On the use of cellular telephony for audio interaction with animals

Dale Joachim* and Eben Goodale†

The Media Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA
*Author for correspondence (joachimd@mit.edu).
†Present address: Department of Zoology, University of Colombo, Colombo 3, Sri Lanka.

Playback is an important method of surveying animals, assessing habitats and studying animal communication. However, conventional playback methods require on-site observers and therefore become labour-intensive when covering large areas. Such limitations could be circumvented by the use of cellular telephony, a ubiquitous technology with increasing biological applications. In addressing concerns about the low audio quality of cellular telephones, this paper presents experimental data to show that owls of two species (Strix varia and Megascops asio) respond similarly to calls played through cellular telephones as to calls played through conventional playback technology. In addition, the telephone audio recordings are of sufficient quality to detect most of the two owl species’ responses. These findings are a first important step towards large-scale applications where networks of cellular phones conduct real-time monitoring tasks.

Keywords: animal surveys; cellular phones; playback; remote monitoring; Voice over Internet Protocol applications; wireless communication

1. INTRODUCTION

Animal vocalizations are useful indicators for assessing biodiversity (Riede 1993) or habitat quality (Slabbekoorn & Peet 2003; Laiolo & Tella 2005). Playback (a controlled mixture of broadcasting and recording) augments the effectiveness of passive recording, particularly when targeting specific species in animal surveys (Johnson et al. 1981; Ogutu & Dublin 1999), and studying animal communication (Falls 1992; Dabelsteen & McGregor 1996). However, while recordings of animal vocalizations are increasingly performed through remote devices (Hobson et al. 2002; Charif et al. 2005) and wireless networks (Wang et al. 1999; Porter et al. 2005), playback experiments are almost always performed with human observers present (Mennill & Ratcliffe 2000; Stoleson et al. 2004). This restriction makes conventional playback experiments a labour-intensive process not well suited to large-scale experiments.

We consider cellular telephony as an emerging technology for conducting remote playback experiments. Indeed, cellular telephony (see review by Rappaport et al. 2002) offers the following benefits over conventional playback methods: (i) less human disturbance, (ii) convenient remote access at all times, and (iii) ability to form large networks of phones that can simultaneously monitor many animals. Cellular phones are increasingly leveraged for tracking animals (McConnell et al. 2004; Sundell et al. 2006), but not for conducting animal vocalizations. The assumed drawback is that cellular phone signal-to-noise quality is insufficient to carry out meaningful experiments. However, we could not find experimental data in the current literature to either refute or support the hypothesis that animals respond to cellular phone grade audio, or that their responses can be detected through cellular phones. We therefore investigated the effectiveness of cellular telephony in owl surveys, a common application of playback (Takats et al. 2001).

2. MATERIAL AND METHODS

We designed cellular phone stations capable of playback, recording or both (figure 1). Playback was achieved through a Nokia N80 GSM cellular telephone, a 9 V amplifier and a 30 W outdoor horn speaker. Similarly, an amplified electret microphone served as the cellular telephone input. A web-based scheduling interface, accessible by cellular phone, allowed the user to select playback and/or recording, the phone with which to interact, the sounds to broadcast and the duration of recording.

We experimented with two owl species, the barred owl (Strix varia) and the Eastern screech owl (Megascops asio), following the published playback protocols (McGalan & Fraser 1985; Smith et al. 1987). We produced nine playback soundtracks per species, making each soundtrack from a separate recording deposited at the Macaulay Library (Ithaca, NY), and converted them to 8 kHz GSM audio compression format. The soundtracks consisted of 20 s recordings of owls repeated for 6 min with 20 s intervals of silence and represented a wide sample of the vocalizations of the species (see table 1). Trials were conducted with an observer present in order to compare the recorded audio to what was detected in the field. We conducted playback at nine locations per species, with the locations chosen to have appropriate habitat and to be at least 0.8 km apart. At each location, we executed trials on consecutive nights, once using the cellular phone to broadcast playback and once using a CD player, with the order of presentation systematically varied; responses were recorded through the cell phone for all trials.

In order to assess the fidelity of audio capture, an independent observer, who did not have previous experience listening to owl recordings other than the playback tapes and who did not have knowledge of the outcome of the experiments, listened to the recordings, inspected spectrograms, and then scored whether an owl had responded or not.

3. RESULTS

Owls responded in a similar manner to the calls broadcast over the cellular phones as compared to the CD player (figure 2). The percentage of trials in which there was a response was similar to cell phones (11 out of 18 trials) as to the CD player (10 out of 18 trials; Fisher’s exact test on response to all owls, $p=0.99$). For the nine locations where owls responded to both the cell phone and the CD player, owls did not differ in their delay of response (repeated sample $T$-test, $T=1.22$, d.f. = 8, $p=0.26$), or their closest approach to the speaker ($T=0.17$, d.f. = 8, $p=0.87$), or the number of calls that they made ($T=1.00$, d.f. = 8, $p=1.00$).

Eighteen out of 21 owl responses (8 out of 9 barred owl calls and 10 out of 12 screech owl calls) were detected by the independent observer from the stored recordings. Those responses that were not detected were faint: we estimated that in the three such trials owls were on average 167 m from the speaker, whereas they were 42 m from the speaker in responses that were detected (independent sample...
At the same time, there were three trials in which the independent observer detected owls that were not present. One of these ‘false positives’ involved detecting an unknown owl-like noise also heard by the field observer; in the other two trials, the independent observer confused the noise of a car with the trill call of a screech owl—a call without frequency modulation and with a similar frequency as the car.

4. DISCUSSION

Our data demonstrate that cellular telephony is a viable method of remote playback for certain applications, like playback to owls. Owls responded to cellular phone grade audio as they do in conventional playback, and their responses were recorded remotely. One problem encountered was that some faint responses were not detected. Although the percentage of responses detected could be increased by greater amplification and improved omnidirectional microphone design (Hobson et al. 2002), very faint responses are still likely to be undetected, as they would in any recording, with the additional constraints of the harsh compression due to the cellular phone encoding method and the relatively high cutoff threshold for silence inside the phone.

For studies or censuses of species other than the owls we experimented with, cellular telephony poses an additional hurdle: the narrow telephony frequency bandwidth may induce too much distortion at higher frequency broadcasts and therefore curtail potential responses. In addition, the user should be aware that audio frequency bands are unevenly represented in cellular compression and optimized for human perception (Painter & Spanias 2000; Kondoz 2004). One may circumvent the audio transmission and compression degradations by recording audio as data files on the cellular phone and then uploading the files. The trade-off of such a procedure is the loss of real-time interaction.

The advantages of cellular telephony become evident in studies in which an area is repeatedly sampled or which cover large geographical areas. Cellular phone systems are appropriate for studies of animal communication in which multiple treatments are conducted at a location at different times, or surveys that measure the effect of environmental variables on animals’ response. Indeed, the cellular phones can be outfitted with sensors that measure environmental variables.
A significant advantage of using a network of cellular phones is manifested in the ability to take ‘snapshots’ in time of environmental sounds throughout a large geographical area. Such networks provide the additional ability to localize or separate animals by position relative to the cellular phone stations, as microphone arrays

Figure 2. Owls responded similarly to the two playback methods. (a) The number of responses at 18 locations. Black columns represent S. varia and white columns M. asio. (b) The average delay in response at nine locations where owls responded to both methods. (c) The closest estimated distance approached by owls. (d) The number of calls owls made at the nine locations. Bars are standard errors.

Table 1. Playback tapes were made from recordings deposited at the Macaulay Library collection, except two exemplars, published by the Cornell Laboratory of Ornithology/Interactive Audio (1990), that were selections from several recordings of the collection. All barred owl recordings were composed of variations on its species-typical hoot.

<table>
<thead>
<tr>
<th>LNS no.</th>
<th>recorder</th>
<th>year</th>
<th>state</th>
<th>contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>barred owls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4546</td>
<td>Allen</td>
<td>1953</td>
<td>New York</td>
<td>hoot</td>
</tr>
<tr>
<td>4554</td>
<td>Reynard</td>
<td>1964</td>
<td>Georgia</td>
<td>hoot</td>
</tr>
<tr>
<td>31 523</td>
<td>Hewitt</td>
<td>1984</td>
<td>Florida</td>
<td>hoot</td>
</tr>
<tr>
<td>110 209</td>
<td>Hershberger</td>
<td>2001</td>
<td>West Virginia</td>
<td>hoot</td>
</tr>
<tr>
<td>125 363</td>
<td>Sander</td>
<td>1992</td>
<td>Oregon</td>
<td>hoot</td>
</tr>
<tr>
<td>125 371</td>
<td>Sander</td>
<td>1992</td>
<td>Oregon</td>
<td>hoot</td>
</tr>
<tr>
<td>128 926</td>
<td>Clock</td>
<td>2005</td>
<td>Arkansas</td>
<td>hoot</td>
</tr>
<tr>
<td>128 926</td>
<td>Clock</td>
<td>2005</td>
<td>Arkansas</td>
<td>hoot</td>
</tr>
<tr>
<td>Cornell Laboratory of Ornithology</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>hoot</td>
</tr>
<tr>
<td>screech owls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4456</td>
<td>Reynard</td>
<td>1958</td>
<td>New Jersey</td>
<td>whiny</td>
</tr>
<tr>
<td>20 424</td>
<td>McIssac</td>
<td>1979</td>
<td>New York</td>
<td>whiny</td>
</tr>
<tr>
<td>20 427</td>
<td>McIssac</td>
<td>1979</td>
<td>New York</td>
<td>trill</td>
</tr>
<tr>
<td>85 307</td>
<td>Hershberger</td>
<td>1997</td>
<td>Maryland</td>
<td>whiny</td>
</tr>
<tr>
<td>100 702</td>
<td>Hershberger</td>
<td>1998</td>
<td>Maryland</td>
<td>whiny</td>
</tr>
<tr>
<td>107 446</td>
<td>Hershberger</td>
<td>2000</td>
<td>Maryland</td>
<td>trill</td>
</tr>
<tr>
<td>Cornell Laboratory of Ornithology</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>whiny and trill</td>
</tr>
</tbody>
</table>

have been designed to do (Freitag & Tyack 1993; Mennill et al. 2006).

Cellular phone technology is increasingly ubiquitous and available in remote regions sometimes lacking standard telephones. Considering the wide—and increasing—use of cellular phones, we envision environmental sound capture through cellular telephony to become increasingly popular and allow community participation in scientific studies and educational projects.

The study followed ABS/ASAB Guidelines for the Treatment of Animals in Behavioral Research and Teaching.

We thank Jennifer Coulson, Tom Sherry and Fred Petry for their discussions and the EECs students who participated in the project while D.J. taught at Tulane University. We thank Jaewoo Chung, Clifton Dassuncao, Grant Elliott, Lalith Manage, Jongu Shin, Yang Yang, Maud Celestin and Greg Budney, Lalith Manage, Jongu Shin, Yang Yang, Maud Celestin and Greg Budney, thank Jaewoo Chung, Clifton Dassuncao, Grant Elliott, Lalith Manage, Jongu Shin, Yang Yang, Maud Celestin and Greg Budney, thank Jaewoo Chung, Clifton Dassuncao, Grant Elliott, Lalith Manage, Jongu Shin, Yang Yang, Maud Celestin and Greg Budney, thank Jaewoo Chung, Clifton Dassuncao, Grant Elliott, Lalith Manage, Jongu Shin, Yang Yang, Maud Celestin and Greg Budney, thank Jaewoo Chung, Clifton Dassuncao, Grant Elliott, Lalith Manage, Jongu Shin, Yang Yang, Maud Celestin and Greg Budney, for their discussions and the EECS students who participated in the project. This study was funded by the National Science Foundation grant IIS-063469.


