Fundamental frequency declination is not unique to human speech: Evidence from nonhuman primates

Marc D. Hauser
Animal Communication Laboratory, Department of Zoology, University of California, Davis, California 95616-8761

Carol A. Fowler
Haskins Laboratories, New Haven, Connecticut 06511 and Department of Psychology, Dartmouth College, Hanover, New Hampshire 03755

(Received 9 September 1990; accepted for publication 10 September 1991)

In human speech, declination of the fundamental frequency (F_0) of the voice spans coherent units of an utterance and, therefore, signals where units begin and end. A rapid final fall at the end of an utterance provides a further indication of an utterance's ending. The occurrence of declination is sufficiently widespread across languages that several investigators have suggested it as a language universal. Language universals may be universal because they are part of a species-specific specialization for language or, alternatively, they may constitute conventionalizations of natural dispositions of the vocal tract that may serve a communicative function. Evidence is offered favoring the latter account for declination and the final fall by showing that vocal productions of vervet monkeys (*Cercopithecus aethiops*) and rhesus macaques (*Macaca mulatta*) show declination, and vervets show clear evidence of a final fall. Interestingly, the fall in F_0 may serve some communicative role in the vocal exchanges of vervets and rhesus, analogous to its signalling function in human language.

PACS numbers: 43.70.Bk, 43.70.Fq, 43.80.Ka

INTRODUCTION

Speech exhibits a number of low-level phonetic regularities that are widespread, if not universal, across languages. Examples include durational lengthening of vowels before voiced as compared to voiceless consonants (e.g., Chen, 1970), devoicing of word-final obstruents (see Locke, 1983, for a review), and declination of fundamental frequency (F_0) in a phrase or utterance (Bolinger, 1978; Chen, 1970). These low-level regularities also serve as resources for the development of phonological processes. [An occurrence that MacNeilage and Ladefoged (1976) call "triggering."] The following phonological regularities are parallel to the foregoing phonetic ones. By some accounts (e.g., de Chene, 1985; MacNeilage and Ladefoged, 1976), English is developing a phonological vowel length distinction before voiced and voiceless consonants as the latter distinction weakens, at least for intervocalic /d/ and /t/; German and Russian have final devoicing rules, whereby word-final consonants must be unvoiced (Kenstowicz and Kisseberth, 1979); it has been proposed that English has intonational "downstep" rules (Pierrehumbert, 1980), and some tone languages have "downdrift" rules for lexical tones produced in sentences (Hyman, 1973).[1]

Some linguistic universals are ascribed to a special-to-human language faculty whereas others are ascribed to their functional utility in communicative exchanges [e.g., Comrie (1981) for a discussion of explanations for language universals]. Two functional explanations offer more plausible accounts of the low-level regularities described above than does an account in terms of a universal language faculty. First, the regularities may reflect dispositional features of the vocal-tract—that is, regularities that are easier to allow to occur than they are to inhibit or to offset. Second, they may enhance the perceptual distinctiveness of otherwise similar phonetic properties in speech. A dispositional account of final devoicing is likely (e.g., Flege and Brown, 1982), and, compatibly, it is characteristic of infants' prelinguistic babbles (Oller *et al.*, 1976). Both articulatory (e.g., Chen, 1970) and perceptual accounts (Javkin, 1976; Kluender *et al.*, 1988; Fowler, in press) of the relation between vowel duration and voicing have been proposed. Likewise, both speaker-based (Gelfer, 1987; Maeda, 1976) and listener-based (Breckenridge, 1977; Cooper and Sorenson, 1981) accounts have been proposed to explain declination. As to the latter, Breckenridge (1977) and Cooper and Sorenson (1981) propose that listeners use F_0 resettings at major phrase boundaries to help them recognize the boundaries as such. Breckenridge speculates, in addition, that listeners extrapolate from the declining contour to predict how long the remainder of the utterance will be.

Our research focuses on declination and on a rapid fall in F_0 that occurs in the terminal portion of an utterance, called the "final fall." In the case of each pattern, most likely both speaker-based and listener-based accounts are correct. Declination is likely to be dispositional in part. For example, in a recent account, 't Hart *et al.* (1990) show that air pressure below the vocal folds (subglottal pressure or P_S)—one of the variables affecting F_0—will tend to fall dispositionally over the course of an expiration. In particular, P_S will fall, and, other things equal, F_0 will fall with it if there is a lag between expiration of some volume of air from the lungs and

 0001-4966/92/010363-07$00.80

the initiation of maneuvers to reduce thoracic-cavity volume. Whereas the reduction in thoracic-cavity volume offsets the lowering of P_S that the reduction in lung volume would otherwise cause, the compensation will be partial given a time lag between the maneuvers. However, as 't Hart *et al.* point out, neither P_S nor F_0 *need* fall during an utterance, and, indeed, some researchers have found flat or nearly flat P_S and F_0 contours in speech (e.g., Ladefoged, 1967; Lieberman, 1967). No one, to our knowledge, has reported a general tendency for F_0 to rise, although this is physiologically possible as well. Accordingly, the two commonly reported patterns are flat or falling, suggesting that either the dispositional tendency for F_0 to fall is counteracted, or it is not.

Other findings reveal that declination is not wholly a dispositional consequence of lung deflation. Lung deflation can be either accelerated or slowed by production of consonants that are associated with characteristically either high (e.g., /fa/) or low (e.g., /ma/) airflow during closure. Gelfer and colleagues (Gelfer *et al.*, 1987) found comparable P_S declinations for utterances composed of reiterant /fa/ or /ma/ despite marked differences in lung volume declination. This indicates that talkers offset the effects on P_S of the rapid decrease in lung volume caused by reiterant /fa/ production; most likely, they did so by recruiting action of the internal intercostal muscles of the respiratory system, which foster expiration (e.g., Ladefoged, 1967). P_S nonetheless did fall over both sets of utterances suggesting, perhaps, that talkers ensure sufficient P_S out to the end of a phrase or utterance, but they do not always attempt to keep P_S level. An additional active contribution to declination is tensing of the cricothyroid muscle of the larynx at the start of a declination contour (Gelfer, 1987). The cricothyroid stretches the vocal folds and increases F_0; therefore, tensing it at the beginning of a declination contour enhances the already high F_0 at the beginning of a phrase or utterance. Similarly, Cooper and Sorenson (1981) reported that their speakers sometimes "reset" the declination contour (that is, they raised F_0 to begin a new declination contour) at the beginnings of major syntactic units where they did not take a breath. In these cases, the resetting must be active; presumably it is designed to signal the beginnings of major syntactic and/or semantic units to listeners.

The following integrated account of declination may handle most of the findings. The origins of declination in vocal production derive from expiration, its effects on P_S, and the effects of P_S on F_0 that will occur unless the talker actively intervenes. However, because talkers do not take breaths (and hence reset F_0) at random locations in an utterance (i.e., they tend to take breaths at locations where they stop to plan ahead, and hence, between major semantic and/or syntactic units) the declination contour tends to span integral units in an utterance. Because it spans integral components of an utterance, it provides redundant information to listeners as to where such units start and end. Insofar as listeners use that information as previously suggested, talkers may tend to provide it intentionally, enhancing the F_0 rise at the beginnings of units and resetting F_0 even at the beginnings of units where they do not take a breath. Having been conventionalized in these ways, the contour is readily noticed by listeners, and this may pave the way for "triggering" in some language communities that then conventionalize the contours further by developing downstepping intonational or tonal patterns.

As for the utterance-final fall in F_0, the rather rapid rate of the fall has suggested relaxation of laryngeal muscles rather than a reduction in P_S to some researchers (Liberman and Pierrehumbert, 1984). As such, of course, it is a natural consequence of utterance termination. However, it can begin rather early in an utterance, and this may suggest augmentation of any dispositional consequences of ending an utterance. Silverman (1987) suggests that it may begin as much as two seconds before utterance termination. Perhaps because the fall can provide useful information to listeners, speakers enhance the fall just as they may enhance the declination contour. Indeed, there is evidence that listeners use a precipitous fall in F_0 as information that a speaker's conversational turn is ending (Beattie *et al.*, 1982; Beattie, 1983). Consequently, falls in the middle of an utterance foster interruption by other participants in the conversation.

If, at its origins, declination and the final fall are dispositional features of vocal production, they may occur widely not only in spoken languages (e.g., Bolinger, 1978; Cooper and Sorenson, 1981; Ohala, 1978), but also in vocal productions of some nonhuman animals that have lungs, a larynx, and voiced calls sufficiently long in duration for lung deflation to have a noticeable effect on F_0.

We should point out before pursuing this idea, however, that some investigators disagree that declination is common in spontaneous speech by human talkers. Lieberman (1967) reports that the "unmarked breath group," produced on standard declarative sentences, has a flat P_S and F_0 out to its end where there is the terminal fall. Using a regression analysis to get a declination slope, Lieberman and his colleagues (Lieberman *et al.*, 1985a) found negative slopes (that is, decreases in F_0 over time) for over 70% of read sentences, but for just 55% [in fact, 63% as 't Hart (1986) points out] of spontaneous utterances. They conclude that declination is not particularly characteristic of spontaneous speech (see also, Umeda, 1982). The methods used by Lieberman *et al.* have been severely criticized (see Repp, 1985, and 't Hart, 1986), and responses to them by Lieberman and colleagues (Lieberman *et al.*, 1985b; Lieberman, 1986). We will not review the controversy here, but only make the following three arguments in favor of declination as a dispositional contour.

First, there are reasons other than that declination is peculiar to read speech as to why declining F_0 contours may be more readily observed in read than spontaneous speech. The surface F_0 contour of any utterance includes contributions from sources other than declination—from the component consonants and vowels of the utterance [that is, variation due to consonant voicing (e.g., Hombert, 1978; Ohde, 1984) and intrinsic vowel F_0 (e.g., Silverman, 1987)] from the intonation contour, which may, but need not be downstepping and even from the emotional tone of the utterance (e.g., Scherer *et al.*, 1984). Neither declination nor the segmental composition of speech is likely to differ between spontaneous and read speech. The intonation contours used, however, as well as the emotional tone of the utterances, are likely to differ. Remez *et al.* (1986) in fact report that spontaneous speech shows more variation in F_0 than does the same speech read aloud. The more sources of F_0 variation that converge with a tendency for F_0 to decline, the more difficult it will be for regression analyses as employed by Lieberman *et al.* (1985a) to detect declination if present. That is, in the absence of independent variables in the regression models to factor out those other influences on F_0, variability that they contribute will serve as noise in the analysis, reducing the detectability of declination. Second, as we indicated earlier, investigators have reported that listeners expect a fall in F_0, and would be surprised if F_0 were not typically declining.[2] Consistent with this claim, 't Hart *et al.* (1990) report that synthesized utterances with a declining F_0 contour are given higher naturalness ratings by listeners than are other contours having no declination. These other contours included utterances having overall larger F_0 falls than rises, but no general downward tilt apart from that. In any case, our claim is that a fall in F_0 is dispositional, not that it must occur. Even the utterances reported by Lieberman *et al.* (1985a), which are least consistent with the declination

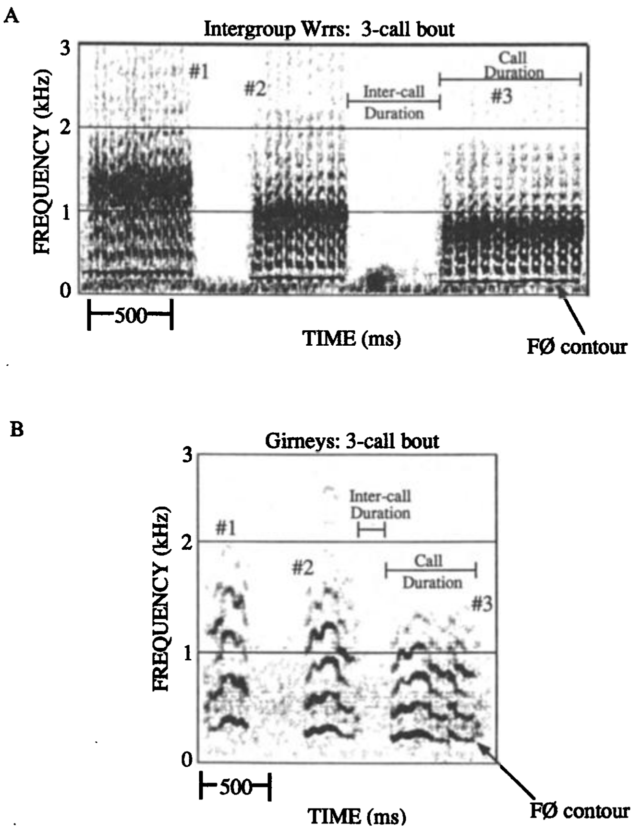

FIG. 1. (a) Narrow-band spectrogram of a three-call bout of intergroup wrrs showing F_0 declination. The F_0 contour of each call is indicated by a dark line. (b) Narrow-band spectrogram of a three-call bout of girneys showing F_0 declination. The lowest band of energy in each call is the F_0 contour. Calls were sampled at 5 kHz and a 256-*pt* FFT used to generate the spectrograms (frequency resolution = 19.5 Hz, time resolution = 51.2 ms).

hypothesis (spontaneous speech with terminal falls removed), contain more negative than positive slopes and a modal value to the left of zero.

In the present research, we investigated the possibility that vocalizations of wild vervet monkeys (*Cercopithecus aethiops*) and rhesus macaques (*Macaca mulatta*) would exhibit declination and a final fall in F_0. Although there are fundamental differences in the vocal tracts of humans and nonhuman primates—particularly in the height of the larynx and, therefore, the relatively small size of the pharynx in nonhuman primates (e.g., Lieberman, 1975; Negus, 1949)—both primates appear to produce most of their vocalizations on an expiratory airflow, and frequently their productions are voiced (Lieberman, 1985; Hauser, 1989). Accordingly, physiological conditions described earlier as favoring declination and the final fall in humans are generally present in these nonhuman primates as well. Having obtained some positive evidence for declination and the fall in F_0 for vocalizations of vervets and rhesus macaques, we attempted to draw inferences as to whether either feature serves a communicative function in exchanges among conspecifics.

To test for declination, we looked for two characteristics of F_0 as it is described in the literature on human speech. First, and most importantly, was there a consistent fall in F_0? Second, was there a tendency for F_0 to start higher in longer utterances, a pattern that is sometimes (e.g., Bruce, 1982; Cooper and Sorenson, 1981), but not always (e.g., Maeda, 1976) found in human speech? To test for the final fall, we asked whether there was evidence of a marked terminal fall in F_0 and whether there was reduced variability in the terminal as compared to the initial F_0 across utterances of different lengths (e.g., Maeda, 1976; Cooper and Sorenson, 1981).

I. MATERIALS AND METHODS

A. Subjects and study area

Six vervet monkey (*Cercopithecus aethiops*) groups were observed in Amboseli National Park, Kenya from August 1983 to June 1985. Approximately 3000 h of behavioral and vocal observations were collected during this period (see Hauser, 1987, 1988, 1989; Cheney and Seyfarth, 1990 for details of the study area and sampling procedures), and both males and females of all ages were sampled. Acoustic analyses focused on the "intergroup wrr" [Fig. 1(a)], a vocalization given during aggressive interactions with a neighboring group (Hauser, 1989).

One group of rhesus macaques (*Macaca mulatta*) living on Cayo Santiago, Puerto Rico (Rawlins and Kessler, 1987) was observed from November 1988 to June 1989. Approximately 1200 h of behavioral and vocal data were collected from adult males and adult females. Acoustic analyses were conducted on the "girney" [Green, 1975; Masataka, 1988; Fig. 1(b)], a call produced during the initiation or maintenance of affiliative social interactions.

For both vervet monkeys and rhesus macaques, vocalizations were recorded with a Sony TC-D5M cassette recorder and a Sennheiser MKH816 directional microphone (foam windscreen and K3U power module). Subjects in both study areas were well habituated to the presence of human observers and thus, most vocalizations were recorded at a distance of 0.5–2 m.

There were three reasons why the intergroup wrr and girney were selected for analyses. First, both calls were often given in pure bouts (i.e., one call type per string of calls). Second, both calls are given during vocal "volleys" or exchanges, and thus appeared to be communicative. Third, both calls tend to have exemplars in which F_0 can be readily resolved from the power spectrum. The particular tokens selected for analyses consisted of data regarding caller identity as well as the identity of all individuals involved in the vocal interaction. Tokens employed in the analyses were limited to those vocalizations in which descriptions of the context and identification of the caller were unambiguous, and the quality of the recording was high. For some individuals, two or more call bouts were available and consequently, means were taken for the purpose of statistical analysis. Descriptions of the context were often incomplete due to poor observation conditions (e.g., dense vegetation) and vocalizations were commonly interfered with by environmental noise (e.g., other animals vocalizing in Kenya and helicopters from a naval base in Puerto Rico). Consequently, although a large number of vocalizations were recorded during the study period, only a limited sample could be used for acoustic analyses.

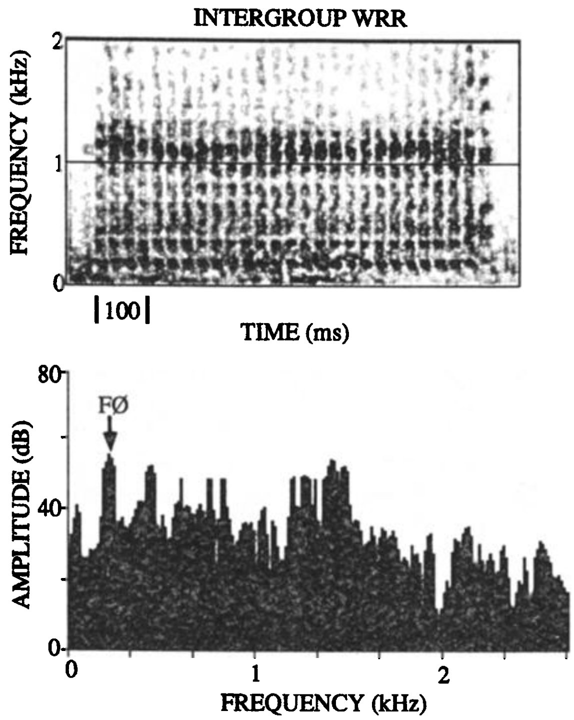

FIG. 2. Top: Narrow-band spectrogram of a vervet intergroup wrr. Bottom: Power spectrum, calculated at the midpoint of the intergroup wrr, with the F_0 peak marked.

B. Acoustic analyses

Acoustic analyses were performed using the SIGNAL digital sound analysis program (Beeman, 1990), which operates on an IBM-compatible 80386 computer. Each analyzed call was first low-pass filtered at 2 kHz and then sampled at 5 kHz. Spectrographic and power spectrum displays were obtained with a 1024-point Fourier transform and a Hanning window; these settings provided a frequency resolution of 5 Hz. F_0 values were obtained from power spectra calculated over a 50-ms interval at the center of each call. As Figs. 2 and 3 illustrate, extracting F_0 from these calls was relatively unambiguous since the F_0 peak was distinct and was readily confirmed by inspection of the waveform. Although wrrs and girneys are given in bouts of one to eight calls, only two and three call bouts are considered; these intermediate bout lengths predominate in the data set.

II. RESULTS AND DISCUSSION

A. Declination in vervet monkeys

Adult vervets ($n = 8$) producing two-call bouts [$n = 23$; Fig. 4(a)] show a significant decrease in F_0 from call one to call two (Wilcoxon signed ranks test: $z = 2.52$, $p < 0.006$). The fall is consistent with declination or with the terminal fall characteristic of human speech. For adults ($n = 11$) producing three-call bouts [$n = 26$; Fig. 4(a)], there was a significant decline from call one to two ($z = 1.98$, $p < 0.02$) and from call two to three ($z = 2.40$, $p < 0.008$). The early fall is consistent with declination, whereas the final fall may reflect a terminal fall in addition. Ninety-eight percent of all adult bouts ($n = 49$) show a decline in F_0 from the first to the last call in the bout. The total

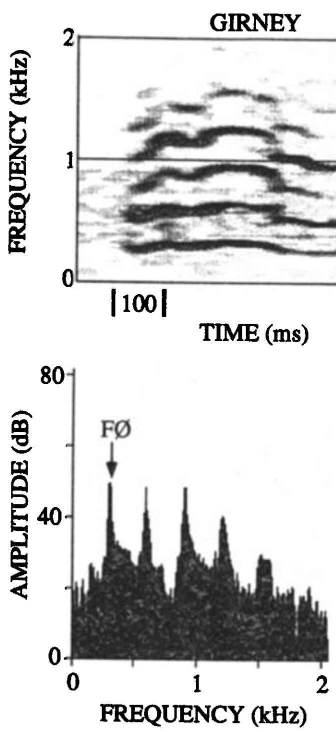

FIG. 3. Top: Narrow-band spectrogram of a rhesus macaque girney. Bottom: Power spectrum, calculated at the midpoint of the girney, with the F_0 peak marked.

(a)

(b)

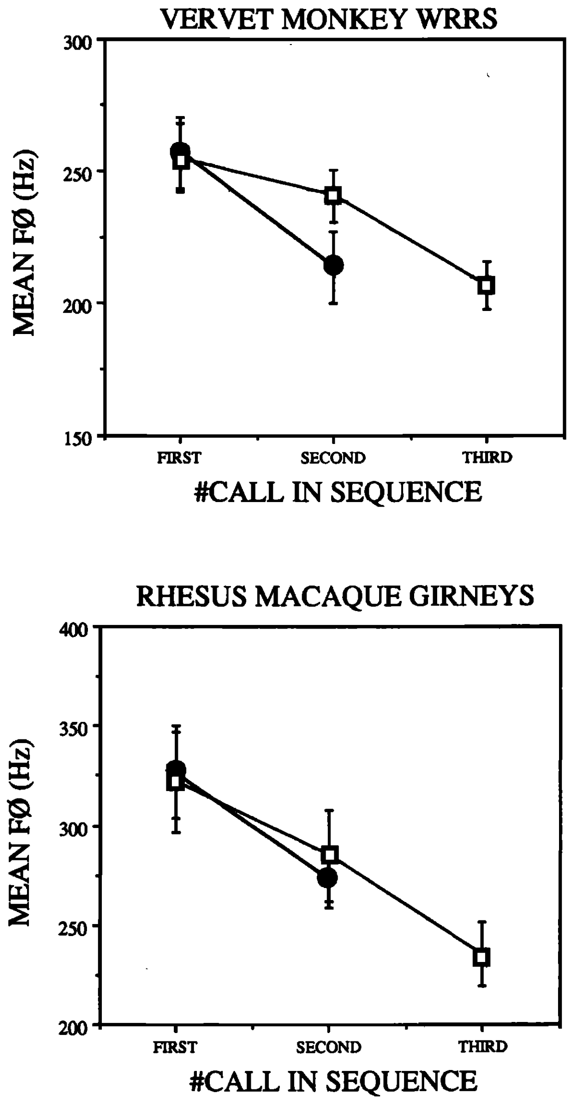

FIG. 4. Changes in F_0 for two-call (closed circles) and three-call bouts (open squares) for (a) vervet intergroup wrrs and (b) rhesus macaque girneys. Standard deviations shown. *Y* axis shows the mean fundamental frequency (F_0 in Hz) and the *x* axis represents the number of the call within the call bout.

fall in F_0 was invariant over two-call (50 Hz) and three-call (51 Hz) bouts—a finding that some investigators (e.g., Maeda, 1976) but not others (e.g., Sorenson and Cooper, 1980) report as characteristic of declination. In contrast to Maeda, Sorenson and Cooper report a higher starting F_0 for longer utterances and so a greater F_0 fall (to an invariant terminal value) for longer utterances. There was a significant correlation between magnitude of the fall in F_0 and bout duration for two-call bouts (rho $= 0.51$, $p < 0.01$), but not three-call bouts (rho $= 0.29$, $p > 0.05$). Similarly, the correlation between starting F_0 and bout duration was positive but not significant. In contrast to adults, juveniles and infants failed to show a consistent pattern of F_0 change in either two- or three-call bouts. However, it is worth pointing out that the inter-call interval was significantly ($z = 7.17$, $p < 0.0001$) longer for juvenile and infant bouts (mean = 623.6 ms, s.d. = 169.5, $n = 19$) than for adult bouts (mean = 273.4, s.d. = 82.8, $n = 49$), suggesting that young animals may have been taking breaths between calls.

Looking now at evidence specifically relating to the terminal fall in F_0, among adults, the evidence is supportive of a terminal fall in two respects. In 78% of three-call bouts, the decline in F_0 between calls two and three was greater than between calls one and two ($z = 2.36$, $p = 0.02$). In addition, there was greater variation (CV = coefficient of variation) in F_0 at the start of the bout (two-call bouts: CV = 26.5%; three-call bouts: CV = 25.9%) than at the end of the bout (two-call bouts: CV = 13.2%; three call bouts: CV = 22.0%).

B. Declination in rhesus macaques

For adult rhesus ($n = 11$), two-call bouts [$n = 24$; Fig. 4(b)] showed a significant decline in F_0 from call one to two ($z = 2.19$, $p < 0.02$). For adults ($n = 6$) producing three-call bouts [$n = 20$; Fig. 4(b)], there was a significant F_0 decline from call one to two ($z = 2.18$, $p < 0.01$) and from call two to three ($z = 2.20$, $p < 0.01$). 72% of all bouts showed a decline from the first to the last call ($z = 2.20$, $p < 0.03$). The total fall in F_0 was nonsignificantly lower for two-call (87 Hz) than for three-call bouts (104 Hz). For neither bout type was there a significant correlation between bout duration and magnitude of the fall or starting F_0.

Turning to evidence for a terminal fall, the evidence is mixed. Just 60% of all three-call bouts showed a greater decline in F_0 between calls two and three than between calls one and two. The difference in the magnitude of the fall was nonsignificant. However, there was slightly greater variation in F_0 at the start of the bout (two-call bouts: CV = 29.5%; three-call bouts: CV = 34.8%) than at the end of the bout (two-call bouts: CV = 23.2%; three call bouts: CV = 20.5%).

C. Does declination serve a communicative function in nonhuman primates?

Thus far, studies of conversational exchanges in nonhuman animals have suggested that temporal cues are of primary importance in determining turn-taking (e.g., Biben *et al.*, 1986; Biben and Masataka, 1987; Bush and Narins, 1989). To determine whether vervets or macaques use frequency information to identify bout termination as humans may do (Beattie *et al.*, 1982), we examined "interruptions" of bouts by other animals. Interruptions were defined as the production of a vocalization by one individual that overlapped with the vocalization of another individual. All of the interruptions occurred toward the terminal portion of the call and therefore did not interfere with our calculations of F_0. The individual responsible for the interruption had to be involved, either directly or indirectly, with the initial communicant. For vervets, we predicted that young animals should be interrupted more frequently than adults, because they had shown no consistent pattern of F_0 change throughout a bout. In agreement with our expectations, only one of

49 adult vervet bouts was interrupted, whereas 6 of 19 bouts by young vervet monkeys were interrupted. Of course, young animals may be interrupted more frequently than adults for reasons other than that listeners cannot determine when a bout is completed. However, five of the six interrupted calls showed precipitous falls (range 36–58 Hz) to a near terminal value between the first two calls of a three-call bout. Only one of the 13 uninterrupted calls showed a similar pattern. Once individuals were interrupted, there were no consistent changes in F_0 for the calls that followed. That is, there was no evidence of resetting. Those bouts without interruptions showed no consistent pattern of F_0 change.

For rhesus, although developmental data were not available, interruptions of adult calls were more common than for vervets: of 44 girney bouts, 13 were interrupted. When individuals were interrupted, either between the first and second call ($n = 10$) or between the second and third ($n = 3$), the F_0 for the call preceding the interruption was lower than the call following the interruption in eight cases. However, it was not lower than F_0 in comparable locations in noninterrupted calls; accordingly, adult rhesus, in contrast to young vervets, were apparently not interrupted because they had inappropriately signaled utterance termination. F_0 declined, as expected, for bouts without interruption. Thus there is some evidence that individuals reset following an interruption. Moreover, the intercall duration for interrupted and noninterrupted bouts was not statistically different, and thus individuals do not appear to be starting a new bout by taking a second breath following interruption.

Is the decline in F_0 the only acoustic parameter that predicts bout termination and interruptions? For wrrs, there was no evidence that other acoustic measures, such as maximum or minimum frequency of the call, bout duration, call duration and intercall duration, signaled bout termination. The same was true of girneys with the exception that the second call in two-call bouts was significantly longer than the first call ($z = 2.2$, $p < 0.01$). The previously listed variables also failed to explain the presence/absence or timing of interruptions.

Our findings suggest two conclusions. First, the fall in F_0 that is universal or nearly so in human vocal communication may be even more widespread still. Declination occurs reliably in the vocal communication of vervet monkeys and rhesus macacques; the final fall was present reliably only in the vocalizations of the vervets. Possibly, both features are disposed to occur in the vocalizations of any animal with a respiratory system and larynx similar to those of humans and a tendency to vocalize on an expiratory airflow. A second conclusion is that a high starting F_0 and a low final F_0 can provide information in human and nonhuman primate vocal communication about utterance beginnings and endings. Speakers and listeners may use such information to guide conversational "turn taking."

ACKNOWLEDGMENTS

Marc Hauser thanks the Office of the President, the Ministry of Wildlife and Tourism, and the Institute of Primate Research for permission to work in Amboseli National Park, Kenya. Thanks also to M. Kessler and J. Berard of the Carribean Primate Research Center, Puerto Rico for access to facilities on Cayo Santiago. We thank J. 't Hart, P. Marler, R. Ohde, K. Silverman, M. Studdert-Kennedy, and an anonymous reviewer for their comments on the manuscript. This research was supported by grants from the Wenner–Gren Foundation, the National Geographic Society (4251-90; co-PI: P. Marler) and a postdoctoral research fellowship from the National Institute of Health (HD-07213). Carol Fowler was supported by NINCDS Grant NS-13617 to Haskins Laboratories.

[1] An example of a downstepping intonation contour is that associated with spoken lists (Pierrehumbert, 1980; Liberman and Pierrehumbert, 1984). In this contour, a level F_0 is maintained on unaccented syllables (where the unaccented syllables are associated phonologically with low tones), whereas decreases in F_0 occur on successive accented syllables (where the accents are realized phonologically by high tones). The decrease in frequency is identifiably not due to declination, because of the way F_0 falls—in clear steps rather than gradually. Tonal downdrift occurs, for example, in Akan (Kenstowicz and Kisseberth, 1979), where a high tone separated from a previous high tone by a low tone is stepped down in frequency.

[2] Other interpretations of these findings have been offered. Silverman (1987) interprets Pierrehumbert's findings and his own differently from the interpretation offered in Pierrehumbert (1979). He proposes that listeners expect the terminal fall at the end of a speech utterance—not necessarily a gradual fall from utterance onset to the terminal fall. An inference from Lieberman's (1967) breath-group theory is that the first peak of the two in Pierrehumbert's contours was interpreted as a prominence peak ($[+P]$) that would cause sufficient lung deflation that F_0 is expected to be particularly low utterance finally. Whereas both of these interpretations are possible for Pierrehumbert's findings, neither apparently explains a finding of Leroy's [1984; cited in 't Hart *et al.* (1990); see also Terken, 1991] that listeners show larger compensations for declination in utterances having falling as compared to monotonous baselines.

Beattie, G. (**1983**). *Talk: An Analysis of Speech and Nonverbal Behavior* (Open U. P., Milton Keyes).

Beattie, G., Cutler, A., and Pearson, M. (**1982**). "Why is Mrs. Thatcher interrupted so often?" Nature **300**, 744–747.

Beeman, K. (**1990**). *SIGNAL User's Guide* (Engineering Design, Belmont, MA).

Biben, M., and Masataka, N. (**1987**). "Temporal rules regulating affiliative vocal exchanges of squirrel monkeys," Behaviour **101**, 311–319.

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

Bolinger, D. (**1978**). "Intonation across languages," in *Universals of Human Language: Phonology*, edited by J. Greenberg (Stanford U. P., Stanford), pp. 471–524.

Breckenridge, J. (**1977**). *Declination as a Phonological Process* (Bell Lab. Technical Memo, Murray Hill, New Jersey).

Bruce, G. (**1982**). "Developing the Swedish intonation model," Phonetics Working Papers, University of Lund **22**, 51–116.

Bush, J., and Narins, P. (**1989**). "Chorus dynamics of a neotropical amphibian assemblage: comparison of computer simulation and natural behavior," Anim. Behav. **37**, 33–44.

Chen, M. (**1970**). "Vowel length variation as a function of the voicing of the consonant environment," Phonetica **22**, 129–159.

Cheney, D. L., and Seyfarth, R. M. (**1990**). *How Monkeys See the World* (Chicago U. P., Chicago).

Comrie, B. (**1981**). *Language Universals and Linguistic Typology* (Chicago U. P., Chicago).

Cooper, W., and Sorenson, J. (**1981**). *Fundamental Frequency in Sentence Production* (Springer-Verlag, New York).

de Chene, B. (**1985**). *The Historical Phonology of Vowel Length* (Garland, New York).

Flege, J. E., and Brown, W. S. (**1982**). "The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance," J. Phon. **10**, 335–345.

Fowler, C. A. (**1991**). "Vowel duration and closure duration in voiced and unvoiced stops: There is no contrast effect here," J. Phon. (in press).

Gelfer, C. (**1987**). "A simultaneous physiological and acoustic study of fundamental frequency declination," Ph.D. thesis, City University of New York, New York.

Gelfer, C., Harris, K., and Baer, T. (**1987**). "Controlled variables in sentence intonation," in *Larnyngeal Function in Phonation and Respiration*, edited by T. Baer, C. Sasaki, and R. C. Scherer (Denver Center for the Performing Arts, Inc., Denver), pp. 113–126.

Green, S. (**1975**). "Variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): a field study," in *Primate Behavior*, edited by L. Rosenblum (Academic, New York), pp. 1–102.

't Hart, J. (**1986**). "Declination has not been defeated—A reply to Lieberman *et al.*," J. Acoust. Soc. Am. **80**, 1838–1840.

't Hart, J., Collier, R., and Cohen, A. (**1990**). *A Perceptual Study of Intonation* (Cambridge U. P., Cambridge).

Hauser, M. D. (**1987**). "Behavioral ecology of vervet monkeys: proximate and ultimate explanations," Ph.D, University of California, Los Angeles.

Hauser, M. D. (**1988**). "How infant vervet monkeys learn to recognize starling alarm calls: the role of experience," Behaviour **105**, 187–201.

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

Hombert, J.-M. (**1978**). "Consonant types, vowel quality and tone," in *Tone: A Linguistic Survey*, edited by V. Fromkin (Academic, New York), pp. 77–112.

Hyman, L. (**1973**). "Consonant types and tone," South. Calif. Occas. Pap. Linguistics **1**, 152–179.

Javkin, H. R. (**1976**). "The perceptual basis of vowel duration differences associated with the voiced/voiceless distinction," Rep. Phon. Lab. Univ. Calif., Berkeley **1**, 78–89.

Kenstowicz, M., and Kesseberth, M. (**1979**). *Generative Phonology* (Academic, New York).

Kluender, K. R., Diehl, R. L., and Wright, B. A. (**1988**). "Vowel-length differences before voiced and voiceless consonants: An auditory explanation," J. Phon. **16**, 153–169.

Ladefoged, P. (**1967**). *Three Areas of Experimental Phonetics* (Oxford U. P., London).

Leroy, L. (**1984**). "The psychological reality of fundamental frequency declination," Antwerp Pap. Linguistics **40** (University of Antwerp, Belgium).

Liberman, M., and Pierrehumbert, J. (**1984**). "Intonational invariance under change in pitch range and length," in *Language Sound Structure*, edited by M. Aronoff and R. Oehrle (MIT, Cambridge, MA), pp. 157–233.

Lieberman, P. (**1967**). *Intonation, Perception and Language* (MIT, Cambridge, MA).

Lieberman, P. (**1975**). *On the Origin of Language* (Macmillan, New York).

Lieberman, P. (**1985**). "The physiology of cry and speech in relation to linguistic behavior," in *Infant Crying*, edited by B. M. Lester and C. F. Z. Boukydis (Plenum, New York).

Lieberman, P. (**1986**). "Alice in declinationland—A reply to Johan 't Hart," J. Acoust. Soc. Am. **80**, 1840–1842.

Lieberman, P., Katz, W., Jongman, A., Zimmerman, R., and Miller, M. (**1985a**). "Measures of the sentence intonation of read and spontaneous speech in American English," J. Acoust. Soc. Am. **77**, 649–657.

Lieberman, P., Katz, W., Jongman, A., Zimmerman, R., and Miller, M. (**1985b**). "Reply to Bruno H. Repp," J. Acoust. Soc. Am. **78**, 1116–1117.

Locke, J. (**1983**). *Phonological Acquisition and Change* (Academic, New York).

MacNeilage, P., and Ladefoged, P. (**1976**). "The production of speech and language," in *Handbook of Perception: Language and Speech*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 75–120.

Maeda, S. (**1976**). "A characterization of American English intonation," Ph.D thesis, MIT, MA.

Masataka, N. (**1988**). "Motivational referents in Japanese monkey communication," Ethology **80**, 265–273.

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

Ohala, J. (**1978**). "Production of tone," in *Tone: A Linguistic Survey*, edited by V. Fromkin (Academic, New York), pp. 5–39.

Ohde, R. (**1984**). "Fundamental frequency is an acoustic correlate of stop consonant voicing," J. Acoust. Soc. Am. **75**, 224–230.

Oller, D. K., Wieman, L., Doyle, W., and Ross, C. (**1976**). "Infant babbling and speech," J. Child Lang. **3**, 1–11.

Pierrehumbert, J. (**1979**). "The perception of fundamental frequency," J. Acoust. Soc. Am. **66**, 363–369.

Pierrehumbert, J. (**1980**). "The phonology and phonetics of English intonation," Ph.D thesis, Massachusetts Institute of Technology, MA.

Rawlins, R. G., and Kessler, M. (**1987**). *The Cayo Santiago Macaques* (SUNY, New York).

Remez, R., Rubin, P., and Nygaard, L. (**1986**). "On spontaneous speech and fluently spoken text: Production differences and perceptual distinctions," J. Acoust. Soc. Am. Suppl. 1 **79**, S26.

Repp, B. (**1985**). "Critique of 'Measures of the sentence intonation of read and spontaneous speech in American English'," J. Acoust. Soc. Am. **78**, 1114–1116.

Scherer, K., Ladd, D. R., and Silverman, K. (**1984**). "Vocal cues to speaker affect: testing two models," J. Acoust. Soc. Am. **76**, 1346–1356.

Silverman, K. (**1987**). "The structure and processing of fundamental frequency contours" Ph.D thesis, Cambridge University, MA.

Sorenson, J., and Cooper, W. E. (**1980**). "Syntactic coding of fundamental frequency in speech production," in *Perception and Production of Fluent Speech*, edited by R. Cole (Erlbaum, Hillsdale, NJ), pp. 399–440.

Terken, J. (**1991**). "Fundamental frequency and perceived prominence of accented syllables," J. Acoust. Soc. Am. **89**, 1768–1776.

Umeda, N. (**1982**). "F_0 declination is situation dependent," J. Phon. **20**, 279–290.