SonifyIt: Towards Transformative Sound for All Robots

Brian J. Zhang[1], Noel Sigafoos[1], Rabecka Moffit[1], Ibrahim Syed[2], Lili S. Adams[1],
Jason Fick[2], and Naomi T. Fitter[1]

***Abstract*— Transformative robot sound yields perceptual, functional, and social benefits in human-robot interactions, but broader research and implementation related to this topic is impeded by the lack of a common sound generation system for robots. Such a system could enable a wide array of situated robot sound studies, smoother collaborations with sound designers than current state of the art methods, and broader adoption of transformative robot sound. Based on other successful open-source projects in the robotics community, we integrated Robot Operating System, a popular robotics middleware, and Pure Data, a visual programming language for multimedia, to enable live sound synthesis and sample playback for robots. This sound generation system synthesized sound in an in-the-wild pilot study with positive qualitative results. Furthermore, an online within-subjects survey study with $N = 96$ showed that the proposed sound system made the robot seem warmer, happier, and more energetic. This work benefits robotics researchers by providing the current sound system as a validated artifact and demonstrating its potential impact on broader robotics applications. We plan to develop this software into an open-source package: SonifyIt.**

I. INTRODUCTION

Transforming a robot's sound profile benefits human-robot interaction (HRI) perceptually, such as by making robots seem socially warmer and more competent [1]; socially, by conveying emotions both independently and in tandem with affective behaviors [2], [3]; and functionally, by helping humans locate hidden robots and understand a robot's intent to stop [4], [5]. For roboticists hoping to implement transformative robot sound, however, little exists in the way of practical guidance or tools. Instead, each robotics research group has developed customized approaches to generating sound, as noted in recent work [6]. In this work, *we aimed to take a first step towards a general tool for designing and deploying transformative sound for robots*.

Preliminary details on sound generation techniques have emerged from prior work on robot sound. Online studies, including our own past robot sound work, primarily use video editing to integrate hand-designed sound samples with robot recordings [1], [7], [8]. On real robots, roboticists have implemented the canned solution of manually cued sample playback [4], [5], [9]. For more complex and variable sound, prior work has proposed sonification based on modulation matrices [10] and live synthesis using parameterized commands [11]. However, the proposed systems have only partial validation and did not publicly release their tools. Thus, the research community cannot fully benefit from these systems.

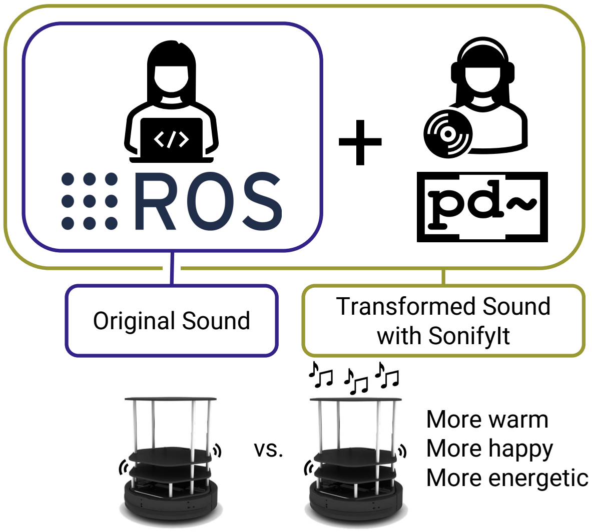

Fig. 1: An overview of the software components of SonifyIt, the system introduced in this work. As highlighted in the lower right, the results of experimental validation confirm the effectiveness of the system's sonification.

This work presents the design and initial validation of SonifyIt, the first publicly available sound generation system for robots. We release an end-to-end working demonstration with this manuscript and plan to roll out the full package, with expanded validated features, in a future publication. We identified current methods for robot sound generation in Section II. Section III presents the system design, informed by prior efforts and work in adjacent fields. To validate the system design, we first conducted an in-the-wild pilot study as a technical test and to gather qualitative feedback, which was generally positive, as explained in Section IV. Next, Section V details an online survey-based study, which confirmed that the system generated sounds with similar effectiveness as overlaid hand-designed sounds used in our previous work, as shown in Figure 1. Lastly, Section VI discusses the validation results, implications for research and practice in robot sound, and future directions. The main contribution of this work is *an early validated demonstration of our robot sound generation system SonifyIt, which will ease future research and use of transformative robot sound.*

II. RELATED WORK

We surveyed recent work on sound in HRI, closely examining methods related to transformative robot sound, which

---

[1] Collaborative Robotics and Intelligent Systems (CoRIS) Institute, Oregon State University, Corvallis, Oregon, USA. `[zhangbr, sigafoof, moffitr, adamsli, fittern]@oregonstate.edu`

[2] Music Department, Oregon State University, Corvallis, Oregon, USA. `[syedi, fickj]@oregonstate.edu`

we define as non-linguistic sounds added to complement a robot's original sonic profile [1].

A. Transformative Robot Sound Studies

Online survey-based studies comprise a significant portion of recent robot sound research. In a study using the Fetch Mobile Manipulator, adding musical sound improved perceptions of a robot compared to harmonic or mechanical transformative sounds [7]. A study using the NAO, a humanoid social robot, showed that participants generally liked more complex sounds (compared to a simpler pattern and the original sound profile) during robot motions [8]. Another NAO-focused study blended additional sounds with the original sound of motions in affective actions to make the behaviors more recognizable [12]. Our prior work showed that transformative sound made five robots of different archetypes consistently seem happier, more energetic, warmer, and more competent [1]. These past online study examples used sound edited onto recordings of robots, including free adjustment of timing and cueing, which does not translate easily to real-world robots.

In-person studies have addressed practical aspects of situated sound generation by robots through manually cued sound samples. For example, a study using the Baxter collaborative robot (cobot) cued music to play while a participant walked by, which made the robot seem more calming, gentle, soft, smooth, friendly, and pleasant [9]. In a study on proxemics with a mobile robot, the robot approached participants while playing different transformative sounds, causing changes in perceived anthropomorphism, animacy, likeability, intelligence, and safety [5], [13]. An interaction during which a mobile robotic teammate moved behind a curtain while playing broadband or tonal sounds increased participants' accuracy, inference speed, perceived noticeability, and perceived localizability compared to when no sounds were played [4]. Lastly, non-linguistic utterances played during idle behaviors on the NAO conveyed recognizable emotions to people [3]. These works played back samples on cue to generate sound, though limited detail was provided on the sound generation systems.

These prior works focused on the effects of transformative sound and used simple and reliable methods to create and study audio manipulations. However, these methods require extensive effort to implement for other robots, as each sound must be manually designed and cued by roboticists. Instead, *we aim to provide a system that can integrate with a robot's existing software to automatically generate sounds*.

B. Robot Sound Generation Methods

Several works have proposed general sound generation systems or provided descriptions of specific sound generation artifacts. Schwenk et al. proposed a robot sonification system using SuperCollider[1], a programming language for audio synthesis and algorithmic composition, to produce sonification [10]. Sonification is the use of non-speech audio to represent information; it can be a form of transformative sound when blended with the original sound of a robot [7], [12]. As a demonstration, the system from Schwenk et al.'s work used joint state and sensor data as inputs to a modulation matrix, which fed into the synthesizer [10]. However, this work focused on the system design rather than validation and did not release the source code. Ritschel et al. also used SuperCollider for sound synthesis with the intent of studying robot sound personalization [6]. Unlike Schwenk et al., Ritschel et al. did not incorporate sonification but instead studied sounds for conveying emotions and intentions. Thus, the sound generation system did not use data from the robot as we propose to do in this work. Furthermore, the source code was not released.

Fernandez et al. proposed the "sonic expression system," a sound generation system for robots that included a synthesizer based on Pure Data (Pd)[2], an open-source visual programming language for multimedia and music creation [11]. This system generated "quasons," or individual notes with a set amplitude, frequency, timbre, and duration, and the software combined quasons into utterances meant to convey emotions, which were manually cued. Validation on three different robots showed mixed results, as participants found some utterances difficult to distinguish. The source code was partially displayed as a figure in the paper manuscript. Zahray et al. used Max/MSP[3], a proprietary visual programming language for multimedia and music creation, to evaluate different robot sonification methods [14]. Unlike in other systems, the robot was controlled using the same programming language as the sound generation system. As in other works, the source code was not released.

These works propose sound generation systems ranging from platform- and study-specific tools to generalizable strategies for all robots, but no work has provided the source code and tutorials necessary for system replication. Furthermore, generalizable systems are not yet comprehensively evaluated, and the evaluations that do exist do not clearly demonstrate that the systems could produce sounds with the same effectiveness as those in prior work. We aimed to *address the gap in practical guidance for robot sound generation systems and ensure that any proposed system met expectations in sound effectiveness from prior work.*

III. SYSTEM DESIGN

While pursuing robot sound research, the authors of this paper, who span the robotics and music technology disciplines, sought a way to rapidly design and iterate on transformative robot sound. Based on our needs and team discussions, we identified three criteria for a new sound generation system that would serve the robotics community:

1) The system should be capable of live sound synthesis and sound sample playback, preferably interchangeably.
2) The system should be accessible to sound designers for ease of collaboration.

---

[1] `https://github.com/supercollider/supercollider`

[2] `https://github.com/pure-data/pure-data`

[3] `https://cycling74.com/`

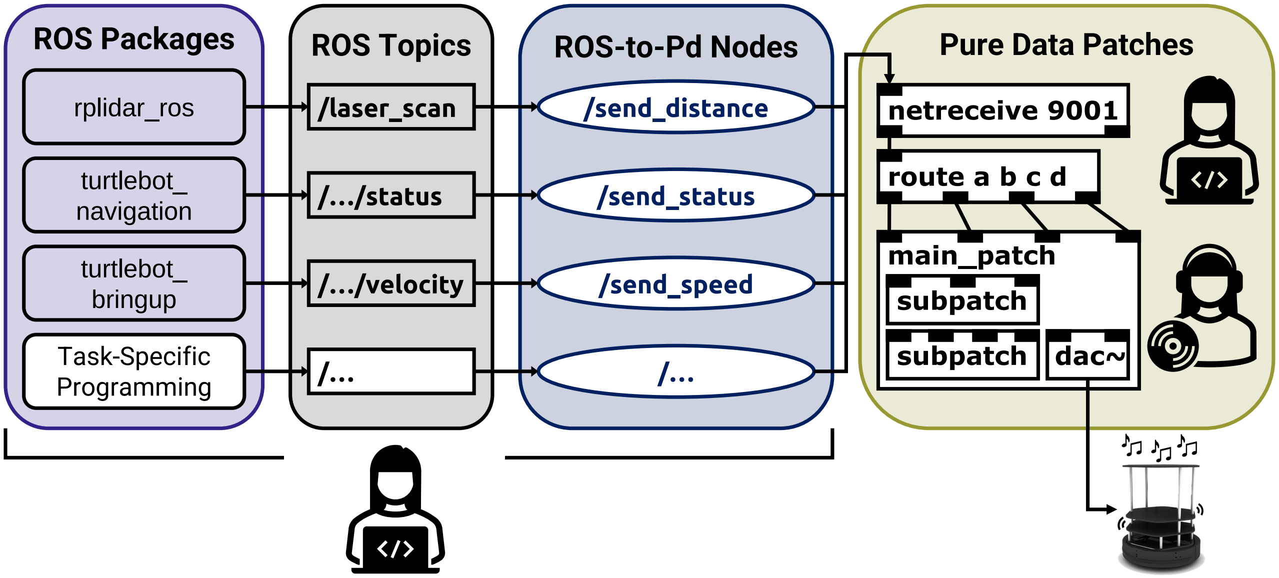

Fig. 2: System diagram showing the separation of programs, data, and responsibilities for the SonifyIt package. Shapes shaded in white require input from roboticists or sound designers for the specific task at hand, while shapes filled with color are open-source contributions within the current demonstration, to be expanded in future work. Shapes filled with "/..." stand for custom data in topics and nodes. After receiving and labeling data via `netreceive` and `route`, a sonification patch or patches can create sound via `dac`. The roboticist and sound designer icons appear by their respective responsibilities.

3) All components of the system should be open-source, preferably working with existing open-source projects.

In particular, we aimed to emulate MoveIt[4], software for robot manipulation natively integrated with Robot Operating System (ROS)[5], a popular robotics middleware. An example roboticist seeking to use MoveIt for a robotic arm would download the main MoveIt software and a package specific to the robotic arm. Then, the roboticist could control the arm through MoveIt using a provided programming interface.

For our planned SonifyIt package, a roboticist would similarly install the main software and a package specific to the robot of interest. In the simplest case, community-contributed sound profiles generated based on ROS topics, or data streams made available to other programs through ROS, would allow a roboticist to add transformative sound without additional programming. For more customized or complex scenarios, a roboticist would modify the robot-specific SonifyIt package to meet their needs. Figure 2 shows how useful ROS topics originate from built-in ROS packages.

In order to meet more complex needs, SonifyIt should allow for live sound synthesis as in [10] and [11]. Synthesis enables greater variation of resulting sounds (i.e., samples do not need to be generated beforehand) as well as direct sonification using robot data. Accessibility for sound designers would also expand the design space and help roboticists collaborate with others who possess ingrained sound design skills. Lastly, to encourage use and community contribution, SonifyIt should be free, open-source software. To begin, we selected an appropriate sound engine in Section III-A, integrated it with robotics software in Section III-B, and developed an initial demonstration in Section III-C, which we release with this paper.

A. Sound Engine Selection

Based on these criteria, we surveyed existing sound synthesis and playback options for robots. After discarding projects that had not received updates within the past two years, three categories emerged:

1) Packages, libraries, or plug-ins for programming languages or software (sound_play[6], openFrameworks[7]).
2) Non-visual programming languages (SuperCollider, Csound[8], Sonic Pi[9]).
3) Visual programming languages (Pure Data, Noisecraft[10]).

The sound synthesis requirement eliminated options such as the `sound_play` package, which natively integrates with ROS but allows only for speech synthesis and sample playback. For sound designer accessibility, we favored visual programming languages, as visual programming languages are designed for novices and may enable learning more effectively when compared to traditional programming languages [15], [16]. Lastly, between Pure Data (Pd) and Noisecraft, Pd integrates more easily with robots as it may be run headlessly (i.e., with no monitor or peripherals) from a shell; Noisecraft is accessed through an Internet browser, which poses difficulties for robotic systems without a desktop

---

[4] `https://moveit.ros.org/`

[5] `https://www.ros.org/`

[6] `https://wiki.ros.org/sound_play`

[7] `https://github.com/openframeworks/openFrameworks`

[8] `https://github.com/csound/csound`

[9] `https://github.com/sonic-pi-net/sonic-pi`

[10] `https://github.com/maximecb/noisecraft`

environment or without Internet access. Thus, we selected Pd as the sound engine for SonifyIt, in a similar way to [11].

Pd patches, or programs, retain important features from traditional programming languages such as subpatching, which is equivalent to function definition. We defined a patch hierarchy including a communication patch, a main control patch, and sound generation subpatches, as seen in Figure 2.

B. ROS Integration

We integrated Pd with ROS such that Pd may effectively "subscribe" to ROS topics through ROS-to-Pd nodes. This integration allows for the transfer of raw or processed data and manual triggers for sound generation. Figure 2 shows a system diagram of the integration and data flow from ROS packages to Pd patches.

To communicate between ROS and Pd, we opted to transmit data over local loopback using the Fast Universal Digital Interface (FUDI) networking protocol. This protocol sends all messages as strings, leading to reduced efficiency for numerical data, but allowing for easier message construction compared to Open Sound Control (OSC). Further performance improvements may be found by using interprocess shared memory libraries, a method used by ROS nodelets, though ROS-to-Pd nodes should avoid transmitting significant amounts of data such as images or point clouds to convey state and sensor information.

Flexibility exists as to the location of processing of ROS topic data to discrete system states, which could include functional or social descriptors such as "moving," "happy," or "alert." For example, a roboticist could add programming to manually publish the state; ROS-to-Pd nodes could classify the state from ROS topics before sending a state to Pd; or the state could be interpreted in Pd before determining sound generation. Similarly, ROS topic messages may contain extraneous or irrelevant data, which may be removed by ROS-to-Pd nodes or be allowed to continue into Pd. To reduce programming burden as much as possible, we recommend processing this information in ROS-to-Pd nodes, which can be distributed in a robot-specific SonifyIt package and activated or deactivated as needed.

C. Provided Demonstration

As part of the provided demonstration and for initial validation, we designed sound for a TurtleBot 2 using SonifyIt. Figure 2 shows the selected ROS topics for state interpretation and sonification, as well as the division of responsibilities for roboticists and sound designers. The multidisciplinary author team followed role breakdowns to create two sound designs: a harmonic sound profile using live synthesis and a musical sound profile using sample playback, where harmonic and musical follow the term conventions defined in [7]. The harmonic sound profile allowed live adjustment in pitch, tempo, and other characteristics. The Pd patches for these sound profiles, as well as accompanying tutorials, are available in the SonifyIt repository[11].

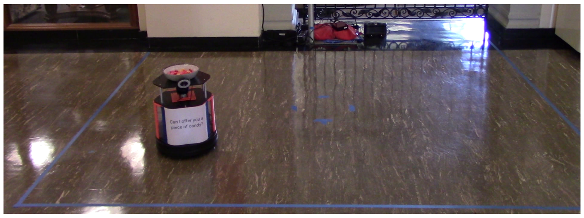

Fig. 3: The in-the-wild study environment and robot.

IV. IN-THE-WILD SYSTEM VALIDATION

We conducted an in-the-wild pilot study to validate the technical capability of SonifyIt and gather qualitative feedback. The procedures for the pilot study were approved by Oregon State University (OSU) under protocol #IRB-2019-0068.

A. Participants

We allowed 22 passers-by to self-select into the study and also randomly sampled 6 staff and other occupants of Oregon State University's Memorial Union building for a total of $N = 28$ participants. No demographic data was collected.

B. Study Design

We deployed the TurtleBot 2 with the harmonic sound profile in a marked study space in Oregon State University's Memorial Union building during normal business hours for one week. The space covered an approximate square 3m×3m large and had signage in English and Spanish to briefly describe the study and informed consent process. Prospective participants read the signage and entered the space to consent and participate in the study. Figure 3 shows the study setup.

The robot ambiently navigated in a U-shaped pattern around the front and sides of the study space, producing consequential sound and harmonic sound from SonifyIt. If a participant entered the space, the robot would detect the human using a 2D Lidar sensor, stop its ambient navigation, and follow the participant within the space. The following routine included stopping when the participant was sufficiently close to the robot. If multiple participants entered the space simultaneously, the robot followed the closest person. Participants were encouraged to take a piece of candy from a bowl on top of the robot by signage on the robot.

After interacting with the robot, participants were asked if they would like to complete a brief recorded interview. On the final day of the study, we solicited interviews from bystanders working in the building who had not yet directly interacted with the robot in the study space to gauge responses to the robot from the perspective of passive and long-term interaction. Participant feedback was solicited in a semi-structured interview format centered around the questions:

- What made you want to interact with the robot?
- Did anything about the robot stand out to you?
- Do you have any other thoughts about the robot?

For bystanders, we asked about how long they had been within the space and whether they had noticed the robot in lieu of question 1 above.

---

[11] `https://github.com/shareresearchteam/sonify_it`

Fig. 4: A cropped keyframe from the online study stimuli.

C. Results

While not all participants mentioned sound, 7 of those who did described the sound as "cute," "nice," "good," "relaxing," "interesting," or "approachable," though 3 users found the sound profile "weird" or "repetitive." Participants who were asked how they might change the sound of the robot mentioned increasing the variability of the sound and including more responsive sounds, such as acknowledging when a person entered the space. On a technical level, SonifyIt enabled modifications to the harmonic sound during the setup process that smoothed the study administration experience, such as changes in pitch and volume to avoid resonance and excess loudness in the study space. No technical issues occurred with the SonifyIt package.

V. ONLINE SURVEY EXPERIMENTAL VALIDATION

We conducted an online survey-based study to determine whether the musical sound profile automatically produced by SonifyIt would yield similar effects as manually overlaid transformative sounds in prior studies [1]. The procedures for the online study were approved by Oregon State University under protocol #IRB-2019-0172. As in [1], we compared the robot with and without transformative sound; however, this iteration of the study used SonifyIt to automatically generate the transformative sound. We carried over the hypothesis:

H1: Adding transformative robot sound will lead to improved perceptions of robot valence, energy level, warmth, and competence.

A. Participants

We randomly selected $N = 96$ adults from the United States via Prolific, an online subject pool similar to Amazon Mechanical Turk (MTurk) but with higher data quality [17]. This sample size mimics the sample size from [1] to support comparisons. Participants ranged between 18 and 82 years of age ($M = 37.8$, $SD = 14.0$), with 55.2% men and 44.8% women. Approximately half (51.0%) of participants had "higher education, formal training, hobbies, or employment" in science, technology, engineering or mathematics, while a smaller proportion (21.9%) had a similar level of experience in music. Participants were screened to ensure normal or corrected-to-normal vision and normal hearing. We compensated participants with USD 3.75.

B. Study Design

Using a Canon Vixia HF R800 camera and a Dayton Audio UMM-6 microphone, we recorded a TurtleBot 2 robot navigating through four motions in an office environment, with and without transformative sound, for a total of eight stimuli. The motions were replicated from [1], but the current study used live generated sound from SonifyIt rather than overlaid hand-designed sounds. Figure 4 shows a frame of a stimulus video; all of the stimuli are available in [18] and a pair of stimuli are featured in the accompanying video.

Using these stimuli, we developed a within-subjects study in which participants completed a 15-minute online survey with the same format as in [1]. After providing informed consent, participants completed an introductory module to calibrate their audio device volume, be introduced to the TurtleBot 2, and respond to a calibration stimulus using the measures described in Section V-C. Next, participants viewed the eight stimuli in a semi-random order such that neither any motion nor sound condition would appear twice in a row. After each stimulus, participants completed the post-stimulus perception questionnaire, and two attention checks were included among the stimuli. Lastly, participants completed a free-response question, manipulation check, attitudes questionnaire, and demographic questionnaire.

C. Measures

The survey questions included the following:

- Post-stimulus perception questionnaire: after each stimulus video, we administered the Robotic Social Attributes Scale (RoSAS) to capture participant perceptions of the robot's *warmth*, *competence*, and *discomfort* by averaging sets of six component attribute responses [19]. Participants rated how much the robot was or was not associated with each attribute on six-point Likert scales. The "happy" component of *warmth* and the appended attribute "energetic" captured participant perceptions of the robot's *valence* and *energy level* (arousal) from the circumplex model of affect [20].
- Free-response question: after responding to all stimuli, participants described the most important factors underlying their responses in at least 200 characters.
- Manipulation check: participants watched two side-by-side clips of the robot, one with and and one without added sound, and were asked to identify which had more sound. The check was failed if the participant answered that both videos were the same.
- Attitudes questionnaire: we used the Negative Attitudes Towards Robots Scale (NARS) to capture general attitude towards robots. Participants indicated how much they agreed or disagreed with fourteen statements on seven-point Likert scales [21].
- Demographic questionnaire: several questions recorded demographic and occupational information, including technical and musical experience.

D. Analysis

We analyzed post-stimulus questionnaire responses using repeated-measures analysis of variance (rANOVA) tests with a Greenhouse-Geisser sphericity correction and an $\alpha = 0.05$ significance level. Each 2×4 rANOVA used factors of *sound condition* and *robot behavior*. As a non-parametric equivalent for two-factor rANOVA tests could not be easily found, we did not test for normality and relied on the rANOVA's robustness to normality assumption violations [22]–[24]. For rANOVAs that indicated a significant effect due to *sound condition*, we applied post-hoc pairwise comparisons with a Holm-Bonferroni correction for Type I errors to determine whether sound condition responses differed for each of the four *robot behavior* levels. We reported effect sizes using η^2_G [25], which is comparable to η^2. $\eta^2 = 0.010$ is considered a small effect, $\eta^2 = 0.040$ a medium effect, and $\eta^2 = 0.090$ a large effect [26]. We used jamovi 2.2, a graphical statistics software using R, for all analyses [27]–[30].

E. Results

rANOVAs showed that transformative sound led to higher *valence* ($p < 0.001$, $F(1, 95) = 17.21$, $\eta^2_G = 0.014$), *energy level* ($p = 0.001$, $F(1, 95) = 10.75$, $\eta^2_G = 0.004$), and *warmth* ($p < 0.001$, $F(1, 95) = 22.78$, $\eta^2_G = 0.010$). For these three measures, Holm-Bonferroni tests showed significant pairwise differences between the original and sonified sound conditions in three of four contexts for *valence*, one of four contexts for *energy level*, and four of four contexts for *warmth*. Figure 5 shows response distributions for all measures and the post-hoc analysis results.

VI. DISCUSSION

Our SonifyIt demonstration created sonification that could be adjusted live and controlled through processed data streams from existing open-source packages. This initial tool can be found in the SonifyIt repository in Section III-C. This flexible sound generation system operated without error during the in-the-wild pilot study. Participants of the in-the-wild pilot study said "I liked his little music" and "it's got a cute song." Repetition had mixed effects, with one participant stating "it's pleasant... relaxing, like repetitive," but another noting "I'd have it less constant because you tune it out after a while." Participants could listen to the robot for an extended period of time; a difference in listening duration may have led to diverging impressions of the repetition.

The experimental validation results mostly supported **H1**. As expected, valence, energy level, and warmth significantly increased. While competence did not significantly increase, the overall effects of the musical sound profile trended with transformative sound effects seen in prior work. Participants noted that "the music helped the robots seem a little friendlier," though free responses also remarked on imperfections, e.g., "the music stopping and starting... was a little bit awkward." Some participants gleaned information from the transformative sound, saying that "the motor sounds coupled with the playing music seems to indicate... that I should make an effort... by getting out of the way." While effect sizes were small, effects of this size are "small at the level of single events but potentially more ultimately consequential," particularly for repeated or one-to-many interactions, and thus remain important [26].

A. Design Implications

The development of the sound profiles using SonifyIt revealed anecdotal insights on the collaborative sound design process. During this process within our team, communication about desired sonification and available data streams proved critical. Based on our experiences, we recommend that roboticists clearly communicate possible data streams, which aids with ideation on possible sonification methods. On the other side, sound designers should describe potential sound designs and the conditions in which they might occur, which can guide roboticists in determining how to convey information from ROS to Pd. Figure 2 indicates the responsibilities of the roboticists and sound designers in a SonifyIt implementation.

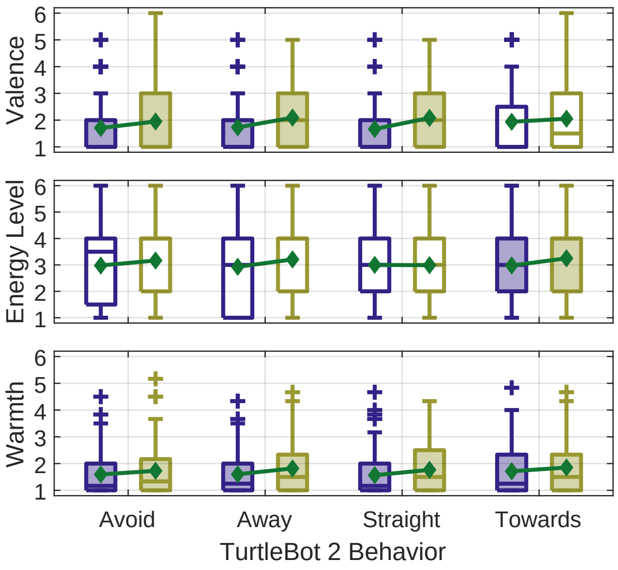

Fig. 5: Results of post-stimulus responses with significant differences. Boxplots include boxes from the 25th to the 75th percentiles, whiskers up to 1.5 times the inter-quartile range, and "+" marks to indicate outliers. *Original* stimulus responses in purple boxplots are placed to the left and *sonified* stimulus responses in yellow boxplots are placed to the right of each *robot behavior* gridline. Boxplot pairs that are filled in indicate significant pairwise differences. Green lines connect the means of each boxplot pair.

Our proposed SonifyIt tool is intended to enable robot sound design and help roboticists and sound designers to more effectively work together and iterate on sound profiles. Based on our evaluations to date, we conjecture that avoiding excess repetition (e.g., by using random variations during synthesis) and increasing sound responsiveness to humans (e.g., using sensor data) may be useful techniques to incorporate in successful sound designs. Sound designers may also benefit from live adjustments to sound designs through synthesis parameters or filtering to suit the audience, to fit the acoustics of a space, or simply to speed up the sound design process. Even with a relatively simple SonifyIt application like the artifact released with this paper, robots can be made to seem more pleasant, energetic, and warm.

B. Key Strengths & Limitations

The strengths of this work include its novelty, as no comparable system to SonifyIt yet exists, and its validation in the form of an in-the-wild pilot study and online survey study, which confirm the proposed tool's technical viability and efficacy for HRI. The structure of SonifyIt enables others to develop and release sound designs for different robots as open-source packages, which is intended to quicken adoption of transformative robot sound. The two forms of validation help to compensate for potential weaknesses, such as a lack of embodied interactions in the online study and insufficient statistical power in the in-person study.

The preliminary state of the SonifyIt artifact, which currently centers on sound designs for the TurtleBot 2, limits the immediate usefulness of this work. Our future expansion of SonifyIt will increase its ability to support the needs of those who aim to implement and study transformative sound. We aim to incorporate key features of prior works, such as the modulation matrix in [10] and the quason generator in [11]. We will also develop additional sound profiles for robots of different archetypes and gather SonifyIt repositories for new robots contributed by community members.

C. Conclusion

In this work, we introduced SonifyIt, a sound generation system for robots that integrates with ROS and uses Pd to perform live synthesis and sample playback based on sensor and state information. Pilot system validation demonstrated the ease of live adjustments and was received well in participant interviews, while experimental validation showed that SonifyIt produces transformative robot sound that is effective at improving HRI. Robot sound researchers will benefit from using SonifyIt to create and share sound designs with the broader robotics community, and SonifyIt may eventually help all robots sound a little better.

ACKNOWLEDGEMENTS

We thank Murphy James for helping with interview transcription and Deb Mott, Myungwoo Crossler-Laird, and the OSU Memorial Union staff for aiding the in-the-wild study.

REFERENCES

[1] B. J. Zhang, N. Stargu, S. Brimhall, L. Chan, J. Fick, and N. T. Fitter, "Bringing WALL-E out of the Silver Screen: Understanding How Transformative Robot Sound Affects Human Perception," in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2021.

[2] R. Read and T. Belpaeme, "Situational Context Directs How People Affectively Interpret Robotic Non-Linguistic Utterances," in *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, 2014, pp. 41–48.

[3] ——, "People Interpret Robotic Non-linguistic Utterances Categorically," *Int. Journal of Social Robotics*, vol. 8, no. 1, pp. 31–50, 2016.

[4] E. Cha, N. T. Fitter, Y. Kim, T. Fong, and M. J. Matarić, "Effects of Robot Sound on Auditory Localization in Human-Robot Collaboration," in *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, 2018, pp. 434–442.

[5] M. Lohse, N. van Berkel, E. M. A. G. van Dijk, M. P. Joosse, D. E. Karreman, and V. Evers, "The influence of approach speed and functional noise on users' perception of a robot," in *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2013, pp. 1670–1675.

[6] H. Ritschel, I. Aslan, S. Mertes, A. Seiderer, and E. André, "Personalized Synthesis of Intentional and Emotional Non-Verbal Sounds for Social Robots," in *Proc. of the Int. Conf. on Affective Computing and Intelligent Interaction (ACII)*, 2019, pp. 1–7.

[7] F. A. Robinson, M. Velonaki, and O. Bown, "Smooth Operator: Tuning Robot Perception Through Artificial Movement Sound," in *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, 2021, pp. 53–62.

[8] A. B. Latupeirissa, P. Claudio, and R. Bresin, "Exploring Emotion Perception in Sonic HRI," in *Proc. of the Sound and Music Computing Conf. (SMC)*. Zenodo, 2020, pp. 434–441.

[9] G. Trovato *et al.*, "The Sound or Silence: Investigating the Influence of Robot Noise on Proxemics," in *Proc. of the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, 2018, pp. 713–718.

[10] M. Schwenk and K. O. Arras, "R2-D2 Reloaded: A flexible sound synthesis system for sonic human-robot interaction design," in *Proc. of the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, 2014, pp. 161–167.

[11] J. Fernandez De Gorostiza Luengo, F. A. Martin, A. Castro-Gonzalez, and M. A. Salichs, "Sound Synthesis for Communicating Nonverbal Expressive Cues," *IEEE Access*, vol. 5, pp. 1941–1957, 2017.

[12] E. Frid and R. Bresin, "Perceptual Evaluation of Blended Sonification of Mechanical Robot Sounds Produced by Emotionally Expressive Gestures: Augmenting Consequential Sounds to Improve Non-verbal Robot Communication," *Int. Journal of Social Robotics*, 2021.

[13] M. Joosse, M. Lohse, and V. Evers, "Sound over Matter: The Effects of Functional Noise, Robot Size and Approach Velocity in Human-Robot Encounters," in *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, 2014, pp. 184–185.

[14] L. Zahray, R. Savery, L. Syrkett, and G. Weinberg, "Robot Gesture Sonification to Enhance Awareness of Robot Status and Enjoyment of Interaction," in *Proc. of the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, 2020, pp. 978–985.

[15] A. Repenning, "Moving Beyond Syntax: Lessons from 20 Years of Blocks Programing in AgentSheets," *Journal of Visual Languages and Sentient Systems*, vol. 3, no. 1, pp. 68–91, 2017.

[16] C.-Y. Tsai, "Improving students' understanding of basic programming concepts through visual programming language: The role of self-efficacy," *Computers in Human Behavior*, vol. 95, pp. 224–232, 2019.

[17] E. Peer, L. Brandimarte, S. Samat, and A. Acquisti, "Beyond the Turk: Alternative platforms for crowdsourcing behavioral research," *Journal of Experimental Social Psychology*, vol. 70, pp. 153–163, 2017.

[18] B. J. Zhang, "Appendix for "SonifyIt: a First Step Towards Transformative Sound for All Robots"," 2022. [Online]. Available: https://www.brianzhang.org/publications/2022-iros-conference/

[19] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The Robotic Social Attributes Scale (RoSAS): Development and Validation," in *Proc. of the ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, 2017, pp. 254–262.

[20] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[21] D. S. Syrdal, K. Dautenhahn, K. L. Koay, and M. L. Walters, "The Negative Attitudes towards Robots Scale and Reactions to Robot Behaviour in a Live Human-Robot Interaction Study," in *Proc. of the Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB)*, 2009, pp. 109–115.

[22] B. J. Feir-Walsh and L. E. Toothaker, "An Empirical Comparison of the Anova F-Test, Normal Scores Test and Kruskal-Wallis Test Under Violation of Assumptions," *Educational and Psychological Measurement*, vol. 34, no. 4, pp. 789–799, Dec. 1974.

[23] J. C. Keselman, L. M. Lix, and H. J. Keselman, "The analysis of repeated measurements: A quantitative research synthesis," *British Journal of Mathematical and Statistical Psychology*, vol. 49, no. 2, pp. 275–298, 1996.

[24] M. J. Blanca, R. Alarcon, J. Arnau, R. Bono, and R. Bendayan, "Non-normal data: Is ANOVA still a valid option?" *Psicothema*, vol. 29, no. 4, pp. 552–558, Oct. 2017.

[25] S. Olejnik and J. Algina, "Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs." *Psychological Methods*, vol. 8, no. 4, pp. 434–447, 2003.

[26] D. C. Funder and D. J. Ozer, "Evaluating Effect Size in Psychological Research: Sense and Nonsense," *Advances in Methods and Practices in Psychological Science*, vol. 2, no. 2, pp. 156–168, 2019.

[27] The jamovi project, "jamovi (Version 2.2) [Computer software]," 2021.

[28] R Core Team, "R: A language and environment for statistical computing (Version 4.0) [Computer software]," 2021.

[29] H. Singmann, "afex: Analysis of Factorial Experiments [R package]," 2018.

[30] R. Lenth, "emmeans: Estimated Marginal Means, aka Least-Squares Means [R package]," 2020.