Page 41 / WISEST Special Issue / Volume 2 / DOI: 10.29173/aar48

Manual versus automatic identification of black-capped chickadee (*Poecile atricapillus*) vocalizations

Vala Ingolfsson[1], William D. Service[1], Carolina Montenegro[1], Christopher B. Sturdy[1]

[1]Department of Psychology, Neuroscience and Mental Health Institute, University of Alberta

Abstract

One time-consuming aspect of bioacoustic research is identifying vocalizations from long audio recordings. SongScope (version 4.1.5. Wildlife Acoustics, Inc.) is a computer program capable of developing acoustic recognizers that can identify wildlife vocalizations. The goal of the current study was to compare the effectiveness of manual identification of black-capped chickadee vocalizations to identification by SongScope recognizers. A recognizer was developed for each main chickadee vocalization by providing previously annotated audio of chickadees. Six chickadees (three male, three female) were recorded in one-hour intervals with and without anthropogenic (i.e., man-made) noise to provide a variety of samples to test the recognizer. These recordings were analyzed via the recognizer and two human coders, with an additional third coder reviewing a random subset of recordings for reliability. Strong agreement was found between the human coders, $\kappa = 0.76$, $p < 0.00$. Agreement between human coders and the recognizer was moderate for fee songs, $\kappa = 0.46$, $p < 0.00$, and strong for fee-bee songs, $\kappa = 0.77$, $p < 0.00$, as well as for chick-a-dee calls, $\kappa = 0.82$, $p < 0.00$. Results showed that male chickadees produced more tseet calls in silence and females produced more gargle calls during noise. No differences were found in vocalizations based on time of day. Our observations also suggest that the chick-a-dee recognizer was capable of identifying gargle and tseet calls along with the intended chick-a-dee calls. Overall, SongScope was effective at identifying fee-bee songs and chick-a-dee calls, but not as effective for identifying fee songs. These recognizers can allow for faster acoustic analyses (by approximately four times) and be continuously improved for greater accuracy.

Key words:
bird calls, SongScope, black-capped chickadee, birds, vocalizations

Cite as: Ingolfsson V., Service W.D., Montenegro C., and Sturdy C.B. 2019. Manual versus automatic identification of black-capped chickadee (Poecile atricapillus) vocalizations. Alberta Academic Review, Vol 2 (2) 41-42, WISEST Special Issue (not peer-reviewed), DOI 10.29173/aar48.

ISSN 2561-5327 (print)
ISSN 2561-5335 (online)

Manual versus automatic identification of black-capped chickadee (*Poecile atricapillus*) vocalizations

Vala Ingolfsson, William D. Service, Carolina Montenegro, & Christopher B. Sturdy
Department of Psychology, Neuroscience and Mental Health Institute, University of Alberta

Sponsored by: NSERC CRSNG PromoScience

Introduction

- Black-capped chickadees (BCCH) are ideal subjects for studying communication and vocal learning.
- Studies often involve "call cutting" (i.e., isolating and identifying vocalizations from hours of recordings).
- **SongScope** is a computer program used to create recognizers that identify specific animal vocalizations.

Figure 1. Black-capped chickadee.

- The current study tests how recognizers built in SongScope compare to manual call cutting.
- In addition we assessed how the time of day and noise impacts vocalizations produced.

(Version 4.1.5. Wildlife Acoustics, Inc.)

Methods

- A recognizer was generated in SongScope for each chickadee vocalization (i.e., *chick-a-dee* call, *chick-a* call, *tseet* call, *tseet* cluster, *gargle* call, *fee* song, *fee-bee* song), using pre-existing samples.

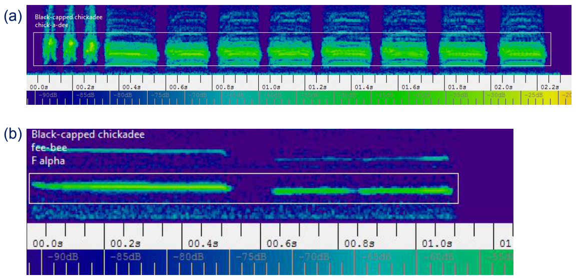

Figure 2. Annotations in SongScope. (a) *chick-a-dee* call; (b) *fee-bee* song.

- By annotating samples from individuals the program develops a model of what each vocalization should look like.
- By adjusting variables such as frequency range, maximum song length, and the length of syllables and the gaps between them, the recognizer can be made more accurate and specific.

Methods

- A total of six chickadees (3 female, 3 male) were recorded in 1 hr intervals, in the morning (08:30) and afternoon (14:30).
- Hours with silence and anthropogenic (i.e., manmade) noise were counterbalanced to provide a variety of samples to test the recognizers on.
- Two coders in addition to the recognizer reviewed the recordings in order to test interrater reliability between coder and SongScope.
- In addition, a third coder reviewed a random sample of recordings to test coder to coder reliability.

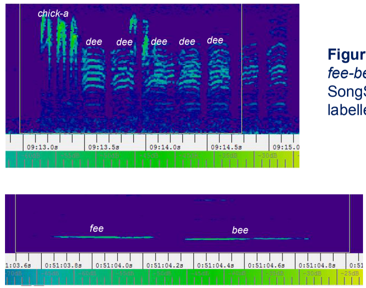

Figure 3. *Chick-a-dee* call and *fee-bee* song identified by a SongScope recognizer (we labelled the syllables).

Results

- Vocalization by sex

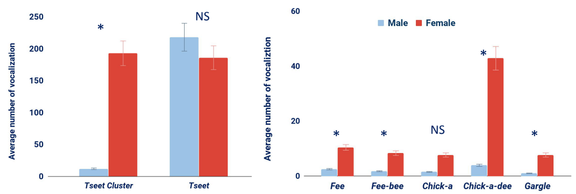

Figure 4. Differences of average vocalizations by sex. * indicate *ps* < 0.05, NS show no significance. Error bars represent 10% error amount for each data point.

- Males produced more *tseets* during Silence, $p < 0.01$, and females produced more *gargles* during Noise, $p < 0.00$.
- There were no significant differences in vocalizations by time of day.

Results

- Coder to coder interrater reliability (IRR)
  - A difference of two vocalizations was allowed for agreement
  - There was strong agreement between Coder 1 and Coder 3, $\kappa = 0.76$, $p < 0.00$, and moderate agreement between Coder 2 and Coder 3, $\kappa = 0.67$, $p < 0.00$, based on a random sampling of recordings.

	Coder 1	Coder 3		Coder 2	Coder 3
Chick-a-dee	122	116	*Chick-a-dee*	22	20
Fee	42	39	*Fee*	5	3
Fee-bee	64	64	*Fee-bee*	3	5

Table 1. Sample of raw vocalization scores between Coder 1 and 2 for one bird and between Coder 2 and 3 for a seperate bird.

- Coder to SongScope IRR
  - *Chick-a-dee* call. There was strong agreement by recording, $\kappa = 0.82$, $p < 0.00$.
  - *Fee-bee* song. There was strong agreement by recording, $\kappa = 0.77$, $p < 0.00$.
  - *Fee* song. There was moderate agreement by recording, $\kappa = 0.46$, $p < 0.00$.

	Fee		*Fee-bee*		*Chick-a-dee*	
	Coder	SongScope	Coder	SongScope	Coder	SongScope
Bird 1	23	2	11	6	0	2
Bird 2	5	0	3	2	40	42
Bird 3	67	33	88	81	250	249

Table 2. Sample of total vocalizations comparing Coder and SongScope coding.

	Silence		Noise		Silence		Noise	
	Fee-bee Coder	*Fee-bee* SongScope	*Fee-bee* Coder	*Fee-bee* SongScope	*Chick-a-dee* Coder	*Chick-a-dee* SongScope	*Chick-a-dee* Coder	*Chick-a-dee* SongScope
Bird 1	2	2	9	4	0	0	0	2
Bird 2	3	2	0	0	22	22	18	22
Bird 3	64	64	24	17	122	120	128	129

Table 3. Sample of total vocalizations by noise condition, type by Coder and SongScope coding.

Discussion

- Coder-coder IRR was found to be satisfactory, and coder-SongScope IRR was strong for *chick-a-dee* calls and *fee-bee* songs, but weak for *fee* songs.
- The *chick-a-dee* recognizer was able to identify *gargle* and *tseet* calls as well as *chick-a-dee* calls, possibly due to structural similarity.
  - Recognizers can be continuously improved for greater accuracy.
- Call cutting by SongScope was found to be much faster (48 hours versus approximately 12 hours) than human call cutting.

Acknowledgements

Special thanks to all the Neuroethology Laboratory members.