American Journal of Primatology 33:31–50 (1994)

The Acoustic Features of Gorilla Double Grunts and Their Relation to Behavior

ROBERT M. SEYFARTH[1], DOROTHY L. CHENEY[2], ALEXANDER H. HARCOURT[3], AND KELLY J. STEWART[3]
[1]*Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania*
[2]*Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania*
[3]*Department of Anthropology, University of California, Davis, California*

Mountain gorillas (*Gorilla gorilla beringei*) give double-grunts to one another in a variety of situations, when feeding, resting, moving, or engaged in other kinds of social behavior. Some double-grunts elicit double-grunts in reply whereas others do not. Double-grunts are individually distinctive, and high-ranking animals give double-grunts at higher rates than others. There was no evidence, however, that the probability of eliciting a reply depended upon either the animals' behavior at the time a call was given or the social relationship between caller and respondent. The probability of eliciting a reply could be predicted from a double-grunt's acoustic features. Gorillas apparently produce at least two acoustically different subtypes of double-grunts, each of which conveys different information. Double-grunts with a low second formant (typically < 1600 Hz) are given by animals after a period of silence and frequently elicit vocal replies. Double-grunts with a high second formant (typically > 1600 Hz) are given by animals within 5 s of a call from another individual and rarely elicit replies.

Key words: mountain gorilla, social behavior, communication, vocalizations

INTRODUCTION

When humans classify the calls of nonhuman primates by ear, they typically lump all similar sounding vocalizations within the same class, labeling them onomatopoetically as coos, grunts, screams, barks, and so on. Recent research, however, has shown that in many species the animals themselves make further discriminations, and distinguish a number of subtly different call subtypes within each of these broad acoustic classes. Playback experiments verifying that such subtypes elicit different behavioral responses have been conducted on several different species [see reviews by Snowdon, 1988; Cheney & Seyfarth, 1990].

The acoustic cues that differentiate call subtypes vary from one species to the next. For example, two of the four subtypes of pygmy marmoset trills differ only in

Received for publication August 5, 1993; revision accepted November 22, 1993.

Address reprint requests to Robert M. Seyfarth, Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104.

duration [Snowdon & Pola, 1978], whereas subtypes of the Japanese macaque coo differ according to the location of the call's peak frequency [Green, 1975; Zoloth et al., 1978], and subtypes of the rhesus and the pigtail macaque scream differ according to a variety of measures, including bandwidth, the presence of short pulses, and the presence or lack of frequency modulation [Gouzoules et al., 1984; Gouzoules & Gouzoules, 1989]. Among vervet monkeys, four acoustically different grunts given in different social situations can be distinguished according to the location and changes in two frequency peaks, one at roughly 240 Hz and the other between 550 and 900 Hz [Cheney & Seyfarth, 1982; Seyfarth & Cheney, 1984]. Owren [1990a, b] demonstrated experimentally that vervet monkeys discriminate between two of their alarm call types on the basis of such resonant frequencies, or formants [Ladefoged, 1982; see also Hienz & Brady, 1988]. Such results are of special interest because formant frequency patterns provide some of the most salient acoustic cues used in the perception of human speech [Lieberman, 1984].

In a variety of different situations—when feeding, resting, playing, or moving from one location to another—free-ranging mountain gorillas (*Gorilla gorilla beringei*) utter a vocalization which Fossey [1972] described as a "belch" or "double-belch," but which is more accurately termed a "double grunt" because of the way in which it is produced (see below). The call consists of two acoustic units, the first shorter than the second (Fig. 1), and is produced by the gorillas with their mouths closed. While double-grunts are by no means the only vocalization produced by mountain gorillas, they are given at a higher rate than any other call type [Marler, 1976a; Harcourt et al., 1986, 1993].

Gorillas often give double-grunts when in full view of one another. In many other cases, however, the vocalizer is wholly or partially obscured by vegetation. Although their precise function has not yet been determined, double-grunts appear in some cases to mediate competitive interactions: subordinates are more likely to stop feeding on the approach of a dominant when grunts are exchanged, than when they are not, or no vocalization is given [Harcourt et al., 1986].

To a human listener, all double-grunts sound more or less the same. Although the double-grunts of silverbacks can be distinguished by their low pitch, no other consistent differences between individuals or between double-grunts given in different social contexts are immediately apparent. The existence of acoustic subtypes in other primate species, however, suggests that gorilla double-grunts may be more complex than they first appear. Previous observers [Fossey, 1972; Harcourt et al., 1986; see also Harcourt et al., 1993] have noted that roughly one-third of all double-grunts are followed immediately by a double-grunt or other vocalization from one or more animals nearby. Such "answering" vocalizations might depend primarily on the caller's or the listener's behavior at the time a vocalization is given, or they might be a function of the longer-term social relationship between caller and recipient. Alternatively, call "sequences" could simply be an incidental outcome of many animals vocalizing independently. Finally, bearing in mind the acoustic sub-types found in other primates (see above), it is possible that gorillas distinguish acoustically distinct subtypes of double-grunts, and that one of these subtypes initiates an exchange, while another functions as an "answer."

In this paper, we present the results of field observations on the use of double-grunts by mountain gorillas, together with waveform and spectral analysis of over 400 double-grunts given by seven different individuals. Our goal is to describe the acoustic features of double-grunts, to determine whether gorillas use different acoustic subtypes in different social situations, and, if so, to clarify which acoustic features might underlie this discrimination.

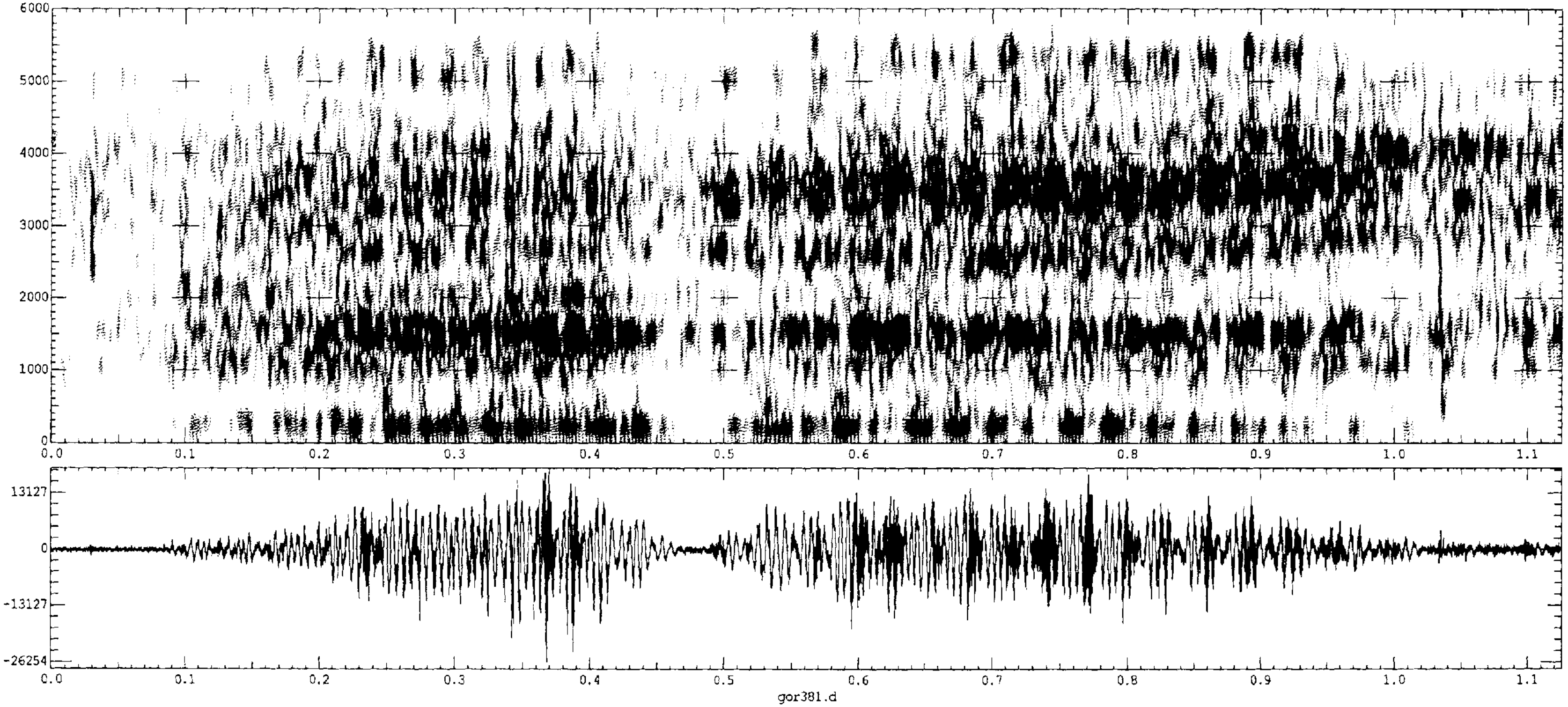

Fig. 1. Wide-band (300 Hz filter) spectrogram (above) and waveform (below) of a gorilla double-grunt. On the spectrogram, X-axis indicates time in seconds, Y-axis indicates frequency in kHz.

METHODS

Field Recording

The double-grunts of gorillas in two social groups, '5' and 'Pn' were recorded during February and March, 1983 in the Parc National des Volcans, Rwanda, and the Parc National des Virungas-Sud, Zaire. Details of the study area are given in Vedder [1984] and Watts [1991]. Details of social relationships among mountain gorillas are reviewed in Stewart and Harcourt [1987]. At the time of observation, Group 5 included two silverback males, one fully adult-sized blackback male, four adult females, and seven immatures. Group Pn was composed of two silverback males and five blackback males.

Study groups were fully habituated to human observers. Vocalizations were recorded at a distance of 0.1 to 3 m using a Sennheiser directional microphone (model ME88) and either a Uher Report Monitor tape recorder (tape speed 9.5 cm/s) or a Sony TCD-5M cassette recorder. During a typical day's observation, two investigators were present with each group. One observer (either Harcourt or Stewart) collected data on social behavior following a sampling regime similar to that described in Harcourt et al. [1993], while the other observer (either Cheney or Seyfarth) recorded double-grunt vocalizations from as many individuals as possible. Only calls from known individuals were retained for acoustic analysis. Whenever a call was tape-recorded, we noted the time, caller's identity, behavior, and the identity and behavior of other individuals within 5 m. We also noted the predominant behavior of individuals in the group at the time. If a double-grunt was given within 5 s of a vocalization from another individual, it was labeled an "answer," otherwise it was labelled a "spontaneous" vocalization. When a double-grunt was given in answer or elicited a reply, we noted the identity and behavior of the other individual involved (in virtually all cases studied, only one individual answered another).

Social Behavior

To supplement field recordings of double-grunts, we used data on rates of vocalization, social behavior, and social relationships from sub-periods of a more than two-year long study of group 5 and another bisexual group, Nk. The data presented here come from more than 12 h of focal animal sampling per animal in group 5 over a period of six months in 1982–1983 (median of 12.5 h/adult), and 17.5 h over five months for adults in group Nk in 1981. (Data for group Nk immatures are not presented, because they double-grunted too infrequently.) Data were collected using the methods described in Harcourt and Stewart [1989] and Harcourt et al. [1993]. Dominance was defined according to the direction of approach–retreat interactions. In both heterosexual groups, silverback males were dominant to blackback males, blackback males were dominant to females, and females were dominant to immatures, and the same relations among males held in the all-male group, Pn; in addition, individuals within each age-sex class could be assigned dominance ranks relative to each other [Harcourt & Stewart, 1989; Stewart & Harcourt, 1987; Yamagiwa, 1987].

Acoustic Analysis

Calls were digitized at a sampling rate of 12,000 Hz using a Sun microsystems IPX workstation and WAVES software for signal processing and analysis, developed by David Talkin at AT&T Bell Laboratories and available commercially from Entropic Speech Research Laboratory, Washington, D.C. From a display of the waveform and spectrogram of each call (Fig. 1), we used a cursor to measure the

TABLE I. The Acoustic Features of Gorilla Double-Grunts Measured in This Study

I. Temporal domain
  1. Length of unit 1
  2. Length of unit 2
  3. Length of inter-unit interval
  4. Total call length (1 + 2 + 3)

II. Frequency domain
  5.–9. Values of F1, F2, F3, F4, and F5, unit 1
  10.–14. Values of F1, F2, F3, F4, and F5, unit 2
  15.–19. Change in F1, F2, F3, F4, and F5 within unit 1
  20.–24. Change in F1, F2, F3, F4, and F5 within unit 2
  25.–29. Change in F1, F2, F3, F4, and F5 between units 1 and 2
  30.–33. Ratios F2/F1, F3/F1, F4/F1, and F5/F1, unit 1
  34.–37. Ratios F2/F1, F3/F1, F4/F1, and F5/F1, unit 2
  38.–41. Values of (F2-F1), (F3-F1), (F4-F1), and (F5-F1), unit 1
  42.–45. Values of (F2-F1), (F3-F1), (F4-F1), and (F5-F1), unit 2

duration of each unit, the inter-unit interval, and the total call length (measures 1–4 in Table I).

Although many double-grunts gave no indication of regular periodicity, in 30–40% of all calls some periodicity was apparent, either in the waveform or in clear vertical striations at regular intervals on the spectrogram (see for example unit 1 in Fig. 2). In the hope of obtaining data on fundamental frequency (FO), we used a cursor to measure pitch periods by eye. Results suggested that the FO of gorilla double-grunts was between 25 and 40 Hz. These values, however, fall outside the limits imposed by our recording equipment. As a result, we cannot be sure of their accuracy, and have not included them in our analysis.

To begin the analysis of spectral characteristics, we first examined a large number of spectrograms using wide-band (300 Hz) and narrow band (75 Hz) filters. Frequency spectra were then computed using discrete Fourier transformations (DFT). To obtain an overall picture of each call, one DFT spectrum was computed for each unit; for these measurements the analysis window was adjusted to match the length of each unit, up to a maximum window size of 300 msec. In cases where a unit (typically, unit 2) was longer than 300 msec, we measured a spectral "slice" of 300 msec duration beginning 20 msec after the unit's onset. Such measurements yielded a frequency resolution of 3–4 Hz and provided the primary source of data on spectral characteristics. Figure 3 shows a DFT spectrum of unit 1 in Figure 1.

To obtain data on changes in frequency peaks over time, a number of spectral slices was computed at regular intervals throughout each call: four spectral slices in unit 1 and six in unit 2. Thus, the length of each slice varied depending upon the length of each unit. For these measurements, the analysis window was held constant at 25 msec, yielding a frequency resolution of 40 Hz. To determine whether a spectral peak rose, fell, or remained stable over time, we calculated the slope of the regression line created by the four (or six) successive measures of a peak in unit 1 (or 2). This slope was either positive, negative, or zero (measures 15–24, Table I).

Finally, as an aid in identifying frequency peaks and measuring their bandwidths, 10th order autocorrelation LPC peaks were computed for each of the spectra described above, yielding a smoothed frequency spectrum such as that shown in Figure 3. Nth order autocorrelation analysis "searches for" the strongest ((N/2)-1) peaks in the temporal waveform [Markel & Gray, 1976; Owren, in press]. Tenth order analysis was chosen because (i) examination of a large number of spectro-

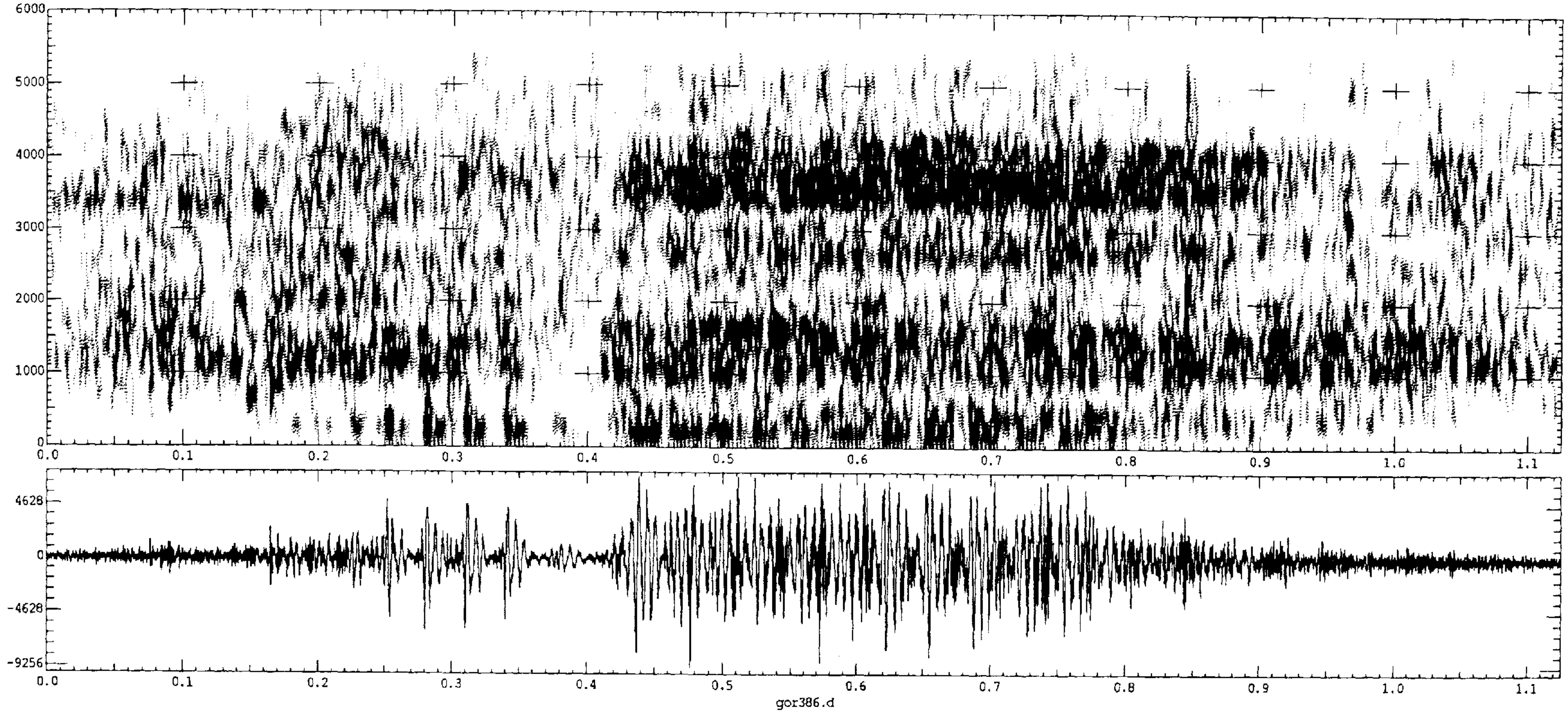

Fig. 2. Legend as in Figure 1.

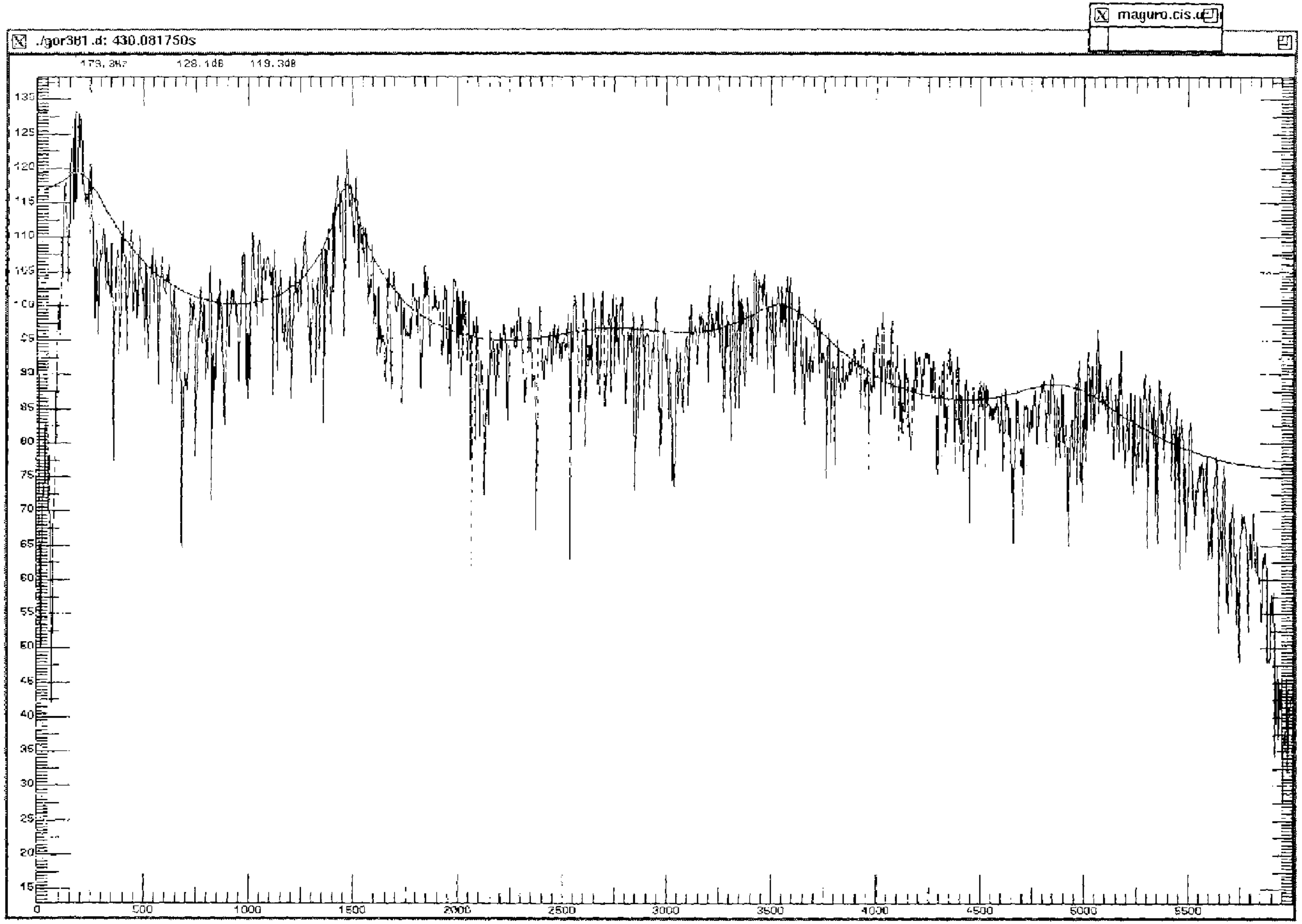

Fig. 3. Discrete Fourier transformation (DFT) spectrum of unit 1 in the double-grunt shown in Figure 1. The smoothed frequency spectrum was obtained using a 10th order autocorrelation LPC subroutine (see text).

grams revealed few, if any, calls with more than four peaks, (ii) the length of the gorilla vocal tract [Lieberman, 1968] and comparative research on human speech lead one to expect, in gorillas, roughly one frequency peak every 1,000 Hz (Ladefoged personal communication), and (iii) in systematic comparisons of 20 randomly chosen FFT spectra with smoothed frequency spectra using values of N from six to 20, tenth order analysis consistently provided the closest visual "fit" between the FFT and the smoothed frequency spectrum (see Owren, in press, for further discussion of applying LPC analysis to nonhuman primate vocalizations). As a first step in analysis, we examined in detail 79 double-grunts given by one adult male, Sr, a member of the all-male group Pn. Sr's double-grunts were first sorted according to the social context in which they occurred: whether Sr was near and/or in sight of other individuals, whether he was feeding, resting, moving, or engaged in any one of a variety of different social behaviors at the time of vocalizing, and whether the group as a whole was feeding, resting, engaged in social behavior, moving, or about to move to another area. We then searched for any acoustical feature that might allow a listener to distinguish a call given in one of these behavioral contexts from a call given in another. If one or more acoustic features differed statistically from one context to the next, we concluded that gorillas might potentially make use of this cue when distinguishing among different call subtypes. We then formulated a hypothetical set of 'rules' (see below, Results section D.1) for distinguishing among the double-grunts of adult male Sr, and tested whether these rules also accurately discriminated among the double-grunts of six other individuals (five males and one female).

RESULTS

Rates of Spontaneous and Answering Double-Grunts

In both heterosexual groups, 5 and Nk, silverback males gave double-grunts at far higher rates than did other group members, and adults gave them more often than did immatures (Fig. 4, 5) [see also Fossey, 1972; Harcourt et al., 1986, 1993; and Marler, 1976a]. In addition, it appeared that high-ranking individuals gave double-grunts more often than did individuals of lower dominance rank, as Harcourt et al. [1986] indicated. Thus, data from all individuals in group 5 showed that the Spearman correlation between dominance rank and grunts/hr was 0.945 ($P < 0.01$); among adults alone, it was 0.815 ($P < 0.05$) in group 5 (see Fig. 4), and 0.732 ($0.1 > P > 0.05$) in group Nk (see Fig. 5). Adults were more likely to be answered than were immatures in group 5 (Fig. 4; too few calls by group Nk immatures for analysis here), but otherwise no relation between competitive ability and being answered was found.

Timing

The majority of the first subject's, Sr's (group Pn), taped calls (74/79, or 94%) consisted of two units. This was also true for each of the six other individuals tested. For all seven individuals tested (N = 271 calls), unit 1 had a mean length of 258.8 msec (SD = 99.4); unit 2 had a mean length of 565.9 msec (SD = 241.0), and the break between units had a mean length of 75.0 msec (SD = 38.3). Unit two was longer than unit one in every vocalization tested. These results are very close to Harcourt et al.'s [1993] values from an independent sample of 30 calls from 13 individuals in groups 5 and Nk.

Spectral Analysis

Formants. Resonant frequency peaks, or formants, in Sr's double-grunts occurred at roughly 230, 1,560, 2,700, 3,900, and 5,000 Hz (Fig. 6). Similar results

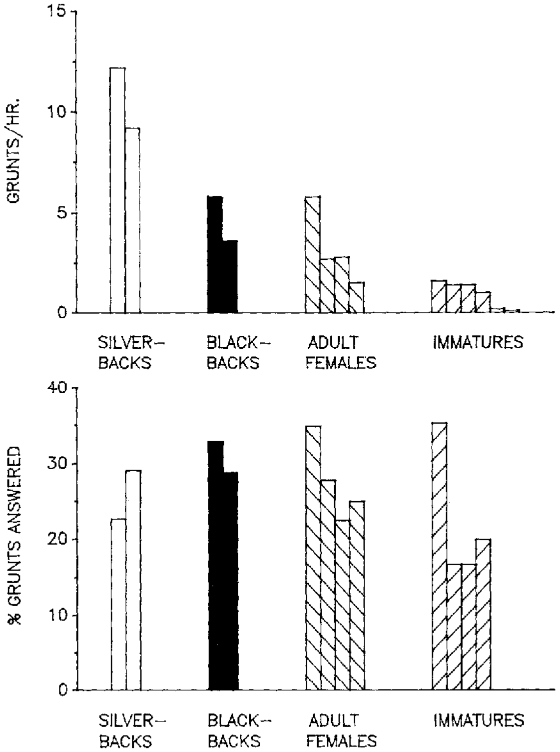

Fig. 4. Rates of double-grunting (top) and rates at which double-grunts were answered in group 5. Animals grouped by age-sex class and arranged, both within and between age-sex classes, in descending dominance rank order from left to right.

were obtained from the six other individuals tested (Table II). Not all formants were present in all calls, however. As data in Table II indicate, while F1 and F2 could be identified in almost all double-grunts, the third, fourth, and fifth formants were often absent.

Formant changes. In many of Sr's double-grunts, there was a shift in formant values from unit 1 to unit 2. Often, formants changed independently of one another, with, for example, F1 rising and F2 falling from unit 1 to unit 2. Apparently, gorillas can change the shape of their supralaryngeal resonant cavity so that it is fixed in one configuration during production of unit 1, and then assumes a slightly different configuration during production of unit 2.

There was little evidence, however, that gorillas change the shape of their vocal tract during a given unit; their double-grunts, in other words, revealed few formant transitions analogous to those found in human speech. In our sample of over 400 double-grunts, we found less than 30 calls with a clear formant transition. While gorillas clearly can produce formant transitions, such transitions seem to play a relatively unimportant role in their production of double-grunts.

The Relation Between Acoustic Variation and Behavior

Sr's double-grunts. To test for any relation between the acoustic features of double-grunts and ongoing behavior, we sorted Sr's calls into groups according to whether he was visible to others or not, and according to whether he was feeding,

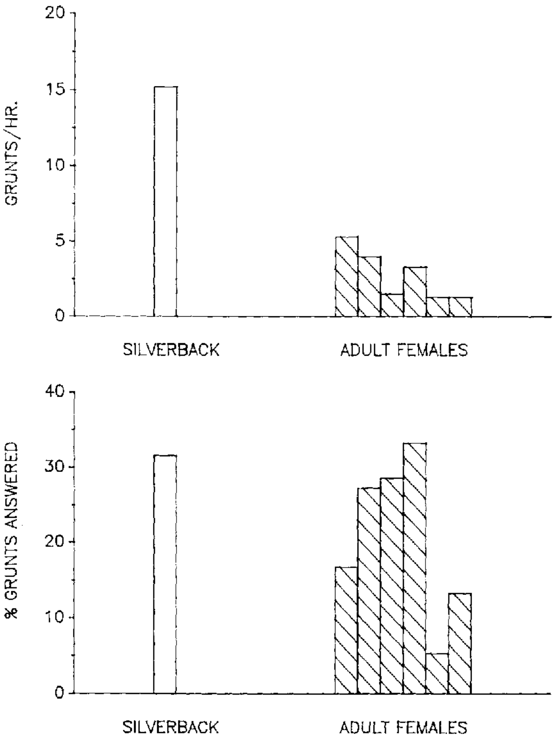

Fig. 5. Legend as in Figure 4, except data from group Nk.

resting, moving, or engaged in a particular sort of social behavior at the time of vocalizing. We then used a Kruskal-Wallis one-way analysis of variance by ranks to compare calls in these different behavioral categories according to each of the acoustic measures listed in Table I. Analysis revealed no relation between Sr's behavior at the time of vocalizing and any acoustic measure. There was also no relation between the acoustic features of Sr's double-grunts and caller visibility, nor did double-grunts differ acoustically depending on the likelihood of imminent behavioral change; that is, the probability that the vocalizer or the group as a whole would switch from one activity to another within 30 s or 1 min after a double-grunt had been uttered.

There were, however, significant acoustical differences between calls that were given spontaneously, and calls that were apparently given in answer to others (i.e., within 5 s of another individual's double-grunt). When Sr vocalized spontaneously, 60% (32/53) of his double-grunts elicited a reply; when Sr gave a double-grunt within 5 s of a vocalization given by another individual, 31% (8/26) of his calls elicited a reply ($\chi^2 = 6.11$, df $= 1$, $P < 0.013$). In other words, calls given spontaneously by Sr frequently elicited replies, calls given within 5 s of another animal's call less frequently elicited replies, and as a result, there were many two-call sequences (A vocalizes, followed by B), but very few three-call sequences (A vocalizes, followed by B, followed by C). This result could, of course, have occurred simply because gorillas closely monitor the pattern of calling within their group, and generally refrain from replying to another individual's double-grunt if

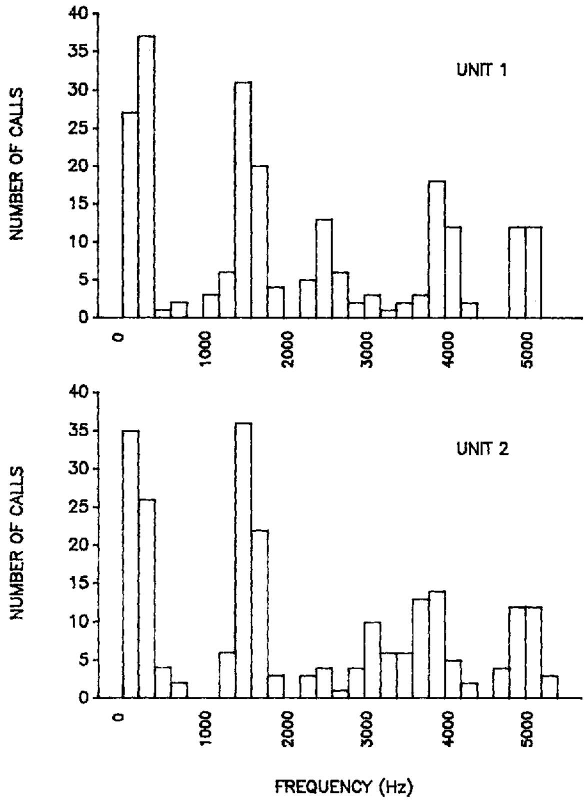

Fig. 6. Number of calls given by adult male Sr that showed resonant frequency peaks at each 200 Hz interval between 0 and 5000 Hz. Based on an analysis of 67 calls; data for units 1 and 2 presented separately.

that individual's call has occurred in the context of an answer. It is also possible, however, that answering grunts differ acoustically from those given spontaneously, and that, in addition to monitoring the pattern of calling within their group, listeners generally refrain from answering a double-grunt that has the acoustic features of a reply.

A variety of data suggest that answering grunts do differ consistently from those given spontaneously. When compared with Sr's spontaneous double-grunts, Sr's answers had a significantly shorter inter-unit interval, lower F1 in unit 1, and a higher F2 in both units (Table III).

From an analysis of Sr's calls, we may therefore hypothesize that gorilla double-grunts include at least two acoustically distinct subtypes: calls given spontaneously and those given in reply to another individual. The subtypes differ according to the acoustic cues listed in Table III. At least in part because calls given spontaneously are acoustically different from calls given in answer, the two double-grunt subtypes lead to different behavioral outcomes. Spontaneous double-grunts elicit a high rate of answering calls, whereas answering calls rarely elicit a reply themselves. To test this hypothesis, at least as it pertains to the acoustic differences among call subtypes, we analyzed the calls of six other individuals, five adult males and one adult female.

The double-grunts of other individuals. In five of the six individuals

TABLE II. Mean Values for Five Formants in Units One and Two of Double-Grunts Given by Six Different Individuals. Values in Each Cell Give Sample Size, Mean Formant Frequency (in Hz), and Standard Deviation (in parenthesis)

Individual		F1	F2	F3	F4	F5
Sr	Unit 1	67	67	30	37	24
		226.8	1558.2	2619.9	3942.3	4993.3
		(96.5)	(155.5)	(280.0)	(158.3)	(90.0)
	Unit 2	67	67	26	43	29
		234.5	1561.2	2899.0	3828.5	4979.7
		(112.0)	(133.3)	(331.2)	(283.6)	(158.9)
Bt05	Unit 1	51	51	12	17	15
		269.4	1591.8	2875.0	3500.9	4786.6
		(115.7)	(257.5)	(347.2)	(139.4)	(173.1)
	Unit 2	51	51	8	22	11
		318.2	1618.3	2867.4	3528.5	4804.3
		(139.7)	(211.5)	(252.1)	(206.5)	(177.8)
Ef	Unit 1	28	28	—	6	2
		248.9	1738.5	—	3656.6	4884.2
		(69.4)	(293.2)	—	(194.6)	(25.8)
	Unit 2	28	28	—	4	4
		312.9	1767.9	—	3503.1	4596.2
		(119.9)	(305.0)	—	(362.6)	(309.5)
Pn	Unit 1	40	40	10	9	1
		201.2	1427.3	2421.9	3880.1	4436.8
		(80.9)	(112.9)	(419.2)	(266.3)	(0.0)
	Unit 2	40	40	7	6	2
		318.5	1420.0	2675.9	3823.2	4701.2
		(162.9)	(158.3)	(97.2)	(253.7)	(284.5)
Ic	Unit 1	27	27	3	6	1
		243.2	1536.9	2640.2	3746.9	4902.4
		(111.1)	(287.7)	(393.3)	(81.3)	(0.0)
	Unit 2	27	27	3	7	—
		233.8	1569.1	2585.4	3692.5	—
		(98.5)	(360.1)	(559.1)	(204.1)	—
Bt	Unit 1	40	40	3	6	2
		217.4	1463.7	2702.0	3828.3	5015.2
		(57.0)	(274.1)	(438.4)	(154.1)	(0.0)
	Unit 2	40	40	7	7	—
		300.8	1487.5	2766.2	3781.3	—
		(134.0)	(259.9)	(241.6)	(337.1)	—
Zz	Unit 1	18	18	1	4	2
		232.8	1462.1	2712.1	3681.8	4930.8
		(59.0)	(356.2)	(0.0)	(75.2)	(73.5)
	Unit 2	16	18	1	4	1
		256.9	1473.7	2712.1	3625.0	4984.8
		(132.7)	(329.2)	(0.0)	(96.1)	(0.0)

tested, their spontaneous calls were significantly more likely to be immediately followed by calls from others than were their calls given in answer (Fig. 7), suggesting that for these animals, too, spontaneous and answering calls differed

TABLE III. Acoustic Cues That Differentiated Adult Male Sr's Double-Grunts Given Spontaneously From Those Given in Reply. Values in Each Cell Give Mean, SD, and Sample Size. Significance Tests Based on Kruskal-Wallis One-Way Analysis of Variance by Ranks

Acoustic feature	Spontaneous double-grunts	Answering double-grunts	Significance test
Duration of break	0.053 s +0.015 N = 45	0.072 s +0.015 N = 21	t = 19.54**
F1, unit 2	225.6 Hz +114.9 N = 46	253.8 Hz +106.0 N = 21	t = 4.81*
F2, unit 1	1534.9 Hz +149.7 N = 46	1609.3 Hz +159.3 N = 21	t = 3.82*
F2, unit 2	1539.1 Hz +137.0 N = 46	1609.5 Hz +113.5 N = 21	t = 3.56*

$^{**}P < 0.01$; $^{*}P < 0.05$.

acoustically. Indeed, two of the acoustic features that distinguished Sr's spontaneous and answering calls, also distinguished the spontaneous and answering calls of other individuals. Answering calls had both a higher F2 in unit 1 and a higher F2 in unit 2 (Table IV) than did calls given spontaneously. Two of the other acoustic features that distinguished Sr's spontaneous double-grunts from his answers—the length of the inter-unit interval and the value of F1 in unit 1—were apparently unique to Sr; few of the other individuals tested exhibited a significant difference according to these measures (Table IV). The data suggest, therefore, that gorillas produce calls with a lower F2 when vocalizing spontaneously and a higher F2 when giving a call in reply, and that listeners might attend to the value of F2—among other behavioral events—when distinguishing between calls given spontaneously and those given in reply.

For most individuals tested, the absolute value of F2 reliably differentiated spontaneous calls from replies. As the data in Fig. 8 indicate, there was relatively little overlap in the distribution of F2 values in the two different call types: in both units, the F2 value for a call given spontaneously typically fell outside the 95% confidence limits for a call given in reply, and vice versa.

On average, calls given spontaneously had an F2 in unit 1 that was 213.3 Hz lower than the F2 in unit 1 for calls given in reply (SD = 88.3, range 74.4–313.9). In unit 2 the mean difference was 206.4 Hz (SD = 91.3, range 70.4–314.1). Although little is known about auditory perception in mountain gorillas, research on chimpanzees (*Pan troglodytes*) and humans suggests that discriminations of roughly 200 Hz are well within the gorilla's sensory capacities. Chimpanzees and humans tested with pure tones between 0.5 and 2.0 kHz exhibit difference thresholds of 10–15 Hz and 2–5 Hz, respectively [Kojima, 1990; Wier et al., 1977; Sinnott et al., 1987]. When tested with synthesized speech in which the value of F2 has been systematically varied between 1 and 2 kHz, humans exhibit a difference threshold between 20 and 90 Hz [Flanagan, 1955; reviewed in Rosen & Fourcin, 1986].

Additional support for the view that a high second formant identifies an an-

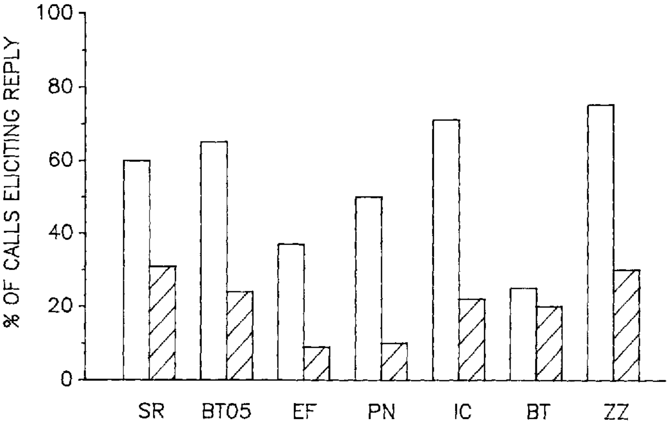

Fig. 7. The proportion of all spontaneous calls (open histograms) and answers (hatched histograms) given by different individuals that elicited a vocal reply from another gorilla. Individuals listed along the X-axis.

TABLE IV. Results of Statistical Tests Designed to Determine Whether the Acoustic Cues That Differentiated Adult Male Sr's Spontaneous vs. Answering Double-Grunts Also Differentiated the Spontaneous and Answering Double-Grunts of Six Other Individuals

	Acoustic feature			
Individual	Break duration	F1, unit 1	F2, unit 1	F2, unit 2
Bt05	NS	NS	t = 4.59*	t = 5.56*
Ef	t = 6.20**	NS	t = 3.88*	t = 3.88*
Pn	NS	NS	t = 11.49**	NS
Ic	NS	t = 3.72*	NS	t = 4.05*
Bt	NS	NS	t = 13.47**	t = 14.94**
Zz	NS	NS	t = 4.74*	t = 4.19*

$^{**}P < 0.01$; $^{*}P < 0.05$.

swering double-grunt comes from an analysis of the small subset of answering calls that did elicit replies. As mentioned earlier, few answering calls elicited replies themselves. However, if gorillas use F2 to distinguish between spontaneous calls and answers, and if a high F2 denotes an answer, then the relatively few answering calls that did elicit replies should have been characterized by lower than normal F2s. This was, in fact, the case. For a significant number of individuals, answering calls that elicited replies had a lower mean F2 in unit 1 than did answering calls that did not elicit replies (Wilcoxon Test, N = 7, T = 3, $P < 0.04$); data from unit 2 were similar, but did not reach statistical significance (Wilcoxon Test, N = 7, T = 6, $0.10 > P > 0.05$) (Fig. 9). For all individuals tested, the probability of eliciting a vocal response declined as the value of F2 increased (Fig. 9).

Individual Identity

In addition to providing information about whether a vocalizer is grunting spontaneously or replying to another, gorilla double-grunts may also signal indi-

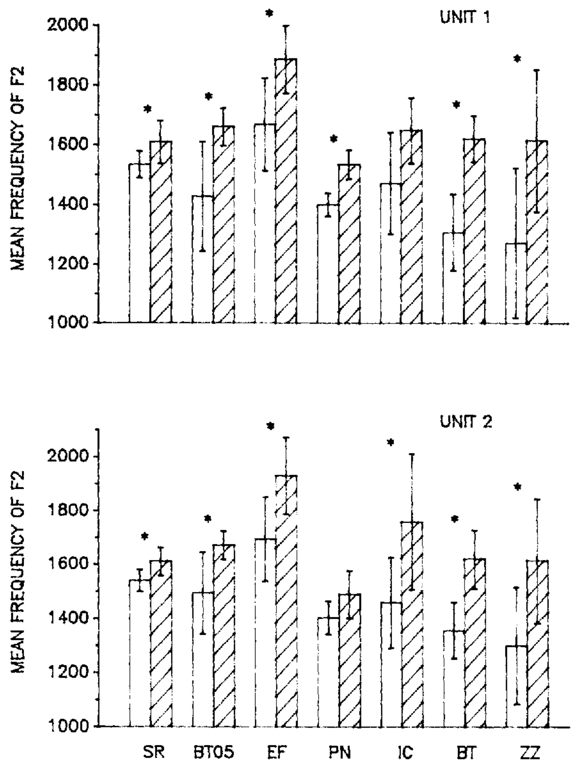

Fig. 8. The mean frequency of F2 in spontaneous calls (open histograms) and answers (hatched histograms) given by different individuals. Individuals listed along the X-axis. Histograms show mean values and 95% confidence limits for F2 in unit 1 (top) and unit 2 (bottom). An asterisk indicates a significant difference between the two distributions (Kruskal-Walis one-way analysis of variance by ranks, $P < 0.05$).

vidual identity. To test for individuality in spontaneous and answering calls, we examined heterogeneity among different individuals for each of the acoustic features listed in Table I, using a Kruskal-Wallis one-way analysis of variance by ranks. To hold group membership and social context constant, we tested each group separately and carried out separate analyses of spontaneous calls and calls given in reply. Where significant overall heterogeneity ($P < 0.05$) appeared, paired comparisons were carried out using Mann-Whitney U Tests. Table V presents, for each group and each call type separately, the acoustic cues that yielded a significant difference in each of these paired comparisons. Two conclusions emerged.

First, in most cases, the individuals within a group differed from one another by at least one acoustic measure. Only the answering double-grunts of three individuals in group 5, Bt05, Ic, and Zz, showed no statistically significant differences according to any of the acoustic measures tested. Gorilla double-grunts thus potentially carry information about individual identity as well as social context.

Second, while spontaneous calls and replies were differentiated acoustically only according to the value of F2, individual identity was differentiated by many other acoustic features, such as the duration of unit 1 (measure 1), the duration of the inter-unit interval (measure 3), the value of F1 (measures 5 and 10), and the difference between F2 and F1 in unit 1 (measure 38). These results suggest that, if gorillas distinguish the calls of different individuals, they might attend to a

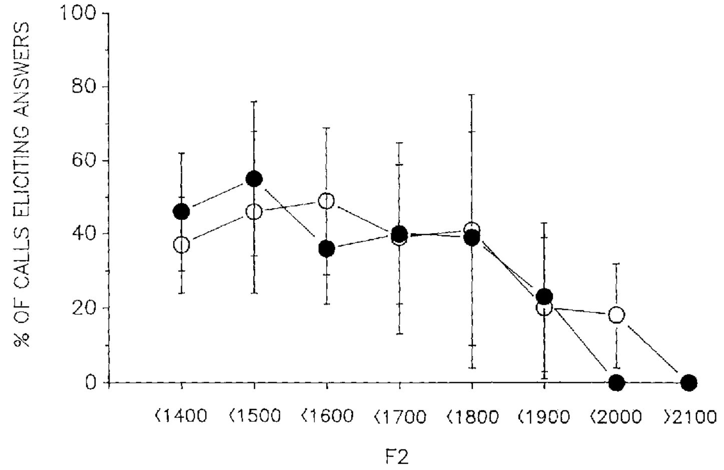

Fig. 9. The relation between F2 in unit 1 (open circles), F2 in unit 2 (closed circles), and the proportion of calls eliciting a reply from other individuals. Values shown are means and standard deviations.

different set of acoustic cues from those they could attend to when distinguishing different call types. Lillehei and Snowdon [1978] obtained similar results in a study of stumptail macaques.

DISCUSSION

Double-grunts are the most common vocalization uttered by mountain gorillas [Marler, 1976a; Harcourt et al., 1986, 1993]. Although they appear to constitute a single, highly variable class of acoustic signals, the gorillas' selective response to calls suggests the potential for greater complexity.

In one group, double-grunts of high-ranking individuals were apparently more likely to elicit a double-grunt in reply than were double-grunts from low-ranking individuals (Fig. 4). There was no other indication, however, that the gorillas' selective responding depended either on differences in behavior at the time of vocalizing or on any features of the social relationship of caller and respondent. Nor did we find any evidence that, for example, calls given in dense vegetation when the caller was out of sight of others were more likely than other calls, to elicit replies. With no obvious context on which the gorillas could rely, it is possible that their selective responses depended on the perception of at least two different acoustic subtypes of double-grunt, each of which could have conveyed different information and hence served a different function [Biben et al., 1986, and Symmes & Biben, 1988 report similar results in studies of squirrel monkeys].

When analyzed according to their acoustic features, gorilla double-grunts were both individually distinctive (Table V) and fell into two classes. One call type had a low second formant (roughly speaking, below 1,600 Hz), and was given by animals after a period of silence. It was described, therefore, as a "spontaneous" double-grunt. This call frequently elicited a double-grunt in reply (Fig. 7). A second call type—called an "answer"—had a higher second formant (roughly speaking, above 1,600 Hz), and was given by animals within 5 s of a call from another individual. This call less often elicited a double-grunt in reply (Fig. 7). In the rare cases when an answer did elicit a reply, the responder may have confused an answer with a spontaneous double-grunt. Answers that elicited replies had lower

TABLE V. Acoustic Cues That Yielded a Significant Difference ($P < 0.05$) Between the Calls of Different Individuals When Group Membership and Call Subtype Were Held Constant. Numbers in Each Cell Refer to the Numbers of Acoustic Features Listed in Table I *

	Pn	Bt	Sr
Pn group, spontaneous calls			
Pn		5,30	3,5,6,10,11 34,38,42
Bt			3,6,10,11, 30,38,42
Pn group, answers			
Pn		34	11,34,38
Bt			2,34,38,42
Bt05	Ic	Zz	Ef
Group 5, spontaneous calls			
Bt05	3,5,10,30, 38	1,3	2,6,11
Ic		1,5,6,11,30	1,3,5,6,10, 11,38
Zz			6,10,11,30
Group 5, answers			
Bt05	NS	NS	1,6,11
Ic		NS	1,6,11
Zz			6,11

*Where a number appears, all individuals in that context showed significant heterogeneity ($P < 0.05$) when tested with a Kruskal-Wallis one-way analysis of variance by ranks, and the two individuals were significantly different ($P < 0.05$) when tested with a Mann-Whitney U test.

second formants, and hence may have been more ambiguous, when compared with answers that did not elicit replies (Fig. 9).

Data on the relation between F2 and the probability of a vocal response suggest at least two hypotheses about the function of double-grunts. On the one hand, double-grunts may mediate competitive situations, as suggested by Harcourt et al. [1986, 1993]. A low F2, for example, may—together with other social signals—indicate the caller's intention to contest a resource, perhaps merely by providing information on intention to stay [Hinde, 1981]. If they are perceived as mildly threatening, calls with low F2 values might be especially likely to elicit vocal replies. The replies might have high F2 values in order to signal non-competitiveness, even submission. Once submission has been conveyed, no further reply is necessary, thus explaining why calls with a high F2 rarely elicit replies.

Such correlations, however, are not uniformly significant, and an alternative hypothesis argues that gorilla double-grunts are not competitive calls but instead function primarily to mediate spacing between individuals. In this case, calls with a low F2—again, together with other social signals—would indicate the caller's location and/or intention to move, whereas calls with a high F2 would function as replies. This hypothesis suggests that high-ranking animals vocalize at higher rates because, for instance, they play a greater role than lower-ranking animals in the initiation of group movement. At present, we have insufficient data to test these two hypotheses.

Not only that, but both could be correct. Given the variety of contexts in which calls are given, it is highly unlikely that one explanation for the function of the calls will suffice [Harcourt et al., 1986, 1993]. Competition is unlikely to be the function of calls exchanged at 30 m, but cohesion is unlikely to be the function of calls exchanged at 3 m. Testing of hypotheses about message, meaning, and function will have take very close account of context [Smith, 1977].

Similarly, at least two hypotheses can be proposed concerning the mechanisms that underlie the perception of gorilla double-grunts. On the one hand, double-grunts may comprise a graded series of calls in which the probability of eliciting a vocal response is determined by the value of F2, as shown in Figure 9. Alternatively, double-grunts may be perceived by listeners as two discrete call types separated by a perceptual 'boundary' somewhere around 1,600 Hz. Calls with relatively low or relatively high F2s are unambiguous, signaling (for example) either a mildly aggressive threat or submission. As a result, they elicit replies at relatively high or relatively low rates. By contrast, calls close to the perceptual boundary are more ambiguous in the information they convey to others, and consequently elicit replies at intermediate rates.

Mountain gorilla double-grunts could, therefore, constitute a graded series of calls that is perceived either continuously or as two or more discrete categories. Although these two perceptual processes can lead to similar behavioral outcomes, they are important to distinguish because of their implications for research both on other nonhuman primate vocalizations and on the evolution of language [e.g. Marler, 1976b; Owren et al., in press]. At present, however, in the absence of detailed perceptual experiments, it is impossible to reach any conclusions about the perceptual mechanisms underlying the response to gorilla double-grunts.

ACKNOWLEDGMENTS

We thank the Governments of Rwanda and Zaire, especially the Office Rwandais du Tourisme et des Parcs Nationaux and the Institut Zairois pour la Conservation de la Nature, for permission to conduct the study. The Karisoke Research Centre field assistants, especially Alphonse Nemeye and Emmanuel Rwelekana, enormously improved the efficiency of our work.

For comments on earlier drafts, we thank Nicola Bessel, Marc Hauser, William Labov, Peter Ladefoged, Mark Liberman, Peter Marler, Michael Owren, Chuck Snowdon, Tom Veatch, Christine Zeller, and three anonymous reviewers.

For financial support, RMS and DLC thank the NSF (grants BNS 82-15039 and 85-21147), the NIH (grant NS 19826), and the University of Pennsylvania; AHH and KJS thank the National Geographic Society and the Harry Frank Guggenheim Foundation. This is publication 127 of the Karisoke Research Centre.

REFERENCES

Biben, M.; Symmes, D.; Masataka, N. Temporal and structural analysis of vocal exchanges in squirrel monkeys (*Saimiri sciureus*). BEHAVIOUR 98:259–273, 1986.

Cheney, D.L.; Seyfarth, R.M. How vervet monkeys perceive their grunts: Field playback experiments. ANIMAL BEHAVIOUR 30:739–751, 1982.

Cheney, D.L.; Seyfarth, R.M. HOW MONKEYS SEE THE WORLD. Chicago, University of Chicago Press, 1990.

Flanagan, J.L. A difference limen for vowel formant frequency. JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 27:613–617, 1955.

Fossey, D. Vocalizations of the mountain gorilla (*Gorilla gorilla beringei*). ANIMAL BEHAVIOUR 20:36–53, 1972.

Gouzoules, H.; Gouzoules, S. Design features and developmental modification of pigtail macaque, *Macaca nemestrina* agonistic screams. ANIMAL BEHAVIOUR 37:383–401, 1989.

Gouzoules, S.; Gouzoules, H.; Marler, P. Rhesus monkey (*Macaca mulatta*) screams: Representational signalling in the recruitment of agonistic aid. ANIMAL BEHAVIOUR 32:182–193, 1984.

Green, S. Communication by a graded vocal system in Japanese monkeys. Pp. 1–102 in PRIMATE BEHAVIOR, VOL. 4. L. Rosenblum, ed. New York; Academic Press, 1975.

Harcourt, A.H.; Stewart, K.J. Functions of alliances in contests within wild gorilla groups. BEHAVIOUR 109:176–190, 1989.

Harcourt, A.H.; Stewart, K.J.; Harcourt, D.E. Vocalizations and social relationships of wild gorillas: A preliminary analysis. Pp. 346–356 in CURRENT PERSPECTIVES IN PRIMATE SOCIAL DYNAMICS. D.M. Taub; F.E. King, eds. New York, Van Nostrand Rheinold, 1986.

Harcourt, A.H.; Stewart, K.J.; Hauser, M.D. Functions of wild gorilla 'close' calls. I. Repertoire, context, and interspecific comparison. BEHAVIOUR 124:89–122, 1993.

Hienz, R.D.; Brady, J.V. The acquisition of vowel discriminations by nonhuman primates. JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 84:186–194, 1988.

Hinde, R.A. Animal signals: ethological and games theory approaches are not incompatible. ANIMAL BEHAVIOUR 29:535–542, 1981.

Kojima, S. Comparisons of the auditory functions in the chimpanzee and human. FOLIA PRIMATALOGICA 55:62–72, 1990.

Ladefoged, P. A COURSE IN PHONETICS, 2nd edition. New York, Harcourt, Brace, Jovanovich, 1982.

Lieberman, P. Primate vocalizations and human linguistic ability. JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 44:1,574–1,584, 1968.

Lieberman, P. THE BIOLOGY AND EVOLUTION OF LANGUAGE. Cambridge, Harvard University Press, 1984.

Lillehei, R.A.; Snowdon, C.T. Individual and situational differences in the vocalizations of young stumptail macaques. BEHAVIOUR 65:270–281, 1978.

Markel, J.D.; Gray, A.D. LINEAR PREDICTION OF SPEECH. Berlin, Springer-Verlag, 1976.

Marler, P. Social organization, communication, and graded signals: The chimpanzee and the gorilla. Pp. 342–399 in GROWING POINTS IN ETHOLOGY. P.P.G. Bateson; R.A. Hinde, eds. Cambridge, Cambridge University Press, 1976a.

Marler, P. An ethological theory of the origin of vocal learning. ANNALS OF THE N.Y. ACADEMY OF SCIENCES 208:708–717, 1976b.

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

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

Owren, M.J. Applying linear predictive coding (LPC) to frequency spectrum analysis of animal acoustic signals. In RECENT TECHNICAL ADVANCES IN THE STUDY OF ANIMAL ACOUSTIC COMMUNICATION. S. Hopp; C. Evans, eds. Berlin, Springer-Verlag, in press.

Owren, M.J.; Hopp, S.L.; Seyfarth, R.M. Categorical vocal signaling in nonhuman primates. Pp. 102–122 in NONVERBAL VOCAL COMMUNICATION: COMPARATIVE AND DEVELOPMENTAL APPROACHES. H. Papousek; U. Jurgens; M. Papousek, eds. Cambridge, Cambridge University Press, 1992.

Rosen, S.; Fourcin, A. Frequency selectivity and the perception of speech. Pp. 373–487 in FREQUENCY SELECTIVITY AND HEARING. B.C.J. Moore, ed. New York, Academic Press, 1986.

Seyfarth, R.M.; Cheney, D.L. The acoustic features of vervet monkey grunts. JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 75:1,623–1,628, 1984.

Sinnott, J.M.; Owren, M.J.; Petersen, M.R. Auditory frequency discrimination in primates: Species differences (*Cercopithecus, Macaca, Homo*). JOURNAL OF COMPARATIVE PSYCHOLOGY 101:126–131, 1987.

Smith, W.J. THE BEHAVIOR OF COMMUNICATING. Cambridge, Harvard University Press, 1977.

Snowdon, C.T. Communication as social interaction: Its importance in ontogeny and adult behavior. Pp. 108–122 in PRIMATE VOCAL COMMUNICATION. Todt, P. Goedeking; D. Symmes, eds. New York, Springer-Verlag, 1988.

Stewart, K.J.; Harcourt, A.H. Gorillas: Variation in female relationships. Pp. 155–164 in PRIMATE SOCIETIES. B.B. Smuts; D.L. Cheney; R.M. Seyfarth; R.W. Wrangham; T.T. Struhsaker, eds. Chicago, University of Chicago Press, 1987.

Symmes, D.; Biben, M. Conversational vocal exchanges in squirrel monkeys. Pp. 123–132 in PRIMATE VOCAL COMMUNICATION. D. Todt; P. Goedeking; D. Symmes, eds. New York, Springer-Verlag, 1988.

Vedder, A.W. Movement patterns of a group of free-ranging mountain gorillas (*Gorilla gorilla beringei*) and their relation to food availability. AMERICAN JOURNAL OF PRIMATOLOGY 7:73–88, 1984.

Watts, D.P. Strategies of habitat use by mountain gorillas. FOLIA PRIMATOLOGICA 56:1–16, 1991.

Wier, C.; Jesteadt, W.; Green, D. Frequency discrimination as a function of frequency and sensation level. JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA 61:178–184, 1977.

Yamagiwa, J. Intra- and inter-group interactions of an all-male group of Virunga mountain gorillas (*Gorilla gorilla beringei*). PRIMATES 28:1–30, 1987.

Zoloth, S.R.; Petersen, M.R.; Beecher, M.D.; Green, S.; Marler, P.; Moody, D.B.; Stebbins, W.C. Species-specific perceptual processing of vocal sounds by monkeys. SCIENCE 204:870–872, 1978.