Acoustic communication is essential in the social relationships of most seabirds, for example mate–mate or parent–chick identification (Falls, 1982). In birds, the best evidence of individual recognition by voice comes from colonial species, where the omnidirectional properties of sound make acoustic signals more reliable than visual signals in crowded colonies (Hutchison et al., 1968; Beer, 1970; White, 1971; Brooke, 1978; Moseley, 1979; Charrier et al., 2001). Among colonial seabirds, penguins identify their mate or parent mainly by an acoustic signal, the display call (Jouventin, 1972, 1982; Proffitt and McLean, 1991; Seddon and Van Heezik, 1992; Aubin and Jouventin, 2002; P. Jouventin and T. Aubin, manuscript in press). In the king penguin _Aptenodytes patagonicus_, one of the largest species of penguin, there is no nest and parents carry the single egg or young chick on their feet. This species breeds in dense colonies of up to 1 million individuals. Mates must find each other so they can shift the egg or chick back and forth as they take turns feeding at sea (Stonehouse, 1960). After this brooding stage, chicks become emancipated, gathering in aggregations (the ‘crèche’ stage), where they wait for their parents to return from foraging trips at sea. The adult (male or female), on return from the sea, makes its way in the colony, calling at regular intervals (Lengagne et al., 1999a). On hearing the display call of one of its parents, the chick calls in reply, and moves towards the parent, often running. Other chicks in the vicinity do not react to this ‘extraneous’ call (Jouventin, 1972, 1982). After acoustic recognition, the adult feeds its chick. King penguins therefore have few, if any, physical landmarks to help in the meeting of individuals, and are thus a good model for the study of acoustical recognition among individuals.

The display call of an adult king penguin (male or female) is a succession of 4–8 similar sound components, the syllables, separated by strong amplitude declines, which coincide with declines in frequency. The structure of the syllable is complex, with two fundamental frequencies (the ‘two-voice’ phenomenon; Grenewalt, 1968) and their corresponding harmonics, strongly modulated in frequency and amplitude (Robisson, 1992a). It has previously been shown that the signal is highly stereotyped within each individual but differs noticeably between individuals. Individuality of information can potentially be supported by either temporal or frequency parameters (Robisson, 1992b; Lengagne et al., 1997); however, even though it is possible for the king penguin to distinguish the acoustic signals of different individuals, how the signal is identified remains unknown. To investigate this

**Localisation of an acoustic signal in a noisy environment: the display call of the king penguin _Aptenodytes patagonicus_**

Thierry Aubin1,* and Pierre Jouventin2

1CNRS-NAMC, UMR 8620, Université Paris-Sud, F-91405 Orsay Cedex, France and 2CNRS-CEFE, UPR 9056, 1919 Route de Mende, 34293 Montpellier, France

*Author for correspondence (e-mail: thierry.aubin@ibaic.u-psud.fr)

Accepted 19 September 2002

**Summary**

King penguin chicks identify their parents by an acoustic signal, the display call. This call consists of a succession of similar syllables. Each syllable has two harmonic series, strongly modulated in frequency and amplitude, with added beats of varying amplitude generated by a two-voice system. Previous work showed that only one syllable of the call is needed for the chick to identify the calling adult. Both the frequency modulation pattern of the syllable and the two-voice system play a role in the call identification. The syllabic organisation of the call, the harmonic structure and the amplitude modulations of the syllables apparently do not contribute to individual recognition. Are these acoustic features useless? To answer to this question, playback experiments were conducted using three categories of experimental signals: (i) signal with only the fundamental frequencies of the natural call, (ii) signal with the amplitude of each syllable kept at a constant level and (iii) signals with only one syllable, repeated or not. The responses of chicks to these experimental signals were compared to those obtained with the calls of their natural parents. We found that these acoustic features, while not directly implicated in the individual recognition process, help the chicks to better localise the signal of their parents. In addition, the redundant syllabic organisation of the call is a means of counteracting the masking effect of the background noise of the colony.

Key words: sound localisation, communication in noise, individual recognition, king penguin, _Aptenodytes patagonicus_.

**Introduction**

Acoustic communication is essential in the social relationships of most seabirds, for example mate–mate or parent–chick identification (Falls, 1982). In birds, the best evidence of individual recognition by voice comes from colonial species, where the omnidirectional properties of sound make acoustic signals more reliable than visual signals in crowded colonies (Hutchison et al., 1968; Beer, 1970; White, 1971; Brooke, 1978; Moseley, 1979; Charrier et al., 2001). Among colonial seabirds, penguins identify their mate or parent mainly by an acoustic signal, the display call (Jouventin, 1972, 1982; Proffitt and McLean, 1991; Seddon and Van Heezik, 1992; Aubin and Jouventin, 2002; P. Jouventin and T. Aubin, manuscript in press). In the king penguin _Aptenodytes patagonicus_, one of the largest species of penguin, there is no nest and parents carry the single egg or young chick on their feet. This species breeds in dense colonies of up to 1 million individuals. Mates must find each other so they can shift the egg or chick back and forth as they take turns feeding at sea (Stonehouse, 1960). After this brooding stage, chicks become emancipated, gathering in aggregations (the ‘crèche’ stage), where they wait for their parents to return from foraging trips at sea. The adult (male or female), on return from the sea, makes its way in the colony, calling at regular intervals (Lengagne et al., 1999a). On hearing the display call of one of its parents, the chick calls in reply, and moves towards the parent, often running. Other chicks in the vicinity do not react to this ‘extraneous’ call (Jouventin, 1972, 1982). After acoustic recognition, the adult feeds its chick. King penguins therefore have few, if any, physical landmarks to help in the meeting of individuals, and are thus a good model for the study of acoustical recognition among individuals.

The display call of an adult king penguin (male or female) is a succession of 4–8 similar sound components, the syllables, separated by strong amplitude declines, which coincide with declines in frequency. The structure of the syllable is complex, with two fundamental frequencies (the ‘two-voice’ phenomenon; Grenewalt, 1968) and their corresponding harmonics, strongly modulated in frequency and amplitude (Robisson, 1992a). It has previously been shown that the signal is highly stereotyped within each individual but differs noticeably between individuals. Individuality of information can potentially be supported by either temporal or frequency parameters (Robisson, 1992b; Lengagne et al., 1997); however, even though it is possible for the king penguin to distinguish the acoustic signals of different individuals, how the signal is identified remains unknown. To investigate this
coding–decoding process, we conducted playback experiments in the field. Birds were tested using temporal or frequency modifications of their natural call (Jouventin et al., 1999; Lengagne et al., 2000, 2001). It appeared that the vocal signature was based on two complementary codes: (1) frequency modulation (FM) shape of the syllable, and (2) the beats generated by the two voices. A chick recognised its parental call and paired mates recognised each other’s calls when only one syllable was played back. Even a modified signal with only the two fundamental frequencies present was still recognised. In addition, king penguins paid little attention to the amplitude modulation (AM) structure of the syllable.

Are the complex AM and harmonic structures completely useless? Is the call organisation in successive syllables unnecessary? In fact, the display call is given in the context of a colony where it is difficult to localise individuals in the moving crowd, and where the acoustics are masked by the noise produced by other individuals. Thus, we hypothesised that these acoustic parameters might aid the emitter (1) to be located easily, and (2) to be discriminated from the background noise of the colony. To test these hypotheses, we played parental calls back to their chicks with either the AM or harmonic or syllabic structures modified. Then, we compared the responses to these signals with those to unmodified (natural) calls.

Materials and methods

Subjects and location

King penguins Aptenodytes patagonicus (Miller 1778) were studied at Possession island, Crozet archipelago (46°25’S, 51°45’E) from December 1995 to January 1996 and at Kerguelen Island (49°14’S, 70°33’E) in March 2000. Field studies were conducted in two large colonies, the first containing about 40000 pairs of adults and 1500 chicks and the second about 60000 pairs of adults and 2300 chicks. Chicks were tested during the ‘crèche’ stage, i.e. when they were more than 4 weeks old. Parental calls were recorded opportunistically from adults on their return to the colony. Each bird was then followed until it achieved a rendezvous with its chick, which after feeding was marked individually by a temporary plastic band around the flipper.

Recording and playbacks

Adult king penguins were recorded using an omnidirectional Sennheiser MD211 microphone mounted on a 2.5 m pole and connected to a Sony TCD10 Pro II DAT recorder (sampling frequency: 44.6kHz, frequency response flat within the range 20–20000Hz). The distance between the beak of the recorded bird and the microphone was about 1 m.

For playback experiments, we used a 4200 Uher tape-recorder (tape-speed 19 cm s⁻¹) connected to a 50 W Audix PH3 selfpowered loudspeaker (frequency response 100–5600 Hz +2 dB). Signals were played at a natural sound pressure level (Robisson, 1993a; Aubin and Jouventin, 1998), about 95 dBsPL (sound pressure level), measured 1 m from the loudspeaker, with a Bruël & Kjaer Sound Level Meter type 2235 (linear scale, slow setting) equipped with a 1 inch condenser microphone type 4176.

Sound synthesis and analysis

Analog signals were digitised with a 16 bit Oros Au21 acquisition card (equipped with a 120 dB/octave anti-aliasing...
Localisation features in penguin calls

Filtering (Mbu-Nyamsi et al., 1994) at a sampling frequency of 16 kHz. Signals were stored on the hard disk of a computer and then examined and modified with the Syntana V1.2 analytical software package (Aubin, 1994).

Three kinds of experimental signals were constructed:

(A) Signal without amplitude modulation

The amplitude modulation (AM) of the adult call (Fig. 1A) was removed by application of the analytic signal calculation (Mbu-Nyamsi et al., 1994). We obtained a signal with the same spectral content, and the same frequency modulation (FM) as those of the parental call, but without any AM (Fig. 1B). In fact, for this signal, the amplitude was kept at a constant level, corresponding to the mean level of the parental call. The calculation also kept the two harmonic series of the two-voice system and consequently the beats were not removed. It should be pointed out, however, that chicks did not receive a true ‘without AM’ signal due to propagation through the colony generating constructive and destructive interference. Nevertheless, after short range propagation (see playback procedure below), the envelope of the signal is weakly modified (Aubin and Jouventin, 1998) and the amplitude modulation shape of the experimental signal is modified at random by the environment, not by following the natural shape of a penguin call.

(B) Signal keeping only the fundamental frequencies

To remove the harmonics, we used the optimal digital filtering method with the fast Fourier transform (FFT) as described by Press et al. (1998) with a window size of 4096 points. As natural calls were strongly modulated in frequency, three or four successive filtration steps in the temporal domain were necessary to suppress the harmonics of the different syllables (Fig. 1C).

(C) Signals with only one syllable

We built a series of signals retaining only the first syllable of the call. The series consisted of four signals with either the first syllable alone, or with the 2, 4 or 8 first syllables repeated (Fig. 2). Eight successive syllables corresponds to the number of syllables regularly encountered in natural parental calls (Robisson, 1993b). For each chick tested and for the signals with repeated syllables, the temporal rhythm syllable/silence corresponded to the mean temporal rhythm measured in the parental call (i.e. a mean ± s.e.m. of 531±11 ms for the syllable and of 55±3 ms for the inter-syllable silence, N=16).

Playback procedure and criteria of responses

Tests were conducted between 10.00 am and 5.00 pm, during clear and dry weather, with a wind speed less than 4 ms⁻¹. The chicks were generally resting in the colony, preening themselves. Series A and series B experiments were conducted at Crozet island and series C at Kerguelen island.

To test the locatability of the signals (experiments with signals of series A and B), the following playback procedure was used. The loudspeaker was first put in the colony at a distance of 12 m from the test chick. At this distance, and with a normal density of birds (about 2–3 individuals m⁻² in a colony of king penguins, according to Barrat, 1976) the chick can detect, recognize and localise the parental call without ambiguity (Aubin and Jouventin, 1998). Then, the parental call or the experimental signal (on average 3.5 s long) was broadcast repetitively with a 7 s time interval of silence between each signal. This repetition rate is similar to that of an adult calling while searching for its mate or chick (Robisson, 1993b). The broadcast ended when either 12 signals had been emitted or as soon as the distance between the tested chick and the loudspeaker was below 2 m. At this distance, we estimated that the ‘meeting’ was realised, that is, that the sound source was accurately found by the chick. The number of calls broadcast was then noted. To avoid habituation, no more than 12 signals were broadcast. If the distance chick-loudspeaker remained greater than 2 m after the broadcast of the whole series of signals, a score of 12 was attributed to this experimental session. For series A, 16 chicks were tested and for series B, 15 chicks. Each chick was tested with two series of signals, the experimental one (series A or series B signals) and the control one (natural parental call series) and on 2 consecutive days, with one experiment each day. The order of presentation of the two series was randomised for the different chicks tested. Hence, the observed responses for the whole group of chicks were not dependent on playback order.
To test the ability of chicks to discriminate the signal from the background noise (experiments with series C signals), the following playback procedure was used. The loudspeaker was placed in the colony at a distance of 7 m from the test chick, which corresponds to the natural calling distance of an adult (Robisson, 1993b; Lengagne et al., 1999a). Series C signals were played to the tested chick only when the ambient noise, measured with the Sound Level Meter, was between 70 and 75 dB SPL, i.e. near the average values usually observed in a king penguin colony (70 dB SPL, Robisson, 1993b; 74 dB SPL, Aubin and Jouventin, 1998). All of the 16 chicks tested received the whole series of four signals. To prevent habituation, each chick was tested only once a day with one of the signals of the series. As previously, the order of presentation of the signals from day to day was randomised for the different chicks. To evaluate how chicks detected and recognized the playback signal, a four-class behavioural scale was used. The scale was ranked as follows: class 0: no reaction; class 1: agitation, head turning; class 2: head turning, calls in reply to the signal broadcast; class 3: head turning, calls in reply to the signal broadcast, approaches in the direction of the loudspeaker and stops in the vicinity (less than 2 m).

This behavioural scale is similar to the one used in our previous playback studies with king penguins (Aubin and Jouventin, 1998, 2002; Jouventin et al., 1999; Lengagne et al., 2000).

**Statistical analysis**

As data were not normally distributed, non-parametric analyses were performed. To compare paired samples (series A and B), a Mann–Whitney U-test was used. To compare paired samples in more than two categories (series C), we used a Kruskal–Wallis test together with exact two-sided P values. As the same marginal distribution was used through several comparisons, we used the Bonferroni Dunn–Sidak method (Ury, 1976) to assess the final significance of the test. All statistics computations were carried out with Statistica V5 and StaXact V3 (Cytel, 1995) softwares.

**Results**

**Series A and B signals**

Both natural and modified signals were recognised by the chick without ambiguity. The chick was quiet and silent before the experiment, systematically called in reply and moved when these signals were emitted. Nevertheless, the number of broadcast calls necessary for the chick to localise the sound source is significantly more important with a signal without the natural AM than with a signal with the natural AM (mean number ± s.d. of broadcast calls: 9.81±2.61 for the signal without natural AM and 5.00±3.20 for the control signal; Z=−3.302, P<0.001, N=16). In the same way, the chicks had significantly more difficulties in localising the signal keeping only the fundamental frequencies than for the natural call with the complete harmonic series (mean number ± s.d. of broadcast calls: 9.27±2.68 for the signal keeping only the fundamental frequencies and 5.47±3.29 for the control signal; Z=−3.199, P<0.001, N=15).

**Series C signals**

The broadcast of only one syllable, against a normal background noise, elicited weak responses. Only one chick approached the sound source, two chicks called in reply and the 13 remaining others were silent and stayed at the same place. With two syllables broadcast, 75% of birds (12/16) called in reply and one chick approached the sound source. This difference was significant compared with when one syllable was broadcast (Table 1). With four identical syllables broadcast, all the chicks called in reply, and 75% of them approached the sound source. Compared to the previous situation, the difference was highly significant (Table 1). Finally, there was no significant difference in the responses between broadcasting 4 or 8 syllables (Table 1).

**Discussion**

To facilitate the meeting of mates or parents and their chick in an extremely crowded and noisy environment, king penguins have developed remarkable acoustic capacities in both sound production and recognition. They use a twin coding system, which allows the production of a huge number of vocal signatures and thus drastically minimises the confusions in the identification process (Lengagne et al., 2001; Aubin and Jouventin, 2002). As mentioned previously, this coding system is based upon the characteristics of the slow frequency modulation of each syllable and of the beats generated by the two voice system. These acoustic parameters propagate well through obstacles, such as those constituted by the large bodies of penguins gathered in a flock (Lengagne et al., 1999b; Aubin et al., 2000). Nevertheless, this ‘anti-confusion’ and ‘anti-obstacle’ coding system is not sufficient by itself to ensure finding the right chick in a crowd. The signal has to be discriminated from the background noise and localised. In the present study, according to the first experiment, the chick needs significantly fewer calls to localise the source of emission for

<table>
<thead>
<tr>
<th>Number of</th>
<th>Intensity of responses</th>
<th>Paired comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>repetitions of the first syllable</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1. Responses of chicks to the first syllable of the parental call played back 1, 2, 4 or 8 times.
the natural signal than for signals with amplitude modulation removed or with only the fundamental frequencies kept. With natural calls, the move is oriented directly towards the sound source, whereas with modified signals the searching behaviour is more erratic. This explains why more calls, and consequently more time, are necessary for the chick to come near the sound source when modified signals are broadcast. Thus, natural calls appear to be more locatable than the same calls with natural AM or with the harmonics removed. These results are in agreement with those obtained in psycho-acoustic experiments: birds can locate wide-spectrum signals better than any pure tone, and can locate sounds with sharp amplitude changes better than sounds weakly modulated in amplitude (Wiley and Richards, 1978, 1982; Dooling, 1982). This ability has been particularly well studied experimentally in the laboratory using the barn owl *Tyto alba*, a species known to be uniquely specialised for the detection of prey (Konishi, 1973, 1977; Shalter and Schleidt, 1997; Knudsen and Konishi, 1979). In fact, it appears that wide spectrum signals with sharp amplitude modulations recur in the vocalisations of many colonial birds (White and White, 1970; Wiley, 1976; Mathevon, 1997; Charrier et al., 2001), i.e. in an environment where the estimation of the direction of a signal is particularly useful.

The display call of the king penguin comprises a series of 4–8 syllables that are more or less similar and, as mentioned previously, only one (for instance the first) is sufficient to elicit recognition. This was demonstrated in a previous study (Jouventin et al., 1999). During these earlier experiments, however, the one syllable signal was broadcast only during relative periods of silence, that is, when adults in the vicinity were not emitting calls. In fact, under more usual conditions, as in the present study (i.e. when the level of the background noise is about 70–75 dB), the broadcast of only one syllable is not sufficient to draw the chick towards the loudspeaker. A majority of birds did not react at all or only raised their heads, while staying in the same place. From two identical syllables broadcast, almost all the birds call in reply, indicating that the call has been discriminated and recognised. Nevertheless, a majority of birds do not move towards the sound source. A possible explanation is that the time duration of the signal (2 syllables of about 500 ms each) is not sufficient to localise the sound source. This explanation seems implausible, since it has been demonstrated that small birds are able to localise a 200 ms duration pure tone or noise with a precision of about 20° in azimuth (Klump et al., 1986; Park and Dooling, 1991). Birds with larger heads, such as the king penguin, are likely to be better for localisation than small birds, since behavioural and physiological studies have shown that large birds use mainly the interaural time difference for localisation in azimuth (studies on the barn owl; Moiseff and Konishi, 1981; Moiseff, 1989). Another possible explanation for this lack of movement towards the source after hearing two syllables could be linked to the gregarious behaviour of the king penguin chicks. Because of predation risks, chicks are reluctant to leave the crèche (Stonehouse, 1960). However, a more insistent appeal, with 4–8 syllables as in our experiment, would encourage the chick to go out of the crèche, the parental signal being recognised and localised with certainty. Effectively, it is only with the broadcast of four syllables that almost all the birds move closer to the loudspeaker. Thus it appears that at least four syllables are necessary and sufficient to ensure a successful adult–chick meeting. From this series of experiments, it can be stated that a single syllable, although containing the whole identity code, is not sufficient to convey with security the information, even at short range (7 m), given the noisy environment of the colony. As predicted by the mathematical theory of information (Shannon and Weaver, 1949), king penguins must repeat the code to ensure the communication in context of the colony. In addition, a phenomenon called temporal summation predicts considerable improvement in the detectability of signals in a noisy environment by increasing the signal duration (Klump, 1996). Temporal summation seems to be widespread in birds, i.e. budgerigars (Dooling and Searcy, 1985), starlings (Klump and Maier, 1990), canaries (McGregor et al., 1997).

In conclusion, king penguins use a particularly efficient acoustic system of communication, allowing recognition and localisation of individuals within a very constraining environment: a colony of thousands of birds, with a high background noise and a lack of visual cues. The acoustic features of the call that are not directly implicated in the coding process, such as the wide spectrum, the strong amplitude modulations and the redundant syllabic organisation, permit better localisation of the emitter and serve to counteract the masking effect of the background noise of the colony.

The study was supported by the Institut Français pour la Recherche et la Technologie Polaires and the Centre National de la Recherche Scientifique. Logistic support was provided by Terres Australes et Antarctiques Françaises. Special thanks to T. Lengagne for help in the field and in statistical analysis and to two anonymous referees for improvement of the manuscript. We are also grateful to F. S. Dobson for comments and improvement of the English.

References


