Does learning affect the structure of vocalizations in chimpanzees?

ANDREW J. MARSHALL*, RICHARD W. WRANGHAM* & ADAM CLARK ARCADI†

*Department of Anthropology, Harvard University

†Department of Anthropology, Hofstra University

We recorded 'pant-hoot' vocalizations from male chimpanzees, *Pan troglodytes*, housed in two captive facilities in the U.S.A., Lion Country Safari and North Carolina Zoological Park. Acoustic analysis revealed significant differences between the two groups in the temporal patterning of the calls. Because the captive males within each group are from diverse origins, within-group similarity in pant-hoot structure could not have resulted from genetic similarity of the callers. In addition, there were no obvious differences in housing conditions that could have caused the between-group differences. Instead, the results suggest that the calls in each group converged in structure as a consequence of vocal learning. Within-group variation in call structure of the captive groups was similar to that found in a group of wild Ugandan chimpanzees (Kanyawara study group, Kibale National Park), suggesting the presence of species-specific constraints on this call within which different populations can converge on local variants. In addition, an acoustically novel pant-hoot variant that was introduced by one male to the Lion County Safari colony spread to five other males in the same colony. This suggests that chimpanzees may also be able to modify the frequency parameters of their calls through learning.

© 1999 The Association for the Study of Animal Behaviour

Learning can affect the development of vocal communication systems in various ways, including the production of calls, their usage and their comprehension (Janik & Slater 1997; Seyfarth & Cheney 1997). Janik & Slater (1997) have suggested that the term 'vocal learning' be applied only to cases in which the acoustic parameters of calls are altered by social experience, whereas changes in the usage and comprehension of calls should be considered 'contextual learning'. In addition, these authors identified two levels of complexity that can characterize vocal learning. The first involves changes in the temporal patterning and amplitude of acoustic elements already present in an animal's repertoire, which can be affected by comparatively simpler modifications in breathing patterns, and may be closely linked to contextual learning. More-complex vocal learning, by contrast, involves changes in the frequency parameters of calls, which must be mediated through the more-complexly regulated musculature of the vocal apparatus. Thus, while the simpler form of vocal learning is probably widespread in

mammals (e.g. Sutton et al. 1973), more complex forms of vocal learning appear to be rare (Janik & Slater 1997; Seyfarth & Cheney 1997).

Vocal learning (sensu Janik & Slater 1997) has been reported in passerine birds (Kroodsma 1982; Marler 1990), cetaceans (Tyack & Sayigh 1997), seals (Ralls et al. 1985) and bats (Jones & Ransome 1993; Esser 1994; Janik & Slater 1997; Boughman 1998). Although most attempts to find vocal learning in nonhuman primates have failed (Newman & Symmes 1982; Snowdon & Elowson 1992; Hauser 1996; Janik & Slater 1997; Snowdon 1997), some recent studies have suggested that it may occur. For example, Elowson & Snowdon (1994) showed that pygmy marmosets, *Cebuella pygmaea*, can alter their vocalizations in response to their social environment. Elowson & Snowdon analysed the acoustic structure of the pygmy marmoset 'trill' vocalization in two unfamiliar, acoustically isolated groups, and they found it to differ significantly. When the two groups were brought into acoustic contact, call duration converged and coextensive shifts were made in two frequency parameters (bandwidth and peak frequency) (Elowson & Snowdon 1994; but see Janik & Slater 1997). In addition, Mitani et al. (1992) reported differences between the pant hoots of chimpanzees, *Pan troglodytes schweinfurthii*, in Gombe National Park and Mahale National Park (both

Correspondence: A. J. Marshall, Department of Anthropology, Harvard University, Peabody Museum, 11 Divinity Avenue, Cambridge, MA 02138, U.S.A. (email: amarshal@fas.harvard.edu). A. C. Arcadi is at the Department of Anthropology, Hofstra University, Hempstead, NY 11549, U.S.A.

in Tanzania), in both the frequency and duration of 'build-up' elements and in the average fundamental frequency of the 'climax' phase. In red-chested moustached tamarins, *Saguinus l. labiatus*, variation was found in the range of frequency modulation and in the peak frequency of each syllable in the long call between two geographically isolated populations (Maeda & Masataka 1987), and playback experiments provided evidence that these differences were perceptually salient (Masataka 1988). Finally, in barbary macaques, *Macaca sylvanus*, frequency differences in the shrill barks of two wild-caught captive groups were detected and shown to be perceptually salient in subsequent playback experiments (Fisher et al. 1998).

Although vocal learning has been implicated in population differences in primate call production, interpopulation differences might be a consequence of genetic or environmental differences (Mitani et al. 1992; Hauser 1996; Janik & Slater 1997). Therefore, our study compares vocalizations in two colonies of captive chimpanzees that have no relevant genetic differences between them. To test whether captivity influences variation in acoustic structure, we also analysed 'pant hoots' collected from a wild community, the Kanyawara study group in Kibale National Park, Uganda. By controlling for genetic and environmental influence, our comparisons were designed to test the hypothesis that interpopulation differences in the structure of chimpanzee pant hoots are the result of vocal learning.

METHODS

Vocalizations

We studied the male pant hoot, a loud vocalization that chimpanzees produce frequently in diverse contexts (Goodall 1986). Typical pant hoots include four sequential phases: (1) the introduction, containing at least one long, unmodulated, low-pitched element; (2) the build-up, containing short inhaled and exhaled elements, produced at a faster rate than the introduction and often rising in pitch as the phase progresses; (3) the climax, containing at least one long, frequency-modulated scream; and (4) the let-down, which resembles the build-up (Goodall 1986; Mitani et al. 1992).

Subjects

Lion Country Safari (LCS), West Palm Beach, Florida, housed 42 chimpanzees in five acoustically connected groups living on separate islands. Each group had two to four adult males, one to six adult females, and one to four juveniles. North Carolina Zoological Park (NCZP), Asheboro, North Carolina, housed 12 chimpanzees (three adult males, seven adult females, and two juveniles) in an open exhibit during the day and in an indoor housing facility at night. Pant hoots were recorded opportunistically from all available adult males (LCS: 11; NCZP: 3). The subspecies and geographical region of origin of these males was largely unknown, although because the majority of trade in wild-caught chimpanzees has

historically occurred in West Africa (Fulk & Garland 1992), it is reasonable to assume that most captive chimpanzees are from the subspecies *P. t. verus* (LCS: four from Sierre Leone, four from unknown African countries, and three born in various facilities in the U.S.A.; NCZP: two from unknown countries, one zoo born). The Kanyawara study group, in Kibale National Park, Uganda, is a community of approximately 50 habituated, unprovisioned wild chimpanzees, *P. t. schweinfurthii*, that has been under continuous observation since 1987. When we made recordings for analysis, the community consisted of the following regularly observed individuals: eight adult males, 12 adult females, and 16 adolescents, juveniles and dependent offspring. We used recordings made from three adult males in the analysis for comparison to the captive communities.

Recordings

Behavioural and acoustic data were collected by A.J.M. at LCS between 6 June and 14 July 1995, and at NCZP between 18 July and 23 August 1995. Pant hoots were recorded using a Sony WM-D3 Professional Walkman, a Sony ECM 909-A directional microphone (with a frequency response range of 100–15 000 Hz), and Maxell XLII CrO_2 tapes. Recordings of the Kanyawara chimpanzees were made by A.C.A. between May, 1988 and December, 1989. Recording equipment, methods and analysis were similar to those described here, and are described in detail in Arcadi (1996).

Analysis

We digitized LCS and NCZP pant-hoot recordings at a sampling rate of 48 kHz using Sound Designer II. We carried out quantitative analysis of calls using Canary 1.2 (Cornell University Bioacoustics Workstation) on a Macintosh Quadra, using 256-point Fourier transformations that resulted in spectrograms with filter bandwidths of 761.3 Hz. Frequency resolution was 187.5 Hz and grid time resolution was 5.33 ms. Pant-hoot phases were discerned visually on audiospectrograms and waveform graphs. We measured the total duration in seconds of each pant hoot with on-screen cursors, in addition to the following parameters for each phase: the number, duration and frequency of elements, and interelement intervals (s). The average fundamental frequency was characterized for the climax phase elements wherever clear harmonic bands were distinguishable on audiospectrograms.

Only calls for which a complete recording was available were analysed. Between seven and 11 calls were analysed for all target males, either all available recordings (≤ 10) or a randomly selected subset. Acoustic variation within individuals in the parameters considered was not related to behavioural context (e.g. feeding, resting, sexual excitement, or displays), but see Clark & Wrangham (1993) for contextual variation in acoustic parameters of wild chimpanzee pant hoots.
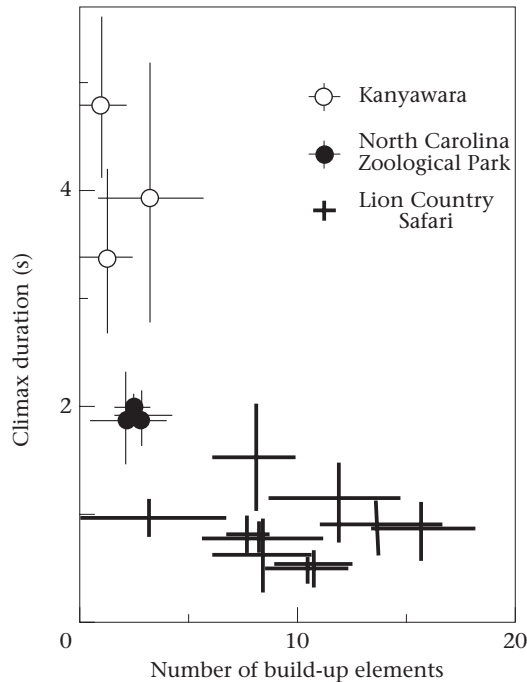
Figure 1. Mean±SD climax duration versus the mean±SD number of build-up elements. Number of pant hoots analysed was 9–10 for each male in Lion Country Safari and North Carolina Zoological Park, and 58, 13 and 10 for the three Kanyawara males. Bivariate plots of other call parameters showed similar clustering of individuals by community affiliation.

Table 1. Major differences in pant-hoot structure at LCS and NCZP

Acoustic feature	LCS	NCZP	KAN
%Pant hoots with introduction present	11	90**	90.7
Duration of introduction (s)	0.9	2.4*	2.6
Number of build-up elements	9.5	2.2*	1.7
Duration of build-up (s)	3.8	0.9**	2.3
Duration of climax (s)	1.1	1.9*	4.1

Values are means (or percentages) of individual means. Sample sizes were: $N=11$ males (LCS), $N=3$ males (NCZP), with 9–10 pant hoots analysed for each male, and $N=3$ males (Kanyawara: KAN), with $N=58$, 11 and 15 for males SY, BF and MS, respectively. All P values are two tailed, between-group differences (NCZP and LCS): *$P<0.05$, **$P<0.02$.
Kanyawara data are included for comparison and were not included in the analyses.

RESULTS

Acoustic Structure

When the two captive colonies were compared, inter-colony differences emerged in each of the first three pant-hoot phases. NCZP males produced introductions more often than LCS males (Mann–Whitney U test: $U=33$, $N_1=3$, $N_2=11$, $P<0.02$) and their introductions were longer ($U=30$, $N_1=3$, $N_2=11$, $P<0.05$). In contrast, LCS pant hoots contained build-up phases more often than NCZP pant hoots ($U=30$, $N_1=3$, $N_2=11$, $P<0.05$). Additionally, in calls that contained build-up phases, LCS males produced more build-up elements than NCZP males ($U=33$, $N_1=3$, $N_2=11$, $P<0.02$; Fig. 1). The build-up phases of the LCS chimpanzees were also longer than those produced by the NCZP chimpanzees ($U=32$, $N_1=3$, $N_2=11$, $P<0.02$). Finally, NCZP pant hoots had longer climax phases (1.85 s) than LCS pant hoots (1.06 s) ($U=30$, $N_1=3$, $N_2=11$, $P<0.05$; Fig. 1). Table 1 summarizes these differences. Mean total duration for pant hoots was 5.85 s (range 4.30–9.60) and did not differ between populations. No difference was detected in mean average fundamental frequency of the climax phase (LCS: 1.35 kHz; NCZP: 1.07 kHz).

The Kanyawara chimpanzee pant hoots were similar in structure to those collected in captivity. Summary data of selected parameters of Kanyawara pant hoots are provided in Table 1 for comparison to the two captive communities. For the three communities as a whole, there was an inverse relationship between climax duration and the number of build-ups (Fig. 1). To compare intrasite variation in the three groups, we calculated coefficients of variation for each of the parameters listed in Table 1, correcting for sample size. The range of coefficients of variation within the Kanyawara pant hoots (4.76–115) was similar to the range within the two captive groups (LCS: 41.3–202; NCZP: 5.02–122), indicating that the captive groups were not abnormal in their range of variation. The variances of the call parameters listed in Table 1 between the captive and wild groups were not statistically different using Levene's test (NS for all tests) (Van Valen 1978). However, the power to reject the null hypothesis was low due to small sample sizes.

'Bronx Cheer' Variant

The Bronx cheer is a 'raspberry'-like sound made by blowing air through pursed lips, producing identical, short, repeated elements that are strung together to generate vocalizations of variable duration (A.J.M., personal observation). Bronx cheers were produced in both captive colonies in contexts of frustration (immediately following thwarted attempts to copulate or gain access to food) or insecurity (following an aggressive act by another chimpanzee directed at either the individual or another individual). Some of the subjects of this study incorporated the Bronx cheer into their pant hoots to produce a novel variant that has never been recorded in the wild. Bronx cheer pant hoots incorporated a modified version of the raspberry sound, either directly before the build-up (increasing in pitch, rate and volume until it led into the build-up elements), or at the end of the pant hoot (decreasing in pitch and volume until the call ended). Presence or absence of raspberry sounds before or after calls was noted for each call recorded. (See Figs 2, 3 for spectrograms of both Bronx cheer and normal pant hoots at LCS.)

Bronx cheer pant hoots were given only by LCS males. They were introduced by one male, NL, a 30-year-old that had arrived at LCS at the age of 10 years (T. Wolf, personal communication) already producing this vocalization (L. Marchant, personal communication). During
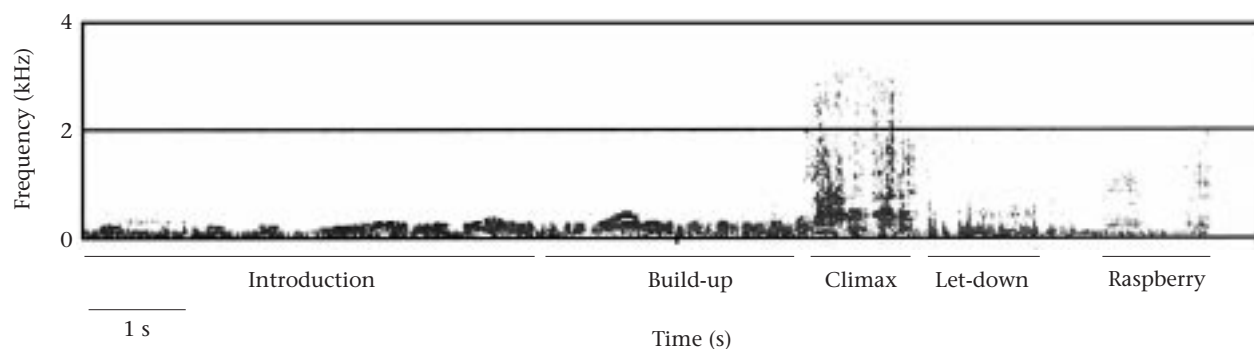
Figure 2. An example of a Bronx cheer pant hoot (NL-6/10-2) from LCS. Different phases of the pant hoot are highlighted, including the novel raspberry sound. The raspberry sound was added at either the beginning or end of the pant hoot. Bronx cheer pant hoots were produced by six males at LCS, and comprised 10–20% of individual pant-hoot production for these males.
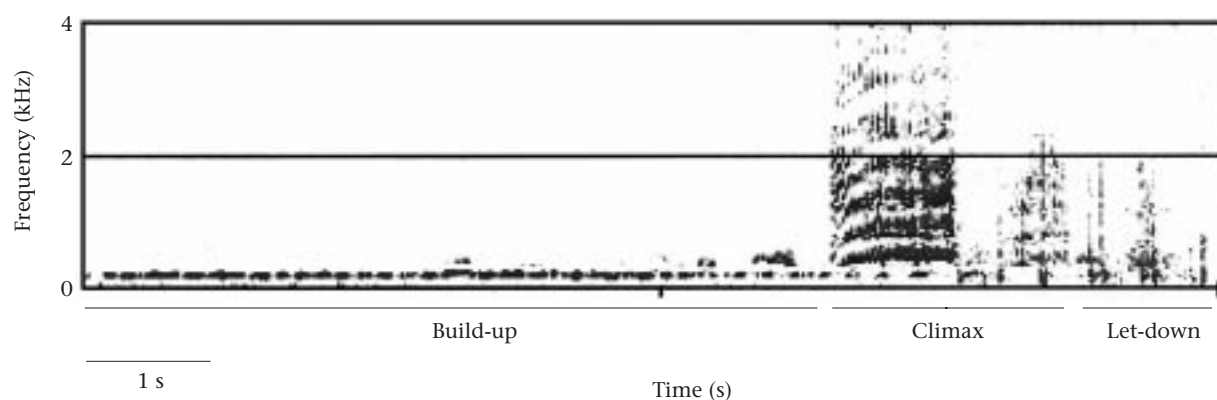
Figure 3. An example of a pant hoot without the Bronx cheer element from LCS (IN-6/22-3). Different phases are noted.

this study, six LCS males (ranging in age from 11 to 30 years) produced Bronx cheer pant hoots, comprising 10–20% of their total output of pant hoots.

DISCUSSION

Our comparison of pant hoots from two populations of captive chimpanzees demonstrates a convergence in temporal call structure within each group, and documents the existence of a novel pant-hoot variant that was introduced by one individual at LCS and that subsequently spread to at least five other individuals. Because we were able to control for genetic and environmental influences on call production, these results suggest that vocal learning in both the temporal and frequency domains may play a role in the acquisition of chimpanzee vocalizations.

Although interpopulation differences in vocalizations ('dialects') can provide evidence for vocal learning (Kroodsma 1982), demonstration of geographical variation alone is not sufficient to demonstrate vocal learning because genetic or environmental differences could also contribute to between-group variation. However, although differences in chimpanzee call structure among wild populations may be due to a variety of causes (Mitani et al. 1992; Arcadi 1996), there are fewer possibilities in captivity. For example, in our comparison of LCS and NCZP, between-group differences cannot easily

be explained by an inherited predisposition towards a certain call structure within each group or by genetically based anatomical differences between groups. The NCZP males were unrelated, only two of the 11 males at LCS were related to other LCS males, and most, if not all, of the animals are derived from West African wild populations of the same subspecies. In addition, there were no obvious social or environmental differences between the two groups that could account for the observed differences in call structure. The age–sex distributions in the two groups were comparable, and feeding schedules and housing conditions were similar.

An additional possibility is that there were motivational differences between the two groups. For example, perhaps the presence (at LCS) or absence (at NCZP) of neighbouring groups caused males to modify the production of their calls. The role of social motivation is impossible to assess independently without experimentally manipulating the social environment, so motivational differences between the two sites cannot be excluded as a potential explanation for the observed variation. However, the range of variation found in the calls of Kanyawara chimpanzees (which are in the presence of other audible groups) is comparable to that observed for both of the captive groups we studied, indicating that such motivational differences are unlikely to explain the significant differences between these two captive groups. Moreover, the similarity of within-group variation across

all groups also suggests that captivity does not distort the processes of vocal development that occur in the wild. Rather, the observed inverse relationships between the presence of the introduction phase and the number of build-ups, and between the climax duration and the number of build-ups, which were consistent across all three populations studied, suggest that there are constraints on the range of variation possible in chimpanzee pant hoots, within which populations may converge on local variants.

The remaining possibility is that differences in pant-hoot structure between LCS and NCZP were the result of learning, which led to convergence in acoustic structure within each colony. In many bird species, learned dialects are characterized by alterations in song phrasing, such as deletions and additions of parts of songs, or changes in temporal properties (Kroodsma 1982). The group differences reported in our study, that is, in the presence or absence and length of existing elements of the pant hoot, are of a similar nature, suggesting the possibility that they too are learned.

One mechanism by which bird dialects may be learned, termed 'action-based learning', involves the selective reinforcement of call variants over time, resulting in a reduction in call variability (Marler 1990). Mitani et al. (1992) suggested that action-based learning might account for the population differences they found in wild chimpanzee pant hoots, but they were unable to support this possibility empirically. Consistent with the requirements of an action-based learning model, Arcadi (1996) found evidence that the pant hoots of young chimpanzees were more variable than those of older chimpanzees. We suggest that action-based learning is a plausible mechanism to account for the differences we found between LCS and NCZP pant hoots, but note that without longitudinal data we are unable to test this hypothesis.

While differences in the temporal patterning of pant hoots provide evidence that a simple form of vocal learning may influence the development of chimpanzee calls, the spread of a novel pant-hoot variant (the Bronx cheer) through LCS indicates that more complex vocal learning affecting the frequency parameters of calls may influence chimpanzee vocal development as well. Vocal improvisation can provide evidence for either vocal or contextual learning (Janik & Slater 1997), depending on the nature of the changes that occur. If pant-hoot variants are merely the combination of sounds already present in the vocal repertoire (i.e. 'raspberries' and pant hoots), contextual learning can explain the formation of novel calls (Janik & Slater 1997). However, if one individual learns different parts of another individual's call and incorporates them into its vocal repertoire in novel ways, this provides clear evidence for vocal learning (Janik & Slater 1997). In the case of the Bronx cheer pant hoot, raspberry sounds produced alone were audibly distinct from those produced at the beginning or end of pant hoots. Although data are unavailable to support this impression empirically, it suggests that the structure of the raspberry sound was altered and incorporated into the pant hoot to produce an innovative form. If so, this would provide evidence of vocal learning in the frequency domain.

Despite many studies, there has been no convincing evidence that vocal learning influences frequency parameters in nonhuman primates (Snowdon 1990; Janik & Slater 1997), although conditioning experiments have demonstrated that temporal aspects of nonhuman primate calls can be modified (Sutton et al. 1973). In nonexperimental settings, vocal learning is difficult to prove, and therefore evidence of vocal learning is often circumstantial or ambiguous (Janik & Slater 1997). Most of the between-group differences in call structure reported here are in the temporal domain, and therefore provide strong support for a simpler form of vocal learning only. However, our documentation of the learned Bronx cheer pant-hoot variant suggests that more complex vocal learning may also play a role in the acquisition of chimpanzee call repertoires. Taken together, our data contribute to mounting evidence that social experience influences chimpanzee vocal behaviour (Mitani & Brandt 1994; Mitani & Gros Louis 1998), and suggest the need for additional experimental studies.

References

Arcadi, A. P. 1996. Phrase structure in wild chimpanzee pant hoots: patterns of production and interpopulation variability. *American Journal of Primatology*, **39**, 159–178.

Boughman, J. W. 1998. Vocal learning by greater spear-nosed bats. *Proceedings of the Royal Society of London, Series B*, **265**, 227–238.

Clark, A. P. & Wrangham, R. W. 1993. Acoustic analysis of wild chimpanzee pant hoots: do Kibale forest chimpanzees have an acoustically distinct food arrival pant hoot? *American Journal of Primatology*, **31**, 99–109.

Elowson, A. M. & Snowdon, C. T. 1994. Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Animal Behaviour*, **47**, 1267–1277.

Esser, K. H. 1994. Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat, *Phyllostomus disolour. NeuroReport*, **5**, 1718–1720.

Fisher, J., Hammerschmidt, K. & Todt, D. 1998. Local variation in Barbary macaque shrill barks. *Animal Behaviour*, **56**, 623–629.

Fulk, R. & Garland, C. 1992. *The Care and Management of Chimpanzees (Pan troglodytes) in Captive Environments*. Asheboro: North Carolina Zoological Park.

Goodall, J. 1986. *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, Massachusetts: Belknap Press.

Hauser, M. D. 1996. *The Evolution of Communication*. Cambridge, Massachusetts: MIT Press.

Janik, V. M. & Slater, P. J. B. 1997. Vocal learning in mammals. *Advances in the Study of Behavior*, **26**, 59–99.

Jones, G. & Ransome, R. D. 1993. Echolocation calls of bats are influenced by maternal effects and change over a lifetime. *Proceedings of the Royal Society of London, Series B*, **252**, 125–128.

Kroodsma, D. E. 1982. Learning and the ontogeny of sound signals in birds. In: *Acoustic Communication in Birds: Vol. 2* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 1–23. New York: Academic Press.

Maeda, T. & Masataka, N. 1987. Locale-specific vocal behaviour of the tamarin (*Saguinus l. labiatus*). *Ethology*, **75**, 25–30.

Marler, P. 1990. Song learning: the interface between behavior and neuroethology. *Philosophical Transactions of the Royal Society of London, Series B*, **329**, 109–114.

Masataka, N. 1988. The response of red-chested moustached tamarins to long calls from their natal and alien populations. *Animal Behaviour*, **36**, 55–61.

Mitani, J. C. & Brandt, K. L. 1994. Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology*, **96**, 233–252.

Mitani, J. C. & Gros-Louis, J. 1998. Chorusing and convergence in chimpanzees: tests of three hypotheses. *Behaviour*, **135**, 1041–1064.

Mitani, J. C., Hasegawa, T., Gros-Louis, J., Marler, P. & Byrne, R. 1992. Dialects in wild chimpanzees? *American Journal of Primatology*, **27**, 233–243.

Newman, J. D. & Symmes, D. 1982. Inheritance and experience in the acquisition of primate acoustic behavior. *Primate Communication*, 259–278.

Ralls, K., Fiorelli, P. & Gish, S. 1985. Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina. Canadian Journal of Zoology*, **63**, 1050–1056.

Seyfarth, R. M. & Cheney, D. L. 1997. Some general features of vocal development in nonhuman primates. In: *Social Influences on Vocal Development* (Ed. by M. Husberger & C. T. Snowdon), pp. 249–273. Cambridge: Cambridge University Press.

Snowdon, C. T. 1990. Language capacities of nonhuman primates. *Yearbook of Physical Anthropology*, **33**, 215–243.

Snowdon, C. T. 1997. Is speech special? Lessons from New World primates. In: *New World Primates: Ecology, Evolution, and Behavior* (Ed. by W. G. Kinzey), pp. 75–93. New York: Walter de Grayer.

Snowdon, C. T. & Elowson, A. M. 1992. Ontogeny of primate vocal communication. *Human Origins*, **1**, 279–290.

Sutton, D., Larson, C., Taylor, E. M. & Lindeman, R. C. 1973. Vocalization in rhesus monkeys: conditionability. *Brain Research*, **52**, 225–231.

Tyack, P. L. & Sayigh, L. S. 1997. Vocal learning in cetaceans. In: *Social Influences on Vocal Development* (Ed. by C. T. Snowdon & M. Hausberger), pp. 208–233. Cambridge: Cambridge University Press.

Van Valen, L. 1978. The statistics of variation. *Evolutionary Theory*, **4**, 33–43.