LUND UNIVERSITY

A Study of Human Perception of Intonation in Domestic Cat Meows

Schötz, Susanne; van de Weijer, Joost

Published in:
Social and Linguistic Speech Prosody : Proceedings of the 7th international conference on Speech Prosody

2014

Link to publication

Citation for published version (APA):
Schötz, S., & van de Weijer, J. (2014). A Study of Human Perception of Intonation in Domestic Cat Meows. In N. Campbell, D. Gibbon, & D. Hirst (Eds.), *Social and Linguistic Speech Prosody : Proceedings of the 7th international conference on Speech Prosody* http://fastnet.netsoc.ie/sp7/sp7book.pdf

Total number of authors:
2

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 10. Nov. 2026

A Study of Human Perception of Intonation in Domestic Cat Meows

Susanne Schötz and 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 domestic cat vocalisations (meows) recorded in two different contexts; during feeding time (food related meows) and while waiting to visit a veterinarian (vet related meows). A pitch analysis showed a tendency for food related meows to have rising F_0 contours, while vet related meows tended to have more falling F_0 contours. 30 listeners judged twelve meows (six of each context) in a perception test. Classification accuracy was significantly above chance, and listeners who had reported previous experience with cats performed significantly better than inexperienced listeners. Moreover, the two food related meows with the highest classification accuracy showed clear rising F_0 contours, while clear falling F_0 contours characterised the two vet related meows that received the highest classification accuracy. Listeners also reported that some meows were very easy to classify, while others were more difficult. Taken together, these results suggest that cats may use different intonation patterns in their vocal interaction with humans, and that humans are able to identify the vocalisations based on intonation.

Index Terms: Animal–Human Communication, Human Perception of Pet Prosody, Prosody of Domestic Cat Vocalisations

1. 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. [1]). However, many animals can produce extensive vocal variation in duration, F_0 and sound pressure level (intensity), and should be able to adopt prosodic patterns similar to those used in human speech. Ohala [2] describes several prosodic features related to the *frequency code*, which are used in animal communication, e.g. low F_0 and resonances to signal large size and dominance.

Despite a recent increase in mammal vocalisation studies (see e.g. [3]), 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 use the frequency code when interacting with humans? Do pets learn to adopt human-like prosodic patterns, such as rising and falling intonation, when signalling different vocal messages to humans? How are prosodic patterns in pet vocalisations perceived by human listeners? Phoneticians who are pet owners can hardly avoid noticing the varied and often human-like prosodic patterns used in human-directed pet vocalisations. This study is an attempt to shed some light on these issues by examining human perception of different intonational patterns in cat vocalisations.

1.1. Cat vocalisations

The cat (Felis catus, Linneaus 1758) was domesticated 10,000 years ago, and has become one of the most popular pets of the world with more than 600 million individuals [5, 6]. Cats are social animals [4], 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. Although there are several descriptions of the communicative social behaviour of the domestic cat (see e.g. [5, 4, 7]), the ones 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 certain vocal message.

The vocal repertoire of the cat is characterised by "an indefinitely wide variation of sound and of patterning". 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 [ɑːou] vowel patterns, and (3) sounds produced with the mouth held tensely open in the same position, i.e. sounds often uttered in aggressive situations, including growls, yowls, snarls, hisses, spits, and shrieks [8, 4].

1.2. The meow

In cat–human communication, the most common vocalisation is said to be the *meow* or *miaow* [9]. Nicastro [10] defines the meow as a quasi-periodic sound with at least one formant and with 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 tonal 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 [11] 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'" [7]. 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. [9, 12]), i.e. that each vocalisation strongly correlates with a certain referent and also that perceiver responses correlate with the vocalisation, then ex-

perienced humans should be able to classify meows produced by unfamiliar cats fairly well.

Nicastro & Owren [9] 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 the attention of humans, but that humans can learn to appreciate meows as they become more experienced.

Schötz [13, 14] made a duration and F_0 analysis of 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.

1.3. Purpose, aims and hypotheses

The purpose of this study was to investigate human listeners' perception of domestic cat vocalisations of the same type (the meow), with similar durations, but 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. A larger goal was to learn more about how humans perceive prosodic cues in cat vocalisations and to increase our understanding of cat–human vocal 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 more often correctly 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.

2. Method

2.1. Material

Three young domestic cats: Donna, Rocky and Turbo (D, R and T; 1 female, 2 males, all three year old siblings from the same litter) were recorded in two different contexts: 1) in a familiar environment, i.e. in their home kitchen while waiting to be fed and 2) in an unfamiliar environment, i.e. in the waiting room (or in a car outside) of a veterinary clinic. The equipment consisted of 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 vocalisations segmented, extracted and normalised for amplitude in Praat [16]. 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 in the experiment. 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 slight variations in the background noise, including a few instances of background human speech, but this was judged to have a neglectable influence on the perception task.

Measures of duration and F_0 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 [11], and were judged as clearly distinguishable from other common cat vocalisation types, including the purr (cf. [15]), the murmur (cf. [13]) and the chirp (cf. [14]). The longer meows were often garnished by short initial trills. Table 1 shows the duration, and the mean, minimum, and maximum F_0 values for the twelve meow stimuli. Figure 1 displays F_0 contours of the meows of the two contexts.

Table 1: *Duration (sec.) and F_0 (Hz) values for the 12 meows in two contexts (Food, Vet) by two cats (D, T).*

meow	duration	mean F_0	min F_0	max F_0	F_0 range
Food D 1	0.78	739	528	939	411
Food D 2	0.91	888	541	1003	462
Food D 3	0.27	797	782	816	34
Food T 1	1.06	532	418	582	164
Food T 2	0.85	539	423	653	230
Food T 3	1.03	567	433	640	207
Vet D 1	1.10	790	715	887	172
Vet D 2	0.80	838	764	924	160
Vet D 3	0.58	915	885	947	62
Vet T 1	1.13	510	451	589	138
Vet T 2	0.87	697	639	737	98
Vet T 3	1.02	540	487	570	83

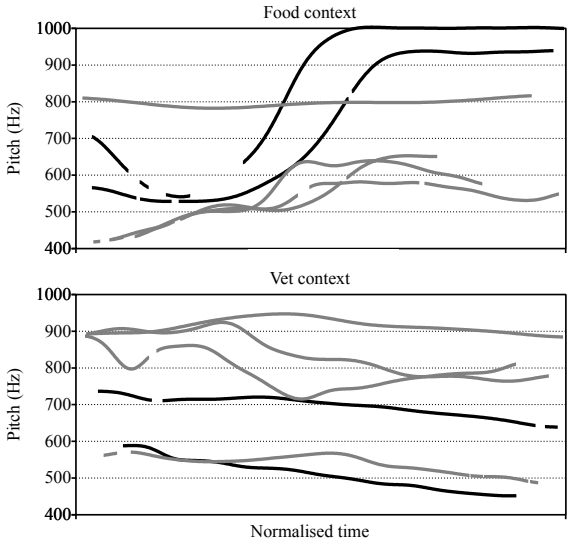

Figure 1: *Time normalised F_0 contours of the food and vet related meows. The two contours of the stimuli that received the highest proportion of correct classifications for each context in the perception experiment are drawn in black.*

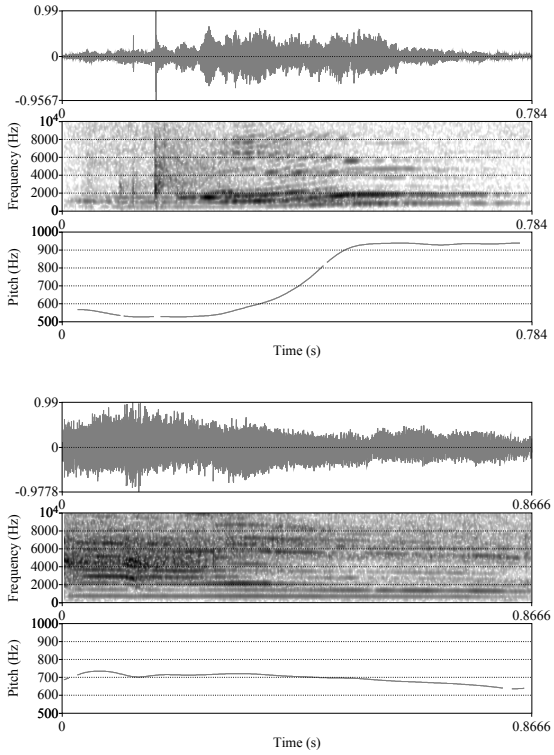

Figure 2: *Waveforms, broadband (300 Hz) spectrograms and F_0 contours of a typical food related meow (top) and a typical vet related meow (bottom).*

2.2. 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 reported 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 an 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.

3. 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.

Table 2: *Percentage 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)
Food D 1	0.83	2342
Food D 2	0.80	2419
Food D 3	0.37	2635
Food T 1	0.54	2944
Food T 2	0.66	2673
Food T 3	0.62	2706
Vet D 1	0.63	3012
Vet D 2	0.57	2904
Vet D 3	0.68	2544
Vet T 1	0.71	2658
Vet T 2	0.71	3127
Vet T 3	0.64	3044

The F_0 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 vet related meows that received the highest number of correct classifications generally display more falling contours.

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 of this analysis 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.

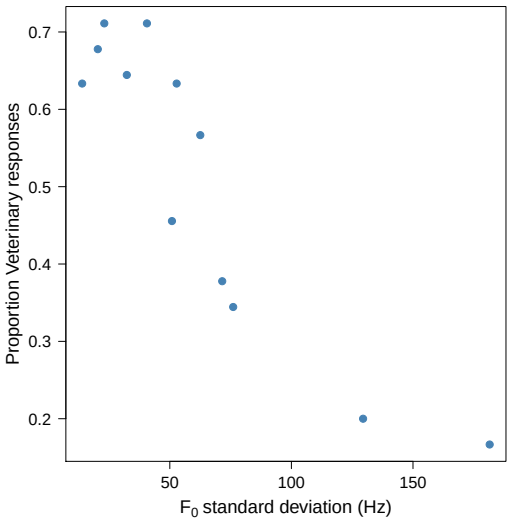

Figure 3: *Relation between F_0 standard deviation and participant choice.*

4. Discussion and future work

The results of the experiment showed that listeners were able to identify domestic cat meows from two different contexts significantly 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 also possible that the listeners were influenced by these differences in F_0 range and interpreted them as expressions of different emotions; food related stimuli as happy with high F_0 range, and vet related stimuli as sad with 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 extremely difficult to classify. Some of the listeners reported that they recognised some of the meows as similar to those of their own cats. This may suggest that different cats produce similar vocalisations in the contexts used in this study. In a future study, we will ask listeners to judge the difficulty of each individual stimulus, and also investigate the phonetic differences between vocalisations that were easy and more difficult to classify.

Several participants reported that they quickly adopted a classification strategy which they used consistently throughout the rest of the experiment even when uncertain of the success rate of this strategy. One strategy would be to listen to the intonational contours of the meows, and judge all rising patterns as belonging to one context, and all falling patterns to the other context. Another possible strategy would be to listen to the vowel quality of the meows. In this study, we did not measure formant frequencies of the vowels included in the stimuli. However, we will examine vowel quality of the cat vocalisations more carefully in future studies, and also systematically study the sound pressure level contours – including the timing of the intensity peaks – of the different meows. It is possible that we will find differences between different types or context meows.

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. [17]). 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 F_0 direction and movement, vowel quality and dynamics (diphthongisation) as well as intensity. We will also examine sounds produced to gain access to desired locations behind an obstacle (cf. [9]), and additional vocalisation types, such as the murmur and the trill, which are common in cat–human communication [13]. We will also try to include cats of a variety of breeds and cats from different countries in order to learn more about the geographical and dialectal variation in cat vocalisations.

5. Acknowledgements

The authors gratefully acknowledge support from the Linnaeus environment Thinking in Time: Cognition, Communication and Learning, financed by the Swedish Research Council, grant no. 349-2007-8695. A warm thanks to the cats Donna, Rocky and Turbo for their patience during the recordings of the material for this study. We are also very grateful to all the participants in our experiment.

6. References

[1] Fitch, W. T., "The phonetic potential of nonhuman vocal tracts: Comparative cineradiographic observations of vocalizing animals", Phonetica, 57, 205—2185, 2000.

[2] Ohala, J., "An ethological perspective on common cross-language utilization of F0 of voice", Phonetica, 41:1–16, 1984.

[3] Håkansson, G. and Westander, J., "Communication in humans and other animals", Amsterdam: John Benjamins, 2013.

[4] Crowell-Davis S. L., Curtis, T. M. and Knowles, R. J., "Social organization in the cat: a modern understanding", Journal of Feline Medicine and Surgery 61:19–28., 2004.

[5] Turner, D. C. and Bateson, P. (eds.), "The domestic cat: the biology of its behaviour", Cambridge: Cambridge University Press, 2000 (2nd edition).

[6] Driscoll, C. A., Clutton-Brock, J., Kitchener, A. C. and O'Brien, S. J., "The taming of the cat", Scientific American, June 2009, 68–75, 2009.

[7] Bradshaw, J., "Cat Sense: The Feline Enigma Revealed", London: Allen Lane, 2013.

[8] Moelk, M., "Vocalizing in the House-Cat; A Phonetic and Functional Study", The American Journal of Psychology 57(2):184–205, 1944.

[9] Nicastro, N. and Owren, M. J.,"Classification of domestic cat (Felis catus) vocalizations by naïve and experienced human listeners", Journal of Comparative Psychology 117:44–52, 2003.

[10] Nicastro, N.,"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, 2004.

[11] McKinley, P. E., "Cluster analysis of the domestic cat's vocal repertoire". Unpublished doctorial dissertation. University of Maryland, College Park, 1982.

[12] Macedonia, J.M. and Evans, C.S.,"Variation among mammalian alarm call systems and the problem of meaning in animal signals", Ethology 93:177–197, 1993.

[13] Schötz, S, "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, 2012.

[14] Schötz, S., "A phonetic pilot study of chirp, chatter, tweet and tweedle in three domestic cats", Proceedings of Fonetik 2013, 14–16 June 2013, Linköping University, Sweden, pp 65–68, 2013.

[15] Schötz, S. and Eklund, R., "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, 8–10 June 2011, 9–12, 2011.

[16] Boersma, P. and Weenink, D., "Doing phonetics by computer" [Computer program], Version 5.3.56, retrieved 15 September 2013 from http://www.praat.org/, 2013.

[17] Burnham, D., Kitamura, C. and Vollmer-Conna, U., "What's new, pussycat? On talking to babies and animals.", Science, 1435–1435, 2002.