on the frequency response of the recording system would have to be higher.

ACOUSTIC THEORY OF SPEECH PRODUCTION

It is both convenient and correct to regard human speech in terms of two quantities: an excitation source and a filter. In the production of the vowel /a/, for example, the quasiperiodic opening and closing motions of the vocal cords generate a periodic excitation source. This laryngeal source is filtered by the supralaryngeal vocal tract. The area function of the supralaryngeal vocal tract determines the filter function of the supralaryngeal vocal tract. The local energy maxima of the supralaryngeal filter function are the "formant" frequencies. The vowel /a/ for an adult male might have formant frequencies of 700, 900, and 2700 Hz. The formants for /i/ for the same speaker might be 300, 2100, and 3200 Hz (Fant, 1960). The speaker could produce either vowel using the same laryngeal excitation. He could, for example, phonate at a fundamental frequency of 130 Hz for both /i/ and /a/. The area function of his supralaryngeal vocal tract would, however, be quite different for these two vowels. The speaker could alternatively phonate the two vowels at different fundamental frequencies, say 130 and 200 Hz. His vocal cords would open and close at different rates in order to produce these two different fundamental frequencies. The phonetic quality of the vowels /a/ and /i/ would, however, be preserved. The speaker could even whisper the two vowels by keeping his vocal cords in an open position and exciting the supralaryngeal vocal tract by means of a turbulent noiselike source. The formants for /a/ and /i/ in all these cases would be a function of the area function of the supralaryngeal vocal tract.

MEASUREMENTS OF NONHUMAN PRIMATE VOCALIZATIONS

Figure 1 shows a reproduced spectrogram from one of the studies of primate vocalizations that we will discuss (Lieberman,

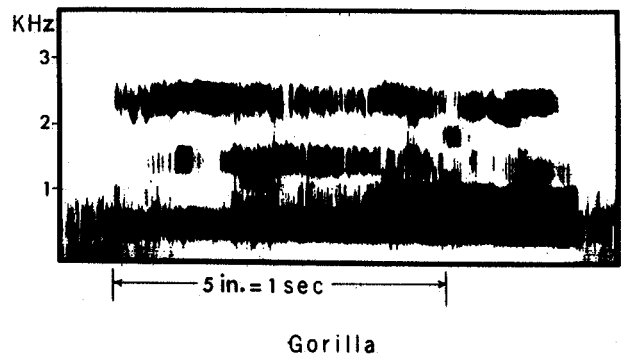


Fig. 1. Spectrogram of cry produced at moderate intensity by 3-year-old gorilla after Lieberman (1968). The bandwidth of the analyzing filter was 300 Hz. The fundamental frequency of phonation ranged from 100 to 120 Hz. The configuration of the gorilla's supralaryngeal vocal tract apparently approximated a uniform tube open at one end, the schwa vowel, since the formant frequencies of the cry occurred at 500, 1500, and 2400 Hz.

1968). This vocalization was produced by a 3-year-old female gorilla at a moderate level of intensity when food was withheld. The spectrogram was made using the "normal" display function on the Voiceprint machine that produces a "conventional" sound spectrogram. Two display options are available on the machine: "normal" and "contour." In both displays energy is displayed as a function of frequency vs time. In the "normal" mode the intensity of energy at a particular frequency is a function of the degree of blackness of the display. In the "contour" display intensity is quantized at 6 dB gradients and a display that resembles a contour map results. We generally preferred to use the "normal" display (the contour display seems to show too much information for visual interpretation). The bandwidth of the spectrograph's analyzing filter was set to 300 Hz, and the frequency preemphasis circuits were set at the "Flat" position, since there is more high-frequency energy in the glottal excitation of the gorilla than is the case for human vocalization. This is also the case for rhesus monkey and chimpanzee vocalizations.

The fundamental frequency of phonation, which can be measured from the vertical striations that appear in the wide-band spectrogram, was unstable and ranged from 100 to 120 Hz. Large pitch perturbations, i.e., rapid fluctuations in the glottal periodicity, occurred from one period to the next. The laryngeal output of the gorilla appears to be very noisy and turbulent. Energy concentration can be noted in Fig. 1 at 500, 1500, and 2400 Hz. Measurements of the skull and mandible of an adult gorilla yield an estimated supralaryngeal vocal tract length of 18 cm. If a gorilla uttered the schwa vowel (the first vowel in the word *about*), that is, a vowel having a vocal tract shape that approximates a uniform tube open at one end, we would expect to find formant frequencies at 460, 1400, and 2300 Hz since the resonances of a uniform tube open at one end will occur at intervals of:

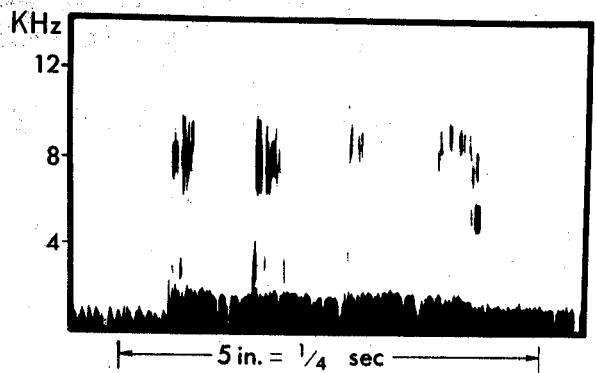
$$\frac{(2k + 1)(C)}{4L}$$

where C = velocity of sound, L = length of the tube, and k is an integer > 0 . We can, therefore, infer that the energy concentrations in this spectrogram reflect the transfer function of the gorilla's supralaryngeal vocal tract in the schwa position. Note that these energy concentrations are not spaced at harmonic multiples of the fundamental frequency.

The main characteristic of this utterance is that the output of the gorilla's larynx is being modified by the resonances of the supralaryngeal vocal tract as is the case for human speech. Note that this is in sharp contrast to the calls of birds, where the fundamental frequency and harmonics of the syrinx's output completely characterize the acoustic nature of the cry (Thorpe, 1961; Greenewalt, 1968).

Note that the bandwidth of the spectrograph's analyzing filter was 300 Hz. Narrow bandwidth analysis would have made it quite difficult to determine the formant frequencies. Narrow bandwidth spectrograms are appropriate for the analysis of bird calls, where the acoustic characteristics of the signal are structured in terms of the fundamental frequency and harmonic structure of the excitation function (the output of the syrinx). They are insufficient, however, when the acoustic characteristics of the signal are determined in part by the transfer function of the supralaryngeal vocal tract's configuration, which acts as an acoustic filter on the excitation function.

The exclusive use of narrow bandwidth spectrograms can lead to descriptions that, although acoustically valid in terms of the narrow bandwidth analysis, are inappropriate in terms of the acoustically and perceptually significant aspects of the signal. Marler and Hamilton (1966), for example, note that, "compared



Rhesus Monkey

Fig. 2. Spectrogram of aggressive "bark" of rhesus monkey, after Lieberman (1968). The bandwidth of the analyzing filter was 1200 Hz. Formant frequencies occurred at 1, 3, and 6 to 8 kHz.

with the calls of birds, many sounds used by primates and other mammals are coarse, lacking the purity of tone and precise patterns of frequency modulation that occur in many passerine bird songs." This statement is true insofar as the primates do not produce cries that can be described in terms of one or two "pure" sinusoidal components. Yet neither can human speech be described in terms of one or two pure tones, "... or precise patterns of frequency modulation..." If the methodology that is appropriate for the analysis of bird calls were used for the analysis of human speech it would be extremely difficult to isolate most of the significant phonologic elements. We would perhaps conclude that human speech employed "coarse" sounds, i.e., sounds that were not inherently musical. The point here is, of course, that the acoustic analysis must be appropriate for the signal. In order to investigate the effects of the supralaryngeal vocal tract we must use analyzing filters that have a bandwidth sufficient to encompass two or more harmonics of the excitation function. This aspect of speech analysis is discussed in detail by Koenig, Dunn, and Lacy (1946).

The sound spectrograph used in this study was manufactured by the Voiceprint Company of New Jersey. Other commercially available spectrographs such as those manufactured by the Kay Electric Company, Pine Brook, New Jersey, would also have been suitable. It is usually not necessary to use the "contour displays" that are available on the Voiceprint machine. When detailed spectral information is necessary, "sections" can be made with either the Voiceprint or Kay Electric machines. It is, however, necessary to maintain adequate bandwidth in the spectrographic analysis if one wishes to determine formant frequencies.

In Fig. 2 a spectrogram of one of the aggressive sounds of a rhesus monkey is presented (Lieberman, 1968). The cry was produced at a moderate degree of vocal effort while the monkey bared his teeth. Six normal monkeys were recorded over a period of 6 months in the monkey colony of the University of Connecticut at Storrs. This particular recording was made with the Sony Type F85 microphone. In Fig. 3 part of the oscillogram of this cry is shown. The first two "bursts" are presented in the oscillogram, which was made as the tape recording was played back at one-fourth speed. The fundamental frequency of phonation is approximately 400 Hz. Note that the fundamental periodicity is very unstable at best. Parts of the waveform appear to be very turbulent. The waveform, in all, looks very much like those associated with pathologic human larynges where a hoarse vocal output results (Lieberman, 1963). The rhesus monkeys, like

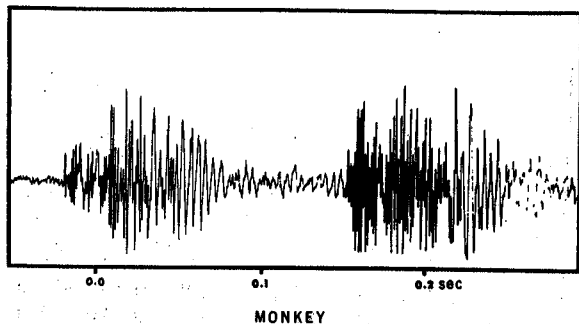


Fig. 3. Oscillogram of same utterance as Fig. 2. The tape recording was played at one-fourth speed. The fundamental frequency of phonation was about 400 Hz. The waveform resembles that characteristic of extremely hoarse human vocalization.

the gorillas and chimpanzees, are unable to produce sustained vocalizations that have a steady fundamental periodicity.

The spectrogram in Fig. 2 was also made from a tape that was played back at one-fourth speed. This procedure increased the effective bandwidth of the spectrograph by a factor of four. The effective bandwidth of the spectrograph was thus 1200 Hz. Energy concentrations occurred at 1, 3, and 6 to 8 kHz. There was approximately 25 msec between each burst and glottal activity seems to have been sustained. Thus the cry is similar to a sequence of voiced stops in intervocalic position.

Unlike voiced stops in human speech, the closure of the vocal tract seems to have been effected by the animal's epiglottis and velum. The monkey's lips were retracted, exposing his teeth throughout the cry, so he could not have used his lips to obstruct his vocal tract. There are also no formant transitions, which would occur if the supralaryngeal vocal tract were momentarily

obstructed by the tongue. The larynx of a rhesus monkey is quite high in contrast to the position of the human vocal tract, and his epiglottis can seal his mouth off at the soft palate (Geist, 1961).

Note that the energy concentrations at 1, 3, and 6 to 8 kHz are again consistent with the resonances of a uniform tube open at one end. We anesthetized a 5-year-old male monkey and measured the length of his supralaryngeal vocal tract. With his lips rounded the length of the supralaryngeal vocal tract was 7.6 cm. The resonances of a uniform 7.6-cm-long tube open at one end are 1100, 3300, and 5500 Hz. We recorded a number of cries that the monkey made with his lips rounded at a low level of vocal effort. The recordings were made in a quiet room using the General Radio 1560-P5 microphone. The average values of F_1 , F_2 , and F_3 were 1300, 3000, and 4400 Hz, respectively. Thus the monkey was producing these cries with a slightly flared supralaryngeal vocal tract.

In Fig. 4 photographs of a casting of the oral cavity of a rhesus monkey are presented. The monkey's tongue and lips were positioned in an approximation of an aggressive "bark" (Rowell & Hinde, 1962), and a plaster-of-paris casting was made shortly after an experiment in which the monkey was sacrificed for other purposes. Note that the vocal tract of the monkey approximates a uniform cross section passage with a flared portion at the laryngeal end. Also note the shallowness of the pharyngeal "bend" and the flatness of the monkey's tongue, which is apparent in the side view. The monkey's tongue fills up the shallow section delimited by the depth of the "bend" at the laryngeal end of the oral cavity.

The nonhuman primates essentially lack a pharyngeal region like man's. In Fig. 5 a schematized view of the pharyngeal and oral regions of the human supralaryngeal vocal tract is presented. Note that the anterior wall of the pharyngeal region is formed by the back of the tongue. The human tongue is thick in comparison with its length. The shape of the pharyngeal region constantly changes during the production of human speech as the tongue moves backwards and forwards. The cross-sectional area of the



Fig. 4. Side and top views of a casting of the oral cavity of an adult rhesus monkey. The monkey's tongue and lips were positioned in an approximation of an aggressive "bark." Note the uniform cross-section of most of the oral cavity.

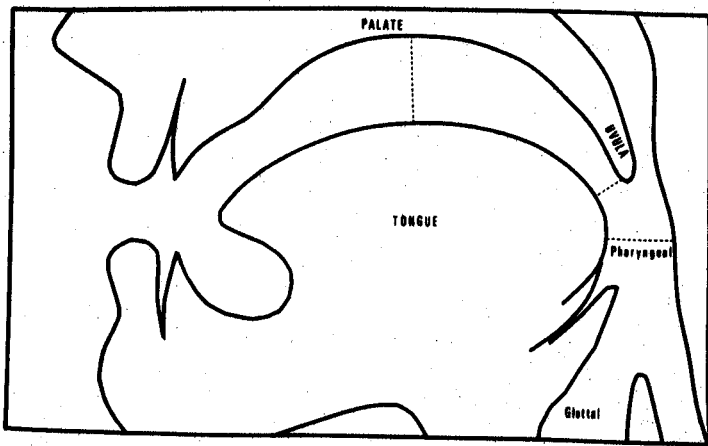


Fig. 5. Schematized view of the human oral and pharyngeal region. Note the relative thickness of the tongue. The anterior wall of the pharynx in man is formed by the tongue, and the cross-sectional area of this back cavity can vary over a 10-to-1 range. A variable pharyngeal region is essential for the production of back vowels and consonants.

pharynx varies, for example, over a ten-to-one range for the vowels /a/ and /i/ (Fant, 1960). The vowel /a/ is produced with a small pharyngeal cross-section while the /i/ is produced with a large cross-section. These variations in pharyngeal cross-sectional area are characteristic for consonants as well as vowels, and they are essential in the production of human speech.

In Fig. 6 a semidiagrammatic representation of the nose, palate, tongue, pharynx, and larynx of a monkey, an ape, and man are reproduced from Negus (1949). Note the relative positions of the palate and larynx. The basis for the nonhuman primates' lack of tongue mobility appears to be anatomical. The pharyngeal region, which can vary its shape in man, has no real counterpart in these animals. Their larynges are positioned quite high compared to the human larynx, almost in line with the roof of the palate. And the tongues of these animals are thin compared to man's. The nonhuman primates do not have a pharynx where the root of a thick tongue forms a movable anterior wall. Zhinkin (1963), for example, in a cineradiographic study of baboon cries, shows that the baboon cannot vary the size of his pharynx. The tongues of the nonhuman primates are long and flat, and their supralaryngeal vocal tracts cannot assume the range of shape changes characteristic of human speech.

COMPUTER-IMPLEMENTED ANALYSIS

The acoustic analysis of primate vocalizations discussed so far was performed limited to the sounds that the animals actually uttered. We used our knowledge of the articulatory basis of human speech to infer that the limitations of these animals' vocal repertoires was anatomical. It is possible to see whether the limitation on the vowel repertoire of a nonhuman primate is actually due to the anatomical constraints imposed by his speech production apparatus (Lieberman, Klatt, & Wilson, 1969). The method used employed a computer-implemented model of the supralaryngeal vocal tract of a rhesus monkey (*Macaca mulatta*) that was systematically manipulated.

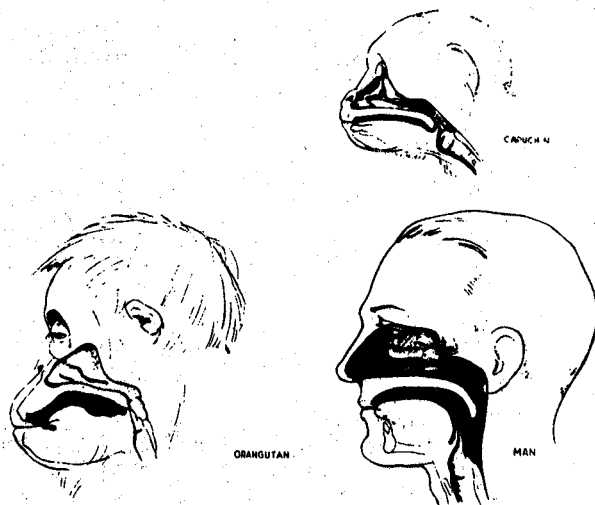
The plaster casting of the oral cavity of a rhesus monkey pictured in Fig. 4 was sectioned at intervals of 0.5 cm and the cross-sectional area was determined by weighing paper tracings of the sections on an analytical balance. This area function is presented as the solid line in Fig. 7.

As noted earlier, the acoustic theory of speech states that the acoustic waveform corresponding to a vowel can be regarded as the output of a vocal tract filter system that is excited by vocal cord vibrations. Temporal effects can be ignored when we differentiate sustained vowels. It is impossible to ignore temporal effects when we consider consonants like the stops /b,d,g/, etc., or diphthongs like /ai/.

The frequency domain transfer function of the vocal tract is

determined by an area function that gives the cross-sectional area of the vocal tract as a function of position along the tract. For frequencies of interest, the vocal tract behaves as a linear solution satisfying the one-dimensional wave equation. A closed-form solution to the wave equation for arbitrary area functions is not known so it is necessary to use an algorithm to find an approximate solution for individual sample area functions.

The algorithm that has been used in the computer program (Henke, 1966) represents the vocal tract by a series of contiguous cylindrical sections, each of fixed area. Each section can be described by a characteristic impedance and a complex propagation constant, both of which are well-known quantities for uniform cylindrical tubes. Junctions between sections satisfy the constraints of continuity of pressure and conservation of volume velocity. The transfer function is calculated directly as a function of frequency. In this fashion the computer program calculated the three lowest formant frequencies. These formant frequencies are presented in Fig. 7.



by Negus

Fig. 6. Semidiagrammatic representation of the nose, palate, tongue, pharynx, and larynx of a monkey and of man from Negus's *Comparative anatomy and physiology of the larynx*. Note the relative position of the palate and larynx in the two diagrams. The monkey lacks a pharyngeal region whose anterior wall can move. The monkey cannot change the configuration of his supralaryngeal vocal tract by means of a thick mobile tongue.

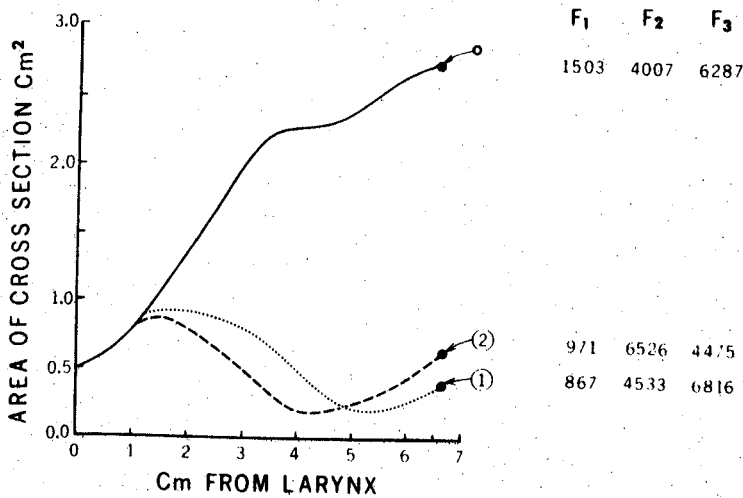


Fig. 7. Area functions of supralaryngeal vocal tract modeled by computer and corresponding vowel formant frequencies. Curve 0 is the unperturbed vocal tract of the rhesus monkey. Curves 1 and 2 are perturbed area functions analogous to human high front vowels. The formant frequencies calculated by the computer are tabulated for each vocal tract configuration (Lieberman, Klatt, & Wilson, 1969).

We systematically explored the possible range of supralaryngeal vocal tract area functions that a rhesus monkey could make by moving his tongue, lips, and jaw. The computer program was used to determine the formant frequencies for each configuration. In Fig. 7 the dashed lines represent vocal tract configurations for two different degrees of tongue height that would most likely lead to the production of an unrounded high vowel. These result from changes in the tract toward that shape of human vocal tract that is characteristic of the production of /i/. We also explored monkey vocal tract configurations that were perturbed in the direction of the human vowels /u/ and /a/. We estimated the range of articulatory maneuvers that are available to a monkey by manipulating the supralaryngeal vocal tract of an anesthetized monkey and by taking into consideration the continuity constraints imposed by the monkey's tongue as well as the effect of different jaw angles and lip rounding. In doubtful cases we allowed greater deviations from the "unperturbed" area function derived from the casting. The computer program calculated the formant frequencies associated with each simulated monkey vocal tract configuration.

In Fig. 8 the first and second formant frequencies of these simulated vocal tract configurations are plotted together with the formant frequencies derived from actual nonhuman primate cries (Lieberman, 1968), and the vowels /a/, /u/, and /i/ measured by Fant (1960) for an adult male human speaker. These three vowels delimit the human "vowel space." We have scaled all the formant frequencies to the length of the rhesus monkey's vocal tract, which was 6.5 cm. Note that the actual monkey and ape cries noted by the letters C (chimpanzee), G (gorilla), and R (rhesus monkey) occupy only part of the vowel space of our computer-generated vowels. The nonhuman primates did not, in fact, use all of the articulatory maneuvers that we simulated for the rhesus monkey on the computer. Note that the computer model further indicates that the possible acoustic vowel space of a monkey is quite restricted compared to the human range. In other words, the vocal apparatus of the rhesus monkey is inherently incapable of producing the range of human speech. The results of the computer simulation (Lieberman, Klatt, & Wilson, 1969) thus are consistent with the analysis of recorded nonhuman primate vocalizations (Lieberman, 1968).

THE PHONETIC CODE, SOME UNANSWERED QUESTIONS

The experiments described herein merely set upper limits on the phonetic code that nonhuman primates might use in their vocal communications. They do not mean that the nonhuman

primates are incapable of communication by means of cries. We have yet to "decode" the communications of the nonhuman primates.

One of the primary characteristics of human language is that the relationship between sound and meaning is arbitrary in language. The difference between a system of cries, even though it may be highly developed, and a language is that the relationship

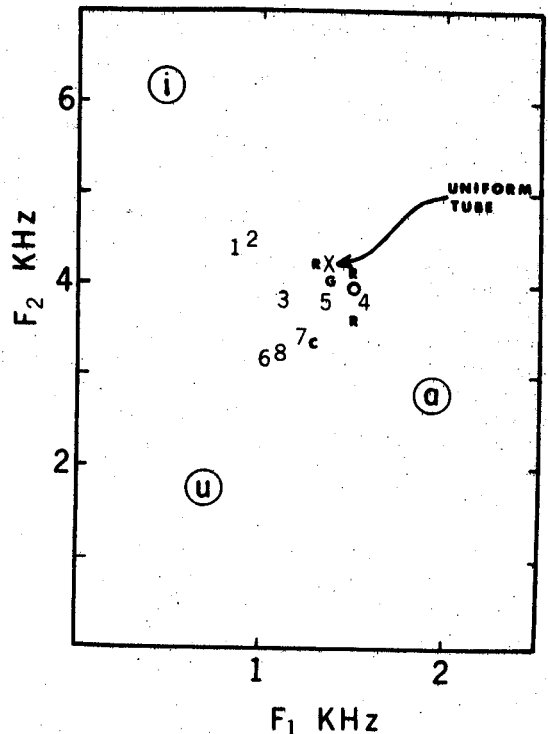


Fig. 8. Vowel "space" of simulated monkey tract relative to human vowel space and measured ape and monkey cries (Lieberman, 1968). Chimpanzee cry noted by Letter C, gorilla by G, and rhesus monkey by R. The formant frequencies that would correspond to a uniform 6.5-cm-length tube terminated at one end are also plotted. All of the formant frequencies have been frequency-scaled towards those of the rhesus monkey to compensate for differences in overall vocal tract length (Lieberman, Klatt, & Wilson, 1969).

between meaning and sound is fixed for cries. A high-pitched /a/, for example, might be the cry of pain for a particular species. It would always "mean" pain no matter what sounds preceded or followed it. In contrast, the sound /a/ in a language may have no meaning in itself, nor might the sounds /m/ and /n/ in isolation. The sound sequence /man/ does have a particular semantic reference or meaning in English while the sound sequences /ma/ and /an/ have other meanings. The sequential coding of sounds in these examples is an essential aspect of linguistic systems.

Most work on animal communication has stressed the temporal ordering of sound sequences (Reynolds, 1968). Human speech is, of course, sequentially coded. But human speech is also a simultaneous code. We have independent control over a number of different "phonologic features." Each feature involves particular maneuvers of man's speech-producing apparatus, and each feature also has its acoustic "correlates." The phonologic features may, in effect, be viewed as matches between the constraints of man's speech-producing apparatus and auditory perception (Lieberman, 1969). The articulatory base of each feature is a maneuver that can readily be executed by man's speech-producing apparatus. The acoustic base of each feature is a signal that can be differentiated and categorized. We apparently "code" and "decode," that is, produce and perceive, speech in terms of these independent phonologic features (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967).

Humans, for example, have independent control over the lips, larynx, and velum (the velum acts as a valve that can connect the nose to the mouth). We can, for example, close our lips or not, adduct our vocal cords or not, open our velum or not, etc. The difference between the sounds /b/ and /p/ in the words *bat* and *pat* is that the vocal cords are adducted when the lips are released in *bat* whereas they are open when the lips are released in *pat*. In a similar way *vat* differs from *mat* with respect to the state of the velum during the first part of the syllable.

We can, therefore, approach the acoustic communications of nonhuman species from at least two independent directions. We can explore the sequential coding of their cries. Birds, for example, do not appear to have control over a number of independent articulatory mechanisms. The time pattern of the fundamental frequency and harmonic content of the syrinx fully specify each bird call. It is therefore appropriate to concentrate on sequential coding in the analysis of these animals' communications systems. The nonhuman primates do not have the ability to produce the full range of human speech. They do, however, have the anatomical ability to control some phonologic features like voicing, nasality, and lip rounding. They have a much greater potential repertoire than do birds.

The question that should be answered is whether any of the

nonhuman primates differentiate their meaningful cries by means of contrast in a simultaneous "feature" code. If apes did communicate by means of cries that were differentiated by phonologic feature contrasts that were a subset of the phonologic features available to man, we would see a link between human language and nonhuman primate behavior. This question, of course, can be resolved only through research that couples acoustic analysis and behavioral techniques. The results should be of interest not only in furthering our knowledge of nonhuman primate behavior but of human linguistic ability and the development of human language.

REFERENCES

- CHIBA, T., & KAJIYAMA, M. *The vowel, its nature and structure*. Tokyo: Phonetic Society of Japan, 1958.
- FANT, C. G. M. *Acoustic theory of speech production*. The Hague: Mouton, 1960.
- GEIST, F. D. Nasal cavity, larynx, mouth and pharynx. In C. G. Hartman (Ed.), *Anatomy of the rhesus monkey*. New York: Hafner, 1961.
- GREENEWALT, C. H. *Bird song: Acoustics and physiology*. Washington, D. C.: Smithsonian, 1968.
- HENKE, W. L. Dynamic articulatory model of speech production using computer simulation. PhD thesis, Department of Electrical Engineering, M.I.T., Cambridge, Mass., 1966. Appendix B.
- KOENIG, W., DUNN, H. K., & LACY, L. Y. The sound spectrograph. *Journal of the Acoustical Society of America*, 1946, 17, 19-49.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- LIBERMAN, P. On the physical bases of phonologic features: Language. In Status Report SR-15, Haskins Laboratories, New York City, 1969, in press.
- LIBERMAN, P. Primate vocalizations and human linguistic ability. *Journal of the Acoustical Society of America*, 1968, 44, 1574-1584.
- LIBERMAN, P. Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *Journal of the Acoustical Society of America*, 1963, 35, 344-353.
- LIBERMAN, P., KLATT, D. L., & WILSON, W. A. Vocal tract limitations on the acoustic repertoires of rhesus monkey and other non-human primates. In Status Report SR-15, Haskins Laboratories, New York City, 1969, in press.
- MARLER, P. F., & HAMILTON, W. J. *Mechanisms of animal behavior*. New York: Wiley, 1966. P. 470.
- NEGUS, V. E. *The comparative anatomy and physiology of the larynx*. London: Hafner, 1949.
- REYNOLDS, P. C. Evolution of primate vocal-auditory communication systems. *American Anthropologist*, 1968, 70, 300-308.
- ROWELL, T. E., & HINDE, R. A. Vocal communication by the rhesus monkey (*Macaca mulatta*). *Proceedings of the Zoological Society of London*, 1962, 138, 279-294.
- THORPE, W. H. *Bird song*. Cambridge: Cambridge University Press, 1961.
- ZHINKIN, N. I. An application of the theory of algorithms to the study of animal speech—methods of vocal intercommunication between monkeys. In R. G. Busnel (Ed.), *Acoustic behavior of animals*. Amsterdam: Elsevier, 1963.