NRI: FND: Creating Trust Between Groups of Humans and Robots Using a Novel Music Driven Robotic Emotion Generator

NSF 2041253 - Poster #150

Gil Weinberg and Richard Savery

Georgia Tech

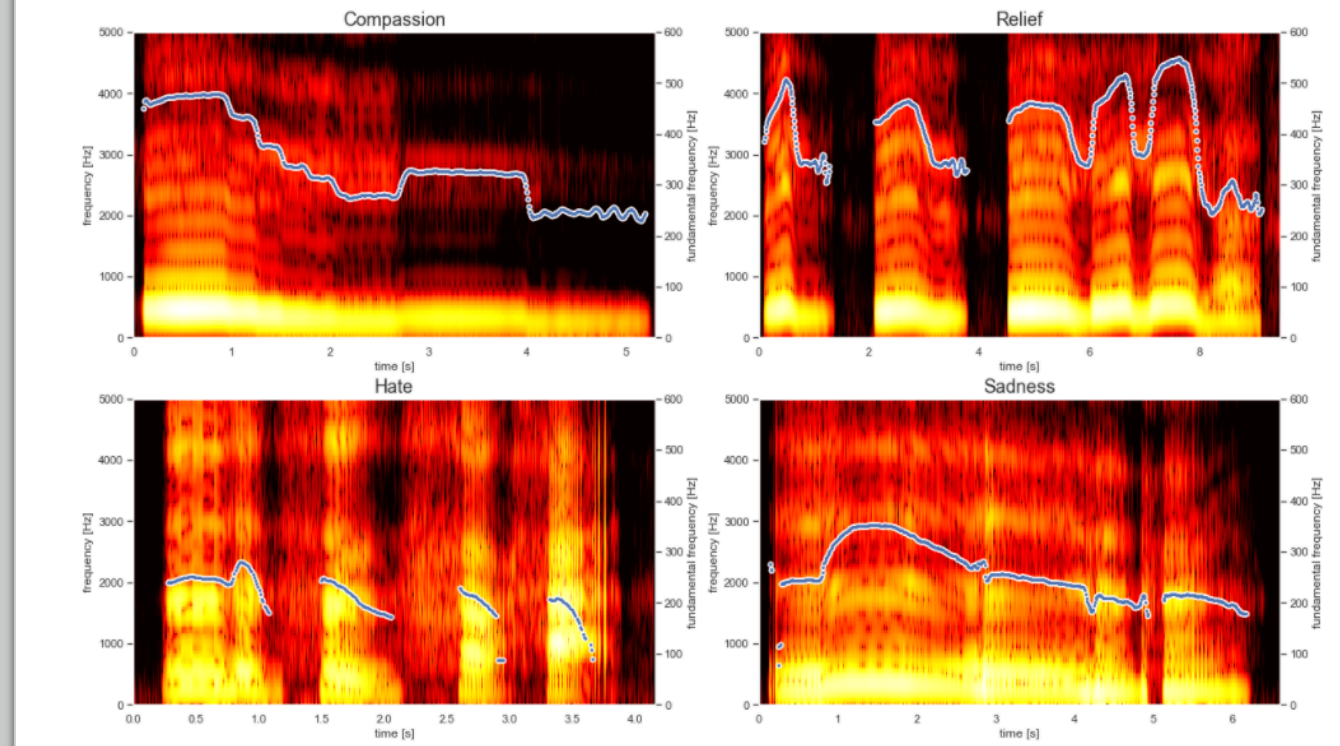
Prosody

The patterns of stress and intonation in spoken language

Conveys meaning, emotional state, social cues, humor, irony and more.

Has been shown to improve trust in human-to-human interaction.

Can be used as a background channel that can be processed peripherally



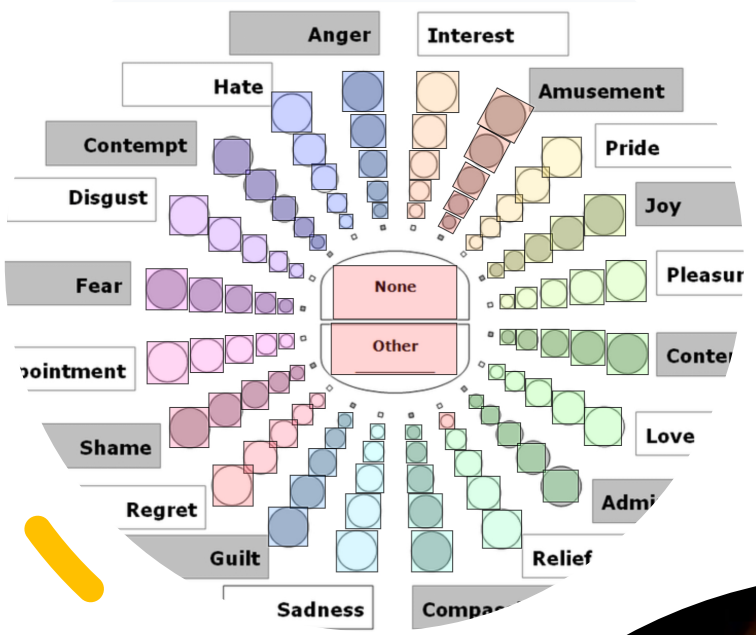
Spectrogram of Musical Prosody Phrases.

The same same sentence is pronounced using different emotional prosodies

The blue line indicates pitch contour

Dataset

4.2 hours of emotionally labeled singing phrases (average of 122 phrases for each emotion)



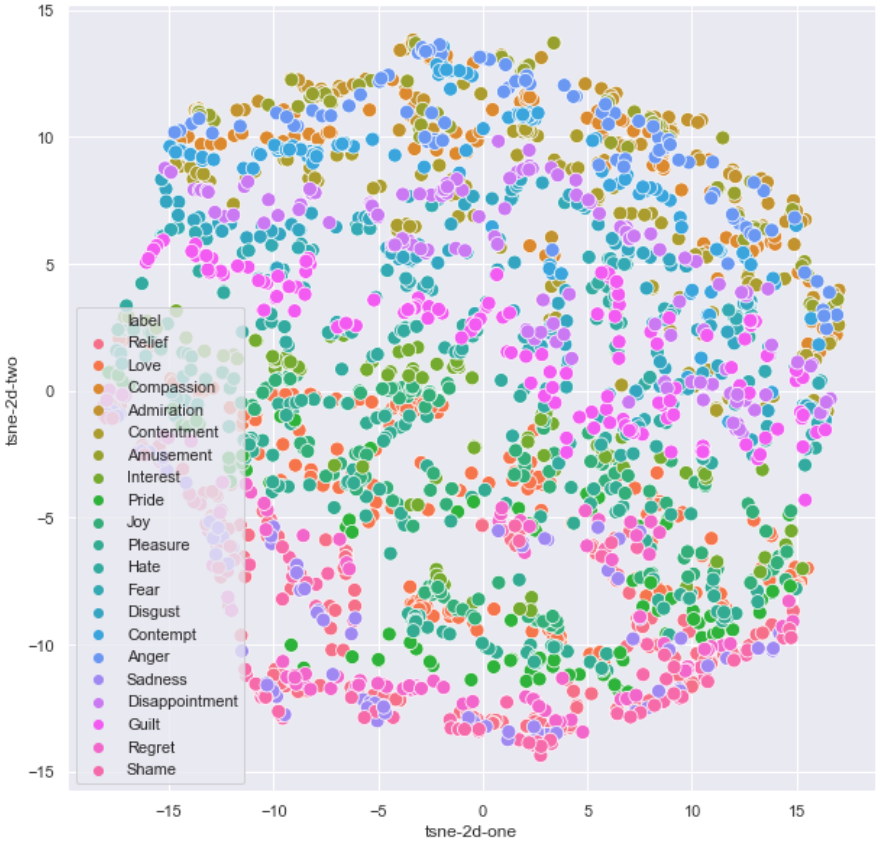
Geneva Emotion Wheel Model

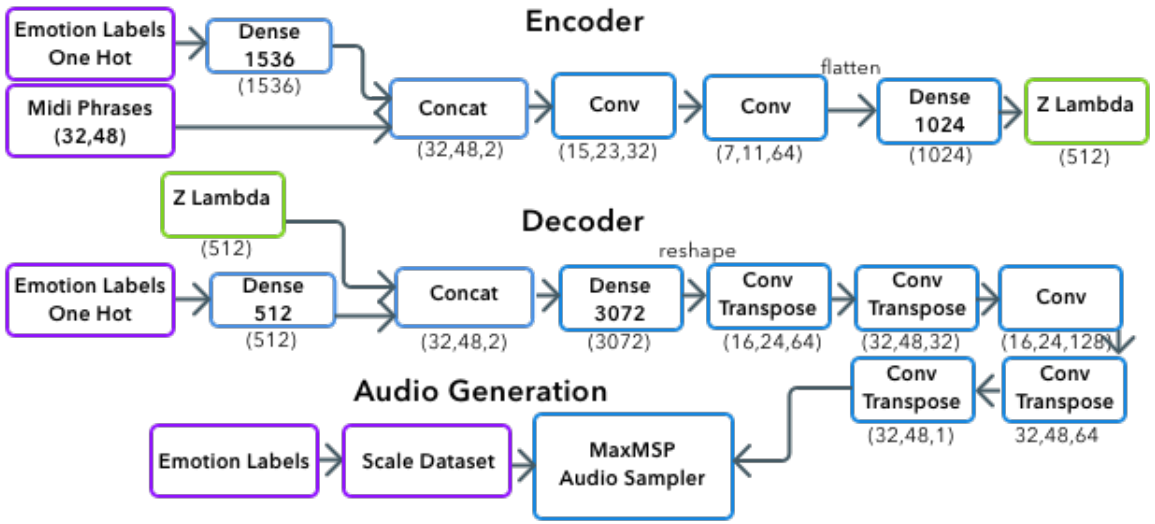
- Continuous valance / arousal
- Discrete emotion descriptors

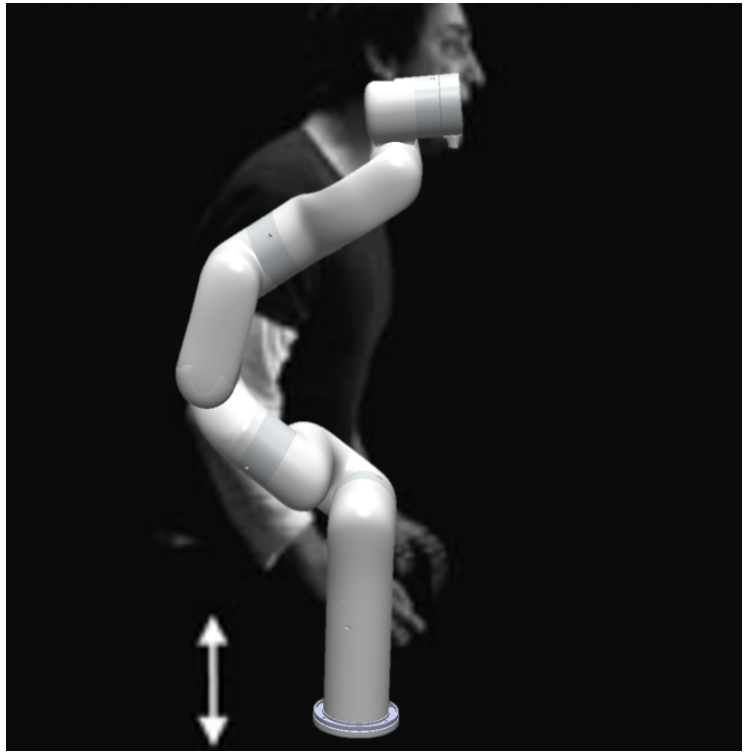
Validation – 45 participants survey

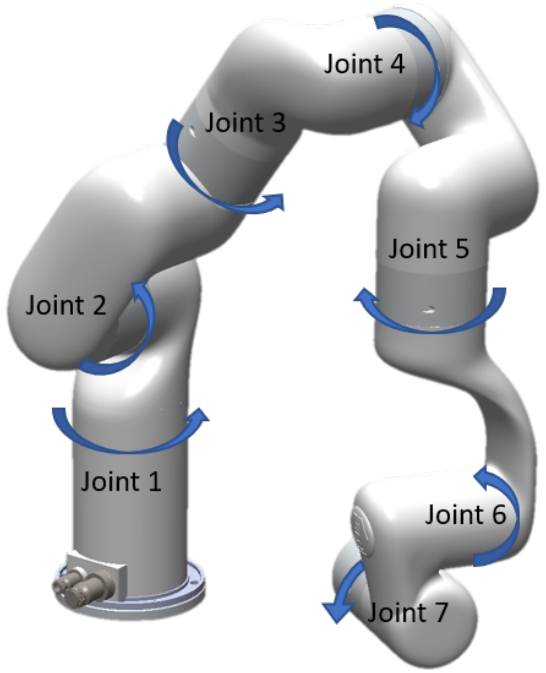
Training and Generation

- ConvNet Conditional Variational Auto Encoder
- Latent space separation
- MIDI and synthesized audio generation









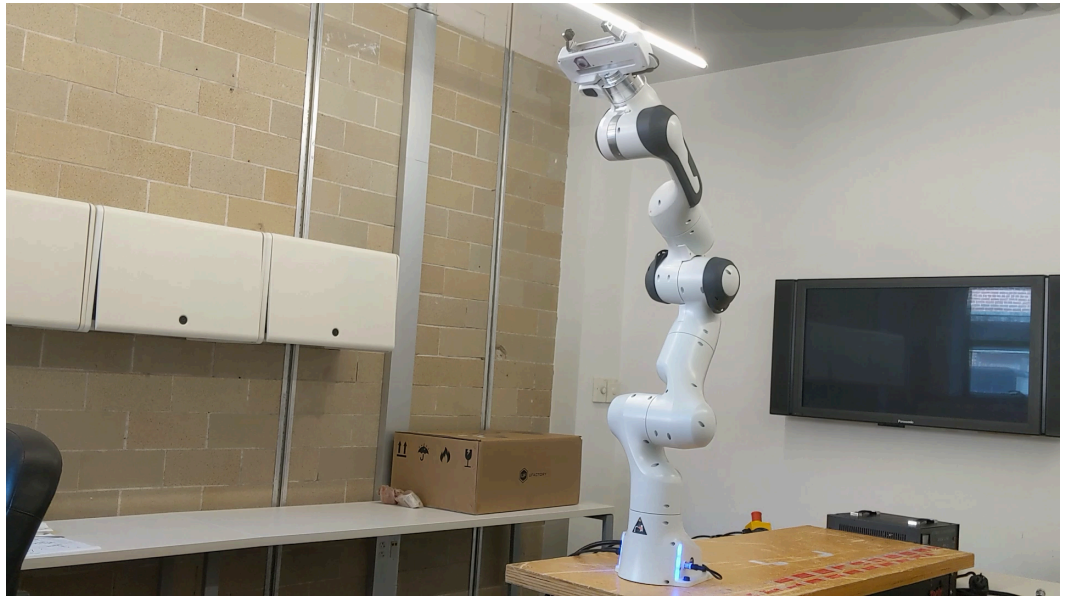
Joy	Head bent up	Joint 6	Fast	Joint tilts end of robot upwards
	Chest bent up			
	Up-down repetative arm motion	Joint 4	Fast	Raises and lowers top half of robot
	corners of lips are drawn back and up			
	Body action: Jumping, Shape change: Expansion			
	Cheek Raiser			
Shame	Head facing down	Joint 6	Medium	Joint tilts end of robot downwards
	Collapsed Shoulders	Joint 4	Slow	Collapses top half of robot down
	CITE (backwards leaning)	Joint 2	Slow	Leans robot away from stimulus
	Raising of inner brows		Fast	

Gesture Design

Walbott (1998) – “Bodily Expression of Emotion”
Dael (2012) – “Emotion Expression in Body Action and Posture”

Sound and Gesture

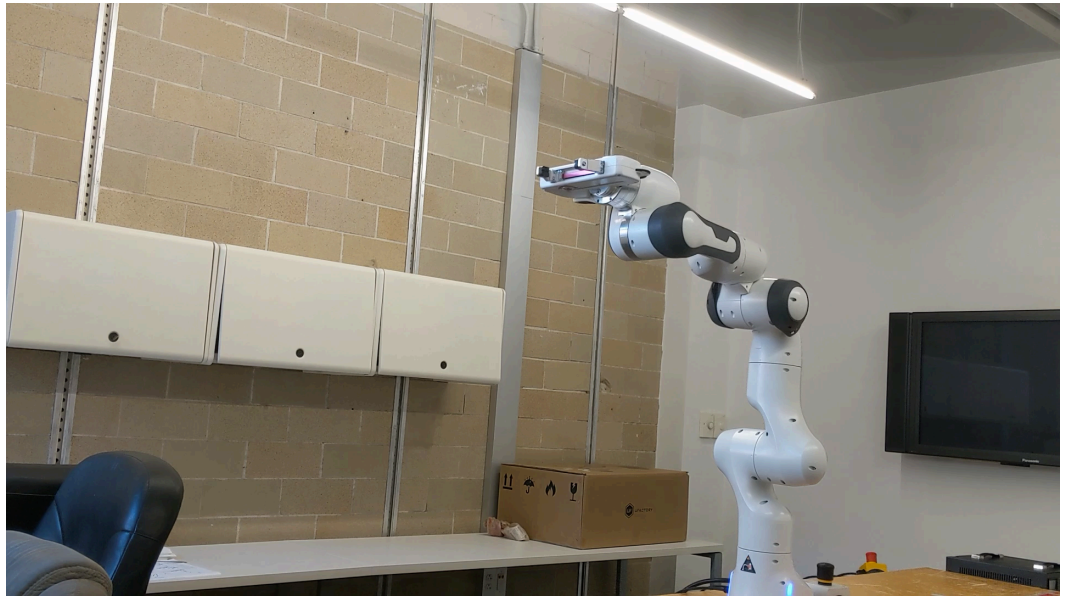
- Implementation on a Panda co-bot arm
- Sound for Joy:
  - Synthesized by our generative system
- Gesture for Joy:
  - Joint 4 bounces up and down
  - Joint 6 tilts upwards



Joy

Sound and Gesture

- Implementation on a Panda co-bot arm
- Sound for Disgust:
  - Synthesized by our generative system
- Gestures for Disgust:
  - Joints 2 and 6 tilt away quickly
  - Joint 5 twists away

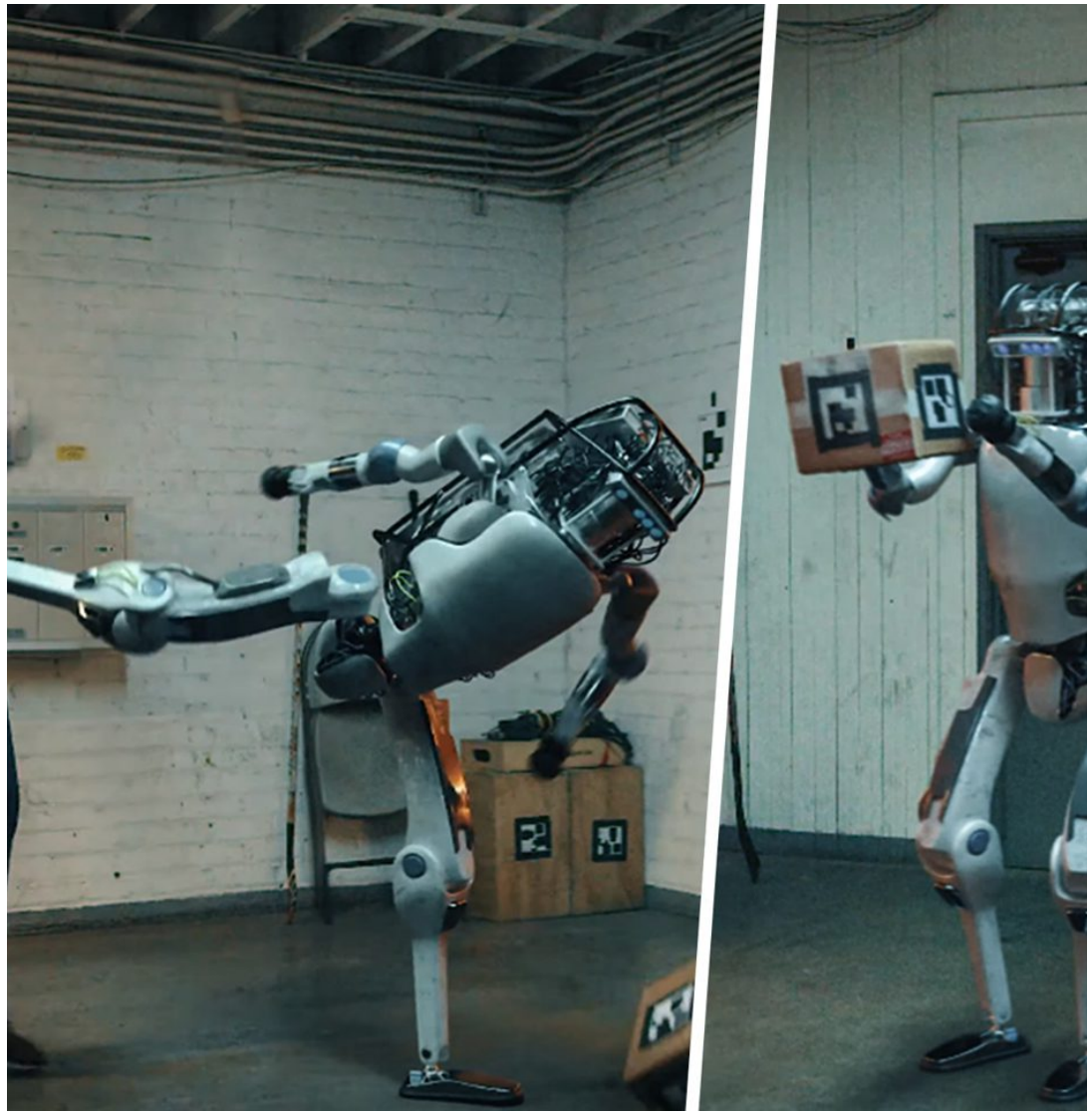


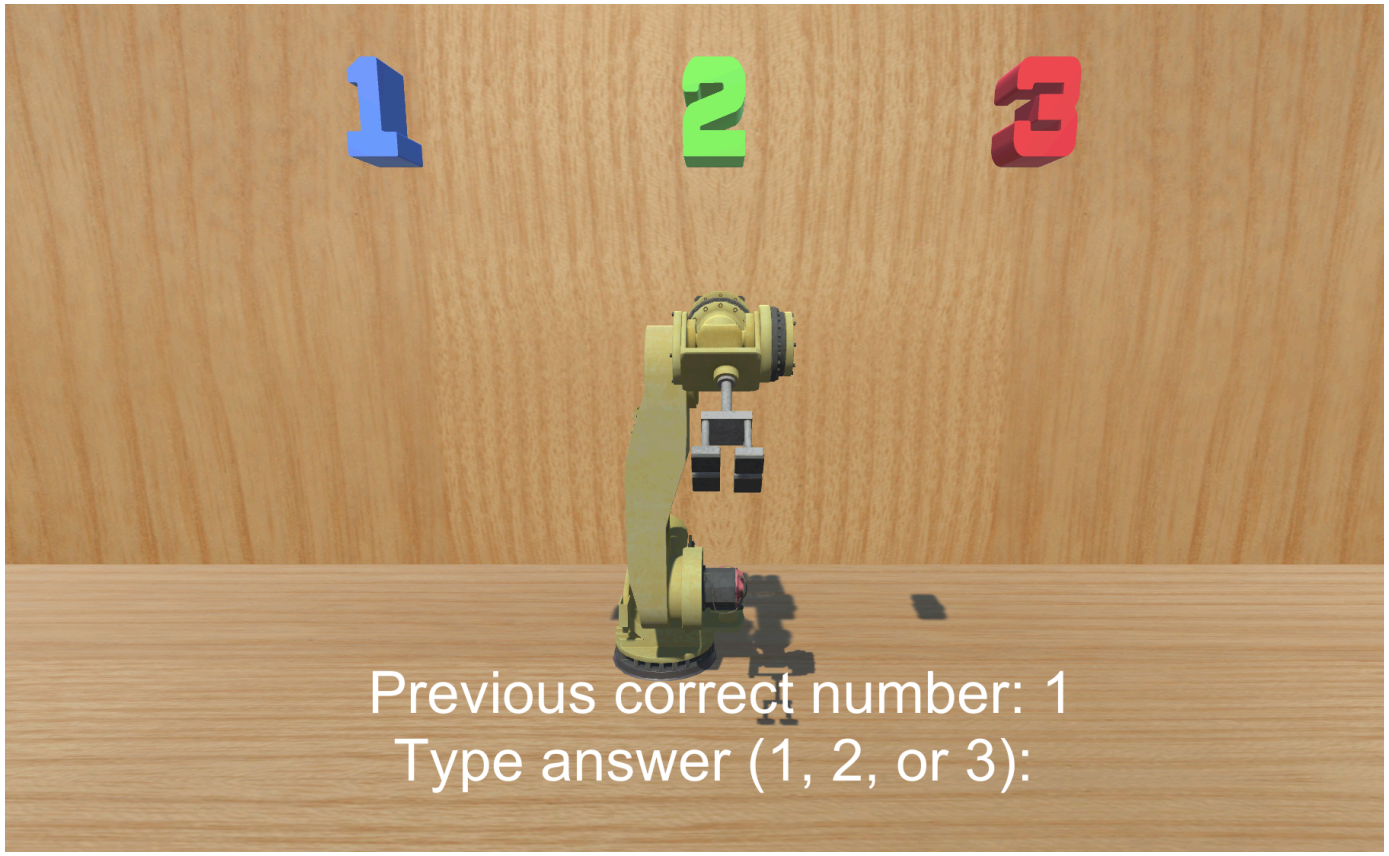
Disgust

Study – HRI Trust

Research Questions

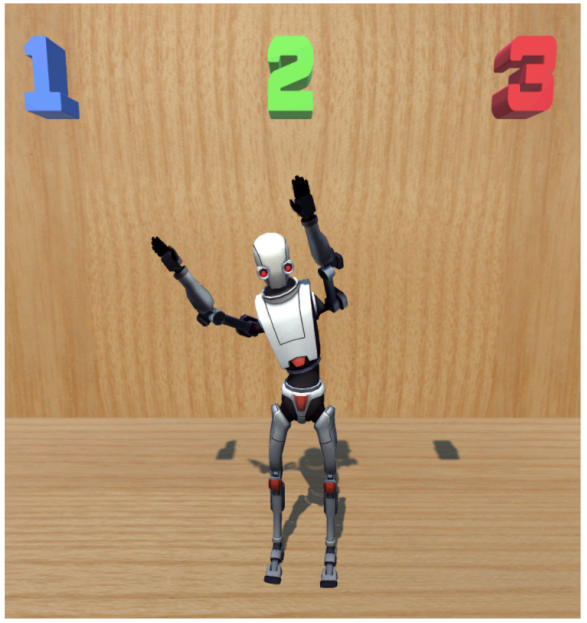
- RQ1 - How can emotional musical prosody alter trust and trust recovery from mistakes, compared to no audio and single-pitch audio
- RQ2 - How can emotional musical prosody alter perceived safety, animacy intelligence and likeability?



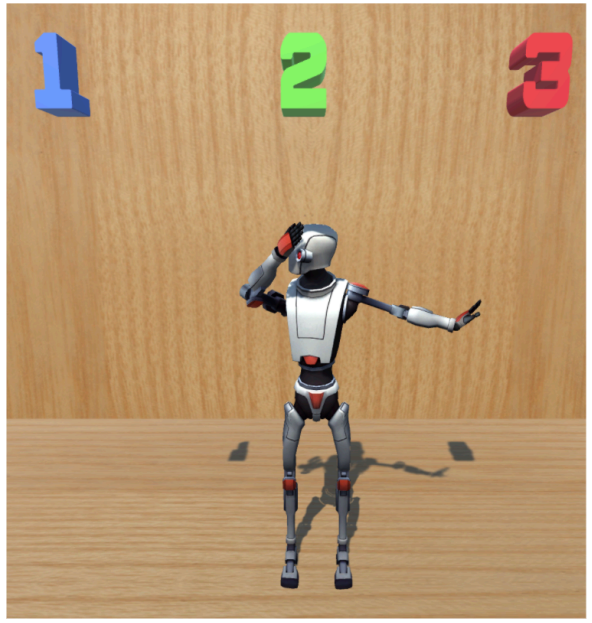


Experiment Design – Robotic Arm

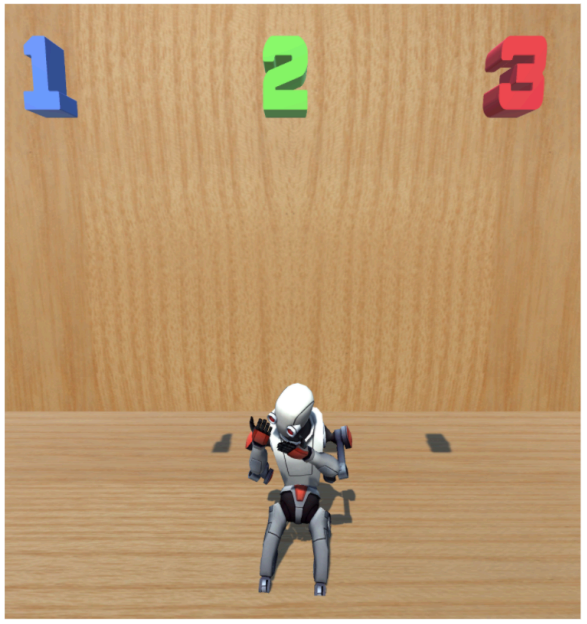
- Dongen's pattern learning and prediction task
- Godspeed measurement for **Anthropomorphism**, **Animacy**, **Likeability**, **Perceived Intelligence**, and the **Perceived Safety** of Robots
- Schaefer's survey for **Robotic Trust**



Joy



Shame

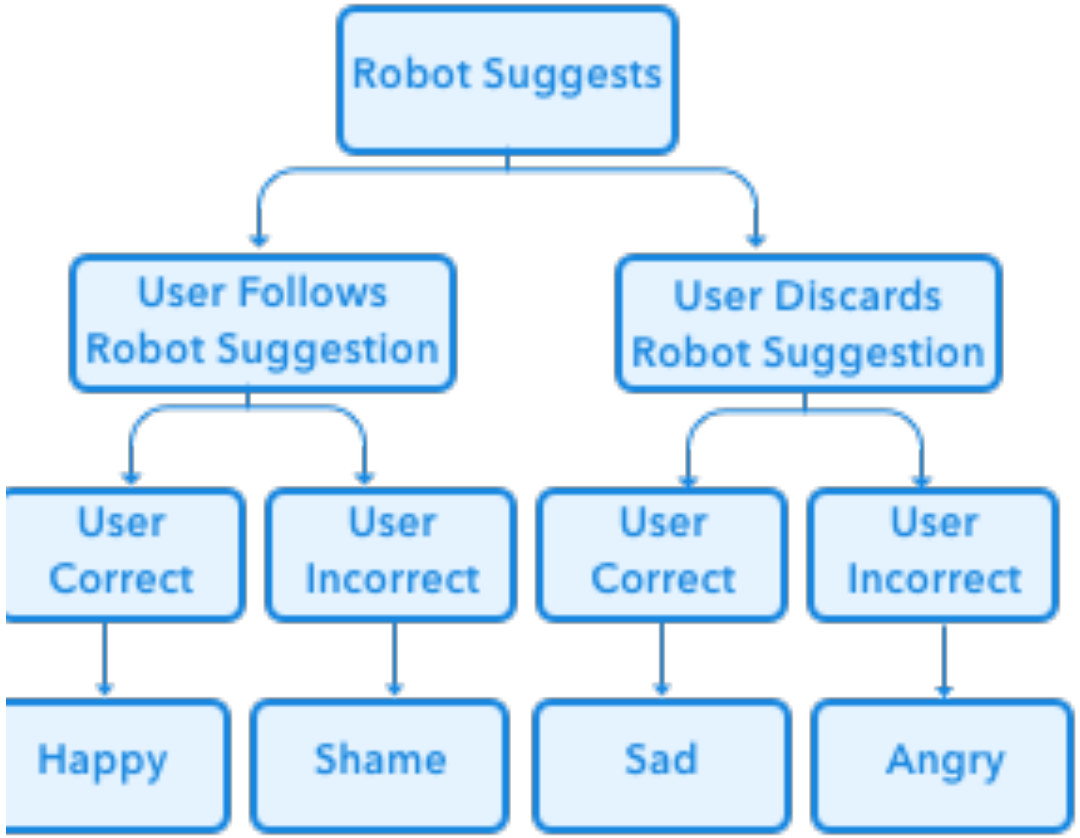


Sadness

Anger

Experiment Design – Hominoid Robot

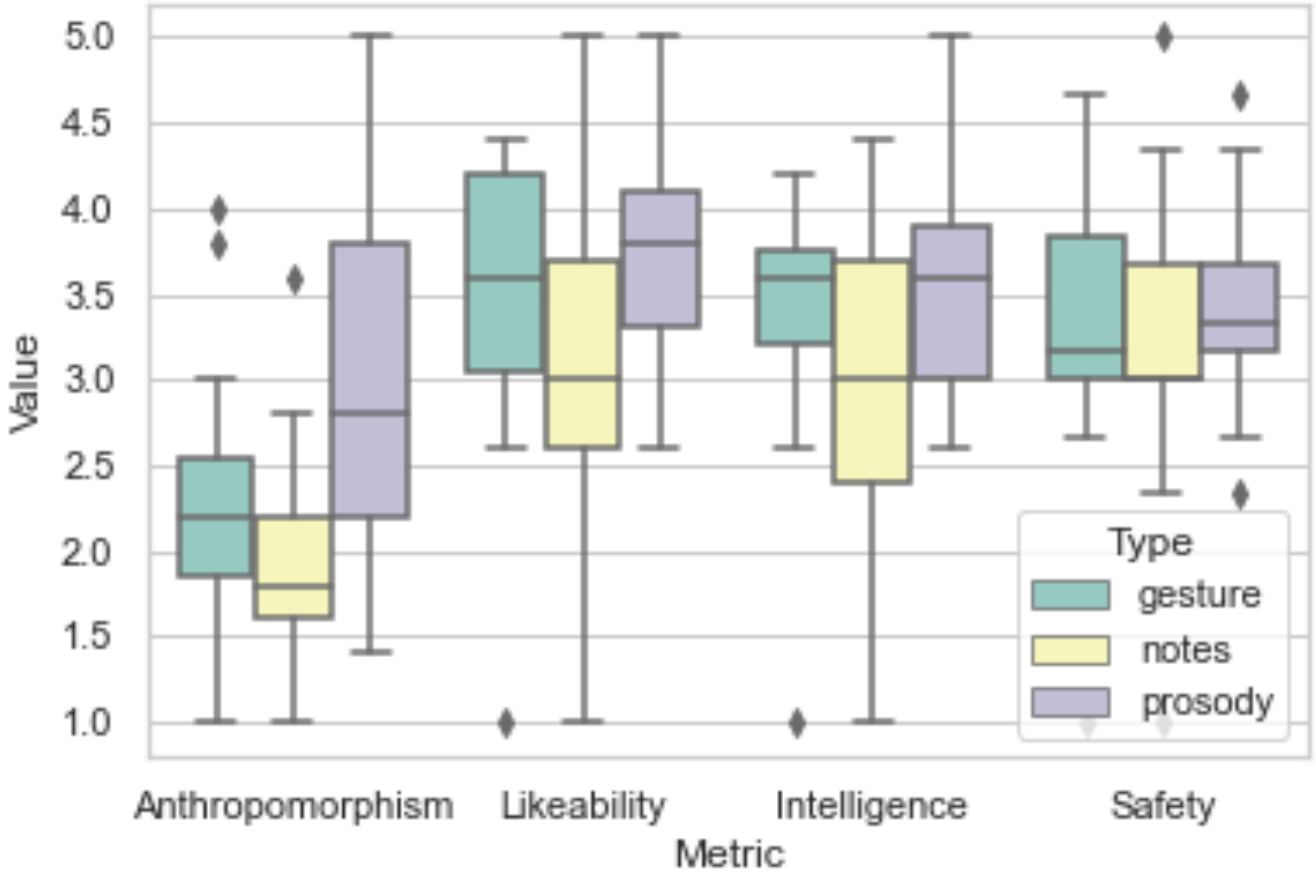
- Dongen's pattern learning and prediction task
- Godspeed measurement for **Anthropomorphism**, **Animacy**, **Likeability**, **Perceived Intelligence**, and the **Perceived Safety** of Robots
- Schaefer's survey for **Robotic Trust**



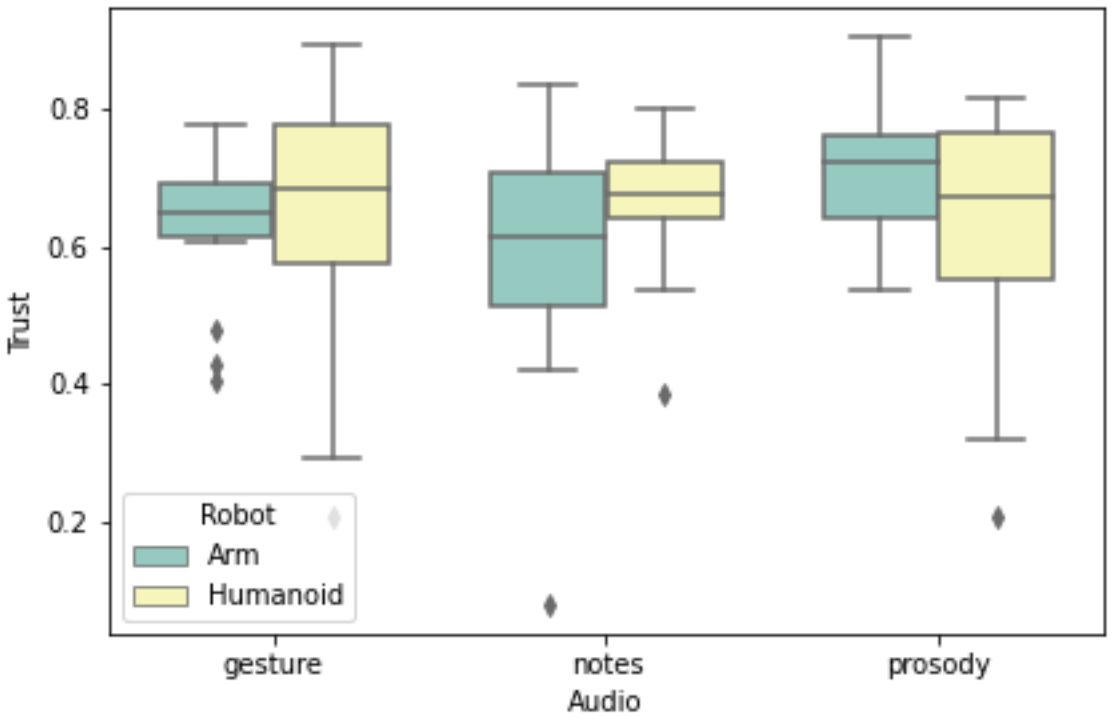
Experiment Design - Process

- Between Group
- 92 participants age 19-49
- Conditions
  - Audio with emotional musical prosody
  - Single-pitch melody
  - No audio

Results – Godspeed Metrics



- Prosody significantly increased perceived anthropomorphism and safety
- Using notes rather than vocalized audio had a negative effect on most metrics



Results - Trust

- Prosody much more effective in supporting trust in robotic arm in comparison to humanoid
- Users' ratings of trust did not strongly correlate with their actual behavior during the task
- When the robot responded with musical prosody users reported higher trust metrics than when the robot responded with single-pitched notes or no audio

Current and Future Work

- Robotic Personality
- New Platforms
  - Mobile Manipulator
  - Social Robot
  - Co-bot arm
- Groups of Robots and Humans