Phonetic Methods in Cat Vocalisation Studies: A report from the Meowsic project

Susanne Schötz[1], Joost van de Weijer[2], Robert Eklund[3]
[1] Dept. of Logopedics, Phoniatrics and Audiology, Lund University, Sweden
[2] Humanities Laboratory, Lund University, Sweden
[3] Dept. of Culture and Communication, Linköping University, Sweden
susanne.schotz@med.lu.se, vdweijer@ling.lu.se, robert@roberteklund.info

Abstract

In the project Melody in Human–Cat Communication (Meowsic) we are using established phonetic methods to collect, annotate, pre-process and analyse domestic cat–human vocal communication. This article describes these methods, and also presents results of meow vocalisations in four different mental states showing variation in fundamental frequency (f0).

Introduction and background

Although phonetics is basically the study of human speech, phonetic methods have increasingly been used in several studies of non-human animal vocalisations, including those of birds (Hunt, 1923), dogs and coyotes (Riede, Mitchell, Tokuda, & Owren, 2005), giraffes (Baotic, Sicks, & Stoeger, 2015), goats and sheep (Ruiz-Miranda, Szymanski, & Ingals, 1993; Morton, Hinch, & Small, 2017), cheetahs (Eklund, Peters, & Duthie, 2010) and domestic cats (Schötz, 2018), just to name a few.

Moreover, speech analysis tools like Praat (Boersma & Weenink, 2019) have been used to explore non-human vocalisations in several studies (e.g. Nicastro, 2004; Riede et al., 2005).

Although some non-humans communicate in the infra- or ultrasound ranges, many animal sounds are audible to the human ear. Moreover, despite there being differences in size and shape of the vocal organs of non-human mammals compared to humans, we tend to produce many sounds in a similar fashion: with an airflow (e.g. from the lungs) which is obstructed in the glottis and/or further along the vocal tract to generate a sound wave, which is modified by resonance (through the shape and size of the vocal tract) by moving our lips, jaws, tongues and other articulators. Birds—despite them having syrinxes instead of larynxes—also share some of these similarities.

Many animal sounds share at least some spectral and prosodic features with human speech, which further motivates the use of similar recording and analysis methods. For instance, Hunt (1923, p. 202) had the following arguments for using phonetic methods when describing bird sound: "It is believed to be the natural and logical system for three chief reasons: (1) it is perfectly flexible, providing for the recognition of all factors in bird sound (pitch, intensity, rate of speed, form, expression, timbre, and phonetic quality); (2) the working symbols are those of the English language, and are therefore common property and do not require special knowledge or technique; (3) these symbols are essentially suited to the expression of bird sounds, since bird sounds are essentially human speech-like."

Phonetic theories and methods have developed further since then, but are still used in many studies of acoustic communication in non-humans, including the domestic cat.

Moelk (1944) was probably the first to describe the vocal repertoire of the domestic cat using phonetic symbols, and the majority of more recent descriptions are based on her observations and classification of the different sound types and patterns. In the project, *Melody in Human–Cat Communication (Meowsic)* we follow this convention, and hope to add to the knowledge about the phonetic characteristics of different cat vocalisation types, although our studies focus on their prosodic properties (Schötz, Eklund, & van de Weijer, 2016). This paper describes the methods used up till now to collect, pre-process and analyse our data, and also presents some preliminary results.

Data collection

Collecting high quality spontaneous and natural audio data from non-humans is often challenging. We tested numerous video and audio recording devices in several settings to find the best and most flexible ones. Our first pilot recordings with small on-animal video cameras (so called pet cams) were found to have a far too poor and noisy audio quality. Several hand-held, tripod-based and head-worn video cameras were tested next, and we finally decided on a flexible set up of one camera on a tripod to capture the whole scene, a head-worn camera to capture close ups of the cats, and small on-animal microphones to capture high quality audio, which were exchanged for small hand-held video cameras whenever the cats preferred not to wear the microphones. Depending on the cat's personality — some cats were social, others shy, and some refused to wear the microphones — we adapted the recording equipment and set-up to each cat, and used several recording devices simultaneously to increase the chances of recording audio suitable for phonetic analysis. Additionally, we often received video clips recorded by the cats' owners. Table 1 lists the equipment we most commonly used in the project, and Figure 1 shows a typical set-up.

Table 1. The different types of recording equipment tested in the project. The main recording equipment is shown in boldface.

Video
Canon HG-10 camcorder with professional Audiotechnica AT813 mono microphone
Canon HG-10 camcorder with Canon DM-50 clip-on microphone
GoPro Hero 4 Session with internal microphone
Sony Handycam HDR-CX730E camcorder with internal microphone
Owner's mobile phone (different brands and models) with internal microphone
Nikon system camera with clip-on microphone
Olympus Tough digital camera with internal microphone
Audio
Roland R-09HR WAVE/MP3 recorder with **Sony ECMAW4** wireless microphone(s)
Marantz PMD660 audio recorder with IMG ECM-302B stage boundary microphone
Roland R-09HR WAVE/MP3 with internal microphone

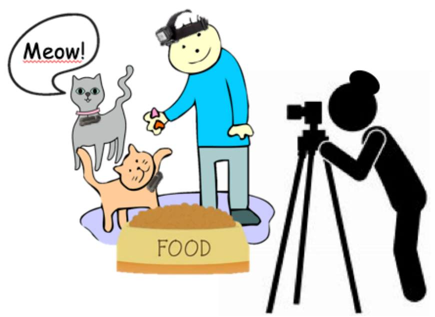

Figure 1. Typical set-up for data collection with the cats wearing wireless microphones on their collars and the owner with a GoPro camera on the head.

Data were collected mainly in and around the cats' homes and in normal everyday contexts like feeding time, play, or wanting to be let out or in. Care was taken not to disturb the communication between cats and owners, but occasionally the authors also interacted directly with the cats.

In addition to recording data, we collected metadata about the cats' age, sex, breed, some of their habits (vocal food, sleep, time spent outside, etc.), and also asked their owners to fill out a personality form.

The recorded video and sound files were transferred to a server, and audio (wav, 44.1 kHz, 16 bit) was extracted from the video files. All data (at present, we have recorded 70 cats from 35 households in the Skåne, Stockholm and Östergötland areas) were anonymized.

Annotation

The general contexts and mental states of the cats were assessed by watching the video files, and the cat vocalisations was annotated by the first author in Praat using a limited set of crude vocalisation types, contexts and mental states. We used a set of mental states we consider uncontroversial, although we are aware of the complications this field contains.

A randomly selected part of the files was annotated by the second author and inter-labeller agreement for annotation of the cat vocalisations was assessed. Results showed varying degrees of agreement between the two labellers with kappa values ranging from 0.43 to 0.97 with an average of 0.70.

Tables 2–4 show the categories used for annotating context, mental state and vocalisation type respectively. An example of an annotation in Praat is shown in Figure 2.

The cat–directed speech was transcribed orthographically.

Pre-processing and analysis

All cat vocalisations were extracted and f0 pitch objects (i.e. the fundamental frequency (f0)) were generated with a user-controlled Praat script which enabled us to use different settings (functions in Praat (toPitch or to Pitch(ac)) as well as different values for the f0 floor and ceiling), depending on the vocalisation type and the average pitch of each individual cat. This way each f0 contour could be manually checked. Voiceless vocalisations (e.g. chattering, hissing and spitting) as well as vocalisations that were too weak or too noisy to reliably extract a plausible or reasonably adequate f0 were discarded and excluded from further analysis.

Table 2. Annotation types for recording contexts used in the project.

Type	*Description/example(s)*
food	Food-related (food soliciting, feeding time, wants treat, etc.)
drin	Drink-related (drink soliciting, wants faucet turned on, etc.)
door	Door/window-related (wants to be let in/out of a room, garden, etc.)
play	Play-related, e.g. play soliciting (except hunt)
cudd	Cuddle-related, while being petted, during cuddling, cuddle soliciting
lift	When a human grabs, lifts (or tries to lift) up and/or holds the cat
tbox	Transport related (entering, travelling in or exiting a cat carrier or transport box)
vets	Veterinary related (at the vet's outside tbox)
rest	When resting calmly or sleeping
terr	Territorial (indoors, outdoors)
gree	Greeting (friendly greeting behaviour)
groo	Grooming (cat, human or other animal)
isol	Isolation (away from mother, siblings, human caretaker or similar)
hunt	Hunting-related (watching, stalking, chasing, catching prey or prey-like object (e.g. toy))
outs	Outside (combined with other context, e.g outside,hunt)
walk	Walking, roaming, digging etc.
bath	Before, during or after bath or shower
othe	Other or neutral, no specific context identified

Table 3. Annotation categories for mental states (numbers indicate level of excitement or strength of mental state: 1: mild, 2: medium, 3: strong) used in the project.

Type	*Description/example(s)*
con1-3	Content, satisfied, happy, friendly
dis1-3	Discontent, annoyed, hostile, distressed
att1-3	Attention soliciting
str1-3	Stressed, anxious, nervous
foc1-3	Focused, alert, interested
aro1-3	Aroused, excited, upset, agitated
neu1-3	Neutral, no specific mental state identified

Table 4. Annotation categories for pure and complex types of domestic cat vocalisations.

Pure types	*Description/example vocalisation types*
Pr (P)	Purr
Me (M)	Meow, Mew, Squeak, Moan (all pure meow sounds)
Tr (R)	Trill, Chirrup, Chirr
Ch (C)	Chirp, Chatter, Tweet, Tweedle (all prey-directed sounds)
Gr (G)	Growl
Ho (H)	Howl, yowl
Sn (S)	Snarl, Cry, Scream, Defence screech, Pain shriek
Hs (Hz)	Hiss
Sp	Spit
Mm	Murmur ([m]-like nasal, no trill, no vowel, mouth closed)
Mt, Tm	Merged/Mixed Tr and Me/Mm (e.g. when open lips)
Ot	Other, Unable to identify/classify this sound
Complex types	*Description/example vocalisation types*
TM, MT	Trill-Meow, Meow-Trill
GH, GT, TG, GS	Growl-Howl, Growl-Trill, Trill-Growl, Growl-Snarl
HG, HL, TH	Howl-Growl, Howl-Lateral (howling with laterals), Trill-Howl (±laterals)
MH, HM	Meow-Howl, Howl-Meow
TP, PT, CT	Trill-Purr, Purr-Trill, Chirp-Trill
MP, PM	Meow-Purr, Purr-Meow

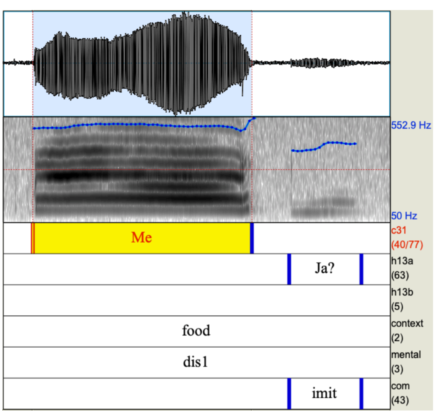

Figure 2. An example annotation of a meow.

To facilitate between-cat comparison, the contours were normalised by setting the minimum f0 for every cat to 0 semitones Mean f0 contours were obtained for each context and mental state by calculating the mean f0 in 100 evenly distributed points in each meow.

Figure 3 shows an example of individual f0 contours and the corresponding mean f0 contour for the context *play*.

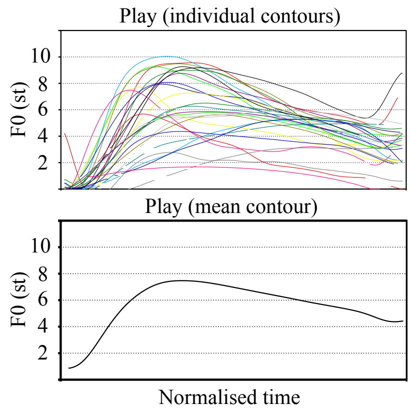

Figure 3. Individual and mean f0 contours for the context *play* (st: semitones).

Finally, measures of duration and f0 (maximum, minimum, mean and standard deviation (sd) for all vocalisations were extracted in Praat.

Results

Description of the material

Currently, the processed dataset consists of 1591 vocalisations produced by 54 cats (45 adults, 9 kittens), an average of a little more than 27 vocalisations per cat. The most frequent vocalisations were meows (57%), trills (10%), trill-meows (9%), growls (6%), purrs (5%) and chirps (4%). All remaining vocalisation types occurred 2% or less in the data that has been processed so far. Many of the cats were recorded in a single context, but a substantial part of the vocalisations was produced by cats that were recorded in two, three or even four different contexts. Most vocalisations were recorded in a *food* context (49%). Other contexts that were relatively frequent were *greet* (11%), *tbox* (11%), *play* (6%) and *door* (6%).

Acoustic measurements

Table 5 shows estimated values of the five acoustic parameters (with 95% confidence intervals) that we have measured so far. The values are based on meows only and have been separated into adult cats and kittens.

Table 5: Estimated acoustic parameters for meows (CI: confidence interval, sd: standard deviation).

Parameter	*Estimate*	*95% CI*
kittens (n = 9)		
duration	0.488	0.420-0.509
f0 mean (Hz)	1238	991-1471
f0 min (Hz)	1138	895-1361
f0 max (Hz)	1307	1035-1566
f0 sd (Hz)	48	26-67
adults (n = 45)		
duration	0.689	0.634-0.740
f0 mean (Hz)	593	551-636
f0 min (Hz)	521	479-563
f0 max (Hz)	659	614-704
f0 sd (Hz)	38	34-42

The values in Table 5 show that meows produced by kittens are much higher in f0 and also much shorter than the ones produced by adults. Within the adult group we looked at effects of cat weight, age and gender on each of the parameters. The adult cats were on average 6.2 years old (range 1-15 years), had an average weight of 4.6 kilos (range 2.5-7.7) and 19 (43%) of them were male cats. Of these three variables, weight and age were somewhat correlated: male cats were on average heavier (5.16 kilos) than female cats (4.15 kilos). The analysis showed that age had a significant effect on mean f0 (EST = -14.053, SE = 6.731, df = 30.098, t = -2.088, p = 0.045), minimum f0 (EST = -13.831, SE = 6.404, df = 28.666, t = -2.160, p = 0.039) and maximum f0 (EST = -15.324, SE = 7.184, df = 30.173, t = -2.133, p = 0.041), but not on f0 sd or on duration. In other words, meows produced by older cats were lower in average minimum and maximum pitch than meows produced by younger adult cats. The kittens were all 0.25 years old, their weight was approximately two kilos, and four of them were male. Since the kittens were few and all had the same weight and age, they were not analysed further here.

F0 contours in different mental states

Figure 3 shows average f0 contours of adult cat meows from four mental states. There is a general tendency for the mean contours produced in the mental states *attention* and *content* to rise, while those produced by *discontent* and *stressed* cats generally display more falling patterns.

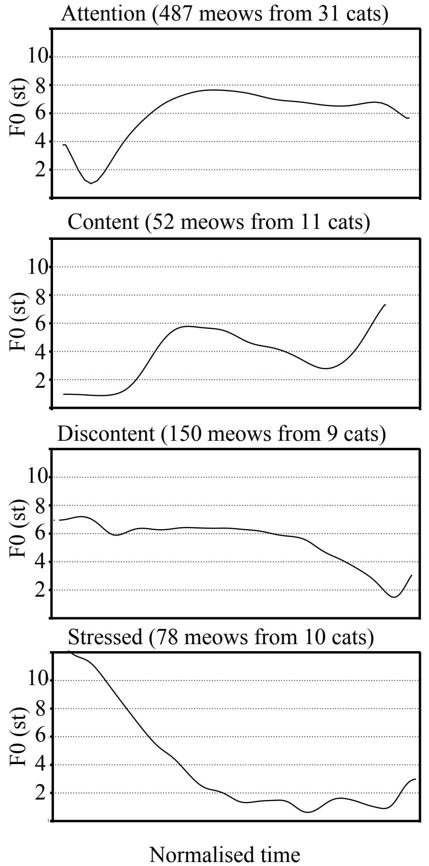

Figure 3. Mean f0 contours of cat meows in four different mental states (st: semitones).

Discussion

The notion of “animal language” is, admittedly, not unproblematic (see e.g. Rendall, Owren, & Ryan, 2009), but for lack of space we will not address this in further detail here. However, humans and other mammals are likely to share similar mental states, experienced and expressed as emotions (see Briefer, 2012). Although animal emotions were until recently controversial, this has recently changed (Bekoff, 2007, p. 42). Our data suggest that 1) phonetic methods can be used to study domestic cat vocalisations, and 2) the physical and mental state of cats influences f0 and duration in their vocalisation.

In future studies we will investigate to what extent cats signal mental state in their voices by analysing a larger number of acoustic-phonetic features in more vocalisation types from a greater number of cats.

Acknowledgements

The authors gratefully acknowledge the Marcus and Amalia Wallenberg Foundation and Lund University Humanities Lab. Special thanks go to all participating cats and their owners.

References

Baotic, A., Sicks, F., & Stoeger, A. S. (2015). Nocturnal “humming” vocalizations: adding a piece to the puzzle of giraffe vocal communication. *BMC Research Notes*, *8*(1). https://doi.org/10.1186/s13104-015-1394-3

Bekoff, M. (2007). Are you feeling what I’m feeling? *New Scientist*, *26 May 2007*, 42.

Boersma, P., & Weenink, D. (2019). *Praat: doing phonetics by computer (Version 6.0.46) [Computer program].*

Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence: Vocal communication of emotions. *Journal of Zoology*, *288*(1), 1–20. https://doi.org/10.1111/j.1469-7998.2012.00920.x

Eklund, R., 1962-, Author, Peters, G., Author, & Duthie, E. D., Author. (2010). An acoustic analysis of purring in the cheetah *(Acinonyx jubatus)* and in the domestic cat *(Felis catus)*. *In Procedings of Fonetik 2010, Lund University*, 17–22.

Hunt, R. (1923). The Phonetics of Bird-Sound. *The Condor*, *25*(6), 202–208. https://doi.org/10.2307/1362681

Moelk, M. (1944). Vocalizing in the House-Cat; A Phonetic and Functional Study. *The American Journal of Psychology*, *57*, 184–205. https://doi.org/10.2307/1416947

Morton, C. L., Hinch, G., & Small, A. (2017). Distress vocalization delay in the neonate lamb as a neurobehavioral assessment tool. *Developmental Psychobiology*, *59*(4), 523–534. https://doi.org/10.1002/dev.21517

Nicastro, N. (2004). Perceptual and Acoustic Evidence for Species-Level Differences in Meow Vocalizations by Domestic Cats *(Felis catus*) and African Wild Cats *(Felis silvestris lybica)*. *Journal of Comparative Psychology*, *118*(3), 287–296. https://doi.org/10.1037/0735-7036.118.3.287

Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, *78*(2), 233–240. https://doi.org/10.1016/j.anbehav.2009.06.007

Riede, T., Mitchell, B. R., Tokuda, I., & Owren, M. J. (2005). Characterizing noise in nonhuman vocalizations: Acoustic analysis and human perception of barks by coyotes and dogs. *The Journal of the Acoustical Society of America*, *118*(1), 514–522. https://doi.org/10.1121/1.1928748

Ruiz-Miranda, C. R., Szymanski, M. D., & Ingals, J. W. (1993). Physical characteristics of the vocalizations of domestic goat does Capra hircus in response to their offspring’s cries. *Bioacoustics*, *5*(1–2), 99–116. https://doi.org/10.1080/09524622.1993.9753232

Schötz, S. (2018). *The secret language of cats: how to understand your cat for a better, happier relationship*. Hanover Square.

Schötz, S., Eklund, R., & van de Weijer, J. (2016). Melody in Human–Cat Communication (Meowsic): Origins, Past, Present and Future. *Proceedings of Fonetik 2016*, 19–24. Stockholm: TMH-QPSR.