Ecological Informatics 27 (2015) 21–25

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Bird-DB: A database for annotated bird song sequences

Julio G. Arriaga [a,1], Martin L. Cody [b,2], Edgar E. Vallejo [a,3], Charles E. Taylor [b,*]

[a] Tecnológico de Monterrey, Campus Estado de México. Atizapán de Zaragoza 52926, Mexico
[b] University of California, Los Angeles, CA 90095-1606, USA

ARTICLE INFO

Article history:
Received 29 August 2014
Received in revised form 25 January 2015
Accepted 28 January 2015
Available online 3 February 2015

Keywords:
Bioacoustics
Bird song
Phrase sequence
Birdsong syntax
Database

ABSTRACT

Projects on the acoustic monitoring of animals in natural habitats generally face the problem of managing extensive amounts of data, both needed for – and produced by – observation or experimentation. While there are many publicly accessible databases for recordings themselves, we are aware of none for annotated song sequences. In this paper, we describe our database system of bird vocalizations and introduce our online sample repository for the community of researchers studying the syntax of bird song.

1. Introduction

Many species communicate with sequences of vocal sounds, analogous perhaps to words. Well-known examples include whales, monkeys, and songbirds, and there are many more described in the recent review by Kershenbaum (2014). The vocalizations are typically thought to be associated with sexual selection, but are often associated with different activities and may be *referential*, in that they allow companions to predict their environment, as predators or food availability (Evans (1997)). In some instances the ordering of such units is associated with different meanings (Arnold and Zuberbühler (2008)). Our understanding of song sequences is growing rapidly as new software and technologies become available, contributing to an emerging field of research — the biology of syntax (Margoliash and Nusbaum (2009)).

A recent workshop of researchers in the biology of syntax (Harvard University (2014)) identified an appropriate, shared database as the single most pressing need for this community. Several excellent online databases are available for general recordings of bird and animal sounds, including Xeno-canto (Planqué and Vellinga (2014)), Cornell's Macaulay Library (Cornell (2014)), Avibase (Lepage (2014)), and the Western Soundscape Archive (Utah (2014)); others that provide metadata about recordings, though the files themselves must be obtained separately from their source, include The British Library (British Library (2014)), Berlin Museum für Naturkunde (Tierstimmenarchiv (2014)), the Borror Library (OSU (2014)), and the Australian National Wildlife Collection Sound Archive (CSIRO (2011)). While providing access to recordings and relevant metadata, these databases do not include annotations about the types or sequences of vocal units comprising them; so they are not suitable for analysis of song syntax. The annotation of phrase types is typically much more time-consuming than the recordings themselves, making a shared, publicly available resource for animal syntax much desired.

In this paper we introduce Bird-DB, a publicly accessible relational database system that contains audio files of bird songs and their annotations for the phrase types they comprise. The format is flexible, though it is intended that the typical entries will be: (a) audio files, e.g., .wav, obtained from nature; (b) annotation files, e.g., Praat .TextGrid files, that indicate the phrase type at particular times; (c) and associated with these are mappings from spectrogram patterns to phrase types, together with metadata files containing information about ecology and location details. A typical pairing of the audio and annotation files is illustrated in Fig. 1.

Over the course of the last 5 years, we have collected over 1000 recordings featuring individuals from more than 30 different bird species pertaining to the California and Western Australia regions. We acquired metadata about each of these audio recordings, modeled after a simplified version of the Macaulay Library metadata records (Bradbury, personal communication). These comprise thousands of songs and tens of thousands of phrases, mostly identified to individual birds. We have found a database to be essential to keep track of these data and for their version control; we have made it accessible to all interested users over the internet.

In this paper, we describe our system for managing the recorded samples of bird vocalizations, their annotation and associated metadata.

* Corresponding author. Tel.: +1 310 825 6850.
E-mail addresses: a00967664@itesm.mx (J.G. Arriaga), mlcody@ucla.edu (M.L. Cody), eevallejo@gmail.com (E.E. Vallejo), taylor@biology.ucla.edu (C.E. Taylor).
[1] +52 1 5553931846.
[2] +1 (209) 296-2514.
[3] +52 1 5543588711.

http://dx.doi.org/10.1016/j.ecoinf.2015.01.007
1574-9541/

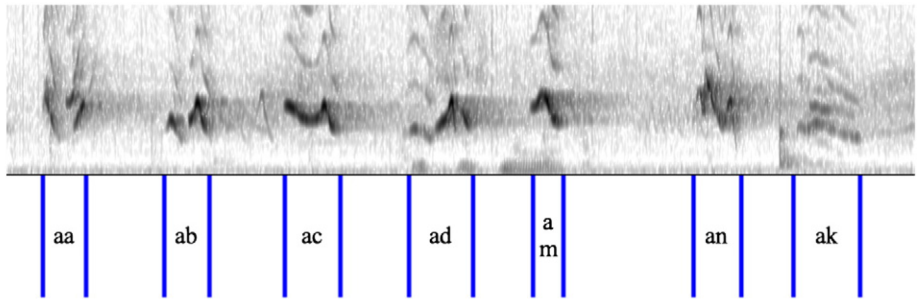

Fig. 1. Audio recording of black-headed grosbeak showing annotation with Praat.

While complicated searches still require commands in MySQL, we have written a PHP interface so that most searching can be done with a simple web-based interface.

We anticipate that the database will find extensive use by others who might want to upload their own data. We will be happy to host this as resources allow.

2. System design and implementation

The system consists of three main parts: a database; an interface for updating and managing the data contained in the database; and an interface for querying the database.

The database was designed around a central element designated as a *TRACK*. A *TRACK* consists of a recording event resulting in one or more audio files (the number of files depends on the hardware used for the recording). Metadata about the recording event is stored as basic attributes in the *TRACK* entity (like the date/time when it was performed and its length) or through relationships to other entities (like the *RESEARCHER* who performed the recording or the *RECORDING_HARDWARE* used). *TRACKS* can be associated with one or more *SUBJECTS*. A *SUBJECT* represents a specific individual and, in turn, can be associated with one or more *TRACKS*. The *LOCATION* entity stores information about the position and environment of the individual in the recording. Additional elements derived by analyzing a *TRACK* can be described through the *ANALYSIS* and related entities. At the time of this writing, our main post-processing focus has been phrase tagging, however the design permits other types of analysis to also be described. New procedures can be described using the *METHOD* entity and combined into one *ANALYSIS* describing the results and linking any relevant files to it. Fig. 2 shows a simplified Entity-Relationship diagram of our database. A more detailed diagram, as well as a data dictionary are included in the Supplementary material.

The database design was performed following standard normalization techniques, satisfying the 3NF conditions (Elmasri and Navathe (2013)). The database was implemented in MySQL residing in an Apache Linux server environment. Data migration was achieved through a combination of Python scripting and manual input.

Our online public interface for bird songs at http://taylor0.biology.ucla.edu/birdDBQuery/ offers limited search capabilities. Users are capable of selecting attributes, relating to several general aspects of our recordings, to be displayed. Queries can also be narrowed down to specific species and individuals. The database returns a list of records meeting those criteria, with links to the appropriate audio and annotation files. Further filtering of these results is possible by using the search box in the results page. Results can also be sorted incrementally or decrementally by any column. See, for example, Fig. 3.

By clicking on the links one can hear the audio file or see the annotation file. The files in .wav or .TextGrid format for Praat can be downloaded as source files; the implementation for such downloads depends on the browser one is using. At present, each file, whether audio or annotation, must be downloaded independently. More advanced capabilities are currently under development.

3. Composition of Bird-DB

Birds-DB currently has 428 files that have been annotated and several hundred more that have not been annotated yet. It currently requires a few terabytes of storage to include all the recordings. The number of annotated phrases is a few tens of thousands. These numbers are growing. The majority of annotations are for California thrashers

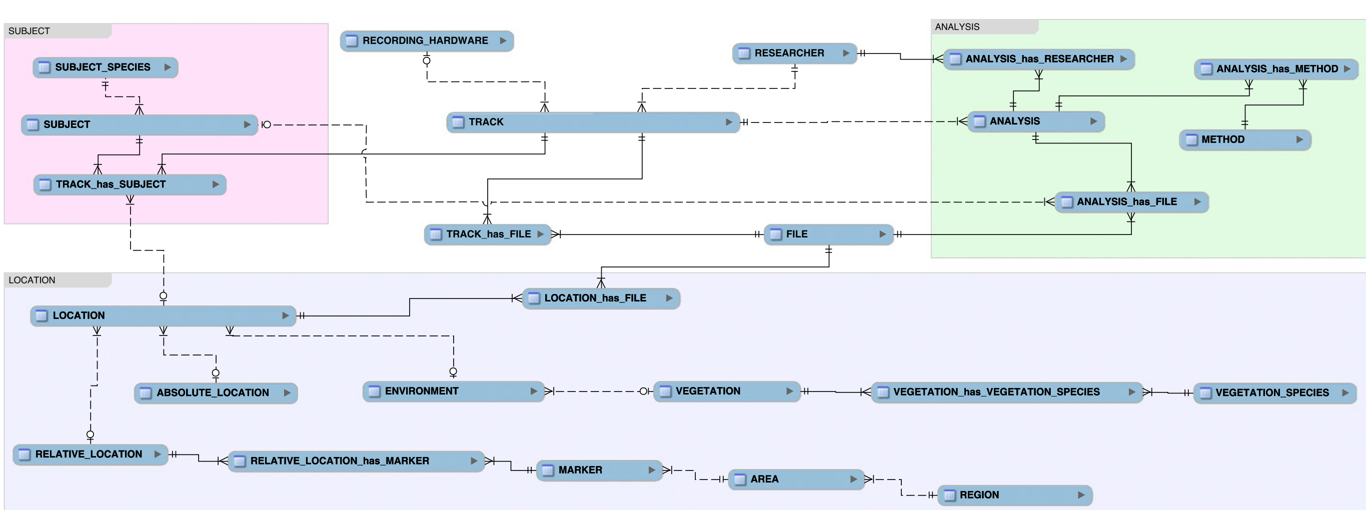

Fig. 2. Conceptual design of Birds-DB, displayed as a simplified Entity-Relationship diagram. Conceptually, the model has been divided into four layers: TRACK, SUBJECT, LOCATION and ANALYSIS.

Show 10 entries

Search:

TrackName	Audio_file	SubjectName	subject_importance	quality_rating	Species_short_name	Subject_species	sex	age_class	identified_by	certainty_of_species	Analysis	Analysis_date	keycode	number_of_phrases	Textgrid_file
tFTANN0512-10	Loading...	BHGBLSP12	Primary subject	4	BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-04-16	fkey_BHGB_2013	373	Files_TextGrids/2012/May/AmadMyJn12-10.TextGrid
tFTANN0512-11	Loading...	BHGBLSP12	Primary subject	4	BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-04-16	fkey_BHGB_2013	332	Files_TextGrids/2012/May/AmadMyJn12-11.TextGrid
tFTANN0512-2	Loading...	BHGBLSP12	Primary subject	5	BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-04-16	fkey_BHGB_2013	59	Files_TextGrids/2012/May/AmadMyJn12-2.TextGrid
tFTANN0512-20	Loading...	BHGBCP12	Primary subject	4	BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-04-16	fkey_BHGB_2013	71	Files_TextGrids/2012/May/AmadMyJn12-20.TextGrid
tFTANN0512-5	Loading...	BHGBUSP12	Primary subject	4	BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-04-16	fkey_BHGB_2013	142	Files_TextGrids/2012/May/AmadMyJn12-5.TextGrid
tFTANN0512-6	Loading...	BHGBUSP12	Primary subject	3	BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-04-16	fkey_BHGB_2013	143	Files_TextGrids/2012/May/AmadMyJn12-6.TextGrid
tFTANN0513-20	Loading...	BHGBW13	Primary subject		BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	full track processed	2013-07-25	fAmadorSpecies		Files_TextGrids/2013/May/1020CHECKED.TextGrid
tFTANN0513-20	Loading...	BHGBW13	Primary subject		BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-07-25	fkey_BHGB_2013	434	Files_TextGrids/2013/May/1020bCHECKED.TextGrid
tFTANN0513-20	Loading...	BHGBTNK13	Background species		BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	full track processed	2013-07-25	fAmadorSpecies		Files_TextGrids/2013/May/1020CHECKED.TextGrid
tFTANN0513-20	Loading...	BHGBTNK13	Background species		BHGB	Black-headed Grosbeak	Unkown sex	Unkown age class	sight and sound	5	BHGB processed	2013-07-25	fkey_BHGB_2013	434	Files_TextGrids/2013/May/1020bCHECKED.TextGrid

Showing 1 to 10 of 107 entries

Previous 1 2 3 4 5 … 11 Next

Fig. 3. Output from database query for recordings about black-headed grosbeaks.

(*Toxostoma redivivum*), Cassin's vireo (*Vireo cassinii*), and black-headed grosbeaks (*Pheucticus melanocephalus*), though a few other species are also present. These are listed in the pulldown menu for the Species field in the online interface.

For our research purposes, annotation is to phrase-type where phrase is defined as a small distinct burst of song less than 1 s in duration, identifiable by its stereotyped delivery, and separated temporally from the preceding and following phrases by a brief silence. Phrases in the species we have annotated most extensively are essentially monosyllabic, with a single extended burst of sound; but many are not, and we used two criteria to define a unit phrase: temporal continuity and consistent association. i) If bi- or tri-syllabic phrase syllables were essentially continuous in time or strictly contiguous, they were regarded as a single phrase. ii) If temporal continuity was ambiguous but the syllables were always found associated in identical sequence, they were interpreted to constitute a single phrase. Of course these definitions may vary from species to species. Our intention is to be consistent with the nomenclature described in Catchpole and Slater (1995). In some species there are longer silent intervals that separate *bouts*, while in others there are not. The database does not indicate such breaks, but the annotation files indicate times, so the user can add that information if they desire.

All the annotated samples in our database have been analyzed, segmented, and phrase-tagged in a consistent way by visual inspection of their spectrogram using the Praat software (Boersma and Weenink (2014)). Consequently, both audio and .TextGrid files (text-based Praat files containing phrase and time information) are available for every recording track.

The mapping from spectrogram to phrase type is arbitrary, though once assigned should be consistent. Our mappings used for each track is shown in the keycode field; it is intended to be unchanging, barring corrections, and is a Keynote (and pdf) file that is in the database, with a link in the keycode field of the output. Angle and Coskun (2014) have provided a detailed description of such a keycode for seasonal variation in a canary.

It is possible for individual users to add their own annotated recordings to the database. At this time our institutional security policies do not allow general access from all users, so the process must be indirect. To add information, first refer to the Data Dictionary file in the Supplementary material, which lists the data requirements for each variable. We have made a Google Docs file, BirdDB-Template, that can be used to fill in the necessary information; it has an example entry for guidance. When this is filled out, the file should be transmitted to the first or last author (JGA or CET). Additional files, such as keycode, .kmz files for location and information about vegetation and dominant plant species should also be provided. The audio and annotation files will need to be made available in Dropbox, Google Drive or some such repository so that can be added as well. In our lab these are typically .wav and .TextGrid, but we sometimes use .flac for audio and .txt for annotations — any open-sourced reasonable alternative is currently acceptable. (Please note that .raw files can become excessively large.) We will then add the material and update the database accordingly.

4. Examples of the use of Bird-DB

For illustrative purposes, the following studies demonstrate the utility of Bird-DB.

4.1. Automated identification of bird individuals

Individual identification of birds by acoustic means has, up to now, relied largely on acoustic properties of the sounds to be classified, for example on vectors of spectral and temporal features of the song. The utility of such measures depends greatly on the environment, including background noise, and on the species being studied, especially for recordings of birds from natural settings.

In preliminary studies we found that phrase composition in song sequences sometimes provides more effective bird individual discrimination and identification than the vectors of acoustic properties that are typically employed. For this research, we used recordings and their annotations for the phrase types provided by Bird-DB from known, banded, Cassin's vireo individuals over two years.

Specifically, a collection of training and validation sets were generated from Bird-DB records to characterize different aspects of the sequences of phrases provided in the annotation files, such as repertoire use and the transition probabilities among different phrases. In addition, we explored how different machine-learning methods can be used to maximize the accuracy of individual identification within a year and across years. In our experiments, we developed machine-learning algorithms for automated classification of Cassin's vireo individuals that achieve >99% accuracy, described in Arriaga et al. (2013) and Arriaga et al. (2014). Automated classification of the phrases themselves using different classification methods adapted from human speech processing, using data from the database, are described in Tan et al. (in press) and Kantapon et al. (in press).

4.2. Network analysis of complex bird song

Song network analysis is an emerging approach to characterize the structure of complex bird songs. In this analysis, a song network is constructed from phrase sequences. A song network may be either an undirected or a directed graph, in which nodes represent different phrases and edges represent transitions between them. The representation of bird songs as networks enables the analysis of the complexity of songs by using tools from established network theory (Newman, 2010).

Previous studies on network-based analysis of complex bird songs have demonstrated the existence of small world architectures in song networks in which subsets of phrases are highly grouped and linked with a short average path length (Sasahara et al., 2012). This contrasts with purely random networks, which would exhibit a short average path length but a small grouping of phrases. Song network analyses enables a more quantitative study of structural properties of complex bird songs than has been previously possible.

For the previous studies, the availability of song recordings and their annotation of phrase types and sequences becomes critical. Bird-DB has proven to be a useful resource for these studies. Additional analysis is possible with the files in the database. In (Sanchez et al. (in press)) we describe a Mathematica Toolbox developed to aid in the further analysis of phrase sequences.

4.3. Limitations

It is important to clarify that Bird-DB focuses on providing annotations for studying the context of songs as well as their structure. Linking this directly to the birds' behavior from the database is currently not possible within Bird-DB. This is because there are currently no standard ontologies, i.e., specific terms to describe classes of behaviors and the relationships among them, for behavior of free-living organisms. Such ontologies are necessarily community efforts, and while such an ontology does exist for model animals in the laboratory (Neurobehavior Ontology, 2014), there is currently none for non-model organisms living in natural environments. There is an on-going series of meetings to arrive at such a consensus ontology, with reason to believe that an appropriate draft will be available before too long (National Science Foundation, 2014); we presume it will be possible to integrate this with Bird-DB when it is available.

A major benefit from the current database will be to compare methods for automatic classification and annotation of bird songs. Many labs are now engaged in automatic classification, most using techniques borrowed and adapted from human speech processing (Brandes, 2008). The Harvard workshop mentioned earlier recommended that a standard toolbox of methods for classification, a platform to compare them, and also to compare different human classifications, would be helpful. This is especially so because the community of speech processing experts is highly specialized and largely separate from those engaged in the study of natural animal behavior. There are several visions about how this might be done (Bradbury, personal communication); we anticipate that the current database could readily accommodate such an expansion.

Finally, compared to other similar online databases, Bird-DB is still lacking many advanced features for searching and displaying data, specifically in-browser integration of different file formats (for example Xeno-canto embeds localization files and spectrograms directly on results pages, making visualization of data clearer and more user-friendly). This does not seem especially daunting, and the project is still in active development; more features along this line will be added as they are completed.

5. Conclusions

The new study of bird song syntax requires extensively annotated files of bird songs recorded in natural settings. These entail a great deal of metadata regarding the environment from which recordings were made, together with large and currently expensive corpora of annotated recordings. We have constructed a database that permits such recordings and annotations as a resource shared by the community. The structure and use of the database are described here. Access to this information will allow researchers the opportunity to analyze the structure of birds' vocalizations, study species' repertoires among several individuals throughout varying periods of time, explore possible effects of seasonal and environmental conditions on a bird's song, among many other possibilities. Our online bird song database, http://taylor0.biology.ucla.edu/birdDBQuery, is publicly available and is the first of its kind to readily offer bird vocalizations consistently phrase-tagged, presenting an increasingly growing collection of samples from several bird species. Data of this nature is fundamental for the study and eventual understanding of the underlying workings of animal vocalizations.

Acknowledgments

This work was supported by the US National Science Foundation under Award Number 1125423 and by Consejo Nacional de Ciencia y Tecnologia under Award Number 1010/214/2012. Neither sponsor had a direct role in the design, the collection, the analysis or the interpretation of this data, the writing of the report or the decision of where to publish.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ecoinf.2015.01.007.

References

Angle, G., Coskun, H., 2014. A complete syllable dictionary for *Serinus canarius*. Ecol. Inform. 20, 67–75.
Arnold, K., Zuberbühler, K., 2008. Meaningful call combinations in a non-human primate. Curr. Biol. 18, R202–R203.
Arriaga, J.G., Kossan, G., Cody, M.L., Vallejo, E.E., Taylor, C.E., 2013. Acoustic sensor arrays for understanding bird communication. Identifying Cassin's vireos using svms and hmms. Advances in Artificial Life, ECAL. vol. 12, pp. 827–828.
Arriaga, J.G., Sanchez, H., Hedley, R., Vallejo, E.E., Taylor, C.E., 2014. Using song to identify Cassin's vireo individuals. A comparative study of pattern recognition algorithms. Mexican Conference on Pattern Recognition 2014, MCPR2014, volume LNCS 8495. Springer-Verlag, pp. 291–300.
Boersma, P., Weenink, D., 2014. Praat: doing phonetics by computer [computer program]. Version 5.3.68. http://www.praat.org/.
Brandes, T.S., 2008. Automated sound recording and analysis techniques for bird surveys and conservation. Bird Conserv. Int. 18, S163S173.
British Library, 2014. British library sound archive. http://www.bl.uk/soundarchive (Accessed: 2014-08-07).
Catchpole, C.K., Slater, P.J.B., 1995. Bird Song: Biological Themes and Variations. Cambridge University Press.
Cornell, 2014. Cornell Library of Ornithology Macaulay Library. http://macaulaylibrary.org (Accessed: 2014-08-07).
CSIRO, 2011. Australian national wildlife collection sound archive. http://www.csiro.au/Organisation-Structure/Divisions/Ecosystem-Sciences/ANWC-Sound-Archive (Accessed: 2014-08-07).
Elmasri, R., Navathe, S., 2013. Fundamentals of Database Systems. Pearson Education, Limited.
Evans, C.S., 1997. Referential signals. Perspect. Ethol. 12, 99–143.
Harvard University, 2014. Understanding bird songs: the role of automated speech processing. http://projects.iq.harvard.edu/understanding_bird_songs (Accessed: 2015-01-07).
Kantapon, K., Tan, L.N., Alwan, A., Taylor, C.E., 2015s. A robust automatic phrase classifier using dynamic time-warping with prominent region identification. Proc. ICASSP (In press).
Kershenbaum, A.a., 2014. Acoustic sequences in non-human animals: a tutorial review and prospectus. Biol. Rev. http://dx.doi.org/10.1111/brv.12160.
Lepage, D., 2014. Avibase — the world bird database. http://avibase.bsc-eoc.org/avibase.jsp (Accessed: 2014-08-07).
Margoliash, D., Nusbaum, H.C., 2009. Language: the perspective from organismal biology. Trends Cogn. Sci. 13, 505–510.
National Science Foundation, 2014. Finalizing an integrated behavior ontology for the behavioral science community. http://grantome.com/grant/NSF/IOS-1439561#panel-abstract.
Neurobehavior Ontology, 2014. The neurobehavior ontology. http://bioportal.bioontology.org/ontologies/NBO/?p=summary.
Newman, M., 2010. Networks, An Introduction. 1st ed. Oxford University Press.
OSU, 2014. Ohio State University Borror Library of Acoustics. https://blb.osu.edu/archive (Accessed: 2014-08-07).
Planqué, B., Vellinga, W.-P., 2014. Xeno-cano.org. http://http://www.xeno-canto.org (Accessed: 2014-08-07).
Sanchez, H., Vallejo, E.E., Taylor, C., 2014s. Pajaroloco: a Mathematica toolbox for the analysis of animal vocal sequences. Artificial Life and Robotics (In press).
Sasahara, K., Cody, M.L., Cohen, D., Taylor, C.E., 2012. Structural design principles of complex bird songs: A network-based approach. PLoS One 7 (9), e44436.
Tan, L.N., Alwan, A., Kossan, G., Cody, M., Taylor, C.E., 2015s. Dynamic time warping and sparse representation classification for birdsong phrase classification using limited training data. J. Acous. Soc. Am. (In press).
Tierstimmenarchiv, 2014. Animal sound archive (tierstimmenarchiv) at the Museum für Naturkunde in Berlin. http://www.tierstimmenarchiv.de (Accessed: 2014-08-07).
Utah, 2014. University of Utah Western Soundscape Archive. http://www.westernsoundscape.org (Accessed: 2014-08-07).