

Human perception of intonation in domestic cat meows

Susanne Schötz, Joost van de Weijer

Lund University Humanities Lab, Centre for Languages & Literature, Sweden
susanne.schotz@ling.lu.se, vdweijer@ling.lu.se

Abstract

This study examined human listeners' ability to classify cat vocalisations (meows) recorded in two different contexts; during feeding time (food related meows) and while waiting at a vet clinic (vet related meows). A pitch analysis showed that food related meows tended to have rising f_0 contours, while vet related meows often had more falling f_0 patterns. 30 listeners judged 6 meows of each context. Classification accuracy was significantly above chance, and listeners with cat experience performed significantly better than inexperienced listeners. The food related meows with the highest classification accuracy showed clear rising f_0 contours, while clear falling f_0 contours characterised the vet related meows with the highest classification accuracy. Our results suggest that cats use different intonation patterns in their vocal interaction with humans, and that humans are able to identify these vocalisations.

Introduction

There is much anecdotal evidence of pets – especially cats and dogs – imitating speech when interacting with humans. This is probably a learned skill used to elicit certain responses or rewards, e.g. food, from their human caretakers. Because of the position of their larynx, nonhuman mammals are able to articulate only a limited number of the vowel and consonant sounds of human language (see e.g. Fitch, 2000). However, many animals can produce extensive vocal variation in duration, f_0 and intensity (SPL), and should be able to adopt human-like prosodic patterns. Gussenhoven (2002) and Ohala (1984) describe pitch features related to *biological codes*, which are used in animal

communication, e.g. the frequency code where low f_0 and resonances signal large size and dominance.

Phonetic studies of pet vocalisations are fairly scarce, and very little is known about the prosodic aspects of pet vocalisations in pet–human communication. To what extent do pets adopt and use human-like intonation in their vocal communication with humans? How are the prosodic patterns of pet vocalisations perceived by human listeners? This study is an attempt to shed some light on these issues by examining human perception of different intonational patterns in cat vocalisations.

Cat vocalisations and the meow

The cat (*Felis catus*, Linnaeus 1758) was domesticated 10,000 years ago, and is one of the most popular pets of the world with some 600 million individuals (Turner & Bateson, 2000; Driscoll et al. 2009). Cats are social animals (Crowell-Davis et al., 2004), and their interaction with humans has over a long time of living together resulted in cross-species communication that includes visual as well as vocal signals. There are several descriptions of the communicative social behaviour of the domestic cat (e.g. Turner & Bateson, 2000; Bradshaw, 2013), but those concerning vocalisations are scarce and often fragmented. It is still unclear how cats combine different sounds, and how they vary intonation, duration and intensity to convey or modulate a vocal message.

Cat vocalisations are generally divided into three major categories: (1) sounds produced with the mouth closed (murmurs), such as the purr, the trill and the chirrup, (2) sounds produced with the mouth open(ing) and gradually closing, comprising a large variety of meows with similar [a:ou] vowel pat-

terns, and (3) sounds produced with the mouth held tensely open in the same position, i.e. sounds often uttered in aggressive situations, including growls, snarls, hisses, and shrieks (Moelk, 1944; Crowell-Davis et al., 2004).

In cat-human communication, the most common vocalisation is said to be the *meow* or *miaow* (Nicastro & Owren, 2003). Nicastro (2004) defines the meow as a quasi-periodic sound with at least one formant and diphthong-like formant transitions. The duration ranges from a fraction of a second to several seconds, and the f_0 contour is generally arch-shaped with the peak marking the maximum mouth opening of the opening-closing gesture. Meows can include atonal features and may be garnished with an initial or final trill or growl.

McKinley (1982) divided the meow type vocalisation into four sub-patterns based on the pitch and vowels included in the sound: the mew, a high-pitched call with [i], [ɪ] or [e] quality; the squeak, a raspy nasal high-pitched mew-like call; the moan, an [o]- or [u]-like opening-closing sound; and the meow, a combination of vowels resulting in a characteristic [iau] sequence.

Cats learn to produce different meows for different purposes, e.g. to solicit feeding, to gain access to desired locations and other resources provided by humans. Each meow is believed to be “an arbitrary, learned, attention-seeking sound rather than some universal cat-human ‘language’” (Bradshaw, 2013). If each cat and owner develop their own arbitrary vocal communication codes, other humans would be less able to identify meows uttered by unfamiliar cats. However, if cat vocalisations contain some kind of functional referentiality (cf. Nicastro & Owren, 2003; Macedonia & Evans, 1993), i.e. that each vocalisation strongly correlates with a certain referent and also that perceiver responses correlate with the vocalisation, then experienced humans should be able to classify meows produced by unfamiliar cats fairly well.

Nicastro & Owren (2003) asked naïve and experienced listeners to judge meow calls from twelve cats recorded in five different behavioural contexts (food-related, agonistic, affiliative, obstacle, and distress). Classification accuracy was modestly (but significantly) above chance, and it was suggested that meows are unspecific, negatively toned sounds that attract human attention, but that we can learn to appreciate meows as we become more experienced.

Schötz (2012, 2013) analysed duration and f_0 in 795 cat vocalisations and found that within each vocalisation type (including the meow) durations were fairly similar, but the overall f_0 variability was high, partly due to the large number of different intonation patterns.

Purpose, aims and hypotheses

The purpose of this study was to investigate human listeners' perception of domestic cat meows with different intonation patterns. By asking listeners to classify a number of meows as belonging to one of two contexts: food related or vet related, our aim was to find out which intonation patterns are more often associated with food related vocalisations and which are more vet related. Further goals were to learn more about human perception of prosody in cat vocalisations and to increase our understanding of cat-human communication.

Based on our own previous experience of these types of meows, as well as on pitch patterns used in human speech and also related to the *frequency code*, we expected the meows of both contexts to be of similar duration and mean f_0 , but we expected a higher number of rising pitch patterns in the food related meows than in the vet related meows. We also hypothesised that experienced human listeners would judge the meows correctly more often than inexperienced listeners and also be more confident in their responses. Moreover, we hypothesised that meows with rising intonation patterns would more often be judged as food related meows than vet related meows

Material and method

Three domestic cats: Donna, Rocky and Turbo (D, R and T; 1 female, 2 males, 3 year old siblings) were recorded in two different contexts: 1) in a familiar environment; in their kitchen while waiting to be fed and 2) in an unfamiliar environment; in the waiting room (or in a car outside) of a veterinary clinic. We used a Sony digital HD video camera HDR-CX730 with an external shotgun microphone Sony ECM-CG50. Audio files (wav, 44.1 kHz, 16 bit, mono) were extracted with Extract Movie Soundtrack, and the meows extracted and normalised for amplitude in Praat (Boersma & Weenink, 2013). Six meows from each context produced by two of the cats (D and T) were selected as material, based on the overall recording quality and on judgements of the owner (one of the authors) of how representative the vocalisations were for each context. As one cat (R) was quiet during the recordings made in the vet context, no meows from this cat were used. An auditive analysis of the material by one of the authors revealed that the food related meows tended to have rising tonal patterns, while veterinary related meows had slightly arched or falling intonation. In addition, we noticed some background noise and one instance of background human speech, but this was judged not to influence the perception task.

Measures of duration and f0 were obtained with a Praat script and manually checked. One meow was significantly shorter than the other vocalisations, but we decided to keep it in order to get a first impression of how stimulus duration would influence the perception results. The other stimuli ranged between 0.58 and 1.13 seconds in duration. All stimuli contained vowels belonging to the meow type, as described by McKinley (1982), and were judged as clearly distinguishable from other common cat vocalisation types, including the purr (cf. Schötz & Eklund, 2012), the murmur (cf. Schötz,

2012) and the chirp (cf. Schötz, 2013). The longer meows were often garnished by short initial trills. Table 1 shows the duration, and the mean, minimum, and maximum f0 values for the twelve meow stimuli. Figure 1 displays f0 contours of the meows of the two contexts.

Table 1. Duration (sec.) and f0 (Hz) values for the 12 meows in two contexts (Food, Vet) by two cats (D, T).

meow	duration	mean f0	min f0	max f0
FoodD1	0.78	739	528	939
FoodD2	0.91	888	541	1003
FoodD3	0.27	797	782	816
FoodT1	1.06	532	418	582
FoodT2	0.85	539	423	653
FoodT3	1.03	567	433	640
VetD1	1.10	790	715	887
VetD2	0.80	838	764	924
VetD3	0.58	915	885	947
VetT1	1.13	510	451	589
VetT2	0.87	697	639	737
VetT3	1.02	540	487	570

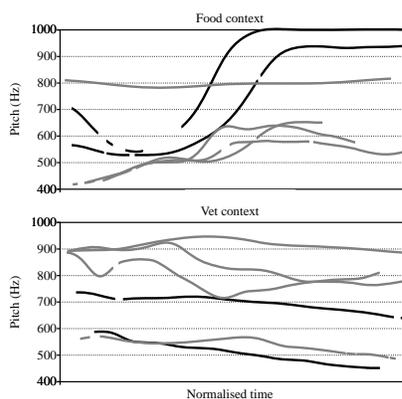


Figure 1. Time normalised f0 contours of the food and vet related meows. The black contours show the two stimuli that received the highest proportion of correct classifications in each context in the perception test.

Procedure

The experiment was designed as a multiple forced choice identification test using the ExperimentMFC function in Praat. A group of 15 men and 15 women volunteered as participants. Their average age was 44 years (range 23 to 69 years). Of the participants, 21 re-

ported being familiar with cats, that is, they either owned a cat at the time of testing, or they had owned a cat prior to the experiment. The time that these participants had owned a cat varied from less than one year to a maximum of 55 years (median 2.5 years). Oral and written instructions were given before the experiment, in which the task was to classify each meow as belonging to either the food context or to the vet context by clicking on the appropriate box on a computer screen. The experiment ran on a MacBook Pro computer in a quiet room. Each of the twelve meow recordings were presented three times in a randomised order through HUMP NF22A speakers or AKG K270 studio headphones at a comfortable sound level. A replay option allowed the participants to listen to each stimulus up to three times. After the test, the participants were asked to make a single judgement of the degree of certainty of their responses on a 5-point scale. Each session lasted about 3-4 minutes.

Results

Of all 1080 responses in the experiment 529 were food related and 551 veterinary related. In total, there were 699 correct responses (65%). The participants who reported familiarity with cats were more often correct (70%) than the participants who did not (54%).

Table 2 displays the proportions correct as well as the average reaction time for every meow stimulus. As shown in the table, there was one meow (Food D 3) that was classified incorrectly considerably more often than the other meows. This meow was exceptionally short compared to the other stimuli (cf. Table 1), and presumably contained too little information for the participants to make good judgements.

The F0 contours of the two stimuli of each context category that received the highest proportion of correct classifications are the ones drawn in black in Figure 1. For the food related meows, these contours show clear rising intonation patterns, while the two vet related

meows that received the highest number of correct classifications display more falling contours.

Table 2. Proportions of correct responses and average response time (RT) for the 12 meow stimuli in the two contexts (Food, Vet) by two cats (D, T).

meow	correct	RT (ms)
FoodD1	0.83	2342
FoodD2	0.80	2419
FoodD3	0.37	2635
FoodT1	0.54	2944
FoodT2	0.66	2673
FoodT3	0.62	2706
VetD1	0.63	3012
VetD2	0.57	2904
VetD3	0.68	2544
VetT1	0.71	2658
VetT2	0.71	3127
VetT3	0.64	3044

We performed a multilevel logistic regression (with random stimulus and subject intercepts) on the results in two steps. In the first step we did not include any predictors of interest other than the intercept. The results indicated that the overall intercept differed significantly from zero ($B = 0.7615$, $SE = 0.2529$, $z = 3.011$, $p = 0.0026$), which suggests that the overall number of correct responses was significantly above chance. In the second step, we added the familiarity predictor to the first model. This predictor had a significant effect ($B = 0.8908$, $SE = 0.3611$, $z = 2.467$, $p = 0.0136$) and overall the second model was significantly better than the first ($\chi^2 = 5.5767$, $df = 1$, $p = 0.0182$). This suggests that the participants who were familiar with cats performed significantly better than those who were not.

We also tested whether the number of years that the participants had owned a cat was a better predictor than the familiarity, but this turned out not to be the case. In fact, number of years had a non-significant effect on the dependent variable, suggesting that participants who owned a cat for a longer period of time did not score better than those who owned a cat for a relatively short time.

The participants who were familiar with cats were not only more often correct in their answers, they were also more confident in their answers. The average confidence rating given by participants familiar with cats was 2.86, whereas that given by the other participants was 1.78. This difference was tested in a linear regression analysis, which showed that it was significant ($B = 1.0794$, $SE = 0.4133$, $t = 2.612$, $p = 0.0143$).

Finally, we examined the relation between the acoustic measurements of the stimuli shown in Table 1 and the judgements made by the participants. Given the high degree of correlation between the different f_0 variables, we used only f_0 standard deviation in combination with duration as predictors of the participant choices in a multilevel logistic regression analysis. The results showed that f_0 standard deviation was a significant predictor ($B = -0.0069$, $SE = 0.0008$, $z = -8.705$, $p = 0.0000$), while duration was not ($B = 0.3969$, $SE = 0.3502$, $z = 1.133$, $p = 0.2571$). The relation between f_0 standard deviation and the listener's judgements is visualised in Figure 3. The lower the f_0 standard deviation of the stimulus, the more often it was classified as a vet related vocalisation.

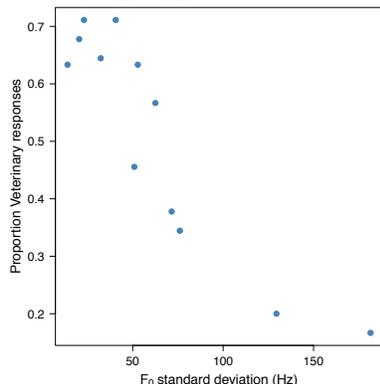


Figure 2. Relation between f_0 standard deviation and participant choice.

Discussion and future work

Our results showed that listeners were able to identify domestic cat meows from two different contexts significant-

ly better than chance, and that experienced listeners were better judges than inexperienced ones. Moreover, there was a tendency to judge meows with rising intonation as food related, and falling intonation as vet related. Our acoustic analysis showed that the food related meows tended to have rising f_0 contours often in combination with high f_0 range, while the vet related meows often had slightly falling f_0 patterns, often accompanied by a low f_0 range. It is possible that the listeners were influenced by the different f_0 ranges and interpreted them as expressions of different emotions; food related stimuli as happy (high f_0 range), and vet related stimuli as sad (low f_0 range).

A majority of the participants made the additional comment that some meows were quite easy to judge, while others were much more difficult. The meow with the shortest duration was often found very difficult to classify. Some listeners reported that they found some of the meows similar to those of their own cats. This may suggest that different cats use similar vocalisations in the contexts used in this study.

Our study suggests that cats can learn to manipulate prosodic patterns in their vocalisations in order to better elicit the desired response from their human companions. Similarly, many humans adapt their speech or speaking style to their pets by using some kind of “pet talk” (see e.g. Burnham et al. 2002). It is not unlikely that pets and their owners together develop a set of different prosodic patterns to improve inter-species communication. We hope to investigate this further in a future phonetic study of pet–human dialogues.

As far as we know this is one of the first phonetic studies of intonation in human-directed cat vocalisations, and there are numerous questions yet to be answered in order to better understand how cats and other pets use prosody in their vocal interaction with humans. Although this study examined a very limited number of meows from only two cats, our hypotheses that humans

can judge similar cat vocalisations that differ in intonation patterns significantly better than chance and that experienced listeners perform better than inexperienced ones were confirmed. In future studies, we intend to investigate other parameters, including f0 direction and movement, vowel quality and dynamics (diphthongisation) as well as intensity.

Acknowledgements

The authors gratefully acknowledge support from the Linnaeus environment Thinking in Time: Cognition, Communication and Learning (Swedish Research Council, grant no. 349-2007-8695). We are also very grateful to all our cat and human participants.

References

- Boersma, P. & Weenink, D. (2013). Doing phonetics by computer, [Computer program] Ver. 5.3.56 retrieved from <http://www.praat.org/>.
- Bradshaw, J. (2000) *Cat Sense: The Feline Enigma Revealed*, London: Allen Lane.
- Burnham, D., Kitamura, C. & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals, *Science*, 1435–1435.
- Crowell-Davis S. L., Curtis, T. M. & Knowles, R. J. (2004). Social organization in the cat: a modern understanding", *Journal of Feline Medicine and Surgery* 61:19–28.
- Driscoll, C. A., Clutton-Brock, J., Kitchener, A. C. & O'Brien, S. J. (2009). The taming of the cat, *Scientific American*, June 2009, 68–75.
- Fitch, W. T. (2000). The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals, *Phonetica*, 57, 205–2185.
- Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. In *Proceedings of Speech Prosody 2002*, Aix-en-Provence.
- Macedonia, J.M. & Evans, C.S. (1993). Variation among mammalian alarm call systems and the problem of meaning in animal signals, *Ethology* 93:177–197.
- McKinley, P. E. (1982). Cluster analysis of the domestic cat's vocal repertoire. Unpublished doctoral dissertation. University of Maryland, College Park.
- Moelk, M. (1944). Vocalizing in the House-Cat; A Phonetic and Functional Study, *The American Journal of Psychology* 57(2):184–205.
- Nicastro, N. & Owren, M. J. (2003) Classification of domestic cat (*Felis catus*) vocalizations by naïve and experienced human listeners", *Journal of Comparative Psychology* 117:44–52.
- 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):287–96.
- Ohala, J. (1984). An ethological perspective on common cross-language utilization of F0 of voice, *Phonetica*, 41:1–16.
- Schötz, S. & Eklund, R. (2011). A comparative acoustic analysis of purring in four cats, In: Quarterly Progress and Status Report TMH-QPSR, Volume 51, 2011. Proceedings from Fonetik 2011. Royal Institute of Technology, Stockholm, Sweden, pp 9–12.
- Schötz, S. (2012). A phonetic pilot study of vocalisations in three cats, Proceedings of Fonetik 2012, Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg, pp 45–48.
- Schötz, S. (2013). A phonetic pilot study of chirp, chatter, tweet and tweedle in three domestic cats, Proceedings of Fonetik 2013, Linköping University, pp 65–68.
- Turner, D. C. & Bateson, P. (Eds.) (2000). *The domestic cat: the biology of its behaviour*. Cambridge: Cambridge University Press.