A pilot study of human perception of emotions from domestic cat vocalisations

Susanne Schötz
Centre for Languages & Literature, Lund University, Sweden
susanne.schotz@ling.lu.se

Abstract

This paper presents preliminary results from a pilot study where 36 human listeners classified 28 cat vocalisations into seven emotion categories. Classification accuracy and between-listener agreement varied considerably between vocalisations. The vocalisations were subdivided into categories based on the emotions perceived by most listeners and compared in an acoustic analysis. Preliminary results suggest that cats vary their intonation to signal different emotions, and that humans perceive them based on cues used to signal emotion in human speech. Surprisingly, the trill vocalisation used for friendly greetings was often misjudged as anger. Future work includes a deeper analysis of the results and also a comparative study of human–directed and cat–directed cat vocalisations.

Introduction

The cat (Felis catus, Linneaus 1758) was domesticated 10,000 years ago, and is one of our most popular pets 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. For instance, they have learned to produce different vocal signals for different purposes, e.g. solicit feeding or gain access to desired locations and other resources provided by humans. 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 and voice quality 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 [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, snarls, hisses, and shrieks (Moelk, 1944; McKinley, 1982).

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

McComb (2009) found acoustic and perceptual differences between happy and food-soliciting cat purring.

Schötz and Eklund (2011) carried out an acoustic study of cat purring, and Schötz (2012, 2013) analysed 795 different cat vocalisations and found that duration varied only somewhat within each vocalisation type. However, f0 variability was high, partly due to numerous different intonation patterns.

Schötz & van de Weijer (2014) examined 30 human listeners' ability to classify cat meows recorded in the two contexts during feeding time (food related meows) and while waiting at a vet clinic (vet related meows). Classification accuracy was significantly above chance, and listeners with cat experience performed significantly better than naïve listeners. A pitch analysis showed that food related meows tended to have rising f0 contours, while vet related meows often had more falling f0 patterns, suggesting that cats use different intonation patterns in their vocal interaction which (experienced) humans are able to identify.

The purpose of this study was to investigate human listeners' perception of emotions from cat vocalisations and compare a number of acoustic features, including measures of f0 and intonation of the judged emotions, and also to human emotions. A larger goal was to learn more about cat–human communication.

Material and method

Vocalisations from five domestic cats were recorded. Three cats were recorded in their home and two cats in an agonistic context in the author's garden. Video was recorded with either a Sony digital HD video camera HDR-CX730 with an external shotgun microphone Sony ECM-CG50 or an Apple iPhone 3G. Audio files (wav, 44.1 kHz, 16 bit, mono) were extracted with Extract Movie Soundtrack. Based on the overall recording quality and on judgements of how representative the vocalisations were for each emotion on judgements of the author, who knew the cats well, 28 different vocalisations were selected as material. A few vocalisations contained background noise, but this was judged to have no influence on the perception task. The vocalisations were segmented, extracted and normalised for amplitude in Praat (Boersma & Weenink, 2014).

Experiment 1: Perception test

Procedure

36 students (22 women, 13 men) of phonetics and general linguistics at Lund University volunteered as participants in a listening experiment. Their average age was 25 years (range 19 to 59 years). Using a seven-point scale, the participants were asked to rate their experience of and attitude towards cats. Their mean experience with cats (1 = none, 7 = extremely good) was 4.05, and their average attitude towards cats (1 = hate, 7 = love) was 5.5.

Oral and written instructions were given before the experiment. The task was to judge the emotion (using seven categories) of 28 cat vocalisations, which were played twice in the same random order on an Apple MacBook Pro computer through HUMP NF22A speakers at a comfortable sound level. To reduce the number of response categories, some emotions were combined into a single category. Sorrow and fear were combined into the category *SorrowFear*, as some vocalisations were judged to signal both emotions. Moreover, all emotions associated with questioning, begging, wanting or needing something (e.g. food or access to a desired location) were combined into the category *Desire*. Furthermore, the category *Other* could be used for any other perceived emotion, and if the listeners were unable to judge the emotion, they were instructed to select the category *Don't know*. The seven categories used in the test were the following:

1) *Joy*: happy or content,
2) *SorrowFear*: sad or afraid,
3) *Anger*: angry or discontent,
4) *Desire*: questioning, begging, wanting or hungry
5) *Neutral*,
6) *Other*,
7) *Don't know*.

After the test, the participants were asked to make a single judgement of the degree of difficulty of the task on a 7-point scale.

Three experiments with 15, 10 and 10 students participating in each were carried out. Each experiment lasted about 20 minutes. Some time after the experiment, the results were presented to the listeners, and they were asked to comment on them, e.g. what listening strategies and/or phonetic cues they had used to make their judgments. Many listeners reported that they had based their judgements on cues of pitch and whether the vocalisation contained elements of noise. They had judged stimuli with a low average pitch and a high degree of noise as *Anger*, and stimuli with a high average high pitch and a low degree of noise as *Joy*. In addition, listeners reported that rising intonation was judged as *Desire*, and falling intonation as *SorrowFear*. This information was used to select features for the acoustic analysis.

Results

Figure 1 shows the total number of responses for the seven categories. Of all 980 responses in the experiment 53 were *Neutral*, 21 *Other* and 72 *Don't know*. The remaining 834 responses were fairy evenly distributed among the four categories *SorrowFear* (200), *Anger* (211), *Joy* (197), and *Desire* (226).

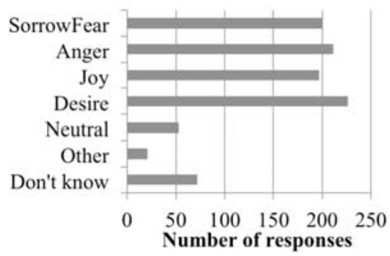

Figure 1. Number of responses for the seven categories in the listening experiment.

Number of correct responses

The emotions of the stimuli were judged (by the author, who had made the recordings and also knew the cats and the contexts in which the vocalisations were produced) and used as preliminary measures of correct emotions for the stimuli. Of the 980 responses in the experiment 350 were correct (38%). Figure 2 shows the percentage of correct responses for each of the 28 stimuli. The two purring stimuli (11 and 20) received the highest number of correct responses, while one low-pitched murmur-meow with clear elements of noise (stimulus 12) received no correct responses, and a low-pitched trill (stimulus 23) only 2 correct responses.

Between-listener agreement

The listener agreement of the responses varied considerably between stimuli. The two purring stimuli showed the highest agreement; over 9/10 of the listeners judged these stimuli as *Joy*. More than half of the listeners perceived the same emotion in nine of the stimuli; four stimuli as *Anger*, two as *Joy*, two as *Desire* and one as *SorrowFear*. Moreover, eight stimuli received the same response from more than 2/5 of the listeners, while five stimuli had over 1/3 of the same listeners responses. Four stimuli had less than 1/3 of the listeners' responses from the same category. Figure 3 shows the distribution of responses from four example stimuli.

Experiment 2: Acoustic analysis

Based on the categories that had received the highest number of responses, the 28 stimuli were subdivided into six emotional categories. Three stimuli had received the highest number of responses for two categories, and these were subdivided into two additional categories: *JoyAnger* and *DesireSorrow*. Measures of duration, f0 and harmonics-to-noise ratio (HNR) were obtained with a Praat script and manually checked. In addition, f0 contours of the stimuli were plotted in six diagrams; i.e. one for each emotion category. Table 1 shows the number of stimuli that was divided into each of the six emotion categories and also the mean results of the acoustic analysis. Figure 4 shows six diagrams with f0 contours of the vocalisations subdivided into each emotion category. The two purring stimuli were categorized as *Joy*, but they were left out of these diagrams, as the f0 in purring is significantly lower than in other cat vocalisations (see Schötz & Eklund, 2011; Schötz, 2012).

Results

Table 1 shows the mean values of duration, f0 (mean, standard deviation, range, minimum, and maximum) as well as mean HNR for six emotion categories containing the 28 stimuli.

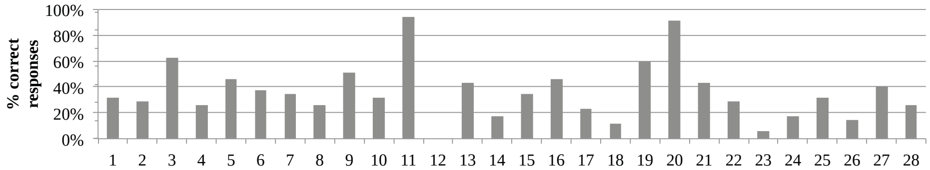

Figure 2. Percentage of correct responses for the 28 stimuli of the listening experiment.

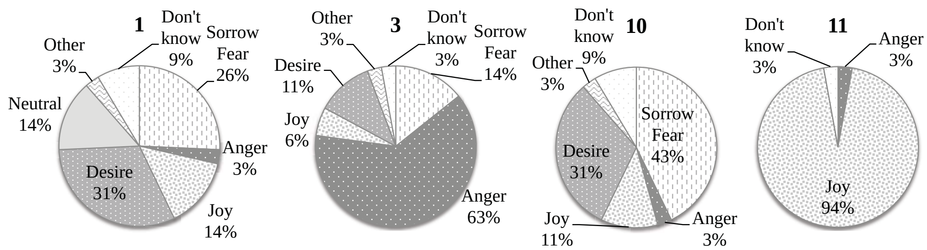

Figure 3. Proportion of listeners' responses (categories) for the example stimuli 1, 3, 10 and 11.

Table 1. Mean duration (sec.), mean f0, f0 standard deviation (stdev), range, minimum and maximum f0 (Hz), and harmonics-to-noise ratio (HNR) of the six emotion categories containing the 28 cat vocalisation stimuli.

Judged category	no stimuli	duration	mean f0	f0 stdev	mean f0 range	min/max f0	HNR
Joy	6	2.18	788	165	501	271/1023	4.1
Anger	9	0.95	415	84	217	211/817	4.8
SorrowFear	5	0.90	694	75	218	300/1102	10.8
Desire	5	0.94	646	108	285	233/886	7.1
JoyAnger	1	0.63	240	21	91	192/283	0.4
DesireSorrow	2	0.66	730	73	244	227/950	10.0
Total (all)	28	1.14	589	96	260	211/1102	6.8

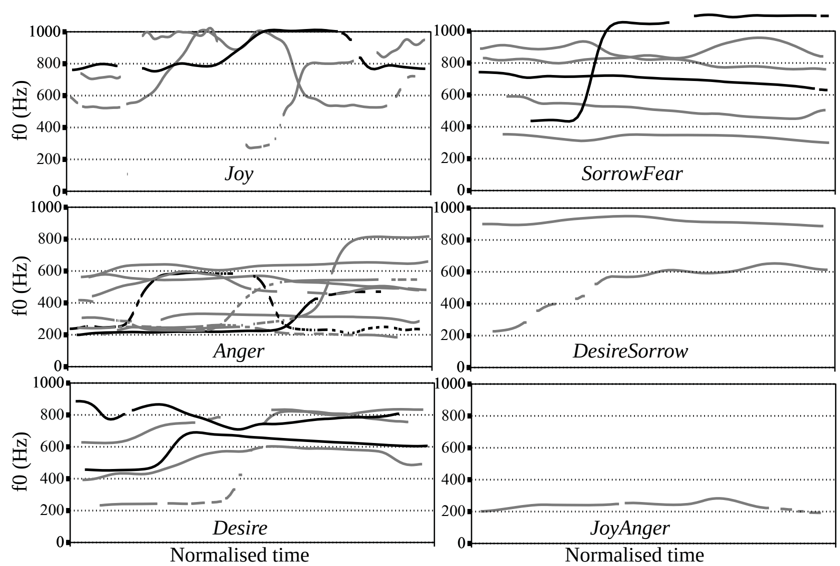

Figure 4. Time normalised f0 contours of cat vocalisations by the categories selected by the majority of the listeners in the perception test.

Duration and HNR

The total average duration was 1.14 sec, and the stimuli judged as *Joy* had the longest duration (2.18 sec.). Clearly shorter durations were found in the stimuli judged as *Anger*, *SorrowFear* and *Desire* (0.90 – 0.95 sec). The categories *JoyAnger* and *DesireSorrow* had the shortest durations (0.63 – 0.66), which may explain why these stimuli had received an equal highest number of responses for two different emotions.

The mean HNR of all stimuli was 6.8, and lower in *Joy* (4.1), *Anger* (4.8) and *JoyAnger* (0.4) than in the other categories. *Desire* had an HNR of 7.1, and the two categories with the highest HNR were *SorrowFear* (10.8) and *DesireSorrow* (10.0), suggesting that the background noise found in these stimuli had not influenced the perception test or the HNR analysis.

f0 values and intonation contours

The vocalisations judged as *Anger* and *JoyAnger* had the lowest mean values for f0, f0 stdev and f0 range. The other categories had clearly higher f0 values, and *Joy* had the highest values of all.

If we exclude the contours with typically initial low f0 followed by a steep rise and ending in very high f0, i.e. the contours of murmur-meows (see Schötz, 2013), the rest of the f0 contours of each category are often similar in shape and range. Moreover, they are not unlike the pitch patterns used by humans to signal the same emotions (see Lindblad, 1992; Rodero, 2011). *SorrowFear* f0 contours are often level and monotonous with a slight fall throughout the vocalisation, which resemble human intonation of sorrow more than of fear. *Joy* has f0 contours characterised by high f0 and a high f0 range with much variation in intonation. *Anger* f0 contours often contain breaks, perhaps due to irregularities or noise, and they are often lower in f0 with either very level intonation or sudden rises and falls, which are in line with the two types of intonation found in human anger.

Discussion and future work

The very preliminary results of this pilot study suggest that human listeners are not very good at judging the emotional state of cat vocalisations, perhaps because they rely on phonetic cues used to signal emotion in human speech.

There was, however, much variation in the agreement between listeners. Some vocalisations, e.g. the purring stimuli (11 and 20) had much higher agreement than others, e.g. the greeting trill stimuli (7 and 15). One explanation may be that the listeners' reported experience of cats varied. About 1/3 of the listeners reported that they were very experienced, while 1/3 had hardly any experience. Another possible explanation is that naïve listeners based their response on biological codes (see Gussenhoven, 2002) and cues for human emotions, as stimuli with high f0 and f0 range were often judged as *Joy*, and stimuli with low f0 and range as *Anger* or *Sorrow*. Naïve listeners may not know what a greeting trill is, and are likely to judge it as *Anger*, as it may sound similar to an agonistic growl. The greeting trill stimuli received about the same number of responses for *Anger* and *Joy*. Moreover, trills are often noisy, and several listeners reported that they had used a high amount of noise as cues for *Anger*.

The results from the acoustic analysis suggest that cats use intonation to signal different emotions. However, although human listeners were fairly good at identifying some emotions, other vocalisations were often misinterpreted. Vocal signals generally cooccur with visual signals, making it easier to distinguish a growl from a trill when it is used to scare off an intruder in one case, but to greet you when you come home from work in the other case. It is likely that humans are much better judges of these types of calls when visual cues are available. Still, vocal signals are important (especially in darkness), and are frequently used by cats in intra as well as in inter-species communication. To be able to communicate better with our cat companions, more phonetic research is needed.

Have cats learned to adapt their vocal patterns (including intonation) to human speech in order to better elicit the desired response from their human companions, or do cats and humans use the same biological codes? The results found by McComb (2009) and the tentative results of this pilot study suggest that cats are able to adapt to human listeners. Future work includes a more thorough analysis of the results found here, and also comparative studies of the phonetic properties of cat-directed and human-directed cat vocalisations.

Acknowledgements

The author is very grateful to all the cat and human participants of this pilot study.

References

Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program] Ver. 5.3.76, retrieved from www.praat.org/.

Bradshaw, J. (2013). *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.

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. In *Proceedings of Speech Prosody 2002*, Aix-en-Provence.

Lindblad, P. (1992). *Rösten*. Studentlitteratur, Lund.

McComb, K., Taylor, A. M., Wilson, C. & Charlton, B. D. (2009). The cry embedded within the purr. *Current Biology* - 14 July 2009 (Vol. 19, Issue 13, pp. R507–R508).

McKinley, P. E. (1982). *Cluster analysis of the domestic cat's vocal repertoire.* Unpublished doctorial 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 listen- ers", *Journal of Comparative Psychology* 117:44–52.

Rodero, E. (2011). Intonation and emotion: influence of pitch levels and contour type on creating emotions. *Journal of Voice*. 25(1):e25–34.

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. In *Proceedings of Fonetik 2011*, Royal Institute of Technology, Stockholm, Sweden, pp 9–12.

Schötz, S. (2012). A phonetic pilot study of vocalisations in three cats. In *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. In *Proceedings of Fonetik 2013*, Linköping University, pp 65–68.

Schötz, S. & van de Weijer, J. (2014). *A Study of Human Perception of Intonation in Domestic Cat Meows*. In *Proceedings of Speech Prosody 2014*, Dublin.