Behav Ecol Sociobiol (2010) 64:1915–1926
DOI 10.1007/s00265-010-1065-2

METHODS

# Bioacoustic distances between the begging calls of brood parasites and their host species: a comparison of metrics and techniques

**Louis Ranjard · Michael G. Anderson · Matt J. Rayner · Robert B. Payne · Ian McLean · James V. Briskie · Howard A. Ross · Dianne H. Brunton · Sarah M. N. Woolley · Mark E. Hauber**

Received: 27 July 2010 / Revised: 31 August 2010 / Accepted: 13 September 2010 / Published online: 25 September 2010


**Abstract** A variety of bioacoustics distance metrics have been used to assess similarities in the vocalizations of different individuals. Here, we provide a detailed analysis of several acoustic similarity indices, some of which have been developed with the specific aim of characterizing the sensory coding of auditory stimuli. We compare different approaches through the analysis of begging calls of several passerine species and specialist brood parasitic cuckoos that putatively evolved to mimic their hosts. The different bioacoustics distances did not provide consistently correlated similarity patterns, implying that they are sensitive to different sound features. However, the encoded spectrogram alignment method was correlated with all other acoustic distance metrics, suggesting that this method

Communicated by M. Soler

Louis Ranjard, Michael G. Anderson contributed equally to the research

**Electronic supplementary material** The online version of this article (doi:10.1007/s00265-010-1065-2) contains supplementary material, which is available to authorized users.

L. Ranjard (✉) · H. A. Ross
Bioinformatics Institute, School of Biological Sciences,
University of Auckland,
Private Bag 92019, Auckland Mail Centre,
Auckland 1142, New Zealand
e-mail: l.ranjard@auckland.ac.nz

M. G. Anderson · D. H. Brunton
Ecology and Conservation Group,
Institute of Natural Sciences,
Massey University,
Albany Campus, Private Bag 102-904,
North Shore Mail Centre,
Auckland, New Zealand

M. J. Rayner · M. E. Hauber
School of Biological Sciences,
University of Auckland,
Private Bag 92019,
Auckland, New Zealand

R. B. Payne
University of Michigan,
1306 Granger Avenue,
Ann Arbor, MI 48109, USA

I. McLean
Department of Zoology, Otago University,
P.O.Box 56 Dunedin, New Zealand

J. V. Briskie
School of Biological Sciences, University of Canterbury,
Private Bag 4800,
Christchurch, New Zealand

S. M. N. Woolley
Department of Psychology, Columbia University,
406 Schermerhorn Hall, 1190 Amsterdam Ave,
New York, NY 10027, USA

M. E. Hauber
Department of Psychology, Hunter College,
City University of New York,
695 Park Avenue,
New York, NY 10065, USA

*Present Address:*
M. J. Rayner
National Institute of Water and Atmospheric Research Ltd. (NIWA),
P.O. Box 99940, Newmarket 1149, New Zealand

provides a consistent approach to use when the perceptually salient sound parameters are unknown for a particular species. Our analyses confirm that statistical similarity of begging calls can be detected in a New Zealand pair of host and specialist parasite species. We also show detectable similarity in two other Australasian host–parasite pairs and another New Zealand system, but to a more limited extent. By examining phylogenetic patterns in the begging call diversity, we also confirm that specialist cuckoos have evolved to mimic the begging calls of their hosts but host species have not co-evolved to modify their calls in response to begging call similarity by the parasite. Our results illustrate that understanding the function and mechanism of behavioral copying and mimicry requires statistically consistent measures of similarity that are related to both the physical aspects of the particular display and the sensory basis of its perception.

**Keywords** Begging call · Bioacoustics distance · Brood parasite · Spectrogram alignment · Spectro-temporal modulation

## Introduction

To understand the evolution of copying and mimicry, we require biologically and statistically relevant and consistent metrics. This is especially important in the absence of detailed knowledge of the salient features of the particular behavioral displays that are imitated and perceived. For example, to quantify acoustic similarity of complex natural sounds, including the vocalizations of birds, several bioacoustic distance measures have been formulated (Tchernichovski et al. 2000; Woolley and Rubel 2002; Cardoso et al. 2007; Vallejo et al. 2007; Guerra et al. 2008; Lesica and Grothe 2008; Sewall 2009; Vyas et al. 2009). Ideally, such metrics of acoustic distance should reflect the distinctive features of the sound that a bird can perceive and produce (Gill et al. 2006; Hauber et al. 2007).

If a distance measure is sensitive to the same characteristics of the sound waves as the bird's sensory system, then one can consider the relationships between the sounds that are described by this measure as biologically relevant. There are many ways to extract parameters for representing a sound wave from the time and the frequency domain. Here, we compare five similarity measures commonly used in bioacoustic studies, including a recently developed encoded spectrogram alignment distance approach (Ranjard and Ross 2008), which might be sensitive to different characteristics of the sound waves of avian vocalizations. Few studies have simultaneously compared different sound similarity metrics to test the consistency, efficacy, and biological relevance of alternative measures (Baker and Logue 2003). Our first aim in this study is to determine the extent to which these different distances are correlated. If two approaches define correlated distances, this means that the corresponding methods are sensitive to the same set of acoustic aspects or a correlated set of different acoustic aspects of the vocalizations. Alternatively, if different metrics are not correlated, it will be necessary to investigate which method is more appropriate for the particular sensory perceptual system of the receiver species under study. Investigating which acoustical properties of the songs are biologically relevant requires field experiments and neurophysiological approaches (Margoliash 1983; Naugler and Ratcliffe 1992; Woolley et al. 2005; Matheron et al. 2008).

Bioacoustic distance metrics have the potential to provide a powerful tool to test the evolutionary hypotheses about avian vocalizations. In the second part of our study, we compare the similarity scores of a suite of hypothesized mimetic begging calls of avian brood parasites and their actual and potential host species. Avian interspecific brood parasites lay their eggs in the nests of other species and thus exploit the parental investment of their hosts (Davies 2000). This form of parasitism has strong reproductive costs for the host, either through reduced growth of their own offspring (Hauber 2003) or complete loss of that breeding attempt through eviction behavior of the parasitic chick (Honza et al. 2007; Anderson et al. 2009a). In response, brood parasite hosts have evolved rejection by discrimination of eggs (Davies et al. 1996; Moskát and Hauber 2007) or nestlings (Grim et al. 2003; Langmore et al. 2003). These responses include recognition-free as well as recognition-based mechanisms (Anderson and Hauber 2007; Langmore et al. 2009; Soler 2009).

Nestlings of some cuckoo species (family Cuculidae) emit begging calls which resemble the host species. By mimicking the host's begging call, the cuckoo nestlings eliminate a potential cue that the parents could use to recognize brood parasites (Grim 2006). As a consequence, the host may be selected to modify its begging call to increase its ability to discriminate between its own young and the brood parasite's offspring, as part of an escalating coevolutionary arms-race process (Langmore et al. 2003). However, Anderson et al. (2009b), using a single approach to estimate bioacoustic similarity, showed that although the begging calls of the shining cuckoo (*Chalcites lucidus lucidus*) closely match those of its host species in New Zealand, the grey warbler (*Gerygone igata*), the begging calls of the grey warbler have not changed in response. In this example, the begging calls of the parasite are closer to its host's than to any other sympatric potential host species. Here, we use several distinct bioacoustic distance metrics to test whether that pattern is robust to different methods of acoustic similarity assessments and to investigate if similar evolutionary outcomes have occurred in other hosts of the

shining cuckoo as well as in the hosts of the Horsfield's bronze-cuckoo (*Chalcites basalis*) in Western Australia, and the long-tailed cuckoo (*Urodynamys taitensis*) in New Zealand. Examples of begging calls for some of the studied species are shown in Fig. 1.

If the begging calls of the cuckoos have evolved to match the begging calls of their hosts, including the ability of an individual cuckoo to vary their begging displays to solicit parental provisioning of their current host species (Butchart et al. 2003; Langmore et al. 2008), it is expected that the bioacoustic distances between the begging calls of cuckoos and their hosts will be smaller than between the calls of other sympatric and phylogenetically related non-host species. For example, the shining cuckoo and the Horsfield's bronze-cuckoo are phylogenetically related (genus *Chalcites*, Christidis and Boles (2008)) and so their begging calls are predicted to be similar under the null hypothesis of evolutionary inertia. However, if a cuckoo chick modifies its begging call to sound as the host species (Langmore et al. 2008), it is expected that acoustic features of its call will approximate those of the host's begging call. As a consequence, these acoustic features will no longer represent the phylogenetic relationships between these cuckoos, but instead, each cuckoo's begging call will reflect the acoustic properties of their host, or at least the sensory cues that the host uses to discriminate foreign nestlings or to provide parental provisions (Hauber and Kilner 2007). In a similar way, it is important to generalize previous conclusions based on a single host species (Anderson et al. 2009b) by comparing the begging calls of the shining cuckoo when it parasitizes different host species. If the shining cuckoo does not change its call to match its actual host, then the begging calls of the Australian and the New Zealand subspecies are expected to be similar due to phylogeny. By using such a comparative approach (described in detail in Anderson et al. (2009b)), we aim to detect discordance between phylogenetic versus the host–parasite-based begging call acoustic relationships.

Overall, to relate different bioacoustic methodologies with the goals of our coevolutionary analyses, we use several metrics to estimate acoustic similarity on the same data set. Our aim is to evaluate the relative efficacy of each biacoustic method to determine the similarity of disparate begging calls, with applications for diverse evolutionary

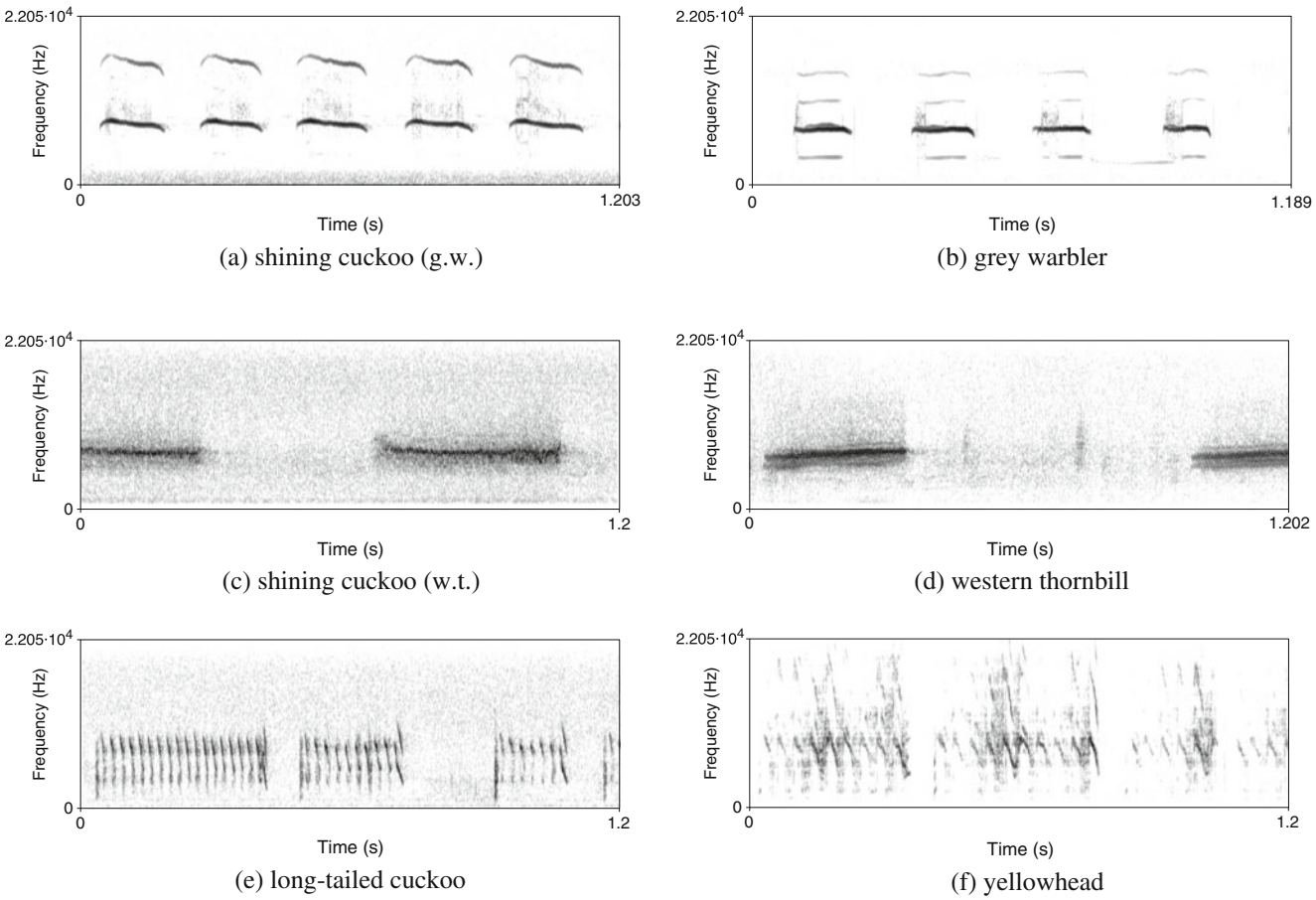


**Fig. 1** Spectrograms of the begging calls of three host–parasite pairs. The names of the shining cuckoo hosts are in brackets and abbreviated for *g.w*, grey warbler and *w.t.*, western thornbill

contexts. This should allow for a more robust method to consistently evaluate the similarity of acoustic signals across diverse species of varied phylogenetic, geographic, and biotic distances.

## Materials and methods

### Study species and sites

We compared the begging calls of the two New Zealand brood parasites, the shining cuckoo and the long-tailed cuckoo, and their respective hosts. To rule out alternative interpretations for begging call similarity due to phylogeny, we included New Zealand native passerines that could be alternative host species (Anderson et al. 2009b). Additionally, we included another subspecies of the shining cuckoo from Australia, *C. lucidus plagosus*, and two of its local hosts, as well as another closely related cuckoo, the Horsfield's bronze-cuckoo, and one of its host species. This begging call data set allows us to investigate whether begging call similarity is due to phylogenetic similarity among brood parasites and their hosts, as well as testing if begging calls are similar amongst closely related brood parasites, or if they are modified to match their respective host species.

Specifically, we recorded begging calls directly from nests of 17 New Zealand native passerine species, shining cuckoos in nests of the grey warbler and two other non-passerine species, the orange-fronted parakeet, *Cyanoramphus malherbi*, and the forest kingfisher, *Halcyon sancta* (see Anderson et al. (2009b) for details about locations and sound recording methods used; Table 1). This previously published data set allowed us to test questions about the evolution of host–parasite begging calls in one particular host–parasite system (Anderson et al. 2009b) as on the main islands of New Zealand, the shining cuckoo is a brood parasite of only the grey warbler (Gill 1983; McLean and Waas 1987).

We then extended this data set by adding recordings of other parasite species and their hosts in New Zealand and Australia (Table 2). Another subspecies of the shining cuckoo occurs in Australia and uses different host species. We used recordings of shining cuckoo nestlings from Western Australia, (Latham 1802), and two of its host species, the western thornbill (*Acanthiza inornata*) and the yellow-rumped thornbill (*Acanthiza chrysorrhoa*; Brooker and Brooker 1989; Langmore and Kilner 2007). The Horsfield's bronze-cuckoo is, amongst other hosts, a brood parasite of the western thornbill, an endemic species of Australia (Payne 2005). Further, we added the begging calls of the other New Zealand brood parasite, the long-tailed cuckoo. The long-tailed cuckoo is a brood parasite of three native species of New Zealand passerines, laying its eggs in the nests of yellowhead (*Mohoua ochrocephala*), whitehead (*Mohoua albicilla*), and brown creeper (*Mohoua novaeseelandiae*; McLean and Waas 1987; McLean 1988). Begging calls were recorded from fledgling long-tailed cuckoos on Little Barrier Island (see McLean and Waas (1987) for more details).

### Recording of begging calls

A total of 462 calls from 47 nests were recorded (Table 1). Begging calls from most New Zealand species were recorded from broods under natural situations during parental feeding visits by setting up a microphone as close as possible to the nest without causing disturbance (usually 20–30 cm; Anderson et al. 2009b). We controlled for nestling development by attempting to record nestlings on the day that primary feathers emerged from the sheaths (Briskie et al. 1999), as determined by either direct inspection or the age of nestlings. However, some instances required nestlings to be recorded opportunistically. If age could not be determined, nestlings from the mid to late stages of development that were responding vocally to parental nest visitations were recorded, as this is when signs of mimicry should be exhibited by nestlings (Langmore et al. 2008). Nestling begging calls were recorded with a Sennheiser ME 66 microphone or a Panasonic RP-VC201 stereo tie-clip microphone, depending on nest accessibility, onto a Sony MZ-NH700 Hi-MD Minidisc, with a sampling rate of 44.1 kHz.

The begging calls of long-tailed cuckoo and the Australian species were recorded onto cassette tape. Details of the sound recording equipment used can be found in McLean and Waas (1987) and Payne and Payne (1998), respectively. Begging calls from cassette tapes were digitized using a Marantz PMD201 portable cassette recorder (by MGA). The cassette recorder was connected to a desktop computer (Digitizing board: Creative™ sound blaster Audigy LS soundcard; 24 bit, 96 kHz) and recorded in Raven as 16 bit, 44.1 kHz.wav files. Recordings were subsequently examined in Raven, version 1.3 (Charif et al. 2007). Sound recordings were digitized and visualized as spectrograms (Hann, window size 5.33 ms, 3-dB bandwidth of 270 Hz, frequency grid DFT size 256 samples and 188 Hz) for quality verification and subsequent analysis.

### Metrics of acoustic distances

We used five different acoustic comparison methods to obtain the bioacoustic distances among the 462 calls. All pairwise distances were then estimated between the 23 species. The distance matrix quantifies similarity among

**Table 1** Number of begging calls recorded for each species

| Species | Number of calls | Number of nests | Location |
|---|---|---|---|
| Bellbird (*Anthornis melanura*) | 10 | 1 | NZ |
| Brown creeper (*Mohoua novaeseelandiae*) | 10 | 1 | NZ |
| Fantail (*Rhipidura fuliginosa*) | 20 | 2 | NZ |
| Grey warbler (*Gerygone igata*) | 21 | 2 | NZ |
| Hihi (*Notiomystis cincta*) | 20 | 2 | NZ |
| Horsfield's bronze-cuckoo (*Chalcites basalis*) | 20 | 2 | WA |
| Kingfisher (*Halcyon sancta*) | 8 | 1 | NZ |
| Kokako (*Callaeas cinerea*) | 20 | 2 | NZ |
| Long-tailed cuckoo (*Urodynamys taitensis*) | 10 | 1 | NZ |
| Orange-fronted parakeet (*Cyanoramphus malherbi*) | 10 | 1 | NZ |
| New Zealand pipit (*Anthus novaeseelandiae*) | 10 | 1 | NZ |
| Rifleman (*Acanthisitta chloris*) | 20 | 2 | NZ |
| North Island robin (*Petroica australis*) | 20 | 2 | NZ |
| Rock wren (*Xenicus gilviventris*) | 10 | 1 | NZ |
| Saddleback (*Philesturnus carunculatus*) | 10 | 1 | NZ |
| Shining cuckoo [g.w.] (*Chalcites lucidus lucidus*) | 10 | 1 | NZ |
| Shining cuckoo [w.t.] (*Chalcites lucidus plagosus*) | 29 | 3 | WA |
| Shining cuckoo [y.r.t.] (*Chalcites lucidus plagosus*) | 30 | 3 | WA |
| Silvereye (*Zosterops lateralis*) | 30 | 3 | NZ |
| Tomtit (*Petroica macrocephala*) | 10 | 1 | NZ |
| Tui (*Prosthemadera novaeseelandiae*) | 20 | 2 | NZ |
| Welcome swallow (*Hirundo tahitica*) | 29 | 3 | NZ |
| Western thornbill (*Acanthiza inornata*) | 29 | 3 | WA |
| Whitehead (*Mohoua albicilla*) | 26 | 3 | NZ |
| Yellowhead (*Mohoua ochrocephala*) | 30 | 3 | NZ |

The name of the shining cuckoo hosts are in brackets and abbreviated for *g.w.,* grey warbler; *w.t.,* western thornbill; and *y.r.t.,* yellow-rumped thornbill

*NZ* New Zealand, *WA* Western Australia

species on the basis of the acoustic analysis of their begging calls. The shining cuckoo was considered independently when it had different hosts in different localities (New Zealand and Western Australia) so that for each method we created a matrix of species pairwise distances (25×25) that were calculated from the begging call distances matrix (462×462; details below):

**Table 2** Host species of the different brood parasites in New Zealand (NZ) and Western Australia (WA) that were used for our analyses

| Parasite | Host |
|---|---|
| Shining cuckoo | Grey warbler (NZ) |
| | Yellow-rumped thornbill (WA) |
| | Western thornbill (WA) |
| Horsfield's bronze-cuckoo | Western thornbill (WA) |
| Long-tailed cuckoo | Yellowhead (NZ) |
| | Whitehead (NZ) |
| | Brown creeper (NZ) |

This is not a complete list of host species used by the shining cuckoo or the Horsfield's bronze-cuckoo

1. Spectrogram alignment (dynamic time warping)

Cepstrum coefficients are commonly used in speech processing to encode the spectrogram of the sound signal (Davis and Mermelstein 1980). They represent the balance of energy between the different frequencies of the spectrum. To compute these coefficients, the spectrum is often modified to map the sensitivity of the human ear by using, for example, the mel scale, leading to the definition of mel-frequency cepstrum coefficients. Even if not originally designed for bioacoustic analyses, such an approach has been shown to be effective for the study of animal vocalizations (Kogan and Margoliash 1998; Trawicki et al. 2005; Clemins and Johnson 2006; Ranjard and Ross 2008; Tao et al. 2008; Brown and Smaragdis 2009).

A sequence of mel-frequency cepstrum coefficients was extracted from each call using the Hidden Markov Model Toolkit (HTK; Young 1994). The first of the 26 filterbank channels started at 1 kHz and the last one terminated at 22.05 kHz. Under a sampling frequency of 44.1 kHz, a Hamming window of 128 samples with 50% overlap was used for computing the spectra and the signal had first-order pre-emphasis applied using a

coefficient of 0.97. Twelve coefficients were calculated and the 0th cepstral parameter was used as the energy component. Two frames before and two frames after the current one were used to estimate the first and second-order temporal derivatives. Energy normalization was implemented by subtracting the maximum value of the energy and adding 1.0. The cepstral coefficients were rescaled by liftering the cepstra using a coefficient of 22 so that they had similar magnitudes.

Next, an average vector sequence was constructed for each species. First, the pairwise distances between all the calls of a particular species were calculated using dynamic time warping as described in Ranjard and Ross (2008). Then, the average sequence between the two most distant calls was computed from the alignment by using a weight of 0.5. This sequence was derived from the warping path (see Ranjard and Ross (2008) for a description of the algorithm). The pairwise distance matrix was recalculated using this average sequence instead of the two previous ones. Therefore, the rank of the matrix was decreased by one at this stage. This process was repeated until only one sequence remained, this sequence representing an average of all call sequences of this species. The method has been implemented in a software package, "DTWave", which can be freely downloaded from http://www.cebl.auckland.ac.nz/DTWave.php. Finally, the species average sequences were aligned by dynamic time warping in order to get the species pairwise begging call distances (25×25 matrix).

2. Spectro-temporal modulation

The modulation power spectrum represents the spectro-temporal modulations of a sound (Singh and Theunissen 2003). More precisely, it indicates the distribution of the specific spectral and temporal modulations contained in the sound recordings being analyzed. This statistical representation of a sound pressure wave is independent of the phase of its different Fourier components. Only the temporal and spectral modulations of these components are presented. Woolley et al. (2005) showed that the zebra finch (*Taeniopygia guttata*) auditory system has selective tuning for spectro-temporal modulation that serves as a mechanism for the discrimination of natural sounds. This indicates that these sound metrics are biologically relevant for the sensory and perceptual systems of this particular species, and perhaps of songbirds in general. Therefore, this method could be a biologically relevant representation of the sound for comparing begging calls.

A modulation power spectrum can be constructed from a group of sounds for which the power density of the components are plotted in a two-dimensional graph. For each Fourier component, the graph represents the temporal modulations on the first axis and the spectral modulations on the second axis. The modulation power spectrum was constructed under Matlab (The Mathworks and Natick 2008) for each species using all the begging calls available for that species. To measure the similarity between species, we computed the two-dimensional correlation coefficients between the resulting spectro-temporal modulation plots. The species pairwise begging call distance was defined here as one minus the correlation coefficient and was used to generate a 25×25 distance matrix.

3. Spectrogram correlation

The computer program Raven is a widely used bioacoustic analysis program, designed to both generate spectrograms and measure sound features. Raven now includes a sound correlation feature that compares two sounds simultaneously and computes a similarity score, based on either the spectrogram or the waveform. This form of analysis has been designed with the analysis of bird song in mind, although the technique should work for any acoustic files.

We created a spectrogram of each call using Raven, version 1.3 (Charif et al. 2007). The discrete Fourier transform was calculated using a Hann window of 235 samples with 50% overlap and 256 frequency bins. A linear scale was used for the power values of the spectrograms, expressing them in squared amplitude units. The distance between each pair of calls was defined as one minus the maximum value of the pairwise spectrogram correlation coefficients, leading to a value ranging between 0, for maximum similarity, and 1. This technique finds the best correlation coefficient by sliding one spectrogram past another so that different time lags are tested. For each possible time lag, the correlation coefficient was calculated and only the highest coefficient was retained. These correlation coefficients were used to construct a begging call pairwise distance matrix (462×462). The species pairwise begging call distance (25×25 matrix) was determined by averaging the call distances for each pair of species.

4. Sound Analysis Pro

The Sound Analysis Pro (SAP) program (Tchernichovski et al. 2000) offers a way to compute a distance between a pair of sounds based on the comparison of different extracted features. This software was developed for the purpose of analyzing bird song, in particular, zebra finch song. As such, the degree of similarity is best for complex sounds, such as comparing two songs, as differences in several acoustic features. This approach compares the sounds by calculating the Euclidean distance between every pair

of song frame feature vectors. The song frames are 70-ms long. The spectral structure of each frame is summarized by the measurements of five features: pitch, frequency modulation, amplitude modulation, Wiener entropy, and goodness of pitch (a description of these features is available in Tchernichovski et al. (2000)).

To compute the Euclidean distance, the units for each feature were transformed to units of statistical distances, the median absolute deviation from the mean. The Euclidean distance was calculated using the symmetric timecourses option so that the highest priority was given to the sequential match of the analysis frames. In other words, this distance is sensitive to the phase of the signal. A slight time warping of 5% was allowed in order to find the highest similarity between the pairs of sound frames. To arrive at an overall score of similarity between calls, the mean Euclidean distance was computed across all pairs of a frame. SAP was used to obtain the begging call pairwise distances (462×462 matrix). The species pairwise distances (25×25 matrix) were computed by averaging the call distances for each pair of species.

5. Sound features (Euclidean distance from feature extraction)

In many bioacoustic studies, the sounds being studied are characterized by measuring sound features of each song or call. These extracted data are then used for subsequent analysis, such as comparisons amongst sounds. We sought to replicate this method using five acoustic features that were extracted from spectrograms in order to represent each call using the program Sound Analysis Pro (Tchernichovski et al. 2000). These parameters were: mean frequency modulation, mean amplitude modulation, mean entropy, mean frequency, and call duration (described in detail in Anderson et al. (2009b)). These sound parameters are typical of many acoustic studies that analyze sound by extracting sound features. Normalization was applied for each of them and the Euclidean distance between each vector representing a call was used as the pairwise call distance. Then, the begging call distances were averaged for every pair of species resulting in the species pairwise distances (25×25 matrix).

### Correlation between bioacoustic distances

By calculating the Pearson product-moment correlation coefficient between the 25 by 25 distance matrices of the five methods, we could verify which distance measures were correlated. For every pair of call distance matrices, Mantel tests of correlation were performed to determine if correlations were statistically significant (Sokal and Rohlf 1995). For each pair of distance matrices, 10,000 random replicates of permutation were performed. The permutations generated a null distribution of correlation coefficients from which a $p$ value was computed for the observed coefficient. The $p$ value represents the percentage of random replicates which returned a coefficient of correlation greater or equal to the observed value. Significant correlation was inferred if the $p$ value was smaller than 0.01. The confidence intervals for each correlation coefficient were also calculated (Manly 1997). If two distances are correlated, one can assume that they are sensitive to the same parameters of the calls.

To provide a visual representation of the correlation between these distance measures, we performed classical (metric) multidimensional scaling (principal coordinates analysis). We used one minus the correlation coefficient as a measure of similarity between the different methods. This representation of the different acoustic distance measures was used to visualize the distance matrix between methods in two dimensions. If two distances are strongly correlated, they will be assigned coordinates in the same region. The maximum error between the distances in the two-dimensional space and the original distances provides an indicator of the quality of the reconstruction. If this error is small in regards to the magnitude of the original distance then one can conclude that the reconstruction is accurate. These analyses were performed with Matlab (The Mathworks and Natick 2008).

Finally, for each method and for each cuckoo species we computed the rank order of each of the other species, according to the five species pairwise distance matrices (25×25). This allowed us to determine the species with the most similar begging call to each cuckoo species. If a cuckoo's call matches the host's begging call then it is expected that the host species will be ranked first. On the other hand, if such a process has not occurred then the begging call similarities should reflect the phylogenetic relationships between the cuckoo species. In this case, the cuckoo species are expected to occupy the first ranks. Typically, the species in question should be ranked first, indicating that its own calls are most similar to each other. If not, it may indicate that there is a larger heterogeneity within our recordings of the begging calls of that single species.

## Results

Three pairs of distance measures showed high correlations: spectrogram alignment with SAP (0.66), spectrogram alignment with sound features (0.65) and SAP with sound

features (0.68). Mantel tests confirmed that most pairs of distances were more correlated than expected by chance (Table 3). However, in four cases the $p$ values were greater than 0.01. Three of them involved correlations between the spectro-temporal modulation distance with SAP, spectrogram correlation, and sound features distances, respectively. The fourth case was between SAP and spectrogram correlation distances. The spectrogram alignment distance measure was correlated significantly with each of the other acoustic distance metrics. Moreover, the spectrogram alignment was the only method to return a distance significantly correlated with the spectro-temporal modulation distance. The spectrogram correlation method did not correlate with estimates obtained using the other distance methods, being only correlated with the spectrogram alignment and the sound features approaches.

Multidimensional scaling provided an acceptable reconstruction in two dimensions with a maximum error of 0.06 between the reconstructed and the original distances. The maximum value of the original distance was 0.85 (error ~0.076%). The results of these correlation analyses among the bioacoustic methods were visualized on the principal coordinate analysis plot (Fig. 2), where the spectrogram alignment distance stands in the middle and the spectrogram correlation and the spectro-temporal modulation distances appear isolated on the edges of the space.

When we used begging calls to assess evolutionary distances, four out of the five distance measures placed the shining cuckoo closest to its grey warbler host (Online Resource 1). This confirms the conclusions of the previous study based on Euclidian sound feature distances (Anderson et al. 2009b). Only the spectrogram correlation distance did not rank the begging calls of the grey warbler as the nearest match to that of the shining cuckoo. Moreover, none of the five methods indicated that the call of the shining cuckoo exhibits close relationships with the calls of the other cuckoo species to which it is otherwise phylogenetically related (congeneric, family Cuculidae).

In contrast to the similarity between grey warbler and shining cuckoos in New Zealand, the begging calls of the western thornbill were not ranked as the closest to its brood parasite, the Australasian subspecies of the shining cuckoo, by four out of the five distance measures (Online Resource 2). Only the spectrogram alignment distance suggested a strong similarity between these calls by ranking the host first after the parasite. The SAP distance and the spectrogram correlation distances did not classify the shining cuckoo as the closest species to itself when it parasitizes the western thornbill. Nevertheless, the spectrogram correlation distance ranked the western thornbill begging calls first which suggests similarity between the shining cuckoo call and its host's call. All distances, apart from the spectrogram correlation, resulted in ranking a cuckoo species as the closest match.

The begging calls of the Horsfield's bronze-cuckoo, another parasite of the western thornbill, also seemed to show some similarity to its host's calls (Online Resource 3). The sound features distance ranked the host as the closest call to the parasite. The spectrogram alignment distance ranked it as the second closest, but in this case the most similar call is that of the shining cuckoo, another brood parasite of the western thornbill. The spectrogram correlation and the SAP methods did not rank the Horsfield's bronze-cuckoo's host as the most similar but still place the western thornbill at the top of the rank order.

The calls of the long-tailed cuckoo likewise show some similarity with the calls of its host species, with four methods ranking them in the first half of the table (Online Resource 4). The spectro-temporal modulation groups the cuckoo and the three host species together, suggesting they have closely related begging calls. The SAP and the sound features distances also seem to detect congruence between these calls with the exception of the whitehead, which was slightly more distant and the pipit which was inserted at the top of the table. The spectrogram alignment technique placed the other cuckoo species (Horsfield's bronze-cuckoo and both shining

**Table 3** Correlation coefficients for each pair of distance measures. The $p$ values from the Mantel test (10,000 replicates) are indicated in brackets and 99% confidence intervals in square brackets

| | Spectrogram correlation | Spectro-temporal modulations | Sound features | Spectrogram alignment |
|---|---|---|---|---|
| SAP | 0.15 [−0.15,0.40] ($p$=0.0819) | 0.23 [−0.07,0.54] ($p$=0.0231) | 0.68 [0.49,0.91] ($p$=0.0001) | 0.66 [0.41,0.83] ($p$=0.0001) |
| Spectrogram correlation | | 0.22 [−0.09,0.50] ($p$=0.0273) | 0.39 [0.11,0.66] ($p$=0.0002) | 0.40 [0.12,0.70] ($p$=0.0003) |
| Spectro-temporal modulation | | | 0.18 [−0.15,0.45] ($p$=0.0743) | 0.47 [0.15,0.74] ($p$=0.0002) |
| Sound features | | | | 0.65 [0.41,0.88] ($p$=0.0001) |

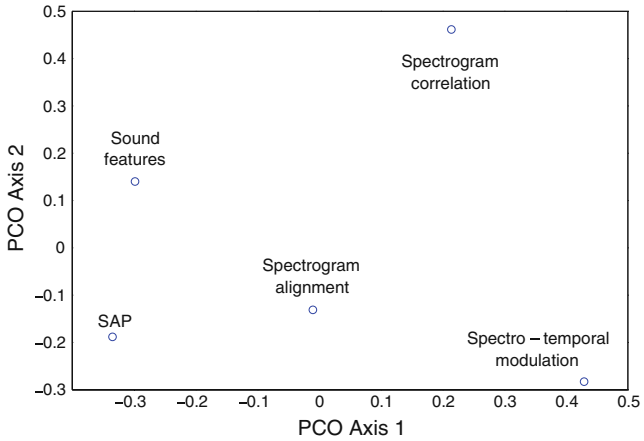


**Fig. 2** Principal coordinates analysis showing the relationships among the different acoustic distance measures. The similarity between these measures is defined as one minus the correlation coefficients of the corresponding distance matrices

cuckoo subspecies) closer to the long-tailed cuckoo than to its host species, revealing some phylogenetic signal present in the calls. By contrast, the long-tailed cuckoo and its hosts were not grouped together by the spectrogram correlation distance nor were the other cuckoo species.

In the rank order of similarities to the begging calls of the New Zealand subspecies of the shining cuckoo (Online Resource 1), the SAP, sound features and the spectrogram alignment methods gave similar ranking to the three other representatives of the genus *Chalcites*. The same observation can be made with the sound features and the spectrogram alignment methods in the tables of Online Resources 2 and 3 and the spectrogram alignment method alone in the table of the Online Resource 4. In these cases, the New Zealand subspecies of the shining cuckoo appears to be isolated from the other subspecies of shining cuckoo and the Horsfield's bronze-cuckoo of Western Australia.

## Discussion

This study revealed heterogeneity across the different bioacoustic distance measures applied to the same set of begging calls in Australasian hosts of brood parasitic cuckoos. Therefore, the same acoustic similarity indices and methods should be used within a study, and their choice should depend on the type of sensory and perceptual information available about the study system. Nevertheless, the comparison of these methods allowed us to find evidence of similarities that could be explained by either phylogenetic relatedness or by begging call mimicry of the hosts of brood parasite nestlings.

### Bioacoustic techniques

The five types of acoustic distance metrics used in this study to measure begging call similarities of diverse species showed both high and low correlations with each other, depending on particular methods being compared. This could be caused by the unique ways these methods quantify and analyze naturally complex sounds. The sound features and the spectro-temporal modulation distances are independent of the phase of the sounds because they both extract parameters from the whole call recording without considering the time point at which they occur. These methods do not show particularly high correlation and appear separated on the principal coordinate analysis plot. This may be because one of the acoustic properties that is included in the sound features measure is pitch, while the spectro-temporal modulation measure does not include carrier frequencies. The spectrogram correlation, the spectrogram alignment, and the spectro-temporal modulation distances all use information, directly or indirectly, extracted from the spectrograms for representing the calls. The SAP and the sound features distances, on the other hand, quantify the similarities between parameters obtained from the raw signal or the spectrum. However, the distance measures did not group according to these characteristics either. Additionally, the different ways with which the species distances were averaged could have had an effect. For example, the shining cuckoo when parasitizing the western thornbill was not ranked first. One can suppose that a large heterogeneity in its begging calls would have had an influence on the average calculation performed to obtain a species pairwise distance matrix. Nevertheless, some methods still ranked this cuckoo species first showing that they were less sensitive to this heterogeneity.

The spectrogram alignment distance measure showed medium to high correlations with all the other distance measures. The diverse distances were not always strongly correlated but the results demonstrate the potential for a metric in agreement with all of them. Overall, our work shows that these distance metrics are not incompatible, and that they might each be sensitive to different aspects of the sounds.

Specifically, our results indicate that the mel-frequency cepstrum coefficients, even if not originally designed for animal vocalizations, provide a generally applicable way of encoding the begging calls that preserves a significant amount of the information included in the sounds. Moreover, dynamic time warping appears to be an efficient algorithm for, first, averaging a set of calls and, second, measuring the similarities between encoded calls. This method was the only distance measure correlated with the spectro-temporal modulation of the sound, which has been shown to relate with neural coding of vocalizations in the

avian auditory system (Woolley et al. 2005) using neurophysiological data (Lewicki 2002; Singh and Theunissen 2003; Hsu et al. 2004). Neuronal function-informed measures of sound distance metrics, thus, can help with selecting the most appropriate method for a specific species under study. However, these experiments are not always feasible, especially for rare species of conservation concern, including New Zealand's endemic hosts and their native cuckoo parasites. The spectrogram alignment method used here is therefore a good integrative method for measuring sound similarities when no or only slight prior neuronal encoding knowledge is available regarding the relevant and salient features of a set of animal sounds under study.

The choice of a particular method should thus depend on the type of data, the existing knowledge available to the investigators, and the accuracy required for a particular analysis. If only general aspects of the sound need to be analyzed, for example, for extracting bioacoustic signals from a constant background noise, the spectrogram correlation and the sound features methods may be appropriate. However, when a greater sensitivity is required to detect fine sound variations, for example, when comparing intraspecific call types, the SAP or the spectrogram alignment methods are more suitable. Moreover, if there is a need for summarizing a set of calls from, for example, the same individual or species, the spectrogram alignment and the spectro-temporal modulation methods could be used as they provide specific tools for such purpose. When previous knowledge is available about the particular aspects of the sound that are relevant for a specific study, the sound feature method can be used provided that it is set up to focus on the corresponding parameters. Furthermore, in the case where the recordings under study are poor with the presence of noise, preprocessing may be useful for filtering a known constant background noise. Overall, it appears that the spectrogram alignment method is a generally applicable and statistically consistent approach to choose when no or only little prior knowledge is available as it is the only distance correlated with all the others.

## Begging call similarities

From a phylogenetic signal perspective, the begging calls of the cuckoos showed some similarities as expected from their shared ancestry. The spectrogram alignment method consistently grouped the Australian cuckoos together (Online Resources 2 and 3). Moreover, the long-tailed cuckoo's calls showed remarkable similarities with these species' calls (Online Resource 4). This could be the consequence of some phylogenetic signal still being present in the begging calls of the shining cuckoo and the Horsfield's bronze-cuckoo. Indeed the distance methods effectively grouped these cuckoo calls together, suggesting that certain sound parameters still retain information about the phylogenetic history amongst cuckoos. Nevertheless, the calls of the New Zealand subspecies of the shining cuckoo parasitizing the grey warbler did not group with the other cuckoo species, showing that this cuckoo lineage has a different begging call, more distant from the calls of the other lineages. This result provides further support to suggest that the begging call of the shining cuckoo has been altered from its ancestral phylogenetic position (Anderson et al. 2009b). Our observations would benefit from further analysis of the phylogenetic signal of brood parasite begging calls and their hosts, on a greater number of host–parasite systems.

These results in some cases also provide evidence that the begging calls of the three cuckoo species have deviated from their phylogenetic affinities (with other cuckoos of the family Cuculidae) to more closely resemble their hosts. Remarkably, the long-tailed cuckoo simultaneously shows phylogenetic signal and evidence for host begging call matching. The spectro-temporal modulation method clearly groups this cuckoo with its host species but the spectrogram alignment method groups it with the Australian cuckoos (Online Resource 4) illustrating that these methods are sensitive to different parameters. The shining cuckoo appears to have modified its begging call to match that of its host species in New Zealand (see above). However, quantitative similarity is less supported in the sample of Australian cuckoos studied here, for whom the similarities between the calls of the hosts and the parasites do not appear as strong: the other Australian host–parasite pairs also showed similarities in their calls but not as consistently across the different methods of bioacoustic distances as is shown for the New Zealand shining cuckoos. This lack of similarity may be due to incomplete sampling of Australian host species (e.g. Malurids), as higher parasitism rates of unsampled hosts may result in a greater extent of adaptation, exhibited through higher detected acoustic similarity. Alternatively, it may be that in these cases, the host species have also modified their calls in response to the mimicry by the brood parasites. Accordingly, the host can facilitate the discrimination between its young and the parasitic nestlings through more discriminable or intraspecifically variable begging calls, followed by a rejection response to the parasite, and so a coevolutionary process would occur with both species altering their begging calls away from their phylogenetic position. However, Anderson et al. (2009b) used randomization techniques to show that the begging calls of the grey warbler do not significantly differ from their phylogenetic position, indicating that this host has not changed its call as a result of the presence of a brood parasite. The finding reported here that some phylogenetic signal is also present in the Australian cuckoo species suggests that such a coevolution process has not

happened to the same extent in these host–parasite pairs either. Therefore, it appears that the shining cuckoo changed its begging call in the New Zealand environment in the context of parasitizing the grey warbler, but that such evolution is not as complete in Australia. This could be explained by the pattern that cuckoo species in Australia parasitize multiple host species and specialization on a single host species is more difficult. Another explanation would be that some cognitive or physiological limitations prevent these parasites from modifying their begging calls closer to their host's (Langmore et al. 2008). Alternatively, similarities of the begging calls of cuckoos from Western Australia may be due to effective mimicry of the more similar calls of their local and, presumably, phylogenetically more closely related hosts. Of course, other alternative explanations for bioacoustic similarity, such as selection for less detectable calls by nest predators or random matching, must also be considered (Grim 2005). Further investigation, including a full phylogenetic analysis combined with the use of a complimentary set of different bioacoustic discrimination indices is required to evaluate these alternative hypotheses.

**Acknowledgements** We thank Phillip Cassey, Brian Gill, Tomas Grim, Rebecca Kilner, Todd Landers, Naomi Langmore, Arnon Lotem, Luis Ortiz Catedral, Kevin Parker, and Allen Rodrigo for discussions and comments. Fieldwork was conducted with permission from the Auckland Regional Council, Department of Conservation, and the Massey University Animal Ethics Committee. We are grateful for field assistance by many helpers, too numerous to mention here (see http://www.massey.ac.nz/dhbrunto/ppl/MichaelAnderson.htm). Funding was provided from a Bright Futures Top Achiever Scholarship, Massey University (M.G. A.), the National Geographic Society and the Human Frontier Science Program (to M.E.H.), and two New Zealand Marsden Fund Grants (M. E.H.; D.H.B. and H.A.R.).

## References

Anderson MG, Hauber ME (2007) A recognition-free mechanism for reliable rejection of brood parasites. Trends Ecol Evol 22(6):283–286

Anderson MG, Moskát C, Bán M, Grim T, Cassey P, Hauber ME (2009a) Egg eviction imposes a recoverable cost of virulence in chicks of a brood parasite. PLoS One 4(11):e7725

Anderson MG, Ross HA, Brunton DH, Hauber ME (2009b) Begging call matching between a specialist brood parasite and its host: a comparative approach to detect coevolution. Biol J Linn Soc 98:208–216

Baker MC, Logue DM (2003) Population differentiation in a complex bird sound: a comparison of three bioacoustical analysis procedures. Ethology 109:223–242

Briskie JV, Martin PR, Martin TE (1999) Nest predation and the evolution of nestling begging calls. Proc R Soc Lond B 266:2153–2159

Brooker MG, Brooker LC (1989) Cuckoo hosts in Australia. Aust Zool Rev 2:1–67

Brown JC, Smaragdis P (2009) Hidden Markov and Gaussian mixture models for automatic call classification. J Acoust Soc Am 125 (6):EL221–EL224

Butchart SHM, Kilner RM, Fuisz T, Davies NB (2003) Differences in the nestling begging calls of hosts and host-races of the common cuckoo, *Cuculus canorus*. Anim Behav 65(2):345–354

Cardoso GC, Gama P, Depraz V (2007) Female and male serins (*Serinus serinus*) respond differently to derived song traits. Behav Ecol Sociobiol 61(9):1425–1436

Charif RA, Clark CW, Fristrup KM (2007) Raven Pro 1.3 User's manual. Cornell Laboratory of Ornithology, Ithaca, NY. edn

Christidis L, Boles WE (2008) Systematics and taxonomy of Australian birds. Collingwood, Vic. : CSIRO Pub.

Clemins PJ, Johnson MT (2006) Generalized perceptual linear prediction features for animal vocalization analysis. J Acoust Soc Am 120(1):527–534

Davies NB (2000) Cuckoos, cowbirds and other cheats. T & A D Poyser, London

Davies NB, Brooke De LM, Kacelnik A (1996) Recognition errors and probability of parasitism determine whether reed warblers should accept or reject mimetic eggs. Proc R Soc Lond B Biol Sci 263(1372):925–931

Davis SB, Mermelstein P (1980) Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE transactions on acoustics, speech, and signal processing 28(4):357–365

Gill BJ (1983) Brood parasitism by the shining cuckoo *Chrysococcyx lucidus* at Kaikoura, New Zealand. Ibis 125:40–55

Gill P, Zhang J, Woolley SMN, Fremouw T, Theunissen FE (2006) Sound representation methods for spectro-temporal receptive field estimation. J Comput Neurosci 21(1):5–20

Grim T (2005) Mimicry vs. similarity: which resemblances between brood parasites and their hosts are mimetic and which are not? Biol J Linn Soc 84:69–78

Grim T (2006) The evolution of nestling discrimination by hosts of parasitic birds: why is rejection so rare? Evol Ecol Res 8:785–802

Grim T, Kleven O, Mikulica O (2003) Nestling discrimination without recognition: a possible defence mechanism for hosts towards cuckoo parasitism? Proc R Soc Lond B 270(suppl 1):S73–S75

Guerra JE, Cruz-Nieto J, Ortiz-Maciel SG, Wright TF (2008) Limited geographic variation in the vocalizations of the endangered thick-billed parrot: implications for conservation strategies. The Condor 110(4):639–647

Hauber ME (2003) Lower begging responsiveness of host versus parasitic brown-headed cowbird (*Molothrus ater*) nestlings is related to species identity but not to early social experience. J Comp Psychol 117(1):24–30

Hauber ME, Kilner RM (2007) Who mimics whom? communication, co-evolution, and chick mimicry in parasitic finches. Behav Ecol Sociobiol 61:497–503

Hauber ME, Cassey P, Woolley SMN, Theunissen FE (2007) Neurophysiological response selectivity for conspecific songs over synthetic sounds in the auditory forebrain of non-singing female songbirds. J Comp Physiol A Sens Neural Behav Physiol 193:765–774

Honza M, Vošlajerová K, Moskát C (2007) Eviction behaviour of the common cuckoo *Cuculus canorus* chicks. J Avian Biol 38 (3):385–389

Hsu A, Woolley SMN, Fremouw TE, Theunissen FE (2004) Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. J Neurosci 24 (41):9201–9211

Kogan JA, Margoliash D (1998) Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden markov models: a comparative study. J Acoust Soc Am 103(4):2185–2196

Langmore NE, Kilner RM (2007) Breeding site and host selection by Horsfield's bronze-cuckoos, *Chalcites basalis*. Anim Behav 74:995–1004

Langmore NE, Hunt S, Kilner RM (2003) Escalation of a coevolutionary arms race through host rejection of brood parasitic young. Nature 422(6928):157–160

Langmore NE, Maurer G, Adcock GJ, Kilner RM (2008) Socially acquired host-specific mimicry and the evolution of host races in Horsfield's bronze-cuckoo *Chalcites basalis*. Evolution 62 (7):1689–1699

Langmore NE, Cockburn A, Russell AF, Kilner RM (2009) Flexible cuckoo chick-rejection rules in the superb fairy-wren. Behav Ecol 20(5):978–984

Latham J (1802) Supplementum indicis ornithologici, sive systematis ornithologiae. G. Leigh, J. & S, Sotheby, London

Lesica NA, Grothe B (2008) Efficient temporal processing of naturalistic sounds. PLoS ONE 3(2):e1655

Lewicki MS (2002) Efficient coding of natural sounds. Nat Neurosci 5:356–363

Manly BFJ (1997) Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, London

Margoliash D (1983) Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. J Neurosci 3:1039–1057

Matheron Nocolas, Aubin T, Vielliard J, da Silva M-L, Sebe F, Boscolo D (2008) Singing in the rain forest: how a tropical bird song transfers information. PLoS ONE 3(2):e1580

McLean IG (1988) Breeding behaviour of the long-tailed cuckoo on little barrier island. Notornis 35:89–98

McLean IG, Waas JR (1987) Do cuckoo chicks mimic the begging calls of their hosts? Anim Behav 35(6):1896–1898

Moskát C, Hauber ME (2007) Conflict between egg recognition and egg rejection decisions in common cuckoo (*Cuculus canorus*) hosts. Anim Cogn 10(4):1435–9448

Naugler C, Ratcliffe L (1992) A field test of the sound environment hypothesis of conspecific song recognition in American tree sparrows (*Spizella arborea*). Behaviour 123(3–4):314–324

Payne RB (2005) The Cuckoos. Oxford University Press, Oxford

Payne RB, Payne LL (1998) Nestling eviction and vocal begging behaviors in the Australian glossy cuckoos *Chrysococcyx basalis* and *C. lucidus*. In: Rothstein SI, Robinson S (eds) Parasitic birds and their hosts: Studies in coevolution. Oxford University Press, New York, pp 152–170

Ranjard L, Ross HA (2008) Unsupervised bird song syllable classification using evolving neural networks. J Acoust Soc Am 123(6):4358–4368

Sewall KB (2009) Limited adult vocal learning maintains call dialects but permits pair-distinctive calls in red crossbills. Anim Behav 77 (5):1303–1311

Singh NC, Theunissen FE (2003) Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am 114(6 Pt 1):3394–3411

Sokal RR, Rohlf FJ (1995) Biometry. W.H. Freeman and Co., New York

Soler M (2009) Co-evolutionary arms race between brood parasites and their hosts at the nestling stage. J Avian Biol 40(3):237–240

Tao J, Johnson MT, Osiejuk TS (2008) Acoustic model adaptation for ortolan bunting (*Emberiza hortulana* L.) song-type classification. J Acoust Soc Am 123(3):1582–1590

Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP (2000) A procedure for an automated measurement of song similarity. Anim Behav 59(5):1167–1176

The Mathworks, Natick, MA (2008) Matlab 7.7.0.471 (R2008b). URL http://www.mathworks.com

Trawicki MB, Johnson MT, Osiejuk TS (2005) Automatic song-type classification and speaker identification of Norwegian ortolan bunting (*Emberiza hortulana*) vocalizations. In: 2005 IEEE Workshop on Machine Learning for Signal Processing, pp 277–282, URL http://ieeexplore.ieee.org/servlet/opac? punumber=10270

Vallejo EE, Cody ML, Taylor CE (2007) Unsupervised acoustic classification of bird species using hierarchical self-organizing maps. In: Randall M, Abbass HA, Wiles J (eds) Progress in artificial life, third Australian conference, ACAL 2007 gold coast, Australia, December 4–6, Springer Berlin/Heidelberg, lecture notes in computer science, vol 4828/2007, pp 212–221

Vyas A, Harding C, Borg L, Bogdan D (2009) Acoustic characteristics, early experience, and endocrine status interact to modulate female zebra finches' behavioral responses to songs. Horm Behav 55:50–59

Woolley SMN, Rubel EW (2002) Vocal memory and learning in adult bengalese finches with regenerated hair cells. J Neurosci 22(17):7774–7787

Woolley SMN, Fremouw TE, Hsu A, Theunissen FE (2005) Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. Nat Neurosci 8 (10):1371–1379

Young SJ (1994) The HTK hidden markov model toolkit: design and philosophy. entropic cambridge research laboratory. Ltd 2:2–44