Understanding Non-verbal Sound of Humanoid Robots in Films

Adrian B. Latupeirissa
Sound and Music Computing
KTH Royal Institute of Technology
Stockholm, Sweden
ablat@kth.se

Roberto Bresin
Sound and Music Computing
KTH Royal Institute of Technology
Stockholm, Sweden
roberto@kth.se

ABSTRACT

People's mental model of robots is of importance since it can influence their expectations of how a robot should appear and behave, which in turn will affect the interaction between human and robot. The current mental model of robots is influenced by the presence of robots in films. Thus, understanding the principles of the design of robots in film would benefit the design of robots in the real world.

This extended abstract presents an ongoing investigation of the use of non-verbal sounds of robot in films. Specifically, the investigation looks into the purpose of the sounds, how they are designed, and how the sound design has changed throughout the history of films. Preliminary result suggests the presence of a categorization of robotic sounds in films: inner workings, communication of movement, and expression of emotion.

While further sound design principles are still being formulated, we would argue that having this historical perspective would benefit the understanding of the current expectation of how a robot should sound, thus laying the groundwork for further research in the use of sound in HRI.

KEYWORDS

HRI, social robotics, sound design, mental model, science-fiction

1 INTRODUCTION

The presence of robots in films informs the audience of how robots can behave and interact with humans. While the film has an end, the memories persist and affect the way people think about robots in general. In turn, people's mental model of a robot will influence their expectations and therefore the interaction between human and robot in the real world [Kiesler 2005]. This brings forth the importance of a good design based on these expectations to help facilitate the interaction.

The current study focuses on the sound design of robot. The study takes inspirations from previous study on the aspect of robot movement, where principles and practices of animation from films have been adapted in a study on robot emotional expressions and movements with positive results [Ribeiro and Paiva 2012]. We propose a similar approach by exploring the elements of robot sound design in films that could be adapted in real world scenario.

This extended abstract presents an investigation into the sound design of robot in films, presented with an overarching hypothesis that sonic representations of robots in films would influence the expectations on sounds produced by real-world robots. The investigation is conducted within the context of SONAO project, introduced in [Frid et al. 2018]. SONAO project aims to improve the comprehensibility of robot non-verbal communication (NVC) through an increased clarity of robot expressive gestures and non-verbal sounds. In previous report, we have presented characteristics of robot voices as well as overall observation of robot sonic presence from select films [Latupeirissa et al. 2019]. We are currently expanding the previous study by dissecting sonic characteristics of robots throughout the history of science fiction films.

2 BACKGROUND

It has been suggested that the display of technologies in sci-fi films *"has played a key role in the advancement of technology in the real world"* [Mubin et al. 2019, p.5:2]. Related to the field of interaction design, it has been argued that interfaces from sci-fi films offer lessons to interaction designers, as sci-fi interfaces reflect current interface understandings in terms of expectations from users [Schmitz et al. 2008; Shedroff and Noessel 2012].

Previous studies related to sound in HRI have been reported to have taken inspirations from films [Bethel and Murphy 2006; Hastie et al. 2016; Jee et al. 2010; Read and Belpaeme 2016; Schwenk and Arras 2014]. A common thread across these studies is that their main inspirations in designing auditory non-verbal expression came from R2-D2 (*Star Wars*, 1977) and Wall-E (*Wall-E*, 2008). The two robots — whose sounds are originally designed by the same person, Ben Burtt [1] — have garnered interests due to their unique mode of sonic communication through a series of beeps and chirps, which were taken by several authors *"as a means to replace or augment synthetic speech lacking paralinguistic cues"* [Hastie et al. 2016, p.257].

The use of sound has been prominent in previous studies related to the emotional expression of robots. Two common strategies are the use of recorded acoustic expressions (such as the sound of crying to express sadness) [Häring et al. 2011] and the use of tone patterns to convey target emotions [Löffler et al. 2018; Song and Yamada 2017]. The presence of sound has also been reported to be crucial in the study of human-robot proxemics [Li et al. 2019; Trovato et al. 2018]. In regard to these studies, it has been noted that *"social interaction between social robots and humans takes place through multimodal interaction"*, and that there is a need to further study the use of robot sound within multimodal interaction [Yilmazyildiz et al. 2016, p.80]. Due to the nature of multimodal interaction, it has been suggested that poorly designed interaction — such as giving a mismatched sound to the appearance or movement of a robot — might introduce an undesirable effect in HRI studies [McGinn and Torre 2019].

Inherently, as robots are made out of servos and motors, the sound of mechanical motions have become an integral part of a robot and the sound has also become a part of people's current mental model of a robot. Moreover, in arguing that the current

[1] https://www.imdb.com/name/nm0123785/

mental model comes from sci-fi films, we are acknowledging the preconception that sounds in films are designed with intention. The sound of robots mechanical workings, for example, is rarely audible when a dialogue is deemed more important in a particular moment. In the real world, the sounds of these mechanical motions are considered as consequential sounds and would be perceived as noise [Langeveld et al. 2013]. The consequential sounds of robots are less explored and little is known on how they could influence interaction. Recent studies on this are reported in [Moore et al. 2017] and [Moore et al. 2019], who have explored the effect of consequential sound on the overall perception of robots. While no design principles have been established, these studies have looked into the aural impressions toward servo motors, which we agree that it would be beneficial as a building block in designing future robotic sound interaction.

3 METHODOLOGY

The list of films was obtained from IMDb[2] by searching for all feature films with the keyword "robot" in the genre of science-fiction, released up to the year 2018. The initial list of 509 films is then narrowed down by considering only non-animated films, involving English-speakers, and involving humanoid robots with agency. With these criteria, our final list consists of 74 titles ranging from *Tobor the Great* (1954) to *Bumblebee* (2018).

For each selected film, video excerpts — each with a total duration of maximum 2 minutes — displaying the robot were isolated by the first author with the help of an assistant where the list of films were evenly divided. The criteria of an excerpt is that the scene shows the main robot of the film engaged in a dialogue or performed a set of movements and gestures, with a clearly audible sound of the robot (i.e without background music or ambient noise). Qualitative content analysis were then performed on the excerpts by examining and annotating the context of the scenes and observing the sound spectrometer. The annotations were then compared across films and grouped by the purposes of the sounds.

4 PRELIMINARY RESULTS

The preliminary results show that the there are three main categories of non-verbal sounds of robot. The categories, along with the highlights, are presented as follows.

4.1 Presence and inner workings

The sound of buzzing machine and electrical noise are commonly used to mark the robots' state of being activated or "alive", as one would have expected when a machine or a device is activated. In films, this type of sound is utilized to communicate the presence of a robot in a scene, even though the robot is not visible on the screen. Another distinctive use is to communicate to the audience that the robot is broken (see Andrew Martin in *Bicentennial Man*, 1999, and Captain S.T.A.R in *The Black Hole*, 1979).

4.2 Communication of movement

Many antagonist robots (most with the characteristics of a single-minded and destructive nature) are often designed with extensive use of low-pitched, mechanical movement sounds and heavy, rigid, and regular-paced foot-steps. The design is meant to threaten the protagonist and the audience. In contrast, the movement of the protagonist robots are often characterized by higher-pitched sounds with a softer sound level, and less aggressive. Aside from supporting the visual appearance and demeanor, the movement sounds are commonly used to augment the expression of emotion through movement, as described in the following point.

4.3 Expression of emotion

We have observed that mechanical sounds show a decreasing in frequency when the robot is communicating sadness, depression, or disappointment (see Marvin in *The Hitchhiker's Guide to the Galaxy*), and an increasing in frequency for expressing attention (see Andrew Martin in *Bicentennial Man*). Repeated high-frequency short sounds in regular interval can be used for expressing happiness (also found in Andrew Martin), while in jerky irregular interval can be used for expressing anxiety (see Chappie in *Chappie*). In a different element, inner workings sounds have been typically used with increasing intensity for instance when showing anger or aggressiveness (see Sonny in *I-Robot*), as well as decreasing intensity to show relaxation (see Bumblebee in *Bumblebee*). All these manipulations of sound cues are reflecting what has been previously documented in several studies on the communication of emotions in speech, music, and body movements (see for example [Bresin and Friberg 2011; Giordano et al. 2014; Juslin and Laukka 2003]).

5 CONCLUSIONS AND FUTURE WORK

The current extended abstract has presented the preliminary findings from our investigation into the sound design of robot in films, in the form of categorization of non-verbal sounds of robot. While the investigation is still ongoing, in the context of understanding current mental model of robot these categorization could offer a glimpse into how people perceive the sound elements of robots in films, and how it might affect the future design of robots. Further examination would be focused on the question of how the sound are designed and how they have been evolved throughout the years related to the technical development in sound production for films.

It is of importance to point out a limitation that all films in the list were produced in the western world. Future study factoring cultural differences would also be of our interest.

[2]Obtained in January 2019 from https://www.imdb.com/

REFERENCES

Cindy L. Bethel and Robin R. Murphy. 2006. Auditory and other non-verbal expressions of affect for robots. *AAAI Fall Symposium - Technical Report* FS-06-01 (2006), 1–5.

Roberto Bresin and Anders Friberg. 2011. Emotion rendering in music : Range and characteristic values of seven musical variables. *Cortex* 47, 9 (2011), 1068–1081. https://doi.org/10.1016/j.cortex.2011.05.009 QC 20190412.

Emma Frid, Roberto Bresin, and Simon Alexanderson. 2018. Perception of Mechanical Sounds Inherent to Expressive Gestures of a NAO Robot - Implications for Movement Sonification of Humanoids. In *Proceedings of the 15th Sound and Music Computing Conference* :. 43–51. https://doi.org/10.5281/zenodo.1422499 QC 20181211.

B.L. Giordano, H. Egermann, and Roberto Bresin. 2014. The production and perception of emotionally expressive walking sounds : Similarities between musical performance and everyday motor activity. *PLoS ONE* 9, 12 (2014), e115587–. https://doi.org/10.1371/journal.pone.0115587 QC 20150126.

Helen Hastie, Pasquale Dente, Dennis Küster, and Arvid Kappas. 2016. Sound emblems for affective multimodal output of a robotic tutor: A perception study. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ACM, 256–260.

M. Häring, N. Bee, and E. André. 2011. Creation and Evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *2011 RO-MAN*. 204–209. https://doi.org/10.1109/ROMAN.2011.6005263

Eun-Sook Jee, Yong-Jeon Jeong, Chong Hui Kim, and Hisato Kobayashi. 2010. Sound design for emotion and intention expression of socially interactive robots. *Intelligent Service Robotics* 3, 3 (2010), 199–206.

Patrik N Juslin and Petri Laukka. 2003. Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological bulletin* 129, 5 (2003), 770.

Sara Kiesler. 2005. Fostering common ground in human-robot interaction. *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication* 2005 (2005), 729–734. https://doi.org/10.1109/ROMAN.2005.1513866

Lau Langeveld, René van Egmond, Reinier Jansen, and Elif Özcan. 2013. Product sound design: Intentional and consequential sounds. *Advances in industrial design engineering* 47, 3 (2013).

Adrian Benigno Latupeirissa, Emma Frid, and Roberto Bresin. 2019. Sonic characteristics of robots in films. (2019), 1–6.

Rui Li, Marc van Almkerk, Sanne van Waveren, Elizabeth Carter, and Iolanda Leite. 2019. Comparing human-robot proxemics between virtual reality and the real world. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 431–439.

Diana Löffler, Nina Schmidt, and Robert Tscharn. 2018. Multimodal expression of artificial emotion in social robots using color, motion and sound. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 334–343.

Conor McGinn and Ilaria Torre. 2019. Can you tell the robot by the voice? An exploratory study on the role of voice in the perception of robots. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 211–221.

Dylan Moore, Tobias Dahl, Paula Varela, Wendy Ju, Tormod Næs, and Ingunn Berget. 2019. Unintended Consonances: Methods to Understand Robot Motor Sound Perception. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 500.

Dylan Moore, Hamish Tennent, Nikolas Martelaro, and Wendy Ju. 2017. Making noise intentional: A study of servo sound perception. (2017), 12–21.

Omar Mubin, Kewal Wadibhasme, Philipp Jordan, and Mohammad Obaid. 2019. Reflecting on the Presence of Science Fiction Robots in Computing Literature. *ACM Transactions on Human-Robot Interaction (THRI)* 8, 1 (2019), 5.

Robin Read and Tony Belpaeme. 2016. People interpret robotic non-linguistic utterances categorically. *International Journal of Social Robotics* 8, 1 (2016), 31–50.

Tiago Ribeiro and Ana Paiva. 2012. The illusion of robotic life: principles and practices of animation for robots. (2012), 383–390.

Michael Schmitz, Christoph Endres, and Andreas Butz. 2008. A survey of human-computer interaction design in science fiction movies. In *Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive enterTAINment*. ICST (Institute for Computer Sciences, Social-Informatics and ..., 7.

Markus Sebastian Schwenk and Kai Oliver Arras. 2014. R2-D2 Reloaded: A flexible sound synthesis system for sonic human-robot interaction design. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication, IEEE RO-MAN 2014, Edinburgh, UK, August 25-29, 2014*. 161–167. https://doi.org/10.1109/ROMAN.2014.6926247

Nathan Shedroff and Chris Noessel. 2012. Make It So: Learning from Sci-fi Interfaces. (2012), 7–8. https://doi.org/10.1145/2254556.2254561

Sichao Song and Seiji Yamada. 2017. Expressing Emotions Through Color, Sound, and Vibration with an Appearance-Constrained Social Robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. ACM, New York, NY, USA, 2–11. https://doi.org/10.1145/2909824.3020239

Gabriele Trovato, Renato Paredes, Javier Balvin, Francisco Cuellar, Nicolai Bæk Thomsen, Soren Bech, and Zheng-Hua Tan. 2018. The Sound or Silence: investigating the influence of robot noise on proxemics. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 713–718.

Selma Yilmazyildiz, Robin Read, Tony Belpeame, and Werner Verhelst. 2016. Review of semantic-free utterances in social human–robot interaction. *International Journal of Human-Computer Interaction* 32, 1 (2016), 63–85.