Anim. Behav., 1993, **45**, 423–433

The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations

M. D. HAUSER*, C. S. EVANS & P. MARLER
Animal Communication Laboratory, University of California, Davis, CA 95616-8761, U.S.A.

(Received 25 October 1991; initial acceptance 20 January 1992; final acceptance 19 March 1992; MS. number: A6201)

Abstract. Many non-human primates have large and acoustically diverse vocal repertoires. However, our understanding of the phonatory mechanisms underlying such acoustic variation is poor. Representative exemplars of some of the articulatory gestures used by free-ranging rhesus monkeys while uttering a range of vocalizations are provided. Results reveal that different call types appear to be associated with characteristic lip configurations and mandibular positions. Quantitative analyses of the 'coo' vocalization indicate that changes in the position of the mandible are reliably associated with changes in dominant frequency (i.e. resonance frequency), but not with changes in fundamental frequency. This finding suggests that rhesus monkeys can modify the spectral properties of the signal, independent of the glottal source (i.e. fundamental frequency). Such articulatory manoeuvres contribute to the animal's potential acoustic space, thereby potentially increasing the array of meaningful vocalizations within the repertoire.

Variation in the morphology of acoustic signals is the product of a number of factors, including changes in respiratory activity (Suthers et al. 1972; Jurgens & Schriever 1991; Hauser & Fowler 1992) and vocal tract configuration (Capranica & Moffat 1983; Nowicki 1987; Suthers et al. 1988; Suthers 1990; Ryan & Drewes 1990). In addition, the structure of a vocalization is likely to be influenced by changes in motivational state (Morton 1977; Smith 1977; Goedeking 1988; Marler 1992).

In this paper, we examine the role of articulatory gestures in the production of acoustic signals by a non-human primate. Understanding how articulation influences sound structure is important because it allows for a more precise assessment of the degree to which animals are capable of controlling call morphology. If acoustic variation is reliably associated with particular production contexts or affective states, then senders have the potential to produce signals with a broad array of meanings for conspecific receivers (Cheney & Seyfarth 1990; Marler et al. 1992). However, relatively little is known about the mechanisms underlying vocal production in non-human primates (but see Bauer 1987; Owren & Bernacki 1988; Hauser 1992; Hauser & Fowler 1992; reviewed in Hauser 1992a). The paucity of information on non-human primate articulation is particularly striking when one considers the comparatively more extensive data for other taxonomic groups such as birds (e.g. Nowicki & Marler 1988; Suthers 1990; Hausberger et al. 1991), bats (e.g. Suthers et al. 1988), cetaceans (e.g. Purves & Pilleri 1983; Amundsen 1991) and especially humans (e.g. Baken 1987; Borden & Harris 1984). We first provide a brief review of some of the articulatory models of sound production, together with a discussion of some relevant empirical work. We then describe the rhesus monkey's vocal repertoire and illustrate the variation in articulatory gestures used. Finally, we present the results of a quantitative analysis of the role of mandibular position in the production of rhesus monkey 'coo' vocalizations. Changes in mandibular position are known to play an important role in human speech (e.g. Lindblom et al. 1979) and in the isolation calls of captive cats (Shipley et al. 1991).

Source-filter Theory and Comparative Vocal Tract Anatomy

The source-filter theory of speech production provides a model for predicting the acoustic morphology of vowels on the basis of vocal tract

---

*Present address: Department of Biological Anthropology, Harvard University, Cambridge, MA 02138, U.S.A.

0003–3472/93/030423+11 $08.00/0

configuration (Fant 1960) and has received considerable empirical support (reviewed by Borden & Harris 1984). This model postulates that the glottal waveform (i.e. fundamental frequency: Fo), or source function, is filtered as a result of the vocal tract frequency response, or transfer function (Fant 1960). An explicit assumption of such models is that source and filter are independent.

To understand the relative contribution of different articulators (e.g. tongue, lips and jaw) to sound production, studies of human speech have pursued two approaches. Some investigators have examined the correlation between changes in the position of a given articulator and the simultaneous change in the spectral properties of the signal. During the production of vowels, for example, changes in lip configuration (i.e. opening and protrusion) are closely associated with changes in the resonance frequencies of the signal (Stevens & House 1955). These spectral differences are largely responsible for the perceptual distinctiveness of different vowel sounds (Chiba & Kajiyama 1941).

A second approach involves direct perturbation of the articulators involved in speech production, thereby allowing a more direct assessment of the role of articulation in modifying acoustic structure. For example, the role of mandibular movements has been explored through the use of a bite block technique (e.g. Lindblom et al. 1979). Here, a speaker is asked to bite down on a block of hard plaster, thereby fixing jaw position during sound production. A small number of perturbation experiments have also been carried out with non-primate animals (e.g. birds: Nowicki 1987; porpoise: Amundsen 1991).

Articulation in Non-human Primate Communication

Owren & Bernacki (1988) examined the acoustic structure of vervet monkey, *Cercopithecus aethiops*, eagle and snake alarm calls to determine whether articulatory changes were responsible for spectral changes in the call. Using an acoustic technique known as 'linear predictive coding', they demonstrated that there were distinct resonance frequencies within each call type which were spectrally independent of the fundamental frequency. On the basis of these results, Owren & Bernacki (1988) argued that vervet monkeys appear to modify vocal tract configuration in order to alter the spectral properties of the signal. Moreover, such spectral differences are perceptually salient to the vervets and appear to be used for recognizing different call types (Owren 1990a, b).

Thus far, only Bauer (1987) has provided an explicit investigation of the association between articulatory changes and changes in call structure. Using 16-mm movie footage collected by P. Marler (see Marler & Hobbett 1975; Marler & Tenaza 1977), Bauer explored the relationship between lip configuration and fundamental frequency (Fo) in a male chimpanzee, *Pantroglodytes*, producing a single sequence of vocalizations, which began with a series of 'screams' given submissively and ended with 'waahbarks' given aggressively. Bauer's results showed that, during the progression from submissive screams to aggressive waahbarks, there was a significant decline in fundamental frequency. Moreover, screams were associated with retracted lips and waahbarks were associated with protruded lips. Together, these results suggest that, in chimpanzees, lip configuration is closely associated with changes in fundamental frequency.

METHODS

Subjects and Study Area

Observations were collected from November 1988 to June 1990 on adult female and male rhesus macaques living in one of the seven social groups (group L) on the island of Cayo Santiago, Puerto Rico. Cayo Santiago is a 15-ha island off the east coast of Puerto Rico (for a detailed description of the island and its history, see Rawlins & Kessler 1987). There were approximately 1200 individuals present during the study and all individuals were habituated to the presence of human observers, thereby facilitating observations of vocal behaviour at close range. Group L was the largest (300–350 individuals) and most dominant social group on the island.

The rhesus monkeys on Cayo Santiago are provisioned, with approximately 50% of their diet being comprised of Purina monkey chow. In addition to chow, they feed on fruit (e.g. coconut), leaves, insects, and soil (for further details, see Hauser 1991, 1992; Hauser & Fowler 1992). There are no predators on the island. The primary sources of mortality are starvation and injury (Berard 1990).

Audio-video Analyses

Audio-video records of the faces of vocalizing monkeys (Fig. 1) were obtained under natural conditions with a high resolution camcorder (>400 lines); subject-to-camera distances were 0·5–2·0 m. The clarity of the video image and associated sound recording dictated the selection of exemplars for both qualitative and quantitative analyses; consequently, one call type, the coo, is over-represented in the data, whereas other call types are not represented. Spectrograms were generated using an Apple Macintosh-based hardware and software package (MacRecorder, Farallon Computing, Inc.) and were based upon overlapping 256-point Fourier transforms (frequency resolution 86 Hz).

Quantitative data are presented on the relationship between mandibular position and the acoustic structure of coos in rhesus macaques. Coos were selected for analyses because a large sample of high quality recordings were obtained and detailed acoustic analyses of call structure have been published elsewhere (Hauser 1991, 1992). Coos examined in the current report were produced by individuals discovering food, including chow and other less commonly encountered items such as coconut; only sequences permitting frame-by-frame measurements of mandibular position and unambiguous measurements of acoustic structure (see below) were used. Although coos are given in non-food contexts (e.g. mother–infant contact and group progression), the acoustic variation observed within coos is not statistically associated with contextual variation but rather, the identity of the caller (Hauser 1991, 1992).

Acoustic analyses were performed using 'SIGNAL', a digital sound analysis program (Beeman 1989) which runs on a 80386-based microcomputer; this system has been previously used for the analysis of rhesus monkey vocalizations (Hauser 1991; Hauser & Fowler 1992). Coos were sampled at 25 kHz (16 bits) and spectrograms were derived from 500 overlapping 256-point Fourier transforms (Hanning window). Power spectra, based on 256-point transforms (frequency resolution 98 Hz), were used to obtain the fundamental frequency and the frequency of the peak with the strongest amplitude (dominant frequency). On average, the dominant frequency had an amplitude 6·4 dB (SD = 1·8, $N=42$) greater than that of the fundamental frequency.

Mandibular position was measured as the vertical distance from the midpoint of the mandible to the midpoint between the nostrils. Measurements were obtained by superimposing computer-generated cross-hairs over the still image of successive video frames, displayed using an editing system and a high resolution monitor. Individual video frames were numbered using a commercial animation programme. An oscillograph of the amplitude waveform was displayed in a 63 × 100 pixel window in order to detect sound onset and offset (Fig. 1h). Distances on the screen were calculated in pixels (resolution = 320 × 200) and standardized to control for head rotation and subject-to-camera distance by dividing each measurement by the subject's nostril diameter. Measurements of five coos produced by an adult female whose head position and distance to the camera did not change during call production indicated that there was little to no variation in nostril diameter during vocal production ($\bar{X} \pm \text{SD} = 16{\cdot}7 \pm 0{\cdot}37$ pixels, $N=167$ frames). To obtain acoustic measurements corresponding to a particular mandibular position, Fourier transforms were performed every 33 ms, corresponding to the interval between video frames. Because the duration of coos was variable (Hauser 1991), the number of frame intervals analysed per call varied, ranging from a minimum of eight to a maximum of 22. Analyses of articulatory gestures are based upon measurements from a total of 1059 video frames and include the period just prior to sound production. Calls occurred during 615 of these frames, yielding a total of 615 measurements of the fundamental and of the dominant frequency.

RESULTS

The Relationship Between Articulation and Call Structure: a Qualitative Assessment

Figure 1 presents spectrograms and corresponding articulatory configurations, taken at the midpoint of a call, for eight vocalizations in the rhesus monkey's repertoire; description of each call type was based on recordings from a minimum of three individuals and a maximum of eight. 'Barks' (Fig. 1a) and 'pant-threats' (Fig. 1b) are given by aggressive individuals; barks, in contrast to pant-threats are more likely to be associated with physical aggression. For both calls, the lips are slightly protruded and maximally spread apart at sound onset. Consequently, the mandible is also at its

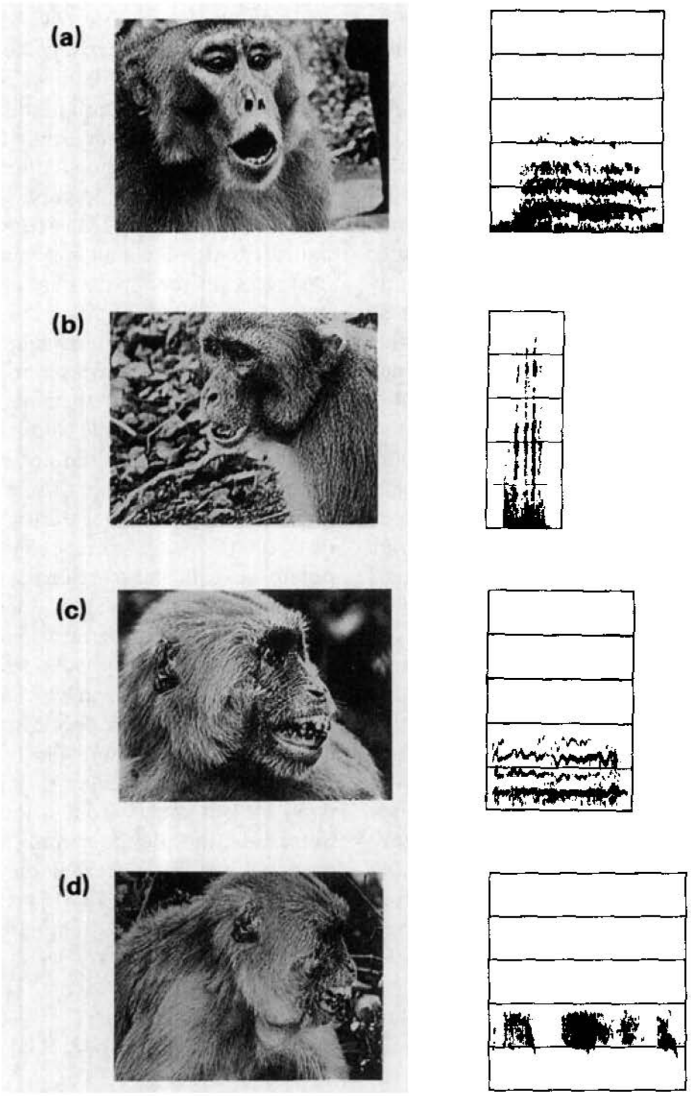


Figure 1.

lowest point when the sound is first uttered. During the production of barks, there is greater separation between the lips and the mandible drops to a lower point than for pant-threats.

'Undulating' and 'noisy' screams (Fig. 1c, d) are produced by individuals being threatened by a dominant member of the group (Gouzoules et al. 1984). Noisy screams are more consistently associated with physical attack than undulating screams. The lips are retracted during both types of screams, thereby revealing the teeth. Lip retraction appears to be more prominent during the production of noisy screams. For both scream types, the teeth are partially separated.

'Copulation screams' (Fig. 1e) are only given by reproductively mature males during copulation (i.e. thrusting with or without ejaculation). The lips are retracted as during production of noisy or undulating screams, but the teeth are clamped shut.

'Grunts', 'girneys' and 'coos' (Fig. 1f, g, h) are all given during affiliative interactions. For grunts, the configuration of the mouth is similar to that used during production of pant-threats, but there is no lip protrusion. In addition, grunts, but not

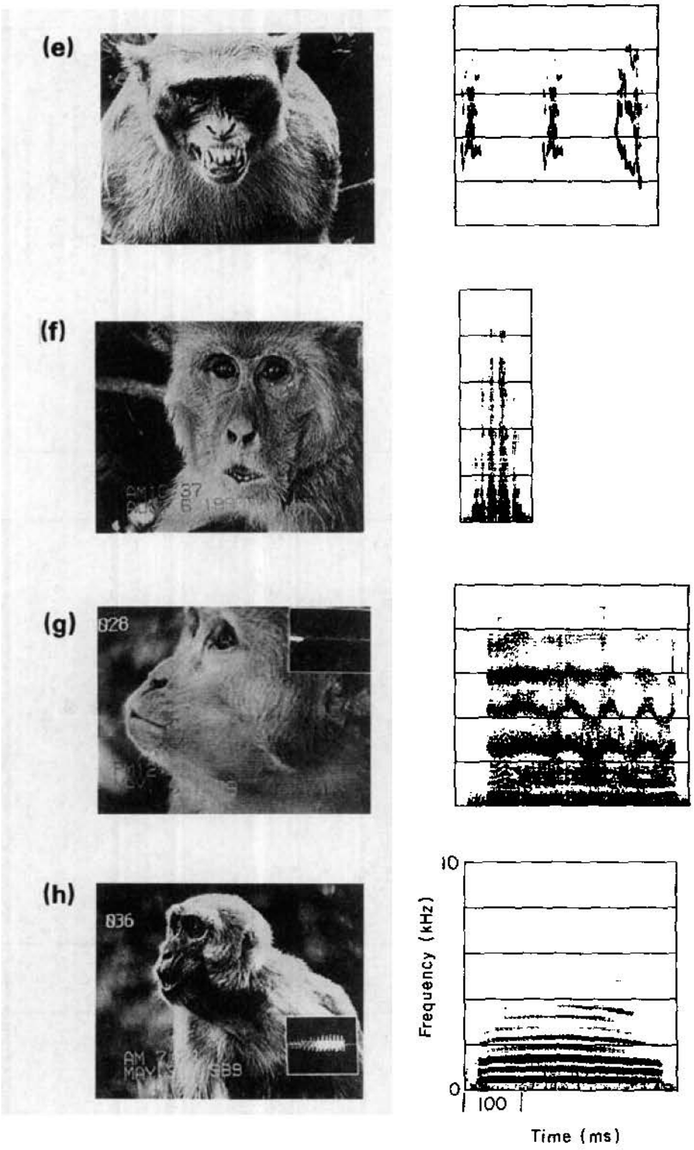
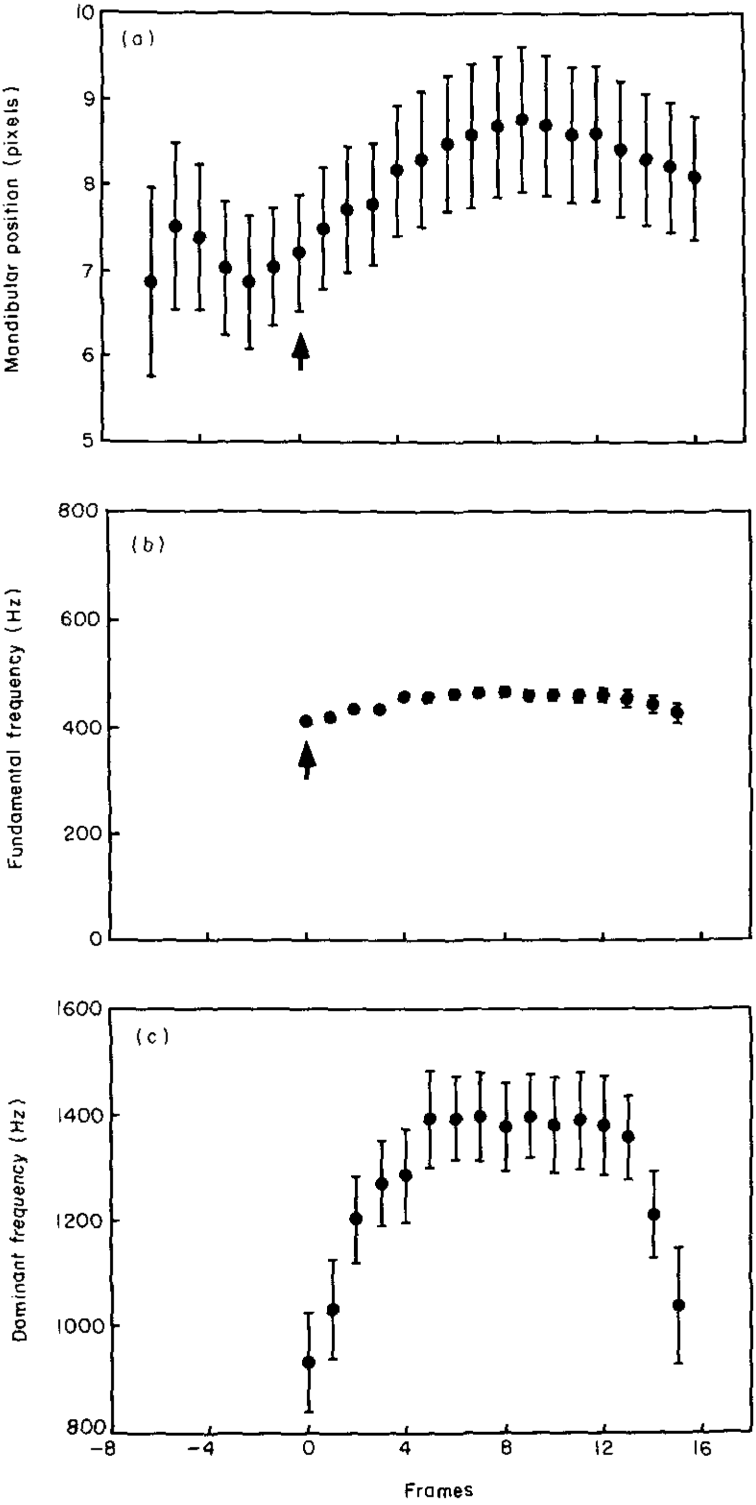


Figure 1. Single video frames of the articulatory gestures used by rhesus monkeys during call production, together with representative spectrograms; each gesture occurred at the mid-point in the call. (a) Bark, (b) pant-threat, (c) undulating scream, (d) noisy scream, (e) copulation scream, (f) grunt, (g) girney, (h) coo. The windows in panels (g) and (h) are oscillographs of the time-amplitude waveform for the articulatory gesture shown; the number in the upper left corner identifies the video frame (see text for further details). The X-axis on the spectrograms is time in ms and the Y-axis is frequency in kHz.

pant-threats, are often produced with negligible lip separation. The spectral and temporal features of grunts and pant-threats are similar. Girneys are produced with what appears to be maximum lip protrusion, minimum lip separation and negligible jaw movement. During sound production, the lips are moved rapidly, as if the animal is chewing. In contrast, coos are produced with a fairly wide separation between the lips and lowered mandible. The extent of lip protrusion appears to be less than for girneys, but greater than for barks and pant-threats.


(a)
Mandibular position (pixels)
10
9
8
7
6
5
(b)
Fundamental frequency (Hz)
800
600
400
200
0
(c)
Dominant frequency (Hz)
1600
1400
1200
1000
800
−8
−4
0
4
8
12
16
Frames

The Relationship Between Mandibular Position and Acoustic Morphology: a Quantitative Analysis of 'Coos'

A sample of 30 coos produced by one adult female was analysed (Fig. 2). Measures of mandibular position (Fig. 2a) revealed that, following sound onset, the mandible gradually dropped, reaching its lowest position toward the middle of the call. As the call was completed, the mandible was elevated. Measures of spectral change, synchronous with individual video frames, revealed that the fundamental frequency (Fig. 2b) exhibited little change, whereas the dominant frequency (Fig. 2c), on average, was characterized by a symmetrical rise to, and then fall from, a frequency maximum. As expected from inspection of Figs 2a, c, changes in mandibular position were associated with changes in dominant frequency (Fig. 3). Although there are no satisfactory statistical tests available for analysing data from one individual, a regression of dominant frequency on mandibular position was significant ($r=0{\cdot}84$, $r^2=0{\cdot}71$, $F_{1,14}=34{\cdot}68$, $P<0{\cdot}0001$).

The relationship between mandibular position and dominant frequency was also examined for three other individuals (two adult males and one adult female). In every case, the fundamental frequency was relatively constant over the course of the call and statistically unrelated ($P>0{\cdot}05$) to changes in the dominant frequency. Results from regression analyses for these three subjects are presented in Table I. These data were consistent with the patterns of change shown in Figs 2 and 3, suggesting that changes in mandibular position accounted for a significant proportion of the variation in dominant frequency. This statistical relationship is all the more striking in light of the variability of coo duration. Consequently, by calculating the average mandibular position for a given frame, some values were taken from the midpoint of a call, whereas others were taken from the endpoint.

To examine more precisely the relationship between mandibular position and dominant frequency, it is necessary to assess each call separately (Table II). Results indicate that 34 out of 42 calls showed a statistically significant relationship between mandibular positions and dominant frequency. Thus, moment-to-moment changes in mandibular position during the production of coos were reliably associated with changes in the dominant frequency.

DISCUSSION

Comparative anatomical work has revealed that there are significant differences between humans and non-human primates in the structure of the vocal tract (Negus 1949; Crelin 1987). Specifically, non-human primates, like human neonates, lack an oropharyngeal chamber and a thick, curved tongue (Negus 1949). As a result, it has been suggested that only simple modifications of the cross-sectional area of the vocal tract, and up-and-down shifts in the position of the larynx can take place (Lieberman et al. 1969; Lieberman 1984, 1990).

Adult humans, in contrast, possess an oropharyngeal chamber and are capable of making substantial changes in the cross-sectional area of their vocal tract, thereby allowing production of a rich array of sounds (Fant 1960; Lieberman 1984). These taxonomic differences form the basis for the argument that non-human primates are incapable of producing the full range of sounds characteristic of human speech (Lieberman 1984) and thus, are fundamentally limited in the number of meaningful utterances that they can generate.

Rhesus monkeys on Cayo Santiago produce a large vocal repertoire. The present study suggests that some of this acoustic variation may be attributable to changes in articulation. Specifically, a preliminary inspection of the repertoire reveals that rhesus monkeys modify lip protrusion, lip separation, teeth separation and mandibular position,

Figure 2. Analyses of mandibular position, fundamental frequency and dominant frequency for the coos of adult female 480 ($N=30$ exemplars). (a) Changes in mandibular position (standard units) over time (frames); note that because mandibular position was measured from the midpoint of the nose to the midpoint of the mandible (see text), an increase in standard units of the Y-axis corresponds to a decrease or drop in mandibular position; (b) changes in the fundamental frequency (Hz) over time; (c) changes in the dominant frequency (Hz) over time. Note that the minimum and maximum frequency values in (b) and (c) are different but that the frequency range is 800 Hz in both cases. Each video frame has a duration of 33 ms. Sound onset is at frame 0, indicated in (a) and (b) by the arrow. Values are means (±SE).

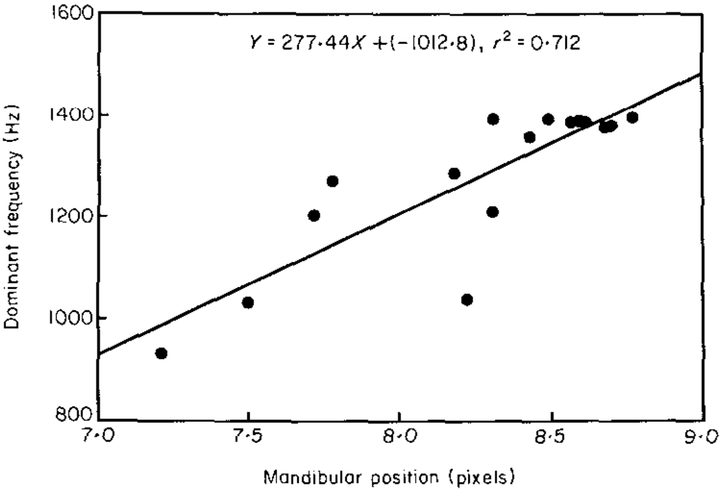


Figure 3. Regression of dominant frequency (Hz) on mandibular position (standard units) for the coos of female 480.

Table I. Regression of mandibular position on dominant frequency for coos

ID	Sex	r	r^2	df*	Number of calls	F	P
845	F	0·87	0·75	1,7	4	21·10	0·003
F24	M	0·80	0·64	1,9	6	16·23	0·003
491	M	0·82	0·67	1,14	2	18·44	0·002

*Represents the number of intervals analysed per call.

and that such articulatory manoeuvres are reliably associated with different call types. In addition to such between-class variation, a quantitative analysis of the 'coo' vocalization suggests a role for articulatory gestures in controlling acoustic morphology within a call type. During the production of coos, changes in mandibular position are statistically associated with changes in the dominant frequency of the call, but not with changes in the fundamental frequency. This result implies that rhesus monkeys have the ability to modify the spectral characteristics of the signal independently of the laryngeal source. A crucial first step in evaluating the social significance of this acoustic variation will be experiments designed to assess the discriminability of coo variants (e.g. Moody et al. 1990).

Mandibular position is unlikely to be the only relevant articulator in non-human primate sound production. Our preliminary sampling of the repertoire indicates that rhesus monkeys modify lip configuration while calling, but we do not yet understand how such changes influence call structure. It is possible that lip configuration initially evolved as a visual display and was then subsequently co-opted for the purpose of acoustic signalling (Ohala 1983, 1984). For example, in most non-human primates, individuals who are threatened retract their lips, producing the characteristic 'fear grimace' (Andrew 1962). In contrast, aggressive individuals often protrude and spread their lips apart. Although lip configuration may have been primarily selected for visual signalling, it none the less will have some influence on the sound generated. Such dominance-related displays should therefore be viewed as multi-modal signals.

In addition to changes in lip configuration and mandibular position, non-human primates may also be capable of modifying the position of the tongue. In human speech, changes in tongue position cause significant changes in the cross-sectional area of the vocal tract (Fant 1960) and consequently contribute directly to the spectral structure of the utterance produced. Thus far, there is no direct evidence from any non-human animal that changes in tongue position play a role

Table II. Correlations between mandibular position and dominant frequency by individual and call

ID	Call number	*r*	*N**	*P*	ID	Call number	*r*	*N**	*P*
480	1	0·97	13	0·001	480	22	0·91	17	0·001
480	2	0·88	12	0·001	480	23	0·92	12	0·001
480	3	0·72	15	0·01	480	24	0·82	16	0·001
480	4	0·84	16	0·001	480	25	0·82	12	0·001
480	5	0·27	13	NS	480	26	0·78	15	0·001
480	6	0·51	19	0·05	480	27	0·76	15	0·001
480	7	0·38	16	NS	480	28	0·83	13	0·001
480	8	0·33	13	NS	480	29	0·28	19	NS
480	9	0·30	15	NS	480	30	0·47	22	0·05
480	10	0·62	18	0·001	845	1	0·83	10	0·01
480	11	0·28	16	NS	845	2	0·88	10	0·001
480	12	0·87	13	0·001	845	3	0·73	15	0·01
480	13	0·70	13	0·02	845	4	0·84	12	0·001
480	14	0·78	16	0·001	F24	1	0·96	15	0·001
480	15	0·78	11	0·01	F24	2	0·73	13	0·01
480	16	0·48	11	NS	F24	3	0·97	11	0·001
480	17	0·15	10	NS	F24	4	0·87	8	0·01
480	18	0·49	18	0·05	F24	5	0·69	20	0·001
480	19	0·92	15	0·001	F24	6	0·86	19	0·001
480	20	0·95	12	0·001	491	1	0·80	16	0·001
480	21	0·93	14	0·001	491	2	0·84	14	0·001

*Represents the number of frames (33 ms intervals) analysed per call.

in sound production. However, observations of vervet monkeys ('wrr' voalization; Hauser 1989), rhesus monkeys ('shrill bark'; Hauser, personal observation), and blue monkeys (*Cercopithecus mitis*; 'trill' vocalization; C. Brown, personal observation) suggest that there is tongue movement during the production of 'trilled' vocalizations. We are currently investigating this possibility further.

In conclusion, our results suggest that rhesus monkeys, like other non-human animals, such as birds, cats and bats (Nowicki & Marler 1988; Suthers et al. 1988; Shipley et al. 1991), may use subtle changes in articulation, both to create an array of call types and to manipulate the fine structure of signals within a class of vocalizations. It appears, therefore, that rhesus monkeys, and most likely other non-human primates as well, can produce a larger set of meaningful utterances than has previously been claimed (e.g. Lieberman 1984). Future work will focus on the contribution of other articulators to sound production and on the extent to which anatomical characteristics of the vocal tract impose significant constraints on repertoire size.

ACKNOWLEDGMENTS

We thank P. Ladefoged and M. Lindau for discussion of the data and, H. Bauer, C. Bickley, U. Jurgens, J. Ohala, M. Studdert-Kennedy and two anonymous referees for comments on earlier versions of the manuscript. Funding for this research was provided by grants from the NIMH (No. MH14651) to P.M., and the National Geographic Society (No. 4251-90) to M.D.H. and P.M. Preparation of this manuscript was supported by post-doctoral fellowships from the National Institute of Health to M.D.H. (HD-07213) and C.S.E. (MH-09683).

REFERENCES

Amundsen, M. 1991. Helium effects on the click frequency spectrum of the Harbor porpoise, *Phocoena phocoena*. *J. acoust. Soc. Am.*, **90**, 53–59.

Andrew, R. J. 1962. The origin and evolution of the calls and facial expressions of the primates. *Behaviour*, **20**, 1–109.

Baken, R. 1987. *Clinical Measurements of Speech and Voice*. Boston: College Hill Press.

Bauer, H. 1987. Frequency code: Orofacial correlates of fundamental frequency. *Phonetica*, **44,** 173–191.

Beeman, K. 1989. *SIGNAL User's Guide*. Belmont, Massachusetts: Engineering Design.

Berard, J. 1990. Life histories of rhesus macaque males. Ph.D. thesis, University of Oregon.

Borden, G. J. & Harris, K. S. 1984. *Speech Science Primer*. Baltimore: Williams & Wilkins.

Capranica, R. R. & Moffat, A. J. M. 1983. Neurobehavioral correlates of sound communication in anurans. In: *Advances in Vertebrate Neuroethology* (Ed. by J. P. Ewert, R. R. Capranica & D. J. Ingle), pp. 701–730. New York: Plenum Press.

Cheney, D. L. & Seyfarth, R. M. 1990. *How Monkeys See the World.* Chicago: Chicago University Press.

Chiba, T. & Kajiyama, M. 1941. *The Vowel: Its Nature and Structure.* Tokyo: Kaiseikan.

Crelin, E. 1987. *The Human Vocal Tract*. New York: Vantage Press.

Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague, The Netherlands: Mouton.

Goedeking, P. 1988. Vocal play behavior in cotton-top tamarins. In: *Primate Vocal Communication* (Ed. by D. Todt, P. Goedeking & D. Symmes), pp. 133–144. Berlin: Springer-Verlag.

Gouzoules, H., Gouzoules, S. & Marler, P. 1984. Rhesus monkey, *Macaca mulatta*, screams: representational signalling in the recruitment of agonistic acid. *Anim. Behav.*, **32,** 182–193.

Hausberger, M., Black, J. M. & Richard, J.-P. 1991. Bill opening and sound spectrum in barnacle goose loud calls: individuals with wide mouths have higher pitched voices. *Anim. Behav.*, **42,** 319–322.

Hauser, M. D. 1989. Ontogenetic changes in the comprehension and production of vervet monkey (*Cercopithecus aethiops*) vocalizations. *J. comp. Psychol.*, **103,** 149–158.

Hauser, M. D. 1991. Sources of acoustic variation in rhesus macaque vocalizations. *Ethology*, **89,** 29–46.

Hauser, M. D. 1992. Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: a learned mode of production? *J. acoust. Soc. Am.*, **91,** 2175–2179.

Hauser, M. D. In press. Non-human primate vocal communication. In: *Handbook of Acoustics* (Ed. by M. Crocker), New York: John Wiley.

Hauser, M. D. & Fowler, C. A. 1992. Declination in fundamental frequency is not unique to human speech: evidence from non-human primates. *J. acoust. Soc. Am.*, **91,** 363–369.

Jurgens, U. & Schriever, S. 1991. Respiratory muscle activity during vocalization in the squirrel monkey. *Folia Primatol.*, **56,** 121–132.

Lieberman, P. 1984. *The Biology and Evolution of Language*. Cambridge, Massachussetts: Harvard University Press.

Lieberman, P. 1990. *Uniquely Human*. Cambridge, Massachusetts: Harvard University Press.

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

Lindblom, B. E. F., Lubker, J. & Gay, T. 1979. Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive stimulation. *J. Phonetics*, **7,** 147–161.

Marler, P. 1992. Functions of arousal and emotion in primate communication: A semiotic approach. In: *Proceedings of the International Primatological Congress, Japan* (Ed. by T. Nishida), pp. 225–234. Tokyo: Tokyo University Press.

Marler, P., Evans, C. S. & Hauser, M. D. 1992. Animal vocal signals: Referential, motivational, or both? In: *Nonverbal Vocal Communication* (Ed. by H. Papousek, U. Jürgens & M. Papousek), pp. 66–86. Cambridge: Cambridge University Press.

Marler, P. & Hobbett, L. 1975. Individuality in a long-range vocalization of wild chimpanzees. *Z. Tierpsychol.*, **38,** 97–109.

Marler, P. & Tenaza, R. 1977. Signalling behaviour of apes with special reference to vocalizations. In: *How Animals Communicate* (Ed. by T. A. Seboek), pp. 965–1033. Bloomington: Indiana University Press.

Moody, D. B., Stebbins, W. C. & May, B. J. 1990. Auditory perception of communication signals by Japanese monkeys. In: *Comparative Perception, Volume II: Complex Signals* (Ed. by W. C. Stebbins & M. A. Berkley), pp. 311–344. New York: John Wiley.

Morton, E. S. 1977. On the occurrence and significance of motivational-structural rules in some bird and mammal sounds. *Am. Nat.*, **111,** 855–869.

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

Nowicki, S. 1987. Vocal tract resonance in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature, Lond.*, **325,** 53–55.

Nowicki, S. & Marler, P. 1988. How do birds sing? *Music Percep.*, **5,** 391–426.

Ohala, J. J. 1983. Cross-language use of pitch: an ethological view. *Phonetica*, **40,** 1–18.

Ohala, J. J. 1984. An ethological perspective on common cross-language utilization of Fo of voice. *Phonetica*, **41,** 1–16.

Owren, M. J. 1990a. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans: I. Natural calls. *J. comp. Psychol.*, **104,** 20–28.

Owren, M. J. 1990b. Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans: II. Synthetic calls. *J. comp. Psychol.*, **104,** 29–40.

Owren, M. J. & Bernacki, R. 1988. The acoustic features of vervet monkey (*Cercopithecus aethiops*) alarm calls. *J. acoust. Soc. Am.*, **83,** 1927–1935.

Purves, P. E. & Pilleri, G. E. 1983. *Echolocation in Whales and Dolphins*. London: Academic Press.

Rawlins, R. & Kessler, M. 1987. *The Cayo Santiago Macaques*. New York: SUNY University Press.

Ryan, M. J. & Drewes, R. C. 1990. Vocal morphology of the *Physalaemus pustulosus* species group (Leptodactylidae): morphological response to sexual selection for complex calls. *Biol. J. Linn. Soc.*, **40,** 37–52.

Shipley, C., Carterette, E. C. & Buchwald, J. S. 1991. The effects of articulation on the acoustical structure of feline vocalizations. *J. acoust. Soc. Am.*, **89,** 902–908.

Smith, W. J. 1977. *The Behavior of Communicating: An Ethological Approach.* Cambridge, Massachussetts: Harvard University Press.

Stevens, K. & House, A. S. 1955. Development of a quantitative description of vowel articulation. *J. acoust. Soc. Am.*, **27,** 484–493.

Suthers, R. A. 1990. Contributions to birdsong from the left and right sides of the intact syrinx. *Nature, Lond.*, **347,** 473–477.

Suthers, R. A., Hartley, D. J. & Wenstrup, J. J. 1988. The acoustic role of tracheal chambers and nasal cavities in the production of sonar pulses by the horseshoe bat, *Rhinolophus hildebrandti. J. comp. Physiol., A*, **162,** 799–813.

Suthers, R. A., Thomas, S. P. & Suthers, B. J. 1972. Respiration, wingbeat and pulse emission in an echolocating bat. *J. exp. Biol.*, **56,** 37–48.