ANIMAL BEHAVIOUR, 2003, **65**, 317–329
doi:10.1006/anbe.2003.2047, available online at http://www.sciencedirect.com

# Long-distance communication of acoustic cues to social identity in African elephants

**KAREN MCCOMB*, DAVID REBY*, LUCY BAKER*, CYNTHIA MOSS† & SOILA SAYIALEL†**
*Experimental Psychology, School of Biological Sciences, University of Sussex
†Amboseli Elephant Research Project, Kenya

*(Received 22 October 2001; initial acceptance 11 December 2001; final acceptance 15 July 2002; MS. number: 7106)*

Research on long-distance vocal communication in mammals has tended to focus on the maximum distances over which a vocal signal might be physically detectable. For example, because elephants and some whales communicate using infrasonic calls, and low frequencies are particularly resilient to attenuation, it has often been assumed that these species can communicate over very long distances. However, a wide range of acoustic characteristics typically carry information on individual identity in mammalian calls, and frequency components crucial for social recognition could be distorted or lost as distance from the source increases. We used long-distance playback experiments to show that female African elephants, *Loxodonta africana*, can recognize a contact call as belonging to a family or bond group member over distances of 2.5 km, but that recognition is more usually achieved over distances of 1–1.5 km. We analysed female contact calls to distinguish source- and filter-related vocal characteristics that have the potential to code individual identity, and rerecorded contact calls 0.5–3.0 km from the loudspeaker to determine how different frequencies persist with distance. Our analyses suggest that the most important frequency components for long-distance communication of social identity may be well above the infrasonic range. When frequency components around 115 Hz become immersed in background noise, once propagation distances exceed 1 km, abilities for long-distance social recognition become limited. Our results indicate that the possession of an unusually long vocal filter, which appears to incorporate the trunk, may be a more important attribute for long-distance signalling in female African elephants than the ability to produce infrasound.



It has commonly been assumed that the maximum distances over which a species can use an acoustic signal to communicate are equivalent to the distances over which components of that signal are physically detectable. For example, because African and Asian elephants (*Loxodonta africana* and *Elephus maximus*, respectively) and some whales (e.g. *Balaenoptera physalus*) have calls with infrasonic fundamental frequencies (less than 30 Hz), it has been suggested that these species can communicate over very long distances (e.g. Payne & Webb 1971; Payne et al. 1986; Garstang et al. 1995). However, although it is theoretically possible that fundamental frequencies in the infrasonic range may still be detectable at large distances from the caller because of the unusual resilience of such low frequencies to attenuation, it is unsafe to conclude that socially relevant information could still be extracted from calls by conspecifics at these distances. In mammalian calls, a wide range of acoustic characteristics typically carry information on individual identity, and frequency components that may be crucial in social recognition could be distorted or lost as distance from the source increases.

Langbauer et al. (1991) obtained responses to playback indicating that female African elephants could detect a variety of infrasonic calls at 1.2 km from the source; males were able to detect the calls at 2 km from the source. Because playback volumes were lower than the maximum sound pressure levels at which some of these calls had been recorded in the wild, the authors extrapolated from their data to conclude that elephants could communicate over distances of at least 4 km. Others (Garstang et al. 1995; Larom et al. 1997a, b) have since used computer modelling based on this estimate to predict that, under optimum atmospheric conditions, elephants could communicate over distances in excess of 10 km. None of the above studies considered whether socially relevant information can be extracted from

*Correspondence: K. McComb, Experimental Psychology, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG, U.K. (email: karenm@biols.susx.ac.uk). C. Moss and S. Sayialel are at the Amboseli Elephant Research Project, P.O. Box 15135, Nairobi, Kenya.*



0003–3472/02/$30.00/0 

calls by other elephants at the proposed transmission distances.

Female African elephants are more vocal than males and communicate using a variety of low-frequency calls (Poole 1995). Although these calls often have infrasonic fundamental frequencies and are referred to as 'infrasonic', harmonics of the fundamental usually extend well into the audible range. In the case of the contact call, which is the low-frequency call most regularly used by females for long-distance communication with social companions, sound energy is often present up to frequencies of at least 1 kHz (K. McComb & D. Reby, personal observation). Previous studies have not considered the potential importance of noninfrasonic frequencies in coding individual identity and social meaning, nor have they given due consideration to whether elephant hearing could, over long distances, readily extract information contained in an infrasonic fundamental frequency contour alone.

In earlier playback experiments, we showed that adult female African elephants in Amboseli National Park, Kenya, are familiar with the contact calls of about 100 other females in the population (McComb et al. 2000). The primary social unit in these elephants is the female family unit, composed of adult females that are usually matrilineal relatives and their immature offspring. Close social ties also exist within bond groups, which are made up of family units that associate often and greet one another when they meet (Moss & Poole 1983). Females can distinguish the calls of female family and bond group members from females outside these categories, and beyond this, they discriminate between the calls of other families based on how frequently the females encounter them (McComb et al. 2000, 2001). In response to playback of the calls of family or bond group members, listening females typically give a distinctive reaction characterized by contact calling and/or approaches to the area from which the call came (McComb et al. 2000). In contrast, on hearing playbacks of calls from families outside these categories, females either listen and remain relaxed if they have associated frequently with the caller, or bunch into defensive formation if they have encountered the caller only rarely, but they do not call back or approach the loudspeaker (McComb et al. 2000, 2001).

To investigate the distances over which social recognition is possible, it is necessary to identify a diagnostic response that shows that subjects have not only detected the call but also categorized the social identity of the caller. To do this, we used the distinctive calling/approach response described above, which is given when subjects have identified a call as belonging to a family or bond group member (McComb et al. 2000). In the present study, calls of family and bond group members were played to subjects from successively closer distances until this recognition response was obtained. Then, to investigate the basis for this social recognition mechanism, we used acoustic analyses based on source-filter theory to identify vocal features that have the potential to advertise individual identity in female contact calls. Finally, to examine how different frequency components in the calls degrade with distance, we rerecorded contact calls at distances from a loudspeaker of 3.0–0.5 km and quantified levels of degradation.

## METHODS

### Study Population

Fieldwork was conducted in Amboseli National Park, Kenya, where data on life histories and association patterns have been obtained for more than 1700 individual elephants over 28 years by the Amboseli Elephant Research Project (Moss & Poole 1983; Moss 1996). The park encompasses 390 km$^2$ and covers a varied ecosystem including open grassland sparsely scattered with *Acacia* trees, dense patches of palms, permanent and semi-permanent areas of swamp and the seasonally flooded bed of lake Amboseli (see e.g. Moss & Poole 1983). All adult individuals in the population are recognized from a combination of natural features, particularly patterns of tears, holes and veins in the ears, and ear and tusk size and shape (Moss 1996).

### Sound Recording and Playback Equipment

Contact calls were recorded on digital audiotape using equipment specialized for low-frequency recording: a Sennheiser MKH 110 microphone linked to either a Sony TCD D10 DAT recorder (with DC modification) or HHb PortaDAT PDR 1000 DAT recorder, through an Audio Engineering Ltd power supply (which incorporated a 5-Hz high-pass filter). With this equipment, the frequency response for recording was flat ( ± 1 dB) down to at least 10 Hz. The system for playback was composed of a custom-built sixth-order bass box loudspeaker with two sound ports (Aylestone Ltd, Cambridge, U.K.) linked to either a Kenwood KAC PS 400M, Kenwood KAC923 or Kicker Impulse 1252 xi power amplifier and a HHb PortaDAT PDR 1000 DAT or Sony TCD D10 recorder (with DC modification), to give a lower frequency limit of 10 Hz and a playback response that was flat ± 4 dB from ca. 15 Hz on one sound port and ca. 20 Hz on the other.

### Social Recognition Distances

Playbacks were conducted between June 1993 and January 2000. All contact calls used as playback stimuli had been recorded from adult female elephants (at least 11 years old), at distances of less than 30 m from the caller and in conditions of low air turbulence (see also McComb et al. 2000, 2001). In each playback, a single contact call was played at peak sound pressure levels of 107 ± 3 dB at 1 m (measured with a CEL-414/3 sound level meter, C weighting), to represent a call given at high volume (McComb et al. 2000, 2001). Playbacks were given only when the female whose call was played was not found within 2 km of the subjects and when subjects were within their home range.

#### *Playback protocol*

Playback experiments were conducted using two vehicles that were in radio contact. A researcher in one vehicle played the calls and used an odometer to measure the distance from the subjects, moving first to the maximum distance at which calls were to be played, then backtracking to successively closer positions (see below). During playback, the vehicle was positioned with its axis on a direct line to the elephants, and recorded vocalizations were played through the rear door, which pointed in the direction of the subjects. Observers in the second vehicle were positioned next to the subjects and monitored their response. Records were kept in the two vehicles on the timing of playback and subject responses, using clocks that had been synchronized immediately before the playback session.

### Playback responses

Responses to playback were observed through binoculars and recorded on videotape or in written notes. From a range of behaviours monitored during earlier playback experiments (McComb 1996; McComb et al. 2000, 2001), we used the behaviours listed below to classify subjects' reactions. Although we did not use a blind observer protocol in these experiments, the key behaviours used to classify a response were unambiguous. Contact calling and approach were diagnostic responses that indicated whether a playback call had been categorized as belonging to a family or bond group member. The other behavioural responses provided contextual information on whether the call had been detected; two of these, bunching and avoidance, also provided an indication that the caller had been categorized as an infrequent associate (McComb et al. 2000, 2001).

*Indicative of call detection.* (1) Listening: any subject holds ears in a stiff extended position. (2) Smelling: any subject uses the tip of its trunk to smell, in lowered, middle, or raised position. (3) Streaming: any subject produces new secretion from the temporal glands, visible as a dark moist spot that was not present before playback.

*Indicative that caller was categorized as infrequent associate.* (1) Bunching: subjects bunch together into defensive formation so that the diameter (estimated in terms of elephant body lengths of the whole group, or of constituent subgroups), decreases. (2) Avoidance: subjects change direction to walk continuously away from the loudspeaker.

*Indicative that caller from family or bond group discriminated.* (1) Contact calling: any subject gives a contact call, usually preceded and followed by periods of listening. (2) Approach: subjects move towards the loudspeaker, often smelling the ground and air as they do so.

### Experimental trials

We first played to subjects the contact call of a family or bond group member from a long distance away, then from successively closer distances. This allowed us to calculate social recognition distance as the distance at which the subjects first gave the diagnostic response of contact calling and/or approach; when this occurred we terminated the experiment. This protocol of moving successively closer to the subjects during the experimental series rather than further away from them avoided potential effects of habituation leading to an underestimate of social recognition distance. Furthermore, because exposure to the call from the greatest playback distances will constitute the weakest signal, this method also reduced possible overall effects of habituation.

Experimental trials could be performed only when a family of subjects was inside its range and without a family or bond group member whose contact call we had in our library of recordings. This situation arose only rarely, particularly between 1997 and 2000 when adult females spent most of their time with other family members. If a particular female was missing from a family or bond group under these conditions, and we had a recording of her contact call, the appropriate call was played first at 2.0 km (first four experimental trials) or 2.5 km (last three experimental trials) from subjects, then at successively closer intervals of 1.0 or 0.5 km (Table 1) until a response of either approach towards the loudspeaker or contact calling occurred within 3 min of playback. Once this diagnostic response had been obtained, the playback trial was terminated.

### Control trials

We had established earlier, with playback distances of 100 m, that subjects typically contact-call in response to playbacks only when they categorize the playback stimulus as belonging to a family or bond group member (McComb et al. 2000). It remained possible that, over longer distances, such a response could be given to a call not because social recognition had taken place, but because it was so indistinct that the identity of the caller remained unknown. We conducted control trials to ensure that false positive responses were not given to the calls of strangers (individuals outside the family or bond group) when these were played from long distances. In these trials, playbacks of the type described for the experimental trials were given to subjects who were not from the same family or bond group as the caller. In these cases, the call was played first at 2.5 km from the subjects, then at successively closer intervals of 0.5 km until we were 0.5 km from the subjects (Table 2).

We conducted all playback trials between the hours of 0700 and 1300 hours.

### Sample sizes and statistical analyses

Seven series of playbacks of family/bond group members' calls were given to seven independent groups of subjects. We used the calls of six adult females (Echo, Esme, Kleo, Kora, Remedios, Ysolde) as playback stimuli. One call was used twice because it represented a family member's call to one group of subjects and a bond group member's call to another. We defined social recognition distance in each of these experimental series as the distance at which the subjects first gave the diagnostic response to playback (calling and/or approach). In control trials, we used these same six calls, and again gave

**Table 1.** Responses to experimental trials, when subjects were played calls of family or bond group members at 2.5–0.5 km

| Family (call played) | 2.5 km | 2.0 km | 1.5 km | 1.0 km | 0.5 km |
|---|---|---|---|---|---|
| EAs (Esme) | No response | — | Listening, streaming, **calling** | — | |
| SBs (Ysolde) | | Listening, smelling | — | Listening, smelling, **calling**, **approach** | — |
| JAs (Ysolde) | | Listening | — | Listening, smelling, **calling**, **approach** | — |
| RAs (Remedios) | | Listening | Listening, streaming, **calling** | — | — |
| EBs (Echo) | No response | Listening | Listening | Listening | Listening, smelling, **calling**, **approach** |
| 1st section KBs (Kleo) | Listening, smelling, streaming | Listening, smelling | Listening, smelling, **calling**, **approach** | — | — |
| 2nd section KBs (Kora) | Listening, smelling, streaming, **calling**, **approach** | — | — | — | — |

Diagnostic response in each trial is shown in bold. Calling=contact calling.

**Table 2.** Responses to control trials, in which subjects were played calls from females who were not family or bond group members at 2.5–0.5 km

| Family (call played) | 2.5 km | 2.0 km | 1.5 km | 1.0 km | 0.5 km |
|---|---|---|---|---|---|
| 1st section FBs (Kleo) | No response | No response | Streaming | Listening, avoidance | — |
| 2nd section FBs (Esme) | No response | Streaming | No response | No response | No response |
| JBs (Echo) | No response | Streaming | No response | No response | Listening |
| AA (Remedios) | Listening | No response | Listening, smelling, streaming | Listening, smelling | Listening, smelling, streaming, bunching |
| PC (Ysolde) | No response | No response | Streaming | Listening, smelling | Listening, streaming |
| PA (Kora) | No response | No response | No response | Listening, smelling | Listening, smelling, avoidance |
| VA (Kora) | Smelling, streaming | No response | Smelling | Listening | Listening, smelling, streaming |

seven series of playbacks to seven independent groups of subjects. We compared the occurrence of the diagnostic response in the experimental trials to that in the control trials in two ways. A two-tailed Fisher's exact test was used to test the 2 × 2 table categorizing whether the 14 independent groups of subjects gave the diagnostic response at any stage during playback trials in relation to whether the caller was or was not a family member. Although the same stimuli were played to subjects in the experimental and control trials, they represent different treatments because in the experimental trials these callers were from the same family/bond group as the subjects, and in the control trials they were not. By using the same stimuli in experimental and control trials we controlled for idiosyncrasies in the playback stimuli, over and above the differences in the category of caller that the stimuli represented to the subjects, which might have contributed to differences in the response. We also used a two-tailed binomial test to calculate the probability of obtaining a positive outcome by chance for all six playback stimuli. A positive outcome would involve obtaining a calling/approach response when the playback stimulus represented a family/bond group member and an absence of this response when it did not.

## Acoustic Cues to Individual Identity

Recordings of 99 different contact calls from 13 adult females were used to examine individual variation in contact call characteristics. Calls were transferred from DAT tapes to the hard disk of a Power PC Macintosh computer using the S/PDIF digital input of an Audiomedia III sound card (48 kHz sampling rate). Digital files were then downsampled to 11.1 kHz and saved in AIFF format (16 bits amplitude resolution). After low-pass filtering, sound files were downsampled to 0.551 kHz, and narrow-band spectrograms (FFT size=512, overlap=50%, filter bandwidth=8.74 Hz, frequency grid resolution=1.077 Hz) of each call were edited and saved using Canary 1.2.4 software (Chariff 1995).

### *Voice production*

A voiced sound is the product of a source signal (generated by vibration of the vocal folds in the larynx) that is subsequently filtered in the cavities of the vocal tract (Fant 1960). The source signal, typically a quasi-periodical wave with a fundamental frequency (F0) and integer multiple harmonics, determines the pitch of the vocalization (Fig. 1). Before radiating through the mouth and nostrils into the environment, the source signal passes through the supralaryngeal vocal tract. Because the vocal tract is effectively a tube of air with natural resonances, it selectively amplifies certain frequencies in the source spectrum. This filtering process thus shapes the spectral envelope of the signal, producing peaks called formants (Fant 1960; Fig. 1). Since characteristics of vocalizations that arise from inherent properties of the filter can vary independently from those that arise from the source, either or both may provide receivers with important information.

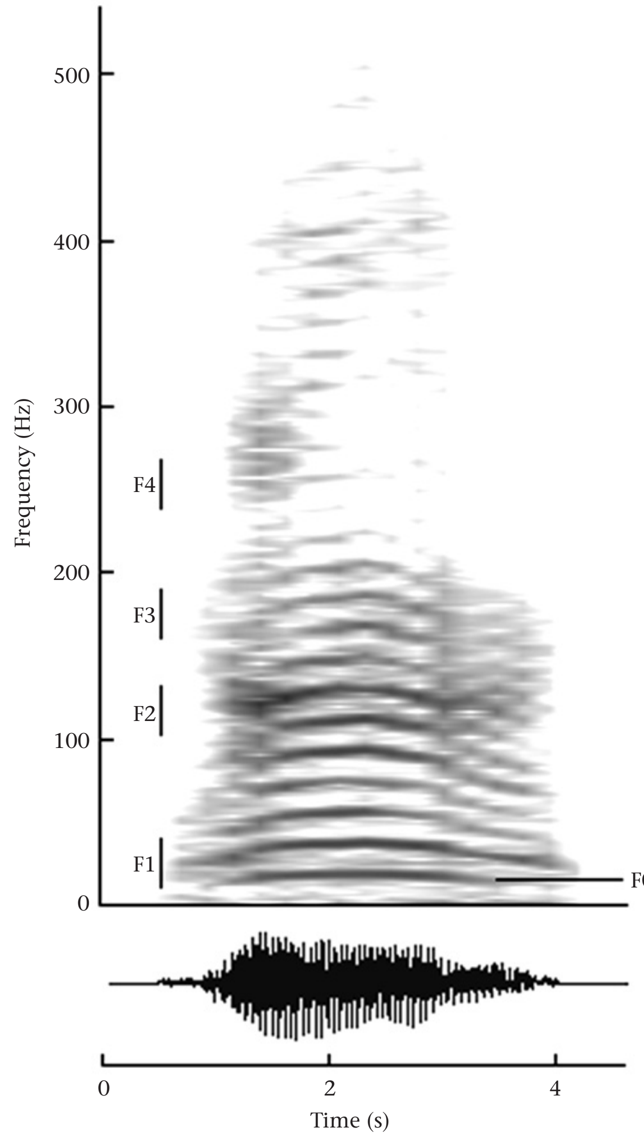


**Figure 1.** Waveform and spectrogram of a female contact call (individual=Esme), showing the fundamental frequency (F0) and harmonics, and the position of the first four formants (F1–F4). Frequency bandwidth: 8.74 Hz; FFT size: 1024 points; overlap: 50%.

We selected characteristics of contact calls for analysis that would reflect acoustic differences generated by both the source and the filter. Source- (fundamental frequency) and filter- (formants) related acoustic features were extracted using PRAAT 3.9.27 DSP package (P. Boersma & D. Weenink, University of Amsterdam, The Netherlands).

### *Extraction of source-related features*

To characterize the source, we measured a number of features from the fundamental frequency contour. An autocorrelation (To pitch (cc) command) algorithm was first used to produce time-varying numerical representations of the fundamental frequency. The time step in the analysis was set at 0.1 s. To prevent octave errors, the expected F0 range was set for each call after a visual

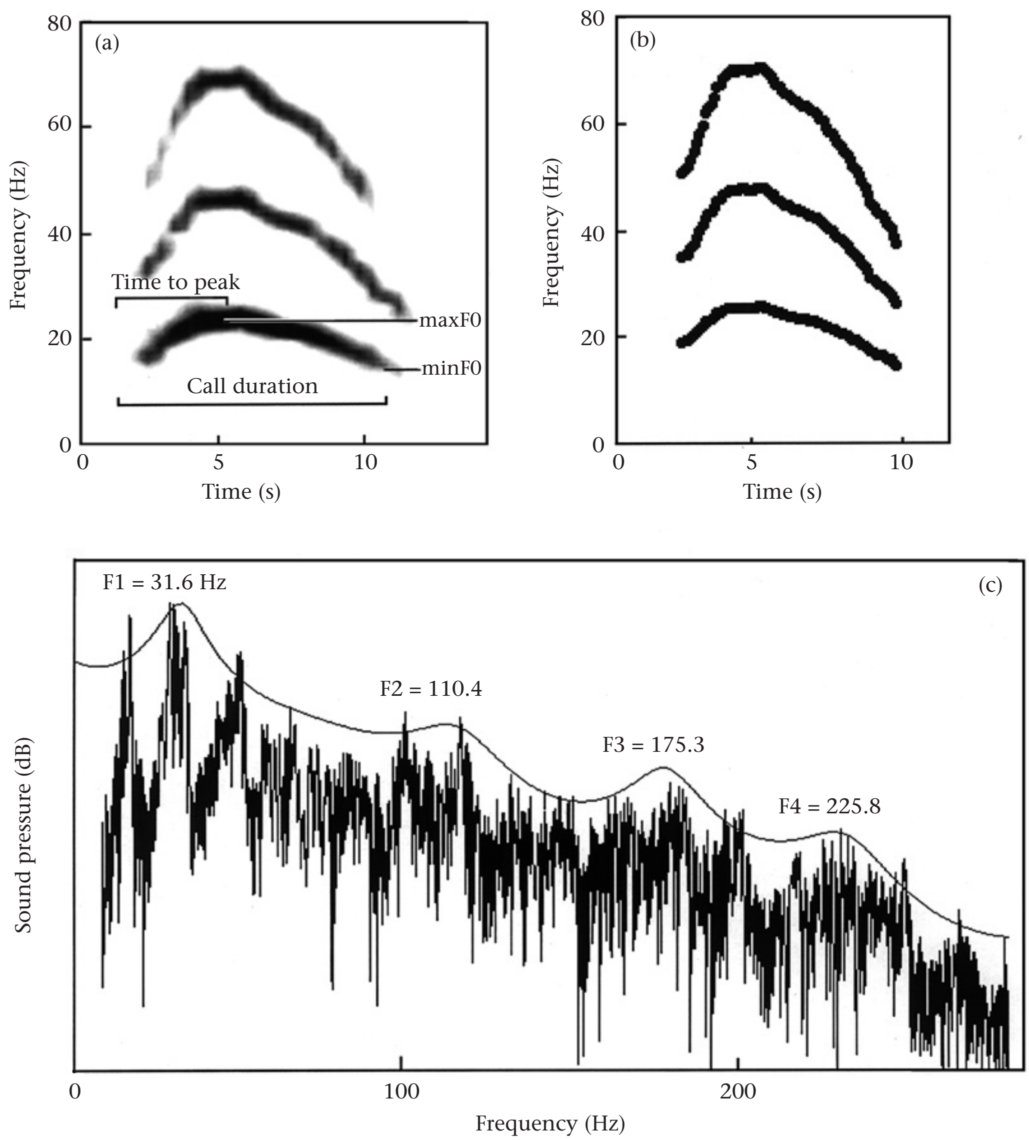


**Figure 2.** Illustration of how the main source variables and filter variables were measured. Source variables: (a) the first 80 Hz of a contact call spectrogram with the fundamental frequency (F0) and the first two harmonics; (b) numerical representation of the fundamental frequency contour (the first two harmonics are extrapolated as integers of the fundamental frequency). (c) Filter variables: overall frequency spectrum of a contact call with linear predictive coding smoothing superimposed. Centre frequencies for formant values are shown.

assessment of fundamental frequency variation in the call using narrow bandwidth spectrographic analysis in Canary. The contour of the fundamental frequency was inferred from onscreen spectrographic examination of the harmonics of the fundamental, thereby excluding any underestimation of the minimum values of F0 caused by a possible attenuation of the frequency components close to the roll-off frequency range of our equipment or by background noise. Typical preset values for the analysis were: 7 Hz (F0 min), 30 Hz (F0 max). The time window in the frequency analysis was variable and was automatically imposed by the preset lower limit for fundamental frequency. The output numerical representations of the frequency contour were transferred to Microsoft Excel to derive the following measurements (Fig. 2a).

(1) Minimum, mean and maximum fundamental frequency in the call (minF0, meanF0 and maxF0, respectively). Where minF0 was close to the roll-off frequency for our equipment, this would not have affected the outcome of the F0 analysis, which computes F0 values on the basis of the general pattern of harmonicity rather than extracting them only from the F0 contour.

(2) Number of inflection points in the fundamental frequency contour, calculated as the average number of inflexions in the fundamental frequency contour (Inflex). The number of inflections was the number of changes in the sign of the derivative (slope) of the fundamental frequency contour, after a three-point average smoothing filter was run to remove rapid variations caused by jitter or analysis imprecision.

(3) The cumulative variation in the fundamental frequency contour during the call, calculated as the sum of the absolute value of the fundamental frequency derivative (Sumvar):

$$\text{Sumvar} = \sum_{n=1}^{T} |(\chi_{n+1} - \chi_n)|$$

where $\chi_n$ is the frequency at point $n$.

(4) The percentage of the call duration that had elapsed when max F0 was reached (%Elapsed).

(5) Call duration (Duration).

### *Extraction of filter-related features*

Characterizing the filter proved more difficult, because the absence of knowledge of the functional anatomy of the elephant vocal tract made a priori hypotheses about the possible range of formant frequencies speculative. Using narrow-band spectrograms, we identified potential formant regions encompassing several harmonics in the lower part of the call spectrum. We used an approach based on linear predictive coding (LPC; Markel & Gray 1976) to characterize the five formant peaks present in the first 275.5 Hz of the call spectrum. An overall spectrum of the call was calculated with the 'To Spectrum' command in PRAAT and an 'LPC smoothing' algorithm was then applied to yield values for the first five potential peaks in the 0–275.5 Hz range (Fig. 2b). The frequency values for these peaks (F1–F5) were extracted with the 'To Formant' command (number of formants=5). Since only the first two peaks (F1, F2) were consistently present in calls, we decided to use only these and to omit F3, F4 and F5 from the analyses.

### *Discriminant analyses*

The importance of source- and filter-related variables in coding individual identity was examined with a stepwise discriminant function analysis DFA (individuals=calls, groups=caller). Three DFAs were carried out, one using source-related variables only, one using filter-related variables only, and one using all available variables.

In all three cases, we tested call membership with both resubstitution and crossvalidation procedures. In the resubstitution procedure, all the calls in the data set are used to build a single model, which is then used to test to which group these calls belonged (in our case, group membership is the identity of the individual that gives the call). This method characterizes the ability of the model to 'recognize' the group membership of the calls. In the crossvalidation, 'all-but-one' or 'leave-one-out' procedure, a different model is built for each call in the data set, using all the calls except the one that is being tested. This procedure is more conservative, and characterizes the ability of the model to predict group membership. The results of the discriminant analyses were expressed as percentages of correct classification. Because some calls (15%) had no value for F1, F2 or both, they were automatically deleted for the calculation of the discriminant functions that involved filter-related variables.

## Attenuation of Spectral Components with Distance

To examine how the spectral structure of contact calls degrades with distance from the caller, we played back calls from five of the six original adult females (Echo, Esme, Kleo, Remedios, Ysolde) and simultaneously rerecorded them at 0.5, 1.0, 1.5, 2.0, 2.5 and 3 km away. All rerecording sessions were conducted between 0700 and 1300, on days when wind speed was low (up to 7 mph). Rerecordings of some or all the five calls were made on four dates: 6 May 1996 (3/5 calls), 2 August 1996 (4/5 calls), 27 July 1999 (5/5 calls) and 24 January 2000 (5/5 calls). Of these dates, July 1999 provided the clearest rerecordings, and rerecordings of contact calls from all five adult females made at this time were analysed and form the basis of the results. Rerecordings of one of these adult female contact calls (from Ysolde) were also examined across the rerecording sessions on all four dates to ensure that observed patterns of call degradation were consistent across recording sessions.

Calls recorded at each rerecording distance were first digitized using the methods previously described, and narrow-band spectrograms (FFT size=512 overlap=50%, filter bandwidth=8.74 Hz, frequency grid resolution=1.077 Hz) of each call were edited and saved using Canary 1.2.4 software (Chariff 1995). We used these spectrograms to assess visually how the different frequency components in the call degraded with distance. To quantify this degradation accurately, we then computed a long-term average spectrum (LTAS) of the call, depicting the energy distribution in the frequency domain, averaged on the duration of the call (frequency analysis window of 15 Hz), yielding 27 quantitative variables (H1–H27) each depicting the amplitude (dB) of 15-Hz frequency slices (H1: 0–15 Hz; H2: 15–30 Hz to H27: 390–405 Hz). We then calculated the LTAS of a segment of background noise immediately preceding or following the call. We subtracted the LTAS of the noise segment from the LTAS of the call to calculate the ratio of the level of the call plus background noise to that of background noise alone (Call to Background Noise ratio) across frequencies at each distance. This information was used to model the attenuation of the lower frequency components with distance in two forms: (1) plots of Call to Background Noise ratio across frequencies at each distance from each of the five adult females rerecorded in July 1999; (2) a plot of Call to Background Noise ratio across frequencies at each distance for one female (Ysolde) averaged over the May 1996, August 1996, July 1999 and January 2000 rerecording sessions.

# RESULTS

## Social Recognition Distance

In the playbacks of family/bond group members (Table 1), the range of distances at which the diagnostic social recognition response (contact calling and/or approach loudspeaker) was given was 2.5–0.5 km ($\bar{X} \pm SD = 1.21 \pm 0.64$ km: modal distance=1 km). The

**Table 3.** Tests of equality of group means between individuals for each source and filter variable used in the discriminant function analysis

| | Range | | | ANOVA | | |
|---|---|---|---|---|---|---|
| | Mean | Minimum | Maximum | Wilk's lambda | *F* | *P* |
| MeanF0 | 16.8 | 11.8 | 24.4 | 0.366 | 10.230 | <0.001 |
| MaxF0 | 19.1 | 13.6 | 26.5 | 0.418 | 8.253 | <0.001 |
| MinF0 | 13.1 | 7.3 | 20.0 | 0.401 | 8.835 | <0.001 |
| Duration | 5.0 | 2.1 | 8.2 | 0.413 | 8.416 | <0.001 |
| Inflex | 4.4 | 1.0 | 14.0 | 0.560 | 4.655 | <0.001 |
| Sumvar | 14.5 | 5.2 | 29.9 | 0.560 | 4.656 | <0.001 |
| Elapsed | 0.38 | 0.12 | 0.72 | 0.667 | 2.952 | <0.002 |
| F1 | 35.1 | 20.4 | 53.9 | 0.453 | 7.142 | <0.001 |
| F2 | 114.1 | 87.1 | 157.4 | 0.417 | 8.279 | <0.001 |

See text for definitions of variables.

subjects typically showed signs of detecting the call (as indicated by listening, smelling or streaming) from distances of 2 and 2.5 km but did not respond as though it came from a family member until playback distances had narrowed to 1.0 or 1.5 km. The control trials confirmed that subjects did not give false positive responses to calls that were not from members of their own family or bond group, regardless of the distance from which we played them (Fisher's exact test: $N$=14, $P$=0.0006; binomial test: $N$=6, $P$=0.031; Tables 1, 2; see also McComb et al. 2000). The occurrence of bunching and avoidance reactions at distances of 1.0 and 0.5 km in these trials (Table 2) was consistent with the calls of infrequent associates (McComb et al. 2000, 2001) having been identified over these distances.

## Individual Identity in Contact Calls

For each of the acoustic parameters in the 99 contact calls that we acoustically analysed, the means differed significantly between individuals (groups in the discriminant function analysis below; Table 3, Fig. 3). Using all available variables, 77.4% of the 84 calls (53.6% in the more conservative crossvalidation method) were correctly attributed to callers. Using source-related variables only (meanF0, maxF0, minF0, Inflex, SumVar, %Elapsed and Duration), 65.7% of calls ($N$=99) were correctly classified (43.4% with crossvalidation). Using filter-related variables only (F1, F2), we found that the percentages of correctly assigned calls ($N$=84) dropped to 40.5% (33.3% with crossvalidation).

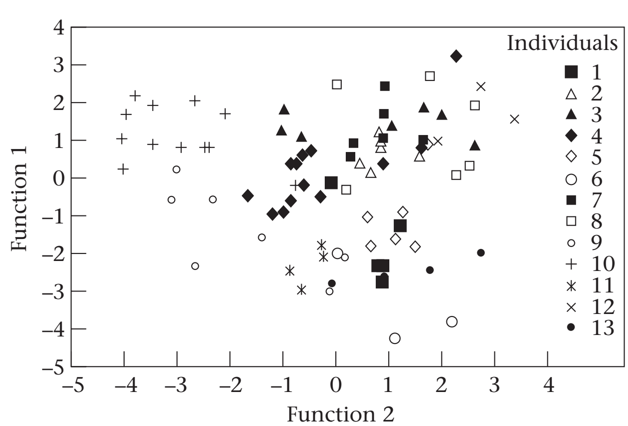


**Figure 3.** Separation of contact calls from 13 adult females based on the first two canonical discriminant functions in the analysis that includes both source and filter acoustic variables.

## Attenuation of Calls with Distance

An example of how a single contact call degrades with distance is shown in Fig. 4. The variation in the Call to Background Noise ratio with frequency for each rerecording distance, for each of the five individuals in the July 1999 session, shows that the frequency peaks in the 115-Hz region (F2) are most prominent and have the highest persistent with distance, decaying at a lower rate than other frequency peaks as distance increases (Fig. 5). When the long-term average spectra for one individual (Ysolde) at each rerecording distance are averaged across all four rerecording sessions, the frequency peaks in the 115-Hz region are again the most prominent and persistent (Fig. 6).

## DISCUSSION

Long-distance playback experiments indicated that social recognition on the basis of call characteristics in our study population of African elephants was possible over distances of up to 2.5 km, but was more usually achieved around 1 km from the playback loudspeaker. Usually subjects listened when the contact calls of family or bond group members were presented from the furthest distances, indicating that they had detected the call. Typically, however, only when the playback distance had narrowed to 1 km did they respond by calling back and approaching in the direction of the loudspeaker, indicating that they had identified the caller as belonging to a family or bond group member. Observations of family members that became separated and used contact calls to relocate each other suggest that females put more, rather than less, effort into calling when distances between

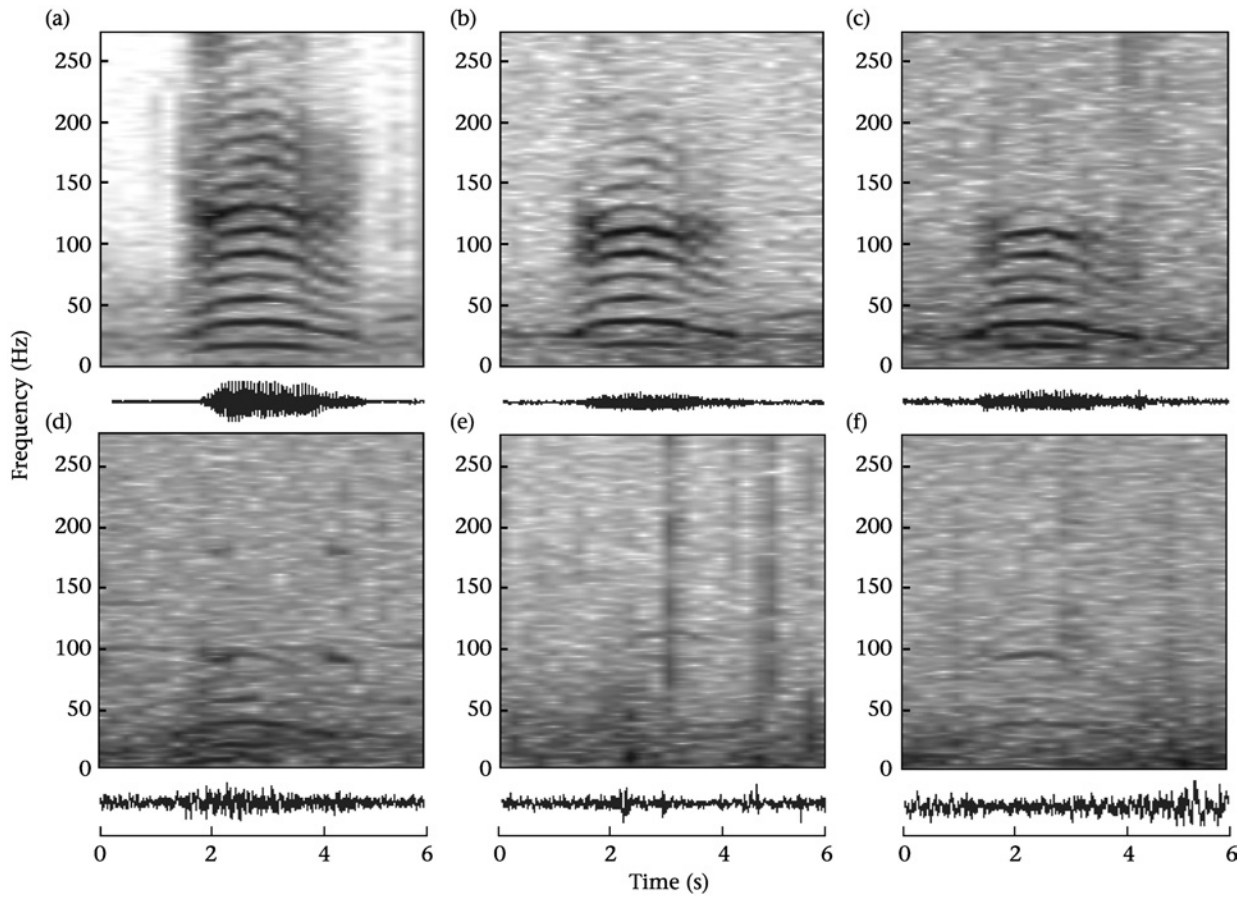


**Figure 4.** Spectrograms showing change in a contact call from one individual (Esme) with distance: (a) original call; (b) rerecording at 0.5 km; (c) rerecording at 1.0 km; (d) rerecording at 1.5 km; (e) rerecording at 2.0 km; (f) rerecording at 2.5 km.

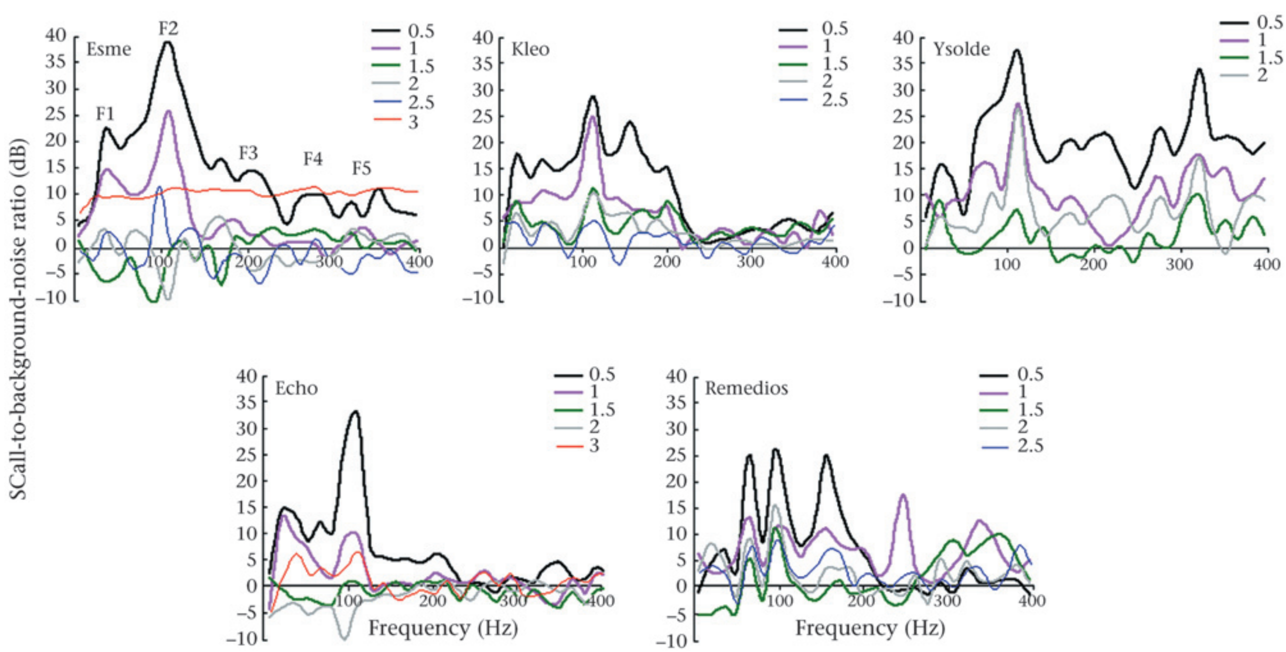


**Figure 5.** Attenuation curves for calls from five individuals showing variation in call-to-background noise ratio with frequency for each rerecording distance (0.5–0.3 km). Missing values for particular distances reflect situations where calls were not detectable in the rerecording.

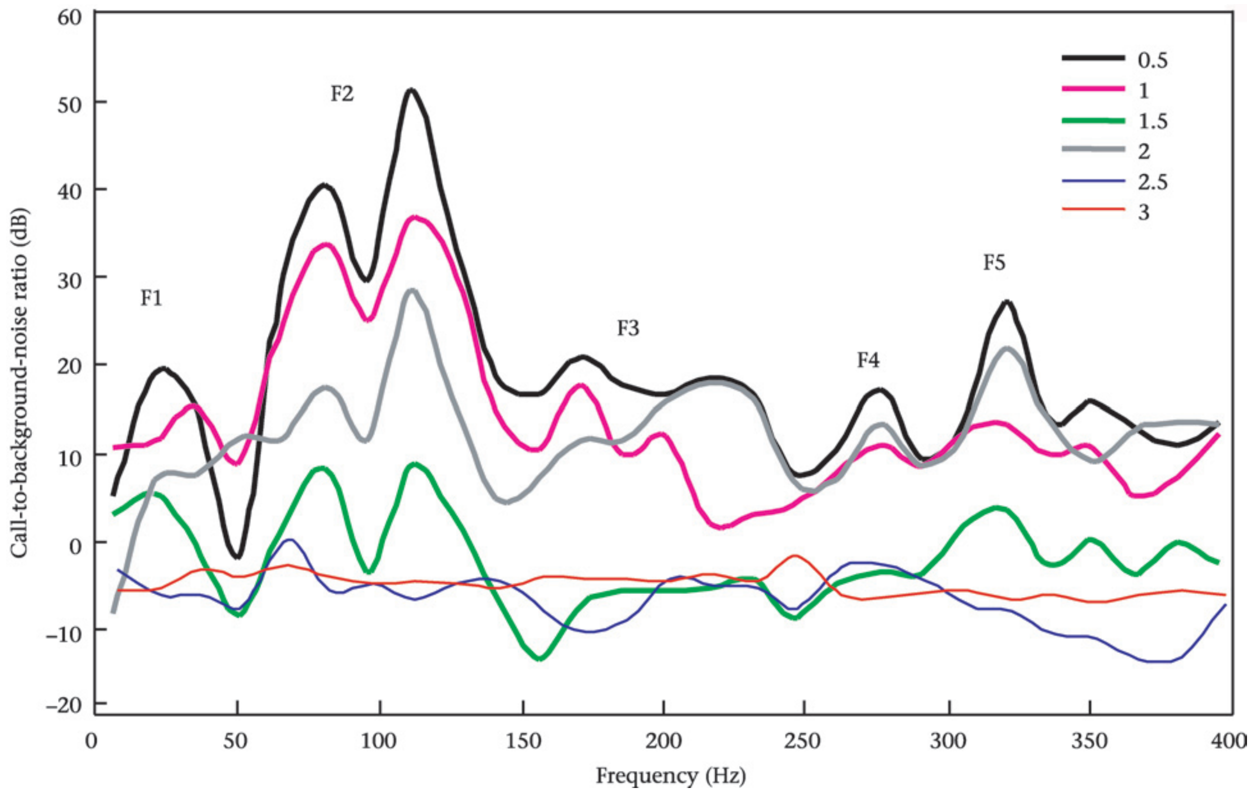


**Figure 6.** Attenuation curve for calls from one individual (Ysolde) averaged across four different rerecording sessions showing variation in call-to-background noise ratio with frequency for each rerecording distance (0.5–3 km).

them are large (Poole et al. 1988; K. McComb, personal observation). We therefore consider it unlikely that subjects recognized the family or bond group member at the distance at which they gave their first listening response, but did not respond because they were waiting for the individual to come closer. In the control trials, where calls played were not from family or bond group members, bunching and avoidance responses indicating that the caller had been categorized as an infrequent associate were obtained at 1 km or less.

Acoustic analyses indicated that all nine acoustic features that we analysed were important for distinguishing individual calls, including variables related both to vocal fold vibration (source-related variables) and to vocal tract resonances (filter-related variables or formants). The range of vocal variation encompassed by the rate at which the vocal folds vibrate, and the modulation of this rate as the call progresses, potentially provides a rich source of interindividual variation. The centre frequencies of the formants might be expected to be less reliable in assigning identity, because they depend on the size and posture of the caller, and different individuals may overlap in these respects (Reby & McComb, in press). Furthermore, while the detailed patterning of a set of formants can provide information on individual identity by reflecting idiosyncrasies in vocal tract shape (e.g. Rendall et al. 1998), this information is more likely to be important for short- and medium-range communication. When female elephants communicate over long distances, the complex pattern of formant frequencies (centre frequencies and bandwidths) is likely to be dramatically altered by attenuation effects that will not be constant across the frequency domain. As a consequence of this distortion of the spectral envelope, resulting in the reduction of formant bandwidths and ultimately in the loss of certain formants, the ability of the formant frequencies to carry information on individual identity over long distances is likely to be severely reduced, as was confirmed by our rerecording measurements.

The average spacing of formant frequencies can provide information on the length of the vocal tract in mammals (Fitch 1997; Reby & McComb, in press). Based on the frequencies of the first four formants, and assuming a vocal tract that is a uniform tube closed at the larynx and open at the radiating end (Reby & McComb, in press), average formant spacing in our analyses is 62.4 Hz. Based on the physical relationship between formant spacing, sound velocity and vocal tract length (Fant 1960; Fitch 1997), this would predict an unusually long vocal tract length for female elephants. Assuming that sound velocity in the vocal tract is 350 m/s (Titze 1994), a formant spacing of 62.4 Hz would predict a vocal tract length of ca. 2.8 m, suggesting that the trunk and possibly a pharyngeal cavity (resulting from a mobile larynx which may be pulled downwards; Gasc 1967; Shoshani 1998) interconnect to form an extended filter. The exceptionally low resonance frequencies resulting from this very long filter accentuate the lower harmonics in the spectrum of the female contact call and are undoubtedly important in facilitating long-distance communication of social identity.

Our analyses of rerecordings suggest an acoustic basis for the loss of social identity cues at the furthest playback distances. These pinpoint a particular vocal tract resonance (F2, the second formant), spanning several harmonics and centred around 115 Hz, that is prominent and persistent, decaying at a lower rate with increasing distance than frequency components below and above it. In our rerecordings, this F2 band of energy dropped to the level of background noise between 1.5 and 3.0 km from the source. The specific distance at which this happened is likely to have been a function of acoustic characteristics of the individual's call and also of the particular wind and atmospheric conditions that prevailed at the instant of rerecording. Given that information on the fundamental frequency contour can be extracted from its harmonics (Houtsma 1995), the second formant highlights several prominent harmonics from which individually specific information about fundamental frequency modulation could be derived by extrapolation. Harmonics in the 115-Hz area may experience less interference from wind noise than the fundamental frequency contour itself or harmonics in the region of the first formant.

The hearing sensitivity of African elephants has not been measured directly, but data exist on hearing in Asian elephants (Heffner & Heffner 1980, 1982). The results of these studies suggest that, although Asian elephants have a lower low-frequency hearing threshold than other mammals (measured as 17 Hz at 60 dB; Heffner & Heffner 1982), they are considerably less sensitive to frequencies below 100 Hz than to those between 100 Hz and 5 kHz (Heffner & Heffner 1982). The hearing curve itself, therefore, provides some indication that elephants may be better adapted for extracting frequency characteristics in the 115-Hz region (F2) than those in the lower part of the contact call spectrum.

Although the volumes at which we broadcast vocalizations are typical of loud contact calls, they are lower than the maximum volume at which contact calls have been reported (Poole et al. 1988). Because of this, we recognize that louder calling, during extreme social excitement, may facilitate larger recognition distances than those described here. Furthermore, there has been recent interest in the possibility that elephant acoustic signals might be transmitted through seismic as well as airborne waves (O'Connell et al. 1997; O'Connell-Rodwell et al. 2000) and that elephants have the potential to sense ground vibrations through bone conduction (Reuter et al. 1998) and mechanoreception (O'Connell et al. 1998). Although it is unknown whether elephants can extract information on individual identity from such waves, the possibility remains that seismic communication may reinforce or extend the distances over which call recognition takes place.

In conclusion, the most important frequency components for airborne long-distance communication of social identity in African elephants may be well above the infrasonic range. Our results indicate that, when these components become immersed in background noise at propagation distances above 1 km, abilities for long-distance social recognition become limited. Social recognition distances are still considerable, reaching a maximum of 2.5 km in our experiments. Information about the fundamental frequency contour may be the key characteristic used to discern identity, but is likely to be extracted from harmonics in the 115-Hz region. Given that the transmission properties of long-distance calls and the hearing abilities of receivers are such that frequencies around 100 Hz seem to be more important than frequencies below 30 Hz, elephants may produce fundamental frequencies in the infrasonic range simply because of their large size (and vocal folds) rather than as an evolved mechanism for long-distance communication. The apparent incorporation of the trunk into the vocal filter, enabling elephants to emphasize low but audible frequencies in the call spectrum, may be more important for facilitating successful communication over long distances. Our results emphasize that it is unsafe to speculate on the distances over which social communication can take place without identifying which signal characteristics are important in coding the relevant social information, how these decay with distance from the signaller, and directly testing how degradation of the signal with distance affects the perceptual performance of the study animals involved.

## Acknowledgments

This study was funded by BBSRC Research Grant 85/S07659 and Royal Society Research Grant number 13 986 (both to K.M.). D.R. was supported by Fondation Fyssen and E.U. Marie Curie Fellowships. Additional financial support and/or equipment was provided by the Amboseli Elephant Research Project, African Wildlife Foundation, Association for the Study of Animal Behaviour, Newnham College (Cambridge), NERC small project grant GR9/1191, the Nuffield Foundation and the Tusk Trust. We are grateful to the Kenyan Office of the President and Kenya Wildlife Services for permission to conduct the work in Amboseli National Park and to Katito Sayialel and Norah Njiriana for assistance with fieldwork. Barbie Allen provided much appreciated logistic support. We are grateful to Chris Darwin for advice and guidance on sound analysis and to Chris Darwin and Stuart Semple for helpful comments on the manuscript.

## References

**Chariff, R. A., Mitchell, S. & Clark, C. W.** 1995. *Canary 1.12 User's Manual*. Ithaca, New York: Cornell Laboratory of Ornithology.

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

**Fitch, W. T.** 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, **102**, 1213–1222.

**Garstang, M., Larom, D., Raspet, R. & Lindeque, M.** 1995. Atmospheric controls on elephant communication. *Journal of Experimental Biology*, **198**, 939–951.

**Gasc, J. P.** 1967. Squelette hyobranchial. In: *Traité de Zoologie. Vol. 16* (Ed. by P. Grassé), pp. 550–583. Paris: Masson.

**Heffner, R. S. & Heffner, H. E.** 1980. Hearing in the elephant (*Elephas maximus*). *Science*, **208,** 518–520.

**Heffner, R. S. & Heffner, H. E.** 1982. Hearing in the elephant (*Elaphas maximus*): absolute sensitivity, frequency discrimination and sound localization. *Journal of Comparative and Physiological Psychology*, **96,** 926–944.

**Houtsma, A. J. M.** 1995. Pitch perception. In: *Hearing* (Ed. by B. C. J. Moore), pp. 267–295. San Diego, California: Academic Press.

**Langbauer, W. R., Payne, K. B., Charif, R. A., Rapaport, L. & Osborn, F.** 1991. African elephants respond to distant playbacks of low-frequency conspecific calls. *Journal of Experimental Biology*, **157,** 35–46.

**Larom, D., Garstang, M., Payne, K., Raspet, R. & Lindeque, M.** 1997a. The influence of surface atmospheric conditions on the range and area reached by animal vocalisations. *Journal of Experimental Biology*, **200,** 421–431.

**Larom, D., Garstang, M., Lindeque, M., Raspet, R., Zunckel, M., Hong, Y., Brassel, K. & O'Beirne, S.** 1997b. Meteorology and elephant infrasound at Etosha National Park, Namibia. *Journal of the Acoustical Society of America*, **101,** 1710–1717.

**McComb, K.** 1996. Studying vocal communication in elephants. In: *Studying Elephants* (Ed. by K. Kangwana), pp. 112–119. Nairobi: African Wildlife Foundation.

**McComb, K., Moss, C., Sayialel, S. & Baker, L.** 2000. Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour*, **59,** 1103–1109.

**McComb, K., Moss, C., Durant, S. M., Baker, L. & Sayialel, S.** 2001. Matriarchs act as repositories of social knowledge in African elephants. *Science*, **292,** 491–494.

**Markel, J. D. & Gray, A. H.** 1976. *Linear Prediction of Speech*. New York: Springer Verlag.

**Moss, C. J.** 1996. Getting to know a population. In: *Studying Elephants* (Ed. by K. Kangwana), pp. 58–74. Nairobi: African Wildlife Foundation.

**Moss, C. J. & Poole, J. H.** 1983. Relationships and social structure of African elephants. In: *Primate Social Relationships: an Integrated Approach* (Ed. by R. A. Hinde), pp. 315–325. Oxford: Blackwell Scientific.

**O'Connell, C. E., Arnason, B. T. & Hart, L. A.** 1997. Seismic transmission of elephant vocalisations and movement. *Journal of the Acoustical Society of America*, **102,** 3124.

**O'Connell, C., Hart, L. A. & Arnason, B. T.** 1998. Comments on elephant hearing [J. Acoust. Soc. Am. 104, 1122–1123]. *Journal of the Acoustical Society of America*, **105,** 2051–2052.

**O'Connell-Rodmell, C. E., Arnason, B. T. & Hart, L. A.** 2000. Seismic properties of Asian elephant (*Elephas maximus*) vocalizations and locomotion. *Journal of the Acoustical Society of America*, **108,** 3066–3072.

**Payne, K. B., Langbauer, W. R. Jr & Thomas, E. M.** 1986. Infrasonic calls of the Asian elephant (*Elaphas maximus*). *Behavioral Ecology and Sociobiology*, **18,** 297–301.

**Payne, R. & Webb, D.** 1971. Orientation by means of long range acoustic signaling in baleen whales. *Annals of the New York Academy of Science*, **188,** 110–141.

**Poole, J. H.** 1995. Sex differences in the behaviour of African elephants. In: *The Differences Between the Sexes* (Ed. by R. V. Short & E. Balaban), pp. 331–346. Cambridge: Cambridge University Press.

**Poole, J. H., Payne, K., Langbauer, W. R. Jr & Moss, C. J.** 1988. The social contexts of some very low frequency calls of African elephants. *Behavioral Ecology and Sociobiology*, **22,** 385–392.

**Reby, D. & McComb, K.** In press. Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*.

**Rendall, D., Owren, M. J. & Rodman, P.S.** 1998. The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *Journal of the Acoustical Society of America*, **103,** 602–614.

**Reuter, T., Nummela, S. & Hemila, S.** 1998. Elephant hearing. *Journal of the Acoustical Society of America*, **104,** 1122–1123.

**Shoshani, J.** 1998. Understanding proboscidean evolution: a formidable task. *Trends in Ecology and Evolution*, **13,** 480–487.

**Titze, I. R.** 1994. *Principles of Voice Production*. Englewood Cliffs, New Jersey: Prentice Hall.