



HAL
open science

Experiences of Personal Sound Technologies

Stine S. Johansen, Peter Axel Nielsen, Kashmiri Stec, Jesper Kjeldskov

► **To cite this version:**

Stine S. Johansen, Peter Axel Nielsen, Kashmiri Stec, Jesper Kjeldskov. Experiences of Personal Sound Technologies. 18th IFIP Conference on Human-Computer Interaction (INTERACT), Aug 2021, Bari, Italy. pp.523-541, 10.1007/978-3-030-85616-8_30 . hal-04196879

HAL Id: hal-04196879

<https://inria.hal.science/hal-04196879>

Submitted on 5 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Experiences of Personal Sound Technologies

Stine S. Johansen¹[0000-0002-8451-5444], Peter Axel
Nielsen¹[0000-0002-0282-7445], Kashmiri Stec²[0000-0003-2393-1126], and Jesper
Kjeldskov¹[0000-0001-7990-4229]

¹Aalborg University, Aalborg, Denmark
stinesl@cs.aau.dk, pan@cs.aau.dk, jesper@cs.aau.dk

²Bang & Olufsen, Struer, Denmark
ksh@bang-olufsen.dk

Abstract. Listening to sound individually while in close proximity of other people is increasingly enabled by a range of technologies. One still in development is sound zone technology that aims to provide personal sound without headphones or other wearable speakers. User-oriented studies in the area of personal listening primarily emerge from the fields of acoustics and sound engineering but are gaining increasing interest within HCI research. In this paper, we present a study investigating the experience of personal sound in relation to different types of situations and personal sound technologies. Our findings show strategies for adjusting personal sound and social interaction, descriptions of sound quality in relation to sound and situation types, and insights into participants' experiences of awareness using personal sound technology. The paper contributes with a thematic characterisation of this type of technology, serving as a foundation for further studies. This furthermore initiates a discussion on personal sound technology and soundscape composition in how situation types affect which sounds to include or exclude, and when.

Keywords: Soundscape · Sound zone · Personal sound · Headphone.

1 Introduction

In this paper, we are concerned with the experience of personal sound technologies. A number of technologies have been developed to provide users with personal listening experiences. Some fit in or around the user's ears, some are worn on the user's neck, and some are stationary speaker systems attached to walls or ceilings. One type that is receiving increasing amounts of attention is sound zone technology [1]. Sound zone technology poses new opportunities for users to modify what they are listening to where they no longer have to wear headphones as a prevention from disturbing other people or when wanting to listen to sound privately [12]. Previous work has shown that wearing headphones has disadvantages such as decreasing the ease of social interaction and becoming uncomfortable when worn for an extended period of time [11]. Therefore, personal listening using sound zones presents an attractive alternative. In the vision for sound zones, multiple users inhabiting the same physical space can

listen to sounds individually without interference between zones, meaning that users in one sound zone are not disturbed by sound from other sound zones [14]. From a user’s perspective, this means that they can experience an individually controlled ‘soundscape’ within a shared space [12]. With soundscape, we refer to the entire acoustic environment as perceived by a person [13].

Previous research elaborates current usage patterns and points toward future wishes for personal sound technology [9]. The relation between an individual and the peripheral environment, including other people nearby, is usually the center of those discussions. Future technological development in this direction includes soundscape filtering strategies that are contextually adapted such as allowing certain kinds of interruptions and masking irrelevant information. An overlap can be seen between those suggested tools and sound zones; both are parts of a broader group that can be called personal sound technologies.

With this paper, we expand on what previous research has called the *phenomenon of personal sound curation* [9]; i.e., the modifications users can apply to their soundscape and a characterization of the resulting experience. We present findings from an investigation into how people use personal sound technologies in different situations, and how we can enable them to talk about the way the technology fits into their daily lives. While technical challenges still exist on the path towards physical spaces with multiple sound zones and limited interference between zones, other technologies enabling personal listening experiences exist and are increasing in variety. Examples include active noise cancellation (ANC) in headphones (such as Bose QC35 or the Sony WH-1000 series) and wearable speakers (such as Bose SoundWear Companion or Zulu Alpha). Currently, personal sound is achievable with these technologies, but at the expense of social interaction, listening in private, or avoiding disturbing other nearby persons. Sound zone technology poses a potential solution to this but with currently limited setups in labs, it is not yet possible to study how. Therefore, it is necessary to create a frame of reference for the experience of sound zone technology by other means to steer future research. We do this by providing a preliminary set of characteristics that describe that experience. Our findings furthermore show that these characteristics are experienced differently in different situations.

The outset of our study is three situations where participants are present in a physical space with personal and shared sound. While wanting a personal listening experience, the social aspects of the situation might change. This leads us to the research question: How are personal sound technologies experienced in various social situations with mixed personal and shared soundscapes?

In answering this, we make two main contributions. First, we provide insights into the experience of personal sound for different situations and sound types. Our findings show different strategies participants devise to deal with a complex soundscape, both for adjusting their own personal sound and for communicating with other participants. They also provide insights into qualities related to sound and situation types as well as participants’ experience of awareness while using personal sound technologies. Second, we provide a set of future directions for conducting user studies related to personal listening experiences.

2 Related Work

In order to clarify the foundation on which we can investigate and characterize the experience of personal sound technology, we outline current work on personal sound technologies, including sound zones. Our general interest is to put sound zone technology into the broader frame of personal sound technology. We also outline existing research on the relation between soundscape and perceived situation appropriateness as a basis for understanding elements of situations that affect the experience of particular sounds. With this previous research, we distinguish between activities and situations. Multiple activities can co-occur in one situation.

2.1 Personal Sound

Within the field of sound studies, Bull analysed the use of the Walkman [2] as an early example of a music player that provided a private listening experience in public places. This was enabled by its portability. In more recent years, noise cancellation has enabled users to control their soundscape to a higher degree. Previous research shows that in order to do so, users employ a number of different strategies, such as adapting the physical fit of the technology or turning up the volume of the personal sound to drown out environmental noise [9]. These findings are based on focus group interviews conducted by showing staged video footage of a person using a set of headphones. The focus group then discussed their own use of similar technologies. In a model derived from the results, Haas et al. describe aspects of personal soundscape curation which includes *intervening conditions*. With this, they refer to social protocols and hierarchies, e.g., the effect of wearing headphones on how they are perceived by other people. Other studies are concerned with users' isolation from the outside world by for example exploring warning systems for headphones to prevent traffic accidents [18].

2.2 Sound Zones

In some cases, users do not want to wear headphones, or they are inhibited from doing so, but personal sound might still be desired. One example of this is a family sharing a living room, where the acoustic environment can become chaotic in the event of everyone engaging with their own activity such as playing a video game and listening to the radio. The family members might want to listen to only their own sound but still be able to interact with each other. Another example is museum exhibitions, where a guest might want to listen to the sound from one exhibit without disturbing other guests at other exhibits. At the same time, they also sometimes want to interact with others at the museum. To solve this, speaker systems where sound is directed towards a predefined area have been developed. These systems aim for total individual control of particular sounds. A system known as a parametric array uses ultrasound waves which become audible through nonlinear media [15]. Such systems, however, perform poorly for low-frequency reproduction of sound, so other methods for directing

sound have been explored. One example is parabolic speakers, e.g., Focus Point FP6030 by Soundtube, where sound is played into a parabolic reflector aimed towards one or more target listeners. SpatialSound Wave by Fraunhofer IDMT is another related example where several sound objects can be positioned in a physical space to give a user the impression that they are in a wanted position in an acoustic scene.

While different methods for implementing sound zone systems have been proposed [8, 14], in conceptual terms, they aim for a similar result. The vision of sound zone technology, as it is currently framed, holds that users can position a number of sound zones in a physical space which allows them to share the space, listen to individual sounds and not wear headphones [1]. This suggests that might be expected in the future when the technology is commercialized, including a possibility for social interaction, low intrusiveness from other sound zones, and increased awareness of the peripheral environment. Awareness does not necessarily entail attention, meaning that sounds from the environment might not necessarily become intrusive. With sound zone systems, users can experience an acoustic environment as they would without headphones but still receive personal sound. In this way, the environment is made up of foreground and background sounds which users can shift their attention between. Different use cases for sound zone systems described in previous work [11] show differences in the way the resulting soundscape is composed. In some cases, users might want complete silence, such as when using noise cancelling headphones, and in other cases, they might want a private listening experience while also hearing the environment around them. In this way, sound zone technology can be viewed as varying forms of interventions into a soundscape [12].

2.3 Soundscape and Situations

We address the relation between personal sound and sound zones as outlined above with research on soundscape and, more specifically, how the experience relates to situations. We refer to the widely recognised definition of soundscape by Schafer as "...the auditory equivalent of a landscape" [16], while also acknowledging later clarifications of the definition that emphasize soundscape as something that is perceived by persons rather than only a physical phenomenon [4]. In other words, the understanding of soundscape behind the study presented here is the total acoustic environment as perceived by participants. Furthermore, Schafer describes a soundscape as composed of different types of sounds. Some reside in the background as 'keynote' sounds while others perform a 'signal' function and capture the attention of listeners. In much later research, Bull similarly described that sound does not possess the same quality as visual objects that can be attended to separate from each other [2]. As such, listeners can attend to particular sounds consciously, but their experience of individual sounds is shaped by all present sounds. Listening to music can add specific meaning to a situation [6], and this has fostered a development of tools to curate listening experiences. Previous work within HCI has investigated social practices emerging around digital music in homes [10]. Findings include a preference for digital

music rather than listening from CDs or records. This enabled new ways of listening to music such as curating shared playlists and discovering new music on web-based platforms.

In a previous study on the relation between soundscape and situations, participants listened to outdoor recordings and were asked to rate the appropriateness of four predefined activities in relation to the sound [13]. Results showed that participants associated particular soundscapes with particular activities. For example, listening to sounds of single vehicles, participants rated bicycling as an inappropriate activity. The authors conclude that a soundscape can afford a certain set of actions. They also found that the activity with which the listener is engaged affects which sounds in the soundscape are paid attention to. Another study, based on questionnaire surveys in urban squares, similarly showed that preferences for particular sounds were dependent on respondents' use of the physical space [17]. At the same time, the soundscape of a physical space influenced how the space was used. Findings further included a tendency that preference for or annoyance with a particular sound depended on the situation in which the sound emerged. An outdoor concert might be perceived positively and attract people whereas the same music played from a store would elicit annoyance. Another aspect of soundscape evaluation is participants' expectations of how a place sounds. Previous research uses Truax's concept 'soundscape competence' to detail how prior experience affects behaviour as a result of characteristics of the soundscape [5]. With soundscape competence, the authors refer to a relation between a sound and the meaning the listener attaches to it. When in a familiar place, persons expect certain sound sources, a certain level of control, particular behaviour from nearby persons, etc.

With these studies in mind, activity and situation are central elements to evaluating technology which modifies users' perceived soundscape. The situations defined for a study of personal sound technologies necessarily affect participants' perception of the appropriateness of specific sounds. In other words, even though the study presented here is focused on experience of soundscapes, we infer from previous work that participants will weigh characteristics of the technologies delivering parts of the soundscape according to the situation they engage with.

3 Mapping Personal Sound Technologies

Personal sound technology is a term which spans across multiple speaker and headphone types. Common for all of them is the fact that they allow for a user to listen to something individually. This distinguishes personal sound technology from personalized sound. Where personalized sound connotes that a sound itself can be customized to each user, personal sound connotes a user controlling one or more sounds.

There are differences, though, in how technologies provide personal sound as detailed in Section 2. Some are worn on the user's body while some are wall-hanging speaker systems. While headphones generally provide users with options

for cancelling or opening up to sounds from the physical environment, this might also soon be possible with speaker systems. Therefore, in order to investigate the experience of listening using personal sound technology in different situations, we first present a mapping of how these types of technologies relate to each other. This provides a systematic overview that we will later use as part of framing the experience of sound zone technology.

Fig. 1 shows an illustration of how different types of personal sound technology relates to the user’s peripheral acoustic environment. For each technology, the user receives a mix between reproduced sound and environment sounds. The left emphasizes the personal sound. The middle equally emphasizes personal sound and peripheral environment, but under the conditions of reproduction of the environment. Finally, the right also emphasises personal sound and peripheral environment, but with nothing covering the user’s ears. For each type, we provide an example of a type of headphone or speaker that accomplishes the particular mix between reproduced sound and environment sound.

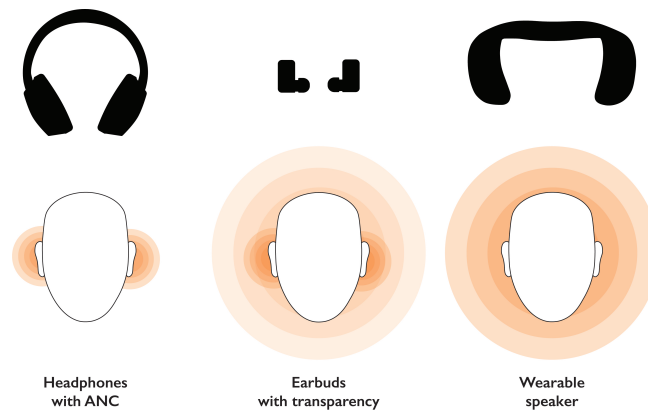


Fig. 1. Each type of technology relates the user to the peripheral acoustic environment differently. Active noise cancellation headphones: Only personal sound. Earbuds: Personal sound and environment sounds through microphone. Body-worn speaker: Personal sound and environment sounds at equal level.

With sound zone technology, the aim is to allow users to modify their soundscape to include or exclude particular sounds as well as limiting the area within which the included sounds can be heard. Different personal sound technologies share functionalities such as creating an individual listening experience, but as we illustrate in Fig. 1, they also result in different relations between the user and their environment. Sound zone technology could be constructed with similar different functionalities, motivated through different use situations [11]. Therefore, characterizing the resulting experience of using different personal technologies is useful as a reference for eventually characterizing the resulting experience(s) of sound zone technology.

3.1 Active Noise Cancellation

On-ear headphones or in-ear earbuds with Active Noise Cancellation (ANC) enable users to only listen to the sound played through the speakers. When the noise cancellation is activated, all sounds from the environment will, to some extent, be cancelled. Loud, high frequency sounds can still be heard in some cases. In relation to sound zones and personal listening, ANC headphones allow users to both isolate themselves from sounds played by other users and restrict their own chosen sound to be played within a private hearing range.

3.2 Transparency

Open earbuds allow users to listen to sound privately while still being able to hear sounds from the peripheral environment. In these cases, users are provided with a ‘transparency’ or ‘hear through’ mode which, instead of cancelling sounds, reproduces sounds from the environment. A microphone attached to the headphones records environmental sounds which are then played back to the user in real time. In relation to Fig. 1, they reproduce all sounds from the environment but, from the users’ perspective, might also reproduce environmental sound recorded and composed prior to listening.

3.3 Wearable Speakers

Wearable speakers are an alternative to headphones and earbuds that allow for personal sound to be reproduced in a larger area than the users’ ears. Users wear the speaker on their shoulders, and two speakers embedded on either side play sounds in an upwards direction. A waveguide is implemented between the speakers to enhance the sound through air pressure. These neck-worn speakers offer users a semi-private listening experience as other people in the same environment may be able to hear the sound from the speakers as well.

4 Method

Our research question to address in this study was: How are personal sound technologies experienced in various social situations with mixed personal and shared soundscapes? To investigate this, we had two particular aims: (1) gain insights into the use and experience of various personal sound technologies according to different situation types, and (2) find out how those technologies affect potential users’ experience of quality and what quality encompasses. This entailed considerations into how we could enable participants to talk about the experience. Answering the research question, we conducted a qualitative study with emphasis on participants’ experience of their soundscape in relation to different situations. In the study, we equipped participants with headphones or speakers to mimic the three relationships between user and environment described in Fig. 1. We also defined three situations in which participants increasingly need

to engage with each other to explore the experience with different demands to their attention. In the following section, we provide details about participants, chosen headphones and speakers, and procedure of the study.

4.1 Participants

As we are concerned with social aspects of personal listening experiences, we conducted the study with two participants at a time to also allow for a discussion on how those technologies affect the social interaction that could occur in users' daily lives. We recruited 12 participants in pairs who knew each other prior to the study. We did so to limit the amount of time necessary for participants to feel comfortable talking to each other. Table 1 presents an overview of participants' age, identified gender, and relationship with each other. We sought differences in age and types of relationships in order to be inclusive to potential differences as relationship type potentially affects the experience of a situation and how participants choose to interact with each other. Only one pair had a particular interest in sound technology (Pair 1). Both expressed interest in buying headphones and speakers of a high quality. For Pairs 3 and 4, all four participants were regular users of one of our chosen technologies. None of the participants were familiar with sound zone technology.

Table 1. Study participants.

Pair #	Participants	Relation
1	Male (28) and female (26)	Romantic partners (7 years)
2	Male (38) and male (54)	Romantic partners (6 years)
3	Female (30) and female (28)	Coworkers (2 years)
4	Male (31) and male (31)	Coworkers (3 years)
5	Female (26) and male (26)	Friends (2 years)
6	Male (29) and male (32)	Coworkers (2 years)

4.2 Technology

We chose three technologies that each unfold an instance visualized in Fig. 1 corresponding to sound zone characteristics. These have assumed differences on three parameters: Participants' ability to (1) interact with each other, (2) hear the other participants' sound, and (3) hear sounds from the environment. We also chose technologies with good reproduced sound quality based on high and consistent frequency and dynamic range. This resulted in using the Bose QuietComfort 35, Beoplay E8 2.0, and Bose SoundWear Companion Speaker.

4.3 Procedure

We conducted the study in a room with a reverberation time matching the average domestic living-room [7]. We did so to match the physical setting with

the situations presented to participants. The study was carried out in three parts. Each part investigated a particular type of technology in relation to three situations. For each part, the two participants wore the same set of headphones, earbuds, or wearable speakers. We counterbalanced the order of situations and technologies between each pair of participants. We defined the situations from two dimensions. First, they should have varying levels of necessary interaction between participants. Second, they should require varying levels of attention from participants. With this in mind, the situations were defined as follows: (a) Draw from a picture (2 minutes), (b) Play a specified card game (3 minutes), and (c) Plan a trip to a specified location (3 minutes).

Participants participated in the first situation individually. We asked them to redraw an image in front of them, thereby lessening the required attention compared to a situation where they were asked to draw from imagination. For the second situation, we asked them to play a card game they were already familiar with. This allowed us to investigate a situation where participants would interact with each other but they were not required to talk. For the third situation, we asked participants to plan a trip to a location they knew. This required them to both think about the elements of planning a trip as well as talk to each other. The inclusion of intense, social situations for all technologies was made to enable participants to discuss the different experiences between situations. The physical setup for each situation can be seen in Fig. 2.

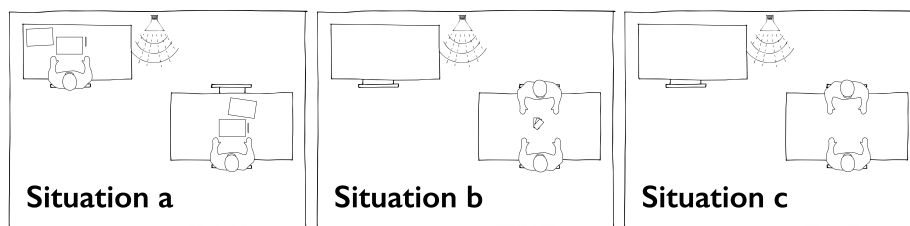


Fig. 2. The physical setup for each situation. Situation A was individual, and situations B and C were social. A speaker provided environment sound.

When participants had engaged with the specified situation, a follow-up, semi-structured interview was conducted. In this interview, we focused on clarifying different elements of the experience, specifically regarding (1) Sound quality, (2) Interacting with each other, (3) Attention to background sound, and (4) Disturbances, but with potential additional comments. We asked participants to reflect upon and describe their experience of each of these topics. The interviews took shape from participants' immediate responses to the initial topics.

Composition of Study Soundscape Present in the study was each participant's personal sound, a kitchen environment sound played from a speaker on a wall of the room, participants' voices, and sounds from when participants or the

facilitator moved. Together, this made up a complex soundscape, experienced by participants as a whole [2], which we then sought to dissect in the interviews. While participants wore the same type of technology in each part, they listened to different types of sound. One participant listened to a podcast featuring only speech by a male and a female voice, continued throughout the study. The other participant could choose a music playlist of three songs from a list of ten genres.

An issue with assessing the quality of an experience of a soundscape is that the assessment itself makes participants draw attention to sounds they would not necessarily have noticed consciously otherwise [4]. For that reason, we did not reveal the presence of sounds such as the background sound by asking about them directly. We counterbalanced the order of the technologies.

Data Analysis The interviews were transcribed and analysed thematically [3]. Based on the mapping in Section 3, we expected findings concerning themes of 'personal sound', 'social interaction', 'awareness', and 'sound quality'. In the first iteration, we coded the data according to these themes as well as emerging themes. This resulted in 43 unique codes that were hierarchically organised into 8 themes. Following a second iteration, we consolidated two themes, 'awareness' and 'intrusiveness', which contained similar analytical points. Two themes, 'error' and 'setup', had very few occurrences and did not offer insights into the research question of this paper and were therefore disregarded.

5 Findings

Table 2 shows an overview of our findings related to characteristics of each type of speaker emerging from the analysis.

Table 2. Overview of findings.

Characteristic	ANC headphones	Open earbuds	Wearable speaker
Personal sound	+ Enabling immersion	- Volume	+ Drowning out environment - Focus on personal sound
Social interaction	+ Hand gestures - Dialogue	+ Dialogue - Sound over participant	+ Dialogue with volume control
Sound quality	+ Reproduction	- Reproduction	- Disturbing sound types
Awareness	+ Personal sound - Environment	+ Environment	+ Other participant
Preference	+ Individual activities - Social activities	+ Limited social interaction	+ Awareness - Comfort

5.1 Personal Sound

Participants modified their personal sound for two purposes: (1) Drowning out environment sounds and (2) achieving higher immersion. We found that they

adapted the use particularly of the ANC headphones and the wearable speakers. For the earbuds, participants in five of the pairs expressed that they played the music or podcast at an undesirably low volume, but the technology did not allow them to increase it to a satisfactory level.

Modifying personal sound mainly had to do with participants' engagement with the individual situation (A). This occurred in instances where participants increased the volume of their personal sound for situation A, the individual drawing situation, but with the purpose of either supporting immersion or of drowning out other sounds. For the ANC headphones, a participant stated that he could increase the volume seeing as social interaction was not necessary. Another instance occurred for the wearable where the participant stated: "*When you [the other participant] moved over and started drawing, I turned up mine to the point where I could no longer hear yours.*" (M2, Pair 4).

The personal sound sometimes affected participants' engagement with a situation by supporting it. One participant described the ANC headphones as built for shutting out the world. He elaborates: "*I really think these allow you to immerse yourself in stuff, because anything else is completely shut out.*" (M1, Pair 6). In this case, the composition of the soundscape is described as very simple and that enables the participant to focus on what they are doing.

These findings regarding personal sound show that the personal sound or use of the technology gets modified according to what participants wish to attend to during a situation; that is, each within their own soundscape.

5.2 Social Interaction

Generally, participants sought to complete the descriptions of the situations, but they did so using different strategies and with different levels of communication according to both situation type and technology type. These findings concern (1) the ways participants chose to interact with each other given the individual soundscapes and the situations of the study, and (2) how use of the speaker technology was modified to ease this interaction.

For the earbuds, most participants described that they were able to speak to each other when engaged with situation C in which dialogue was necessary. However, some sought to use other strategies for communicating with each other for situation B where they interacted with each other, but dialogue was not required. One participant described her use of eye contact with the other participant as a way to negotiate the rules of the card game.

Similarly, for the ANC headphones, a participant stated for situation B that: "*You could just talk with your hands.*". He continues to say for situation A: "*You couldn't do that for the first situation.*". Not only did the participant choose to discontinue speaking with the other participant, he also chose to turn up the volume of his personal sound when situation B did not require any dialogue between participants.

As expected, the noise cancelling function of these headphones made dialogue difficult for participants. Elaborating on her experience, one participant described that, to her, this type of headphone would mainly be useful when

alone. Furthermore, several participants either lowered the volume of their sound or raised their voice to be able to speak to each other. In these instances, participants prioritised being able to have a dialogue with each other over listening to their personal sound.

Participants in three pairs described speaking to each other as easy with the wearable speakers. They utilized the possibility to do so for both situation B and C which stands in contrast to the earbuds where participants could speak to each other but primarily only did so when necessary in situation C. Deciding to speak meant that participants typically chose to focus on the social interaction with the other participant as opposed to listening to their own sound even when the situation did not require them to do so. One participant said the following in relation to the card game (situation B) and the wearable speaker: *“It became an entirely different kind of game, because we were actually able to discuss whether or not this and that was right... I actually experienced looking at you [other participant] more instead being completely closed off.”* (F2, Pair 3).

For easing social interaction, we found three instances where a participant turned off their sound or removed the ANC headphones when engaged with situation C in which participants jointly planned an imaginary trip. One participant in Pair 2 elaborated on this, saying that he did not see any benefit of listening to music while talking to the other participant. In other words, participants adapted their use of the technology to fit the social context of situation C when they could. In contrast, when they were wearing the earbuds and could not turn up the volume, they expressed a desire for higher volume of the personal sound, thereby placing lower priority on the social interaction.

In summary of our findings related to social interaction, participants would mainly choose to speak when possible; thereby prioritizing the other participant over sound. In some instances, participants developed other non-verbal strategies to successfully communicate with each other and prioritised the sound over the other participant.

5.3 Sound Quality

Sound quality was described by participants in terms of (1) sound reproduction and (2) type of sound. Only participants with a high interest in HiFi and speaker technology were affected in the way that the sound quality impacted their ability to focus on the sound. When listening to the podcast, one participant in Pair 1 said that the high sound quality of the ANC headphones helped him better focus on simultaneous activities in the situation; i.e. listening to the contents of the podcast and having a conversation with the other participant.

The way participants referred to their personal sound changed. Regarding the earbuds, one participant initially described her sound as a news podcast, stating that she was able to speak with the other participant. However, a few minutes later in the interview, the participant said: *“There shouldn’t be any noise when you have to talk to each other”* (F2, Pair 3). Another participant provided the following comment about environment sounds while wearing the earbuds: *“There was a lot of background noise with an almost metallic timbre. I mean,*

you can hear a sort of input when people are talking. But with a metallic timbre, and I almost think, for example for the sound of putting the pen on the table, that it became louder and more distorted in a way. And that really disturbed the music.” (F1, Pair 1).

Our findings show that participants do not only refer to sound quality in terms of personal sound reproduction. They also relate quality to sound type. Determining the quality of what they listen to is dependent on these two factors.

5.4 Awareness

Our findings show that awareness of sounds was described by participants in (1) positive, indifferent, and negative ways based on how the personal sound and environmental sounds affected them, and (2) how it related to the situations with which they were engaged. The descriptions referred to awareness of both personal sound and background sounds, and how it was dependent on the particular combination of sounds. Different examples of how one type of sound demanded more attention than the other were given. This happened whether the demanding sound was their own or the other participant’s.

Participants experienced that their level of tolerance changed according to type of situation. During situation A, they were aware of but not disturbed by the other participant’s sound, but it became demanding during social situations for the wearable speaker. Described by one participant: *“For this [individual situation], you are only using visuals and your motor functions, but you don’t have to hear anything, whereas if we are having a conversation about [trip destination] or [card game], our auditory sensory system is already divided.”* (M1, Pair 6).

In some instances, participants intentionally ignored sounds. One participant described a situation where she felt in control of her attention in relation to her own personal sound; in this case a podcast: *“...if it had been someone yelling, or if it had been a voice I recognised, then I would have been disturbed. Then I would have directed my attention... Because I could hear it. But actually it wasn’t interesting enough, so I just let it slide to the background.”* (F2, Pair 3). In this instance, the participant is aware of an unwanted sound, but it is not intrusive to the experience. One participant commented about the podcast: *“It was easy to follow what they were saying. Even without focusing on it while simultaneously doing activities that demanded attention.”* (M1, Pair 1). This instance also describes the personal sound as unintrusive, but not without the participant attending to it.

As expected, no participants noted any environment sounds for the ANC headphones. This was, unexpectedly, also the case for the wearable speakers. For the earbuds, some participants described the sounds from the peripheral environment as intrusive to their ability to listen to the personal sound. One participant explained: *“I didn’t really focus on the quality of the music, but rather on anything else going on in the room, because I actually think that was louder. And also the sound of my own breathing. Suddenly, I sounded like Darth Vader while wearing them.”* (F1, Pair 1). The other participant in Pair 1 continued: *“All the input that you’re making gets transferred back to you through the headset*

which then plays it even louder.” (M1, Pair 1). Only one participant noticed the background sound that was added to the environment through a speaker. The participant mentioned this in the interview of the third part of the study but added that as soon as she noticed the sound consciously, she was also able to remember the sound being played previously.

These examples show that personal sound and environment sounds can both demand attention or be ignored. In some instances, such as the one described by M1 in Pair 1, the personal sound can be perceived as an activity in parallel to situations A, B, and C, whereas in other instances the personal sound resides on the same level as environment sounds.

5.5 Preference

Preference for particular speaker technologies depended more on the situation than technical characteristics. We found that preference was either expressed in relation to (1) situations, (2) the type of personal sound participants listened to, or (3) the comfort of the technology. One participant prioritised his preference the following way: *“I would say the [brand] over-ears for drawing. That way you could concentrate on that. And for talking, that neck thing.”* (M1, Pair 5). This statement confirmed the expectations of how the technologies would perform in relation to individual versus social situations. Some participants even preferred noise cancellation over being able to talk to the other participant. In one instance, a participant experienced that she could achieve a satisfactory level of communication with the other participant using hand gestures, making her prioritise the comfortable wear of the ANC headphones over interaction with dialogue. Some also preferred the ANC headphones due to comfort, practicality of wear, and sound quality. Other participants’ preferences differed from that.

One participant described his least favourite technology in relation to the type of sound he was listening to (a podcast): *“It wasn’t that I felt distracted by sounds in the room, but because someone was talking, you could doubt whether it was someone from the outside or if it was something I heard in my ear. For that kind [podcast], I don’t think those [earbuds] are good.”* (M2, Pair 2). For other participants, the fact that the earbuds reproduced environmental sounds made them prefer these above other speaker technologies. One participant had experienced that, for the individual situation, the earbuds had enclosed her ears but for the social situations, she was able to focus on talking to the other participant.

The wearable speakers were preferred by some participants as a speaker. A number of participants described that they would use those as a regular speaker and not for personal sound. One participant, however, preferred the wearable speakers for the way they allow listeners to be open to the peripheral environment: *“I actually think I liked the neck thing best for most things. I think it was nice, because I also felt that those [ANC headphones] are really closed off. That makes the neck thing nice; the fact that it’s more open.”* (F1, Pair 5).

Our findings reveal that participants’ preference is mainly determined by the situation they were engaged with. For personal sound technologies, the social aspect of the situation is especially essential.

6 Discussion

The aim of this study has been to investigate the experience of personal sound technologies in various situations with mixed personal and shared soundscapes. The findings show that the three technologies provide users with very different experiences. These findings are summarised in Table 2. The juxtaposition in the table allows us to compare and contrast characteristics between technologies.

Some of the differences between technologies that participants experienced confirmed expectations from previous research. Our findings show a trade-off in the awareness of environmental sound and the quality of the personal sound. When users cannot remove environmental sound, they are subjected to different types of sound that can be perceived as disturbing. When isolated from the environment to a larger extent, the load on their awareness is minimised, but they also lack the ability to speak and carry out social activities in a situation. This confirms the mapping in Fig. 1.

6.1 Studying the Experience of Personal Sound

Our findings create a frame of reference for future work on personal sound technology. Achieving a personal listening experience can be done using a number of different technologies where sound zone technology is one potential option in the future. While these technologies are different, the experience can be studied from the same characteristics.

The vision of sound zone technology is to provide an experience of personal sound in which a user can choose which sounds should be present in their soundscape. Based on previous work describing what sound zone systems aim to offer users, cf. Section 2, and our findings, we can define five general characteristics of personal sound technology which are achieved in different ways by current products. First, sound zones offer users personal sound; i.e., a private listening experience [1, 8]. Second, sound zones allow social interaction [11]. Third, the quality of the reproduced sound should be comparable to that which is reproduced through headphones [16]. Fourth, they allow users to be aware of their peripheral environment since they do not require headphones [11]. And finally, preference for these characteristics could be dependent on the context of use [4].

Using the technologies in our study as a proxy for sound zones, we illustrate the potential relations between user, environment, and other people as a continuum in Fig. 3. On the continuum, three instances are marked. In simple terms, these represent the extent to which the user can hear the peripheral environment. In more complex terms, they can be described by unfolding general characteristics of personal sound technology. One aim of the study presented here was to make such a characterisation. Rather than asking how much of the sound from the outside environment to let in to a user's personal listening experience, we expand the discussion on which sounds to include or exclude, as well as when. This is achieved through the themes of our findings.

A question following the representation in Fig. 3 is: In which ways are the chosen technologies representative of sound zone technology? Participants in the

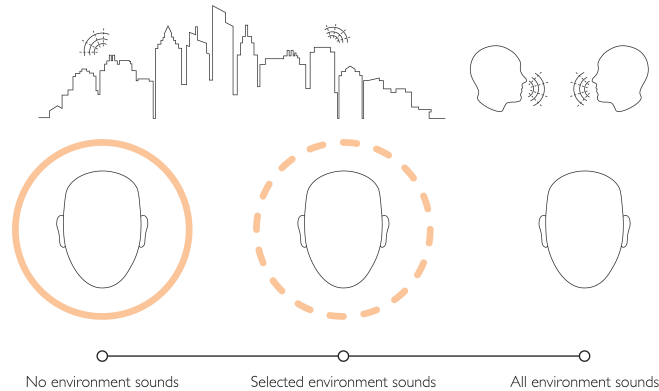


Fig. 3. The relation between user and environment, illustrated as a cityscape and dialogue, on a continuum from closed to open. Three instances are marked, representing no environment sounds, selected environment sounds, and all environment sounds.

study had to wear all three types of technologies which would not be the case for sound zone technology. However, each technology represents a different way sound zones could be envisioned and understood in terms of how closed off from the environment users would be. As such, our findings have provided a richer picture of how sound zones can be used in setups for personal listening.

The earbuds represent a place on the sound zone continuum where some sounds from the environment are selected, and some are not. In our study, the earbuds reproduced all environment sounds which could have had an effect since participants did not experience a choice between wanted and unwanted sounds from the environment. Being presented with that choice can lead users to become aware of sounds they would not otherwise have noticed as previously described in [16], cf. Section 2. Adding to that, our findings showed that except for one, participants were not aware of the background sound being played even though they were physically able to hear it. For future research, it would be interesting to provide users with a selection of background sounds to further study how their own choices on composing the resulting soundscape affects their awareness. Furthermore, valuable insights into how the background sound is delivered can be gained. For example, background sound played through headphones or earbuds might demand a different level of awareness compared to background sound played from a speaker located in a different position than the user.

6.2 Composing for Shared and Individual Soundscapes

Headphones modify users' perceived soundscape by introducing additional sound or removing sound from the environment. The three speaker types in our study enabled participants to modify the soundscape in different ways by either removing, reproducing, or allowing environment sounds. We observed how participants behaved during situations and discussed with them how the speakers had

affected their behaviour. Our findings show that participants' ability to modify their soundscape affected what they wanted to achieve in certain situations. This adds to previous research describing the relation between soundscape and sound zone technology as an intervention [12].

Haas et al. found and described that users sometimes want to modify their soundscape by removing intrusive sounds when engaged in a situation [9], cf. Section 2. In their study, users described noise cancelling headphones as feeling unnatural and that, in many cases, they did not want complete silence. Therefore, they often filled the silence by playing music or other sounds that would be non-intrusive. The participants in our study were listening to sound but had the ability to turn it off. As such, silence was not the point of departure for the participants, and they had to make a conscious decision to turn off the sound or lowering the volume. Some participants chose to do that, and with our findings we add to the study by Haas et al. by showing that personal sound has different purposes such as enabling immersion in a specific activity and drowning out sounds from the peripheral environment.

In our findings, some participants were able to consciously filter out unwanted sounds, and they explained that the situation and surroundings are factors when they determine what is intrusive and what is not. This confirms previous research on urban soundscapes that showed a relationship between prior experience and expectations for which sounds should be present in a specific place [4, 15]. Our findings add to this by demonstrating that, for personal sound technology, participants' filtering abilities are determined by the situations they are actively engaged in and are not limited to the physical space. This was also true for the cases where participants could not filter out unwanted sounds. The same sound could be referred to as music and noise. The situation affected participants' ability to filter out sound; in some cases, this was supported and in others, it was not. Participants might have had expectations to which sounds would be present in the room, resulting in a certain perception of the soundscape.

Our study included both individual and social elements to mimic a situation in which the level of social interaction changes. Compared to the study by Haas et al. [9], we show how these types of technologies fit into a social situation where usage is not only based on activities in which the user is alone or surrounded by strangers with whom they would not usually interact. In a domestic context, users can be surrounded by other people and, for that reason, do not necessarily want to be closed off completely to the peripheral environment. Haas et al. found that headphones signified to other people that the user does not want to be approached. Our findings showed this was different for wearable speaker technology. For example, one participant looked at the other participant more when the technology allowed her to hear the participant and the environment more. While some participants lowered their volume to better focus on speaking to the other participant, no one intended for their own sound to be intrusive to the other participant. As expected, for the individual situation, participants opted to turn up the volume of their own sound or tune out the sound.

Furthermore, the tolerance participants have for certain sounds changes according to situations. Participants defined some situations to be common experiences meant to be shared, and this resulted in a lower tolerance of sounds that make it difficult to interact verbally with each other, such as dialogue. The level of tolerance was also described in relation to how the sounds fit together. With this insight, we suggest that how closed a personal sound technology is to the environment should be determined by the type of sound listened to by individual users. This adds to previous research on appropriateness of activities [13] with the element of personal sound and importance of whether or not the situation was social. Our findings suggest that activity appropriateness is not only determined by the soundscape but also by engagement with other people.

The setup did not always represent a real situation. This was an expected comment due to the characteristics of the technologies which make it more or less inviting to speak to other people. While this conflict enabled participants to talk about what a real situation might look like, a direction for future work is to conduct similar studies in environments with which participants are familiar, such as their own homes or work places. This will allow for other situations that relate to a specific context for using the technology. It will also provide insights into the effect of expected and unexpected sounds on the experience of using personal sound technology.

7 Conclusion

In this paper, we have presented a qualitative study investigating experiences of personal sound technologies to make a frame of reference for future studies. We chose three technologies that relate the user to the environment and other people differently by either being closed or open. The result is a thematic characterisation of these speaker technologies based on users' experience of personal sound, social interaction, sound quality, awareness of peripheral environment, and preference determined from situation type.

As more variations of headphones and speakers for personal sound are developed, research on the effect on social interaction and how that can be supported by these technologies becomes increasingly important. The findings presented here suggest further work in this area. First, the study can be elaborated by using different types of sounds. Participants in our study described differences in how the sound affected their attention and ability to tune out the sound. Second, further research could limit the number of variables to assert causality. Third, the situation that participants engage in can be more closely related to real life with a setup that takes place in their own homes and during their daily routines. Fourth, different technologies that relate users differently to their environment can be investigated in order to describe other instances on the continuum. Finally, the characteristics emerging from this study can be quantified in more controlled lab settings.

References

1. Betlehem, T., Zhang, W., Poletti, M. A., Abhayapala, T. D.: Personal Sound Zones: Delivering Interface-Free Audio to Multiple Listeners. *IEEE Signal Processing Magazine* **32**(2), 81–91 (2015)
2. Bull, M.: *Sounding out the city: Personal stereos and the management of everyday life*. Berg, Oxford, United Kingdom (2000)
3. Clarke, V., Braun, V.: Thematic Analysis. In: *Encyclopedia of Critical Psychology*, pp. 1947–1952. Springer, New York, NY (2014).
4. Brown, A. L., Kang, J., Gjestland, T.: Towards standardization in soundscape preference assessment. *Applied Acoustics* **72**(6), 387–392 (2011)
5. Bruce, N. S., Davies, W. J.: The effects of expectation on the perception of soundscapes. *Applied Acoustics* **85**, 1–11 (2014)
6. DeNora, T.: *Music in Everyday Life*. Cambridge University Press, Cambridge, United Kingdom (2000)
7. Díaz, C., Pedrero, A.: The reverberation time of furnished rooms in dwellings. *Applied Acoustics* **66**(8), 945–956 (2005)
8. Druyvesteyn, W. F., Garas, J.: Personal Sound. *Journal of Audio Engineering Society* **45**(9), 685–701 (1997)
9. Haas, G., Stemasov, E., Rukzio, E.: Can't you hear me? Investigating Personal Soundscape Curation. In: *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia (MUM'18)*, pp. 59–69. ACM Press, New York, NY (2018). <https://doi.org/10.1145/3282894.3282897>
10. Leong, T. W., Wright, P.: Revisiting Social Practices Surrounding Music. In: *Proceedings of the Conference on Human Factors (CHI'13)*, pp. 951–960. ACM Press, New York, NY (2013). <https://doi.org/10.1145/2470654.2466122>
11. Lundgaard, S. S., Nielsen, P. A.: Personalised Soundscapes in Homes. In: *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS'19)*, pp. 813–822. ACM Press, New York, NY (2019). <https://doi.org/10.1145/3322276.3322364>
12. Lundgaard, S. S., Nielsen, P. A., Kjeldskov, J.: Designing for Domestic Sound Zone Interaction. *Personal and Ubiquitous Computing*, 1–12 (2020)
13. Nielbo, F. L., Steele, D., Guastavino, C.: Investigating soundscape affordances through activity appropriateness. In: *Proceedings of Meetings on Acoustics ICA2013* **19**(1), 8 pages (2013). <https://doi.org/10.1121/1.4800502>
14. Nielsen, J. K., Lee, T., Jensen, J. R., Christensen, M. G.: Sound Zones as Optimal Filtering Problem. In: *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pp. 1075–1079. IEEE, New York, NY (2018)
15. Pompei, J. F.: The use of airborne ultrasonics for generating audible sound beams. *Audio Engineering Society Convention* **47**(9), 726–731 (1998)
16. Schafer, R. M.: *The Soundscape: Our Sonic Environment and the Tuning of the World*. Alfred A. Knopf, New York, NY (1994)
17. Yang, W., Kang, J.: Soundscape and Sound Preferences in Urban Squares: A Case Study in Sheffield. *Journal of Urban Design* **10**(1), 61–80 (2016)
18. Yoon, S., Lim, Y., Lim, H., Kim, H.: Architecture of Automatic Warning System on Urgent Traffic Situation for Headphone Users. *International Journal of Multimedia and Ubiquitous Engineering* **7**(2), 421–426 (2012)