I Have No Mouth, Yet I Must Scream: Towards Situated Robot Acoustics in HRI

Eric Hansen
ejhansen@mines.edu
Colorado School of Mines
Golden, Colorado, USA

Ross Mead
ross@semio.ai
Semio AI, Inc.
Los Angeles, California, USA

Tom Williams
twilliams@mines.edu
Colorado School of Mines
Golden, Colorado, USA

ABSTRACT

While there has been a wealth of research on the syntax, semantics, and pragmatics of human-robot communication, there has been less attention paid to the *acoustics* of robot communication. In this work, we define two key problems central to the acoustics of HRI: (1) Proxemics-Guided Dyadic Acoustics, and (2) Acoustics-Guided Dyadic Proxemics. We present an initial mathematical solution to these problems, and present proof-of-concept validation of our solution to the Proxemics-Guided Dyadic Acoustics problem.

1 INTRODUCTION

Language-capable robots hold substantial promise in pro-social areas of national interest [18], including eldercare [2], education [5, 15], and disaster response [1]. While there has been a wealth of research on the syntax, semantics, and pragmatics of human-robot communication, there has been less attention paid to the *acoustics* (i.e., the properties of how sound is transmitted) of robot communication, a dimension of communication that is fundamentally entwined with a robot's situated, embodied nature.

Only recently have acoustic concerns begun to attract significant attention within the human-robot interaction (HRI) community; however, even these recent explorations have missed fundamental questions of robot acoustics that are most critical to real-world human-robot interactions, such as how loudly robots need to speak given the distance to their interlocutor and contextual factors, such as environmental interference, individual sensory differences, context-based acoustic norms, and privacy concerns.

In this work, we thus specifically define two key problems central to the acoustics of HRI:

Proxemics-Guided Dyadic Acoustics: Given the distance between the robot and its intended interactant, and how loudly it wishes its speech to be perceived by that interactant, *how loudly should the robot speak?*

Acoustics-Guided Dyadic Proxemics: Given how loudly a robot plans to speak, and how loudly it wishes its speech to be heard by an interactant, *where should the robot position itself with respect to that interactant?*

2 RELATED WORK

2.1 The Use of Sound in HRI

While there is a vast literature on natural language communication in robotics [18], only recently has there been an active community of researchers exploring the use of sound in HRI. Some of this work can be seen in a 2023 special issue on *Sound in HRI* in the *ACM Transactions on Human-Robot Interaction* [16], which focused on design strategies, evaluation methods, and applications for sound in HRI. For example, researchers publishing in that special issue began to explore the detection of social presence from acoustic features [3], designing the acoustics of robots' nonverbal cues [14, 20], motions [4, 12] and background soundscapes [19], and the design of sounds that communicate things *about* robots [7].

Also relevant to our aims is the robot proxemics work of Ross Mead [8–11], which explored where a robot should position itself so that it can hear interactants through its automated speech recognition system (rather than where it should position itself so it can be heard by an interactant, which is a focus of our work). Similarly relevant to our aims is the privacy-sensitive proxemics research of Sihui Li, who explored where a robot should position itself so that it *cannot* hear interactants, for privacy reasons [6]; however, Li's research assumed pre-defined privacy regions around interactants, rather than actually modeling the acoustic transmission of sound.

2.2 Relevant Acoustics Research Beyond HRI

There is a vast literature on sound design considerations outside of HRI, especially in areas like health and safety [17], as well as studies of issues like the impact of sound interference on conversations. Much of this work is tailored to understanding sound design in well-understood built environments, where factors such as sound directivity [13], and the size and absorption parameters of that environment; however, for mobile robots traveling through dynamic environments, these parameters might be difficult or impossible to accurately acquire or estimate.

3 TECHNICAL APPROACH

Proxemics-Guided Dyadic Acoustics and Acoustics-Guided Dyadic Proxemics each fundamentally involve relationships between (1) the distance between a robot and its interactant, (2) the volume at which the robot speaks, and (3) the volume at which the robot wishes to be heard by its interactant.

More formally, these problems involve a relationship between the *volume setting* V_R of the robot's speech output, the desired Sound Pressure Level (SPL) of the robot's speech at the human's position SPL_H, and the distance between human and robot $D_{H,R}$.

To solve the Proxemics-Guided Dyadic Acoustics problem, we must specifically ask: given the desired Sound Pressure Level SPL_H of the robot's speech at the human's position, as well as the distance between the robot and human $D_{R,H}$, how can we calculate V_R?

The first step is understanding, given SPL_H and $D_{H,R}$, what sound pressure level at the *robot's* SPL_R position will result in that desired SPL_H. To solve this, we can use Equation 1:

$$SPL_H = SPL_R + 20Log_{10}(\frac{r_2}{r_1}) \tag{1}$$

Here, r_1 is the distance at which SPL_H is to be measured (i.e., $D_{H,R}$), and r_2 is the distance at which SPL_R was determined from the source of the robot's speech.

Subtracting the second term from both sides, we get:

$$SPL_R = SPL_H - 20Log_{10}(\frac{r_2}{r_1}) \tag{2}$$

Our next step is understanding the relationship between a selected volume level V_R on the robot platform, and the Sound Pressure Level SPL_R that would be measured at standard distance r_2 from the robot at that volume level.

To solve this, we can use Equation 3, which calculates the SPL_R from V_R (operationalized as the percentage of the robot's maximum volume setting ω), and χ_{r_2}, a coefficient determined through calibration at standard distance r_2.

$$SPL_R = \chi_{r_2} \times ln(V_R) + \omega \tag{3}$$

Inverting Equation 3, we receive an equation for converting from a desired SPL_R at the robot to a volume setting V_R that will produce that Sound Pressure Level:

$$V_R = e^{\frac{SPL_R - \omega}{\chi_{r_2}}} \tag{4}$$

Combining Equations 2 and 4, we thus get an equation for the Proxemics-Guided Dyadic Acoustics problem:

$$V_R = e^{\frac{SPL_H - 20Log_{10}(\frac{r_2}{r_1}) - \omega}{\chi_{r_2}}} \tag{5}$$

Conversely and conveniently, we can rearrange the equation to solve the Acoustics-Guided Proxemics problem:

$$r_2 = \frac{r_1}{10^{\frac{SPL_H - \omega - \chi_{r_2} \times ln(V_R)}{20}}} \tag{6}$$

4 CALIBRATION

To employ either solution, parameters χ_{r_2} and ω must be fit through calibration. This can be achieved by positioning an Sound Level Meter (SLM) at distance r_2 from the robot's speakers, running a 440 Hz tone generator at 0%, 10%, 20%, ... 100% volume, recording a set (in our case, 20) of readings at each step, averaging the readings at each step after removing outliers, and then solving for χ_{r_2}.

Our own calibration procedures at a distance of $r_2 = 0.06m$ produced values of $\chi_{r_2} = 31.004$ and $\omega = 104.89$ ($r^2 = 0.9761$).

5 EVALUATION

To validate our approach, we performed tests with a mobile robot positioned at 1, 2, 3, and 4 meters from an SLM in a relatively clear, somewhat sound-isolated, square-shaped room of about $24m^2$. At each position, a 440 Hz sine pattern was played by the robot, with a desired SPL_H set to 60 dB_{SPL} (the approximate SPL of a human voice). Using Equation 5, SPL_R values of 84.4, 90.5, 94.0, and 96.5 dB_{SPL} were calculated, resulting in V_R values of 63.3%, 72.4%, 78.4%, and 82.9%, respectively. As shown in Fig 1, results were extremely accurate at 0 m, and slightly too loud at farther distances, with the robot overcompensating for the distance to its ostensible interlocutor. Also of note is the non-monotonic increase in measured SPL, indicating a potential shift in background noise that suggests a need for future replication of our experiments.

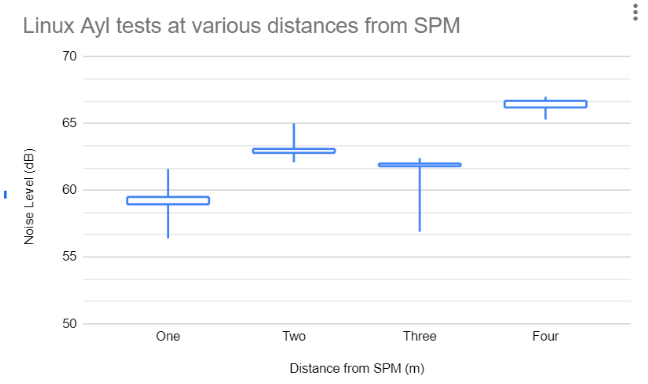

Figure 1: Preliminary Validation Results

6 CONCLUSION

We have defined two new problems that need tackling in the field of HRI: (1) Proxemics-Guided Dyadic Acoustics, and (2) Acoustics-Guided Dyadic Proxemics. We have briefly presented a technical definition of a solution to these problems, and a proof-of-concept highlighting the success of this method for the Proxemics-Guided Dyadic Acoustics problem. Of course, future work is needed to expand on these preliminary investigations, as well as to explore beyond the initial dyadic formulations of these problems to handle multi-agent contexts.

ACKNOWLEDGMENTS

This work has been supported in part by the Office of Naval Research grant N00014-21-1-2418.

REFERENCES

[1] Michael Beetz, Matthias Scheutz, and Fereshta Yazdani. 2017. Guidelines for improving task-based natural language understanding in human-robot rescue teams. In *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 000203–000208.

[2] Juan Fasola and Maja J Matarić. 2015. Evaluation of a spatial language interpretation framework for natural human-robot interaction with older adults. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 301–308.

[3] Nicholas C Georgiou, Rebecca Ramnauth, Emmanuel Adeniran, Michael Lee, Lila Selin, and Brian Scassellati. 2023. Is Someone There Or Is That The TV? Detecting Social Presence Using Sound. *J. Hum.-Robot Interact.* (aug 2023). https://doi.org/10.1145/3611658 Just Accepted.

[4] Adrian B Latupeirissa, Claudio Panariello, and Roberto Bresin. 2023. Probing Aesthetics Strategies for Robot Sound: Complexity and Materiality in Movement Sonification. *ACM Transactions on Human-Robot Interaction* (2023).

[5] Iolanda Leite, Marissa McCoy, Monika Lohani, Daniel Ullman, Nicole Salomons, Charlene Stokes, Susan Rivers, and Brian Scassellati. 2015. Emotional storytelling in the classroom: Individual versus group interaction between children and robots. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 75–82.

[6] Sihui Li, Sriram Siva, Terran Mott, Tom Williams, Han Zhang, and Neil Dantam. 2023. Command Rejection in Privacy-Sensitive Contexts: An Integrated Robotic System Approach. In *Proceedings of the IEEE International Symposium on Robot-Human Interactive Communication (RO-MAN)*.

[7] Maria Mannone, Valeria Seidita, and Antonio Chella. 2023. The Sound of Swarm. Auditory Description of Swarm Robotic Movements. *ACM Transactions on Human-Robot Interaction* (2023).

[8] Ross Mead and Maja J Matarić. 2015. Toward Robot Adaptation of Human Speech and Gesture Parameters in a Unified Framework of Proxemics and Multimodal Communication. In *Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA 2015) Workshop on Machine Learning for Social Robotics*.

[9] Ross Mead and Maja J Matarić. 2016. Perceptual models of human-robot proxemics. *Experimental Robotics. Springer Tracts in Advanced Robotics (STAR)* 109 (2016), 261–276. https://doi.org/10.1007/978-3-319-23778-7_18

[10] Ross Mead and Maja J Matarić. 2016. Robots Have Needs Too: How and Why People Adapt Their Proxemic Behavior to Improve Robot Social Signal Understanding. *Journal of Human-Robot Interaction (JHRI)* 5, 2 (2016), 48–68. https://doi.org/10.1007/s10514-016-9572-2

[11] Ross Mead and Maja J Matarić. 2017. Autonomous human–robot proxemics: socially aware navigation based on interaction potential. *Autonomous Robots* 41, 5 (2017), 1189–1201.

[12] Elias Naphausen, Andreas Muxel, and Jan Willmann. 2023. New Design Potentials of Non-mimetic Sonification in Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction* (2023).

[13] Manuelina Porto Nues Navarro and Roberto Leal Pimentel. 2007. Speech interference in food courts of shopping centres. *Applied Acoustics* 68, 3 (March 2007), 364–375. https://doi.org/10.1016/j.apacoust.2006.01.006

[14] Bastian Orthmann, Iolanda Leite, Roberto Bresin, and Ilaria Torre. 2023. Sounding Robots: Design and Evaluation of Auditory Displays for Unintentional Human-Robot Interaction. *J. Hum.-Robot Interact.* (aug 2023). https://doi.org/10.1145/3611655 Just Accepted.

[15] Aditi Ramachandran, Chien-Ming Huang, Edward Gartland, and Brian Scassellati. 2018. Thinking aloud with a tutoring robot to enhance learning. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*. 59–68.

[16] Frederic Robinson, Hannah Pelikan, Katsumi Watanabe, Luisa Damiano, Oliver Bown, and Mari Velonaki. 2023. Introduction to the Special Issue on Sound in Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction* (2023).

[17] Samuele Schivaoni and et all. 2022. Airborne Sound Power Levels and Spectra of Noise Sources in Port Areas. *Int. J. Environ. Res. Public Health* 19, 17 (Aug. 2022). https://doi.org/10.3390/ijerph191710996

[18] Stefanie Tellex, Nakul Gopalan, Hadas Kress-Gazit, and Cynthia Matuszek. 2020. Robots that use language. *Annual Review of Control, Robotics, and Autonomous Systems* 3 (2020), 25–55.

[19] Ziming Wang, Ziyi Hu, Björn Rohles, Sara Ljungblad, Vincent Koenig, and Morten Fjeld. 2023. The Effects of Natural Sounds and Proxemic Distances on the Perception of a Noisy Domestic Flying Robot. *ACM Transactions on Human-Robot Interaction* (2023).

[20] Brian J Zhang and Naomi T Fitter. 2023. Nonverbal Sound in Human-Robot Interaction: a Systematic Review. *ACM Transactions on Human-Robot Interaction* (2023).