I want to talk to my Cat (Bubbles)

Feline Kin and Nurturing Understanding within the Inner Prompting

Mika Zoé Rosenberg
Universität der Künste
MA Design & Computation
mika.rosenberg@gmail.com

Abstract

This paper presents a foray into bioacoustic research at home, focusing on our domestic companion, the cat (Felis catus). The research is conducted through fieldwork, involving audio data collection and processing from a recording device attached to one cat subject, named Bubbles. The aim is to collect cat vocalizations by marking their occurence with a hotword-marker. Subsquently, hotword-detection speech-to-text models (e.g. Whisper) allow us to extend dataset of animal vocalizations. With such an extended dataset at hand, meow-classifier is trained by means of classical ensemble learning (Random Forest) technique. The paper offers a comparison of these methods and establishes a minimally invasive protocol for future data collection of in vivo and in situ vocalizations of domestic animals.

I. State of the Art

In the realm of human-pet relationships, we find ourselves as novice listeners in a complex dialogue. Our companion species are reflexive participants in conversations in a language not their own, yet they engage, respond, and, in their own way, converse. This dynamic unveils a unique facet of our coexistence, where empathy transcends spoken language, but in this symphony of multispecies sound, how can we decode understanding [1]?

My feline consorts have grown up with me, more than once. This journey with Bubbles, my cuddly bewitching mischief-maker, who was born June 24th, 2023, has been a looking glass into the nuanced communication between humans and their feline companions. Observation and research involving myself, Mika, and Bubbles, have shed light on the intricate vocalizations of domestic cats that extend past meowing [7] and the attempts to, classify and interpret acoustic signaling [25].

These vocal patterns, distinct from feline-to-feline communication and almost virtually absent in non-socialized cats [4], have evolved and been primarily used for cat-to-human communication [2]. Understanding these vocal cues opens a new frontier in enhancing human-domesticate interactions and welfare [16].

This story is about a cohabitation that began 9,500 years ago [28] and the evolution of domestication by natural selection over artificial selection [3]. Over the centuries, cats have become precious companions, integral to our social fabric [13].

Recent advancements in interspecies communication research have opened new avenues in understanding this intricate relationship. Groundbreaking studies employing automatic analysis techniques in bioacoustic research have provided insights into the communication patterns of various species, including sperm whales, dolphins, bats, and numerous bird species [24, 5, 6]. Parallel to these developments, significant progress has been made in decoding cat vocalizations and their specific communication with humans, notably the work led by Ntalampiras at the University of Milan [19, 18]. It points towards a future no longer tuned to a taciturn channel, where we can coexist and engage in meaningful communication with our non-human kin.

The field of bioacoustics involves the study of animal sounds and their characteristics [8]. With the decrease in the cost of recording devices [27], there has been a significant increase in the amount of data available. The manual processing, labeling, and analysis of these large datasets have become a bottleneck [27]. Machine learning techniques, particularly deep neural network models, have become increasingly popular in bioacoustics research [1] as they offer the potential to automate bioacoustic tasks.

Nonprofits like Earth Species Project [1] are already underway with tackling these complex tasks, like creating benchmarks in machine learning for bioacoustic research [10] and working towards self-supervised audio representation models for encoding animal vocalizations [9]. In line with the ever-expanding research in this field, I am conducting a study in collaboration with bubbles, with a companion species perspective and a multi-species existence in mind. This project invites reflection on the use of AI in animal communication and explores ideas of kinship and understanding [12].

II. Larger Aim

I hypothesize that acoustic cues may develop expressly in reference to cats' unique relationship with their companion(s).

Vocal cues may be learned or developed through ritual exposure and adaptation. Very few studies have been conducted exclusively on "ontogenetic ritualization" [26] or conventionalization, on cats and their human companions. Arguing for a study of the individual relationship between cat and their human caretakers- capturing the development of a relationship, and the study of the specific call and response patterns of the cat and human.

In their research, Nicastro and Owren suggest that with repetition of interactions during relational development, cats could learn to adopt acoustic cues that have specific referential value to their owner [16, 15, 17], or as I refer to as companion(s), and that those specific companions can better attach meaning to the acoustic qualities of calls produced in specific contexts.

Ontogenetic Ritualization is defined by Halina et al in context to gestures in bonobo primates, "in which individuals learn their gestures in the context of regularly occurring dyadic interactions such that parts of fully functional social behaviors become ritualized [11]". Variations in affective valence (positive vs. negative) might lead to a "bootstrapping" ritualization process, where owners grow more in tune with their cats' emotional states and specific requirements in various contexts [11].

For these reasons, this research acts as a pilot study of one companion household, myself and Bubbles, to establish a protocol for the study of individual relational development. In the upcoming chapter, I will provide a detailed explanation of the problem's parameterization.

III. Problem Parametrization

This research aimed to create a comprehensive dataset of natural cat vocalizations within a home environment. The primary feline subject was a 4 to 9-month-old British Shorthair female kitten, Bubbles, not neutered. A 27-year-old female, Mika, was the human participant. Random recordings were made, capturing the everyday activities and routines of the home environment, ensuring a diverse collection of vocal samples. The audio data, ranging from a few seconds to several hours, was recorded using a small device attached to the cat's harness. This method captured a broad spectrum of vocaliza-

[1] https://www.earthspecies.org/

tions in different domestic settings.

Two kinds of recordings were developed for cat vocalization detection: Passive Acoustic Monitoring (PAM) [23] and Active Acoustic Monitoring (AAM). Both methods were employed, with AAM focusing on hotword detection and PAM designated for the Random Forest classifier or future applications. These methods were later evaluated for their effectiveness in automated detection of cat vocalization.

The study's initial goal was to develop an automated system for isolating specific audio segments containing "hotwords," like "bubbles," from extensive audio files. This was achieved by utilizing the Whisper speech recognition model [22] [2], a self-supervised pre-trained ASR Model, to detect these "hotwords". In AAM, hotwords were preselected to be mentioned post-vocalization. These included terms like "bubbles" and combinations such as "bubbles + food," "bubbles + cuddle," "bubbles + milk," etc., to provide additional context for future classification. Once a targeted hotword was identified by Whisper, the corresponding audio segment was automatically isolated.

The second aim involved compiling a significant dataset through this automated hotword detection process. This dataset, combined with A publicly-available dataset of Cat Vocalizations named CatMeows [14] [3], was then used to train a Random Forest classifier, an ensemble method incorporating multiple decision trees. The use of this classifier in the training phase was intended to enhance the precision of cat vocalization detection, aiming to build a model capable of recognizing these sounds without the need for specific human voice commands.

The research's objective was to assess and compare the effectiveness and data yield of two automated methods: hotword detection using OpenAI's Whisper and the Random Forest classifier.

IV. Optimization method(s)

Figure 1: *Bubbles and her Harness*

I. Data Acquisition

During the preliminary design phase, I developed a method to record the subject, Bubbles, vocalizations in a non-invasive manner. Technical and practical problems regarding the process had to be addressed and solved, in terms of meeting the specification requirements- prioritizing the comfort and wearability for the user while not minimizing potential influences of the device being worn, changing or affecting behavior patterns.

The recording device's relative position to the sound source had to be fixed. Capturing sound as close to the source as possible was the simplest approach. I opted for a wearable device attached to a harness worn by the cat. The chosen harness had been worn by the subject on several occasions, ensuring the cat was accustomed to it and could wear it comfortably for extended periods. However, it's crucial to recognize that not all animals readily accept wearables; some may require conditioning or training to adapt.

Limitations of my approach - the device at times at times captured disturbances or unintentional noise due to the placement. For example, the cat rolling around the floor or playing, caused a lot of unwanted noise which may have corrupted some of the recordings.

[2] https://openai.com/research/whisper

[3] https://zenodo.org/records/4008297

Other research adopted alternative design methods of data collection. In both cases a small, lightweight microphone placed under the cat's throat through a collar, the collar to which the animal in their study was already familiar, or became accustomed during a period of training. (recording devices were small Bluetooth music headsets with built-in microphones, headsets were then paired to smartphones with audio recording applications [19, 21])

Although replication of their design was considered in my approach, my subject was not accepting of the neck collar, ripping it off after 5 to 20 minutes of wear, as the collar had a breakaway function for safety she quickly discovered the quick release capability and used it to her advantage. More training on wearing the collar would need to be done until she feels comfortable wearing it for long periods or permanently. In addition, the built-in recording capability offered in the dictaphone was an advantage in my case and the presence of Bluetooth earbuds with microphones may have proved a challenge in simplicity of use.

In the study, audio recording was conducted using the *Mini Dictaphone TDW Mini Recording Device 8GB with Microphone Digital*, a device capable of recording at a sample rate of 48kHz, which was purchased from Amazon[4] for € 32.99.

II. Bubbles Extractor

The objective of this study was to develop an automated process to identify and extract specific audio segments containing the hotwords, example " bubbles ", from larger audio files collected from the recorder using OpenAI's Whisper model.

The analysis began by importing necessary Python libraries, including Whisper for audio transcription and PyDub for audio manipulation. The Whisper model, specifically the " large " variant, was loaded for high-quality transcription. To manage computational resources effectively, I specified the use of a particular GPU using Torch's " cuda.set_device " method.

The audio file was converted from WAV format to a more compatible format using FFmpeg, a powerful multimedia framework. This conversion ensured compatibility with Whisper's transcription requirements. Whisper, an automatic speech recognition model, transcribed the audio file. The transcription process segmented the audio and provided a text representation for each segment, including the start and end times of these segments.

A custom function "find_bubbles_segments" was developed to identify segments containing the keyword "bubbles," including variations like 'Bubbles talk.' This function parsed the transcription results, locating segments where the specified keyword was mentioned, and recorded their start and end times. The "extract_preceding_audio" function was then used to extract between 2.5 and 5 second audio segment(s) preceding each identified keyword instance. Each extracted segment was saved as a separate WAV file, with filenames indicating their start and end times relative to the original recording.

Lastly, a tab-separated values (TSV) file was created to log and organize the extracted segments. This file, stored in a designated directory, provides a structured and easily accessible record of all segments extracted during the analysis. Each entry in the file corresponds to an extracted audio segment, facilitating data management and further analysis. The methodology included steps for validation of segment extraction accuracy, though details of these steps will be elaborated in the Results section.

III. Random Forest Classifier

In this study, I utilized a dataset collected from the whisper program aggregate with a pub-

[4] https://www.amazon.de/

[5] https://zenodo.org/records/4008297

licly available dataset,[5] consisting of audio files featuring a mixture of cat vocalizations. We manually labeled each audio file with either ”T” (indicating the presence of a meow) or ”F” (indicating the absence of a meow). These files were processed using the Librosa library, a Python package adept for audio analysis. Our primary function, ”extract_features”, extracted Mel-frequency cepstral coefficients (MFCCs) to capture the power spectrum of audio signals. For each file, I computed the MFCCs using Librosa's 'mfcc' function, setting the count of coefficients to 10. These coefficients were averaged over time to form a feature vector representative of each audio file.

Dataset Splitting and Normalization

We divided the dataset into training and testing sets, adhering to a 90:10 split ratio, utilizing scikit-learn's ”train_test_split” function. The ”StandardScaler” class from scikit-learn was employed to normalize the feature data by removing the mean and scaling to unit variance. This normalization step is essential for the Random Forest algorithm used in our study, which is sensitive to the scale of input features.

Model Training

We opted for the Random Forest classifier, an ensemble learning method combining multiple decision trees to output the class that is the mode of classes of the individual trees. The ”RandomForestClassifier ” from scikit-learn was implemented with 100 trees. This model was then trained on the scaled training dataset.

Performance Evaluation

To evaluate the effectiveness of the model, we used scikit-learn's ”classification_report”. This report provided a detailed view of the classifier's performance across various metrics, including precision, recall, and F1-score, for each label (T and F) in the dataset.

Real-time Audio / Testing

For processing longer audio recordings, a real-time analysis feature was integrated into the algorithm. This feature processed audio in discrete chunks, each 5 seconds long. Each chunk underwent feature extraction and scaling, analogous to the procedure for the training data. The trained Random Forest model then made predictions for each audio chunk. Particular attention was given to detecting the label ”T” (presence of meow), and upon such detection, the corresponding audio waveform was visualized and the audio chunk played using matplotlib and IPython's ”Audio” display tool.

Additionally, for comprehensive data collection, audio chunks containing vocalizations and their corresponding spectrograms were saved for further analysis. This was achieved using librosa for spectrogram generation and soundfile (sf) for saving audio chunks. We also compiled a record of detected events, including timestamps and filenames, which was then exported to a CSV file for structured data management. This process of audio analysis, visualization, and data logging provided an in-depth insight into the patterns and occurrences of cat vocalizations in the audio file. The results of the analysis are presented in the following chapter.

V. Results

In this section, we delve into the empirical results obtained from the investigation into the realm of feline vocalization.

I. Quantitative Analysis of Vocalization Data for Bubbles Extraction

The collected dataset comprised of 122.03 minutes of AAM audio recordings from Bubbles. Upon processing and classification, the following quantifiable results were obtained:

Metric	Value
Total Number of Recordings	31
Total Duration of Recordings	122:03 min
Total Number of Segments	157
True Positive Vocalizations	32.5%
False Positive Vocalizations	67.5%

Figure 2: *Summary of Vocalization Data (Hotword Supervised Detection)*

Note: True Positives are where cat vocalizations were present. False Positives are when cat vocalizations are not present. False Positives were triggered usually due to user error, or false labeling, in which the word was spoken without a vocalization present, incorrect transcription from Whisper, or hallucination of the program.

II. Performance of Random Forest Classifier

A training set was created for testing the performance of the Random Forest Classifier, combining the True Positives and False Positives detected by the whisper supervised detection method and the aforementioned public dataset, CatMeows.

Metric	Value
Total Number of Samples	346
Total Duration of Samples	721.4 sec.

Figure 3: *Summary of Training set for Random Forest Classifier*

The performance of the Random Forest Classifier was evaluated using standard metrics:

	Precision	Recall	f1-score
F	0.80	0.89	0.84
T	0.75	0.60	0.67
Accuracy	X	X	0.79
Macro avg	0.78	0.74	0.75
Weight avg	0.78	0.74	0.78

Figure 4: *Performance Metrics of Random Forest Classifier on Test Set*

III. Summary of Vocalization Data (trained Random Forest classifier)

A new set of PAM recordings was created, in which hotwords were not present. The new audio file was processed on the trained Random Forest classifier, the following results were obtained:

Metric	Value
Total Duration of Recordings	28:44 min
Total Number of Segments	11
True Positive Vocalizations	54.1%
False Positive Vocalizations	45.9%

Figure 5: *Summary of Vocalization Data (trained Random Forest Classifier*

IV. Comparative Analysis

Detection Rate Improvement: A direct comparison between the initial Hotword Supervised Detection method and the trained Random Forest Classifier reveals an enhancement in detecting true cat vocalizations. Specifically, the True Positive rate escalated from 32.5% in the initial dataset to 54.1% in the dataset analyzed by the Random Forest Classifier. This suggests that the machine-learning approach is more effective than supervised adaptation of general speech recognition models for specific bioacoustic tasks.

*False Positive Reduction:*There is also a decrease in False Positive rates (from 67.5% to 45.9%), indicating an overall improvement in model performance.

Methodological Considerations: While comparing results derived from fundamentally different datasets and methods, adjustments were made to account for variations in data structure and quality. Techniques such as data normalization and statistical controls were employed to ensure that the comparison remained valid and meaningful. Additionally, potential biases introduced by disparate data sources were acknowledged and mitigated through analysis and cross-validation techniques.

In summary, the results presented herein mark a stride in the understanding of feline vocalizations. The adoption of a Random Forest classifier has evidently advanced our capability to discern true cat vocalizations in the computational space, as indicated by the notable improvement in True Positive rates and the reduction in False Positives, when juxtaposed with the Whisper hotword detection method. This leap in accuracy not only underlines the effectiveness of machine learning techniques tailored for bioacoustic studies but also opens new avenues for future explorations into the nuanced dynamics of human-cat communication.

VI. Discussion

This study has taken a novel approach to bioacoustic research, specifically focusing on the unique interspecies communication between humans and domestic cats. By employing a blend of machine learning methods, primarily OpenAI's Whisper and a Random Forest classifier, the research has demonstrated a advance in the automated detection and classification of cat vocalizations within a domestic environment. One of the notable achievements of this study is the development of a protocol for capturing and analyzing the vocalizations of domestic cats in their natural home setting. This was accomplished by creating a dataset of natural cat vocalizations, utilizing a non-invasive recording method that respected the comfort and natural behavior of the cat. The integration of machine learning tools to process and analyze these vocalizations marks a critical step in the field of bioacoustics, potentially transforming our approach to interspecies communication.

The comparative analysis between the results of the Whisper hotword detection and the Random Forest classifier indicates the superiority of the latter in terms of reducing false positives and improving the accuracy of vocalization detection. This outcome emphasizes the importance of tailoring machine learning models to the specific needs of bioacoustic research, as opposed to relying on general-purpose models.

However, the study also highlights several challenges and limitations inherent in this kind of research. The occurrence of false positives due to environmental noise, human error, or overlapping calls, underscores the complexity of accurately interpreting animal vocalizations in a domestic setting. Furthermore, the current focus on detection rather than classification leaves room for further exploration in the field. Future research will be aimed at refining these models for the classification of vocalizations, which could provide more detailed insights into the communicative behaviors of cats.

VII. Where do we go from here ?

In conclusion, this study marks a leap forward in deciphering the intricacies of human-cat interactions. We stand at the threshold of a new era in bioacoustic research, where advanced machine learning methodologies, such as the self-supervised learning techniques exemplified by AVES[9], provide promising avenues for further exploration. The incorporation of such methods not only enriches our understanding but also enhances the accuracy and depth of our interpretations. Furthermore, exploring the DAG-HMM-based classification scheme, as suggested by Ntalampiras et al.[19], presents an intriguing alternative to the current Random Forest approach. This methodology could refine our ability to distinguish and categorize cat vocalizations with greater precision.

A pivotal aspect of this progression lies in the analysis and comparison of spectrograms. By examining the True Positive vocalization samples extracted from this study and juxtaposing them with the comprehensive array of 21 distinct cat vocalizations cataloged by Tavernier et al.[25], we venture closer to a more nuanced categorization of feline vocal expressions. This comparative approach holds the potential to not just identify and classify general cat vocalizations but to segment them into specific classes such as meows, chatters, chirps, and hisses, each within its contextual framework. Such advancements are key to unraveling the layered meanings embedded within these vocalizations, ushering in an era where we can more closely interpret and understand the communicative nuances of our feline companions.

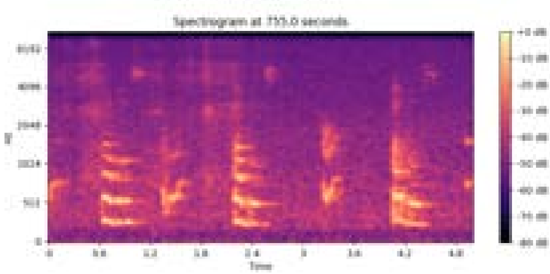

Figure 6: *Example of Spectogram from Bubbles*

This could ultimately enhance our understanding and connection with our feline companions, fostering a deeper empathetic relationship[20] between species that have coexisted for thousands of years. This study, therefore, not only contributes to the field of bioacoustics but also invites us to reflect on the broader implications of AI in fostering interspecies understanding and empathy. This project not only investigates the potential of AI in bridging empathetic connections with feline companions but also prompts us to reflect on the ethics of AI in animal communication. It revisits the idea of kinship, stemming from a long history of domestication.

In future creative iterations, I want to traverse beyond the realms of established linguistics, where human language interprets nature and the world. Here, cat vocalizations are not decoded on human terms; rather, they transcend into the physical and visual spheres.

Instead of meow representing a human language like “Trillll == I am happy to see you “or “meoooawww == I am hungry! soup time was suppose to happen 2 seconds ago!” These classification and interpretations intercepted through machine learning could be represented through colors or patterns of light, in any case, other physical visual auditory queues. Moving away from conventional forms of linguistic communication in to physical, tangible communication patterns.

I envision my bedroom lit up with a symphony of colors, patterns and soft lights that in real time becomes a token of the feelings and intention of bubbles based of her chatter, purrs, murmurs and meows. Representations that are forever evolving, dissolving and metamorphosing into new celestial meaning in our our ontogenetic rituals.

VIII. Acknowledgments

The author extends heartfelt gratitude for the invaluable support received during the course of this study. Special thanks are due to Prof. Dr. Daniel Hromada from Berlin University of the Arts, whose expertise has been indispensable. The Earth Species Project also deserves recognition for laying the foundational work in bioacoustic and machine learning research. Sincere appreciation is extended to Bastian Thürich, not only for his creative support but also for being a caring partner and the partner in care of Bubbles. The colleagues from the Optimization course have offered invaluable insights and camaraderie, for which the author is deeply thankful.

The warmest of thanks are reserved for Bubbles, Herrmann, and the rest of the cosmic cat family. Their patient participation in this pilot study was not only essential but also a source

of joy and inspiration. To these feline companions, a heartfelt Mmmeeow!, in acknowledgment of their unique contribution.

Figure 7: *Baby Bubbles and the Cosmic Cat Family*

References

[1] Karen Bakker. *The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants*. Princeton University Press, 2022.

[2] John Bradshaw and Charlotte Cameron-Beaumont. "The signalling repertoire of the domestic cat and its undomesticated relatives." In: (2000).

[3] John WS Bradshaw. "Sociality in cats: A comparative review". In: *Journal of veterinary behavior* 11 (2016), pp. 113–124.

[4] Charlotte Lucinda Cameron-Beaumont. "Visual and tactile communication in the domestic cat (Felis silvestris catus) and undomesticated small felids". PhD thesis. University of Southampton, 1997.

[5] Kevin R Coffey, Ruby E Marx, and John F Neumaier. "DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations". In: *Neuropsychopharmacology* 44.5 (2019), pp. 859–868.

[6] Florence Erbs et al. "Towards automated long-term acoustic monitoring of endangered river dolphins: a case study in the Brazilian Amazon floodplains". In: *Scientific reports* 13.1 (2023), p. 10801.

[7] Jaciana Luzia Fermo et al. "Only when it feels good: Specific cat vocalizations other than meowing". In: *Animals* 9.11 (2019), p. 878.

[8] Neville H Fletcher. "Animal bioacoustics". In: *Springer handbook of acoustics* (2014), pp. 821–841.

[9] Masato Hagiwara. "Aves: Animal vocalization encoder based on self-supervision". In: *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2023, pp. 1–5.

[10] Masato Hagiwara et al. "BEANS: The benchmark of animal sounds". In: *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2023, pp. 1–5.

[11] Marta Halina, Federico Rossano, and Michael Tomasello. "The ontogenetic ritualization of bonobo gestures". In: *Animal cognition* 16 (2013), pp. 653–666.

[12] Donna J Haraway. *Staying with the trouble: Making kin in the Chthulucene*. Duke University Press, 2016.

[13] Eileen B Karsh and Dennis C Turner. *The human-cat relationship*. 1988.

[14] Luca A Ludovico et al. "CatMeows: A publicly-available dataset of cat vocalizations". In: *International Conference on Multimedia Modeling*. Springer. 2021, pp. 230–243.

[15] Nicholas Nicastro. "Perceptual and Acoustic Evidence for Species-Level Differences in Meow Vocalizations by Domestic Cats (Felis catus) and African Wild Cats (Felis silvestris lybica)." In: *Journal of Comparative Psychology* 118.3 (2004), p. 287.

[16] Nicholas Nicastro and Michael J Owren. "Classification of domestic cat (Felis catus) vocalizations by naive and experienced human listeners." In: *Journal of Comparative Psychology* 117.1 (2003), p. 44.

[17] Nicholas Stephen Nicastro. *Evolution of a domestic cat vocalization under anthropogenic selection*. Cornell University, 2004.

[18] Stavros Ntalampiras, Danylo Kosmin, and Javier Sanchez. "Acoustic classification of individual cat vocalizations in evolving environments". In: *2021 44th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE. 2021, pp. 254–258.

[19] Stavros Ntalampiras et al. "Automatic classification of cat vocalizations emitted in different contexts". In: *Animals* 9.8 (2019), p. 543.

[20] Nur Okutan. "Development of the Animal Empathy Scale Short Form". In: *Psikiyatride Güncel Yaklaşımlar* 15.Supplement 1 (2023), pp. 177–191.

[21] Emanuela Prato-Previde et al. "What's in a meow? A study on human classification and interpretation of domestic cat vocalizations". In: *Animals* 10.12 (2020), p. 2390.

[22] Alec Radford et al. *Robust Speech Recognition via Large-Scale Weak Supervision*. 2022. arXiv: `2212.04356 [eess.AS]`.

[23] Samuel RP-J Ross et al. "Passive acoustic monitoring provides a fresh perspective on fundamental ecological questions". In: *Functional Ecology* 37.4 (2023), pp. 959–975.

[24] Christian Rutz et al. "Using machine learning to decode animal communication". In: *Science* 381.6654 (2023), pp. 152–155.

[25] Chloé Tavernier et al. "Feline vocal communication". In: *Journal of veterinary science* 21.1 (2020).

[26] Michael Tomasello and Josep Call. *Primate cognition*. Oxford University Press, USA, 1997.

[27] Devis Tuia et al. "Perspectives in machine learning for wildlife conservation". In: *Nature communications* 13.1 (2022), pp. 1–15.

[28] J-D Vigne et al. "Early taming of the cat in Cyprus". In: *Science* 304.5668 (2004), pp. 259–259.