ELTE | FACULTY OF INFORMATICS

ETHOROBOTICS

Beáta Korcsok

DEPARTMENT OF ETHOLOGY
EÖTVÖS LORÁND UNIVERSITY

ELTE | IK
DEPARTMENT OF ARTIFICIAL INTELLIGENCE

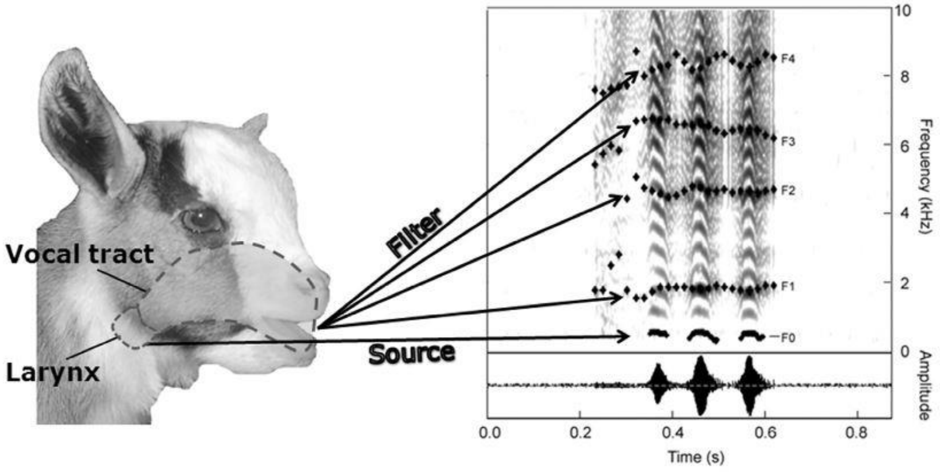
BOSCH

HUN-REN
Hungarian Research Network

Acoustic emotion expression

!

- Humans can assess the emotionally expressive sounds of other species, e.g.
  - Dogs (Pongrácz et al., 2006; Faragó et al., 2014)
  - Pigs (Maruščáková et al., 2015)
- Dog vocalizations (Faragó et al., 2014):
  - High frequency, long call length → fearful
  - Low frequency, long call length → aggressive
  - Short call length → positive inner states
- Similar voice production (source-filter theory, Fant, 1960; Taylor et al., 2010)



Briefer, 2012

Acoustic emotion expression

- Simple coding rules
  - Fundamental frequency & call length → acoustic cues for the listeners → inner state of the vocalizing individual

- Artificially generated sounds → vocal features from animal vocalizations

- Systematic approach → increasing complexity
- Assesment of perceived intensity and valence by humans

Questions and hypotheses

- Are the simple coding rules of fundamental frequency and call length of vocalizations also apply to artificially generated sounds?
- Similarly to animal vocalizations, human listeners perceive artificial sounds:
  - Higher fundamental frequency → more intense
  - Longer calls → more negative valence
- Does the complexity of the artificial sounds influences their effect?
- Same effect in simple sine wave sounds as in more complex sounds

Acoustic emotion expression

- Online questionnaire
- Modified version of Russell's circumplex model (1980)
- Online questionnaire adapted from Faragó et al. (2014)
- N = 237, 95 Hungarian speaking (60 ♀ 35 ♂, age = 36.3 ± SD 11.8 years) and 142 English speaking (122 ♀ 20 ♂, age = 39.9 ± SD 11.7 years)
- 70 sounds/participant
- Sound stimuli:
- Total 588 sounds in 7 categories
- Artificially generated sounds (Praat)

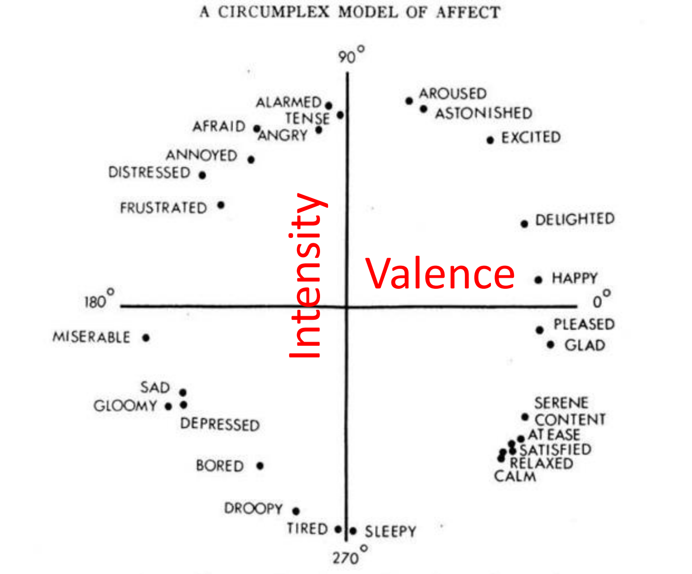
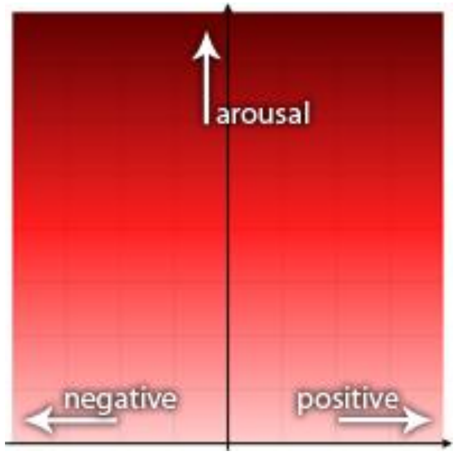
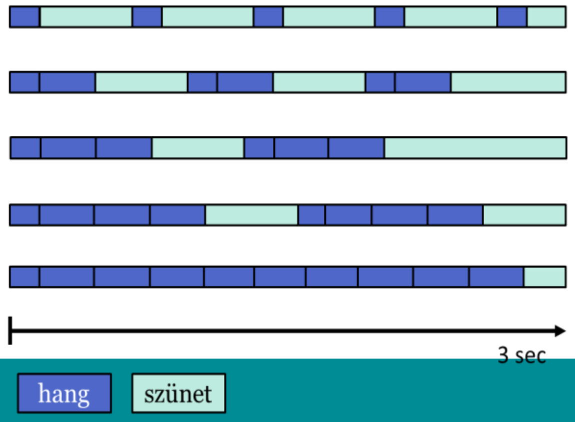


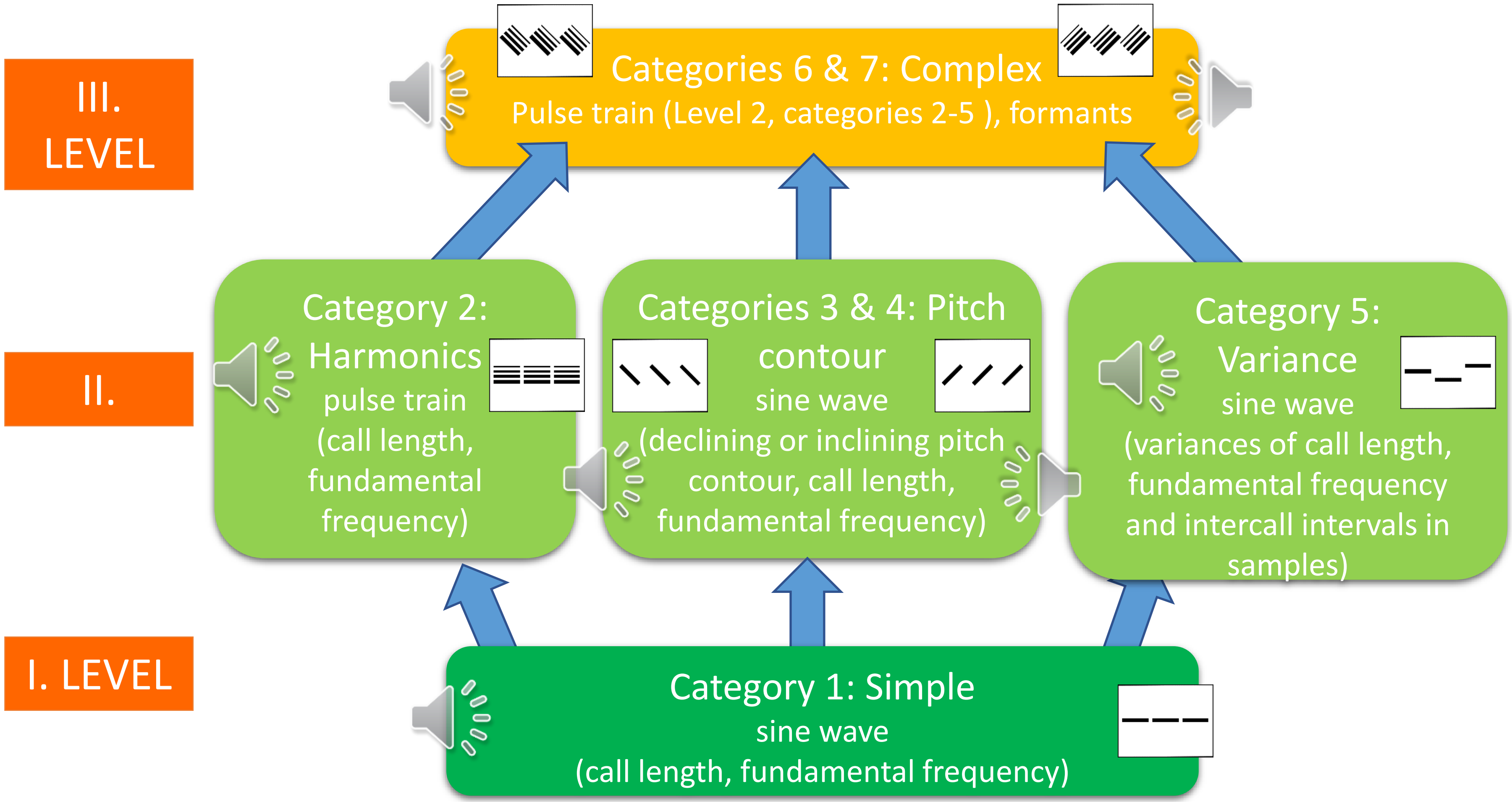
Figure 2. Direct circular scaling coordinates for 28 affect words.





Korcsok et al., 2020

Acoustic emotion expression



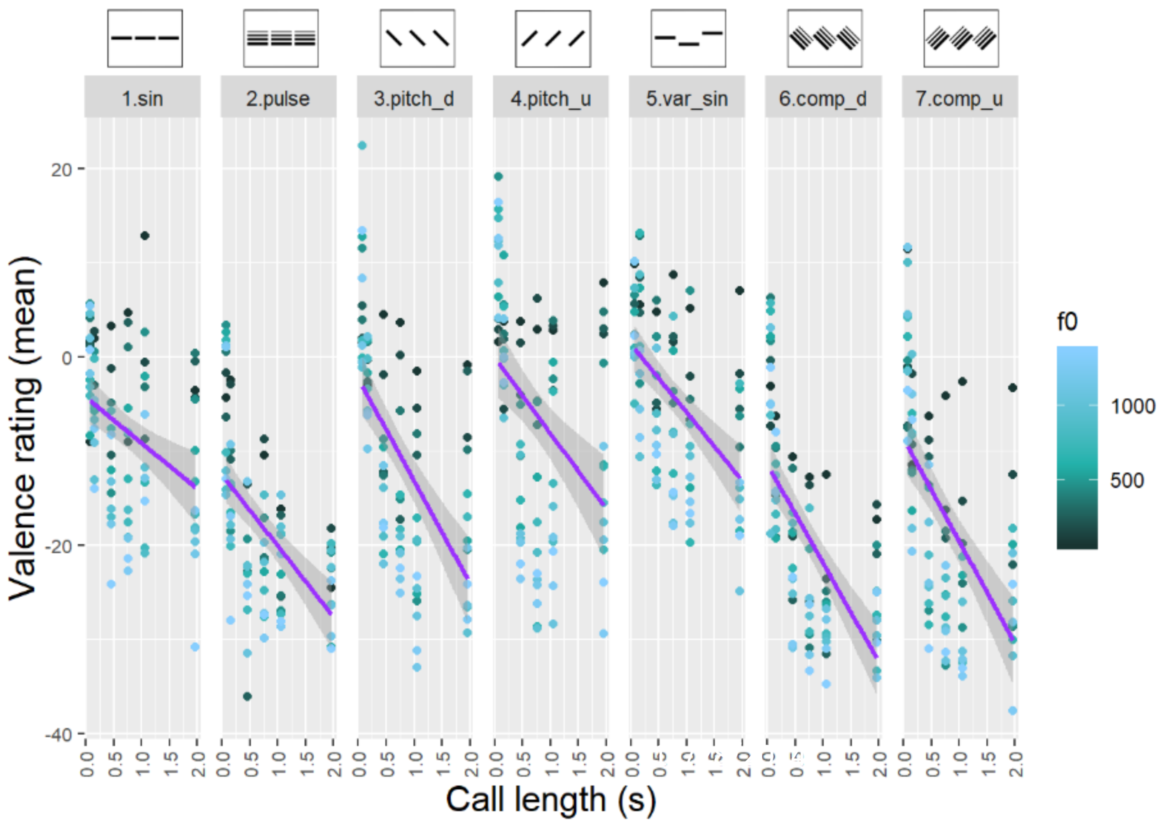
Acoustic emotion expression

Parameters	Value or Range (across all samples)	Variance in categories 1, 2, 3, 4	Variance in categories 5, 6, 7
Fundamental frequency (f_0)	65Hz - 1365 Hz		uniformly distributed random value, ±5% of f_0
Total length (call length + interval length)	~2 s (+ silence until 3 s total duration)		
Call length	0.07; 0.16; 0.46; 0.76; 1.06; 1.96 s		uniformly distributed random value, ±25% of call length
Intercall interval length	0.2 s	uniformly distributed random value, ±25% of interval length	uniformly distributed random value, ±50% of interval length
Pitch contour change in categories 3, 4, 6, 7	uniformly distributed random value, ±10% of f_0		
Vocal tract length in categories 6, 7	20 cm		
Number of formants in categories 6, 7	10		
First formant (f_1) in categories 6, 7	550 Hz		

Valence

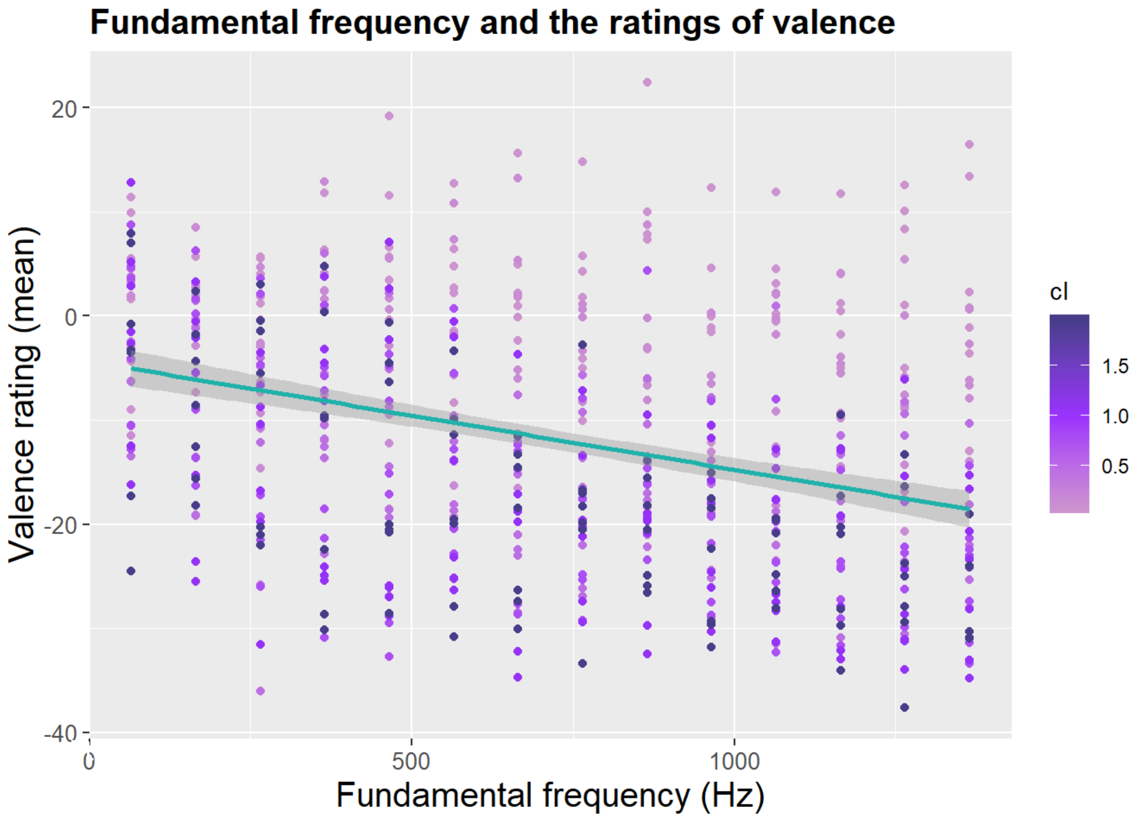
Call lenght

Longer call → more negative



Fundamental frequency

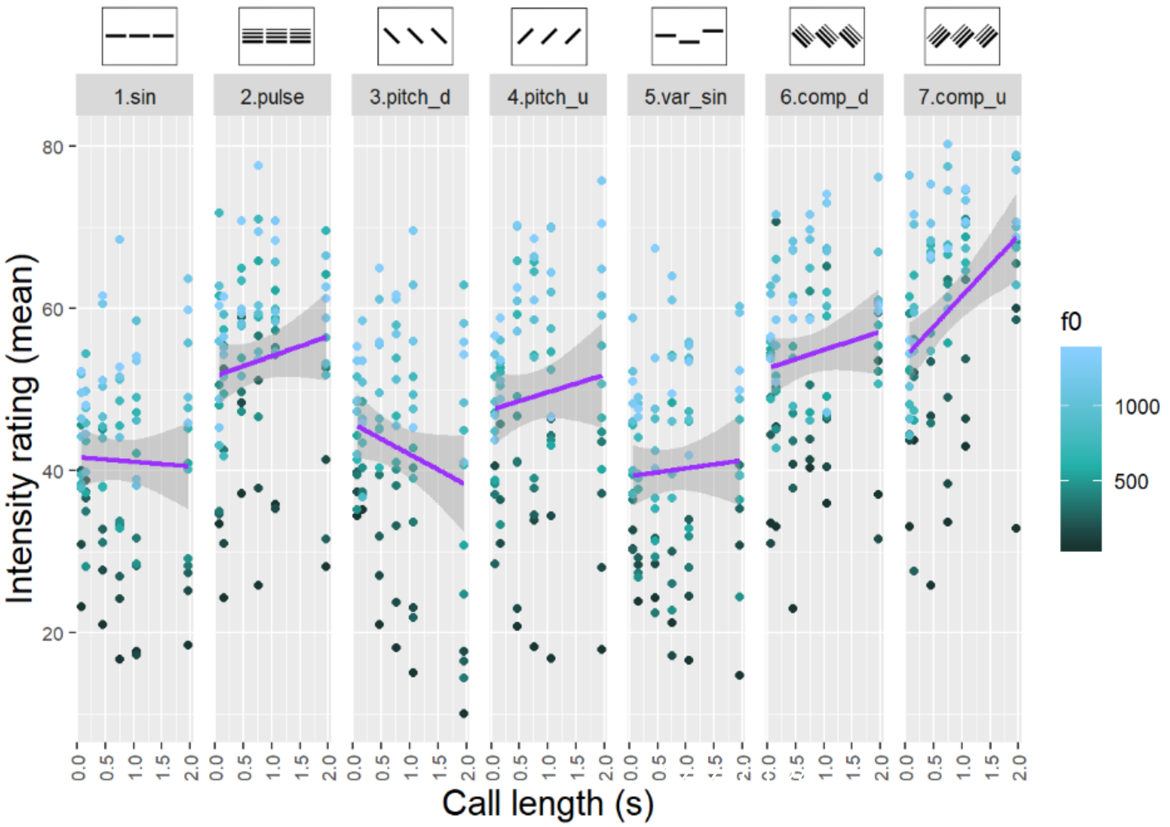
Higher pitch sound → more negative



Intensity

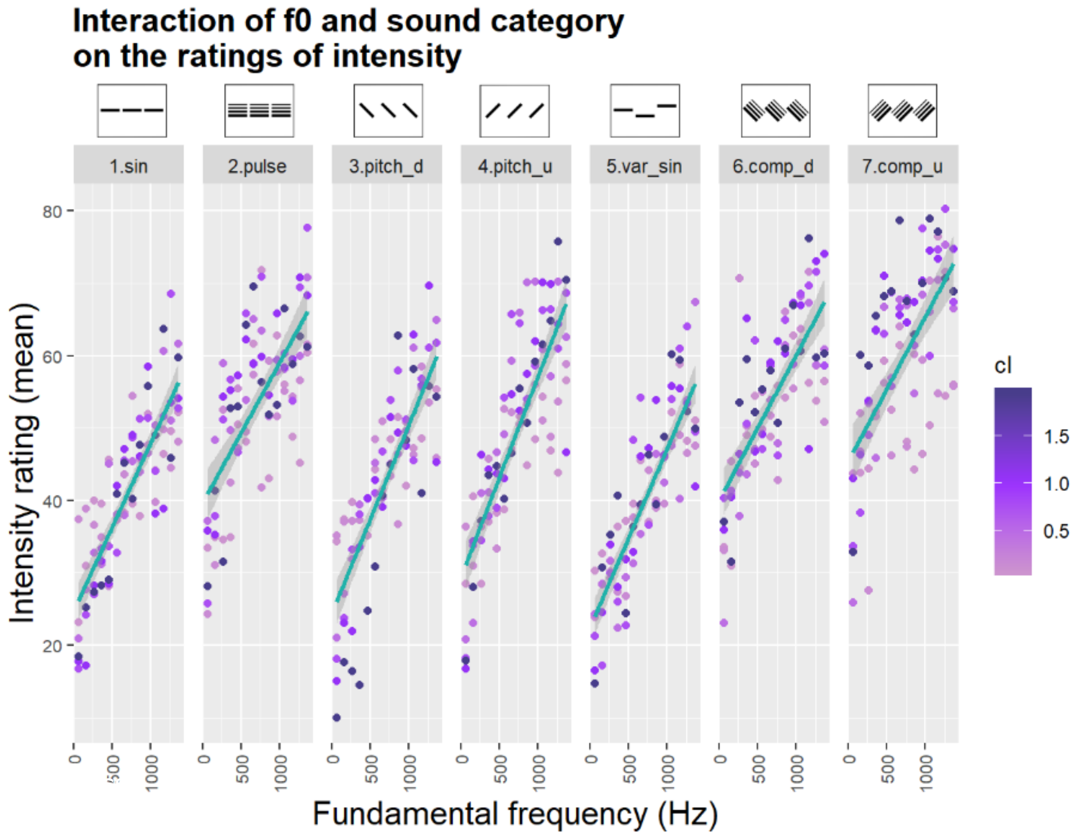
Call lenght

Call lenght → no acoustic rule

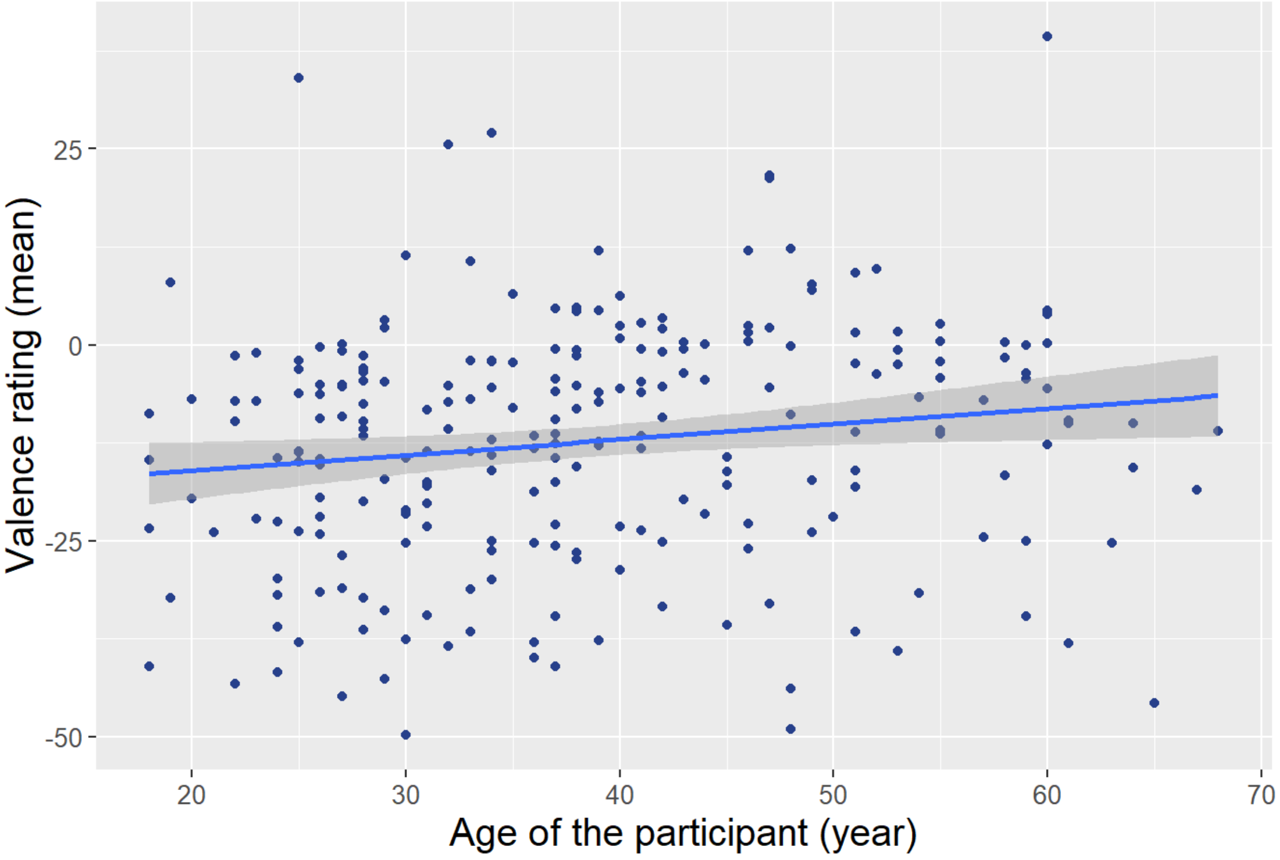


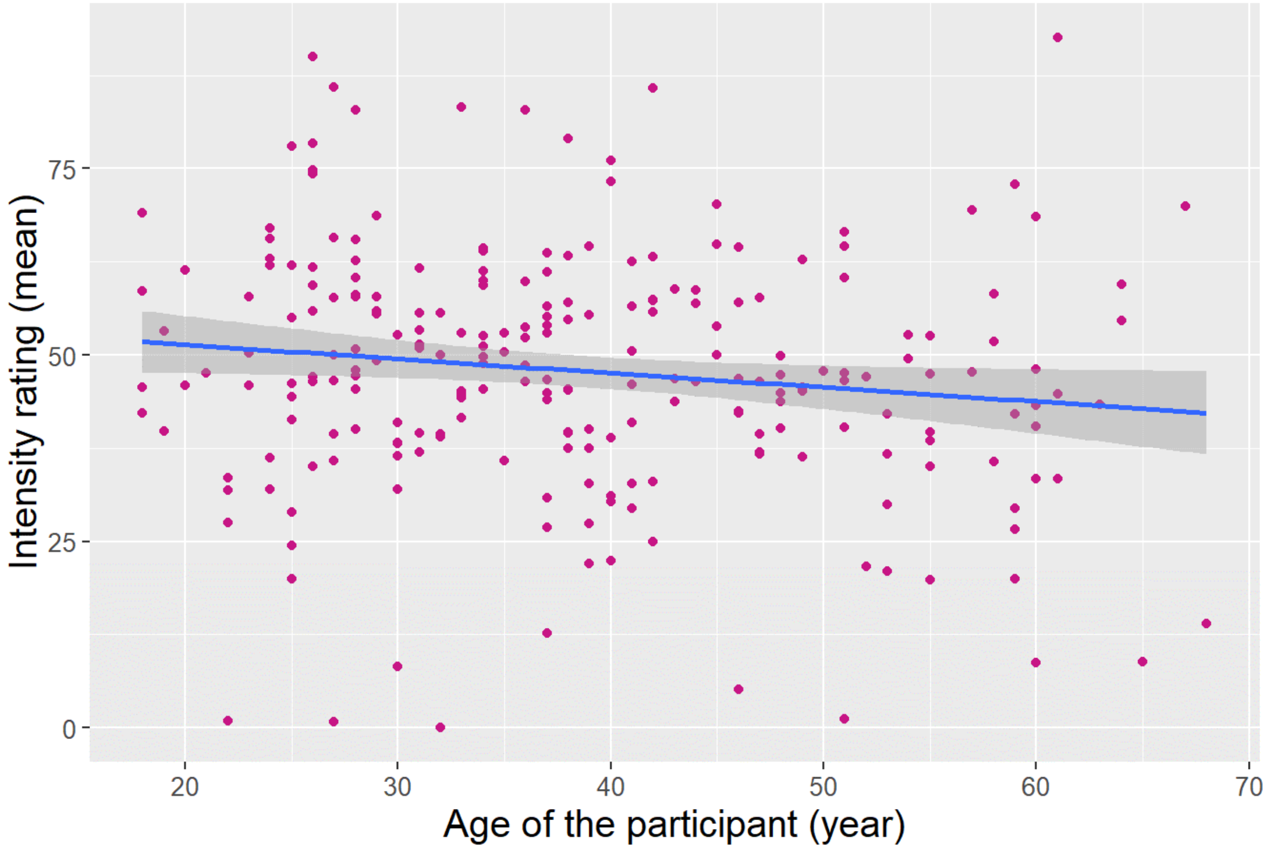
Fundamental frequency

Higher pitch sound → more intense



Age


Valence rating (mean)
25
0
-25
-50
20
30
40
50
60
70
Age of the participant (year)


Intensity rating (mean)
75
50
25
0
20
30
40
50
60
70
Age of the participant (year)

Discussion

- The artificially generated sounds are able to mimic some of the basic coding rules that are present in animal vocalizations

!

  - Higher f_0 → more intense
  - Longer calls → more negative
  - Higher f_0 → more negative
  - Older participants → all sounds less intense, more positive
- Compexity: the direction of association stayed the same in all categories

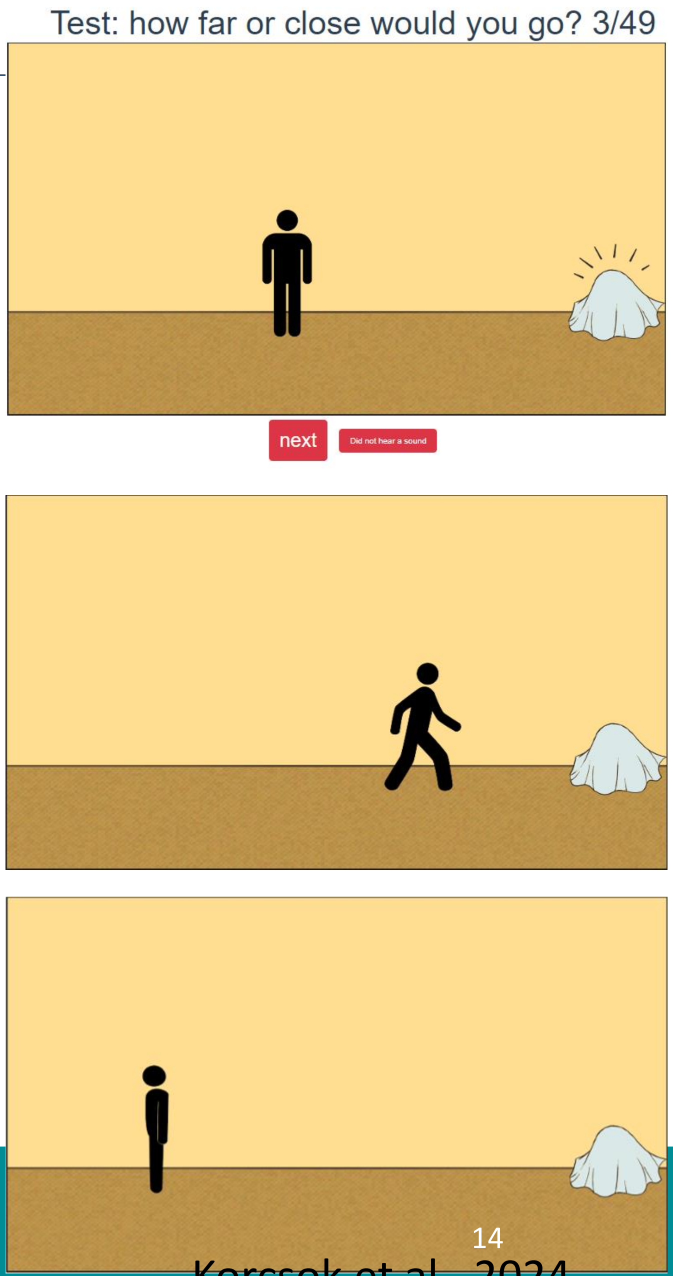
BUT! anger vs. fear?

Social message

- Functional, social message dimension
- Inner state → expression → evoked reaction (Andics & Faragó, 2018)
- Approach-avoidance reactions

Social message

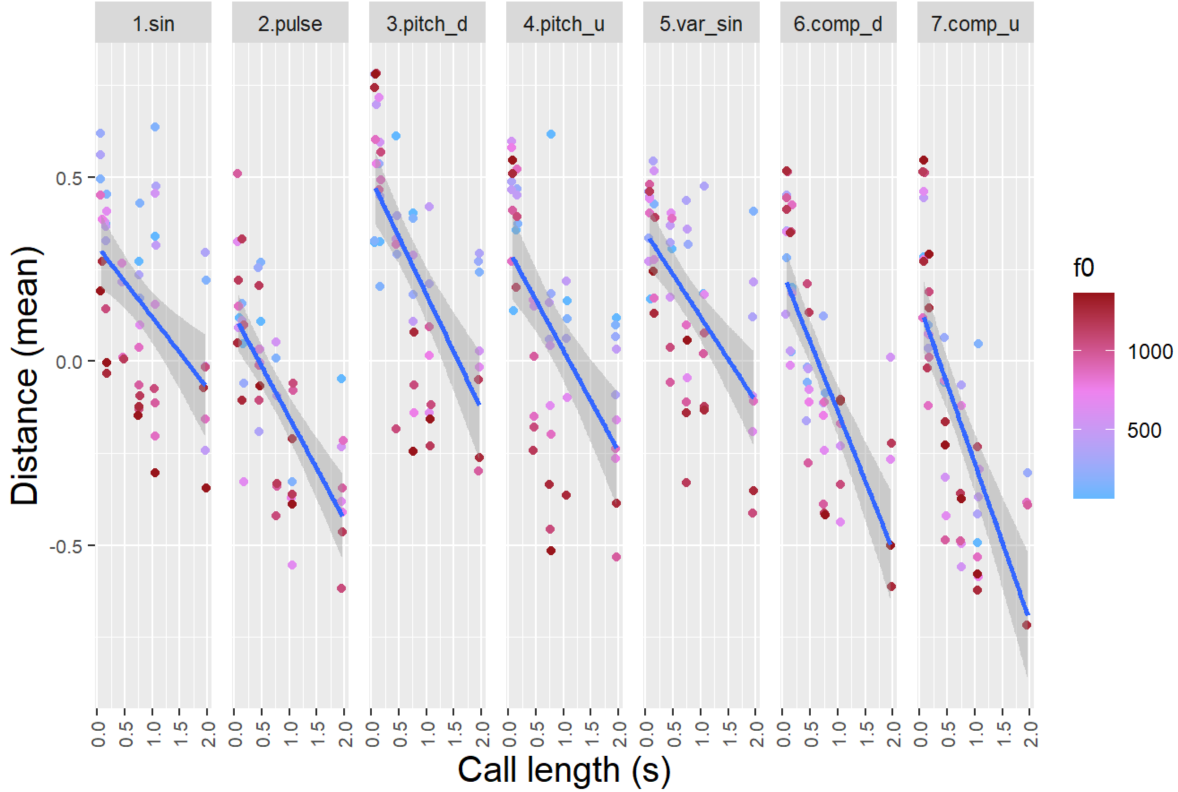
- Artificially generated sounds → vocal features from animal vocalizations
- Subset of sound samples artificially generated via Praat in a previous study investigating valence and intensity (Korcsok et al., 2020)
- Total 343 sounds (49 sounds selected from all 7 categories)
- Online questionnaire with a manikin task (Krieglmeyer & Deutsch, 2010)
- 49 sounds/subject
- N = 172, 92 Hungarian speaking (53♀, 39 ♂, age = 38 ± SD 11 years) and 80 English speaking (55♀, 25 ♂, age = 41.1 ± SD 12.2 years)



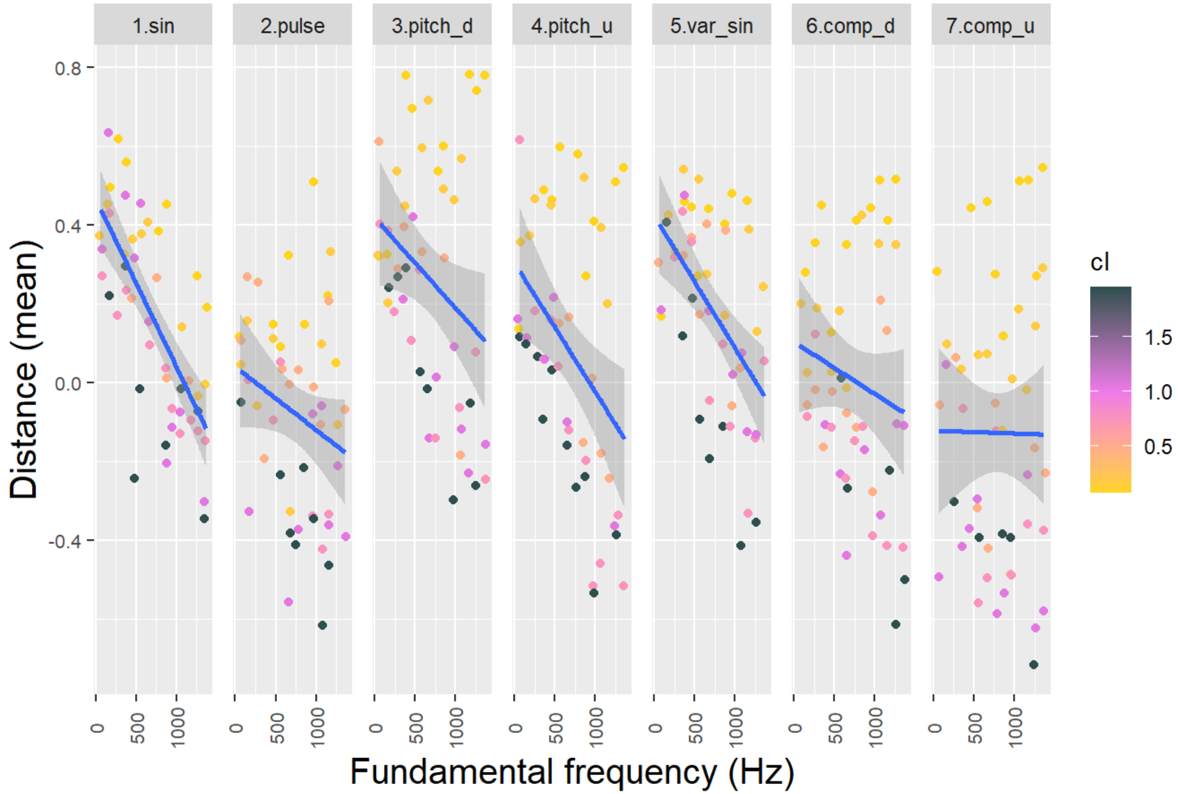
Korcsok et al., 2024

Social message

- Short sounds → approach

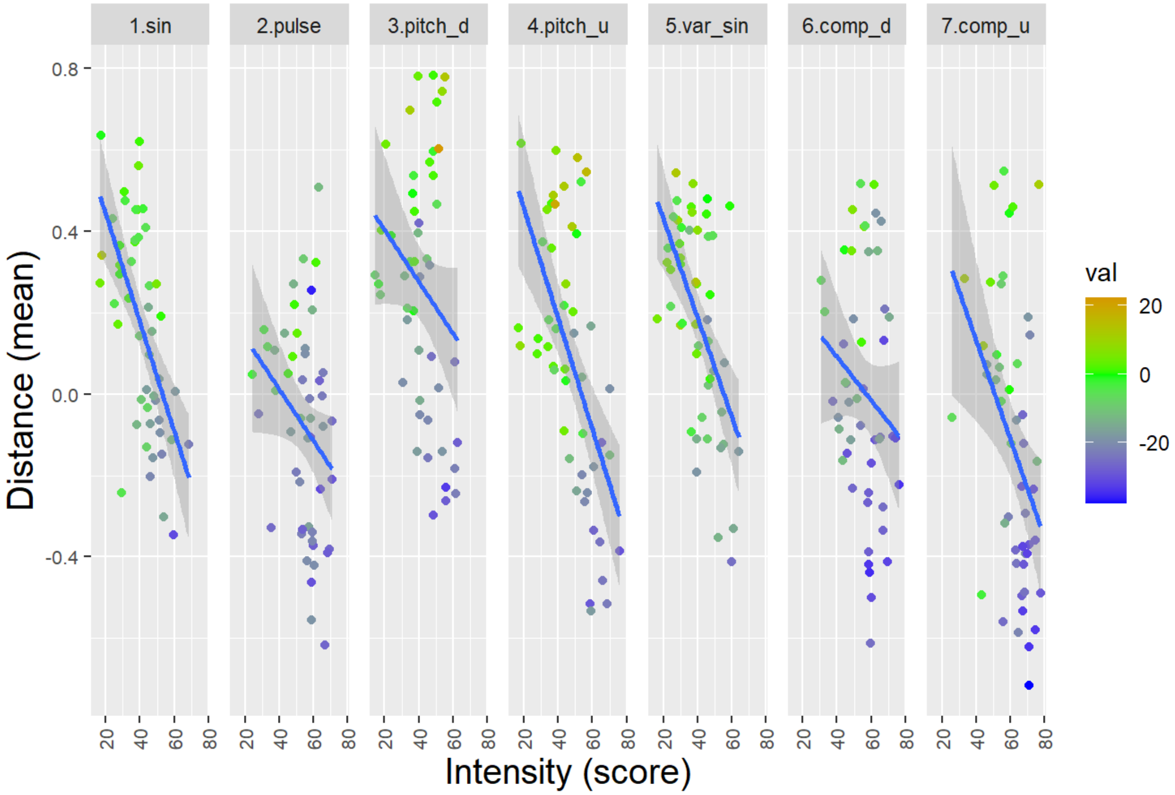


Lower pitch sounds → approach
Lower pitch does not elicit approach in the biologically complex categories



Social message

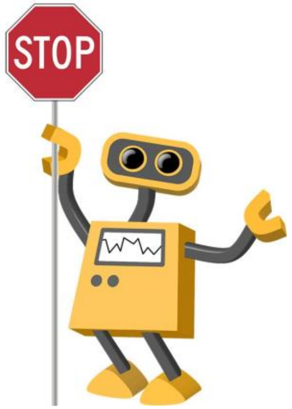
- Positive valence or low intensity sounds → approach
- High intensity sounds → avoidance in case of negative valence sounds

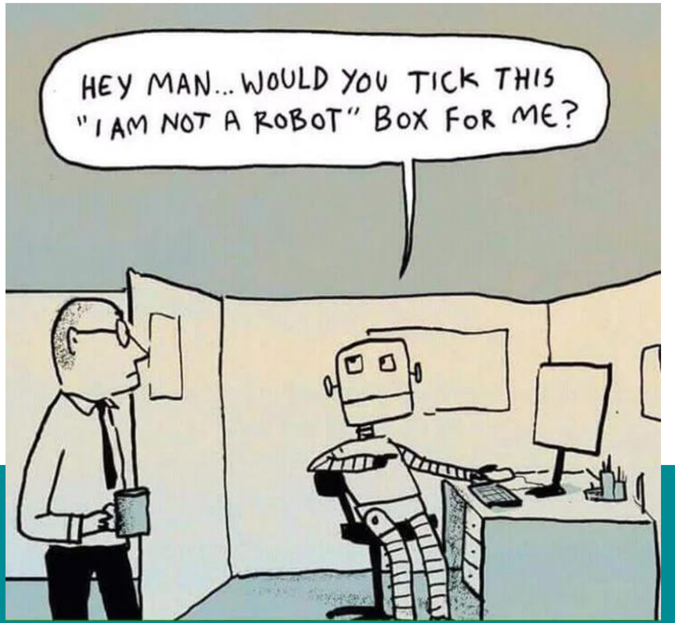


Approach	Avoid
Low pitch + No formants	Low pitch + With formants
High pitch + Soft sound	High pitch + Loud sound
Soft sound	Loud sound
Short calls	Long calls
Long calls + Low pitch	Long calls + High pitch
Positive valence sounds	Negative valence sounds + High intensity sounds
Low intensity sounds	

Social message

- Humans attribute social message to artificially generated sounds, and can react to them with approach or withdrawal
- The elicited approach-avoidance reactions are linked with acoustic features
- Connection with the perception of valence and intensity
- Artificial sounds for social robots → sounds with specific acoustic parameters for specific communicative functions in HRI
- Sounds implemented in a social robot





Leading to sound source

Figure 2. Example of a hearing dog using visual communication signals to guide its owner to the front door

Koay et al., 2013

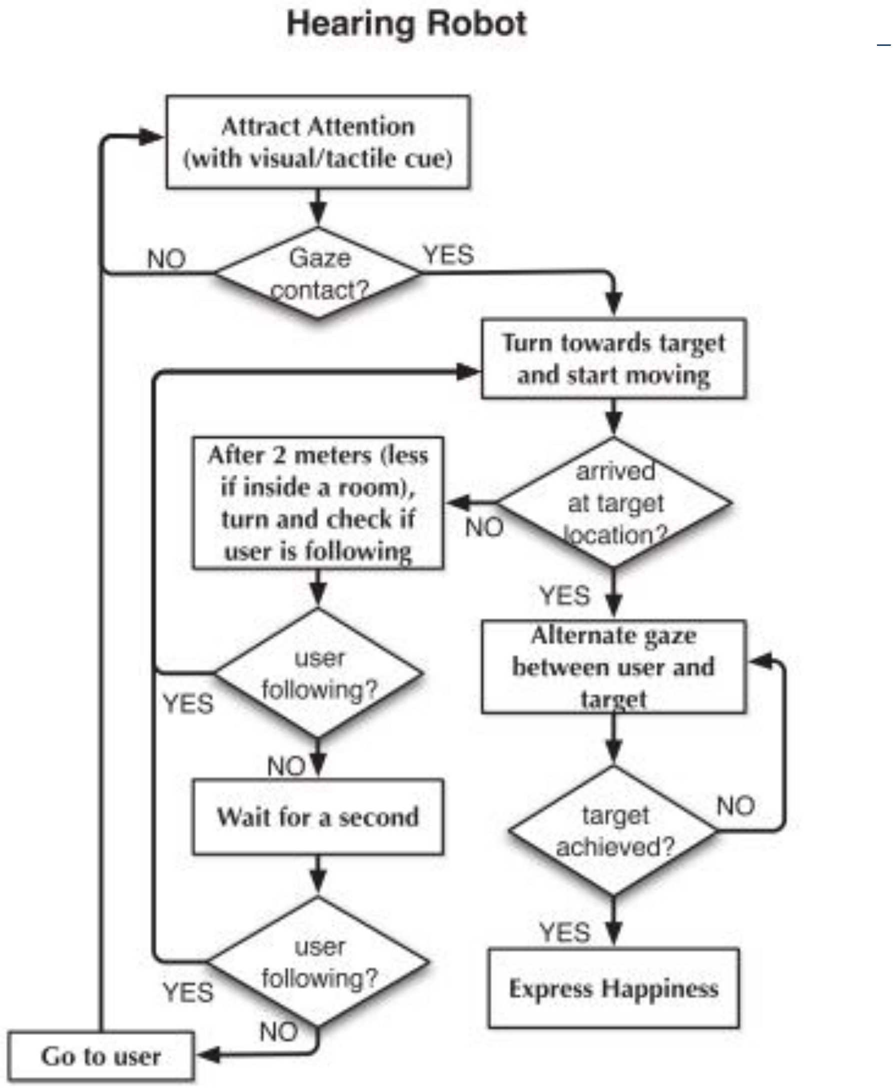


Figure 3. Flowchart of dog-inspired hearing robot behavioral strategy

Leading to sound source

Figure 1. The Sunflower Robot in the UH Robot House.

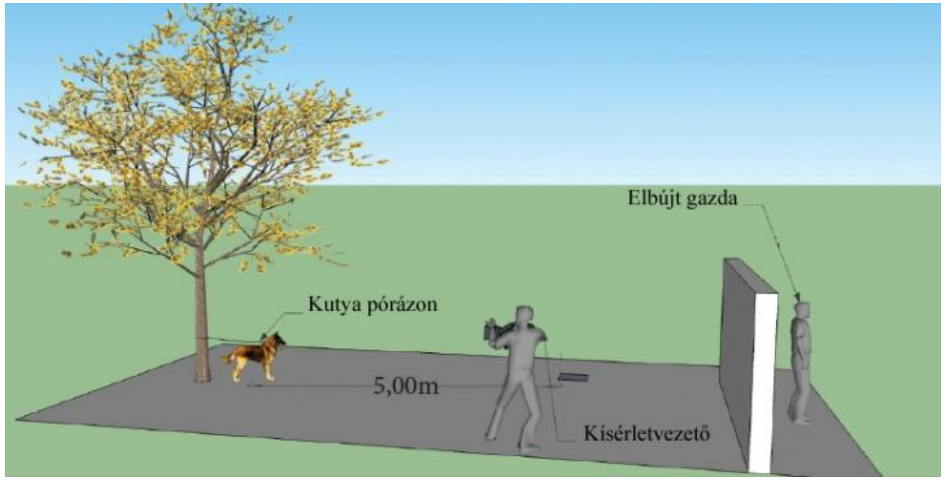
TABLE IV. IMPORTANCE OF ROBOT BEHAVIORS FOR COMMUNICATING INTENTIONS

Category	No. of Responses	% of Responses
Gaze	10	62.5%
Head Movement	8	50.0%
Body Movement	3	18.8%
Feedback	5	31.3%
Lights	4	25.0%

Human-dog → human-robot interaction

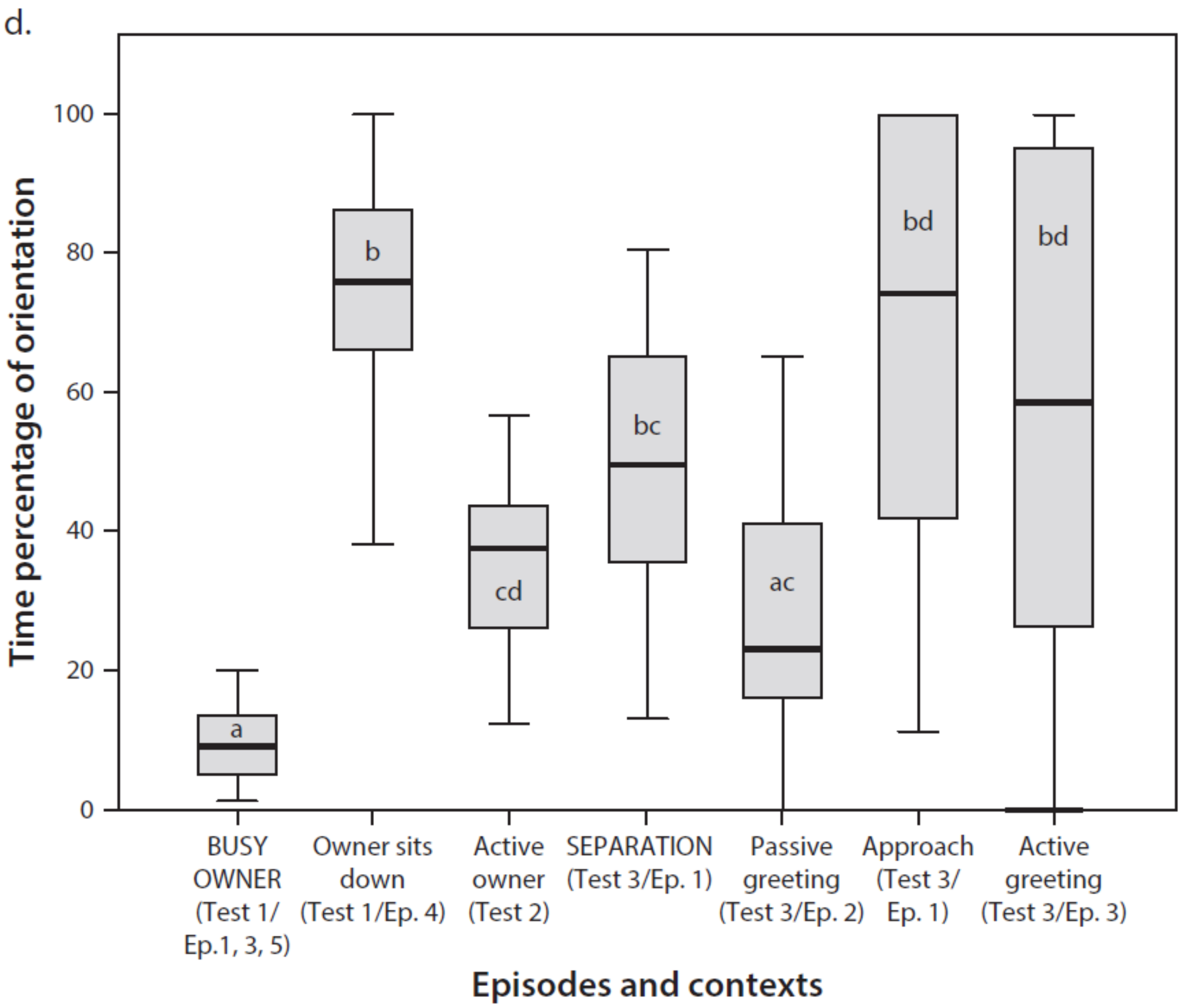
- 29 dog-owner pairs

- 3 tests, several different situations:
- Greeting (separation + approach + active greeting)
- Busy owner (filling in the questionnaire) + passive greeting, sitting down
- Active owner (carrying objects)



Faragó et al., 2014

Human-dog → human-robot interaction



Human-dog → human-robot interaction

Situation dependent behaviours

- Orientation
- Activity
- Exploration
- Tail wagging

In what situations should the robot look at the person?
When should it be active?
Communication, expression of emotions

Individual dependent behaviours

- Proximity seeking
- Tail wagging
- Activity

Robot personality

Robots optimised for individual people

Social monitoring

- Monitoring without disturbing
- Low-level communicative signals
- Readiness for interaction
- „Alive-ness”

!

Attachment

- **Criteria (Rajecki et al., 1978)**

1. Distinguishing „attachment person” (secure base)

!

2. Preference towards „ attachment person” (maintaining proximity)

3. Sensitivity to separation (specific behavioural changes)
   - when the attachment person is missing (pl. searching, passivity)
   - during reunion (greeting)

Attachment

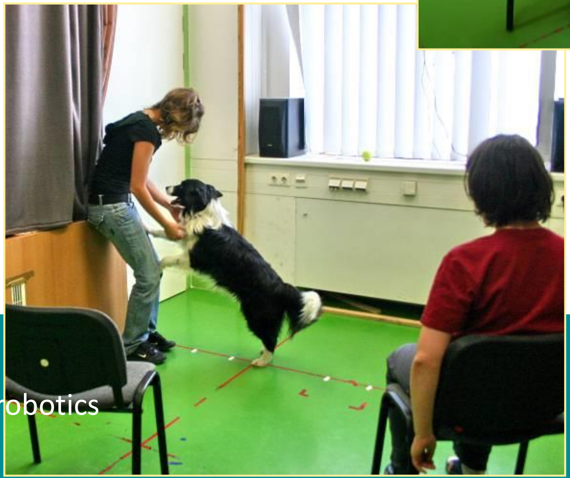
- Species with a complex social structure (e.g. humans)
- Conditions (abilities):
- Individual recognition
- Social monitoring
- Sensitivity to separation
- Proximity seeking
- In a dangerous situation the attachment behaviour is activated

Ainsworth's Strange Situation Test (Ainsworth et al., 1978)

Attachment

Observed behaviour data:
How long does the dog:

- Play?
- Wait (beside the owner, beside the door)?
- Explore?
- How many times does it initiate contact?



Topál et al., 1998

Attachment - 'EtoMotor'

With the cooperation of the University of Miskolc

Fuzzy Rule-based interpolation (if … then … rules)

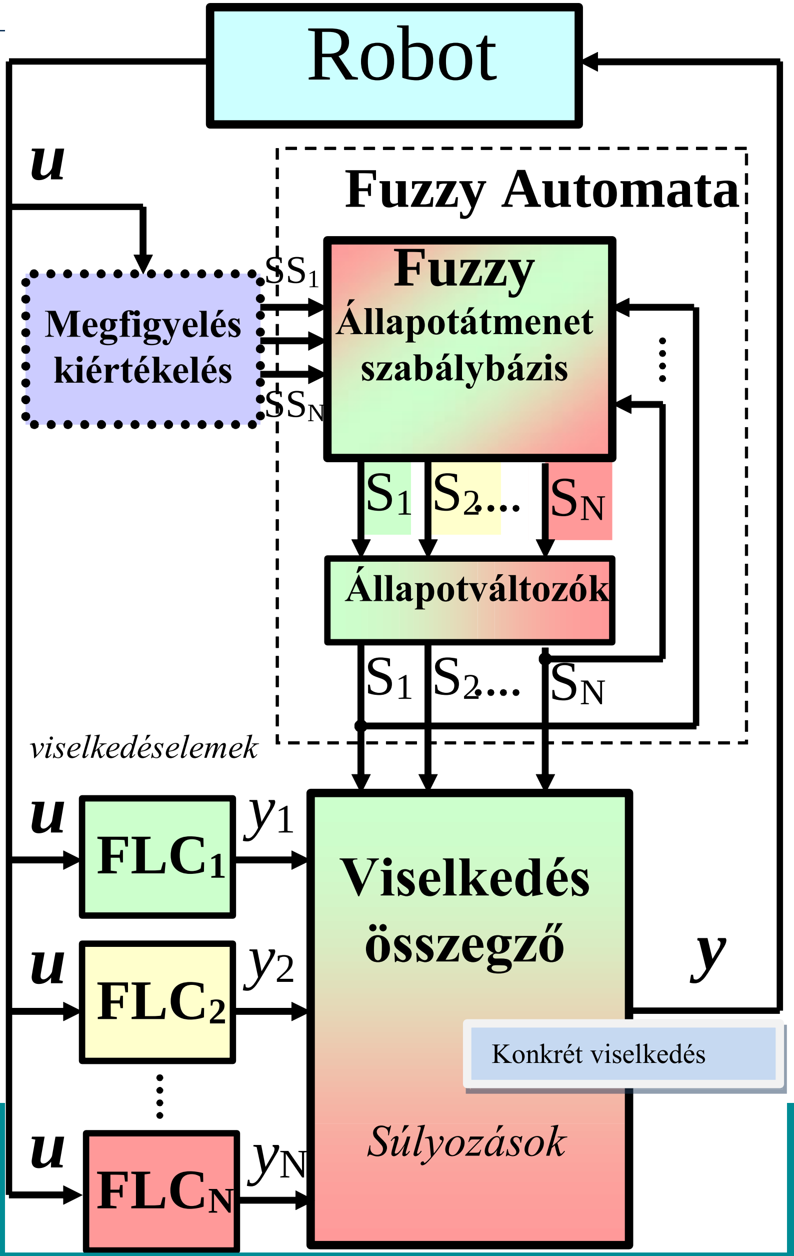
Continuous transition between states

There are parallel and mutually exclusive states

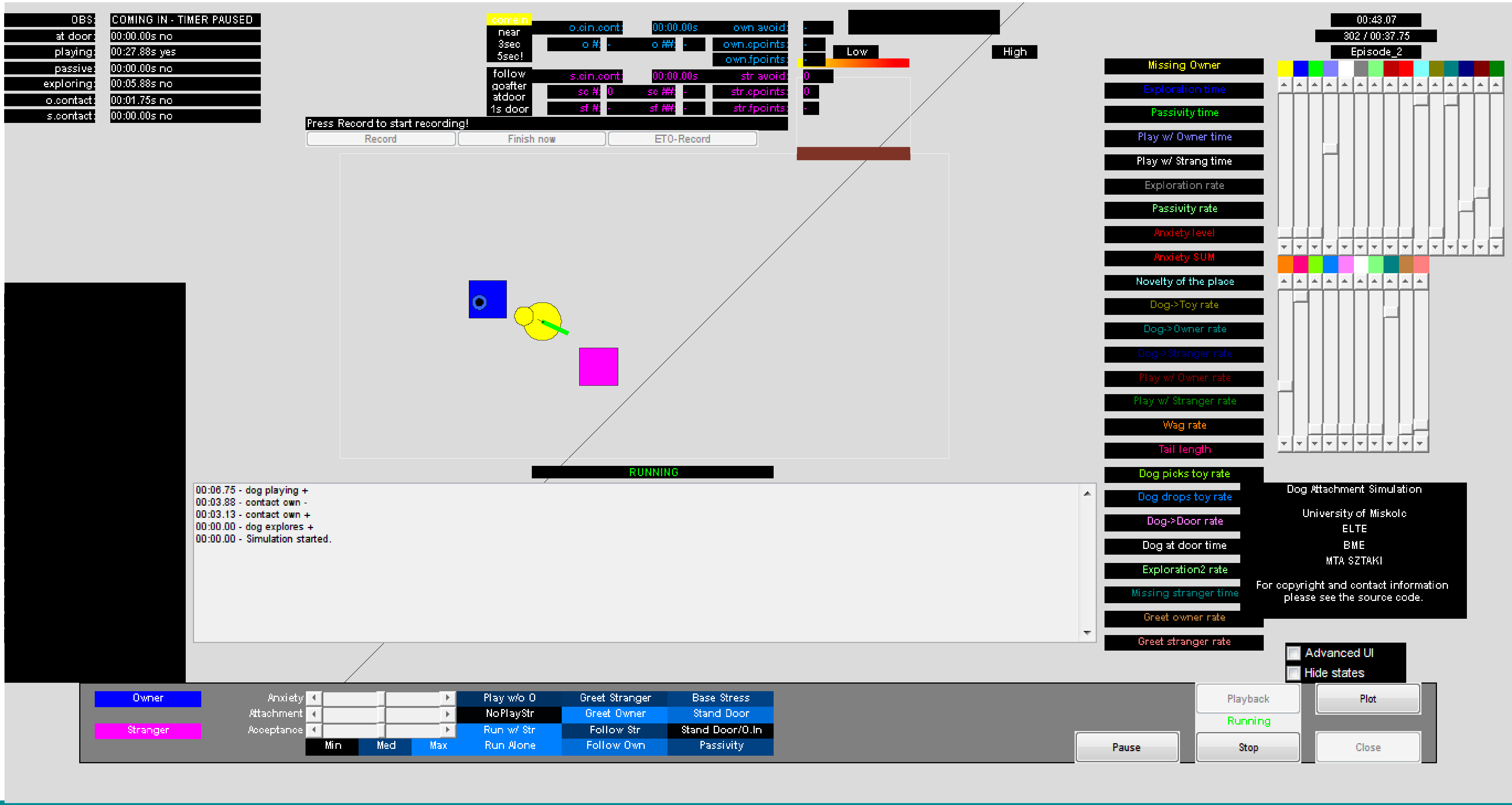
Examples:

"If the owner is close and the stranger is far away, the stress is reduced

If the owner threw the ball and the stress is low, the enthusiasm for the game increases"



Attachment



Kovács et al., 2011

Waiter assistant robot - Biscee

- Possibility to examine human-robot interaction: in real-life situations, spontaneous interactions
- Robotnik base, unique upper structure
- Autonomous operation (navigation, head movement, face tracking, sound)
- ROS (Robot Operating System)
- 8 hour operating time
- RGBD, LIDAR
- Stereo cameras on the head
- UH sensors
- Interactive skills

HUN-REN Comparative Ethology Research Group

Pizza test

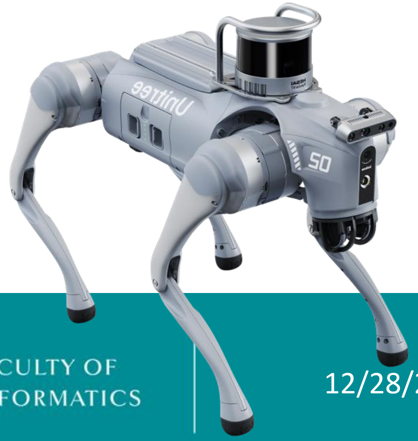
- Guests are greeted by the robot at the door:
- Greeting behaviour: looking at the face + emitting sound
- Leading to the table and serving:
- Attention-getting signals: gaze alternation between person and target + looking at face + vocalization
- Sending away with an ultrasonic sensor
- Experimental groups: interactive; minimally interactive (ultrasonic sending only)

Spontaneous interactions with Biscee

- Hostess work: handing out candies at events
- Comparison with human: candy distribution speed, orientation at the agent, touch, avoidance, standing in its way, waving, pointing…)

Unitree vs. Biscee

- We do not know how children and animals react to commercially available robots
- Do these elicit fear or aggression?
- Robots with different embodiments: animal-like (Unitree) and not animal-like (Biscee)
- Comparison with other agents (unfamiliar dog and human)
- Our new robot, Ethonic: Unitree Go2 Edu



References

- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). Strange situation procedure. Clinical Child Psychology and Psychiatry.Billard, A., Robins, B., Dautenhahn, K., & Nadel, J. (2006). Building robota, a mini-humanoid robot for the rehabilitation of children with autism. The RESNA Assistive Technology Journal, 19(ARTICLE).
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. Journal of Zoology, 288(1), 1-20.
- Ehret, G. (2005). Infant rodent ultrasounds–a gate to the understanding of sound communication. *Behavior genetics*, *35*(1), 19-29.
- Fant, G. (1960). On the acoustics of speech.
- 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
- Faragó, T., Andics, A., Devecseri, V., Kis, A., Gácsi, M., & Miklósi, Á. (2014). Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations. Biology letters, 10(1), 20130926.
- Gácsi, M., Kis, A., Faragó, T., Janiak, M., Muszyński, R., & Miklósi, Á. (2016). Humans attribute emotions to a robot that shows simple behavioural patterns borrowed from dog behaviour. Computers in Human Behavior, 59, 411–419. https://doi.org/10.1016/j.chb.2016.02.043
- Koay, K. L., Lakatos, G., Syrdal, D. S., Gácsi, M., Bereczky, B., Dautenhahn, K., Miklósi, A., & Walters, M. L. (2013). Hey! There is someone at your door. A hearing robot using visual communication signals of hearing dogs to communicate intent. *IEEE Symposium on Artificial Life (ALIFE)*, *2013-Janua*(January), 90–97. https://doi.org/10.1109/ALIFE.2013.6602436
- 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), 1-13.
- Korcsok, B., Faragó, T., Ferdinandy, B., Miklósi, Á., Korondi, P., & Gácsi, M. (2024). People follow motivation-structural rules when they react to synthetised sounds. *Scientific Reports*, *14*(1), 17656.
- Kovács, S., Gácsi, M., Vincze, D., Korondi, P., & Miklósi, Á. (2011). A novel, ethologically inspired HRI model implementation: Simulating dog-human attachment. In *2nd International Conference on Cognitive Infocommunications* (pp. 1-4).
- Knutson, B., Burgdorf, J., & Panksepp, J. (2002). Ultrasonic vocalizations as indices of affective states in rats. *Psychological bulletin*, *128*(6), 961.
- Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance behaviour: The manikin task vs. two versions of the joystick task. *Cognition and emotion*, *24*(5), 810-828.
- Maruščáková, I. L., Linhart, P., Ratcliffe, V. F., Tallet, C., Reby, D., & Špinka, M. (2015). 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. Journal of Comparative Psychology, 129(2), 121.
- Miklósi, Á., Korondi, P., Matellán, V., & Gácsi, M. (2017). Ethorobotics: A new approach to human-robot relationship. *Frontiers in Psychology*, *8*(JUN), 1–8. https://doi.org/10.3389/fpsyg.2017.00958
- Pongrácz, P., Molnár, C., Miklósi, A., & Csányi, V. (2005). Human listeners are able to classify dog (Canis familiaris) barks recorded in different situations. Journal of comparative psychology, 119(2), 136.
- Rajecki, D. W., Lamb, M. E., & Obmascher, P. (1978). Toward a general theory of infantile attachment: A comparative review of aspects of the social bond. *Behavioral and Brain Sciences*, *1*(3), 417-436.
- Russell, J. A. (1980). A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161.
- Topál, J., Miklósi, Á., Csányi, V., & Dóka, A. (1998). Attachment behavior in dogs (Canis familiaris): a new application of Ainsworth's (1969) Strange Situation Test. *Journal of comparative psychology*, *112*(3), 219.

Thank you for your attention!

Korcsok Beáta
korcsokbea@gmail.com