Human-Sound Interaction (HSI) workshop

Balandino Di Donato
Informatics Department
University of Leicester
Leicester, United Kingdom
balandino.didonato@leicester.ac.uk

ABSTRACT

The design of interactions with sound and audio processes is a seminal activity in the creation of a performance, installation, a virtual sound environment, or interface for musical expression. The interaction design is often fixated by the interface without taking into account human factors and our diverse abilities to perceive the sound and interface affordances. Human-Sound Interaction (HSI) look at human-centered interaction design aspects that determine the realisation and appreciation of musical works (installations, composition and performance), interfaces for sound design and musical expression, augmented instruments, sonic aspects of virtual environments, interactive audiovisual dance performances.

Author Keywords

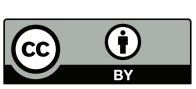
Human-Sound Interaction, Human-Centred Interaction Design, Sonic and Music Affordances, diversity, inclusivity

CCS Concepts

•**Human-centered computing** → **Interaction design theory, concepts and paradigms;** *Interaction design process and methods; Accessibility design and evaluation methods;*

1. BACKGROUND

Performers demonstrate a high level of interactivity with sound through their ability to physically control musical nuance of the instrument. This interaction is facilitated by physical feedback response from the instrument such as string resistance, vibration, pressure and kinematic of keys being pressed. The fixed mechanical structure of the instrument and the available high-frequency feedback loop channels, give performers the sound control at a performative level. Through interaction design, we can amplify the possibilities of musical instruments, especially in concern of timbre and sound-and-gesture articulation properties. In recent years, the evolution of technology has made possible simultaneous development of meta-instruments, hyper-instruments or augmented instruments [2, 19, 10] attracting increased attention from musicians and scientific researchers.



NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

Interactivity is a theme widely explored in the field of Sonic Interaction Design (SID) by placing sound at the centre of the design [23]. This field of research reflects on several aspects of sonic interaction such as perceptual, cognitive and emotional, product sound design, auditory display and sonification. SID aims to explore *"ways in which sound can be used to convey information, meaning, and aesthetics and emotional qualities in an interactive context"* [9] in different contexts which can vary from critical functionalities of a product that relies on auditory feedback to communicate to its users, to artistic musical works such as performances or installations.

Whilst SID highlights sonic aspects, in this workshop, we propose the idea of Human-Sound Interaction (HSI) which focuses on the investigation of human factors in the interaction with sound.

In designing modes of interaction with sound, various factors should be considered. For example, the genre of performance, the interface, the acoustic instrument, the available sensing technology, the computational power required [4], and the affordances of the interface [29] and the sound. It is the latter aspect that is particularly relevant to this workshop in affecting the interaction design.

The term affordance was first used by Gibson [12], and its principles rely on the possible action that each object evokes based on the characteristics of the objects and the capabilities of the subjects [13]. Godøy [14] demonstrates that music can invite certain gestures that are often encouraged by timbral and dynamic qualities of the sound, by mimicking the action that might have produced them or gestures evoked by the music which might not necessarily refer to the production or sound qualities. From here we derive the concept of *affordance of musical sound.*

A previous study by Godøy et al. [15], participants were asked to trace gestures along with sounds. The work concluded that the majority of gestures are evoked by sound morphology. For example, a sound with an ascending pitch would be described with a gesture tracing an ascending curve.

Caramiaux et al. [3] repeated the same experiment exploring differences between gestures evoked by "causal" and "non-causal" sounds, where the "causality" is the degree within which the listener can distinguish the sound's environmental cause. While results suggested that people's gestures aimed to mimic the cause of the sound, there was inconsistency across participants. Participant's gestures were similar when tracing gestures along with non-causal sounds, which described the sound morphology.

Although the aforementioned studies identify important information about the gesture-sound relationships, they do not consider the potential impact of the gesture tracking device's affordance on the sound-tracing exercise.

Differently, Tanaka and Altavilla [29] explored gestural af-

fordance with sound in relation to different devices. They realised a study in which participants were asked to use three controllers, an Axivity Wax, iPhone and Wii-remote, to generate three different types of audio feedback: (i) to trigger a snare drum and a bowed violin sound, (ii) to control a granulation effect over violin sample and modify its pitch and loop speed, and (iii) to control the same granulation fed with a voice sample in a loop. From the results, they concluded that the combination of input devices and output feedback leads to the construction of highly complicated and sophisticated affordances, which require an equal complex construct of affordance.

Gaver [11] expands the concept of affordances, separating them from the perceptual information available. He distinguished them between (i) *correct rejections* when there is no affordance and it is not perceived; (ii) *perceived*, the affordance is present and perceived; (iii) *hidden*, the affordance is present but it is not perceived; and (iv) *false affordances* an affordance is perceived but it does not exist (see Figure 1).

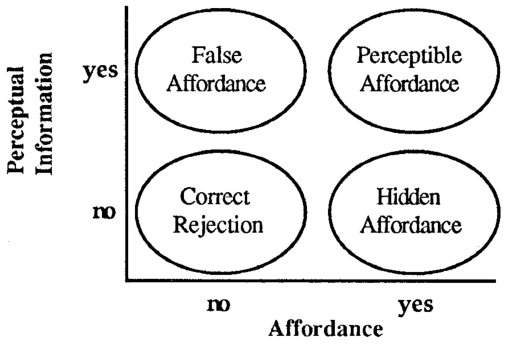


Figure 1: Distinction between different types of affordance. Adapted from [11]

Being our perception of sound a crucial to determine the affordances of a sound, it is important that we consider the diversity of peoples' hearing profile (otological normal, hearing impaired and D/deaf) and how this can impact the design of HSI. These principles are widely investigated by the field of Aural Diversity, the study of sound and music that addresses the full range of human hearing types [18]. Sound a music perception can be affected by cross-modal cues; for example, the impact of vision on the perception of music tempo [31], tone duration [26], and source localisation [27]. As in for the sense of hearing, visual disabilities can affect the perception of music and sound [5, 22].
Especially for musicians, the haptic feedback of a musical instrument covers a vital role in the perception of the musical quality of an instrument [24], and can also be used to make interfaces accessible for musical expression to both visually [30] and hearing impaired [25].
This workshop will question participants' HSIs design practice about the accessibility of their interfaces and the actions that these evoke through auditory, visual and haptic feedback.

The perception of affordances is also influenced by our cultural background and pre-existing knowledge. The constraints that cultural implications pose on affordances have been explored by Oshlyansky et al. [21]. They have realised a study where UK and USA based higher education students were asked to make judgements about the behaviour of abstracted Western-like objects such as a switch. Results showed that while one group of people judged the switch in an ON state, the other group did the opposite. Based on this evidence, the diversity in participants of this workshop and the activity themselves encourage a discussion on different aspects of designing HSIs that might be related to one's culture.

The discovery of affordances and interactivity are driven by a continuous action-perception loop [28]. As well as perceptual and cognitive of an individual's capabilities, physical affordances of an interface are crucial to designing interfaces [17]. To this extent, in the design of HSIs for a diverse audience, the workshop invites participants to reflect on the wide range of people's motor skills, to address suitable interaction models and technological solutions.

2. AIMS AND OBJECTIVES

Referring to principles of *affordance of musical sound*, and the diversity of abilities and human factors to perceive them, this workshop invites participants to take part in a collaborative activity for the exploration of Human-Sound Interaction (HSI), defined as direct, engaging, natural and embodied interaction with sound [6, 7]:

- direct as the impression or a feeling about an interface capable of being described in terms of concrete actions [20]
- engaging as fostering the "feeling of directly manipulating the objects of interest", where "the world of interest is explicitly represented and there is no intermediary between user and world"[20]
- natural *"as being marked by spontaneity"* [16]
- embodied as an extension and incorporation of human skills and abilities within the interaction design of a system [8]

This workshop aims to explore HSI principles and practices in an interdisciplinary, diverse, interactive and collaborative setting. Musicians, technologists and sound artists from different backgrounds will be invited to ideate and design modes of interaction with sound, and reflect on our capabilities to perceive the affordances of sound and of interaction with an interface.
The HSI workshop invites scientific and artistic contributions on topics that include:

- Human-Centred Design approaches for sonic interaction design
- Interactions designed by/with/for diverse musicians
- Fostering diversity in music and sound art through interaction design
- Sonic and cross-modal music affordances
- Embodied interaction with sound
- Interaction design for augmented instruments
- Adaptive frameworks for interaction design
- Interaction design for audio/musical interfaces
- Interaction design for interactive sound art and installations
- Impact of interaction design in artistic practice, culture, societal and gender aspects
- Impact of human, artistic, cultural, societal and gender factors in designing interaction with sound
- Inclusive interaction design strategies with sound

- Sonic interactions in Virtual (VR), Augmented (AR) and Mixed (MR) realities
- Interactive audio-visual dance performance

This is part of a series of workshops across different communities that aim to build the knowledge through the experience of practitioners at the intersection of music, sound art, performance and Human-Centred Interaction Design, and to investigate how to make sound experience more inclusive and diverse through HSI. Findings will contribute to a body of knowledge on Human-Sound Interaction shared with the community.

3. WORKSHOP PROGRAM

In this half-day workshop, participants are invited to take part in the following interaction design activities: warm-up, divergent, incubation break no.1, div-interaction, incubation break no.2, and convergent.

For these activities participants will be asked to use any tool that they are comfortable using for sketching and design ideas. For example: pen and paper, or software tools such as miro[1], Lucidchart[2] Microsoft Office[3] tools, etc.
They will also be provided with a storage space where to upload their ideas and share them instantaneously with the workshop attendees. A week before the workshop, the organisers will send participants a welcome digital package containing detailed instructions about the workshop program and requirements (see also Section 4).

Introduction (30 minutes).

In this phase, the organiser and participants present themselves. After the Human-Sound Interaction concept and the workshop schedule will be introduced.

Warm-up (10 minutes).

Each participant to generate five or more extremely bad ideas around HSI. They can use any design approach, random metaphors, or sketching absurd musical or sound design contexts. *"To invent, you need a good imagination and a pile of junk."* Thomas Edison, quoted in [1].

Divergent (20 minutes).

Participants will be asked to select ideas from the warm-up session, or ideate between 2 and 4 new HSIs, without thinking at the potential interface that fosters such interactions. After, participants will focus on the implementation of their HSI in the real world situation. Through sketching, they will have to think design an interface, installation, software mapping solution or any other artefact that can potentially satisfy the human-sound interaction they have previously designed. For the sake of this workshop, this last one will be called "the prototype".

Incubation break no.1 (20 minutes).

Participants will have the opportunity to let their ideas sink-in and have a rest. Participants can do anything they wish, which does not require any particular cognitive load, but that allows them to keep their mind in a free flow state (making coffee, washing up, rest, a nap, tidying up, etc.).

Div-interaction (30 minutes).

At this point, participants will enter a collaborative design phase. Participants will work in pairs (*Pa* and *Pb*). In the first 15 minutes, *Pa* will share with *Pb* the prototype ideas only, thus, *Pb* will design an HSIs based on *Pa*'s prototypes; and, *Pb* will share with *Pa* the HSI only, thus, *Pa* will have to design an interface prototype upon the HSI. This process is then repeated with *Pa* sharing with *Pb* the prototype, and *Pb* sharing with *Pa* the HSI.

This phase is crucial to the whole workshop. Here themes of Human-Interaction Design in a collaborative and iterative process will generate through a natural exchange of ideas. The order within which each participant will exchange the HSI and the prototype will allow the organisers to observe the adaptive, transferable and transformative aspects of the design process.

Incubation break no.2 (20 minutes).

As Incubation break no.1 (see Section 3).

Convergent (40 minutes).

Participants will receive back their original HSI and prototype designs from the participant they paired up in the *Div-interaction* phase, and go through a convergent design process, in which they will finalise their HSI and interface prototype. More to the point, in this phase they will have to finalise their HSI designs and implementation.

Wrap-up (10 minutes).

Closing and remarks of the session.

4. TECHNICAL SETUP

This workshop will be online delivered by a video conferencing software provided by the conference organisers to compliant with the conference format.
Participants will be provided with workshop materials and instructions a week before the workshop. All materials will be produced meeting accessibility and ethical standards, at the core of the University of Leicester practices. Live captioning and any other accessibility request will be taken into account.

5. TIMING AND PARTICIPANTS

This is a half-day online interactive workshop lasting three hours. The HSI workshop welcomes any number of participants, preferably an even number of participants due the nature of the div-interaction phase, within the limits of the online platform complaint with the conference policy.

6. ORGANISER

6.1 Balandino Di Donato

Balandino Di Donato is a Digital Artists, Researcher and Lecturer in Creative Computing at the University of Leicester. During his PhD at Royal Birmingham Conservatoire (BCU), he explored the design of embodied interactions with audiovisual processes during music performance. He worked as Research Assistant at Goldsmiths, University of London on the realisation of EMG-interface and -driven AI as part of the ERC-funded project: BioMusic; and, as Research and Artistic Assistant at Centro Ricerche Musicali in Rome. He authored and contributed towards the development of software and interfaces for musical expression (Myo Mapper, Interga Live, TUI Metis). His research focuses on Digital Arts and Human-Centred Interaction Design.

[1] https://miro.com/
[2] https://www.lucidchart.com
[3] https://www.microsoft.com/en-gb/microsoft-365

Ethical Standards

This workshop will be conducted complying with all University of Leicester's Research Ethics Policy[4]. It provides the framework and guidelines for conducting research with integrity and promoting good practice in all aspects of research.

7. REFERENCES

[1] R. C. Arkin, R. C. Arkin, et al. *Behavior-based robotics*. MIT press, 1998.

[2] L. Bianchini, M. Lupone, S. Lanzalone, and A. Gabriele. Augmented instruments at crm – centro ricerche musicali of rome: Feed-drum, skinact, windback and resoflute. In *Proceedings of the International Computer Music Conference*, ICMC '19, New York, NY, USA, 2019.

[3] B. Caramiaux, P. Susini, T. Bianco, et al. Gestural embodiment of environmental sounds : an experimental study. In *Proceedings of the International Conference on New Interaces for Musical Expression*, NIME '11, pages 144–148, Oslo, Norway, 2011.

[4] P. R. Cook. Principles for designing computer music controllers. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 3–6, Seattle, WA, 2001.

[5] A.-A. Darrow and J. Novak. The Effect of Vision and Hearing Loss on Listeners' Perception of Referential Meaning in Music. *Journal of Music Therapy*, 44(1):57–73, 03 2007.

[6] B. Di Donato, C. Dewey, and T. Michailidis. Human-sound interaction. In *Proceedings of the 2020 Conference on New Interfaces for Musical Expression*, NIME'20, Birmingham, United Kingdom, 2020. Royal Birmingham Conservatoire, Birmingham City University.

[7] B. Di Donato, C. Dewey, and T. Michailidis. Human-sound interaction towards human-centred cross-modal interaction with sound. In *Proceedings of the 7th International Conference on Movement and Computing*, MOCO '20, New York, NY, USA, 2020. Association for Computing Machinery.

[8] P. Dourish. *Where The Action Is: The Foundations of Embodied Interaction*. MIT Press, Oct. 2001.

[9] K. Franinović and S. Serafin. *Sonic interaction design*. Mit Press, 2013.

[10] A. Gabriele, M. Lupone, P. Inverardi, and P. Pelliccione. Ad-opera: music-inspired self-adaptive systems. In *Proceedings of the FSE/SDP workshop on Future of software engineering research*, pages 119–124, 2010.

[11] W. W. Gaver. Technology affordances. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pages 79–84, New York, NY, USA, 1991. ACM.

[12] J. J. Gibson. *The senses considered as perceptual systems*. Houghton Mifflin, 1966.

[13] J. J. Gibson. The theory of affordances. *Hilldale, USA*, 1:2, 1977.

[14] R. I. Godøy. Gestural affordances of musical sound. In R. I. Godøy and M. Leman, editors, *Musical Gestures: Sound, Movement, and Meaning*, pages 115–137. Routledge, New York, New York, 2010.

[15] R. I. Godøy, E. Haga, and A. R. Jensenius. Exploring music-related gestures by sound-tracing: A preliminary study. In *Proceedings of the 2nd International Symposium on Gesture Interfaces for Multimedia Systems*, GIMS2006, Leeds, United Kingdom, 2006.

[16] S. A. Grandhi, G. Joue, and I. Mittelberg. Understanding naturalness and intuitiveness in gesture production. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI, pages 821–824, Vancouver, BC, Canada, 2011.

[17] R. Hartson. Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology*, 22(5):315–338, 2003.

[18] A. Hugill. Aural diversity website, 2020. Available at: `http://auraldiversity.org/`. Accessed: 09/05/2020.

[19] M. Lupone, S. Lanzalone, A. Gabriele, and M. De Luca. Forme immateriali by michelangelo lupone: structure, creation, interaction, evolution of a permanent adaptive music work. In *Proceedings of the Electroacoustic Winds 2015 Conference*, Aveiro, Portugal.

[20] D. A. Norman and S. W. Draper. *User Centered System Design: New Perspectives on Human-computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1986.

[21] L. Oshlyansky, H. Thimbleby, and P. Cairns. Breaking affordance: Culture as context. In *Proceedings of the Third Nordic Conference on Human-Computer Interaction*, NordiCHI '04, page 81–84, New York, NY, USA, 2004. Association for Computing Machinery.

[22] L. Pring, K. Woolf, and V. Tadic. Melody and pitch processing in five musical savants with congenital blindness. *Perception*, 37(2):290–307, 2008. PMID: 18456928.

[23] D. Rocchesso, S. Serafin, F. Behrendt, N. Bernardini, R. Bresin, G. Eckel, K. Franinovic, T. Hermann, S. Pauletto, P. Susini, and Y. Visell. Sonic interaction design: Sound, information and experience. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '08, page 3969–3972, New York, NY, USA, 2008. Association for Computing Machinery.

[24] C. Saitis, H. Järveläinen, and C. Fritz. *The Role of Haptic Cues in Musical Instrument Quality Perception*, pages 73–93. Springer International Publishing, Cham, 2018.

[25] S. C. Sanayakkara, L. Wyse, S. H. Ong, and E. A. Taylor. Enhancing musical experience for the hearing-impaired using visual and haptic displays. *Human–Computer Interaction*, 28(2):115–160, 2013.

[26] M. Schutz and S. Lipscomb. Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, 36(6):888–897, 2007. PMID: 17718367.

[27] B. Shelton and C. Searle. The influence of vision on the absolute identification of sound-source position. *Perception & Psychophysics*, 28(6):589–596, 1980.

[28] D. Svanæs. *Understanding interactivity: steps to a phenomenology of human-computer interaction*. Norges teknisk-naturvitenskapelige universitet, 2000.

[29] A. Tanaka, A. Altavilla, and N. Spowage. Gestural musical affordances. In *Proceedings of the 9th International Conference on Sound and Music Computing*, Copenhangen, Denmark, July 2012.

[30] A. Tanaka and A. Parkinson. Haptic wave: A cross-modal interface for visually impaired audio

[4] `https://www2.le.ac.uk/institution/ethics/policy`

producers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 2150–2161, New York, NY, USA, 2016. Association for Computing Machinery.

[31] O. Vitouch, S. Sovdat, and N. Holler. Audio-vision: Visual input drives perceived music tempo. In *Proceedings of the 9th International Conference on Music Perception & Cognition*, pages 86–89, Bologna, Italy, 2006.