

ABSTRACT:

This chapter provides an introduction into the world of sonification, why it exists, and how we may benefit from it. It presents the material in a way that requires minimal prior knowledge and defines key terms necessary for the content comprehension. A series of common real-world scenarios are presented and revisited throughout the chapter to illustrate key traits and the underutilized sonification potential. They are followed by some of the common-sense strategies that have emerged from this nascent field of research. Further, the chapter explores the full potential of sonification to help us quantify the limits of human ability to perceive and interpret spatial audio streams while also sidestepping some of the key limitations of the current virtual approaches to spatializing sound.

Keywords: sonification, data, information, spatial audio, representation

8

Introduction to Sonification

Ivica Ico Bukvic

8.1 Introduction

Imagine you are driving a car. As an experienced driver you may be able to identify and monitor a number of aural cues, such as the car engine and wind noise, as well as the dashboard sounds designed to better inform you of the current driving conditions. Whether you are driving a vintage car with a manual gear shift or an automatic, listening to the engine pitch as it increases its rotations per minute (RPMs) may help you pinpoint the optimal time to switch gears. Or, it may help you identify a potential problem with the transmission that suddenly fails to shift into a higher gear, leaving the engine stuck at screeching high RPMs. As you speed up, the wind noise increases, reinforcing the sense of speed. Concurrently, your dashboard plays a symphony of cues or earcons (Oswald 2012), reminding you that you may have forgotten to fasten the seat belt, fully close one of the doors, or that you may be running low on fuel or electrical charge.

It starts to rain. As raindrops begin hitting the windshield, the sudden increase in their rate reminds you of summer storms and the ensuing flash floods that often lead to dangerous driving conditions. You instinctively slow down.

Now snap back out of our imaginary car ride and consider just how many different aural cues were present in the scenario. We had the engine noise, the wind noise, multiple dashboard sounds and then the rain. We could also dig deeper and consider other noises that may have punctuated various actions, like shifting gears. Perhaps the car was old and every time it shifted gears the transmission generated a disturbing clunk. Or, moving the gear shift offered a satisfying thump as it reassuringly landed into a new position.

All of these sounds can be seen as a form of sonification. Some of them are simply byproducts of natural laws that govern our universe, such as the wind noise, engine pitch, or the raindrops hitting the windshield. Others were designed and engineered by humans for humans, such as the dashboard earcons whose purpose is to remind the driver of items that may require their prompt attention.

8.2 So, What Is Sonification and Why Should You Care?

The origins of the term *sonification* are not entirely clear. Arguably its first official mention in an academic publication dates back to 1990 (Rabenhorst et al. 1990) when it referred to a complementing aural counterpart to data visualization. Since then, its definition has continually evolved, resulting in a number of similar yet distinct definitions, many of which continue to be used concurrently. Worrall, in his overview of sonification (Worrall 2009), presents a comprehensive review of the term's evolution, while also offering the all-encompassing definition:

Data sonification is the acoustic representation of data for relational interpretation by listeners, for the purpose of increasing their knowledge of the source from which the data was acquired.

Another, considerably simpler, version offered by Kramer et al. (1999) defines sonification as “the use of non-speech audio to convey information.” In this chapter, allow me to propose what may be arguably the simplest and most inclusive version, which defines sonification as “audio that conveys information,” thereby leaving all the possibilities for sonification open, including human speech, music and even environmental sounds, as is the case with the rain and wind noises in our imaginary car ride example. Unlike the dashboard sounds that have been engineered by humans, the environmental sounds simply exist due to the laws of nature that

govern our universe. As such, they could be seen as unintentional but nonetheless useful byproducts whose usefulness is attained through repeated exposure and growing sensitivity to their nuances.

While this definition stretches the conventional boundaries of sonification that are limited to solely intentional manipulation of sound by humans for humans, it offers unique opportunities for us to observe and learn from how we interact with the environment. Such observations may in turn serve as an inspiration for optimal ways to design an intentional data sonification.

When considering any of the aforesaid definitions, it is important to understand the difference between data, information, and knowledge. Data can be seen as a measurement of change in a phenomenon that commonly exists regardless of our awareness. Consider today's weather monitoring systems that are capable of providing incredible amounts of weather data from all over the globe. This data is simply human-readable representation of changes in the environment that would exist regardless whether we are consciously monitoring them, or whether we are doing so physically (e.g. by being physically present at a particular location and using our sensory mechanisms) or through technology. We could probably spend an entire chapter on the philosophical discussion of whether the data can exist without conscious human involvement. This, however, is not our focus. Instead, I will close this thought by repurposing the old adage: if a tree falls in a forest and no one is there to account for it, does it generate data?

Back to the global weather data example—if we feel compelled to do so, we could go and read all the exciting (or boring) temperature, humidity, pressure and wind speed and direction data from any place on Earth. Most of the time, however, we don't do this. Instead, we focus on the data that most immediately affects us, namely that related to our current location, or possibly a travel destination. When observed, such data becomes information. By reviewing it, we acquire information, such as the temperature or humidity levels, or the percentage of the cloud cover. And, if we have adequate prior experience that we may be able to cross-correlate with the newly acquired information, we may utilize such knowledge to make appropriate observations and

decisions, such as it will be warm and that we may need to dress lightly or pack appropriately, or that we may want to bring an umbrella.

So, what does all this mean for someone who has potentially no prior experience with sonification and more importantly, why should one care? For this, we need to take a side trip to the world of perception and cognition.

8.3 Human Perception

Consider that your understanding of the surrounding environment is rooted in your ability to perceive it and interact with it. We perceive the world through our senses of sight, hearing, taste, smell and touch. In addition, we can also sense temperature, balance, movement, position, and (with varying aptitude and consequently punctuality) the passage of time. On an existential level our very understanding of reality is inextricably tied to our individualized senses and their inherent and unique idiosyncrasies. Yet, what we all have in common is that we strive towards a state where we can process more information more efficiently, so that we can better anticipate change and therefore make better decisions. Every sense plays an important role in this effort and offers its own *unique affordances*. Studying such unique affordances may help us uncover new ways to perceive data and extract information, as well as *broaden our cognitive bandwidth*, thereby allowing us to consume and interpret more data more efficiently. Doing so will empower us to reinforce existing knowledge and build entirely *new knowledge*. Such knowledge will in turn enable us to spot previously learned patterns or uncover new ones, and apply them in other contexts. Below, we will briefly focus on each of the three aforementioned key elements of human perception and, consequently, sonification, namely unique affordances, the broadening of cognitive bandwidth, and new knowledge.

8.3.1 Unique Affordances

When it comes to experiencing data, we most often resort to a visual rendition, a chart or a bar graph. Yet, our vision, as mighty as it may be, is not the end-all solution. After all, we have multiple senses for a reason—to cover aspects the other senses are either unable or not particularly good at addressing. When compared to human vision, one of the unique affordances of human hearing is a much higher data resolution we can consciously process per second. Consider when playing a fast-paced video game, arguably one of the more broadly accessible time-critical vision-centric activities. In most situations we are perfectly satisfied with 60 pictures per second (aka frames per second or FPS). And, if we were ever to experience a second where we only received 59 frames instead of the usual 60, possibly because our computer's video card or graphics processing unit (GPU) was unable to render all 60 frames in time, chances are we would never notice such a drop in the overall frame rate, let alone pinpoint the moment in which such a frame drop may have occurred. When it comes to sound, even the now dated compact disc (CD) quality offers 44,100 points of data per second and while we may not be able to continually detect every subtle data point individually, when combined with neighboring data points, even an untrained ear is capable of identifying changes in pitch, timbre, loudness and location. More importantly, an impulse whose length is $1/44,100^{\text{th}}$ of a second is surprisingly easily identified as a simple “click.” Although we are clearly not comparing apples to apples, consider the stark contrast between the aforesaid click that takes place for only a fraction of a second to one dropped frame among 60 that remains by and large unnoticed. This is in good part why in our imaginary car ride the aural cues emanating from the dashboard can be incredibly brief—consider the sound of a turn signal typically going tick-tock that is not unlike the aforesaid impulse, versus a considerably more sluggish rate of the blinking LED that serves as its visual counterpart.

While not entirely unique, another particularly potent affordance of hearing perception is our ability to recognize patterns. Consider a CD that is stuck looping a small chunk of music due to damage to its surface and how quickly you are capable of identifying the ensuing skipping or looping just by listening. In this context, the looping audio signal is a repeating data set consisting of potentially thousands of data points, much like a repetition one may encounter in a graph, such as the “head and shoulders” pattern commonly encountered in stock market data that may signal looming change from bullish to bearish market and vice versa. Our hearing is incredibly sensitive to such temporal patterns offering a unique opportunity to apply this same strength towards other forms of data, including abstract data, charts, bar graphs and beyond.

8.3.1.1 Psychoacoustic Properties

When perceiving a sound, we are typically capable of identifying pitch (or lack thereof, as is the case with noisy sounds, such as the sound generated by saying “shhh”), loudness, timbre or sound color and duration, as well as location in respect to our position. All of these dimensions of the sound are psychoacoustic interpretations of its acoustic properties. The dominant frequency or the rate of repetition of a periodic waveform per second expressed in hertz (Hz) is an acoustic property we can interpret as pitch. Here we use the term *periodic* to suggest the sound in question has a clearly repeating waveform pattern, which in turn allows us to perceive the dominant frequency. It is interesting to consider that our perception is logarithmic in nature (Varshney and Sun 2013). Consider that the note identified as A4 in Western musical notation (~880 Hz) is double the frequency of A3 (~440 Hz). At the same time, A5 (~1,760 Hz) is double that of A4 (~880 Hz) and therefore effectively double the frequency range between A3 and A4. Yet, in our ears we hear relationships between A3 and A4, and A4 and A5 to be equivalent, both distances being perceived as an octave apart. Pitch is not necessarily always perceivable, as is the case in the aforesaid “shhh” example, or the wind noise in our imaginary car ride scenario. Both examples resemble spectral composition (or timbre) of a white noise. In such cases there is no distinct dominant pitch, a property that in and of itself allows us to identify such sound and

separate it from others that may be pitched. We refer to such sounds as non-pitched or aperiodic. Most sounds in nature, however, exist somewhere in between these two extreme cases where they possess pitched qualities while also being noisy. Consider, for instance, different cymbals in a drum set, each of which may have a different implicit pitch.

Loudness is derived from the sound waveform's amplitude. The greater the amplitude, the louder the sound. Just like pitch, it is logarithmic in its nature, although our perception resolution of different loudness levels is arguably less than that of pitch, as is evident in a considerably sparser choice of dynamics in music notation (commonly *pianississimo*, *pianissimo*, *piano*, *mezzo-piano*, *mezzo-forte*, *forte*, and *fortissimo*, and *fortississimo*). Timbre is a curious beast. Essentially, any sound you have ever perceived is simply a combination of a large number of individual frequencies with distinctive and continually varying loudnesses and whose cumulative result is a unique color. This is why a note A3 played by a violin and a piano sound different and clearly distinguishable, despite the fact both sounds have the same dominant pitch. Human voice can generate a rapidly changing collection of dominant frequencies in terms of both their pitch and loudness, producing vowels, consonants, syllables and words. On a purely philosophical level, silence can be seen as a form of sound that is devoid of any perceivable frequencies or a sound whose loudness is zero, whereas the white noise is composed of all perceivable frequencies. As such, when comparing these outlier cases with the visual domain, one may regard silence as the color black versus the white noise as the color white.

Another implicit perceived property of sound is its duration, or what is more broadly defined as its temporal component. The reason we do not want to limit ourselves to using time solely for measuring duration is because extreme changes in duration may potentially dramatically alter other sound's properties. Consider a pitched sound, an A4 sustained note played by a violin. If we were to splice a recording of such a sound into a tiny snippet, thereby using solely the temporal dimension, not only would we lose the ability to perceive the sound's pitch, but would also potentially significantly alter the perceived timbre, ostensibly converting it into an impulse-like "click." Consequently, it is important to consider the inextricably

intertwined codependence of the different acoustic and, by extension, psychoacoustic properties of sound. Just like the aforesaid example in which the duration affects the perceived pitch and timbre, consider how loudness changes may also alter aspects of the sound perception. Consider the following theoretical example where the sound's loudness has been made zero. Under such circumstances the sound has lost all its identifying properties, becoming effectively synonymous with silence. As we will soon learn, this seemingly dubious example may prove particularly helpful when considering sonification.

While all of these dimensions are unique affordances of sound perception, perhaps the most compelling unique affordance of human hearing is the ability to locate a sound. Unlike vision, which is predominantly anterior in its nature, meaning we only see what is in front of our eyes, human hearing is effectively spherical. We can hear with varying levels of accuracy in all directions around us even though we only have two ears. Pinnae, or the cartilage some of us pierce to hang earrings, is what gives us that third dimension by filtering sound that comes from behind, above and below. From the moment our ears begin to broadcast signals to our brain, the brain begins to learn and refine its ability to differentiate sound source locations. Given that each of our pinnae is unique, our individual brains are also individually calibrated to account for such deviations in our anatomies.

The ability to hear all around us is also a critical survival skill. Whether it is a saber tooth cat that may have been preying on our ancestors, possibly trying to sneak from behind, or a bus whose horn caught our attention at the very last second helping us avoid becoming a hood ornament, our all-encompassing spherical spatial perception is something truly unique to our hearing. In our imaginary car ride, distinguishing the location of the dashboard sounds, windshield raindrops, strange noise that may be emanating from the rear left wheel well or a plethora of other inherently spatial aural cues is what further reinforces our ability to differentiate them. Sound spatialization also appears to be the last frontier in the emancipation of sound properties. For at least the past two millennia, human civilization has spent time refining pitch, loudness and eventually timbre, allowing each of these dimensions to drive musical expression.

Consider how in the twentieth century we encounter for the first time a large volume of music whose structure is driven primarily by timbre, particularly the percussion ensemble and computer music. With this in mind, the twenty-first century may very well be defined by the emancipation of the spatial component where musical motives may be rooted in sound's location, rather than its well-established dimensions of pitch, loudness, timbre and the inevitable temporal component, or duration. The similar trend is also apparent within the field of sonification. Even though we have a large number of spatialization algorithms at our disposal that allow us to simulate sound motion across multiple loudspeakers, their use remains relatively limited.

Consider once again our imaginary car ride and how even luxury cars continue to sound out earcons from some dubious location on the dashboard, instead of utilizing one or more of its typically rich six-channel (or greater) surround-sound system. Wouldn't it be a lot easier to locate which door may be ajar if we were to simply play a particular earcon from the loudspeaker mounted on that particular door? Instead, more often than not, we find ourselves taking our eyes away from the road, the very thing we are not supposed to do, and staring down at the dashboard where a pretty picture of our car is meant to show us what door may be ajar. Or, worse yet, we struggle to decipher what the earcon is even supposed to mean simply because the car manufacturer never bothered to design different earcons for different events and warnings. Indeed, spatialization is an oft neglected aspect of human hearing, and consequently sonification, that offers ample opportunities for broadening the cognitive bandwidth. So what does that mean?

8.3.2 Broadening Cognitive Bandwidth

Broadening cognitive bandwidth by combining multiple senses is a well-established concept within cognitive load theory ([Paul A. 2002](#)), which has seen resurgence in the recent big data movement. There are two ways to observe its implications for cognitive bandwidth. One way is to reinforce a particular cue using multiple sensory inputs. For instance, in our imaginary car ride

scenario, the turn signal offers both visual and aural cues to make it easier for the driver to be aware of the fact they have a turn signal on. On the other hand, we could also consider there is only so much we can process at any given moment using any single sense. If we couple two or more senses, it is likely we will have greater capacity to process more information simultaneously where each sense may be assigned to a different data source. This may not necessarily mean the ensuing bandwidth or the new whole will be greater than the sum of its parts. Yet, in time-critical situations, even a fraction of an improvement is still clearly better than perception bandwidth through any single sense. Consider the imaginary car ride we discussed at the beginning of this chapter. Most, if not all of the sonification elements occur concurrently while you are driving a car, an action that requires vision, touch, motor skills, a sense of motion, and potentially even a sense of smell that may help us detect that strange odor coming from the engine. Therefore, we can easily process an aural cue from the dashboard informing us that we may be running low on gas while still keeping an eye on the road and controlling the vehicle. Similarly, we may be listening to the radio, which, unless the sound levels are so loud that they prevent us from being able to perceive environmental aural cues (akin to what a blinding light does to our vision), is a straightforward matter that otherwise has no adverse effect on our ability to drive the car.

It may be useful to further differentiate between cognitive bandwidth achievable by cross-pollinating two or more senses and the two or more aspects within the single sense. For instance, perceiving two data sources, one visually, and one aurally, may yield greater cognitive bandwidth than when trying to pack both data through solely sound or sight. On the other hand, one could also consider projecting two different data sources by using pitch and loudness. Lastly, in either example we could use one and the same data point to drive both dimensions, thereby lowering the ensuing cognitive load, or at least increasing chances of us being able to perceive them in a timely fashion. For instance, if the car we are driving is in a heavy rainstorm, the ensuing raindrop noise may occlude sound generated by the car blinkers. In this situation, having

an accompanying visual cue on the dashboard may help us minimize chances of having a blinker stuck in the on mode long after we have already changed a lane or taken that turn.

It may now seem obvious that one of the greatest advantages of sonification is its coupling with visual and other sensory inputs, with the recent revolution in virtual/augmented/mixed reality technologies being an obvious target. However, despite decades of research, sonification is still a nascent field (Bukvic and Earle 2018), and as such stands to benefit from the early-stage exploration and codifying of its unique affordances in isolation from other sensory inputs that may conflate any such assessment. As was the case with visualization, doing so may help us identify new approaches to pattern recognition and knowledge building that in turn may be applicable to other domains.

8.3.3 Knowledge Building

Remember the previous skipping CD example and our ability to near instantaneously spot such an anomaly simply by listening? Now consider how once we are exposed to a particular instrument, for instance, a snare drum or a guitar, how well we may have memorized its timbre, and have become capable of identifying the same instrument in a potentially complex soundscape, such as an orchestra performance or a rock band song. This is where our brain's ability to recall vivid details of previously learned timbres and cross-correlate them in real-time against perceived patterns really comes to focus. How about a few more examples? How well do you still remember a song you grew up with or got married to and how well you can anticipate its moments even after only a few repeat listenings? Or, that album that you listened to multiple times in the same order to the point where when one song ends, isn't it fascinating how precisely you know what is coming up next or just how brief are the pauses between the songs? Unlike the skipping CD example where we may be facing an unfamiliar sound and are still able to quickly spot a repeating pattern, the aforesaid scenarios are examples of our brain's powerful ability to

recall complex data patterns stored in the form of music and/or sound and voice color memories, much like our ability to spot and recognize familiar faces.

Human hearing indeed offers ample opportunities for exploring such knowledge building and in turn pattern recognition in a rich, immersive, spatially aware environment. What is particularly exciting is that, much like our recognition of familiar faces in a complex image, this process is inherently subconscious and our brains allow us to detect these patterns quickly and with no added conscious effort. Coming back to our imaginary car ride, consider while driving in a potentially noisy environment (raindrops on the windshield, wind and engine noise, radio, etc.), detecting a familiar dashboard sound (aka an earcon) occurs naturally to us even though we may have made no conscious effort to monitor for its potential occurrence.

While the idea of using sonification to identify an entirely new pattern in a known data set (e.g. a stock market or geospace data) seems like the proverbial Holy Grail that would validate the proposed value and importance of sonification, it is worth considering that sonification may be also seen as “an under-utilised dimension of the ‘wow!’ factor in science engagement multimedia” (Ballora 2014). As such, regardless of our ability to generate entirely new knowledge, on a purely educational and engagement level sonification may prove an indispensable vehicle for facilitating understanding of known concepts.

To summarize, sonification is a nascent field that offers unique affordances that may help us uncover new knowledge in the existing data. When coupled with other senses, it has a potential to broaden the user’s cognitive bandwidth and increase efficiency. It can also serve as a potent way of representing known concepts to promote learning and engagement.

8.4 Is It Art or Science?

For sonification to be possible it needs a data source. Consider that the data can be found everywhere around us as well as in anything we do or create. It may be the number of hours we slept last night, or how many times we chewed each bite of our sandwich at lunch. Consequently,

it can be extracted from any creative domain, whether it is the sciences, arts, engineering, design or, as is more common, a combination thereof. While generating data from a scientific experiment may seem straightforward, imagining the same in music, for instance, may at first seem strange. Consider for a moment all the different ways we can analyze music using music theory concepts, such as harmonic motion, the recurrence of the theme, or the overall form. The ensuing analysis offers ample data points much like a scientific experiment. Moreover, the very audio recording can be seen as a form of data sonification where the data are the notes in the score. Another way to look at this strange and rather malleable relationship between the arts and sonification is through the lenses of a growing number of art works where the aural and/or visual content is generated by sonifying and/or visualizing some sort of a data stream. The ensuing artifact, whether it is an interactive art installation, sound art, or a musical composition, is effectively a sonification of a data set shaped by an artist's aesthetic choices.

8.5 Classifications

Considering that sound perception is time-based, sonification is by and large focused on rendering continuous data stream over time, such as the changes in weather or stock market data, or engine noise. The sonification timeline does not have to match the normal passage of time and, like video, it can be played faster or slower, as well as forwards and backwards. Rather than continually outputting sound, sonification can also focus on punctuating a specific threshold, as is the case in our imaginary car ride where a single sound is played once the fuel level reaches a previously designated low point. As a result, such a cue or an earcon can be seen as a subset of sonification. Unlike the continuous data stream examples, however, earcons are typically interpretations of previously known patterns or conditions (e.g. a previously designated low fuel tank level), whereas continuous streams tend to be raw and unprocessed and left up to our brains to interpret and spot any potential patterns, as is the case with our skipping CD example. Furthermore, the continuous data streams can be generated in two different ways—by taking the

data and feeding it directly into an audio output (e.g. a loudspeaker) with minimal processing, or by using it to modify a property of an engineered sound, such as its pitch, timbre, loudness, location, or a combination thereof.

Audification is a subset of sonification. It is essentially taking data and with minimal processing converting it into sound. The reason processing may be necessary is because some data ranges may not produce a usable audio signal. For instance, we may want to audify changes in daily temperature. In order to do so, we may need to normalize the data ranges, meaning take the lowest and highest possible temperatures and map them onto the waveform ranges that in a digital domain typically span from -1 to 1. We may further need to adjust its frequency or the speed by which we feed it to the loudspeaker, in order to generate a waveform that is audible to human ears. This, by convention includes frequencies between 20Hz and 20,000Hz, although in practice there are a few outlier cases where the perception of frequencies below 20Hz and 20,000Hz may be also possible. While potentially indispensable, this kind of sonification has a fairly specific, if not limited, utility.

Let's assume that our observed temperature is measured only hourly. If we had to generate even a single second of sound, we would require 44,100 such data points, which means we would need at least 1,837.5 days worth of temperature data. The ensuing single second of sound would effectively form an example of audification. In this case, however, given such dense packing of data, despite the high resolution of human hearing, listening to a short snippet is unlikely to allow us to accurately observe what may have occurred within one afternoon, which at this point would amount to a mere six samples of audio or barely 1/10,000th of a second. One way to address this is to slow down the playback of such events and allow the same afternoon to take place across several seconds. This, however, may result in a rather inefficient experience of the ensuing data, requiring minutes, if not hours of listening, needless to mention it may also generate a subsonic waveform, or a waveform that has a resulting frequency lower than 20Hz, that is simply not audible to humans. This is where sonification steps in to save the day. Rather than normalizing or slowing down the data, we may want to use it to drive the frequency

of a carrier signal, whether that be a synthesized (e.g. an oscillator) or a sampled sound (e.g. the sound of violin). Depending on the temperature ranges and the context, we may want to transpose it to make it more easily audible, or even multiply its temperature changes to make them more easily perceived. This kind of sonification indeed offers greater flexibility over audification. The temperature can also affect the loudness, timbre, or position. For instance, the resulting sound may get louder and “brighter” as a result of using a low pass filter, as well as move from left to right.

When considering earcons, another way to think about them may be that the monitored data is continually sonified, except that its loudness curve has an extreme exponent and that only when it reaches the peak of its normalized range does it make any perceivable sound. Of course, in practice this is typically not the case. However, as a theoretical construct it may offer a way to reconcile the perceived differences between a momentary cue and a sonification of a continuous data stream.

8.6 Sonifying Data

A growing body of sonification research, whether a conscious scientific study or a byproduct of an artistic installation, sound art, or a composition, has yielded a broad array of approaches to sonifying data. Considering there are near infinite ways one may sonify a particular data set, and given that this area of research is still in its infancy, it should not come as a surprise that the results tend to be by and large case-specific and their lessons difficult to abstract. Nonetheless, over time several common-sense observations have emerged. They are listed below in no particular order.

8.6.1 Ability to Perceive Multiple Concurrent Psychoacoustic Properties

We are capable of concurrently perceiving multiple psychoacoustic properties. This means one could map different data streams to each psychoacoustic property of a sound. For instance, in our weather pattern data example the temperature could be used to drive sound's pitch, while changes in humidity could be perceived as changes in timbre. Naturally, there are limits to what we can concurrently observe. Such limits typically depend on many variables, such as the choice of sound, variance in the data, the mapping strategies, potential hearing perception limits of our target audience, and environmental conditions. As a result, this remains by and large a case-specific iterative trial-and-error design process.

8.6.2 Single Variable Sonification Using Multiple Psychoacoustic Properties Can Lead to Improved Perception

Humans are generally more capable of perceiving differences in sound if such differences manifest concurrently through different psychoacoustic properties. For instance, in the previous example, rather than mapping both temperature and humidity onto pitch and timbre, respectively, we may want to map temperature both on the pitch and timbre. This typically results in making changes to temperature easier to perceive, but there are once again limitations to what combinations of parameters yield useful results. For instance, having a low pass filter set too low on a high-pitched sound may render such a sound inaudible, thereby making its pitch and pitch-variance imperceptible to the listener. Another obvious downside is that by doing so we limit the number of variables we can concurrently present using one sound.

8.6.3 Ability to Perceive Multiple Concurrent Sounds

We have an innate capacity to perceive multiple concurrent sounds. Once again, the limits of our capacity vary depending on the aforesaid variables, namely the choice of sounds, variance in the data, the mapping strategies, potential hearing perception limits of our target audience, and

environmental conditions. For instance, having two sounds that have similar spectral composition or timbre may make one sound become masked by another, thus making it difficult if not impossible for a listener to identify and separate them. The opposite may be true, as well, whereby one sound may be so rich and spectrally variant that at times our brain may play tricks on us, making us believe we may have heard a sound that simply wasn't there. Has it ever happened to you that while taking a shower you may have heard a faint familiar sound, like your smartphone ringing, and as you scrambled out of the shower to pick up the phone you found out that the phone never rang? Such events are more common than you think, particularly in the aforesaid scenario, where the shower noise closely mimics white noise, or the presence of all perceivable frequencies at continually random and varying rates and intensities. Consequently, as different frequencies randomly vary in their strength, it is not unlikely for the sound to momentarily resemble another familiar sound, like that of our smartphone's ringtone.

8.6.4 Capacity to Spot Patterns in Complex Soundscapes

As we continue concurrently mapping different variables onto different psychoacoustic properties of a sound and/or using multiple concurrent sounds, we will eventually cross the threshold beyond which we lose the capacity to monitor all the individual data streams. Yet, all is not lost beyond this threshold. We still have an opportunity to leverage human ability to spot patterns, such as in the skipping CD example. Under such circumstances we may not be able to identify exactly what is going on with each individual variable, but we may be able to spot a cumulative timbre we may have heard before, particularly if it happens to have a distinct feature (e.g. it may be loud, "brighter" than the rest, may have an interesting interaction between two or more sounds due to phasing, and so on).

8.6.5 Difficulties in Tracking Steadily Changing Data

Spotting specific milestones in the sonification of steadily changing data is not particularly easy, especially as the rate of change and its range increase. Consider for a moment a violin that glissandos continually up and down. In situations like these, akin to the low fuel earcon in our imaginary car ride, such milestones can be further reinforced by punctuating them with added sounds. Another way to think of this is as if the said earcons were equivalent to the tick marks on a ruler.

8.6.6 The Cocktail Party Effect and the Auditory Scene Analysis

In designing a sonification, consider the cocktail party effect ([Arons 1992](#)) that allows us to carry a conversation even in noisy environments by focusing our hearing to a specific sound timbre or color. Bregman's auditory scene analysis ([Bregman 1994](#)) further elaborates on this concept by providing insight into how our mind is capable of partitioning multiple concurrent sounds based on sequential and simultaneous grouping. According to Bregman, sequential grouping pertains to a continuous stream coming from one source based on its acoustic parameters, including the perceived source location. Simultaneous grouping on the other hand focuses primarily on the perceived pitch. This, in combination with the spatial component, may help broaden our capacity to perceive multiple concurrent events. This is in part possible because our brain can bundle sounds we deem undesirable and relegate them to a background noise, like the sound of a computer fan, or a dishwasher. Naturally, there are limits to how much we can discriminate between the wanted and unwanted sounds, and as the noise floor rises, it becomes increasingly harder to spot individual sounds. In other words, as the environmental sounds we choose to relegate to a background noise increase in loudness and as their cumulative spectra approaches white noise, our capacity to perceive desired sounds diminishes.

8.6.7 Listener Fatigue

One particularly important aspect of sonification is the concern with listener fatigue. The choice between a pleasing and annoying sound may mean the difference between a useful sonification and a pounding headache. Moreover, improper scaling of the data may mean the lack of noticeable difference in the data over the long term, which may lead to a listener simply tuning out such sounds, subconsciously labeling them as background noise, much like the noise of an air vent or a computer fan. To minimize the chance of fatigue, sonification research often resorts to the use of tonal relationships. For instance, mapping data to the diatonic (tonal) pitches of a scale or using data variance to induce previously vetted changes in the rhythm of an engineered sound or a sound layer may be perceived as more appealing than atonal and seemingly random rhythmic sonifications. Similarly, leveraging of sounds we already naturally perceive as soothing even when observed over an extended period of time, such as the sound of a babbling brook, rain, wind chimes, or evening crickets, may be a step towards minimizing the chance of listener fatigue. Other consideration is our greater sensitivity to the higher ranges that we may perceive as being more piercing.

8.6.8 Context Is Everything

Not all target audiences have the same hearing ability. As we age, we tend to lose the ability to perceive higher frequencies. Those who work in noisy environments may lose their hearing ability faster. It is important to consider who the sonification is for and what may be the conditions under which such a sonification may be perceived.

8.6.9 Minimize the Arbitrary

When considering data sources, it may be useful to focus on those with properties that limit the need for their arbitrary assignment. For instance, consider geographical or geospatial data that lends itself to spatialization given such data is inherently spatial. Spatializing such data to reflect its location in space, in turn, limits the need for the sonification to introduce yet another

arbitrarily assigned spatial mapping or a variable that may need to be accounted for in the sonification's assessment and whose potentially varied implementation may adversely affect the human ability to understand the data and its patterns.

8.6.10 Ecological Perception

Lastly, when studying sonification from a scientific perspective, we may want to approach it by studying individual acoustic properties in isolation. In doing so, we need to be aware that, depending on the chosen context and conditions, such an approach can potentially lead to misleading or outright inaccurate conclusions. For instance, one may want to consider using a sine tone, the most basic sound as a means to study the human ability to perceive spatial sound. Yet, our ability to spatialize sound depends on the sound being spectrally rich. In other words, a sound needs to consist of more than one concurrent frequency to offer a decent chance of having its location accurately identified by a listener. As a result, a sine tone offering a single frequency is all but impossible to consistently locate in space and as such, any study that leverages sine tone for this purpose may lead to an inaccurate conclusion that humans are not capable of accurately localizing sounds in space. Another way to describe this is the need to ensure that the sonification offers ecological perception and validity ([De Gelder and Bertelson 2003](#)) that leverages all the affordances of our environment ([Gibson 2014](#)) and thereby sidesteps any potential pitfalls of studying and leveraging acoustic properties in isolation, while exposing and engaging the full perception potential of a perceiver or, more precisely, listener.

8.7 Leveraging the Psychoacoustic Meaning of Environmental Sounds

Throughout this chapter we have revisited examples of environmental sonification, or the unintentional byproducts of our physical world that provide us with invaluable information about

our immediate surroundings, such as the wind and engine noise in our imaginary car ride. We now return to these examples as an inspiration for a sonification that leverages preexisting conditioning we have acquired over the course of our lifetimes to facilitate the understanding of the underlying data. For instance, consider a sonification that represents weather forecast through sound. Rather than relying on an abstract synthesized sound, we may want to use the sound of rain to signify the rain forecast. Moreover, we could use various kinds of rain sounds to suggest the anticipated amount of rain. A relatively silent and steady rainfall sounds significantly different from a violent summer storm, and these are events we can easily recognize provided we have been exposed to them in the past. If we expect wind gusts, we may want to use the wind noise. Or, if we expect both high winds and rain, we could use both. By doing so, we are effectively leveraging the psychoacoustic or conditioned meaning of an environmental sound and thereby facilitating the connection between the data and the sound. In other words, we are eliminating the intermediary step in interpreting the sonification where a listener needs to learn an entirely new vocabulary to be able to connect a potentially abstract aural event with a specific data stream.

Any environmental sound in this kind of sonification can be further manipulated through its psychoacoustic dimensions. This kind of an approach when compared to a more conventional sonification using abstract sounds requires additional considerations. For instance, we may want to use a particular data stream to affect the pitch of the rain sample. Unless executed slowly and subtly, such a change immediately begins to sound unnatural and therefore potentially distracting, or may lead to a sound that loses its original recognizable association and meaning. On one hand, this may be a powerful way to grab one's attention. On the other, it may also break the sense of seamless integration into our everyday environmental soundscape and therefore increase chances of listener fatigue. While altering pitch or detuning may be tricky, filtering and/or changes in loudness are more natural, as they mimic changes in intensity or a separation between the listener and the source by a wall or some other physical and acoustic obstacle.

It is worth noting not all sounds can be portrayed in this way, so it may be necessary to leverage movie or even cartoon-like approaches to sonifying elements that otherwise do not produce perceivable sound. For instance, consider a sunny day. The sun itself does not generate any sound, or at least none that we can directly perceive. It, however, affects other elements in our environment that do produce perceivable sound, such as the crickets and cicadas that try to cool themselves off during a hot afternoon. Or we can resort to human-centric activities that may take place on such a day and that have an implicit relationship with the sun, such as a sound of a hot, sizzling steak on a grill, as was the case with one of the studies I conducted back in late 2000s (Bukvic, Gracanin, and Quek 2008) (Bukvic and Kim 2010)

To date, sonification research has typically shied away from this kind of an approach in favor of more controlled lab-like experiments that focus on synthesized sound, in part because psychoacoustic meaning may have been deemed difficult to quantify and defend within the scientific community. As such, it remains an exciting and woefully underutilized research opportunity.

8.8 A Case for Immersive Exocentric Sonification

Our interaction with the world depends on multiple sensory inputs to create a more comprehensive picture of our environment. Moreover, our experiences are often reinforced when perceived concurrently through multiple senses. For instance, reading a text aloud may help us retain more information than silently reading “in our heads” in part because we are engaging multiple sensory mechanisms and therefore reinforcing learning by concurrently activating multiple areas of our brain. Likewise, consider eating an onion with your nose closed (yes, that came out of the blue). Suddenly, this flavorful vegetable seems tasteless and bland. It is the combination of taste and smell that makes an onion seem not only flavorful but also at times so potent our eyes may tear up.

It appears that studying the way we perceive the world needs to be inherently holistic or, as Gibson put it, ecological ([Gibson 2014](#)). In other words, the ecological perception suggests that none of our senses can reach their full potential on their own—eyes need eye, head and body muscles to move to be able to attain a full picture of our immediate surrounding. Similarly, to localize sounds more accurately we need to move our head and body, therefore improving both our spatial image. Likewise, by positioning ourselves closer to the desired audio source (remember the cocktail party effect?) we can improve the desired signal-versus-background-noise loudness ratio and consequently the sound's clarity and, in case of human voice, intelligibility. Yet, such a holistic approach is not always possible, particularly considering how complex and multivariate such a study may become when factoring in all the senses. Inevitably, the early efforts in studying human perception have gravitated towards a lab-like experiment approach where scientists isolate a particular sense or focus on a small subset of its affordances. This has yielded important information about individual sensory mechanisms and has paved a way towards more holistic approaches to studying perception that have built upon this foundational knowledge. Today, we are witnessing a growing focus on the holistic study of human perception and its impact on various scenarios ([Lindstrom 2005](#)), ([Bukvic and Earle 2018](#)).

Similarly, much of the research in data sonification conducted so far has focused on applying data on one or a few acoustic sound properties in isolation from other sensory inputs. This may be in part due to the fact the sonification is still a nascent field in need of unpacking its multiple dimensions before it can be cross-pollinated with other senses. It may also be due to the technical limitations of the existing audio infrastructure. For instance, as I write this chapter, there are only a handful of facilities around the world that offer access to a large number of loudspeakers, or what we refer to as high density loudspeaker arrays (HDLAs) capable of providing a high-resolution immersive sound environment ([Lyon et al. 2016](#)). As a result, most sonification studies have focused on a subset of psychoacoustic properties. Many have either avoided the use of spatial audio, or have resorted to a simple stereophonic rendition using headphones. Some have tried using the binaural approach that offers greater spatial resolution

but currently suffers from an inability to consistently account for variance in the anatomy of individual pinnae, as well as its inability to accurately portray sounds that emanate in front of a listener. A handful of more recent research projects have tried to integrate hearing with other senses, such as head rotation, or body position, both of which are indispensable variables to accurately locating sound sources. A few studies have dabbled with using HDLAs to avoid the challenges and idiosyncrasies inherent in simulating spatial audio ([Bukvic and Earle 2018](#)).

It is worth noting that the spatial dimension of the human hearing capacity is inextricably intertwined with our other senses, particularly head rotation and position. For instance, to minimize the so-called cone of confusion for sounds to the left and right of our head where there is no pinna to offer filtering variance, or the front-back confusion, we commonly rotate our head to change our perspective to the sound source and use pinnae and interaural time difference (ITD, or a difference in time it takes for a sound to arrive to each ear due to different physical distance between the source and each ear) to better pinpoint a sound location. The inability to do so vastly limits our capacity to locate and consequently discriminate sounds in a more complex soundscape (again, remember the cocktail party effect) and may yield misleading results. Consequently, studies that do not account for this capacity may conclude that our spatial hearing capacity seems lower than it actually is.

Even when studying human sight in isolation from other senses, we do not prevent users from moving their eyes and in most cases their head. These are, after all, critical elements that enhance our ability to see because our eyes only have a very narrow area where we can clearly perceive details and the rest is generally blurry and only sensitive to motion and change. It is our brain that constructs a holistic image as our eyes scan around and whose range is further enhanced by our head motion. It only makes sense that we should also allow for our hearing to be studied in consort with head movement. Moreover, by moving around in space, we are capable of further refining our ability to locate sound sources. Have you ever misplaced a smartphone in your home? As you go around looking for it, you may opt for calling it from another phone and then listening to where its ring may be coming from. The ability to move

throughout the space enables you to use the loudness of the ringtone to your advantage and thereby narrow down the area where it may be located. As you get closer to the smartphone, you have fewer walls and obstacles between you and the phone playing acoustic tricks with sound reflections that may deceive your ability to localize the sound source. This is where you can finally realize the full potential of your spatial aural perception and locate the phone with relative ease.

Coupling head and body movement and position with our hearing perception brings us closer to the way we interact with our environment, something that comes natural to us. We don't have to train for it. We simply interact. Now, consider a walkable HDLA environment with loudspeakers all around you, including above and below. This three-dimensional immersive space populated with numerous aural cues is much like a cocktail party, except instead of being surrounded by people, you are surrounded with multiple sonified data sources. As you move around the space, you are able to get closer to a particular sound, thereby naturally attenuating the perception of other sounds while concurrently using your spatial memory to generate a more comprehensive spatial map. In such an environment, you can interact with the data in the same way you interact with the real world. We can refer to such an environment as the immersive exocentric environment. Consequently, we can refer to the sonification that takes place in it as immersive exocentric sonification (Bakvic and Earle 2018). Unlike the use of the term *exocentric* within the context of immersive environments that suggests an out-of-body experience (Dede 2009), here we refer to a sonification that emanates from a perimeter of the space and does not require the user to remain in the center of the space where the distance and the ensuing loudness between individual loudspeakers and the listener's ears may be near identical.

In our review of hearing perception's unique affordances we briefly touched upon the spatial component. Here we revisit it as a critical element of the aforesaid immersive exocentric sonification. Spatialization algorithms and approaches are indeed numerous and include:

- The ubiquitous stereophonic spatialization that simply utilizes either a pair of loudspeakers or stereo headphones;
- The binaural spatialization that has seen a growing popularity within the gaming and more recently virtual, augmented and mixed reality contexts, in good part due to its limited hardware needs ([Carty and Lazzarini 2009](#));
- Ambisonics that can also scale to a large number of loudspeakers ideally arranged to form a dome-like icosphere or its subset ([Furness 1990](#));
- Wave field synthesis (WFS) that recreates a waverfront using Huygens's principle and typically requires a large number of tightly packed loudspeakers ([Boone et al. 1995](#));
- Vector base amplitude panning (VBAP) that benefits greatly from a large number of loudspeaker arrangements that ideally form a series of triangles, and cumulatively some subset of an icosphere ([Pulkki 1997](#));
- The manifold-interface amplitude panning (MIAP) algorithm that also targets icosphere-like structures ([Seldess 2014](#)); and,
- Most recently, layer based amplitude panning (LBAP) designed in part for the exploration of the immersive exocentric sonification and the spatial mask concept with loudspeakers separated into layers that do not require any particular shape or arrangement ([Bukvic 2016](#)).

These are merely a subset of a much larger ecosystem of algorithms and approaches to sound spatialization that may be utilized within the context of the immersive exocentric sonification and beyond, each of which offers its own unique advantages and limitations. Their proliferation in the twenty-first century points to the anticipated musical emancipation of the spatial component of human hearing we mentioned earlier.

When considering the aforesaid immersive exocentric sonification, it is worth noting that the varying distance between the loudspeakers and a listener as they navigate the space should not be seen as a detriment, but rather as an advantage that more closely resembles the way we interact with the world. There may be phasing artifacts and other unforeseen interactions. Yet, all these anomalies are something we commonly experience in the real world and as such our brains are typically accustomed to either ignoring or leveraging them to improve our spatial resolution. This is also true of visual stimuli. Consider the way we perceive movies on cinema screens that, depending on where we sit, may be so large we cannot fit all of them within our peripheral vision. Under such circumstances, some of the areas on the edge of the large screen are clearly much farther from our eyes than the areas closest to where we may be sitting. This may lead to unusual distortions to the image's perspective with respect to our vantage point. Yet, our brain is capable of treating the ensuing projection as a flat image and we happily reconstruct it as such in our heads without any concern for its physical distortion. By extension, in the immersive exocentric environment, particularly when using a layered approach to grouping loudspeakers inherent to the aforesaid LBAP spatialization algorithm, the loudspeakers within the same layer share the same elevation, much like pixels on a screen that share the same row. As we move around such space to explore the immersive spatial aural content, loudspeakers that are closer to us may have a perceived higher elevation versus those that are farther from us, despite all of them sharing the same elevation with respect to the ground. Despite this anomaly, once we have the overall awareness of the visual (e.g. TV) or aural (e.g. loudspeaker perimeter or front) canvas, our brains are capable of compensating for such physical inconsistencies in favor of maintaining consistent relationships between the individual loudspeakers, or, as is the case with its visual counterpart, between the pixels on a screen.

The immersive exocentric sonification focuses solely on the space's perimeter generated by the loudspeakers. While there are ways of simulating sounds within the space, particularly when utilizing WFS and ambisonics, they are prone to idiosyncrasies that limit the human ability to move and study such sources from different vantage points. As such they may be seen as

being equivalent to studying our ability to visualize data while making certain elements visual only under certain circumstances. If we were to consider the ensuing loudspeaker front or perimeter as one canvas, we could project and move data across it. The most obvious data choices may be inherently spatial data, such as geographical or geospace data that limit the need for arbitrary assignment of variables to the spatial component (Bukvic and Earle 2018). Much like the aforesaid leveraging of the psychoacoustic meaning of environmental sounds, the immersive exocentric sonification is an entirely new opportunity in the sonification research that begs further exploration.

8.9 Where to Go From Here?

In this chapter we have introduced ourselves to the world of sonification, why it exists and how we may benefit from it. We have also reviewed some of the common-sense strategies that have emerged from this nascent field of research. Lastly, we have proposed two new areas of research that may allow us to explore the full potential of sonification, the leveraging of the psychoacoustic meaning of environmental sounds and a more natural immersive exocentric approach to sonification that may help us quantify the limits of the human ability to perceive and interpret spatial audio streams while sidestepping some of the key limitations of the current virtual approaches to spatializing sound. Unsurprisingly, what has been covered in this chapter merely scratches the surface of a much larger body of knowledge. Below is a list of references and suggested follow-on reading that may provide additional insight in the various aspects of sonification and related disciplines.

References

Arons, B., 1992. A review of the cocktail party effect. *Journal of the American Voice I/O Society* 12, no. 7, pp. 35–50.

- Ballora, M., 2014. Sonification, science and popular music: In search of the ‘wow’. *Organised Sound* 19, no. 1, pp. 30–40.
- Boone, M. M., Verheijen, E. N., and Van Tol, P. F., 1995. Spatial sound-field reproduction by wave-field synthesis. *Journal of the Audio Engineering Society* 43, no. 12, pp. 1003–1012.
- Bregman, A. S., 1994. *Auditory Scene Analysis: The Perceptual Organization of Sound* [Online]. MIT Press. Viewed October 25, 2016
<https://books.google.com/books?hl=en&lr=&id=jI8muSpAC5AC&oi=fnd&pg=PR11&dq=Auditory+Scene+Analysis:+The+Perceptual+Organization+of+Sound&ots=SFrZH9ABvD&sig=o5uUgvlJZHz_hdi4c1CQQIfG_qk>.
- Bukvic, I., Gracanin, D., and Quek, F., 2008. Investigating artistic potential of the DREAM interface: the aural painting. Presented at the International Computer Music Conference, Belfast, UK.
- Bukvic, I., and Kim, J., 2010. Perception and Interpretation of concurrent aural shapes using DREAM interface. Presented at the International Conference on Auditory Display, Stony Brook, NY.
- Bukvic, I. I., 2016. 3d time-based aural data representation using d4 library’s layer based amplitude panning algorithm. *Presented at the International Conference on Auditory Display, Canberra, Australia*. Viewed November 9, 2016
<www.icad.org/icad2016/proceedings2/papers/ICAD2016_paper_10.pdf>.
- Bukvic, I. I., and Earle, G., 2018. Reimagining human capacity for location-aware aural pattern recognition: A case for immersive exocentric sonification. *Presented at the International Conference on Auditory Display, Houghton, MI*.
- Carty, B., and Lazzarini, V., 2009. Binaural HRTF based spatialisation: New approaches and implementation. *DAFx 09 Proceedings of the 12th International Conference on Digital Audio Effects, Politecnico Di Milano, Como Campus, Sept. 1–4, Como, Italy*. Department

- of Electronic Engineering, Queen Mary University of London, pp. 1–6. Viewed February 21, 2016 <<http://eprints.maynoothuniversity.ie/2334>>.
- Dede, C., 2009. Immersive interfaces for engagement and learning. *Science* 323, no. 5910, pp. 66–69.
- De Gelder, B., and Bertelson, P., 2003. Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences* 7, no. 10, pp. 460–467.
- Furness, R. K., 1990. Ambisonics-an overview. In: *Audio Engineering Society Conference: 8th International Conference: The Sound of Audio*. Audio Engineering Society. Viewed November 9, 2016 <www.aes.org/e-lib/browse.cfm?elib=5417>.
- Gibson, J. J., 2014. *The Ecological Approach to Visual Perception: Classic Edition*. Psychology Press.
- Kramer, G., et al., 1999. *The Sonification Report: Status of the Field and Research Agenda. Report Prepared for the National Science Foundation by Members of the International Community for Auditory Display*. Santa Fe, NM: International Community for Auditory Display (ICAD).
- Lindstrom, M., 2005. *Brand Sense: How to Build Powerful Brands Through Touch, Taste, Smell, Sight and Sound*. London: Kogan Page Ltd.
- Lyon, E., et al., 2016. Genesis of the cube: The design and deployment of an hdl-a-based performance and research facility. *Computer Music Journal* 40, no. 4, pp. 62–78.
- Oswald, D., 2012. *Non-speech Audio-semiotics: A Review and Revision of Auditory Icon and Earcon Theory*. Georgia Institute of Technology.
- Paul, A. K., 2002. Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction* 12, no. 1, pp. 1–10.
- Pulkki, V., 1997. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society* 45, no. 6, pp. 456–466.
- Rabenhorst, D. A., et al., 1990. *Complementary Visualization and Sonification of Multi-Dimensional Data*. IBM Thomas J. Watson Research Division

- Seldess, Z., 2014. MIAP: Manifold-interface amplitude panning in Max/MSP and pure data. In: *Audio Engineering Society Convention 137*. Audio Engineering Society. Viewed January 31, 2016 <www.aes.org/e-lib/browse.cfm?conv=137&papernum=9112>.
- Varshney, L. R., and Sun, J. Z., 2013. Why do we perceive logarithmically? *Significance* 10, no. 1, pp. 28–31.
- Worrall, D., 2009. *Chapter 2: An Overview of Sonification. Sonification and Information: Concepts, Instruments and Techniques*. PhD Dissertation, University of Canberra, Canberra, Australia.