

Assessment of the Emotional State in Domestic Dogs Using a Bi-dimensional Model of Emotions and a Machine Learning Approach for the Analysis of its Vocalizations

Humberto Pérez-Espinosa^{1,2}, Verónica Reyes-Meza³,
María de Lourdes Arteaga-Castañeda³, Ismael Espinosa-Curiel²,
Amando Bautista³, Juan Martínez-Miranda^{1,2}

¹ Mexican National Research Council (CONACyT), Mexico

² CICESE-UT, Tepic, Nayarit, Mexico

³ Universidad Autónoma de Tlaxcala,
Centro Tlaxcala de Biología de la Conducta, Mexico

hperez@cicese.mx, vrmeza@gmail.com

Abstract. Artificial intelligence techniques have been used to classify automatically animal vocalizations. The researchers have found distinct acoustic patterns from the analysis of the relationship between regular and irregular components of the signal. In this work, we analyzed the feasibility of training models for bark classification according to the emotional state of the dog using a model designed to represent human emotions. For this, we induced positive and negative emotions in 10 Schnauzer and 10 Chihuahua dogs reacting to the presence of a stranger and playing with its owner. We recorded and analyzed almost six thousand audio samples of dog vocalizations. Human evaluators assessed the emotional status of dogs. We also found a strong correlation between the evaluators' assessments and the automatic estimation of valence and arousal. Our study underlines the validity of a new technique to classify dog's emotional states using an automatized acoustic analysis of its vocalizations that could be useful to model the relationship between sound properties and emotional content.

Keywords. Automatic barks classification, dog vocalizations analysis, dog's barks classification.

1 Introduction

The emotional perception is a critical trait during the life of social mammals; this trait is fundamental for eliciting approach or provoking rejection towards stimuli [3]. In humans, as in other mammals, emotional states have been studied using behavioral components. However, it is possible to measure emotional valence (positive or negative) from vocalizations [14, 5, 8, 4]. Studies in human voice had

found that it is possible to infer emotions from their affective prosody [20], even some models of speech production applied to animal vocal communication, give information about negative or positive characteristics of vocal production and their acoustic structure. For humans, one of the most common animal vocalization is dog barks which have acoustic features such as frequency, amplitude, pitch, and rhythm. Using these parameters we can recognize barks associated with particular contexts (*e.g.*, the presence of a stranger, stay alone, play with the owner) [10, 28, 29, 22]. In a research work conducted by [23], human listeners rated the emotional state of dogs from bark samples using a 5-item emotional scale (aggressiveness, fearfulness, despair, playfulness, and happiness). In most of the cases, there was a correspondence between the categories assigned and the emotional state of the dog. In a similar study [21], children associated barks to emotional facial expressions. The fact that humans can judge the emotional state of dogs based on the acoustic characteristics of its bark is not surprising since we use similar acoustic characteristics to express anger, fear, happiness, and other emotions [1]. The analysis of the dogs' vocalizations can be automated through computer systems to improve the care for pets and even use them for safety-oriented applications or assistance. Some researchers have applied artificial intelligence techniques for automatic classification of barks. They have trained pattern recognition models to differentiate barks previously recorded in different contexts. In a study conducted with over 6,000 bark samples of Mudi dogs, a Bayesian classifier was trained to solve two tasks, recognition of dogs from its barks and barks classification according to their context [17]. They built a set of acoustic descriptors using an evolutionary algorithm and feature selection techniques. The algorithm classified barks according to the context (recording scenario) with an accuracy of 43%. In the same study, the authors trained a model for identifying individual dogs. In this case, they obtained an accuracy of 52%. The contexts with the best classification accuracy were *the presence of a strange* and *dogs fight*. Using the same database, [16] made a comparison of four supervised machine learning methods for the classification tasks: sex, age, context and identification of the dog. In the case of context classification (alone, ball, fight, food, play, stranger, walk), they tested two learning schemes, a single model for all dogs and a model for every dog. Through these experiments, the researchers have proven the feasibility of training models to recognize and classify different aspects of barks. However, the obtained accuracy seemed not enough to create robust applications. It is necessary to further research in data generation, training and artificial intelligence techniques to achieve better results. In the present study, we describe an experimental procedure to generate a corpus of dog vocalizations and to test if it is possible to develop automatic classifiers of dog vocalizations using a bi-dimensional model designed to classify human emotions [19].

2 Materials and Methods

The first step was to generate a database of dog vocalizations. We recorded audio and video from 23 dogs of two breeds: 11 Schnauzer (seven males and four females), and 12 Chihuahua (five males and seven females). The ages ranged from 5 to 84 months.

2.1 Experimental Procedure

Based on related works [21, 9], we designed a protocol for capturing negative emotions such as fear or anger and positive emotions such as happiness or joy. After several pilot tests, we chose the stimuli related to negative (aggression to unfamiliar people) and positive emotions (playfulness). We recorded the dogs in their habitual environment. Before recording, the experimenter contacted the owners and explained to them the objective of the recording session and the procedure. When the owners allowed the experimenter to leave the camera recording at their homes, we filmed the dogs for a couple of hours with the objective of obtaining the barks spontaneously when a stranger knocked at the door, or when the dog played. Other dogs were recorded only during the induction of stimuli (15-25 min).

Aggression to unfamiliar people We divided this procedure into two parts:

Aggr1: The experimenter activates the recording camera before knocking the door. Usually, the dog starts barking when it realizes the presence of a stranger at the door. After the door is open, the experimenter stimulates aggressive barks by doing some threatening movements in front of the dog while he continues recording until the dog stop barking.

Aggr2: To record the second stimulus, the experimenter moves away from the owner's house for ten minutes and then returns. The experimenter pretends to be a thief, hits the door and tries to force it to open. In this case, the barking is more intense. The recording stopped until the dog becomes quiet.

Playfulness We recorded this context after dogs became familiar with the experimenter. Such as in previous studies [26], the experimenter asked the owner to play with the dog using a toy or its favorite object. Then, the experimenter records the interaction between the owner and the dog. However, not all dogs responded to this stimulus.

2.2 Post-processing of Recordings

The post-processing phase included three steps:

1) **Manual material selection:** the experimenters selected the parts of the audio recordings to include in the study according to their quality. They discarded the parts with excessive ambient noise or low sound volume. We used

extracted the audio from videos. **2) Automatic segmentation:** we implemented a method in Matlab R2014b to extract audio segments from the original recordings. This method takes as input the energy and the spectral centroid thresholds to detect audio activity, generating segments of variable lengths that range from very short individual bark sounds (0.4 seconds) to longer segments (4 seconds) that contain a group of 10 individual bark sounds emitted very quickly without any pause. **3) Manual classification:** the experimenters classified each generated audio segment and stored in a particular directory according to each of the nine classes listed in Table 1.

The experimenters who carried out this classification process knew beforehand what stimuli caused each vocalization. Based on this knowledge, they performed the manual classification process. The only category where they had to use their subjective assessment was to discriminate between aggressive and very aggressive barks. The rest of the categorization tasks were performed objectively. At the end of this process, we obtained 10:43:27 hours of audio-visual material. The corpus consists of 5,805 audio segments: 167 whines, 61 growls and 3,209 barks (1,822 aggressive, 1,036 very aggressive, 192 in the context of play and 159 that were not induced by the protocol).

Table 1 shows the classes of sounds from the analyzed segments:

Table 1. Classes of sounds.

<i>Vocalization</i>	<i>Description</i>
Positive-growl	Growl generated by positive stimuli.
Negative-growl	Growl generated by negative stimuli.
Positive-whine	Whine generated by positive stimuli.
Negative-whine	Whine generated by negative stimuli
Aggressive-bark	Bark occurred due to the presence of a stranger.
Very-Aggressive-bark	Bark occurred due to the presence of a stranger and imminent threat.
Play-bark	Bark occurred while the dog is playing.
Other-bark	Bark that was not produced by some of the planned stimuli.
Other-sound	All other sounds from the background (TV, human voices, bird sounds, etc).

2.3 Human Assessment of Recordings

To assess the emotional content in the vocalizations, we used a continuous scale of emotional vocalizations inspired in the Self-Assessment Manikin (SAM) scales [2]. The SAM is based on a three-dimensional model of human emotions. The dimensions are the arousal, valence, and dominance. Some authors named these dimensions as emotion primitives [27, 11]. They argue that emotions are built on a small number of basic emotion descriptors that are not themselves emotions, but all emotions correspond to combinations of these primitives [6]. We worked with two of these primitives, arousal, and valence which are the most commonly used. We decided to use only two primitives to simplify the model. Arousal describes the perceived vigilance as a physiological and psychological condition.

For example, boredom is supposed to have low arousal, and ecstasy is supposed to have high arousal. Valence describes the positive or negative feeling caused by a situation, an object or an event (*e.g.*, anger is supposed to have a negative valence, and joy is expected to have positive valence). We are interested in finding out these emotion primitives in dog vocalizations.

Participant judges 101 human judges participated in listening sessions. 77 female and 24 male, from 17 to 55 years old, mean age of 21. All the participants had lived with dogs from one to 30 years (mean 11 years). The vocalizations samples were played back on a computer by an assistant. At the beginning of the evaluation session, we distributed the answer sheets to each judge. The arousal and valence scale was explained to the judges, clarifying its use with examples.

Playback stimuli We selected the ten dogs with more samples from each of the two breeds. From the original recordings, we randomly extracted four 10-seconds-long segments for each dog. We excluded samples with environmental noise that could interfere with the assessment of emotion primitives (*e.g.*, knock the door, human voice, and sounds from TV). In total, we used 40 segments of each breed. 50 human judges listened only to the Chihuahua samples, and 51 listened only to the Schnauzer samples (in the same order). After listening to each audio segment, the judges filled a sheet that shows the dog emotional scale. The judges had to assess both valence and arousal for each playback. We performed the playback test in small groups (3 - 6 people). It is important to mention that the judges do not know what stimulus provoked the barking they are hearing.

2.4 Acoustic Characterization of Dog Vocalizations

In any classification problem, we need to identify, define and explain the attributes that will characterize the samples presented to the learning algorithm. In this section, we describe the information used to acoustically characterize all the audio samples (the 10-second-long samples employed in the playback test and the manually classified segments) to process them automatically using machine learning techniques that help us to build classification models. We used the openSMILE software to extract Low-Level Descriptors (LLDs). We extracted the LLDs included in the large openSMILE emotion feature. The LLDs were computed using a frame size of 25 ms and a frame step of 10 ms. A moving average filter for smoothing data contours was applied. To have the same number of attributes for each single vocalization recording, regardless of its duration, it was computed 39 statistical functions over the values of the LLDs and their deltas and double deltas coefficients in each frame of the recording. We obtained a total of 6,552 attributes for each single audio sample. After an experimentation stage, we decided to use the *Relief Attribute* feature selection method as implemented in Weka. This method showed the best accuracy rates when we took the 350 best-ranked attributes. We selected these features from the original set of 6,552

features to obtain the best ones and reduce the dimensionality of the features vector.

2.5 Classification and Validation Methods

We used the acoustic characterization described in the section 2.4 and applied the machine learning technique *Support Vector Machines* (SVM) using a polynomial kernel to classify different dog vocalizations and various types of barks. We selected SVM given that this technique has been successfully used to classify human emotions employing a similar acoustic feature set [18]. Unlike the assessment presented in section 2.3, where we based the analysis on the human evaluations, in this section, it is done using automated tools. This analysis allows an objective assessment of the data. To measure the impact and dependency on individuals in the classification, we validated the models by two methods:

10 Fold Cross Validation (10FCV) In this validation scheme, a classification model is trained using the 90% of the samples in the dataset. We tested it with the 10% left out. This validation round is repeated ten times, each time leaving out a different set of samples. We used this validation scheme to have a baseline accuracy. However, given that we extracted several samples from the same recording, they could generate an effect of pseudo-replication.

Leave One Dog Out Validation (LODOV) We implemented a scheme of cross validation where a classification model is trained using all the samples of N-1 dogs and tested it with the one left out. N is the total number of dogs in the data set, 23 dogs in our case. This validation round is repeated N times, each time leaving out a different dog. This scheme ensures that the trained model is dog-independent. We calculated the accuracy statistics from the accumulated confusion matrix.

3 Results

3.1 Generated Database

This database contains almost eleven hours of audio and video recordings of two breeds Chihuahua and Schnauzer. This database could be an interesting resource from the perspective of the computer science and also from the point of view of the animal behavior research areas. There is currently a trend towards the study of the application of computational systems in the interaction with dogs [13, 30, 24, 12]. In this research context, a bark collection like the generated in this work could be interesting in order to explore the dog-computer interaction through interfaces that take barks as input.

3.2 Agreement of the Participating Judges

It is important to evaluate the agreement among the assessments of the judges because a good agreement means that the emotions primitives expressed in

dog’s vocalization are perceived similarly by humans. Therefore, these emotions primitives are well suited for the modeling of dog’s emotions. We calculated an agreement measurement among judges when they evaluated the playback material. We used the Cronbach’s alpha coefficient [7] that is widely used to describe the reliability of tests and evaluations when there is no missing data (as in our case). We obtained a good agreement among the assessments of the judges. Table 2 shows the absolute agreement among the answers of the judges. We can see that, in general, we obtained a good agreement.

Table 2. Absolute agreement between the answers of participants measured using alpha coefficients for both primitives and both breeds. The fourth column shows the interpretation of the alpha coefficient.

<i>Breed</i>	<i>Dimension</i>	<i>Alpha</i>	<i>Reliability</i>
Chihuahua	Arousal	0.838	Good
	Valence	0.858	Good
Schnauzer	Arousal	0.869	Good
	Valence	0.775	Acceptable

3.3 Statistical Analysis

We performed statistical tests to the data obtained by the subjective assessment of recordings. To complete this analysis, we used the software *IBM SPSS Statistics Version 24*. We applied a Spearman’s rho test to the combination of variables to detect a correlation between them. An interesting research question is how strong is the correlation between valence and arousal. As in the case of human emotions, this two-dimensional representation, on average, present a weak but consistent V-shaped relation of arousal as a function of valence [15]. The more negative or positive emotions usually rated to have higher arousal level too, while more neutral ones are low on arousal too. It is important because these variables are two emotional dimensions that must be weakly correlated to map a variety of emotions. We can see that the Spearman’s rho correlation between these two variables is -0.044. This result indicates that there is a very weak correlation between the perception of both primitives. Therefore, it is possible to use them as emotion descriptors. We used the non-parametric test chi-square to measure the dependency between the nominal variables: emotion primitives and type of vocalization. We found a significant statistical association between the vocalizations of dogs (aggressive bark, very aggressive bark, growl, and whine) and the value of arousal annotated by the participants. We identified that judges perceived growls as very negative and active; and whines as negative and very passive. We also found a significant correlation (0.347) between these same variables. The comparison between vocalization and valence also shown a significant association and a significant correlation (0.431). This result indicates that judges perceived differences in the emotion primitives when they listened

to a different category of dog vocalization (*i. e.*, the perceived emotions changed according to the type of vocalization). Aggressive barks were evaluated mainly as a very negative and active vocalization. Very few judges evaluated aggressive barks as very positive and very passive vocalizations. This fact confirms that the proposed scale makes sense for the assessment of emotions in dog vocalizations.

3.4 Automatic Classification of Dog Vocalizations

The tables 3 and 4 show the performance of the classification experiments in terms of precision, recall and F-measure. The F-measure is metric for classification performance that is calculated as the harmonic mean of precision and recall. The F-measure score reaches its best value at 1 when precision and recall are 1, and worst at 0 when precision or recall is 0 (*i.e.*, the closest to 1, the most accurate is the classification). As we can see in Table 3, we obtained good results classifying the three different vocalizations and contrasting them against to other domestic noises captured during the recordings.

Table 3. Automatic classification of dog vocalizations.

<i>Class</i>	<i>Precision</i>	<i>Recall</i>	<i>F-Measure</i>
Whine	0.899	0.856	0.877
Growl	0.907	0.803	0.852
Bark	0.951	0.950	0.950
Sound	0.933	0.940	0.937
Weighted Avg.	0.942	0.942	0.942

We included the domestic sounds category because it is relevant to build real-world applications based on automatic dog’s vocalization recognition. When the classification system works in realistic environments, it will have to discriminate between dog’s vocalizations and other sounds before classifying types of vocalizations. Table 4 shows the results of classifying only barks according to the emotional state. In this case, we have four classes: aggressive, very aggressive, playfulness and other (barks not generated by the stimuli tasks). The playfulness barks have the best F-measure. We applied two validation schemes to measure the impact of individual dog acoustic properties in the modeling of emotion recognition. We observed a significant decrease in recognition performance when we applied a LODOV, from an average F-measure of 0.79 to 0.49. This fact tells us that each dog has particularities in the way it vocalizes. It implies that recognition models should include a way to adapt to the individual peculiarities to maintain an acceptable accuracy.

3.5 Automatic Assessment of Emotion Primitives in Dog Vocalizations

Finally, in this section, we compare the classification made by the system and the classification made by the human judges. We trained regression models with

Table 4. Automatic classification of dog vocalizations (A) *stratified 10-FCV*, and (B) *leave one speaker out* validation schemes.

<i>Response</i>	<i>Precision</i>		<i>Recall</i>		<i>F-Measure</i>	
	A	B	A	B	A	B
Very aggressive	0.75	0.27	0.62	0.16	0.68	0.20
Aggressive	0.79	0.59	0.87	0.73	0.83	0.65
Playful	1.00	0.75	0.99	0.98	0.99	0.85
Other	0.82	0.27	0.78	0.18	0.80	0.21
Weighted Avg.	0.79	0.48	0.79	0.53	0.79	0.49

the audio samples used in the subjective assessment of emotion primitives. The trained models automatically estimate a continuous value in the same range of emotion primitives than the training samples. In this case from 1 to 5 that are the possible values of arousal and valence given by the judges using the SAMS scale. We used the objective information provided by the acoustic features described in section 2.4 and the algorithm *SMOreg Support Vector Machines for Regression*. We validated by 10FCV. We found that it is possible to train a pattern recognition model to recognize emotions in dog vocalizations with a good estimation performance. For the primitive valence, we obtained a Pearson's correlation coefficient of 0.6341, mean absolute error 0.3031 and root mean squared error 0.3031. For the primitive arousal, we obtained a Pearson's correlation coefficient of 0.7048, mean absolute error 0.3788 and root mean squared error 0.4782. These results are very encouraging and confirm the viability of building automatic systems for dog emotion recognition.

4 Discussion

In the previous sections, we presented a methodology to create a database of emotional dog vocalizations. We used the collected data to validate a method for emotion assessment on dog vocalizations based on subjective assessments by human listeners. We trained pattern recognition models to identify dog vocalizations from other domestic sounds, to classify barks by the context of inductions and to estimate the emotional content of dog vocalizations. The obtained results showed that the primitives arousal and valence are pertinent to assess dog's emotions given that humans correctly perceived emotion information using this bi-dimensional model. These primitives revealed to be independent of between them across all dogs, which allows us to map a variety of emotions. We also identified that automatic emotion recognition in vocalizations is affected by individual acoustic patterns. Inducing and recording barks in dogs is a difficult task. We were able to obtain a lot of data during the aggression to unfamiliar people stimuli but few during the playing sessions.

The presence of the experimenter seems to affect the habitual dog's behavior when the owner tries to play with them. The perceived emotion primitives reflect this situation. The percentage of negative and very negative vocalizations was greater than the proportion of positive and very positive vocalizations. On the

other hand, it is hard to obtain very passive vocalizations given that aggression and playfulness are mainly active and very-active emotions. It is hard to obtain 10-second long segments for all the vocalizations in the database to carry on the emotion evaluation of vocalizations. For this reason, did not include all the types of vocalizations in the human assessment of emotions. The participants in the emotional evaluation of dog vocalizations felt comfortable with the graphical five-points emotion scale that we designed. This instrument was beneficial for our purposes. We found, in general, good agreement between the answers of participants in the emotional evaluation of dog vocalizations.

Arousal had a better agreement than valence in Chihuahua dogs. On the other side, valence had a better agreement than arousal for Schnauzer dogs. We also found that this task is very subjective. The individual reliability of a single participant was poor in all cases. Valence and arousal showed a weak correlation. However, we identified that many barks with active and very active arousal have very negative valence. We identified that judges perceived growls as very negative and active; and whines as negative and very passive. Also, they recognized aggressive barks as very negative and active. Unfortunately, we did not have enough playfulness samples to include in the emotion evaluation study. Finally, we were able to train accurate models to recognize and classify dog vocalizations automatically.

We also trained accurate models to classify the different types of barks. We found that classification models for dog vocalizations could be dog dependent. When we used a scheme validation that leaves out one dog from the training samples, the classification performance decreased drastically. Even though the playfulness class had few samples, it was the most recognizable type of barks. The classes very aggressive and aggressive had a high level of confusion between them. The classifier accurately differentiated the dogs' vocalizations from other domestic sounds. The pattern recognition models that we developed obtained an acceptable accuracy in the estimation of emotion primitives. Similarly to other works [25, 19], we found that estimating valence from acoustic features is harder than estimating arousal. The methodology presented can be a significant contribution to the automatic analysis of dog's barks. The machine learning approach linked with the human assessment of emotions is a novel and interesting methodology that can open the way for new opportunities in the applied field and also can have high commercial interest too.

5 Conclusions

In this work, through a subjective evaluation of recorded vocalizations, we observed that human judges correctly perceived variations of primitive emotional vocalizations of different dogs. We conducted experiments using machine learning techniques with which we found a particular way of expressing emotions vocally by each dog. Finally, we trained a model for acoustic pattern recognition aimed at estimating emotional primitives in dog vocalizations. The performance obtained by this method suggests that it is possible to build computer systems able to

assess the emotional state in dogs based on their vocalizations. Using this technology, developers could design applications oriented to the welfare of dogs, for example, alerting the owners when the dog shows a constant negative emotional state and/or low arousal levels. Other applications could be for domestic security, for example, an app that sends an alert to the owners when they are not at home (or activate a security mechanism), informing that barks with high arousal and negative valence.

Ethical Considerations

Our research is based on non-invasive procedures for evaluating dog behaviour. During recordings we observed that dogs live in optimal health conditions with food and water ad libitum. The owners signed a consent form to voluntarily permit their dogs participate in the present study, and that the resulting media can be used in publications.

Acknowledgements. This work has been carried out in the context of the “Cátedras CONACyT” program funded by the Mexican National Research Council (CONACyT). We thank to Doris Contreras, Fernando Flores, Kenia Rivera, Hector Cruz, Dulce Nolasco and Abraham Galicia for the assistance with the building of the database. This work was partially funded by CONACYT under the Thematic Networks program (Language Technologies Thematic Network project 281795).

References

1. Banse, R., Scherer, K.R.: Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology* 70(3), 614 (1996)
2. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry* 25(1), 49–59 (1994)
3. Briefer, E.: Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology* 288(1), 1–20 (2012)
4. Briefer, E.F., Tettamanti, F., McElligott, A.G.: Emotions in goats: mapping physiological, behavioural and vocal profiles. *Animal Behaviour* 99, 131–143 (2015)
5. Burgdorf, J., Panksepp, J., Moskal, J.R.: Frequency-modulated 50kHz ultrasonic vocalizations: A tool for uncovering the molecular substrates of positive affect. *Neuroscience & Biobehavioral Reviews* 35(9), 1831–1836 (2011)
6. Cowie, R., Cornelius, R.R.: Describing the emotional states that are expressed in speech. *Speech communication* 40(1), 5–32 (2003)
7. Cronbach, L.J.: Coefficient alpha and the internal structure of tests. *psychometrika* 16(3), 297–334 (1951)
8. Faragó, T., Andics, A., Devecseri, V., Kis, A., Gácsi, M., Miklósi, Á.: Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations. *Biology letters* 10(1), 20130926 (2014)

9. Faragó, T., Pongrácz, P., Range, F., Virányi, Z., Miklósi, Á.: The bone is mine: affective and referential aspects of dog growls. *Animal Behaviour* 79(4), 917–925 (2010)
10. Feddersen-Petersen, D.: Vocalization of european wolves (*canis lupus lupus* l.) and various dog breeds (*canis lupus f. fam.*). *Archiv fur Tierzucht* 43(4), 387–398 (2000)
11. Grimm, M., Mower, E., Kroschel, K., Narayanan, S.: Combining categorical and primitives-based emotion recognition. In: *Signal Processing Conference, 2006 14th European*. pp. 1–5. IEEE (2006)
12. Hirskyj-Douglas, I., Read, J.: The ethics of how to work with dogs in animal computer interaction. In: *Proceedings of the Animal Computer Interaction Symposium. Measuring Behaviour* (2016)
13. Hirskyj-Douglas, I., Read, J.C.: Who is really in the center of dog computer design? In: *Proceedings of the 2014 Workshops on Advances in Computer Entertainment Conference*. p. 2. ACM (2014)
14. Knutson, B., Burgdorf, J., Panksepp, J.: Ultrasonic vocalizations as indices of affective states in rats. *Psychological bulletin* 128(6), 961 (2002)
15. Kuppens, P., Tuerlinckx, F., Russell, J.A., Barrett, L.F.: The relation between valence and arousal in subjective experience. *Psychological Bulletin* 139(4), 917 (2013)
16. Larranaga, A., Bielza, C., Pongrácz, P., Faragó, T., Bálint, A., Larranaga, P.: Comparing supervised learning methods for classifying sex, age, context and individual mudi dogs from barking. *Animal cognition* 18(2), 405–421 (2015)
17. Molnár, C., Kaplan, F., Roy, P., Pachet, F., Pongrácz, P., Dóka, A., Miklósi, Á.: Classification of dog barks: a machine learning approach. *Animal Cognition* 11(3), 389–400 (2008)
18. Pérez-Espinosa, H., Pérez-Martinez, J.M., Durán-Reynoso, J.Á., Reyes-Meza, V.: Automatic classification of context in induced barking. *Research in Computing Science* 100, 63–74 (2015)
19. Pérez-Espinosa, H., Reyes-García, C.A., Villaseñor-Pineda, L.: Acoustic feature selection and classification of emotions in speech using a 3d continuous emotion model. *Biomedical Signal Processing and Control* 7(1), 79–87 (2012)
20. Pollermann, B.Z., Archinard, M.: Acoustic patterns of emotions. *Improvements in speech synthesis* p. 237 (2002)
21. Pongrácz, P., Molnár, C., Dóka, A., Miklósi, Á.: Do children understand man’s best friend? classification of dog barks by pre-adolescents and adults. *Applied animal behaviour science* 135(1), 95–102 (2011)
22. Pongrácz, P., Molnár, C., Miklósi, Á.: Barking in family dogs: an ethological approach. *The Veterinary Journal* 183(2), 141–147 (2010)
23. Pongrácz, P., Molnár, C., Miklósi, A., Csányi, V.: Human listeners are able to classify dog (*canis familiaris*) barks recorded in different situations. *Journal of Comparative Psychology* 119(2), 136 (2005)
24. Pons, P., Jaen, J., Catala, A.: Envisioning future playful interactive environments for animals. In: *More Playful User Interfaces*, pp. 121–150. Springer (2015)
25. Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., Wendemuth, A.: Acoustic emotion recognition: A benchmark comparison of performances. In: *Automatic Speech Recognition & Understanding, 2009. ASRU 2009. IEEE Workshop on*. pp. 552–557. IEEE (2009)
26. Svartberg, K., Forkman, B.: Personality traits in the domestic dog (*canis familiaris*). *Applied animal behaviour science* 79(2), 133–155 (2002)
27. Wu, D., Parsons, T.D., Narayanan, S.S.: Acoustic feature analysis in speech emotion primitives estimation. In: *INTERSPEECH*. pp. 785–788 (2010)

28. Yin, S.: A new perspective on barking in dogs (*canis familiaris*). *Journal of Comparative Psychology* 116(2), 189 (2002)
29. Yin, S., McCowan, B.: Barking in domestic dogs: context specificity and individual identification. *Animal Behaviour* 68(2), 343–355 (2004)
30. Zeagler, C., Gilliland, S., Freil, L., Starner, T., Jackson, M.: Going to the dogs: towards an interactive touchscreen interface for working dogs. In: *Proceedings of the 27th annual ACM symposium on User interface software and technology*. pp. 497–507. ACM (2014)