Honda Research Institute

When, What, and How: Designing Empathic Behaviors for Social Robots

Invited Lecture — AFAR Lab, University of Cambridge

Christian Arzate Cruz
November 17, 2026

Empathic Multimodal Robots

Design and implement **empathic robots** that **sense** multimodal cues and **express** using verbal and nonverbal behaviors.



User

When?
the user seeks empathy

while Haru listens

while Haru speaks

What to say?
empathic text

nodding,
vocal gestures

How to show it?

voice, face,
gestures



Haru

Why Empathy?

Empathic robots foster

- Emotional connection [1]
- Understanding [1]
- Engagement [2]

Definition

To experience (**affective**) and understand (**cognitive**) the feelings of another, and act according to those phenomena (**behavioral**) [3].

[1] Lamia Alam. "Examining Cognitive Empathy Elements within AI Chatbots for Healthcare Systems". PhD thesis. Michigan Technological University, 2022. DOI: 10.37099/mtu.dc.etdr/1437 [2] Zhou et al. (2020), *Computational Linguistics* [3] Clark et al. (2019), *Journal of Organizational Behavior*

Presentation Content

Robot Behavior

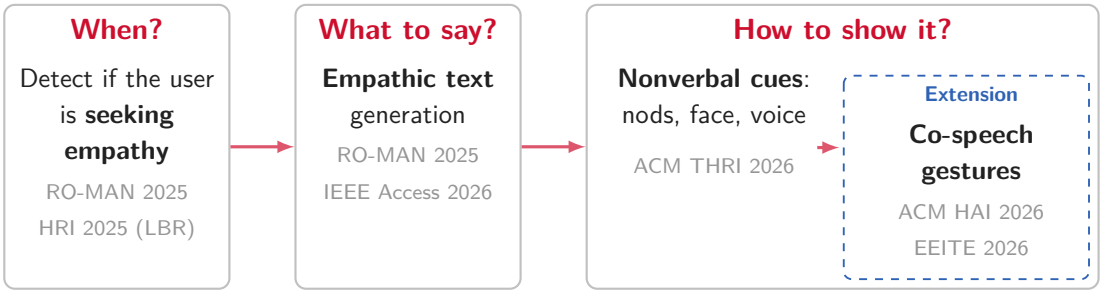
- **Multimodal Empathic Cues for Social Robots** — ACM THRI 2026 *(conditionally accepted)*
- **Rhythmic–Semantic Co-Speech Gestures** — ACM HAI 2026
- Efficient Emotion-Aware Iconic Gesture Prediction — EEITE 2026

Human Understanding

- **When and How to Express Empathy in Human-Robot Interaction Scenarios** — IEEE ROMAN 2025
- **Empathetic Robots Using Empathy Classifiers in HRI Settings** — ACM/IEEE HRI 2025 (LBR)
- Behavioral Empathy Classification in Dyadic Interactions — IEEE Access 2026

How the Pieces Fit

Each project answers one question an empathic robot must solve.



PART 1

Responding Empathically

When to Be Empathic

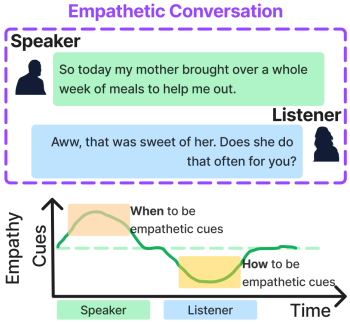
Respond empathically only when the user is **seeking empathy** [1, 2].

Why not always?

- **Empathy fatigue**: constant empathic exchanges can exhaust people [3–5]
- **Indiscriminate empathy**: LLMs may empathize even with harmful viewpoints [6]

How

- **Data-driven**: we learn the cues of empathy-seeking from public datasets [7–9]



[1] Arzate Cruz et al. (2025), *IEEE RO-MAN* [2] Arzate Cruz et al. (2025), *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* [3] Thomas et al. (2007), *Journal of General Internal Medicine* [4] Omdahl et al. (1999), *Journal of Advanced Nursing* [5] Nourriz et al. (2024), *Empathy-Centric Design Workshop* [6] Cuadra et al. (2024), *CHI* [7] Hosseini et al. (2021), *AAAI* [8] Hosseini et al. (2021), *Findings of EMNLP* [9] Montiel-Vázquez et al. (2024), *IEEE Access*

Empathic Text Generation

We developed methods to guide LLMs toward more empathic text through **integration of empathy cues** [1] using:

- **Variable conditioning** [2, 3]: the prompt tells the LLM **how** to respond

Our prompt asks for all three EPITOME mechanisms at the **strong** level.

[1] Montiel-Vázquez et al. (2025), *IEEE Access* [2] Arzate Cruz et al. (2025), *IEEE RO-MAN* [3] Arzate Cruz et al. (2025), *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*

How Is Empathy Expressed in Text?

Empathic responses combine **three mechanisms** [1].

Emotional reactions	Expressing feelings toward the person *"I feel sad for you"*
Interpretations	Showing understanding of their experience *"This must be terrifying"*
Explorations	Asking to learn more about their feelings *"Are you feeling alone right now?"*

[1] Sharma et al. (2020), *EMNLP*



How Robots Can Express Empathy

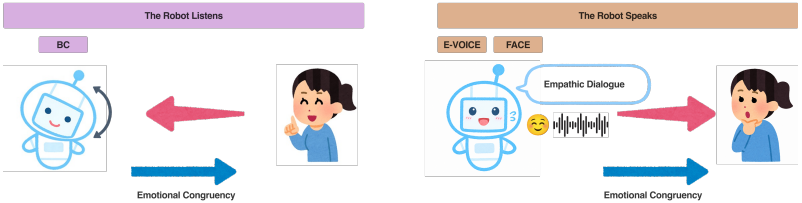
Which behaviors make sense depends on the robot's **role**: listening or speaking.

	Verbal	**Nonverbal**	
		Emotional	**Cognitive**
Listening user speaks	Verbal backchannels ("mm-hmm", "I see")	Facial mimicry of the user's emotion	**Head nods** **Forward posture** Attentive gaze
Speaking robot replies	**Empathic text generation**: acknowledge, validate, ask	**Facial & motor expressions** **Expressive voice** Co-speech gestures	Well-timed replies Turn-taking gaze

Arzate Cruz et al., "Multimodal Empathic Cues for Social Robots", ACM THRI 2026 (conditionally accepted).

Multimodal Cues

We propose and evaluate a **generic framework** for robots that use **backchanneling** (BC), **facial and motor expressions** (FACE), and **expressive voice** (E-VOICE) to express empathy multimodally.




Seeking Empathy: Not Detected

Multimodal Cues Impact — Study Design

How to show it?

Within-subjects, full factorial, 40 participants.

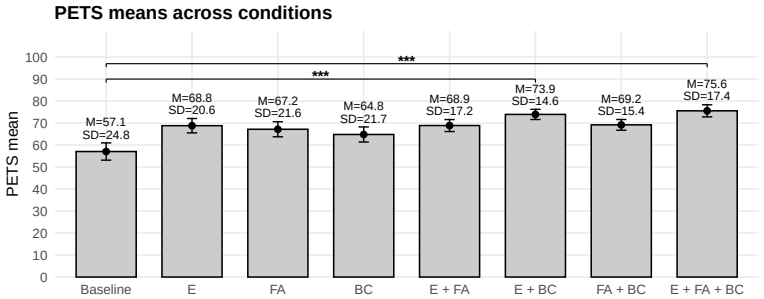
- **8 conditions** = BC × FACE × E-VOICE (baseline = all off)
- **8 emotional scenarios** from Plutchik, balanced with Latin squares
- 3 conversational exchanges per condition
- **Measure:** PETS, rescaled to $[0, 100]$ (ER = emotional responsiveness, UT = understanding & trust)
- **Analysis:** ART ANOVA (data non-normal) + ART-C contrasts
- Gaze identical in all conditions

Which cues matter? | Do they interact? | Which combination is best?

Results — Which Configuration Wins?

How to show it?

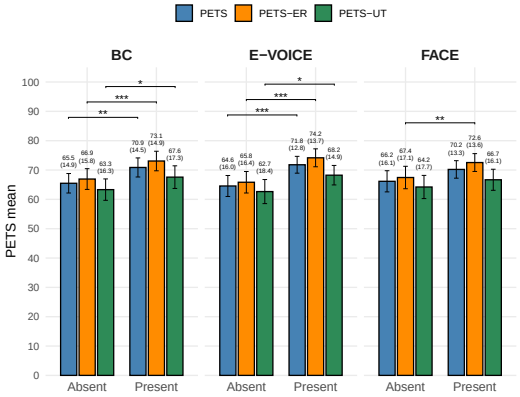
Only **combinations** beat the baseline, not single cues.



Baseline 57.1 → E-VOICE+BC 73.9 → all three 75.6.

Results — Which Cue Does the Work?

- **E-VOICE** and **BC** show significant main effects on perceived empathy.
- **FACE** is significant only on the emotional-responsiveness subscale — weaker, not absent.
- **No interactions:** effects are approximately additive.



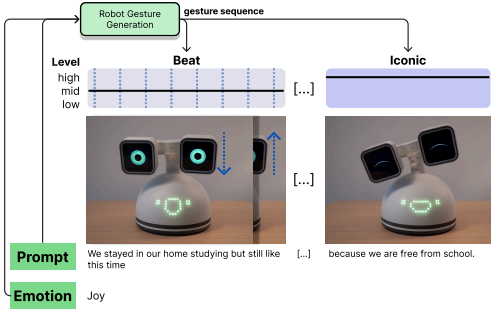
FACE was weakest: it came **late** (after a 4–5 s reply delay) and only at **discrete** moments.

PART 2

Expressing Emotion

Co-Speech Gestures

To keep emotion visible **throughout** the reply, we generate **iconic** and rhythmic **beat** gestures from text and emotion.



Training data

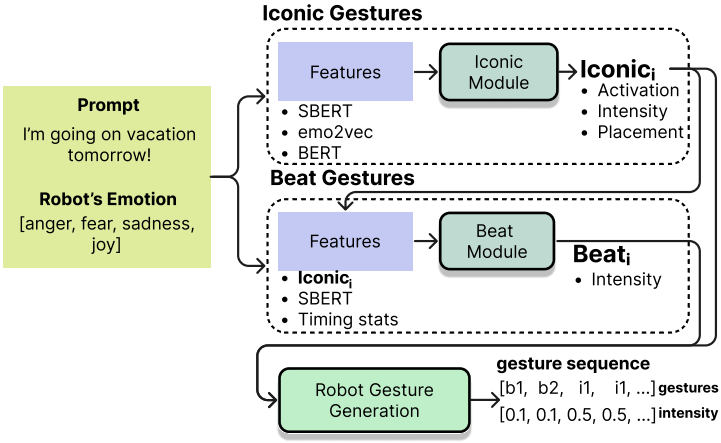
BEAT2

60 h of motion-captured co-speech gestures, 25 speakers, a wide range of emotions

Liu et al., CVPR 2024

The Method

Prompt + emotion → iconic module → beat module → gesture sequence.





User Study

Online, within-subjects, 60 participants (Prolific).

Baselines

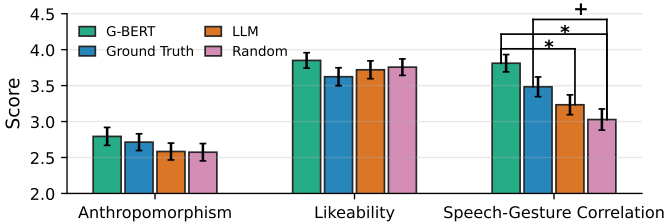
- Ground truth (BEAT2 labels)
- LLM prompting
- Random
- **G-BERT** (ours)

Measures (5-point bipolar)

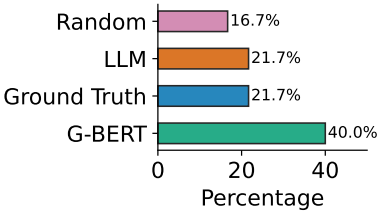
- Anthropomorphism
- Likeability
- Speech–gesture correlation
- Overall ranking

15 participants per emotion; method order randomized; attention checks applied.

Results



User study scores by method



Overall ranking

G-BERT wins on **speech–gesture correlation** and is ranked best by **40%** of participants; anthropomorphism and likeability are unchanged.

Summary

Responding empathically

When: only when the user **seeks** empathy.

How: mainly through **voice** and **listening behavior**; the cues add up.

Expressing emotion

Co-speech gestures keep emotion visible **while speaking**, improving speech–gesture alignment, and could raise **perceived empathy**.

Honda Research Institute

Slides, videos & papers

Christian Arzate Cruz