Thesis Booklet

ROBOT BEHAVIOUR DESCRIPTION AND ARTIFICIAL EMOTION EXPRESSION IN ETHOROBOTICS

Beáta Korcsok

Supervisor
Péter Korondi, DSc

Professor

Budapest University of Technology and Economics
Faculty of Mechanical Engineering

Department of Mechatronics, Optics and Mechanical Engineering Informatics

Budapest, 2023

1. Introduction and State of the Art

The research of social robotics, especially the field of HRI (Human-Robot Interaction, the investigation of interaction and communication between robots and humans), is a relatively new, interdisciplinary research area, which is expected to advance considerably in the next decades.

Social robots, especially service and assistive robots, are expected to become an everyday part of our future e.g., in the field of hospitality [1] or as assistive elderly care robots [2]. With the growing importance of social robots and other artificial agents (including software-based, virtually displayed agents), the development of adequate communication in Human-Robot and Human-Computer Interaction (HRI and HCI) is becoming imperative. A common approach in developing the communicational signals of social robots and other artificial agents is to base them on human communication [3] e.g., on speech [4] and human-specific gestures [5]. Human-like communication seems to be a natural way of interaction for social robots, as human languages can convey high complexity in sharing information [6], and e.g., facial gestures can express a wide variety of affective states [7]. However, this approach is frequently undermined by technological limitations relating to the perceptive, cognitive, and motion skills implemented in the agent, which can become more obvious during the course of interaction, leading to disappointment [8], [9].

Overt similarity can also cause aversion towards human-like robots, as predicted by Mori's Uncanny Valley hypothesis [10], [11]. Mori argued that close, but not 'perfect' resemblance to humans does not increase likeability, rather it leads to even stronger aversion [11]. Since then, studies were conducted to investigate this hypothesis, both with morphed pictures (faces: [12], [13], videos of real robots [14] or actual interactions with them [15]. Although the results are mixed and the debate is not over, the issue is worth considering during the design of experiments and the development of social robots. Interestingly, a similar effect can be found in animals (long-tailed macaques, *Macaca fascicularis*; [16]), as the behaviour of the macaques towards generated monkey faces (that are realistic, but do not reach the realism of actual faces) indicate the effect of the Uncanny Valley. However, there is also some data suggesting that such aversion toward artificial agents may be modified during development through socialisation in childhood, and that this issue may be circumvented by a different robot building strategy [10].

Furthermore, the proposed functions (tasks or roles) of specific robots do not always require the level of complexity found in human communication [8], or their capabilities and functions are not in line with that of humans (e.g., no need for head-turning with 360° vision [10], no morphological limitations in sound production). To avoid these issues, another approach can be considered.

Ethorobotics − created from the word '*ethology'*, (the field of biology committed to the analysis of human and non-human animal behaviour) and '*robotics'* − considers HRI as interspecific interaction in which the artificial agent is regarded as a separate species, and only has to be equipped with a basic level of social competence and communicational skills that are aligned with its function [10]. Social competence in the case of social robots should follow a bottom-up approach, in which robots do not have to possess the same or similar levels of social skills as humans do, rather a minimum set of skills with enough commonalities to that of humans' [10].

The importance of the topic was further increased by a recently started and ongoing phenomenon, according to which, in parallel to the continued rise of classic industrial robots, a paradigm shift is taking place in the industry, predicting an increased industrial interest in social and service robots alongside the grand-scale industrial applications [17]. While the 3rd and 4th Industrial revolution brought forth the popularity of automatization and cyber-physical systems, the 5th Industrial revolution puts humans, and the interaction of humans and technology in its centre. In contrast with classical industrial robots, that function in environments separated from humans, collaborative robots already require the development of adequate interactive capabilities in industrial production. This is a more pressing issue in the case of social robots, that instead of the robot-optimised environments of factories, have to function in dynamically changing human environments and have to interact, or even cooperate with humans as part of their basic functions, requiring the implementation of various social skills and interactive competences [18]. **With the appearance of Industry 5.0, the robots equipped with social skills can appear in the factories** [19].

Alongside social robotics, new technologies such as Information and Communication Technologies (ICT) have been on the rise. Artificial agents used in ICT devices could be viewed by users as somewhat living beings similarly to how children regard social robots differently than strictly inanimate objects [20]. A biological approach in the design of these agents' embodiment and behaviour could enhance the interaction between humans and the artificial agents. Emotion expression is an important aspect, which could be used as a means of communication for alerts and notifications (e.g., low battery) but it could also be used to make the device or robot more acceptable to humans and easier to interact with [21].

2. Establishing the research topic and the scientific challenge

The present dissertation – similarly to a significant portion of contemporary scientific pursuits – is the result of interdisciplinary research, in which the perspectives of engineering and biology appear simultaneously. The dissertation encompasses the development of nonverbal communicational signals for robots, and a method for the objective description of robot behaviour in the field of social robotics.

Based on research activities and the devices entering the market, the fastest developing branch of robotics is social robotics. However, this is the area that is surrounded with the most misconceptions. The target audience of social robots are people who do not possess robotics expertise (with whom the robots have to communicate and cooperate with), who frequently overestimate the capabilities of robots, while others are also unaware of their current physical and computational skills. We also cannot predict the time when robots become a natural part of our homes, and some might even start considering these robots as family members, similarly to family pets.

Strictly speaking, robots have already appeared in our homes, as while the simple cleaning robots do not yet take notice of a person entering the room, the natural demand can easily arise for it to ask us if it should continue working, or if it should pause while we are in the area. From this, it is only a step forward for us to require added social skills, for example for the robot to greet us as we enter.

The development of physical capabilities and the aligned interactive competences is parallel in social robots. The advancement of physical capabilities is clearly the enterprise of engineers, while the investigation of the central aspects of social interactions, their understanding and definition goes beyond engineering competences, but it is not independent of it. The advancement of social robotics developments and the spread of the technology in everyday life may be facilitated if we investigate the social skills and communicational possibilities necessitated by the function of the robot already during the design phase. Ethorobotics provide a scientifically grounded approach with a strong biological basis and a unified concept, complementing the engineering expertise.

3. The objective of the dissertation

During the writing of the dissertation, the subject area changed so much that the objectives of the dissertation had to be fundamentally reformulated, but I did not experience this as a difficulty, but rather as a strengthening of the importance of the dissertation.

A challenge of interdisciplinary research stems from the issue that the cooperating research areas often use different terminologies in differing interpretations of the same phenomena, or use the same terminology in different manners. Behaviour, as discussed in an engineering viewpoint in engineering systems, is drastically distant from the interpretation of behaviour by ethologists in biological systems to an extent, which can even hinder cooperation between engineers and ethologists. Furthermore, the scientific approach of engineers is more characterised by synthesis, while ethologists are more analysis-oriented, which results in the fundamentally different description of behaviour in these two areas. The replicability of experiments is important in science. If we describe the behaviour of an animal, it is generally sufficient to focus on a few behavioural elements depending on the current scientific question, since in a repeated experiment the animal is likely to also perform those behavioural elements that we did not record in our description. In contrast, when describing the behaviour of the robot every detail must be determined, as the robot does not do anything by itself. **One of the goals of the dissertation is to present a novel method developed specifically for the description of robot behaviour, that can be accepted not only by engineers but also by ethologists, and which is suitable for describing the behaviour of different robot platforms in such a way that the experiments on these robot behaviours can be replicated and compared by others.**

An important conceptual message of the thesis is that the robot should not imply more than what it is actually capable of. Therefore, we aim to avoid the extensive use of verbal communication until robots can meet the expectations set by humans towards their cognitive understanding capabilities. At the beginning of our research we focused on non-verbal communication in order to circumvent the rudimentary verbal communication available for robotics at the time. A significant change in verbal communication technologies took place in the last phase of writing this thesis with **the development of Chat GPT and subsequent programs providing artificial intelligence services based on LLM (Large Language Model) methods**. An increasingly wide range of verbal communication skills utilisable in social robotics is based on Chat GPT; for example, our most important industrial partner in the field of social robotics, Enjoy Robotics, has also converted to use Chat GPT as the basis of the verbal communication of its robots.

Notwithstanding, the driving force behind my research that I present in the thesis is the claim that verbal communication in itself is generally not sufficient. It is known that during verbal communication between people, information of fundamental importance passes through metacommunicative channels. Non-verbal communication covers all behaviours of communicative importance, all of which is not the textual content of the speech itself [22], for example the spatial position or locomotive behaviour (proxemics, e.g., to how close and at what speed one person approaches another), hand gestures, facial expressions, touch, body posture, gaze behaviours, non-verbal signals independent of, or accompanying speech (e.g., laughter, sighs, prosody of speech, timbre), turn-taking, etc. [23]. That is why we often insist on personal discussions instead of written ones. Although social robots can produce the generated verbal content with the help of speech synthesis software, the realisation of human-robot interaction in everyday life will also require other components of communication, which can be expressed by robots, i.e., artificial agents with a "body", a physical presence.

Special fields of application can also require effective and faster-than-verbal communication in e.g., noisy environments, which also prioritizes non-verbal communication. A potential example is the collaborative work of robots and humans as part of the Industry 5.0 environment, in which humans could react to visual signals or short sound effects faster than the time required to say words. In human non-verbal communication the visual and acoustic modalities are of great importance, which modalities are also fundamental in developing the communicative skills of robots. Accordingly, my thesis also focuses on the investigation of visual and acoustic signals.

The other goal of the thesis is to explore new communicative prospects that can aid social interactions between robots and humans. As the main goal is to serve robotics, the topic is more aligned with the doctoral school of engineering than with the doctoral school of biology. The achievements of other scientific fields (ethology, sociology, psychology) must be made accessible and applicable to the experts working in the field of engineering. **The dissertation follows this approach: it investigates general rules observed in biological systems, that can be implemented in abstract forms in engineering systems.**

4. The applied research methodology

In the field of engineering, the formulation of a hypothesis is via the writing of an equation, and proving the hypothesis is achieved by establishing the mathematical truth of the equation, or by conducting measurements. The present thesis utilises the second approach, although the measurements differ from the conventional measurements that are more frequently used in engineering. Measurements are always afflicted by noise, so ensuring the reliability of the measurement results is extremely important for every measurement. There are accepted mathematical methods to achieve this, and we did so in all cases. The questionnaire measurement methodology used during the experiments is also commonplace in research related to comfort theory, which is one of the important areas of the doctoral school.

The usability of the new robot behaviour description method was validated by applying it to the description of the behavioural elements of a robot in an experiment specifically dealing with human-robot social interactions.

Engineering work is mostly based on synthesis, while ethology on analysis, therefore the methodology of scientific process in these two fields differs substantially. It is important to emphasize that in scientific analysis, the literature research is part of the scientific work, and not just an introductory chapter, as it usually appears in dissertations dealing with synthesis. In this dissertation, the literature research related to the theses is always an important part of the thesis itself.

This difference in approach justifies the fact that, after stating the theses, I briefly address the methodological differences of these two scientific fields in the formulation of my own scientific contribution.

5. The structure of the dissertation

In the dissertation, following the introductory chapter, I focused on the methodological aspects of ethorobotics and created the concept of a robot ethogram (a catalogue describing the robot's behavioural elements) that meets both engineering and ethological considerations (Chapter 2).

In Chapter 3 and 4 I delve into the topic of artificial emotion expressions. A long-term goal of ethorobotics is to develop multimodal communicational signals for artificial agents based on general biological rules, but first individual modalities also have to be explored. In Chapter 3 I introduced a study in which I investigated the recognition of visual emotionally expressive behaviours via an abstract virtual agent. The visual emotion expression displays were based on the emotionally expressive behaviours of a wide range of animal taxa. Similarly, the artificially generated emotionally expressive sounds presented in Chapter 4 were also created following the acoustic coding rules found in animal vocalizations. Finally, in Chapter 5 I present some ethical issues and conclusory ideas.

6. Theses

Thesis 1.

Methodological problem statement

The elements that are most directly evident to a person who interacts with the robot – following its embodiment – are the actions, the *behaviour* of the robot. All HRI research on some level has to deal with these behavioural elements that consist the continuous behaviour of the robot, as the basis of interaction between the robot and the human. Ethologists have developed the *ethogram*, which serves to catalogue and describe behavioural elements to aid observational and experimental investigations [24]. In the field of social robotics, engineers and experts from other areas investigate a wide variety of social skills and communicational signals with the use of diverse robot platforms. The cataloguing and detailed description of these behavioural elements increases the comparability of HRI research results, and facilitates the objective investigation of behaviours during robot development and experiments. In my work connected to Thesis 1 my goal was to adapt the method of ethograms in a form that is acceptable for engineers.

Formulation of Thesis 1.

The description of the behavioural elements of social robots that takes both ethological and engineering scientific criteria into account is possible with the help of ethograms that fulfil the following requirements:

a) **The position and form have to be described in relation to the angles of articulations, the directions of movements, the necessary degrees of freedom,**

and the intended kinematic chain of the motion (Denavit-Hartenberg kinematic description).

b) **The proxemic parameters have to be specified via the spatial positions and orientations.**
c) **The robot orientation has to be defined according to the function of the behavioural (e.g., orienting towards human face).**
d) **The intensity of the behavioural element is to be defined via the used quantities of relevant parameters (speed, acceleration, duration, exerted force, sound pressure level of vocalizations, etc.).**
e) **Variability of the behavioural element parameters are to be specified via the ranges of these variations.**
f) **The behavioural elements have to be defined as overlapping or mutually exclusive.**

These conditions ensure that the described behavioural elements are comparable and replicable.

Related publications: [P1], [P2], [P7], [P8], [P10]

My own scientific contribution

- I introduced the usability of the ethogram in robotics in the publication [P1], in which I was co-first author with Dr. Judit Abdai. The content of the article reviews the methodological difficulties of two separate topics, Human-Robot Interaction (HRI) and Animal-Robot Interaction (ARI). The ethogram description created by me in the HRI section of the article is introduced here for the first time.
- My further work supporting the thesis, in which I elaborate in detail the steps, background and advantages of using the ethogram as a method in social robotics, is presented in the following article: Korcsok, B., Korondi, P. How do you do the things that you do? - Ethological approach to the description of robot behavior, Biologia Futura (accepted, In press.).
- I corroborate the applicability of the ethogram method by presenting its use in a HRI experiment. Beyond the development of the ethogram for the experiment, I also contributed to the research by creating the supporting structure and modular cover of

the robot and its tray, by developing the experimental protocol, and by conducting the experiments. The main programmer of the robot was Ferdinandy Bence, who also participated in the testing.

Application of Thesis 1.

Ethograms are created by observing the continuous behaviour of the studied species and defining discrete behavioural elements that comprise the behaviour repertoire of the animal, based on their form and sometimes their presumed function. These elements can be described on different levels, (e.g., relating to only a muscle, a body part or the whole body), and can contain different modalities (e.g., vocalizations, see the ethogram of bottlenose dolphins (*Tursiops truncatus*) [25]). Ethograms usually contain a detailed description of position and form (e.g., limb raised at a certain angle [26]), spatial position (e.g., behavioural element occurring near conspecifics [27]), orientation (e.g., head orientates towards food), and sometimes even intensity (e.g., loudness of vocalization, speed of locomotion) for the specific behavioural elements.

In the dissertation after the overview of descriptive methods used in ethology and robotics, I am presenting an ethogram format that can be easily applied to differing modalities and robot platforms, and that is based on the ethograms utilised for various animal species.

I validated the method's usability via its application in a research study. I have cooperated with the researchers of the Department of Ethology at ELTE to conduct an experiment, in which a social robot served food to participants as a waiter assistant robot in a mock-up restaurant (Figure 1). Note: one of the most promising areas of application for social robotics is waiter assistant robots. At the time of writing this thesis, an ordinary person could mostly encounter social robots in restaurants or hotels.

The goal of the study was to provide an objective analysis on how certain communicational signals (orientation with gaze, vocalizations, etc.) affect humans' interactions with the robot and their opinions and attitudes regarding robots. The two experimental groups of the study were the 'interactive' and 'minimally interactive' groups, which differed in the communicational signals used by the robot. In the minimally interactive group the robot could be sent away from the table by the guests via an ultrasound sensor, but the robot did not perform other interactive behaviours. In the interactive group the robot interacted with humans via multiple communicative signals and behavioural elements (greeting, real-time generated artificial sounds based on rules found in animal vocalizations [28], orienting at the faces of humans and performing gaze alternation between them, looking down at its tray and the use of attention getting sounds, as well as a farewell sound, sending the robot away via an ultrasound sensor).

Figure 1. The social robot leads the participants to their table.

As part of the creation of the ethogram, I have defined these communicative behavioural elements, and all other behavioural elements exhibited by the robot during the experiment, in a table format. The descriptions of behavioural elements include: the name and unique code of the behavioural element; the experimental group in which the behavioural element occurs; the behaviour complex that the behavioural element belongs to (locomotion, orientation, vocalization, interaction); the verbal description of the behavioural element; the direction of movement for those behaviours for which it is applicable; the variability of the behavioural element (e.g., the randomness of fundamental frequency that occurs during vocalization); the description of form and position with the help of codes used in ethology; the autonomy of the element; the process that activates the behavioural element (e.g., is it activated by an environmental stimulus or the experimenter); whether the behaviour takes place without feedback (Fixed Action Pattern); the possible overlap of the behavioural element and its simultaneous appearance with other elements; its kinematic description with the Denavit-Hartenberg parameters used in the field of engineering; and finally the illustration of the behavioural element using pictures and videos. In total I have described 19 behavioural elements in the ethogram (6 in the minimally interactive group and 13 in the interactive group). For an example, see Figure 2. The full ethogram is available online[1].

Behaviour	Behaviour code	Robot interactivity	Behaviour complex	Verbal description	Direction of movement	Variability	Presentation
Leaving table	**ILLT**	Interactive	Locomotion	The robot turns away from the table with its body	Towards the base checkpoint	Variability in head movement is the result of the goal location. Head turning range is normally 107° (between 237°-130°), extreme ranges can occur (between 248°-119°); tilt is constant (235°). Head movement speed is set to 28.6 deg/s (0.5 rad/s). Speed of turning away from the table with the body is set to 2.86 deg/s (0.05 rad/s).	

Trunk (body) orientation	Head orientation	Autonomy	Occurrence / activation	Fixed Action Pattern	Exclusory / overlapping	Presentation (video)
R	F --> F [O]	Autonomous	Controlled by navigation	No, path autonomously generated by the navigation stack, object avoidance can occur	Can overlap with head orientation behaviour (**IHOL**)	https://doi.org/10.5281/zenodo.8286061

DoF	Kinematic description of joints θ1	θ2
3	237°-130°	235°

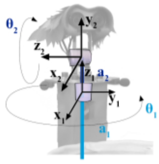

Figure 2. A section from the ethogram. The figure shows the description of a behavioural element, leaving table (***ILLT***).

Thesis 2.

Main research question

Given an abstract agent with a dynamic (changeable) geometric formation (Figure 3). If we move and change this abstract geometry in accordance with the emotionally expressive behaviours common to biological organisms and readily observed across different species of animals (we increase its size, rotate it, change its colour, the speed of its movement, etc.), will humans be able to recognize the emotional expressions based on biological rules that we intend to display via the artificial agent?

[1] https://doi.org/10.5281/zenodo.8349912

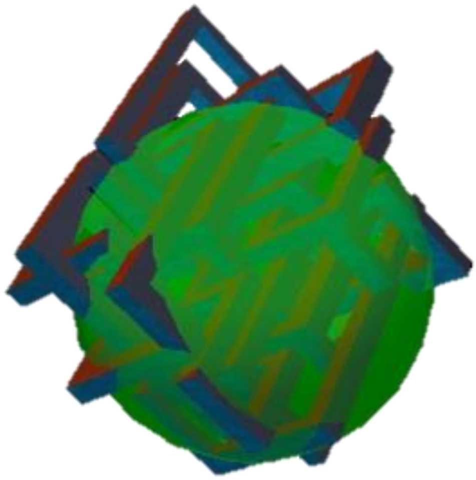

Figure 3. The abstract virtual agent used for the investigation of artificial emotion expressions. The agent is comprised of a sphere and a labyrinth-like, angular part, “grid”. I have used this previously created agent in my work. The agent was created by György Persa and Péter Baranyi [29], [30], and the emotionally expressive displays were implemented by Márta Gácsi.

Hypothesis:

Participants recognize the generalized simple expressive behaviours observed in animal species as displayed by an abstract artificial agent, and give the highest score to the respective emotions in multiple choice questions.

Formulation of Thesis 2.

It is provable with statistical analysis conducted with the previously created Emotion Display Agent, that if a dynamically changing complex abstract artificial agent is equipped with movement and colour dynamic properties based on the behaviour of animal species that indicate their internal emotional states, then humans are able to associate basic emotions with the artificial agent according to the following rules:

- **Neutral state: The agent retains its original size and position. Rounded and jagged parts are both visible in the outline of the agent. The angular subunit moves at a constant medium speed. The colour of the agent is green, the colour of the angular subunit changes randomly.**

- **Surprise: The normal movement of the angular subunit of the agent stops, the size of the agent increases slightly. The angular subunit turns 90° in the opposite direction of its normal movement, then the movement stops again. The angular subunit becomes slightly brighter.**
- **Fear: The agent moves towards the edge of the movement area. The movement speed of the angular subunit increases. The overall size of the agent and the size of the angular subunit decreases, the outline of the agent becomes fully rounded. The colour of the agent becomes pale and bright.**
- **Anger: The size of the agent and the angular subunit increases. The outline of the agent becomes more jagged. The movement speed of the angular subunit increases. The agent becomes red and the angular subunit takes on colours that contrast with the colour of the agent.**
- **Happiness: The agent moves towards the bottom of the movement area, its size increases, the outline of the agent becomes fully rounded. With rapid movement, the agent moves slightly from the bottom of the movement area toward the centre, and then returns to the bottom of the movement area. The agent turns orange, the colour of the angular subunit becomes similar the colour of the agent.**
- **Sadness: The agent moves slowly towards the side of the movement space. The movement speed of the angular subunit decreases significantly. The overall size of the agent and the size of the angular subunit decreases, the outline of the agent becomes fully rounded. The colour of the agent fades and becomes mostly transparent.**

Related publications: [P3], [P4], [P6], [P9]

Note: an important part of the scientific work has been the literature research of animal behaviour patterns, therefore the list of articles supporting the thesis can be considered part of the thesis.

My own scientific contribution

The thesis is based on analysis, and it declares **scientifically established statements after the scientifically required statistical analysis** of the provided stimulus material questionnaire data, **and after extensive literature research**. Some of the statements could be considered true intuitively, however, in a scientific work with an analytical approach scientific proof of all statements is required. The provided material has been heuristic from the point of scientific analysis, that is, its effect on humans was not scientifically verified before. The result of my work, based on literature research and mathematical statistics, is the verification that the characteristics of the emotionally expressive animations – that are based on general biological rules and the connected behavioural characteristics observable in the animal world – provide emotion expressions recognisable by humans. I made a decisive contribution to the research on which the thesis is based on by thoroughly analysing and interpreting the collected data using descriptive and mathematical statistical methods, by establishing and exploring the scientific literary background of the previously created animations and by writing the manuscript.

Scientific importance of Thesis 2.

The artificial visual agent investigated in the thesis can already be used for the expression of emotions in its given form, but the thesis also provides guidance on the principles with which new artificial emotion expressions can be created. The results of the thesis inspired the start of a study at the University of Debrecen, in which the emotional expression of a geometrical agent – with a significantly different appearance from the artificial visual agent used in the thesis – is investigated.

In my thesis work, a multi-questionnaire study was carried out on the assessment of the displays with both Hungarian and Japanese participants. In the first, open-ended questionnaire the participants could freely express their opinion on what mood, feeling they considered the agent displaying the animation had, while in the second questionnaire after viewing each display again, they scored all animations in terms of all the emotions on a 5-point Likert scale. I have conducted both descriptive and statistical analyses (statistical analysis: 1. comparison of the success of recognition in each display to chance level with binomial tests, 2. Friedman tests and Wilcoxon signed rank tests, with Benjamini-Hochberg corrections, 3. Ward's hierarchical cluster analysis, 4. Generalized Linear Mixed Models analysis).

The results show that the recognition of the displayed emotions was significantly higher than chance level in all emotion expression displays. In most cases the subjects could successfully recognize the emotions that the agent's display was meant to convey: the scores given to the correct emotions were significantly higher than the scores given to all the other emotions in four displays (Fear, Surprise, Sadness and Anger). In the Happiness display the scores of happy and angry were not significantly different. The Hungarian participants recognized the Anger display correctly based on both the descriptive and statistical analysis, and the certainty of their answers was the highest in this display. Participants mixed emotion displays more easily which had similar behavioural attributes relating to their intensity (happiness - anger) or valence (sadness - fear). When we look at cultural differences, we can see that in most cases the participants from the two cultures recognized the displays similarly, but differences were also found. The Anger and Surprise displays received similar scores in the two groups, but Japanese participants were more successful in the recognition of Sadness, while also being less successful in the recognition of Happiness than Hungarians. Although the differences show the need for cultural fine tuning, the majority of displays received the highest scores for the correct emotions, indicating that the general rules used for creating the emotion expressions can be considered a good basis for emotion expression in artificial agents.

Thesis 3.

Main research question

Humans are able to recognise animal vocalizations as emotionally expressive sounds, and rate their valence and intensity based on acoustic parameters such as the fundamental frequency or call length [31]–[33]. These connections are consistent with the existence of simple coding rules of affective vocalizations that are a commonly found, shared characteristic of terrestrial tetrapods [34], [35]. The research question was whether the simple coding rules connected to the fundamental frequency and call length of vocalizations can be applied to artificially generated sounds?

Hypothesis:

The simple coding rules observed in animal vocalizations apply to artificial sounds as well, the direction of the effects of the acoustic parameters on the emotion ratings are the same on the distinct biological complexity levels. Human listeners will rate artificial sounds with higher fundamental frequency as more intense and sounds with longer call lengths as having more negative valence.

Formation of Thesis 3.

It is provable on a statistical basis that the simple acoustic coding rules observed in emotionally expressive animal vocalizations can be abstracted to non-biological sounds as well, as humans are able to evaluate the intensity and valence of artificially generated sounds according to these rules (the perceived intensity increases with the fundamental frequency, and the emotional valence becomes more negative with the increase of call length).

Related publication: [P5]

My own scientific contribution

My contribution was fundamental throughout the research in the development of the main concept, the generation of the artificial sounds, the data analysis and the writing of the manuscript.

Scientific importance of Thesis 3.

The generated artificial sounds can be used already in their current form, but the thesis also provides guidance with which new artificial sounds can be created. The strength of the thesis is shown by the fact that it inspired several further studies: the generated sounds were investigated using fMRI methods by measuring the brain activity of dogs and humans, as well as in a questionnaire study in which we examined the approach-avoidance reactions elicited by the sounds. The sounds were also used as the vocalizations of artificial agents: in the HRI experiment related to Thesis 1, the real-time generated sounds of the robot were created based on the parameters and results of this study and Thesis 3, and the sounds are also being used in dog-robot interaction research as the vocalizations of small remote controlled artificial agents.

Following the previously outlined concepts, the artificial sounds were created based on general acoustic features of emotionally expressive vocalizations of humans and non-human animals. The specific ranges of the various acoustic parameters were mostly based on dog vocalizations, as these have been previously studied with a similar methodology [31]. I have generated the sound samples in seven categories, on three complexity levels using the Praat 6.0.19 software. The simplest sound category used sine-wave sounds, as these are singlefrequency sounds that rarely occur naturally, but which are frequently used in artificial signals of machines. Then, starting with the simple sine-wave sounds, new acoustic features (pitch contour changes, harmonics, variations of call properties within a sound sample, formants) were added that are characteristic of animal vocalizations, to make more complex and biologically more congruent samples (Figure 4). In each category, I systematically changed the fundamental frequency and call length of the sounds to cover the relevant acoustic ranges of these parameters. The total number of created stimuli consisted of 588 sound samples, 84 in each category.

An online questionnaire was used to measure the perceived emotional valence and intensity of the sounds in a two-dimensional model of emotions, with a total number of 237 human participants. The questionnaire used a modified version of the two-dimensional model of emotions by Russell [36]. The questionnaire measured the values the participants gave for the sounds on the valence and intensity axes. After a sound was played, the participants had to indicate the valence on a horizontal axis and the intensity on the vertical one with one click.

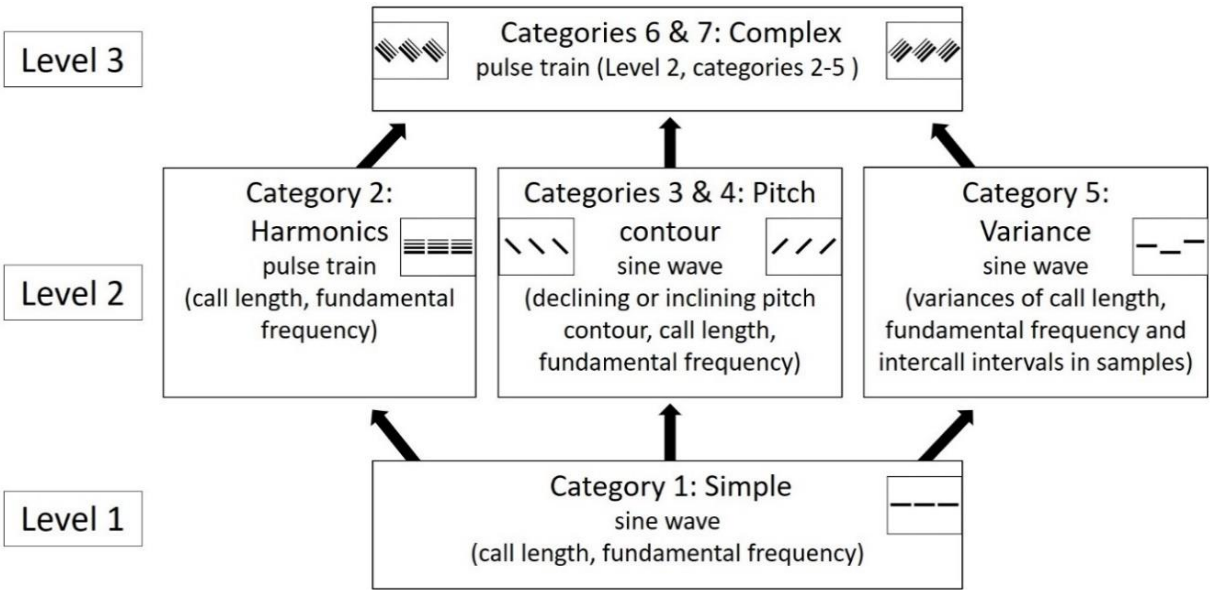

Figure 4. The categories of the artificial sounds across three levels of complexity

Statistical analysis was conducted in the R statistical environment, with Linear Mixed Modeling with Tukey post-hoc tests. The models of category 1 were also used to predict the valence and intensity ratings of the other categories: comparisons between the predicted and actual valence and intensity ratings were conducted with Pearson’s correlation.

The results show that the artificially generated sounds are able to mimic some of the basic coding rules that are present in animal (mammalian) vocalizations. The fundamental frequency of the sounds affects the perceived intensity, that is, sounds with higher fundamental frequency were perceived as more intense, while sounds containing longer calls were rated as more negative across all categories. Additionally, fundamental frequency also affected valence, as sounds with higher fundamental frequency were rated as more negative.

As the basic coding rules related to fundamental frequency and call length were also present in the artificial sounds with no added biological features, we can infer that these effects might originate from a more fundamental component of sound processing.

Publications

[P1] Abdai, J., Korcsok, B., Korondi, P., & Miklósi, Á. (2018). Methodological challenges of the use of robots in ethological research. *Animal Behavior and Cognition*, *5*(4), 326–340. https://doi.org/10.26451/abc.05.04.02.2018

[P2] Faragó, T., Miklósi, Á., Korcsok, B., Száraz, J., & Gácsi, M. (2014). Social behaviours in dog-owner interactions can serve as a model for designing social robots. *Interaction StudiesInteraction Studies Social Behaviour and Communication in Biological and Artificial Systems*, *15*(2), 143–172. https://doi.org/10.1075/is.15.2.01far

[P3] Korcsok, B., & Korondi, P. (2023). Érzelemkifejezések háttere és szerepe a szociális robotikában. *Recent Innovations in Mechatronics*, Accepted, In Press.

[P4] Konok, V., Korcsok, B., Miklósi, Á., & Gácsi, M. (2018). Should we love robots? – The most liked qualities of companion dogs and how they can be implemented in social robots. *Computers in Human Behavior*, *80*, 132–142. https://doi.org/10.1016/j.chb.2017.11.002

[P5] Korcsok, B., Faragó, T., Ferdinandy, B., Miklósi, Á., Korondi, P., & Gácsi, M. (2020). Artificial sounds following biological rules: A novel approach for non-verbal communication in HRI. *Scientific Reports*, *10*(1), 7080. https://doi.org/10.1038/s41598-020-63504-8

[P6] Korcsok, B., Konok, V., Persa, G., Faragó, T., Niitsuma, M., Miklósi, Á., Korondi, P., Baranyi, P., & Gácsi, M. (2018). Biologically inspired emotional expressions for artificial agents. *Frontiers in Psychology*, *9*(JUL), 1–17. https://doi.org/10.3389/fpsyg.2018.01191

[P7] Korcsok, B., Korondi, P., & Miklósi, Á. (2014). Az etológia szerepe a szociális robotok fejlesztésében. *ARES Konferencia Kiadvány*, 66–70.

[P8] Korondi, P., Korcsok, B., Kovács, S., & Niitsuma, M. (2015). Etho-robotics: What kind of behaviour can we learn from the animals? *IFAC-PapersOnLine*, *48*(19), 244–255. https://doi.org/10.1016/j.ifacol.2015.12.041

[P9] Nagy, B., Korcsok, B., Nánási, J., & Niitsuma, M. (2019). EthoCar, Human-Dog Interaction Inspired Human-Autonomous Car Communication. *CAX TECHNOLOGIES JOURNAL*, *7*, 5–12.

[P10] Korcsok, B., & Korondi, P. (2023). How do you do the things that you do? - Ethological approach to the description of robot behaviour. *Biologia Futura*, Accepted, In Press.

References

[1] J. Bowen and C. Morosan, "Beware hospitality industry: the robots are coming," *Worldw. Hosp. Tour. Themes*, vol. 10, no. 6, pp. 726–733, Dec. 2018, doi: 10.1108/WHATT-07-2018-0045.

[2] J. Abdi, A. Al-Hindawi, T. Ng, and M. P. Vizcaychipi, "Scoping review on the use of socially assistive robot technology in elderly care," *BMJ Open*, vol. 8, no. 2, p. e018815, Feb. 2018, doi: 10.1136/bmjopen-2017-018815.

[3] N. Mavridis, "A review of verbal and non-verbal human-robot interactive communication," *Rob. Auton. Syst.*, vol. 63, no. P1, pp. 22–35, 2015, doi: 10.1016/j.robot.2014.09.031.

[4] M. Bennewitz, F. Faber, D. Joho, and S. Behnke, "Fritz - A humanoid communication robot," *Proc. - IEEE Int. Work. Robot Hum. Interact. Commun.*, pp. 1072–1077, 2007, doi: 10.1109/ROMAN.2007.4415240.

[5] R. Meena, K. Jokinen, and G. Wilcock, "Integration of gestures and speech in human-robot interaction," *3rd IEEE Int. Conf. Cogn. Infocommunications, CogInfoCom 2012 - Proc.*, pp. 673–678, 2012, doi: 10.1109/CogInfoCom.2012.6421936.

[6] F. Pellegrino, C. Coupé, and E. Marsico, "Across-Language Perspective on Speech Information Rate," *Lang. (Baltim).*, vol. 87, 2012.

[7] P. Ekman and W. Friesen, *Unmasking the face: A guide to recognizing emotions from facial clues*.

MALOR BOOKS, 2003. [Online]. Available: https://www.amazon.fr/Unmasking-Face-recognizing-emotions-expressions/dp/1883536367

[8] Á. Miklósi and M. Gácsi, "On the Utilization of Social Animals as a Model for Social Robotics," *Front. Psychol.*, vol. 3, no. MAR, pp. 1–10, 2012, doi: 10.3389/fpsyg.2012.00075.

[9] R. Rose, M. Scheutz, and P. Schermerhorn, "Towards a conceptual and methodological framework for determining robot believability," *Interact. Stud. Stud. Soc. Behav. Commun. Biol. Artif. Syst.*, vol. 11, no. 2, pp. 314–335, 2010, doi: 10.1075/is.11.2.21ros.

[10] Á. Miklósi, P. Korondi, V. Matellán, and M. Gácsi, "Ethorobotics: A new approach to human-robot relationship," *Front. Psychol.*, vol. 8, no. JUN, pp. 1–8, 2017, doi: 10.3389/fpsyg.2017.00958.

[11] M. Mori, "The Uncanny Valley," *Energy*, vol. 7, 1970, doi: 10.1038/nn.2647.

[12] K. F. MacDorman, R. D. Green, C. C. Ho, and C. T. Koch, "Too real for comfort? Uncanny responses to computer generated faces," *Comput. Human Behav.*, vol. 25, no. 3, pp. 695–710, 2009, doi: 10.1016/j.chb.2008.12.026.

[13] J. Seyama and R. S. Nagayama, "The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces," *Presence Teleoperators Virtual Environ.*, vol. 16, no. 4, pp. 337–351, Aug. 2007, doi: 10.1162/pres.16.4.337.

[14] K. F. MacDorman, "Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley," *ICCS/CogSci-2006 long Symp. Towar. …*, no. January 2006, pp. 26–29, 2006, doi: 10.1093/scan/nsr025.

[15] C. Bartneck, T. Kanda, H. Ishiguro, and N. Hagita, "My robotic doppelgänger - A critical look at the Uncanny Valley," *Proc. - IEEE Int. Work. Robot Hum. Interact. Commun.*, vol. 31, no. 0, pp. 269–276, 2009, doi: 10.1109/ROMAN.2009.5326351.

[16] S. A. Steckenfinger and A. A. Ghazanfar, "Monkey visual behavior falls into the uncanny valley," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 43, pp. 18362–18366, 2009, doi: 10.1073/pnas.0910063106.

[17] D. Belanche, L. V. Casaló, and C. Flavián, "Frontline robots in tourism and hospitality: service enhancement or cost reduction?," *Electron. Mark.*, vol. 31, no. 3, pp. 477–492, Sep. 2021, doi: 10.1007/s12525-020-00432-5.

[18] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Rob. Auton. Syst.*, vol. 42, no. 3–4, pp. 143–166, 2003, doi: 10.1016/S0921-8890(02)00372-X.

[19] R. Gervasi, K. Aliev, L. Mastrogiacomo, and F. Franceschini, "User Experience and Physiological Response in Human-Robot Collaboration: A Preliminary Investigation," *J. Intell. Robot. Syst.*, vol. 106, no. 2, p. 36, Oct. 2022, doi: 10.1007/s10846-022-01744-8.

[20] P. H. Kahn, B. Friedman, D. R. Pérez-Granados, and N. G. Freier, "Robotic pets in the lives of preschool children," *Interact. Stud. Soc. Behav. Commun. Biol. Artif. Syst.*, vol. 7, no. 3, pp. 405–436, Nov. 2006, doi: 10.1075/is.7.3.13kah.

[21] E. Hudlicka, "To feel or not to feel: The role of affect in human-computer interaction," *Int. J. Hum. Comput. Stud.*, vol. 59, no. 1–2, pp. 1–32, 2003, doi: 10.1016/S1071-5819(03)00047-8.

[22] J. A. Hall, T. G. Horgan, and N. A. Murphy, "Nonverbal Communication," *Annu. Rev. Psychol.*, vol. 70, no. 1, pp. 271–294, Jan. 2019, doi: 10.1146/annurev-psych-010418-103145.

[23] M. L. Knapp, J. A. Hall, and T. G. Horgan, *Nonverbal communication in human interaction*. Cengage

Learning, 2013.

[24] N. Tinbergen, “Comparative Studies of the Behaviour of Gulls (Laridae): a Progress Report,” *Behaviour*, vol. 15, no. 1–2, pp. 1–69, 1960, doi: 10.1163/156853960X00098.

[25] R. Hickey, S. Berrow, and J. Goold, “TOWARDS A BOTTLENOSE DOLPHIN WHISTLE ETHOGRAM FROM THE SHANNON ESTUARY, IRELAND,” *Biol. Environ. Proc. R. Irish Acad.*, vol. 109, no. 2, pp. 89–94, Jan. 2009, doi: 10.3318/BIOE.2009.109.2.89.

[26] W. M. Schleidt, G. Yakalis, M. Donnelly, and J. McGarry, “A Proposal for a Standard Ethogram, Exemplified by an Ethogram of the Bluebreasted Quail (Coturnix chinensis),” *Z. Tierpsychol.*, vol. 64, no. 3–4, pp. 193–220, 1984, doi: 10.1111/j.1439-0310.1984.tb00360.x.

[27] N. Greenberg, “An ethogram of the blue spiny lizard, Sceloporus cyanogenys (Reptilia, Lacertilia, Iguanidae),” *J. Herpetol.*, pp. 177–195, 1977.

[28] B. Korcsok, T. Faragó, B. Ferdinandy, Á. Miklósi, P. Korondi, and M. Gácsi, “Artificial sounds following biological rules: A novel approach for non-verbal communication in HRI,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–13, Dec. 2020, doi: 10.1038/s41598-020-63504-8.

[29] G. Persa, A. Csapo, and P. Baranyi, “A pilot application for ethology-based CogInfoCom systems,” in *2012 IEEE 10th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, Jan. 2012, pp. 477–482. doi: 10.1109/SAMI.2012.6209014.

[30] P. Baranyi, G. Persa, and Á. Csapó, “Definition of cognitive infocommunications and an architectural implementation of cognitive infocommunications systems,” *World Acad. Sci. Eng. Technol.*, vol. 5, no. 58, pp. 501–505, 2011.

[31] T. Faragó, A. Andics, V. Devecseri, A. Kis, M. Gácsi, and Á. Miklósi, “Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations,” *Biol. Lett.*, vol. 10, no. 1, 2014, doi: 10.1098/rsbl.2013.0926.

[32] I. L. Maruščáková, P. Linhart, V. F. Ratcliffe, C. Tallet, D. Reby, and M. Špinka, “Humans (Homo sapiens) judge the emotional content of piglet (Sus scrofa domestica) calls based on simple acoustic parameters, not personality, empathy, nor attitude toward animals.,” *J. Comp. Psychol.*, vol. 129, no. 2, pp. 121–131, May 2015, doi: 10.1037/a0038870.

[33] C. Tallet, M. Špinka, I. L. Maruščáková, and P. Šimeček, “Human Perception of Vocalizations of Domestic Piglets and Modulation by Experience With Domestic Pigs (Sus scrofa),” *J. Comp. Psychol.*, vol. 124, no. 1, pp. 81–91, 2010, doi: 10.1037/a0017354.

[34] P. Filippi *et al.*, “Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: Evidence for acoustic universals,” *Proc. R. Soc. B Biol. Sci.*, vol. 284, no. 1859, pp. 1–9, 2017, doi: 10.1098/rspb.2017.0990.

[35] A. Andics and T. Faragó, “Voice Perception Across Species,” in *The Oxford Handbook of Voice Perception*, S. Frühholz and P. Belin, Eds. Oxford University Press, 2018, pp. 362–392. doi: 10.1093/oxfordhb/9780198743187.013.16.

[36] J. A. Russell, “A circumplex model of affect.,” *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980, doi: 10.1037/h0077714.