

# LISTENER PERCEPTION OF SPATIALISED AUDIO FOR EMBODIED INTERACTION IN SONIFICATION

*Andrew Beadling and Paul Vickers*

Department of Computer and Information Sciences  
Northumbria University  
Newcastle-upon-Tyne, UK

{andrew2.beadling, paul.vickers}@northumbria.ac.uk

## ABSTRACT

One underexplored method for examining sonification perception is through spatialised audio parameter mapping. To lower the entry threshold for listeners unfamiliar with sonification, binaural rendering of Ambisonic audio is employed as a spatialisation technique. A Pure Data patch was developed and a listener study was conducted to investigate participants' experience of listening to the spatialised sonification of city air quality data. The results indicate there is a link between listener experience and quality of perception, indicating that future research should focus on developing listener skill. Spatial audio appears to have some influence on perception for musically-trained listeners, but this appears to be most effective as an extra factor supplementing the manipulation of other sonic parameters such as frequency. Binaural rendering was useful for introducing uninitiated listeners to sonification, but there was difficulty in replicating the experience of physical space.

## 1. INTRODUCTION

In considering debates surrounding user engagement there remains a question as to how to regard sonification as an embodied experience [1, 2, 3, 4, 5]. Our everyday experience is of sound in a 360° three-dimensional space. Perhaps for pragmatic reasons, most sonification research has relied on stereo audio. Stereo is easy to generate and to playback on commonly available equipment. The three-dimensional spatialisation of sonification has tended to require specialist (and expensive) loudspeaker arrays, or HRTF-based solutions for headphones.

Ambisonics is a full-sphere surround sound audio format developed in the 1970s by Michael Gerzon [6] at the UK's National Research Development Corporation but met with limited adoption. Now that key patents associated with the technology have expired and the processing power required to encode and decode Ambisonic signals in real time is easily managed by current personal computers, it has become much easier to generate convincing three-dimensional spatialised sonifications. Furthermore, the flexibility offered by Ambisonics of decoupling the encoding process from decoding and reproduction means that a sonification can be spatialised in a full three-dimensional manner and then decoded for playback on devices with different spatialisation capabilities with relative ease by the interested researcher.



This work is licensed under Creative Commons Attribution – Non Commercial 4.0 International License. The full terms of the License are available at <http://creativecommons.org/licenses/by-nc/4.0/>

Examining spatialisation as a sonic parameter for adding potential extra layers of meaning, this study explores how spatial dynamics could be utilised to inform the experience of interpretation by assessing listener perception through listener experiments. Boehringer *et al.* [7] report ongoing studies involving trained listeners including sonification researchers, sound artists, and musicians with an investment in the field using a high-end specialist loudspeaker. However, the technical requirements of producing Ambisonic material (including installation of Ambisonic encoder, decoder, and effects plugins) with such equipment still presents a high entry barrier for engagement with general users. The binaural rendering of Ambisonic audio presents a much more accessible route for exploring spatialised sonification with general users (including listening remotely online).

This paper reports on an initial exploration of spatial sonification via the binaural rendering of Ambisonic material. The work involves the design and development of an accompanying Pure Data (Pd) patch that can be used to sonify data and prepare it for spatial rendering in a digital audio workstation. The patch has been developed to present a simplified method for researchers who wish to explore sonification as a tool for data exploration. In this study the patch was used to sonify weather data.

## 2. SONIFICATION CONCERNS

Despite a generally agreed upon recognition of sonification's uses and fundamental qualities, it has been suggested that there is no single coherent theory of sonification [8, p. 5], owing to its status in relation to other sonic media. Whilst sonification has been regarded as a purely information driven tool, focusing on empiricism, others consider sonification as an essentially dialectical exchange between scientific method and aesthetics. In describing sonification functions, Walker and Nees [9, p. 12] extended their previous view of sonification to include art, entertainment, sports, and exercise as potential functions for use. Vickers and Hogg asserted that the adoption of an aesthetic approach to sonification can help to facilitate ease of listening, which in turn would 'increase the communicative and expressive capability of the auditory displays' [10, p. 210], a view held by others regarding to the use of elements such as timbre and spatialisation in parameter mapping [11, 12].

Parameter mapping sonification is a common technique for perceiving changes in a dataset. This allows users to explore changes in a dataset by assigning data variables to individual sonic parameters and using the 'change in acoustic dimension' [9, p. 16] to demonstrate movement between two points.

Positioning sonification as an interaction between sound and listener, it has been argued that sensory perception should be given greater consideration, with Vickers *et al.* noting that ‘our perception of sound is as much determined by our sedimented knowledge of sound production as by the sense of hearing, we cannot bypass all that innate knowledge we have about sound in the real world’ [4, p. 98]. Sonification does not exist in a vacuum: listeners apply context that is culturally informed. This view aligns with Harrison *et al.* [13] in their assessment of HCI, as they suggest that a ‘phenomenologically situated paradigm’ is a lesser reified but nonetheless integral area to be considered in HCI.

## 2.1. Spatialisation

If embodiment is regarded as the ‘common way in which we encounter physical and social reality in the everyday world ... that we encounter directly rather than abstractly’ [14, p. 100], then the concept serves as a useful framing for exploring sonification and spatial audio. Using spatial parameters for mapping allows researchers to situate listeners in a shared perceptual space, ‘the space in acousmatic music within which the perceptions of composers, scientists and audience intersect’ [15, p. 3]. Whilst parameter mapping sonification has tended to be concerned with mapping variables to factors such as frequency or amplitude, spatialisation creates potential for added dynamic layering in the exploration of sonified data. Primarily seen as a tool for facilitating exploration of other sonic parameters by making it ‘easier to decode multiple simultaneous sounds’ [16, p. 192], spatialised sonification has been utilised to foreground or reduce audio evoking the Cocktail Party Effect [17], which directs their perception of sound in space through filtering to the most relevant sound. The use of spatial parameters is therefore regarded as an extra dynamic, with advances in technology allowing sonification to act as ‘virtual auditory environments’, with spatial audio being used as a form ‘spatial cueing’ [18, p. 73], foregrounding other elements of sonification to draw a listener’s attention to a particular attribute of a piece; the extent to which spatial dynamics can be regarded as a reliable primary method for inference has been less explored. Research presented at ICAD has explored this through the conducting of listening experiments with experts with experience in sonification across various fields including computer sciences, music, and arts practice [7] paying attention to perception and interpretation. This has provided a point of entry for discussion of sonification and perception with regards to space; however, the focus on expert analysis therefore centralises responses from those with prior investment in the field, leaving room for discussion of perspective of untrained listeners and general users.

## 2.2. Ambisonics

One issue for exploration of sonification using Ambisonics has been the high entry threshold for users. The development of virtual environments has provided an alternative mode for exploring Ambisonic audio rendered binaurally, allowing users to replicate three-dimensional sound through a listener’s headphones. The difficulty facing spatial sonification with this technology is the removal of the listener from the physical context of a real-world environment. Dourish emphasises how even in immersive virtual reality environments, users are still disconnected from worlds they do not inhabit [14, p. 102]. From this perspective, whilst the addition of spatial dynamics can potentially enhance or change a user’s understanding, binaural sonification lacks verisimilitude by

removing listeners from the context in which they normally intersect with spatial sound. The use of head trackers offers the ability to maintain sound sources in a fixed external position regardless of the listener’s head position, but this is beyond the scope of the current study.

It has been argued that perception is ‘depend[ent] on a complex and dynamic interplay between the sound attributes, the environment, and the listener’ [19, p. 154]. Reliable perceptions of information relayed through sonification is dependent on contextual factors that lie beyond the sonification itself. It is assumed that understanding is dependent on how experienced a listener is in various contextual aspects relating to both sonification itself and the data being sonified. Consequently, a musician’s ear will be more attuned to the subtle nuances that may occur over the course of sonification due to a likely familiarity with shared elements relevant to both musical study and auditory display such as changes in pitch, amplitude, timbre, spatial configuration tempo or rhythm. This has implications for sonification’s wider usefulness: without context, a sound may be loaded with differing meaning based on a listener’s prior engagement which may result in poor communication of data or interpretation outside of that intended by the designer. This split between trained and untrained listeners draws distinctions with sociological perspectives including Bourdieu’s writing on cultural capital, which views the siphoning of aesthetic objects as a distinctly education- or class-based phenomenon [20, p. 71].

This raises questions about the usefulness of sonification for untrained listeners, with perception and listener experience therefore presenting a significant problem in sonification’s wider employment. It may be reasonable to advocate for simplicity in sonification, however, Vickers has noted that rather than focusing on ease, greater consideration should be given to developing trained listeners, observing that ‘it has long been accepted in other fields that sound-based exploratory tools require skill to use well’ [5, p. 740]. This view suggests that complex mappings for sonification may be of use in developing the field by moving towards more ambitious methods of mapping, including spatial arrangement.

## 3. AMBISONIC SETUP

To enable general users to listen to sonifications spatialised using Ambisonics, a Pd Vanilla patch was developed. The patch’s audio was routed into REAPER v.6.67 for Ambisonic spatialisation. The Pd source files and REAPER project files can be found at the project’s online repository [21]. REAPER has been a favoured environment for previous research using spatialised audio, and as a result there are several demo files for demonstrating how to use IEM’s Ambisonic plugins [22]. The audio was encoded into Ambisonic B-format using Kronlachner’s *ambix* and *mxfc\_convolver* plugins for placing and panning the sounds in three-dimensional space [23]. The binaural decoder plugin was used to decode the Ambisonic signal into a binaural sound file suitable for playback over headphones. Internal audio connections between Pd and REAPER on the host macOS system were made with the BlackHole audio loopback driver [24].

### 3.1. The Pd Patch

A Pd patch (Fig. 2) was developed to enable the sonification of data using a series of synthesisers and output channels with a view to producing sonified datasets which can be transferred to a digital

audio workstation for Ambisonic spatialisation. The patch contains a default setup of an example used in this sonification study, using open-source weather data from Newcastle University's Urban Observatory [25]. Data is provided to the patch via a tab-delimited text file (e.g., see Fig. 1).

SensorName	Date	Timestamp	Humidity	Pressure	Temperature
PER_AIRMON_MESH1965150	19/07/2022	05:00:00	60.2	1007.6	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:01:00	60.2	1007.4	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:02:00	60.2	1007.4	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:03:00	60.2	1008	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:04:00	60.2	1007.8	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:05:00	60.3	1007.3	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:06:00	60.2	1007.9	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:07:00	60.2	1007.2	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:08:00	60.2	1007.4	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:09:00	60.2	1008	22.2
PER_AIRMON_MESH1965150	19/07/2022	05:10:00	60.2	1006.9	22.3
PER_AIRMON_MESH1965150	19/07/2022	05:11:00	60.3	1007.5	22.3
PER_AIRMON_MESH1965150	19/07/2022	05:12:00	60.3	1007.7	22.3
PER_AIRMON_MESH1965150	19/07/2022	05:13:00	60.3	1007.1	22.3
PER_AIRMON_MESH1965150	19/07/2022	05:14:00	60.2	1007.4	22.3
PER_AIRMON_MESH1965150	19/07/2022	05:15:00	60.2	1007.2	22.3
PER_AIRMON_MESH1965150	19/07/2022	05:16:00	60.2	1007.3	22.4

Figure 1: Weather Data in a tab-delimited text file

Data variables are linked to channel objects which contain the audio synthesis objects that sonify the data. In the weather data example three variables were mapped to seven channels: channels 1–4 sonify temperature, and channels 5–6 sonify atmospheric pressure, each channel adding and removing different harmonic elements to coincide with data going above or below thresholds. Modulation is controlled by the movement of variables, and is linked to a number of sonic parameters including pitch and amplitude and, in the case of channels 5–6, includes modulation of a bandpass filter. The addition of multiple audio objects for each variable has been included in the example to contribute to a sense of overcrowding or dissonance, which is inherently linking to perceptual associations with weather data. Channel 7 sonifies humidity data using percussive audio simulating a cymbal; this has been done to demonstrate the varied capabilities for sonified audio synthesis and to provide an example of an object that sits at a higher frequency point on the auditory spectrum that can be clearly differentiated from other variables. Changes in the sonic parameters of this object are controlled by movement in variable 3, modulating the decay length and filter cut-off of the audio output.

The `amplitudeModulation` object contains the default routing of variables rescale settings to the correct channel. Amplitude modulation for each channel is determined by the changes in the variable that it is linked to a value in the range 0.3–1.0. When the variable level is raised or lowered, the amplitude follows this.

The `catchVariables` object catches audio output from each audio channel and routes it to the audio mixer and VU meter GUI objects to show peak levels. It is also used to route all audio to one place for ease of digital to audio conversion.

The `notes` object contains a `select` function that receives user messages from the key selector object in the main interface. When a user selects a key, note values are selected from the corresponding key objects which are linked to MIDI note values. These notes are then sent to the channel objects to determine how variables are mapped to relative note values. In the instance of added harmonies, an addition function adds intervals of a major third and a perfect fifth where applicable.

The `outputConfig` object contains the routing necessary to output audio in either stereo or multichannel format. It contains instructions detailing how to switch between the two; the default setting is stereo output.

The `rescaleSettings` object sets the default rescale set-

tings for each variable, meaning that data parameters can be determined to match the maximum and minimum values for both the data variable, and the output. This is useful when rescaling for audio output as it allows users to position the variables within a scale relevant to human audio perception. In the example, all parameters have been set to load as default, but these can be changed by changing the message box values.

The `setTempo` object is used to set the rate at which the patch cycles through text file rows. In the weather data example tempo has been set to modulate along with the temperature variable, meaning that the rate at which Pure data sends bang messages, used to control the triggering of sounds, increases and decreases along with the changing data in variable 1. The object also contains a device to create a variable decay length based on the current tempo. This has been done to achieve a more natural sounding decay time as the tempo changes and reduce overlapping of sounds, with longer decay time as the tempo reduces and shorter decay as it increases.

### 3.2. Dataset

The study used weather data from the Urban Observatory [25] which holds the largest set of publicly available urban sensor data in the UK, gathering data in the city of Newcastle-Upon-Tyne and its surrounding areas in real-time through sensors located across the city. The chosen data related to air quality data on 19th July 2022 05:00–23:59, the day of the highest temperature on record in the UK [26]. The excerpt used in the study covered the period from 14.00–16.00, during which daily levels peaked.

The sonification represented three air quality variables: carbon monoxide (CO<sub>2</sub>), nitrogen dioxide (NO<sub>2</sub>), and trioxxygen (O<sub>3</sub>). Whilst the patch can process data for variables specified by the user, the choice to limit use to three variables was made so as not to overwhelm uninitiated listeners with information. A different selection of variables from the same dataset was used for the interview participants. With both sets of data used in the study it should be acknowledged that there is connotative significance in using data relating to any grouped dataset. In the case of this study listener perception is inextricably linked to assumptions that may be attached to discussions of record high temperatures.

### 3.3. Sonification Mapping

The incoming data were scaled and quantised to the integer range 1..16 thereby mapping each data point to one of 16 possible frequencies spanning two octaves. Variables were separated by pitch shifting to avoid significant overlapping of frequencies to maximise participants' chance of discriminating between variables. The rescaling decision was taken with the aesthetic sensibilities of the piece in mind. Rescaling allows disparate ranges to be transposed to an equal number of segments such that all are comfortably within a normal hearing range. In the study the frequencies were mapped to a C minor scale over two octaves, but a user could choose any other key.

Custom Pd objects were built for the sound synthesis, designing sounds to replicate the sound of synthesised string instruments. Multiple phasor objects create polyphonic synthesiser sounds, and the use of long attack and release parameters produce an envelope that can be likened to more natural sounding timbre. An amplitude modulator changes the level in proportion to the movement of data. Finally, the metronome triggering the sequential listing of

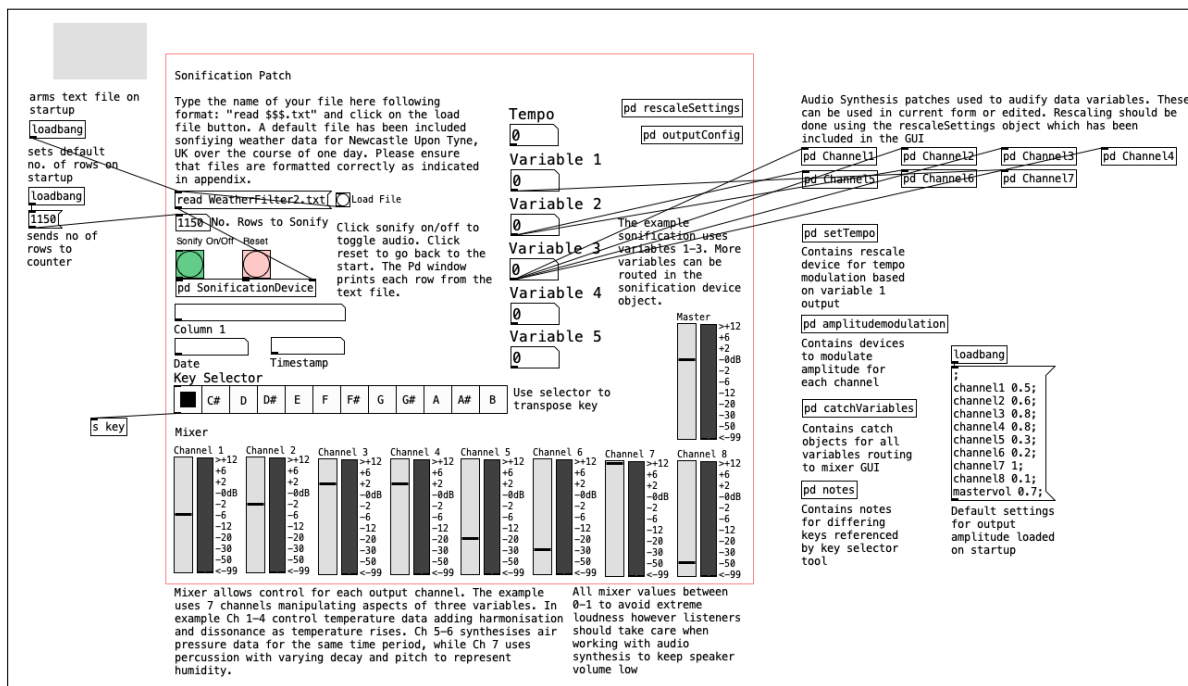


Figure 2: Pure Data Sonification Patch GUI

the data is controlled by a modulator that increases and decreases the tempo based on a fourth variable, particle count. The three variables used in the study are mapped to different iterations of the synth string object across three different pitch ranges to differentiate them from one another. CO<sub>2</sub> is mapped to the mid-range from C4 to D6 (261.63–1174.66 Hz); O<sub>3</sub> is mapped to the lower range of C2 to D4 (65.41–293.66 Hz); and NO<sub>2</sub> is mapped in the upper range of C6 to D8 (1046.50–4698.64 Hz). The sonification polarity was positive.

Rescaling is a deliberate manipulation of the original dataset as it transposing frequencies to a particular key signature as both involve rounding data points to the nearest relevant frequency. From a design perspective, this focuses a heuristic process centred on usefulness [27], rather than an empirically-informed direct mapping. As sonification presents a dialectical relationship between aesthetics and empiricism there is value in presenting rescaled data: though a sonified dataset may not map data directly to a corresponding frequency, the temporal shift between points still works to provide an approximation that details how data is changing.

### 3.4. Ambisonic Rendering

Kronlachner’s ambiX plugins [23] were used to sculpt the spatial arrangement of each track, with the plugins replicating movement in three-dimensional space. Spatialisation can be achieved either manually by setting user-generated automation points over the course of a track, or through automation controlled by MIDI or audio waveform following, the latter being used in this instance. Spatialisation is applied to each track through the ambiX encoder plugin using sidechain automation to alter the states of the FX parameters. Sidechain automation functions by using an audio source to trigger a change in the processing of a specified FX pa-

rameter when passing a user-defined threshold. This was applied to the elevation and azimuth parameters of the ambiX plugin for each track. The range, width, height, and direction of motion differ between variables to provide difference.

The CO<sub>2</sub> parameter is located at the front centre of the sphere and was automated with a clockwise motion which swells towards the outer width of the sphere as the CO<sub>2</sub> levels increase. The speed of motion increases with higher levels to give a propeller-like effect. O<sub>2</sub> levels sit in the lower front half of the sphere, alternating between the left and right side. The CO<sub>2</sub> spatialisation uses an attack of 197 ms and release of 300 ms to present a smooth transition of movement. O<sub>2</sub> has a quicker attack of 130 ms, creating a jittering effect which becomes more erratic as levels increase. The lower height positioning of the O<sub>2</sub> has been chosen in line with the frequency; O<sub>2</sub> was mapped to the lowest frequency in the chosen range. NO<sub>2</sub> is similarly applied with a jitter effect which is set to travel around the entire face of the sphere, with azimuth panning at a greater rate depending on NO<sub>2</sub> levels.

There are challenges to spatial parameters using Ambisonic plugins at the mixing stage. Here, the spatial mapping occurs outside of Pd, meaning that sonified data has already been rendered as audio and therefore spatialisation is applied to it post hoc. The benefit of sidechain automation over manual changes is that it allows data variables to have an influence on spatial parameters even though changes are applied outside the initial sonification process.

## 4. LISTENING TEST

To investigate how the spatialisation of variables influenced the way people listened to them a pilot listening study was conducted.

#### 4.1. Study Design

Nineteen participants were recruited via social media channels to listen to an excerpt of data that was sonified with the Pd patch, spatialised in REAPER with the Ambisonic encoding, and rendered binaurally for playback over headphones. They were asked questions about the sonification mapping techniques used: amplitude, tempo, pitch, and spatial arrangement. Participants were asked to categorise their previous experience of engaging with sonification to ascertain how individual contexts inform the usefulness of sonification, as well as their understanding of musical theory. Listener skill is a crucial consideration [5]: experiments in sonification are often evaluated based on ease of understanding, with an assumed lack of expertise on the part of the listener.

Three participants agreed to take part in a one-to-one interview to gain further insight into the responses to the questionnaires gathered in the listening test. Based on ideas proposed by Vickers [5], interviews were used to assess the impact of spatial attributes including direction, foregrounding and movement.

Because the listening test was conducted online optimal conditions for listening could not be guaranteed. The experiment instructions stipulated the wearing of headphones but it is impossible to know if participants followed this guidance. Similarly, variation in headphone quality may also have implications for a listener's ability to adequately perceive changes in spatial arrangement. Design and implementation were performed using Audio-Technica ATH-M50x closed-back dynamic monitor headphones, specifically designed for audio engineering. It is assumed that not all headphones used by participants met the same specifications, and as a result this may affect the quality of audio playback.

#### 4.2. Results

Participants categorised their prior experience of sonification as one of 'I have no prior engagement with sonification' (nine participants), 'I am aware of sonification as a concept but have not engaged with it before' (seven participants, or 'I am aware of sonification as a concept and have listened to sonified data before' (three participants). They were also asked to rate their understanding of concepts relating to music theory to establish whether a listener could be described as having 'trained ears'. Twelve of the nineteen participants had little or no understanding with the remaining seven having at least some understanding (Table 1). There was no correlation between those participants who reported having previous engagement with sonification and their musical background. The overall variance in experience levels gives a broad range of backgrounds for understanding sonification: ten varying combinations of experience level have been cited, with the most common response being 'I have no previous engagement with sonification' and 'very little or no understanding of music theory'. Most participants, therefore, could be regarded as untrained, i.e., without previous experience of sonification or music.

#### 4.3. Responses to the sonification

Participants indicated how they perceived the sounds changing, establishing whether they could identify if and how the musical parameters changed over the course of the piece, and whether distinctions could be drawn between response and experience. The most frequently cited word used describing change was "pitch". Twelve of the nineteen participants observed some change in fre-

Table 1: Listener understanding of music theory and previous engagement with sonification

Music Theory Knowledge	Score (n=19)
No understanding	5
Little understanding	7
Some understanding	2
Strong understanding	4
Very strong understanding	1
Prior engagement with sonification	
No prior engagement	9
Aware but no prior engagement	7
Aware and have listened to sonifications before	3

quency over the course of the piece, making it the most influential factor for listener recognition of auditory change (see Table 2).

By contrast, a third of listeners noted changes to the spatial dynamics of the sounds making it the least noticed parameter. Among those who identified spatial parameters, some noted changes relating to motion, citing "clockwise" movement and "smooth undulating" throughout the piece, whilst others referred to sounds moving binaurally between headphones, and concentration from a wide panorama to narrow concentration as variable values increased. These responses affirm Walker and Nees's suggestion that listeners generally agree upon the strengths of different sonification parameters for differing data types [9, p. 24]).

Regarding how changes in sound related to changes in the data, seventeen participants identified that changes in sonic parameters could be linked to increased levels of variables over time, while only one felt that the sonification did not discernibly reflect any changes, with one respondent suggesting the piece "did not reflect any peaks or troughs". Six out of the seven musically-trained listeners and eleven of the twelve untrained participants alike felt that at least one parameter change could be applied to changes in the variables, suggesting that parameter mapping sonification is useful for a large proportion of listeners irrespective of previous sonification or musical experience.

Examining the written responses, thirteen participants applied their understanding of the data using context directly, positing how the changes in musical parameter could be reflected in the air quality due to real-world factors. The most cited interpretation attributed changes over time to harmful chemical compounds in the air increasing due to increasing temperature. Traffic congestion was also suggested as a factor, whilst it was also posited that the increase in pitch and amplitude to uncomfortable listening levels could be interpreted as indicating dangerously high levels as the data peaked. This appears to confirm that listeners build their interpretation of auditory data through their preconceptions of sound based on real-world schemata. As the data is related to air quality, and specifically the levels of chemical compounds in a city centre, listeners applied the sonic dissonance from the sonification to negative connotations relating to harmful chemicals in the environment, suggesting that the context supplementing the sonification may implant a new sense of 'urgency' [9, p. 13].

#### 4.4. Parameter importance

Participants were asked to rank from 1 to 4 the importance they attached to each parameter in understanding changes in the dataset. The parameter considered to be overwhelmingly the most important in understanding data movement is pitch, which was rated first by twelve participants, followed by tempo (five participants) (see

Table 2). While volume was also considered to be an important factor there was much less agreement overall as a third of the participants (six) rated it as the least important parameter. Similarly, spatial arrangement was considered the least important factor but a third of respondents ranked it the second most important factor. This variance between participants points to a high degree of subjectivity involved in the meaning-making process when looking at factors beyond pitch, particularly with regards to spatial dynamics, with some users finding their application beneficial and others considering them near redundant.

Table 2: Listener ratings of how parameters affect understanding of data and their overall perceived importance.

Affects understanding:	Volume	Pitch	Tempo	Space
Strongly	4	9	5	5
Somewhat	11	7	10	6
Not much	4	3	3	4
Not at all	–	–	1	2
Don't know	–	–	–	2

Importance:	Volume	Pitch	Tempo	Space
1st	1	12	5	1
2nd	8	2	3	6
3rd	4	4	7	4
4th	6	1	4	8

Overall, listeners appeared more attuned to recognising changes in pitch as an indication of movement. Connotatively, frequency change may be easier for listeners to transfer visual understandings from data visualisation, with the increase in pitch being the most obvious aural equivalent to an upwards movement. Whereas changes to volume and tempo may be more subtle for untrained listeners, frequency movement takes a central position within the sonification and therefore may be the clearest indication of change. With regards to interpretation of sonification and accessibility for different user experience levels, Vickers [3] draws distinctions between a tendency to favour simplistic methods of mapping to popular music listening: if frequency is the parameter that can be most clearly linked to a musical melody, then this is something that general listeners will likely be attuned to. By contrast, popular music is less likely to explore drastic changes in amplitude or tempo, which may contribute to the perceived importance from general users.

#### 4.5. Understanding of spatialisation

For trained listeners, it is assumed that a background in music will be make them more adept at listening out for subtleties present in the timbral qualities of sonification than general listeners. Whilst attributing more importance to spatialisation, amplitude, and tempo than untrained listeners, the near-universal favouring of frequency suggests that this was the dominant mode through which sonification was understood by most participants. Particularly with spatialisation, but also to a certain extent with amplitude and tempo, drastic changes to these variables are not common in many musical traditions; it may be that listener attention has been diverted by frequency as the most immediate parameter with which they are familiar.

Considering further the spatial properties (Table 3), five of the twelve participants with little or no understanding of musical theory rated spatial arrangement as the fourth most important factor

Table 3: Listener rankings of *spatial* parameter usefulness.

	Overall (n=19)	Understanding of music theory	
		At least some (n=7)	Little/none (n=12)
1st	1	–	1
2nd	6	3	3
3rd	4	1	3
4th	8	3	5
<b>Total</b>	<b>19</b>	<b>7</b>	<b>12</b>

and only one rated it as the most important. Three of the twelve musically-educated participants rated spatial arrangement as the second most important factor in understanding the data with the same number rating it as the least important. Overall, spatialisation was considered the least important factor with nearly half (eight) of the participants ranking it bottom. This could indicate that for trained listeners spatial dynamics can offer some level of enhanced experience of sonification. When considering untrained listeners the role of spatialisation is less clear at this time.

In addition, participants were asked to identify where they thought the sound was arranged spatially in their headphones to assess if any changes associated with movement were more useful than others. They were asked to mark one or multiple areas on a visual representation of the binaural listening space where the sounds were concentrated, or movement was most clear. Overall, participants felt that the greatest concentration of sound was in the top left and left of centre areas (Fig 3). Comparing responses from those with musical backgrounds and those without revealed no meaningful difference between the two groups. Overwhelmingly, the areas where sounds were harder to identify were across the lowest areas.

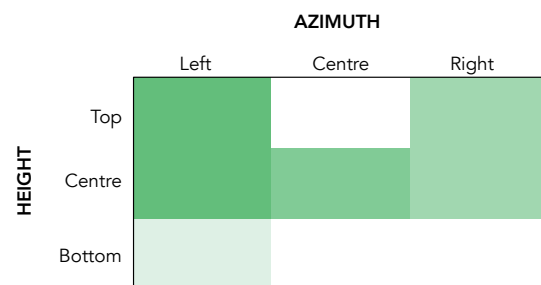


Figure 3: Overall listener perspectives on spatial distribution of sonification.

Listener perceptions differed somewhat from the actual spatial arrangement of the sonification. During the course of the sonification there was movement from all three variables which took differing spherical routes. Given the constant movement and differing paths of the three variables, there was concentration in all areas of the listener’s headphones at various points, meaning that across the piece the sound was distributed relatively evenly across the space. This suggests that listeners were able to perceive sounds more easily on left and right sides of their headphones or in the centre but were less able to identify whether a sound was occurring centrally in wider areas based on height. Similarly, listeners appeared to be less able to identify sounds concentrated in lower areas.

Finally, listeners were asked how spatial movement is reflected

by spatial dynamics. Six participants reported that they were unsure how spatial changes could reflect changes in the data, indicating that space is less obviously translated to conventional perceptions. Eight listeners suggested in some way that movement was connected to the increases and decrease in air quality levels, and six attempted to give an assessment of how this was reflected. The most cited change in spatial dynamics was left-to-right panning between headphones, with seven listeners noting this to some degree. By contrast, only one participant was able to identify the arrangement of each variable in a different height and width within their headphones and suggested that each variable carried its own path of motion which was influenced by changes over time. Three participants identified non-linear movement, referring to a circular clockwise motion at the widest points. Overall, whilst width was the most widely cited area of change noted by listeners, height was mentioned the least. This suggests that whilst listeners were not able to perceive spatial changes, they were able to use applied knowledge of sonification from the experiment to make suggestions about how space could be manipulated to reflect changes in a significant number of instances.

#### 4.6. Interview responses

Interviews were conducted with three participants with varying levels of experience to supplement the questionnaire responses. Interviewees were asked to listen to a full length sonification measuring three variables during the period sonified in the listening test: temperature, air pressure, and humidity. Interviewees' responses were focused on the experience of spatial arrangement and how this informed their understanding alongside other audio parameters. As with the previous listening study responses, interviewees considered that pitch was the most obvious indicator of change. When asked to consider how spatial dynamics influenced understanding, interviewees indicated that the perception of the spatial dynamics was heavily informed by other parameters. One reported 'when the pitch went higher, I thought of how the sound went up spatially too'. This implies that over the course of the sonification the height of the melody rose within the headphones as the frequency increased, which could be linked directly to the raised temperature from the dataset. This shows that for some listeners spatial dynamics are most useful when used in tandem with other parameters. It should be noted that while one variable's spatial movement was linked to height, temperature was not. In this sense, the listener's perception of upwards movement in spatial dynamics was informed more by preconceptions about frequency where high pitch is equal to high spatial placement.

This raises an important consideration within sonification about retrospectively attributing meaning — more so than visualised data, which is often static — audio, including sonification, is experienced in the continuous present. Therefore, much of the meaning making process occurs after the listening has taken place. Whereas listeners appear to be able to recall changes more readily in frequency as well as drastic changes in amplitude and tempo, participants were less assured in noting where a sound was focused spatially, even if they could tell spatial changes were occurring. Spatial arrangement does have positive implications in its contribution to sonification, however these are best understood as an extra parameter to bolster the use of other sonic parameters, especially frequency. Interviewees posited that one difficulty in understanding the spatial aspects could be the lack of Ambisonics in day-to-day experience. One interviewee reasoned that:

... the way that people read sound is linked to context so that if, for example, you're watching a horror film, an intense moment would be signified by an increase in the pitch of the music in that film. Whereas people have less experience of reading sound in a space than sound as a sort of flat medium.

This suggests that with certain sonic parameters listeners are training through general everyday listening which may guide the watcher to indicate their experience through frequencies or volume. By contrast, spatial dynamics are less commonly explored in everyday contexts, particularly through headphone use. When listening to music, spatial rendering is usually applied after sound has left the speakers and is determined by the literal space a listener inhabits. The same participant suggested that spatial rendering may have been easier to ascertain meaning the use of surround sound in a physical space such as a gallery space or concert hall. Applying spatial dynamics for binaural listening therefore appears to be difficult even with trained musical listeners — movement within headphones can be identified but understanding of how and why appears to be more complicated for users to establish. This aligns with Vickers' assertion of embodied interaction, arguing that interaction as much includes the environment and space in which the listening takes place as it does the experience of the sonification itself [5]. This suggests that for listeners with varying experience, the listening environment is a significant factor in informing their experience of spatial dynamics in sonification.

## 5. CONCLUSIONS

This work explored how listeners with differing levels of expertise experience sonification, and to what extent the spatial arrangement of audio can contribute to sonification practice. The results suggest that perception of spatialised audio can contribute to embodied interaction, and that the spatial mapping worked most successfully in providing an extra dynamic layer to parameter mapping sonification that also makes use of other sonic qualities such as frequency change when rendering data. In terms of listener experience contributing to perception, musically trained listeners were more likely to cite the importance of more subtle factors including spatial arrangements than untrained listeners, however, interviews with participants of differing experience levels cited the difficulty of binaural rendering to adequately perceive spatial arrangement. Untrained listeners can perceive spatial dynamics at play when listening to sonified data, however, there appeared to be less logical application of how the spatial arrangement could be manipulated to interpret the data. In this regard, embodied perception would appear to be intrinsically linked to physical space, which the binaural audio did not adequately replicate. In terms of experience levels overall, responses indicated that sonification still occupies a niche space in wider consciousness beyond researchers within the field, and as a result further training may be required to train uninitiated listeners in how to interpret data aurally.

The benefit of binaurally rendering Ambisonic audio is useful for sonification researchers in future research as the study demonstrated the ability to engage with general users without the high barrier for entry posed by technical requirements and can, it is hoped, be used to train or enrich the experience of sonification listening for participants in future studies. Given that the results indicated that untrained listeners with little knowledge of sonification found perception most difficult, future studies should interrogate the experience of spatial dynamics in sonification on more

targeted user groups, particularly those with no engagement in the field. Additionally, it may be useful to test spatial sonification using non-time-based datasets, as the study showed that the use of temporal data has implications for listener perception that are related to other parameters such as “movement” of frequency. By contrast, Ambisonic sonification based on spatial datasets including those of geological or cartographical varieties may be less tied to preconceptions surrounding sonic parameters.

## 6. REFERENCES

- [1] H. Ballweg, “Ways of guided listening: Embodied approaches to the design of interactive sonifications,” Ph.D. Thesis, Northumbria University, Dept. of Computer and Information Sciences, Newcastle upon Tyne, UK, June 2018.
- [2] S. Roddy and D. Furlong, “Embodied aesthetics in auditory display,” *Organised Sound*, vol. 19, pp. 70–77, 4 2014.
- [3] P. Vickers, “Sonification and music, music and sonification,” in *The Routledge Companion to Sounding Art*, M. Cobussen, V. Meelberg, and B. Truax, Eds. Oxford: Routledge, 2017, pp. 135–144.
- [4] P. Vickers, B. Hogg, and D. Worrall, “Aesthetics of sonification: Taking the subject-position,” in *Body, Sound and Space in Music and Beyond: Multimodal Explorations*, ser. SEMPRE Studies in The Psychology of Music, C. Wöllner, Ed. Routledge, 2017, ch. 6, pp. 89–109.
- [5] P. Vickers, “Sonifications sometimes behave so strangely,” in *Bloomsbury Handbook of Sonic Methodologies*, ser. Bloomsbury Handbooks, M. Bull and M. Cobussen, Eds. New York: Bloomsbury Academic, 12 Dec. 2020, ch. 47, pp. 733–743.
- [6] S. Thornton. (2009) Ambisonics. Available: <https://www.michaelgerzonphotos.org.uk/ambisonics.html>
- [7] J. Boehringer, J. Bowers, P. Vickers, T. Shaw, G. K. Sharma, B. Hogg, M. Pietruszewski, and J. W. Newbold, “The many ways of listening: A provisional analysis of talk in critical listening sessions to inform the research agenda of sonification,” in *ICAD 2021 — The 26th Meeting of the International Conference on Auditory Display Meeting of the International Conference on Auditory Display*, A. Andreopoulou, M. Droumeva, and J. W. Newbold, Eds. ICAD, 25–28 June 2021.
- [8] T. Hermann, A. D. Hunt, and J. Neuhoff, “Introduction,” in *The Sonification Handbook*, T. Hermann, A. D. Hunt, and J. Neuhoff, Eds. Berlin: Logos Verlag, 2011, pp. 1–6.
- [9] B. N. Walker and M. A. Nees, “Theory of sonification,” in *The Sonification Handbook*, T. Hermann, A. D. Hunt, and J. Neuhoff, Eds. Berlin: Logos Verlag, 2011, pp. 9–39.
- [10] P. Vickers and B. Hogg, “Sonification abstraite/sonification concrète: An ‘aesthetic perspective space’ for classifying auditory displays in the ars musica domain,” in *ICAD 2006 - The 12th Meeting of the International Conference on Auditory Display*, T. Stockman, L. V. Nickerson, C. Frauenberger, A. D. N. Edwards, and D. Brock, Eds., London, UK, 20–23 June 2006, pp. 210–216. Available: <http://hdl.handle.net/1853/50641>
- [11] T. Hermann, “Taxonomy and definitions for sonification and auditory display,” in *Proc. 14th Int. Conf. Auditory Display (ICAD 2008)*, B. Katz, Ed. Paris, France: ICAD, 24–27 June 2008.
- [12] L. M. Brown, S. A. Brewster, R. Ramloll, M. Burton, and B. Riedel, “Design guidelines for audio representation of graphs and tables,” in *ICAD '03 9th International Conference on Auditory Display*, Boston, MA, USA, July 2003, pp. 284–287.
- [13] S. Harrison, D. Tatar, and P. Sengers, “The three paradigms of HCI,” in *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems San Jose, California, USA*, 2007, pp. 1–18.
- [14] P. Dourish, *Where the Action Is: The Foundations of Embodied Interaction*. The MIT Press, 09 2001. Available: <https://doi.org/10.7551/mitpress/7221.001.0001>
- [15] G. K. Sharma, “Composing with sculptural sound phenomena in computer music,” Ph.D. dissertation, Kunst Uni Graz, Graz, Austria, 2016.
- [16] E. Childs and V. Pulkki, “Using multi-channel spatialization in sonification: a case study with meteorological data,” in *ICAD '03 9th International Conference on Auditory Display*, E. Brazil and B. Shinn-Cunningham, Eds. Boston, MA: ICAD, 2003, pp. 192–195.
- [17] E. C. Cherry, “Some experiments on the recognition of speech with one and with two ears,” *Journal of the Acoustical Society of America*, vol. 25, no. 5, pp. 975–979, 1953.
- [18] J. G. Neuhoff, “Perception, cognition and action in auditory displays,” in *The Sonification Handbook*, T. Hermann, A. D. Hunt, and J. Neuhoff, Eds. Berlin: Logos Verlag, 2011, pp. 63–86.
- [19] B. N. Walker and G. Kramer, “Ecological psychoacoustics and auditory displays,” in *Ecological Psychoacoustics*, J. G. Neuhoff, Ed. London: Elsevier Academic Press, 2004, pp. 149–174.
- [20] P. Bourdieu, “Cultural reproduction and social reproduction,” in *Knowledge, education, and cultural change*. Routledge, 1973; 2018, pp. 71–112.
- [21] A. Beadling, “nuson-ambison: Ambisonic Sonification,” May 2023. Available: <https://github.com/nuson/ambison>
- [22] (2023) IEM Plug-in Suite. Available: <https://plugins.iem.at/>
- [23] M. Kronlachner. (2014, Jan.) ambix v0.2.10 – Ambisonic plug-in suite. Available: <http://www.matthiaskronlachner.com/?p=2015>
- [24] (2023) Existential Audio BlackHole: Audio loopback driver. Available: <https://github.com/ExistentialAudio/BlackHole>
- [25] Newcastle University. Urban Observatory data repository. Available: <https://urbanobservatory.ac.uk>
- [26] Met Office. Record high temperatures verified. Available: <https://www.metoffice.gov.uk/about-us/press-office/news/weather-and-climate/2022/record-high-temperatures-verified>
- [27] S. Barrass and P. Vickers, “Sonification design and aesthetics,” in *The Sonification Handbook*, T. Hermann, A. D. Hunt, and J. Neuhoff, Eds. Berlin: Logos Verlag, 2011, pp. 145–172. Available: <http://bit.ly/3r9tsof>