

Towards Sound Accessibility in Virtual Reality

Dhruv Jain
Microsoft Research, Redmond,
Washington; University of
Washington, Seattle, Washington
djain@uw.edu

Sasa Junuzovic
Microsoft Research, Redmond,
Washington
sasajun@microsoft.com

Eyal Ofek
Microsoft Research, Redmond,
Washington
eyalofek@microsoft.com

Mike Sinclair
Microsoft Research, Redmond,
Washington
sinclair@microsoft.com

John R. Porter
Microsoft, Redmond, Washington
joporte@microsoft.com

Chris Yoon
Microsoft, Redmond, Washington
chyoon@microsoft.com

Swetha Machanavajhala
Microsoft Research, Redmond,
Washington
swmchan@microsoft.com

Meredith Ringel Morris
Microsoft Research, Redmond,
Washington
merrie@microsoft.com

ABSTRACT

Virtual reality (VR) leverages sight, hearing, and touch senses to convey virtual experiences. For d/Deaf and hard of hearing (DHH) people, however, information conveyed through sound may not be accessible. While prior work has explored making every day sounds accessible to DHH users, the context of VR is, as yet, unexplored. In this paper, we provide a first comprehensive investigation of sound accessibility in VR. Our primary contributions include a design space for developing visual and haptic substitutes of VR sounds to support DHH users and prototypes illustrating several points within the design space. We also characterize sound accessibility in commonly used VR apps and discuss findings from early evaluations of our prototypes with 11 DHH users and 4 VR developers.

CCS CONCEPTS

• **Human-centered computing** → Accessibility.

KEYWORDS

accessibility, virtual reality, Deaf, hard of hearing, sound, haptic, visualization

ACM Reference Format:

Dhruv Jain, Sasa Junuzovic, Eyal Ofek, Mike Sinclair, John R. Porter, Chris Yoon, Swetha Machanavajhala, and Meredith Ringel Morris. 2021. Towards Sound Accessibility in Virtual Reality. In *Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21)*, October 18–22, 2021, Montréal, QC, Canada. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3462244.3479946>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ICMI '21, October 18–22, 2021, Montréal, QC, Canada
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-8481-0/21/10...\$15.00
<https://doi.org/10.1145/3462244.3479946>

1 INTRODUCTION

Virtual reality (VR) environments are filled with a rich diversity of sounds ranging from sounds that provide critical notifications (e.g., enemy footsteps) to those that increase realism (e.g., wind blowing) [2, 5]. For many d/Deaf and hard of hearing (DHH) people, these sounds and the information they convey may not be accessible, which may limit or inhibit their VR experience (e.g., by missing critical cues in games or conversations in social apps). In this paper, we investigate accessible visual and haptic augmentations for sounds in VR.

While prior work has explored visual [26, 30, 37, 49] and haptic [43, 50, 69] substitutes of everyday sounds for DHH users, this approach has not been examined for virtual reality—a radically different context due to possibilities for fictional worlds and exaggerated representations of the real world [11]. Some VR games have begun incorporating accessibility features for DHH people (e.g., [73, 74]); however, these are one-off efforts, which have not been formally explored in the literature. Our focus is on making mainstream VR accessible, which can lead to lower development costs, increased availability, and better social acceptability over specialized experiences [53, 57].

To arrive at accessible substitutes for VR sounds, we began with articulating a design space containing nine dimensions to map different types of sounds and their characteristics specific to VR [4] to visual and haptic feedback. We used *morphological search* [54], a technique inspired from linguistic fields, to find syntactic and semantic relationships between sound and visual/haptic design dimensions.

To investigate the potential of our design space, we used it to meet two expected goals expected of a successful design space [9, 10]: contextualizing relevant prior work and generating new prototypes. Towards the first goal, since the literature work on VR for DHH users is scarce, we instead investigated commercial apps and games containing accessibility features for DHH users (e.g., captions, sound location indicators). We found that our design was able to successfully contextualize all sound-to-visual/haptic

mappings in 78 apps we analyzed. Our findings also identified the potential to generate new mappings for several inaccessible sounds.

Based on these insights, we then used our design space to develop novel sound-to-visual/haptic mappings for some inaccessible sounds and sound characteristics (our second goal). We instantiated six of these mappings by building four visual and two haptic sound feedback prototypes for DHH users (see our supplementary video). As an example, the two haptic prototypes include a novel haptic belt to convey multiple sound characteristics (*e.g.*, pitch, persistent, loudness) of ambient sounds and a smartwatch vibration app to convey location of critical information sounds. Compared to prior work, which has only built specialized VR experiences for DHH users (*e.g.*, for storytelling [63]), our prototypes are generalizable and can be incorporated by developers into any new or existing VR app using our code scripts. Preliminary evaluations with 11 DHH users and 4 VR developers suggest that our prototypes are useful, customizable, and can be easily integrated to VR apps. We close by discussing design considerations for integrating sound feedback in VR apps.

In summary, our research contributes: (1) a design space for developing visual and haptic augmentations of VR sounds to support DHH users, (2) findings that contextualize sound accessibility across commonly used apps and games, (3) six prototypes instantiating novel sound to visual and haptic mappings, and (4) insights from preliminary evaluations of these prototypes with 11 DHH users and four VR developers.

2 RELATED WORK

We cover visual/haptic sound feedback work for non-VR settings and VR accessibility work for other disabilities.

2.1 Sound Visualization

Communicating acoustic information through visuals has been explored in many domains such as audio debugging [17], oceanography [12], appliance repairs [3], ecological surveys [35, 52], and medical imaging [23, 72]. We focus on prior work in music and real-life sound visualizations, which closely relate to our work.

While detailed music visualizations are not our focus, the mappings of different musical characteristics to visuals helped inform our design space. Basic music visualizations include 2-D time series displays (*e.g.*, waveforms, spectrograms) for characteristics such as pitch or loudness [13, 34, 60]. In terms of 3D visualizations, Smith and Williams [60] constructed a 3-D version of the classic piano-roll visualization, using colors to mark timbre. Miyazaki *et al.* [40] expanded this visualization by additionally conveying pitch as hue. For DHH users specifically, Nanayakkara *et al.* [42] created an abstract display that changed size, color, and brightness in correlation with the harmonics to enhance musical experiences of DHH people. Cooper Union In New York City designed a wall display to convey musical frequencies through light for teaching music to DHH children [75].

Our work is also influenced by prior work exploring visualizations of everyday sounds to support DHH people [21, 26, 27, 49]. For example, Jain *et al.* [26] investigated visualizations for localizing sounds on augmented reality glasses, identifying dimensions such as egocentric vs. exocentric views, directional granularity (*e.g.*,

4-way vs. 8-way), and icons (*e.g.*, arrows vs. pulses). Another work [24] explored visualizations for speech captioning, delineating contextual (*e.g.*, 1:1 vs. group meetings) and UI design (*e.g.*, placement of captions in 3D space) dimensions. *HomeSound* [29, 30] investigated sound visualizations to provide sound awareness in the homes of DHH people, suggesting display form-factor (*e.g.*, smartphone, tablets), sound types to display (*e.g.*, alarms, alerts), and sound characteristics to display (*e.g.*, identity, location, duration) as the primary design dimensions.

While promising, the above work has not been studied in the context of VR, which contains different sounds than real-life and uses devices with potentially a greater field-of-view than non-VR devices. Consequently, our design space is much broader, covering sounds (*e.g.*, interaction sounds, non-localized speech) and sound characteristics (*e.g.*, persistence, priority) with no counterparts in the real-world, and new dimensions to convey sound feedback (*e.g.*, visual elements, VR specific form-factors).

2.2 Haptic Sound Feedback

While no prior work has explored haptic sound feedback for VR context, our work is informed by prior research on haptic feedback for real-life sounds intended to support DHH users. This research can be categorized into three domains: speech perception, environmental sound awareness, and music feedback. Most efforts have focused on speech therapy by conveying voice tone or frequency—for example, through electro-tactile stimulators that use complex patterns of low-intensity electrical currents to deliver haptic sensation to the skin [18, 55, 65]. For environmental sound feedback, prior work has examined wrist-worn devices that use a single vibration motor to convey either the occurrence of a sound [55, 62] or loudness by vibrating with an intensity proportional to the ambient sound level [25].

For music, haptic devices have conveyed tempo or rhythm via chairs fitted with multiple vibration motors [32, 43] and, more recently, using wearables such as a vibrotactile shoe [67]. Building on this work, we examine an electromagnetically actuated haptic belt to convey music and ambient sounds in VR (the first of our two haptic prototypes). Besides dedicated haptic devices, three works have explored conveying vibration in tandem with visual feedback on smartphones [7, 38, 58]. Here, vibration was used as a secondary feedback intended to alert the user to look at their phones rather than to encode sound as a haptic signal. While useful, participants who evaluated these prototypes also wanted additional information (*e.g.*, loudness) through vibration. Informed by this finding, Goodman *et al.* [20] explored vibratory patterns to indicate sound direction on a smartwatch; we extend this work to VR by examining a smartwatch app for localizing VR sounds (our second haptic prototype).

Finally, researchers have also tried methods to completely substitute hearing with haptic sensation (*e.g.*, [18]), but with little promise. Our work is complementary; we aim to augment, not substitute, sounds.

2.3 VR Accessibility

Explorations of accessibility in VR have only recently gained traction. For example, the first two symposiums for virtual, augmented,

and mixed reality accessibility (XRAccess [76]) were held in 2019 and 2020. Informed by the symposium discussions, the World Wide Web Consortium (W3C) published the first working draft of VR accessibility requirements in Feb 2020 [77]; relevant needs for DHH users include customizable subtitles, descriptions of important sound events, and availability of binaural audio recordings.

Prior work in VR accessibility has largely addressed people with visual [59, 64, 71] or mobility impairments [19, 45]. For DHH users, researchers have explored specialized VR apps, for example, to assist with storytelling [15, 63] or teaching [44, 47]. While not formally evaluated in the literature, some commercial VR games also offer accessibility features for DHH users such as subtitles [73] or sound direction indicators [78]. Our work focuses on mainstream VR accessibility, which has advantages of lower cost, increased availability, and better social acceptability over specialized VR experiences; the latter risk stigmatization and tech-abandonment [46, 53, 57].

3 GENERATING THE DESIGN SPACE

To create our design space, we used *morphological search* [54]—a technique inspired from linguistics used to explore relationships in a multi-dimensional setting. In HCI, morphological search has been used to generate the design space of input devices [10] and information visualization [9]. Conceptually, sound to visual/haptic feedback can be modeled as an interaction (in an artificial language) between the dimensions (or a vocabulary) of sound to the dimensions of visual/haptic feedback. Identifying these dimensions is key to modeling the design language of sound to visual/haptic feedback. As with any language, we included both the syntactic and semantics dimensions for sounds, visual, and haptic vocabularies.

3.1 Sound Vocabulary

The dimensions of sounds were inspired from our DIS2021 paper [28] where, informed by game audio typologies [16, 22, 61] and interviews with sound designers, we articulated a taxonomy of sounds covering two dimensions: *source* (*that is*, the origin of the sound—*e.g.*, a character, an object) and *intent* (*that is*, the impact of a sound on the player’s experience—*e.g.*, to increase realism, convey critical information).

For the sound source dimension, we used the concept of *diegesis* [61] to differentiate the sounds emanating from objects in the VR world (*e.g.*, a river flowing) from those playing in the background (*e.g.*, music). In total, we included nine mutually exclusive categories to represent sounds from localized and non-localized sources:

- *Localized speech*: spatially positioned speech (*e.g.*, a character speaking)
- *Non-localized speech*: ambient speech (*e.g.*, a narrator, player thinking aloud).
- *Inanimate objects*: sounds from non-living objects (*e.g.*, weapons, appliances)
- *Animate objects*: non-speech sounds from living beings (*e.g.*, footsteps, animal calls)
- *Interaction sounds*: sounds from interaction between multiple objects (*e.g.*, player touching a menu or punching an enemy)
- *Point ambience*: spatialized ambient sounds that are diegetic—that is, belong to the game world (*e.g.*, a river on one side of the player)

- *Surrounding ambience*: diegetic non-spatialized ambient sounds (*e.g.*, a crowd)
- *Notification sounds*: non-diegetic critical alerts (*e.g.*, low on ammunition, end of a player’s turn).
- *Music*: non-diegetic background music

The second dimension, *sound intent*, contains six categories:

- *Sounds for conveying critical information*: all sounds that are critical for progression in an app (*e.g.*, enemy footsteps, low on ammunition, end of a player’s turn).
- *Sounds for increasing realism*: ambient or objects sounds that increase immersion (*e.g.*, river, vehicles).
- *Sounds for rhythm or movement*: (*e.g.*, music in an exercise or dancing app).
- *Sounds for generating an affective state*: emotional sounds (*e.g.*, stressful sounds when approaching the end of a level, calm music in a meditation app).
- *Sounds for aesthetics or decoration*: non-critical sounds that increase beauty (*e.g.*, background music, sounds accompanying decorative visuals).
- *Sounds for non-critical interaction*: interaction sounds that are not critical to game progression (*e.g.*, picking up a decorative object).

The above two dimensions (source and intent) represent sounds as an abstraction. To successfully substitute sounds, we also need to consider the basic constituents of a sound, each of which can be mapped to a visual or haptic feedback. Thus, we included a third dimension, sound characteristics, containing:

- *Occurrence*: to contextualize a potential visual or haptic mapping that aims to convey just the occurrence of a sound (*e.g.*, through a haptic beep),
- *Identity*: to represent a mapping that conveys identity of a sound (*e.g.*, “enemy footsteps” in text),
- *Location*: to represent a mapping that may convey the sound source location (*e.g.*, a visual arrow),
- *Loudness*: to represent a mapping that may convey the loudness of a sound (*e.g.*, a loudness bar),
- *Pitch*: to represent a mapping that conveys the sound pitch (*e.g.*, through colors),
- *Persistence*: to represent a mapping that conveys the duration of a sound (*e.g.*, through a waveform),
- *Priority*: to represent a mapping that conveys priority of a sound over others (*e.g.*, a narrator speech could be at a higher priority than the background music, which can be conveyed through a priority number written in front of each sound. Sound “priority” is a configurable setting for a Unity sound object.),
- *Environmental effects*: to convey environmental effects such as reverb (*e.g.*, through visual animation)

3.2 Visual Vocabulary

To develop the visual vocabulary, similar to what we did for sound in our earlier sound taxonomy work [4], we started by articulating a dimension to describe the basic constituent or the ingredient of visual feedback. To do so, we relied on past work in real-world sound feedback for DHH users [24, 26, 29] as well as online game forums and blogs listing sound-to-visual feedback suggestions from

DHH gamers [4, 79–82]. We collected the possible ways in which sounds have been represented visually, and then followed an *axial* coding process [14] to group similar visual representations together (e.g., waveform and spectrograms were combined into “time series”). In all, we identified five distinct visual elements:

- *Text*: using text to name sounds (e.g., “footsteps”, “gunshots”)
- *Icon*: using icons such as pulses, arrows to represent sounds
- *Map*: using a map of the location, like a damage indicator in shooting games, to localize the sounds
- *Time series*: a time series visualization to represent variation in a sound characteristic such as loudness (e.g., through a waveform) or pitch (e.g., through a spectrogram) over time
- *Abstract*: any abstract animations for sounds (e.g., sparks originating from a bullet when it hits a wall)

Our second dimension, inspired from a past work in augmented reality based real-world sound visualizations [26], is called *visual view*. It contains two categories based on whether a visualization is (1) affixed to the user’s gaze (called an *overlay*) or (2) placed in the 3D space (called a *cue*).

Beyond the elements and views, the type of VR device heavily affects how and what kind of visual feedback is delivered. For example, to effectively place a visualization in the 3D space, a cue visualization requires a stereo display [1], which is not available on all VR devices. To articulate the many forms of visual VR devices, we added a third dimension called *visual form factor*, containing three categories: (1) *headset* (e.g., Oculus Rift), (2) *smartphone* (e.g., Google Cardboard), and (3) *room-scale* VR.

3.3 Haptic Vocabulary

Compared to visuals, haptic feedback of sounds is much less explored. To articulate visual dimensions, we relied on recent work in real-life sound visualizations and drew parallels for VR. Such a strategy was not possible for haptic feedback because work in this space is more than 20 years old and used outdated form factors (e.g., large haptic belts [55], neck-worn bracelets [18]) which have little resemblance to modern VR haptic devices such as controllers and gamepads. Thus, we mainly relied on the VR experiences of our two DHH authors of this paper, and several online blogs detailing haptic sound feedback suggestions from DHH VR games [4, 79–82]. We articulated three haptic feedback dimensions, with an aim to both represent the current VR haptic sound feedback possibilities as well as to inspire future work in new interfaces for this emerging medium:

First, the device on which the haptic feedback is delivered (called *haptic form factor*), containing four kinds of devices based on emergence:

- VR controller and gamepads (e.g., Oculus controller)
- Other commodity devices that deliver haptic feedback and can be used with VR (e.g., smartwatches)
- Emerging VR haptic feedback device (e.g., haptic belts, vests, and gloves)
- More forward-looking devices (e.g., phantom [83], pneumatic [70])

Second, the location on the body where the haptic feedback is delivered, called *haptic feedback location*, including five categories: (1) *palm*, (2) *arm*, (3) *torso*, (4) *head-mounted*, and (5) *legs*.

Finally, inspired from other HCI haptic work (e.g., [36, 48]), the basic element of delivering haptic feedback, called the *haptic elements*, containing:

- *Intensity*: that is, amplitude or strength of the haptic feedback
- *Timbre*: that is, the sharpness of the feedback, also known as pitch [36],
- *Rhythm*: that is, the spacing between each individual feedback, also known as beats [36].

Note that our overarching goal is to augment, and not substitute sounds. Completely substituting high-bandwidth auditory sense with a low-bandwidth haptic feedback is very challenging, and though past work has attempted solutions (e.g., [18]), none have found practical utility.

3.4 Design Space of VR Sound Accessibility

The design space of VR sound accessibility is the set of possible operations (or mappings) from the dimensions of sound to the dimensions of visual or haptic feedback. We visualize these mappings in a table, such that the columns represent the different sound, visual, or haptic dimensions. Each row is a mapping and represents a possible combination of categories across dimensions. For example, Table 1 shows a sound-to-visual feedback mapping example, called captions for narrator speech, with its targeted categories under different dimensions. Likewise, Table 2 shows an example sound-to-haptic feedback mapping: conveying the loudness of a bomb exploding nearby through vibration intensity on a haptic controller. A designer can generate new mappings by using a different permutation of categories within the dimensions. Of course, not mappings are feasible. Thus, implementation practicalities of new category combinations need to be carefully considered.

4 SOUND ACCESSIBILITY FEATURES IN EXISTING APPS

To evaluate the potential of our design space and to identify gaps for visual and haptic prototype construction, we used our design space to contextualize sound to visual/haptic mappings in existing VR apps and games.

Method: As literature work on VR sound feedback is scarce, we analyzed commonly used commercial apps (e.g., *Tilt Brush*, *Guided Meditation*) and games (e.g., *Minecraft*, *Make Sail*) that contain accessibility features for DHH users. We also investigated non-VR games (PC, smartphone, and video console games) because accessibility integration in VR apps only started recently and some accessibility features from non-VR games may translate well into VR apps. To find these apps, we subscribed to accessibility and games newsletters as well as performed a web search with keywords such as “games deaf accessibility,” and “VR Deaf”.

While searching, we documented the sound to visual/haptic mappings in these apps on a spreadsheet. To find these mappings, the DHH first author either briefly used the apps or read their online documentation. The search concluded after reaching our saturation criteria—that is, when the first author could find any new sound to visual/haptic mapping in the last 10 surveyed apps. In total, 78 apps (51 VR, 27 non-VR) were documented, containing 114 unique mappings. These mappings were either well documented by the

Table 1: An example of sound-to-visual mapping generated using our design space.

Mapping name	Sound source	Sound intent	Sound characteristic	Form factor	Visual View	Visual element
e.g., captions for narrator speech	e.g., non-localized speech	e.g., critical information	e.g., identity (spoken content)	→ e.g., headset	e.g., overlay	e.g., text

Table 2: An example of sound-to-haptic mapping generated using our design space.

Mapping name	Sound source	Sound intent	Sound characteristic	Form factor	Location	Haptic element
e.g., intensity of bomb explosion	e.g., inanimate object (bomb)	e.g., critical information	e.g., loudness	→ e.g., controller	e.g., palm	e.g., intensity

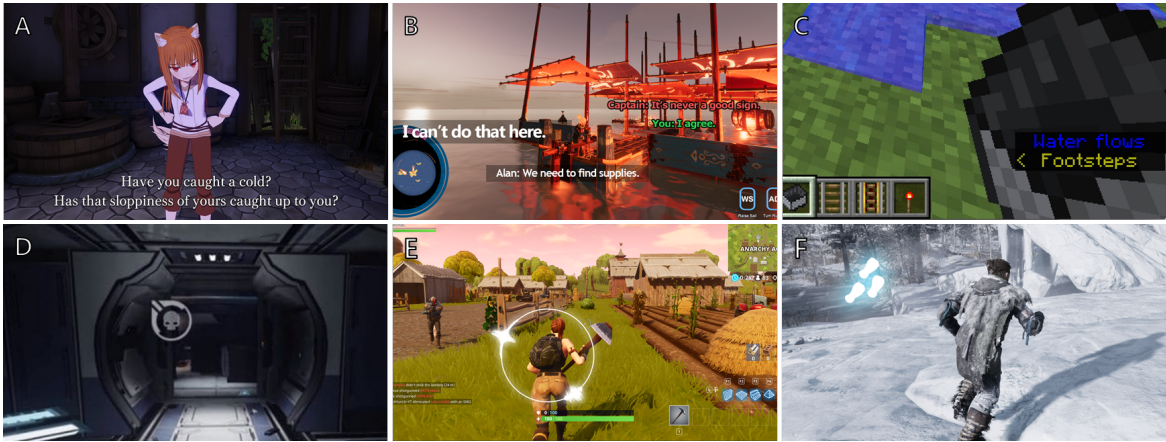


Figure 1: Example VR apps with accessibility features for DHH users: (A) subtitles in *Spice and Wolf VR* [73], (B) spatial captions in *Make Sail* [85], (C) text display for critical sounds in *Minecraft* [86], (D) icons for gunshots in *Persistence* [78], (E) HUD for ambient sounds in *Fortnite* [85], and (F) positional icons for footsteps in *Fade to Silence* [84].

developers or were straightforward to identify, hence additional annotators, beyond the first author, were not needed.

To characterize a mapping, the author used our design space to represent each mapping within the respective sound, visual, or haptic dimensions. For example, visual directional indicators for gunshot sounds in a headset-based app (e.g., Figure 1D) were represented as follows: for “inanimate object” sounds (dimension: sound source) meant “for conveying critical information” (dim: sound intent), the “spatial location” (dim: sound characteristics) was mapped to “headset” (dim: visual form factor), “icon” (dim: visual element), and “cue” (dim: visual view); see Table 3 “positional icons” for a more readable representation.

Results: In sum, our design space was able to cover all mappings we found in the 78 apps. Table 3 and 4 detail the common mappings (appearing in ≥ 5 apps) with examples apps. In summary, the most common visual mappings included non-localized (Figure 1A) or localized (Figure 1B) captions for speech sounds, and text (Figure 1C) or icons (Figure 1D) for non-speech sounds. As expected, haptic mappings were less common than visuals, the most prominent

being the controller vibration for conveying critical information (e.g., [84]) or an interaction cue (e.g., [82]).

5 NEW VR SOUND ACCESSIBILITY PROTOTYPES

Besides contextualizing prior mappings, a design space should also enable new mappings. We used our design space to develop six novel sound-to-visual/haptic mappings, which were turned into VR prototypes.

5.1 Identifying Novel Sound-to-Visual/Haptic Mappings

To identify the gaps for developing novel mappings, we first consulted our findings from the preceding study and online forums with feedback from DHH games [4, 79–82] to determine what sounds and sound characteristics are already accompanied by visual and haptic feedback.

Table 3: Common sound-to-visual mappings in 78 apps we analyzed. We omitted the “visual form-factor” dimension since it was difficult to determine all supported form-factors of an app (e.g., headset, phone) from online survey. *approx. mapping

Custom mapping name	Sound source	Sound intent	Characteristic		Visual element	Visual view
Subtitles (e.g., [87])	non-localized speech	critical information	identity (speech content)	→	text	overlay
Spatially positioned captions (e.g., [85])	localized speech	critical information	identity (speech content)	→	text	cue
Non-speech sounds in text (e.g., [86])	inanimate (e.g., bombs) or animate objects (e.g., footsteps), point ambience (e.g., water)	critical information, realism	identity	→	text	overlay
Colors on the edge of screen to locate danger zones (e.g., [85])	inanimate (e.g., bombs) or animate objects (e.g., enemy footsteps)	affective	identity, location	→	abstract (spatial color gradient)	cue
HUD (e.g., [85], Figure 1E)	multiple	realism	location, loudness, persistence	→	map	overlay
Positional icons (e.g., [84], Figure 1F)	animate (e.g., footsteps) and inanimate (e.g., gunshots) objects	critical information, realism	Identity, location	→	icon (e.g., arrows, arcs, pulses)	cue, overlay
Abstract visuals for critical sounds (e.g., [4])	inanimate objects (e.g., barrel drop, grenade)	critical information	location, loudness	→	abstract (e.g., color streaks)	cue
Sign language avatar (e.g., [88])	non-localized speech	critical information	identity (speech content)	→	icon*	cue
Interaction cue (e.g., [82])	interaction sounds (e.g., punching)	interaction cues	loudness (strength), location	→	abstract / icons (e.g., pulses)	cue

Table 4: Sound to haptic mappings found in the 78 apps we analyzed.

Mapping name	Sound source	Sound intent	Characteristic		Form factor	Location	Haptic element
Controller vibration on occurrence of a sound (e.g., [79])	inanimate objects (e.g., bomb explosion), interaction (e.g., hitting a wall)	critical information, interaction	occurrence	→	gamepad, controller	palm	intensity
Haptic jacket for music beats (e.g., [74])	music	rhythm or movement	loudness	→	custom (jacket)	torso	intensity, pitch, rhythm

We found that, for visual cues, interaction sounds, point ambience, and object-based sounds (animate or inanimate) are—by definition—always visible (that is, an object or a punch can be seen), unless they are outside the field-of-view. Although rare, speech sounds are shown as localized transcriptions (for localized speech—e.g., see Figure 1B) or subtitles (for non-localized speech such as a narration, e.g., Figure 1A). In contrast, surrounding ambience and music sounds are not visible except when a subtle visual indication is present (e.g., leaves rustling), which are difficult to notice. Finally, while notification sounds are sometimes visualized, such as by a text indicating the end of a player’s turn, in cases where a visual indicator was absent (e.g., in a 3D chess), it led to a critical performance breakdown.

Regarding haptics, sounds were rarely accompanied by haptic feedback, except in a few important cases—for example, to convey

gunshots in shooting games or critical interactions (e.g., a punch), both using controller vibrations. One game developer offered an option to purchase a haptic jacket for conveying music beats [74].

Within the eight sound characteristics, the visual cues almost always convey identity or location, and in some cases, loudness or persistence (e.g., visualization in Figure 1E shows persistence of ambient sounds). We could not find any examples of conveying pitch, priority, or environmental effects, which are important elements in VR sound design. The haptic cues mainly convey the occurrence, or in case of the haptic jacket, loudness.

Based on these insights, we identified 13 novel sound-to-visual/haptic mappings for less represented sounds (surrounding ambience, music, notification sounds) and sound characteristics (e.g., pitch, priority, persistence). Then, based on discussions with the VR developers and the two DHH researchers on our team, we

Table 5: The four novel sound-to-visual mappings identified from our design space. Each mapping can be prototyped for multiple visual form factors (e.g., headsets, smartphones).

Mapping name	Sound source	Sound intent	Characteristic		Visual element	Visual view
Text to display notification sounds	notification sounds	critical information	identity	→	text	cue
Text to display priority of all sounds	all	all	priority	→	text	overlay
Waveform for loudness of ambient sounds	surrounding ambience	increasing realism	loudness	→	waveform	overlay
Abstract visualization for music	music	rhythm or movement	loudness, persistence	→	abstract	cue

Table 6: The two novel sound-to-haptic mappings identified from our design space.

Mapping name	Sound source	Sound intent	Characteristic		Form factor	Location	Haptic element
Wideband actuators for ambient sound feedback	surrounding ambience, point ambience, music	increasing realism	loudness, location, persistence, pitch	→	custom	torso	intensity, timbre, rhythm
Directional beeps to localize sounds	multiple	critical information	location	→	commodity (smart-watch)	arm (wrist)	rhythm

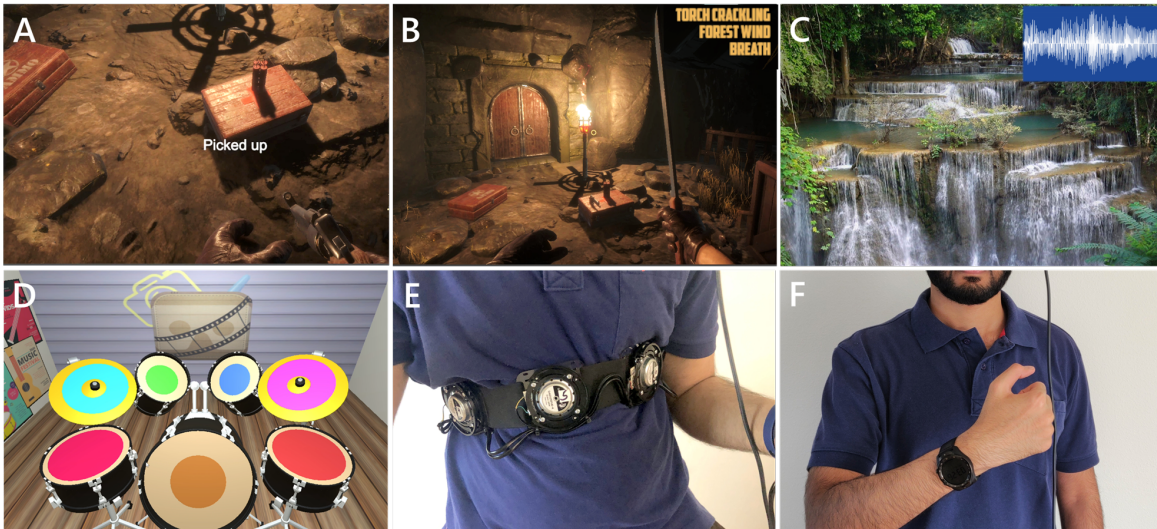


Figure 2: Example use cases of our four visual (A–D) and two haptic (E–F) prototypes to support sound accessibility: (A) text display for notification sounds, (B) text display for all sounds that are currently playing, arranged vertically in the order of high to low priority (priority is a default Unity metadata [89]), (C) waveform of ambient sounds, (D) abstract visualization (circles) for rhythm sounds, (E) wideband haptic belt for ambient sounds, and (F) smartwatch app that vibrates to localize critical information sounds. See our supplementary video for working examples of our prototypes.

filtered out seven mappings, leaving six (four sound-to-visual, two sound-to-haptic) that seemed useful and feasible to prototype with the current technology. These six mappings appear in detail in Table 5 and 6 and include, in brief: text to display notification sounds, text

to display priority of all sounds, a waveform to visualize surrounding ambience sounds, an abstract visualization for music sounds, wideband haptic feedback to convey multiple characteristics (e.g., persistence, pitch), and directional vibrational beeps to convey location of sounds.

5.2 Visual and Haptic Prototypes

To instantiate the six new mappings, we designed visual and haptic prototypes for DHH users (Figure 2). These prototypes are designed as augmentations—that is, they can retrofit to any Unity-based VR app. The four visual prototypes append sound visualizations (text, icons, or an abstract visualization) to an app. The two haptic prototypes use external hardware, a custom haptic belt or a commodity smartwatch, to deliver wideband or vibrational haptic feedback, respectively. These prototypes are customizable—that is, developers can customize the visual or haptic feedback rendering to harmonize with the aesthetic or design of their VR app. To assess robustness and scalability, each prototype was briefly tested by a DHH team member on three different VR apps. The code for our prototypes is available on our project page: aka.ms/TowardsSoundAccessibilityInVR.

Visual prototypes: We built four visual prototypes as C# scripts: (1) identity of notification sounds to text, (2) priority of all sounds to text, (3) loudness of ambient sounds to a waveform, and (4) beats of rhythm sounds to abstract visualization (see Figure 2A–D and our video). Developers can use a prototype by attaching its script to a sound object (called *AudioSource* [90]) in Unity. When a script is attached, a default visualization is generated, which can be customized in two ways: (1) by configuring the visualization’s metadata (e.g., size, color, shape) in the Unity editor, or (2) by making a new, custom visualization by extending our base C# class.

Haptic prototypes: We built two haptic prototypes (see Figure 2E, 3F and video): (1) a custom haptic belt to convey the complex pattern (pitch, persistent, loudness) of ambient sounds, and (2) an app for commercially available smartwatches to convey direction of critical information sounds.

Our haptic belt, worn on the chest or abdomen, uses eight wideband (voice coil) actuators [6] arranged in tandem: two each on the front, right, left, and rear side of the torso. Our approach of using wideband haptic actuators for sound feedback is novel; prior work has only explored single-frequency vibrotactile motors (e.g., LRA) to convey simple characteristics (i.e., occurrence, loudness) of environmental sounds [25, 39]. In contrast, wideband actuators deliver high acceleration over a wide frequency range to provide very granular haptic feedback [91]. Similar to the visual prototypes, our haptic belt is driven by C# scripts that can be attached to any sound object in Unity. We offer three modules for developers to customize the feedback for their VR apps: (1) a non-directional module which triggers all haptic actuators with the same intensity proportional to the volume of a sound, (2) a directional module that map actuators based on the direction of a sound (e.g., if a sound source is to the left of the user, the left actuators’ intensity will be greater than the right one), and (3) a base class that can be extended to design custom more sophisticated haptic feedback.

We assembled the haptic belt using off-the-shelf components. For the actuators, we use eight *Dayton 24W Exciters* [92] attached at equal distances to a Velcro belt. Each actuator pair (left, right, front, and rear) is driven through a 2-channel 12V amplifier [93] and a USB sound card [94]. These actuator pairs appear in Unity as stereo audio output devices and can be configured using our custom Unity environment script.

The second haptic prototype is an Android app for commercial smartwatches that uses vibration patterns to convey the location of critical information sounds—a key desired feature by DHH people in online game forums [79, 81, 82]. Use of a commercial device enabled us to explore a simple, more portable alternative to the haptic belt at the cost of feedback granularity. Because commercial watches only contain a single vibration motor, we used a strategy devised in Pielot *et al.* [51] to convey directional feedback: one vibration beep indicates a sound source to the left of the user, and two beeps indicates a source to the right. These watch beeps are controlled, over a WiFi connection, through a C# script attached to any Unity sound object.

5.2.1 Implications and Comparison to Prior Work. Our prototypes first and foremost demonstrate the potential of our design space to generate new mappings. Furthermore, while prior work has investigated VR accessibility for DHH users, the focus has been on making VR experiences for specialized tasks such as storytelling [15, 63] or teaching [44, 47]. Our generic visual and haptic prototypes can be augmented to any VR app. Finally, our haptic belt prototype offers a fundamentally new way to map sounds to VR haptic feedback, using wideband actuators to convey multiple sound priorities (pitch, persistence, loudness, location); prior work for DHH users has only used small single-frequency vibration motors to convey either loudness [25, 62] or just the occurrence of a sound [18, 31, 50].

5.3 Preliminary Evaluations

To evaluate our prototypes, we conducted preliminary studies with 11 DHH individuals and four VR developers.

5.3.1 Evaluation with DHH individuals. Due to the COVID-19 pandemic, in-person studies were not feasible. As well, remote studies with VR hardware are difficult. Thus, we made example videos of our visual prototypes and collected feedback over email.

Participants: We recruited 11 DHH participants (P1–P11, five men, four women, two non-binary) using email lists and snowball sampling. Participants were on average 40.4 years old ($SD=15.0$). Four participants identified as hard of hearing, four as deaf, and three as Deaf. Seven participants had used a VR headset before.

Procedure: We emailed a brief description of each visual prototype, two example videos of its application (see supplementary video) and asked two open-ended questions on potential usefulness and improvements.

Data analysis: We conducted an iterative coding process [8] on the email responses. One researcher scanned the responses, developed an initial codebook, and iteratively applied the codes to the responses while refining the codebook. After convergence, a second researcher coded all responses using the final codebook. The interrater agreement, measured using Krippendorff’s alpha [33], was on average 0.81 ($SD=0.13$, $range=0.69-1$, $raw\ agreement=94.4\%$). Conflicting code assignments were resolved through consensus.

Findings: Most DHH participants (9/11) found two of our four visual prototypes—“priority of all sounds to text” (Figure 2B) and “loudness of ambient sounds to waveform” (Figure 2C)—to be useful. For the other two prototypes, “identity of notification sounds to text” (Figure 2A) and “rhythm sounds to abstract visualization” (Figure 2D), participants (5/11) were skeptical of their interference with

the aesthetics of the VR apps, which could diminish immersion. For example,

“I am not sure but this text-pop up [of notification sounds] could take me out of the scene and diminish immersion. [Also], what if there are a lot of sounds and we have a big text box which looks awkward. . .” (R3)

Another set of three participants were concerned that these prototypes may reduce the game challenge:

“So, it’s nice to have this notification [to text]. But I wonder if showing too much information may diminish the competition. [...] There could be a very subtle sound for a gun shooting from behind me [...] Maybe that’s how the [game] developer wanted it to be. And if you show me this stuff [in text], I may know too much.” (P7)

Regardless of these concerns, all participants weighed accessibility as more important than the concerns of aesthetics or challenge and provided suggestions to overcome these limitations. For example, P7 added: *“one idea is to have a bar that specifies the [game] difficulty level. [...] As the level goes up, show me less amount of information.”* Similarly, for “notification sounds to text prototype” (Figure 2A), P9 said: *“the text might be annoying only if there are a lot of things to [notify]. So, [...] having an option to turn off when I want will help.”*

In the above two quotes, P7 and P9 also reinforced the need for end-user customization. Indeed, customization (e.g., enable/disable prototype, change color/size) was a key feature desired by all participants. For example, *“the breath [sound] is nice to know if I am walking in a jungle or something, but when I am shooting at someone, I only want [to know about] the sounds relevant to the game play.”* (P2) She added: *“but then I am hard of hearing, and I can make out the breath if I am paying attention. I am curious how Deaf folks feel about the noises they need to be aware of.”* Indeed, P8, who is profoundly deaf, wrote: *“I can’t hear the breath sound. So, it was very interesting for me to know that I am breathing.”* These examples also show that end-user customization can help accommodate the varying hearing levels of DHH users.

5.3.2 Preliminary Evaluation with VR Developers. To examine the usability and generalizability of our prototype scripts, we obtained feedback from four VR developers on their experience with integrating our prototypes in their Unity apps.

Participants: We recruited four hearing VR developers (D1-D4, two men, two women). Participants were on average 36.2 years of age ($SD=4.8$, $range=31-44$), and had at least one year of VR development experience.

Procedure: For each visual and haptic prototype, we emailed the C# script, a brief description, and instructions to incorporate it into a Unity app. Participants added the six prototypes into at least one Unity-based VR app they were working on and reported on their experience over email.

Data analysis: We followed the same iterative coding process as for the study with DHH users. Average Krippendorff’s alpha between the two coders was 0.76 ($SD=0.18$, $range=0.67-1$, $raw\ agreement=88.7\%$).

Findings: None of the participants had ever thought about accessibility whilst working on a VR project, highlighting the scarceness of this important space. In fact, they were unaware of any guidelines or tools to support VR accessibility. When evaluating our prototypes, all four developers found all six prototypes easy to understand and integrate into VR apps with minimal additional workload. D2 was excited and exclaimed: *“Can I continue to use them after the study!?”* D2 and D3 mentioned further use cases of our prototypes: for VR audio debugging and for hearing users—for example, *“to visualize music beats when I am not wearing headphones.”*

All developers particularly liked the option to customize the prototype visualizations, mentioning, for example, *“these will help me configure the design [to provide] the best experience for my clients.”* D1 and D3 suggested releasing a toolkit—e.g., as a Unity prefab [95]—for easy and wide adoption of sound accessibility into VR apps.

6 DISCUSSION

Too often, accessibility is only considered as an afterthought, resulting in inaccessible or sub-par user experiences [66]. As Mott et al. point out in a recent position paper [41], VR technologies are at a crossroads in time where there is still an opportunity to codify accessibility best-practices into this emerging medium. While researchers have begun to consider making VR accessible to those with diverse visual [59, 64, 71] and motor [45] abilities, the needs of DHH users are as-yet-unexplored. In this work, we have presented a first comprehensive look at sound accessibility in VR. Our contributions include: (1) a novel design space for VR sound accessibility that was able to all contextualize prior sound-to-visual haptic feedback mappings in recent VR apps and games as well as generate new mappings, (2) six visual and haptic sound feedback prototypes that can augment to any VR app unlike prior work, (3) and insights from preliminary evaluations of these prototypes, yielding tensions among accessibility, aesthetic, game challenge, and immersion, which can generalize to any VR accessibility work. Below, we reflect on further implications and key limitations of our work.

Further evaluations: We refer to the evaluations of our prototypes as “preliminary” because their format (email-based evaluations through videos and code sharing) was constrained by the COVID-19 pandemic. In addition, due to the high cost and emerging nature of VR, very few people own VR technologies (and fewer still from disability demographics, due to the inaccessibility of most mainstream VR apps), which makes distributed usability studies of VR challenging. Nevertheless, our prototypes were informed by two DHH HCI researchers on our team and we are planning further in-person studies once the social distancing guidelines are relaxed. Our preliminary findings helped identify more concrete research questions to explore in in-person studies, which are: (1) How do the sound-to-visual/haptic prototypes change the VR experience of DHH users? (2) What is the effect on cognitive overload, particularly when combining multiple feedback prototypes? and (3) How much is the original experience (e.g., immersion, game challenge) preserved?

Prototype explorations: We prototyped simple sound to visual/haptic mappings to demonstrate the potential of our design

space. Future work should investigate more sophisticated prototypes. For example, our haptic belt, which directly mapped raw ambient sound through sound cards, can be extended by supporting more granular mappings of sound characteristics (e.g., pitch, priority, reverb) to the haptic actuators. Another unexplored area is “affective sounds”, which are important to convey emotional state in VR but are rarely represented by an alternative modality. Finally, our evaluations explored each prototype in isolation, and how best to combine feedback for multiple, simultaneous sounds is an open question. A key consideration for simultaneous feedback is to avoid cognitive overload, particularly when visual and haptic feedback are jointly delivered. One possibility is to explore high-attentional haptic feedback for initial necessary alerts (e.g., notifying about sound occurrence) and high-bandwidth visual feedback to convey additional information (e.g., identity).

Personalization: Another avenue for exploration is options for personalization of prototypes. VR developers can customize the default visualizations or haptic feedback of our prototypes to match the app’s aesthetics and goals; however, it may then be desirable for the end-user to have some additional customization options based on their preferences and hearing abilities (e.g., what sound frequencies they can hear best, asymmetric abilities across ears). Pairing of VR hardware with users’ assistive devices, such as hearing aids or cochlear implants, to provide a seamless experience that customizes itself to the capabilities of that hardware (e.g., optimized for certain frequencies) is another possibility. However, personalization comes with tradeoffs, including end-user effort in creating a profile of preferences, privacy considerations around revealing one’s hearing abilities, and security concerns in trusting third-party devices to pair safely with personal assistive technologies.

Applications for other domains: While we targeted accessibility for DHH users, our work can also support other disabilities, such as those with cognitive or visual impairment. For example, people with visual impairment are often overloaded with sonic information [68] and may appreciate mapping some sounds to haptic feedback. Our work can also benefit hearing users in cases of situational impairments [56] (e.g., in noisy environments) and cognitive overload (e.g., when a user is feeling exhausted). Finally, sound feedback can support important applications in other domains such as home automation, wildlife surveys, and appliance repairs.

Limitations: While our design space showed promise in contextualizing prior mappings and generating ideas for new mappings, we do not claim that it is exhaustive. Indeed, VR sound design technology is still in its infancy, and as technologies evolve, so should related ontologies. We welcome future work that refines, extends, or reimagines our design space. Furthermore, we acknowledge our work may not be desired by all DHH people, since some do not want sound feedback [13, 41]. At the same time, we argue that the DHH community is diverse [13] and past large-scale surveys [8, 23] as well as the experiences of two DHH authors of this paper suggest that many DHH people appreciate sound information. Nevertheless, future work should also explore non-sound related accessibility features desired by DHH users such as background blurring to focus visual attention [117].

7 CONCLUSION

Ensuring that mainstream VR applications are accessible to people with a spectrum of hearing capabilities is an important and largely unexplored research challenge [41]. In this paper, we used multiple methods (including analysis of *status quo* VR apps, development and refinement of ontologies, prototyping, and online user studies) and engaged a variety of stakeholders (including DHH and hearing research team members, DHH end-users, and Unity developers) to formally characterize the state of sound accessibility in VR as well as lay a groundwork for accessible sound representations in this emerging medium. Our work advances sound accessibility in VR by developing a design space for mapping sound to visual and haptic feedback and innovating six visual and haptic VR prototypes to support DHH users.

REFERENCES

- [1] Kurt Akeley, Simon J Watt, Ahna Reza Girshick, and Martin S Banks. 2004. A stereo display prototype with multiple focal distances. *ACM transactions on graphics (TOG)* 23, 3: 804–813.
- [2] Matilda Annerstedt, Peter Jönsson, Mattias Wallergård, Gerd Johansson, Björn Karlson, Patrik Grahn, Åse Marie Hansen, and Peter Währborg. 2013. Inducing physiological stress recovery with sounds of nature in a virtual reality forest—Results from a pilot study. *Physiology & behavior* 118: 240–250.
- [3] Beatriz Pozo Arcos, Conny Bakker, Bas Flipsen, and Ruud Balkenende. 2020. Practices of fault diagnosis in household appliances: Insights for design. *Journal of Cleaner Production*: 121812.
- [4] Morgan Baker. Gamasutra: Deaf Accessibility in Video Games. Retrieved September 6, 2020 from https://www.gamasutra.com/blogs/MorganBaker/20200720/366615/Deaf_Accessibility_in_Video_Games.php
- [5] Durand R Begault. 2000. *3-D Sound for Virtual Reality and Multimedia*. Retrieved from <https://ntrs.nasa.gov/api/citations/20010044352/downloads/20010044352.pdf?attachment=true>
- [6] Bill Black, Mike Lopez, and Anthony Morcos. 1993. Basics of voice coil actuators. *PCIM-VENTURA CA*: 19: 44.
- [7] Danielle Bragg, Nicholas Huynh, and Richard E. Ladner. 2016. A Personalizable Mobile Sound Detector App Design for Deaf and Hard-of-Hearing Users. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, 3–13.
- [8] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2: 77–101.
- [9] Stuart K Card and Jock Mackinlay. 1997. The structure of the information visualization design space. In *Proceedings of VIZ'97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*, 92–99.
- [10] Stuart K Card, Jock D Mackinlay, and George G Robertson. 1990. The design space of input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 117–124.
- [11] David J Chalmers. 2017. The virtual and the real. *Disputatio* 9, 46: 309–352.
- [12] Kai Håkon Christensen, Johannes Röhrs, Brian Ward, Ilker Fer, Göran Broström, Øyvind Saetra, and Øyvind Breivik. 2013. Surface wave measurements using a ship-mounted ultrasonic altimeter. *Methods in Oceanography* 6: 1–15.
- [13] Peter Ciuha, Bojan Klemenc, and Franc Solina. 2010. Visualization of concurrent tones in music with colours. In *Proceedings of the 18th ACM international conference on Multimedia*, 1677–1680.
- [14] Juliet M Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13, 1: 3–21.
- [15] Sigal Eden and Sara Ingber. 2014. Enhancing storytelling ability with virtual environment among deaf and hard-of-hearing children. In *International Conference on Computers for Handicapped Persons*, 386–392.
- [16] Troels Folmann. 2004. Dimensions of game audio. Unpublished. Available at <http://www.itu.dk/people/folmann/2004/11/dimensions-of-game-audio.html> [Accessed October 28, 2005.].
- [17] Simon Franco. 2016. In-Game Audio Debugging Tools. In *Game Development Tools*. AK Peters/CRC Press, 161–184.
- [18] Karyn L. Galvin, Jan Gimis, Robert S. Cowan, Peter J. Blamey, and Graeme M. Clark. 2001. A Comparison of a New Prototype Tickle Talker™ with the Tactaid 7. *Australian and New Zealand Journal of Audiology* 23, 1: 18–36.
- [19] Kathrin Gerling and Katta Spiel. 2021. A Critical Examination of Virtual Reality Technology in the Context of the Minority Body. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA.

- [20] Steven Goodman, Susanne Kirchner, Rose Guttman, Dhruv Jain, Jon Froehlich, and Leah Findlater. Evaluating Smartwatch-based Sound Feedback for Deaf and Hard-of-hearing Users Across Contexts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1–13.
- [21] Ru Guo, Yiru Yang, Johnson Kuang, Xue Bin, Dhruv Jain, Steven Goodman, Leah Findlater, and Jon Froehlich. 2020. HoloSound: Combining Speech and Sound Identification for Deaf or Hard of Hearing Users on a Head-mounted Display. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, 1–4.
- [22] Sander Huiberts. 2010. Captivating sound the role of audio for immersion in computer games. University of Portsmouth.
- [23] John W Hunt, Marcel Arditi, and F Stuart Foster. 1983. Ultrasound transducers for pulse-echo medical imaging. *IEEE Transactions on Biomedical Engineering*, 8: 453–481.
- [24] Dhruv Jain, Bonnie Chinh, Leah Findlater, Raja Kushalnagar, and Jon Froehlich. 2018. Exploring Augmented Reality Approaches to Real-Time Captioning: A Preliminary Autoethnographic Study. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems*, 7–11.
- [25] Dhruv Jain, Brendon Chiu, Steven Goodman, Chris Schmandt, Leah Findlater, and Jon E Froehlich. 2020. Field study of a tactile sound awareness device for deaf users. In *Proceedings of the 2020 International Symposium on Wearable Computers*, 55–57.
- [26] Dhruv Jain, Leah Findlater, Christian Volger, Dmitry Zotkin, Ramani Duraiswami, and Jon Froehlich. 2015. Head-Mounted Display Visualizations to Support Sound Awareness for the Deaf and Hard of Hearing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 241–250.
- [27] Dhruv Jain, Rachel Franz, Leah Findlater, Jackson Cannon, Raja Kushalnagar, and Jon Froehlich. 2018. Towards Accessible Conversations in a Mobile Context for People Who are Deaf and Hard of Hearing. In *ASSETS 2018 - Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 81–92.
- [28] Dhruv Jain, Sasa Junuzovic, Eyal Ofek, Mike Sinclair, John Porter, Chris Yoon, Swetha Machanavajhala, and Meredith Ringel Morris. 2021. A Taxonomy of Sounds in Virtual Reality. In *Designing Interactive Systems (DIS) 2021*, 160–170.
- [29] Dhruv Jain, Angela Carey Lin, Marcus Amalachandran, Aileen Zeng, Rose Guttman, Leah Findlater, and Jon Froehlich. 2019. Exploring Sound Awareness in the Home for People who are Deaf or Hard of Hearing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 94:1-94:13.
- [30] Dhruv Jain, Kelly Mack, Akli Amrous, Matt Wright, Steven Goodman, Leah Findlater, and Jon E Froehlich. 2020. HomeSound: An Iterative Field Deployment of an In-Home Sound Awareness System for Deaf or Hard of Hearing Users. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*, 1–12.
- [31] Dhruv Jain, Hung Ngo, Pratyush Patel, Steven Goodman, Leah Findlater, and Jon Froehlich. 2020. SoundWatch: Exploring Smartwatch-based Deep Learning Approaches to Support Sound Awareness for Deaf and Hard of Hearing Users. In *ACM SIGACCESS conference on Computers and accessibility*, 1–13.
- [32] Maria Karam, Carmen Branje, Gabriel Nespoli, Norma Thompson, Frank Russo, and Deborah Fels. 2010. The emoti-chair: An interactive tactile music exhibit. 3069–3074.
- [33] Klaus Krippendorff. 2018. *Content analysis: An introduction to its methodology*. Sage publications.
- [34] Arpi Mardirossian and Elaine Chew. 2007. Visualizing Music: Tonal Progressions and Distributions. In *ISMIR*, 189–194.
- [35] Tiago A Marques, Len Thomas, Stephen W Martin, David K Mellinger, Jessica A Ward, David J Moretti, Danielle Harris, and Peter L Tyack. 2013. Estimating animal population density using passive acoustics. *Biological Reviews* 88, 2: 287–309.
- [36] Martyn Reding. Designing haptic responses. Retrieved September 6, 2020 from <https://medium.com/@martynreding/basics-of-designing-haptic-responses-63dc6b52e010>
- [37] Tara Matthews, Janette Fong, F. Wai-Ling Ho-Ching, and Jennifer Mankoff. 2006. Evaluating non-speech sound visualizations for the deaf. *Behaviour & Information Technology* 25, 4: 333–351. <https://doi.org/10.1080/01449290600636488>
- [38] Matthias Mielke and Rainer Brueck. 2015. Design and evaluation of a smartphone application for non-speech sound awareness for people with hearing loss. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 5008–5011.
- [39] Mohammadreza Mirzaei, Peter Kan, and Hannes Kaufmann. 2020. EarVR: Using ear haptics in virtual reality for deaf and Hard-of-Hearing people. *IEEE Transactions on Visualization and Computer Graphics* 26, 5: 2084–2093.
- [40] Reiko Miyazaki, Issei Fujishiro, and Rumi Hiraga. 2003. Exploring MIDI datasets. In *ACM SIGGRAPH 2003 Sketches & Applications*. 1.
- [41] Martez Mott, Edward Cutrell, Mar Gonzalez Franco, Christian Holz, Eyal Ofek, Richard Stoakley, and Meredith Ringel Morris. 2019. Accessible by design: An opportunity for virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 451–454.
- [42] Suranga Chandima Nanayakkara, Lonce Wyse, S. H. Ong, and Elizabeth A. Taylor. 2013. Enhancing Musical Experience for the Hearing-Impaired Using Visual and Haptic Displays. *Human-Computer Interaction* 28, 2: 115–160.
- [43] Suranga Nanayakkara, Elizabeth Taylor, Lonce Wyse, and S H Ong. 2009. An enhanced musical experience for the deaf: design and evaluation of a music display and a haptic chair. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 337–346.
- [44] S C W Ong and S Ranganath. 2005. Automatic sign language analysis: a survey and the future beyond lexical meaning. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27, 6: 873–891.
- [45] Shanmugam Muruga Palaniappan, Ting Zhang, and Bradley S Duerstock. 2019. Identifying Comfort Areas in 3D Space for Persons with Upper Extremity Mobility Impairments Using Virtual Reality. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 495–499.
- [46] Phil Parette and Marcia Scherer. 2004. Assistive Technology Use and Stigma. *Education and Training in Developmental Disabilities-September 2004*: 217–226.
- [47] David Passig and Sigal Eden. 2001. Virtual reality as a tool for improving spatial rotation among deaf and hard-of-hearing children. *CyberPsychology & Behavior* 4, 6: 681–686.
- [48] Mark Paterson. 2017. On haptic media and the possibilities of a more inclusive interactivity. *New Media & Society* 19, 10: 1541–1562.
- [49] Yi-Hao Peng, Ming-Wei Hsu, Paul Taele, Ting-Yu Lin, Po-En Lai, Leon Hsu, Tzu-chuan Chen, Te-Yen Wu, Yu-An Chen, Hsien-Hui Tang, and Mike Y. Chen. 2018. SpeechBubbles: Enhancing Captioning Experiences for Deaf and Hard-of-Hearing People in Group Conversations. In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, Paper No. 293.
- [50] A J Phillips, A R D Thornton, S Worsfold, A Downie, and J Milligan. 1994. Experience of using vibrotactile aids with the profoundly deafened. *European journal of disorders of communication* 29, 1: 17–26.
- [51] Martin Pielot, Benjamin Poppinga, Wilko Heuten, and Susanne Boll. 2011. A tactile compass for eyes-free pedestrian navigation. In *IFIP Conference on Human-Computer Interaction*, 640–656.
- [52] Ilyas Potamitis, Stavros Ntalampiras, Olaf Jahn, and Klaus Riede. 2014. Automatic bird sound detection in long real-field recordings: Applications and tools. *Applied Acoustics* 80: 1–9.
- [53] Marti L Riemer-Reiss and Robbyn R Wacker. 2000. Factors associated with assistive technology discontinuance among individuals with disabilities. *Journal of Rehabilitation* 66, 3.
- [54] Tom Ritchey. 2011. General morphological analysis (GMA). In *Wicked problems—Social messes*. Springer, 7–18.
- [55] Frank A Saunders, William A Hill, and Barbara Franklin. 1981. A wearable tactile sensory aid for profoundly deaf children. *Journal of Medical Systems* 5, 4: 265–270.
- [56] Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. 2003. When computers fade: Pervasive computing and situationally-induced impairments and disabilities. In *HCI international*, 1298–1302.
- [57] Kristen Shinohara and JO Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 705–714.
- [58] Liu Sicong, Zhou Zimu, Du Junzhao, Shangguang Longfei, Jun Han, and Xin Wang. 2017. UbiEar: Bringing Location-independent Sound Awareness to the Hard-of-hearing People with Smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2: 17.
- [59] Alexa F Siu, Mike Sinclair, Robert Kovacs, Eyal Ofek, Christian Holz, and Edward Cutrell. 2020. Virtual Reality Without Vision: A Haptic and Auditory White Cane to Navigate Complex Virtual Worlds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13.
- [60] Sean M Smith and Glen N Williams. 1997. A visualization of music. In *Proceedings, Visualization '97 (Cat. No. 97CB36155)*, 499–503.
- [61] Axel Stockburger. 2003. The game environment from an auditive perspective. *Level Up*: 4–6.
- [62] I R Summers, M A Peake, and M C Martin. 1981. Field trials of a tactile acoustic monitor for the profoundly deaf. *British journal of audiology* 15, 3: 195–199.
- [63] Mauro Teófilo, Alvaro Lourenço, Juliana Postal, and Vicente F Lucena. 2018. Exploring virtual reality to enable deaf or hard of hearing accessibility in live theaters: A case study. In *International Conference on Universal Access in Human-Computer Interaction*, 132–148.
- [64] Ryan Wedoff, Lindsay Ball, Amelia Wang, Yi Xuan Khoo, Lauren Lieberman, and Kyle Rector. 2019. Virtual showdown: An accessible virtual reality game with scaffolds for youth with visual impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15.
- [65] Janet M Weisenberger, Susan M Broadstone, and Frank A Saunders. 1989. Evaluation of two multichannel tactile aids for the hearing impaired. *The Journal of the Acoustical Society of America* 86, 5: 1764–1775.
- [66] Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-Based Design: Concept, Principles and Examples. *ACM Trans. Access. Comput.* 3, 3: 9:1–9:27. <https://doi.org/10.1145/1952383.1952384>
- [67] Lining Yao, Yan Shi, Hengfeng Chi, Xiaoyu Ji, and Fangtian Ying. 2010. Music-touch Shoes: Vibrotactile Interface for Hearing Impaired Dancers. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '10)*, 275–276.

- [68] Koji Yatani, Nikola Banovic, and Khai Truong. 2012. SpaceSense: representing geographical information to visually impaired people using spatial tactile feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 415–424.
- [69] Eddy Yeung, Arthur Boothroyd, and Cecil Redmond. 1988. A Wearable Multi-channel Tactile Display of Voice Fundamental Frequency. *Ear and Hearing* 9, 6: 342–350.
- [70] Kyle T Yoshida, Cara M Nunez, Sophia R Williams, Allison M Okamura, and Ming Luo. 2019. 3-DoF Wearable, Pneumatic Haptic Device to Deliver Normal, Shear, Vibration, and Torsion Feedback. In *2019 IEEE World Haptics Conference (WHC)*, 97–102.
- [71] Yuhang Zhao, Cynthia L Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. 2018. Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, 1–14.
- [72] Jie Zhu, Johan Christensen, Jesper Jung, Luis Martin-Moreno, X Yin, Lee Fok, Xiang Zhang, and F J Garcia-Vidal. 2011. A holey-structured metamaterial for acoustic deep-subwavelength imaging. *Nature physics* 7, 1: 52–55.
- [73] Spice & Wolf VR Leaves You Wanting For More Of Lawrence And Holo's Sweet Interactions. Retrieved September 8, 2020 from <https://www.siliconera.com/spice-wolf-vr-leaves-you-wanting-for-more-of-lawrence-and-holos-sweet-interactions/>
- [74] Beat Saber and SUBPAC Bring VR Rhythm Game to Hearing Impaired. Retrieved September 8, 2020 from <https://about.fb.com/news/2020/03/vr-game-for-hearing-impaired/>
- [75] How Visuals Can Help Deaf Children "Hear" | Live Science. Retrieved August 18, 2020 from <https://www.livescience.com/47004-visuals-help-deaf-children-experience-sound.html>
- [76] Home | XR Access Initiative. Retrieved September 19, 2020 from <https://xraccess.org/>
- [77] XR Accessibility User Requirements. Retrieved September 19, 2020 from <https://www.w3.org/TR/xaur/>
- [78] The Persistence for PS VR - PlayStation.Blog. Retrieved September 20, 2021 from <https://blog.playstation.com/2018/10/11/the-persistence-for-ps-vr-gets-huge-free-update-october-18/>
- [79] Valorant: Deaf Accessibility Case Study – Morgan L. Baker. Retrieved September 8, 2020 from https://leahybaker.com/valorant_access/
- [80] Sign Language In VR 'Worth Exploring' As Hand Tracking Improves. Retrieved September 8, 2020 from <https://uploadvr.com/sign-language-vr-asl/>
- [81] Enabling VR for the visual and hearing impaired. What are methods to improve their experiences?: oculus. Retrieved September 8, 2020 from https://www.reddit.com/r/oculus/comments/22u0z6/enabling_vr_for_the_visual_and_hearing_impaired/
- [82] Insights Into VR Gaming and Esports from a Deaf Player. Retrieved September 8, 2020 from <https://www.vrfitnessinsider.com/insights-into-vr-gaming-and-esports-from-a-deaf-player/>
- [83] Phantom Premium Haptic Devices. Retrieved April 12, 2021 from <https://www.3dsystems.com/haptics-devices/3d-systems-phantom-premium>
- [84] Deaf Game Review - State of Decay 2. Retrieved September 8, 2020 from <https://caniplaythat.com/2019/02/11/deaf-game-review-state-of-decay-2/>
- [85] How Can Developers Build Accessibility Bridges for Deafness and Gaming? Retrieved September 6, 2020 from <https://www.dualshockers.com/how-can-developers-make-games-accessible-for-deaf-gamers/>
- [86] Minecraft Accessibility | Minecraft. Retrieved September 8, 2020 from <https://www.minecraft.net/en-us/accessibility>
- [87] Microsoft Flight Simulator Accessibility Review – Flying with Assistance. Retrieved September 8, 2020 from <https://caniplaythat.com/2020/08/17/microsoft-flight-simulator-accessibility-review-pc/>
- [88] Gaming's favorite VR mouse uses sign language in the cutest way. Retrieved September 8, 2020 from <https://www.polygon.com/2017/8/3/16089720/moss-vr-sign-language-in-games>
- [89] Unity - Scripting API: AudioSource.priority. Retrieved September 6, 2020 from <https://docs.unity3d.com/ScriptReference/AudioSource.html>
- [90] Unity - Scripting API: AudioSource. Retrieved September 6, 2020 from <https://docs.unity3d.com/ScriptReference/AudioSource.html>
- [91] Elevating Haptic Technology with Lofelt Wave. Retrieved September 6, 2020 from <https://assets.ctfassets.net/qku6b8nf1d79/76s0L6xb9GBUdSMb3828AH/64695fe11161bd683e73ca757997e139/Elevating-Haptic-Technology-with-Lofelt-Wave.pdf>
- [92] Dayton Audio 25mm Exciter. Retrieved September 6, 2020 from <https://www.amazon.com/Dayton-Audio-DAEX25FHE-4-Efficiency-Exciter/dp/B00PY66ZCC>
- [93] Kinter MA170+ 2-Channel Mini Amplifier. Retrieved September 6, 2020 from <https://www.amazon.com/KINTER-MA170-2-Channel-Amplifier-Treble/dp/B07C1Q1FPT>
- [94] Sabrent USB External Stereo Sound Adapter. Retrieved September 6, 2020 from <https://www.amazon.com/Sabrent-External-Adapter-Windows-AU-MMSA/dp/B00IRVQ0F8>
- [95] Unity - Manual: Prefabs. Retrieved September 13, 2020 from <https://docs.unity3d.com/Manual/Prefabs.html>