AI & Soc (2012) 27:213–222
DOI 10.1007/s00146-011-0341-7

ORIGINAL ARTICLE

Aesthetic strategies in sonification

Florian Grond · Thomas Hermann

Received: 18 April 2011 / Accepted: 1 August 2011 / Published online: 30 August 2011

Abstract Sound can be listened to in various ways and with different intentions. Multiple factors influence how and what we perceive when listening to sound. Sonification, the acoustic representation of data, is in essence just sound. It functions as sonification only if we make sure to listen attentively in order to access the abstract information it contains. This is difficult to accomplish since sound always calls the listener's attention to concrete—whether natural or musical—points of references. Important aspects determining how we listen to sonification are discussed in this paper: elicited sounds, repeated sounds, conceptual sounds, technologically mediated sounds, melodic sounds, familiar sounds, multimodal sounds and vocal sounds. We discuss how these aspects help the listener engage with the sound, but also how they can become points of reference in and of themselves. The various sonic qualities employed in sonification can potentially open but also risk closing doors to the accessibility and perceptibility of the sonified data.

Keywords Sonification · Aesthetics · Historic context

1 Introduction

Sonification, today, is an interdisciplinary practice ranging from scientific applications to sound art and composition. Following an early definition by Kramer (1994), it is often understood in practical terms as "representing data with non-speech sound". This characterization, however, is rather coarse and merely addresses challenges in designing, composing or programming sonifications. From a musical perspective, the sonification of data continues from where mimesis, in the tradition of program music and in the indexical function of sound recording, ends. From the perspective of the theory of knowledge and the history of science, we can speak of sonification when sound is used as a medium that represents more than just itself. In other words, sound becomes sonification when it can claim to possess explanatory powers: when it is neither solely music nor serves as mere illustration. Although there are historic and contemporary examples in which sound is the medium used to promote scientific insights, the word 'insight' itself shows how much language metaphors related to knowledge production point to the primacy of vision. We will examine historic examples in which sonic representations or sonic manifestations of natural phenomena were used to arrive at general conclusions, extrapolating and abstracting from the actual phenomena themselves.

The data substrate for sonification today increasingly originates from our virtual or digital surroundings. Sonification therefore often aims to represent something that lacks a natural sonic reference point. It is challenging to find sounds that are at all suitable for these abstract virtual entities, because any perceived sound always refers to earlier sounds that we have experienced in our natural or cultural contexts.

This is a problem shared by scientific as well as artistic approaches to sonification. It is interesting to observe how artists and researchers who employ sonification are careful to maintain or renegotiate links with everyday listening habits and cultural listening experiences.

In this article, we investigate these links to listening contexts and discuss their uses as aesthetic strategies in sonification. The notion of aesthetics to which we refer

F. Grond (✉) · T. Hermann
CITEC Cognitive Interaction Technology Centre of Excellence, Bielefeld University, Universtätsstrasse 21-23, 33615 Bielefed, Germany
e-mail: fgrond@techfak.uni-bielefeld.de

looks at the palette of important properties of sound and sonification scenarios that define how we perceive what we hear. Sonification encompasses many different disciplines ranging from the sciences to the arts, both of which are often linked with technology. Consequently, no traditional genre or canonical form exists for sonification. The loss of canonical categories is similar to the challenges faced when music is extended into sound art, as discussed by Licht (2007). In sound art, each aspect of the work (including those beyond sound itself, such as visual, haptic or performance aspects) influences our perception of it. It thus becomes an aesthetic choice to determine which of these aspects to include, and most of all, how.

This palette of properties of sonifications and sonification scenarios is not restricted to particular disciplinary angles that tie them to the physiological factors of human listening capacities or to auditory gestalt formation from perception theory, for instance. Instead, we have extracted them from sonification practices by looking at historic and contemporary artistic examples as well as established scientific sonification methods. These properties certainly play different roles in the analogic—symbolic continuum of sonification proposed by Kramer (1994): they apply differently to alarm sounds than to long audification streams, for instance. It should also be noted that, as discussed by de Campo (2007), the data structure itself has a significant influence on the form a sonification takes.

The challenge of sound design in sonification is to negotiate between the given constraints of the data and the qualities of sounds and its context. The aesthetic strategies, which are based on the aforementioned palette of properties of sonifications and their scenarios, can be understood as guidelines that critically help to assess whether sonification is an appropriate medium to represent a given data set. In brief, the collection of aesthetic strategies in this article addresses the question:

> *What circumstances enable us to listen to sound as sonification?*

2 Elicited sounds

Everyone involved in sound synthesis knows about the effect of attack, decay, sustain, release (ADSR) envelopes. As amplitude envelopes, they shape the gain of the underlaying sound waves, thereby providing them with a distinct articulation, which reminds the listener of the physical origin of the sound waves. This particularly applies to the first part of the envelope, the attack, which points directly at the action that caused the sound.

Such an articulated sound can for instance be a plucked string, as in the early and often-cited example of a scientific discovery through sound: the monochord. While playing this instrument, Pythagoras of Samos concluded, after analytic listening, that the world must be organized along harmonic ratios. His conclusion, starting from a single string instrument to explain the entire universe, might seem farfetched. But beyond these aspects of discovering harmonious ratios of intervals, there is one thing about the monochord that continues to fascinate us today. It is the childlike joy we experience when plucking a string or moving the bridge underneath. Imagined actions can also connect us with the cause of a sound, they do not need to be performed.

Compared with Pythagoras, Joseph Leopold von Auenbrugger had what we would today call a concrete research question. As a medical doctor, he faced the challenge of assessing internal organs. He did so by developing the diagnostic method of percussion. He attributed his discovery to his boyhood experience of watching his father tap kegs to determine the amount of fluid in them.[1] Inspired by this experience, Auenbrugger tapped his patient's chest as if it were a vessel, then analysed the different sounds it produced. In both these examples, the tight relation between action and perception is important for our engagement with the sound: the sound is clearly anchored to a physical cause. Additionally, this closed loop allows us to correctly interpret the information carried by the impulse response in relation to the impact.

Elicited sounds and the sonic feedback they induce later found analogue uses in echolocation and related sonification metaphors and models. In a data sonogram for instance, a virtual shockwave is sent out into a data cloud. The data points respond through a decaying oscillation in the sequential order of their distance to the epicenter of the shock wave. Although the action itself is virtual, the impulse response follows invariant dynamics inspired by physical laws, see Hermann (2002). These approaches are part of the new developments in sonification research that follow the paradigm of interactive sonification, as introduced by Hermann and Hunt (2005).

A further interesting example from this field is the application *Reim* by Bovermann (2009). In this application, the attack sound produced when hitting a key on the keyboard is picked up by a contact microphone in real time. The sound is then processed through filters, whose parameters are determined by weather forecast data. In *Reim*, the physical origin of the contact sound leads to a particularly intriguing interaction experience.

From the field of mobile phone applications, there is the proof of concept called *shoogle* by Williamson et al. (2007), in which shaking a phone simulates the sound of

[1] For an in depth overview on sonification in medical diagnostics as discussed in Sects. 2 and 3, please confer Baier and Hermann (2008).

pebbles in a box, and each pebble represents an unread SMS in the inbox.

Interactive sonification comes very close to the field of new musical interfaces. As the example of Pythagoras shows, sounds produced by an instrument can be listened to analytically, beyond their qualities as musical expression. This is due to the tight connection between the characteristics of the sound and the action that elicited it. The listener tries to infer the cause of the sound, and the cause amounts to what the sound is referring to.

3 Repeated sounds

Many of the examples in the category above carry another important characteristic of sonification, namely repetition. The perception of repetitions that take on a rhythmic quality is an example in which our listening capacities exceed vision by far. As an aesthetic strategy in sonification, repetition allows us to establish similarity and difference. Repeatability is a necessary condition for scientific experiments. But it is also an aesthetic choice, since through repetition, sonic events begin to refer to each other.

An early example of repetition is given by Rene Theophile Laennec (1781–1826), who continued acoustic diagnostic research after Auenbrugger. Laennec developed the method of auscultation by listening through a stethoscope, which is ironically a medical icon today. Auscultation comes from the latin word *ausculture, careful listening*, and was developed to diagnose heart, lung and bowel sounds. Heart and lungs in particular produce a repetitive sound. While the heartbeat clearly has rhythmic qualities and auditory gestalt formation applies, breath and its length have more similarities with an extended musical phrase, and the focus is rather on sound texture. Auscultation does not induce any sound in the form of impulses, as with percussion. It is interesting to note that the human body can be diagnosed like an instrument, by listening to its resonance patterns as explained in Sect. 2. But it can also be diagnosed through the sounds that it produces autonomously. In both cases, repetition is key to understanding and evaluating the acoustically perceived information.

Another prominent example of sonification in which repetition played an important role was an experimental setup by Galileo Galilei, described by Drake (1980). In this experiment, little bells were fixed along an inclined plane. If their positions along the plane were adjusted so that they rang in equal time intervals, the law of free fall could be demonstrated by measuring the distance of the bells along the plane. This is an early record that links empiricism and theory through a sonic display.

John Eacott's piece *Hour Angle* is a good musical example to illustrate this thought. *Hour Angle* is the real-time conversion of the sun's trace over the surface of the earth. The trace progresses very slowly in time, if we consider the duration of a music piece as the reference time frame. Mimicking the slow progression, the sound is a minimalistic, slowly progressing repetitive structure. The listener experiences this progress not as a big emotional movement, as in Smetana's *Vltava* (1874), but rather as a rational procedure of measuring and counting. The sound serves the same purpose as did the bells in Galilei's inclined plane.

As developed above, repeatability is an important condition for scientific sonification. But from a musical perspective, repetition poses an interesting question: what are the categories that establish identity and difference? If we hear a sound, for instance, we can identify it as a well-known tune, or we can recognize it as a certain performer's interpretation. The sheer fact of repetition immediately gives the listener the possibility to assess.

The repetition of the same or of a similar sound is the simplest way to convince a skeptical listener that no particular musical skills, such as perfect pitch or the recognition of intervals or chords, are required in order to contextualize what was heard. The reference can be found in the repeated sound itself. Repetition suggests that we experience more than just a unique event which, due to its singularity, can fascinate but cannot be understood.

4 Conceptual sounds

A successor in the tradition of Pythagoras is Johannes Kepler. In his book *Harmonices Mundi* (1619), he postulates a relationship between the movement of the planets and musical intervals. From today's perspective, Kepler resolves the struggle of sonification and missing sonic references like a piece of conceptual art. The source of thought is a mathematical equation derived from actual observations, the score is a book and the projected sound resonates in our imagination. Kepler says it openly:

> Ich möchte hierüber auch das Ohr befragen, jedoch so, daß der Verstand aussprechen soll, was natürlicherweise das Ohr zu sagen haben würde. Kepler (1967)[2]

Kepler presents a notion that evokes sonification, despite the fact that it remains inaudible.

[2] Translation from Grond and Schubert-Minski (2009): I would also like to consult the ear on this, though in such a way that the intellect articulates what the ear would naturally have to say.

Another conceptual sonification without sound comes from Rilke (1919). In a 1919 essay, the poet develops thoughts on the similarity between the gramophone and the human skull. He finds that the oscillations of the traces on early LPs are similar to the wiggling of the coronal structures of the skull where the plates are connected. Hence, if one would scratch along this seam with an appropriate needle—just like the one from the gramophone—a sound should be heard. This sound originating from our skull would not be an arbitrary sound but the *Ur-Geräusch* (the *primal sound*), as Rilke called it.

In both the examples of Kepler and Rilke, sound is neither actually produced nor is it at all reproducible; rather, it acts as a metaphor to guide our imagination or, to avoid a visual metaphor, to provoke our inspiration. Sound becomes a vehicle of thought and claims its rightful place next to images and language.

Beyond these examples from science and literature, conceptual aspects can also be found in sonifications which where not conceived as such, but can be thought of as data-inspired music. In these cases, sound is not meant as sonification with the intention of systematically identifying inherent structures in underlying data. Nonetheless, these works cannot be reduced to music, either, since they challenge the notion of composition as a conscious and willful act of arranging artefacts.

Like Kepler, the first example refers to the sky and the stars. It is John Cage's work *Atlas Eclipticalis* (1961). Cage plots celestial maps on a musical score and develops distinct playing instructions out of these. Apart from the degrees of interpretational freedom, the piece uses the inherent structure of the starry sky as its compositional basis. Although the very literal depiction of the natural reference within the score makes one think of contemporary sonifications from data sets, the mapping strategy does not aim at a particular articulation of inherent relations within the data. In *Atlas Eclipticalis*, the role of the underlying data is rather a substrate that poetically questions the motivation of choices and the origin of decisions in musical compositions.

The second example, similar to Rilke, is concerned with the human head and live signals picked up from its surface. It is the 1965 piece *Music for Solo Performer*, an artistic appropriation of EEG measurements by composer Alvin Lucier. In this piece, Lucier amplifies his brain waves and converts them into acoustic signals. Various percussion instruments standing on loudspeakers are caused to vibrate by the strength of the audio signals emitted by the loudspeakers connected to the source. In *Music for Solo performer*, data are not used to make brain activity understandable but rather to emphasize the nature of thoughts and mental processes, which are, like music and sound, movements in time—a fact that contrasts nicely with the physically inactive, sitting performer.

In *Music for solo performer*, the conceptual aspect of sound as a metaphor or medium is reinforced by the fact that we do not hear the sound directly from the speaker but only its reinterpretation through the percussion instrument. This illustrates why the conceptual aspects of sonification are important to be considered. They point at the properties of sound as a medium that is, to a certain extent, not to be confused with the message.

5 Technologically mediated sounds

In some cases, it is not the sound itself that catches our attention, but the technological apparatus that produces or projects it.

The strong influence of technology on our perception can be found as early as measurements of electric potentials were connected to an acoustic display. When the telephone was invented in the late nineteenth century, fantasies of progress were soon projected onto it and it was quickly used as a scientific display.

Let us examine two sonifications from this time, which Dombois (2008) refers to as the first era of sonification. The first example, by Jacques Arsène d'Arsonval, dates back to 1878:

> Le téléphone est un instrument d'une sensibilité exquise. J'ai été amené à le comparer avec le nerf qui était considéré comme le réactive le plus sensible de l'électricité depuis les célèbres expériences de Galvani. d'Arsonval (1878)[3]

D'Arsonval's experiments were apparently successful since he could conclude that even a bad telephone is more sensitive than a nerve.

In 1883, Nikolai Wedenski published *Die telephonischen Wirkungen des erregten Nerven* (The telephonic effect of the excited nerve):

> Das vor kurzem bewiesene Vermögen des Telephons, die negativen Schwankungen des Muskelstromes anzuzeigen, veranlasste mich, zu prüfen, ob möglicherweise das Telephon auch zum Nachweis der galvanischen Wirkungen des erregten Nerven dienen kann. Wedenskii (1883)[4]

[3] Translation by the author: the telephone is an instrument of exquisite sensitivity. This fact led me to compare it with the nerve, which was considered to be the most sensitive reactant for electricity since the famous experiments of Galvani.

[4] Translation by the author The recently demonstrated capacity of the telephone, to indicate the negative oscillations of the muscle current, motivated me to verify if the telephone could possibly serve to indicate the galvanic response of the excited nerve.

Interestingly, the main point of the argument in the opening of the article is the telephone and not the sounding data. Wedenskii then downplays the expectations to make the results look even more convincing. Later in the article, however, Wedenskii elaborates what he heard in detail and explains how it allowed him to distinguish various phenomena of interest. His article is thus the first to put the phenomenology of sonification into the focus of attention.

There are also interesting contemporary examples in which sonification addresses technology first and foremost. An example that explicitly plays with the idea of conceptual sonification is the G-Player (2004) or its portable version, the G-POD (2006), by Jens Brand[5]. This work sonifies the minute distance variations between a satellite and the earth's surface. It literally scratches across the earth's topography. G-POD is not only a sonification but also a performance which Brand opens with the claim: "the earth is a disk!" The audience is led to believe that listening to the earth is as natural as listening to a whole evolution of technological gadgets: walkman, disk-man, iPod. Inspired by technology, the G-POD also has a strong conceptual element, but unlike Kepler's *Harmonices Mundi*, we can actually listen to it.

Valentina Vuksic makes us listen to technology in a unique fashion that differs from the examples above. Her performance *tripping through runtime* mediates the boot processes of a collection of laptops through electromagnetic pickups. This is comparable to the aforementioned examples only insofar as it turns what was unheard before into sound. However, the sensor device, the electromagnetic pickup, is simple and remains in the background. The audience immediately experiences intimacy with the technological object itself. The mediating interface remains unspectacular and does not get in the way when listening to the boot process.

Similar in its technological basis, Christina Kubisch's electric walks allow the audience to explore electromagnetic fields in the environment through receiver coils that are integrated in headphones. In this work, we also hear the technological artifacts that surround us, but compared with *tripping through runtime*, we are more likely to find specific causes for what we hear. The intimate connection is also there but it is mostly created through the experience of hearing sound through headphones.

Technology remains an enabling factor in modern sonification. The works above illustrate that technology is an attractive point of access to the work, sometimes even a fetish. As such, it creates a strong context that influences how we perceive the sound.

[5] http://g-turns.com/.

6 Melodic sounds, cultural aspects

Melody and music are categories that naturally come to mind when talking about sonification. They are also the categories that most closely relate to aesthetics. Pitch and harmonic relations in sound, as discovered by Pythagoras, remain important factors for sonification since they are amongst the most salient, if not the most salient property of sound. They establish structure and hierarchy and allow to easily create distinguishable sonic events. Since these categories are already discussed in depth by Vickers (2005) and Vickers and Hogg (2006), we give only some examples to illustrate their roles for sonification.

The typical uses of auditory icons and earcons illustrate how melodic structures often fill a gap and become a proxy for nonexisting objects. Auditory icons are defined as sounds that naturally occur with a certain action or are readily associated with it. A common example is the sound of paper thrown into the bin, which is played when a file is deleted on the computer desktop. In contrast, earcons are sonic proxies for actions that have no natural sounding reference. In most cases, they are designed with a melodic or harmonic structure. Computer users know these sounds as a tone sequence with a raising melodic structure after a successful login, and a falling one when logging out. The lack of a natural acoustic equivalent to these actions is compensated for through a melodic structure.

Although melodic sounds are popular for earcons, they are not an absolutely necessary choice. Even when they are not melodic, sounds, in many cases, have a more or less pronounced pitch that defines their relation to other sounds. It is difficult to find parameters other than pitch and harmony that allow for the manipulation of sound in such a continuous and wide range.

While harmonic and melodic structures are more related to acoustic gestalt formation and perception, with the musical sounds in sonification, we refer to structural similarities with music and composition. Good examples are the numerous works of DNA sonification. An interesting example from composition can be found by Gena and Strom (1995), and an excellent review of scientific DNA sonifications was compiled by Garcia-Ruiz and Gutierrez-Pulido (2006). One wonders why this data type is particularly appealing for artists and scientists interested in sonification. An answer can be found in the book *Gödel Escher Bach* Hofstadter (1980), in which the process of reading and translating the sequence of DNA was compared with a musical canon for the first time. The idea was that sound which is notated as a sequence establishes many cross-references as it is played. As a result, something greater than the single tones emerges. Similarly, the DNA strand is translated via RNA into an amino acid sequence which folds into a functional complex, an enzyme or protein for instance, which operates on a higher level of functionality than its constituent entities. Here, the sonification of underlying data faces the challenge of arranging information in a time sequence so that it creates a balanced amount of relations and references within the sequence. Broadly speaking, this challenge can be compared with composition.

The cultural dimension of music is also an important factor for sonification, as we find for instance in *Antarktika*[6] by Frank Halbig. *Antarktika* is the sonification of ice-core drilling data from the arctic ice shelf. In this work, the data are translated into a score following a strict and well-chosen set of rules. The resulting score was interpreted in 2006 by the Helios string quartet. *Antarktika* additionally shows a montage of film footage from ongoing antarctic expeditions. This visual dimension of the work has a pronounced documentary character, which contrasts with the classic reception of a string quartet. These two elements, the cultural associations tied to documentary film elements and the well-established musical practice, make the listener oscillate between scientific and musical expectations. In brief, they establish a web of possible references, and hence profoundly influence what we perceive.

7 Familiar sounds

The familiarity of a sound seems to be a vague concept at first, since any sound can become familiar with time. Here, however, we use the notion of familiarity to refer to sonifications that we can relate to without ever having heard their sounds before. If we hear those sounds for the first time, we can appreciate them as something new yet specific, and originating from a yet to be discovered source. The most convincing example of the use of familiar sound comes from the sonification technique of audification in combination with a specific data substrate.

Audification is the most direct conversion of measurements into sound. The data are usually loaded into a sound buffer and directly sent to the digital/analogue converter. A change in playback speed offers the most important degree of freedom for manipulating sound in audification. This is in fact often necessary since many processes which are measured do not happen on a time scale that would lead to perceivable sound signals for the human ear. The direct conversion of data into sound is a good choice if the data of interest exhibit the following properties: first, they have one dimension that can be interpreted as time. Second, the sampling rate along the temporal dimension is sufficiently high to adequately represent the underlying dynamical process. If both requirements are met, then the variations in the data such as dynamical properties like oscillations and transients are often recognizably "translated" into the perceived sound. Further, the result sounds familiar in the sense defined above mostly if the underlying dynamics are of physical origin, meaning that they represent a deterministic dynamical process. In fact, the data recording sensors can in these cases be thought of as microphones, which are sensitive above and mostly below the audible range. In this case, sonification has an indexical function similar to that of sounds from field recordings[7].

If the process is stochastic, however, like stock market data, the resulting sound of an audification is sonically more related to noise. Hence, it might be familiar, but it lacks specificity. In order to shed more light onto this argument, let us examine parameter mapping sonifications. If we map data features to prerecorded sounds of physical instruments, as often used in MIDI-based sonifications, the appeal of the physicality of the sounds points to the instrument, not to the underlying data.

Appealing examples of such familiar sounds are found with the audification of earthquakes in Dombois (2002) and by Abenavoli[8]. Earthquakes are usually inaudible phenomena, apart from the noise of destruction that accompany them. Despite the shocking nature of the event, the physical processes at work are still too slow to be perceived by our ears. Audification takes the digitized data from the seismograph and transfers them into sound by increasing the playback speed by a factor of 1,000 or more. Interestingly, the whole shockwave of the earthquake is well perceivable as such, and different playback speeds lead to different auditory gestalt formations. Less accelerated playback unfolds the sequence of events during the beginning of the shockwave. The beginning exhibits characteristics of an attack and therefore clearly refers to elicited sound as in Sect. 2. Faster playback compresses the whole process and allows to focus on characteristic post-pulse oscillations from aftershocks that are reminiscent to echoes. Many of these phenomena can be encountered in the installation Circum Pacific 5.1, in which Florian Dombois explores earthquakes by means of audification. In this work, 3 weeks' worth of seismological recordings from five stations around the Pacific are presented as audifications of 15 min in a 5.1 surround setup, which mimics the spatial relation of measurement stations. Over extended intervals, the haptic sound quality is reminiscent of the noise of rushing water interrupted more or less regularly by clicks, while the earth's tremors are comparable to the sound of a low-pitched gong. The familiarity of the sound is strongly related to the physical processes at

6 http://www.antarktika.at/.

7 Indexicality in sonification has also been discussed by Vickers and Hogg (2006).

8 http://www.fondation-langlois.org/html/e/page.php?NumPage=268.

work, which have analogous counterparts in our everyday listening experiences.

Dynamic processes causing sounds can sometimes refer to our technological experience, as an example from electromyography (EMG) demonstrates. EMG was sonified as soon as loudspeakers became available. For instance, as noted in the Sect. 5, early telephone loudspeakers were used for EMG sonification. A characteristic myopathic condition leads to an abrupt audible decrease in muscle tension; the amplitude and frequency of the potentials both wax and wane. This change produces a characteristic sound in the audio output of the electromyograph due to the corresponding change in pitch, which has been likened to the sound of a dive bomber[9]. Although a dive bomber sound will hopefully disappear from our acoustic memory 1 day, it works well in the case of EMG since it is a very charged reference. It therefore increases the recognizability of the phenomenon of interest in the sonification stream.

Although our acoustic perception is certainly shaped by evolving in a natural environment, physicality does not suffice to establish the familiarity of a sound. A good counterexample is the audification of electroencephalograms (EEG), where the changing electric potential is measured on the scalp and converted into sound. The basis of this oscillating electric potential is a natural physiological process. Yet, as we cannot directly experience the dynamics of electric potentials, the resulting listening experience remains unfamiliar. Only physical processes, similar to those we are constantly surrounded by, can be said to be familiar. As a consequence, we are accustomed to correctly interpreting their minute variations since, in our experience, they are consistently linked to their source. The establishment of a similarly consistent link between data and sound beyond physicality, as discussed by Hermann and Ritter (2004) and Hermann (2002), was one important motivation for the development of model-based sonification by the second author. This method offers several approaches to consistently connect high-dimensional data to sound rendering processes, even if these data do not have an inherent temporal dimension.

If we look at the physical origin of familiar sounds from a sound synthesis perspective, similar aims are met by physical modelling. The difference, however, is that physical models are usually controlled by a limited set of parameters. In model-based sonification, sounds are influenced by all data points from a potentially high-dimensional embedding data space.

In the example of the dive bomber discharge with EMG, familiarity primarily refers to the ability to recognize a sound as something that has been encountered before. But familiarity more generally implies the identification of a sound with a process rather than with an object. The process itself can turn into a stable acoustic gestalt, and hence, can be recognized. As a consequence, all its variations begin to carry information.

8 Multimodal sounds

In most of the cases mentioned above, sound referred to listening experiences, cultural experiences, familiar dynamical processes or, as in the case of repetition, it was a reference to itself. In many instances, sonification comes together with visualizations. In these cases, it is important to look at how the two modes, sound and vision, interact: how sound refers to vision, and vice versa. There are interesting examples for both the influence of vision to sound and of sound to vision. Both directions demonstrate that our perception constantly attempts to integrate all input streams, thereby bending the "objective" input data from one modality if necessary.

Experiments by Guttman et al. (2005) show that when rhythmical stimuli are involved, the ear overrules what the eye sees. A more familiar example of sound influencing our perception of the visual can be found in cinema, where sound often acts on an emotional level. Movie soundtracks prime our emotions long before the visual scene reveals a surprise, a shock or comic relief. These influences have been extensively researched in the field of film studies, notably by Chion (2008) and Flückiger (2001).

There are two striking examples demonstrating how vision changes the perception of sound. The first is the McGurk effect, named after the author from the often-cited publication *Hearing Lips and Seeing Voices* McGurk and MacDonald (1976). In this case, the visual stimulus of lip movements changes what we hear when our brain has to process conflicting information. For example, the visual stimulus represents a syllable starting with a consonant formed with open lips, like *ta*. However, the sound that is played at the same time is a syllable starting with a consonant formed with closed lips, like *ba*. Yet most subjects report that they perceive the sound *ga*. The second example comes from the use of sound in film. Voices in a dialogue are usually placed in the central loudspeaker, even if the person speaking is not located in the centre of the screen. This misplacement does not disturb the viewing experience; however, because the sound is correctly attributed to the moving lips of the speaker. Both cases demonstrate how tightly audiovisual experiences are coupled. Indeed, they amount to more than the sum of their parts. In the case of sound and vision, the spectator/audience always tries to find or even construct a visual cause for what they hear.

[9] For an overview on electromyography and auditory display see Bonner and Devleschoward (1995) and Walton (1952).

In audiovisual works, the artist is always confronted with the question of whether to use both media in a parallel, mutually illustrative way, or rather as counterpoint, as discussed by Beller (2011). In audiovisual installations with elements of sonification, an interesting tension between these two poles of representation can be found, as in the following artworks:

Brilliant Noise is a contemporary work that employs the sonification of video images in a multimedia installation. The sound consists in the conversion of radio waves emitted by the sun, which are further modified by particular parameters taken from existing satellite video material of the sun. Since the visual dynamic and the sound material are unfamiliar to the audience, the combined audiovisual experience creates ongoing suspense by oscillating between illustration and counterpoint.[10]

The installation *Intermittent*, developed by the first author together with Claudia Robles in 2006, is an audiovisual loop that is edited in real time according to the time intervals given by a dynamical process known as the logistic function. The rhythm that comes from the logistic function exhibits intermittent behaviour, which is an irregular alternation between periodic and chaotic oscillations. Although they come from different origins, both the video and audio sources are subjected to the same rhythm. Spectators inevitably establish a strong link between both, and repeatedly try to find the visual cause for the perceived sounds in spite of the fact that they remain unrelated.[11]

It is difficult to ensure that the audio and visual modes are equally present when creating audiovisual works. As such, they present a particular challenge in the process of creation. In the case of sonification, however, visualization provides the added benefit of creating a context for the sound. This context allows for a certain amount of control in determining what the sound is pointing to and which thoughts it might evoke in the listener. But as shown by examples, it can also lead to unexpected and surprising interpretations when the audience/spectator attempts to integrate both modes.

9 Vocal sounds

Although an early definition sonification was set as the use of non-speech sound to convey information, the appeal of the human voice still holds an important place. An interesting historical reference to voice in a monitoring function can be found in Worrall (2009), where Worral points to the origin of the word auditing, which refers to the practice of comparing two copies of one account. Back in the day, the comparison was made by two people reciting a different copy of the account. Thus, inconsistencies could be identified by listening to differences without being able to read. If we look at speech in auditory displays, we find text-to-speech engines as screen readers for the blind community, for instance. Blind users can speed these engines up to such an extent that sighted users would have a hard time understanding the transported content. A similar concept is used in auditory display through the means of spearcons, as introduced by Walker et al. (2006). Spearcons aim to bridge the gap between the specificity of auditory icons and the generality of earcons by speeding up and altering text-to-speech menu items. Spearcons have the advantage of maintaining a structural similarity between menu items for similar actions, such as *save* and *save as*. Hence, they carry information on both sonic and symbolic levels.

It is known that sonification works better in auditory display if no speech is present at the same time. This is due to the fact that our attention is immediately drawn away from sound and towards speech. The powerful capacity of the human voice to attract attention can, however, be used as an advantage for sonification, since it shows how much our cognitive capacities are tuned into its perception. As a result, we can distinguish even the minutest variations in vocal utterances.

Vowels are an important part of human speech and are also very accessible from a sound synthesis perspective. In order to synthesize the basic set of vowel sounds (a-e-i-o-u), we only need a set of 2 formant filters and the ability to control their position, gain and width. With this, we gain access to a continuous dimension in timbre space that is orthogonal to pitch and loudness. The resulting potential to design sounds is usually less accessible and therefore less systematically explored in the case of the simple mapping of parameters or the playback of prerecorded sounds.

In 2004, Cassidy et al. (2004) explored vowel-based sonification as a means to support the diagnosis of hyperspectral colon tissue images. Vowel-based sonification has also been extensively used for EEG sonification by Hermann et al. (2006). In this case, the dynamic nature of the underlying data with all its characteristic transients applies particularly well to the nature of human articulation. In this sonification, specific traits are first extracted from the EEG and are subsequently used, by means of selective parameter mapping, to control the synthesis of vowel sounds. They thus generate auditory forms that correspond with the traits of the multivariate EEG signals. This is achieved through the variation in two formants, filtered out of white noise or periodic impulses. The sounds vary continuously, especially between the vowels a–e–i, and in their characteristics, such as tone colour and sonority. Since vocal sonifications can be easily replicated by the human voice,

[10] http://www.semiconductorfilms.com.

[11] http://www.grond.at/html/projects/intermittent/intermittent.htm.

they also open the possibility to vocally point out specific elements in a sound stream.

But why should speech-related sounds constitute a category on their own, despite the relatively easy access they provide to timbre space? For our investigation about sonification strategies and sonic points of references, the answer is closer to us than in all the other categories. Vocal sounds point to ourselves and to our capacity to make vocal utterances. Therefore, a vocal sonification does not constitute a technological artifact from the outside world. It finds its origin in everyone who is listening to it.

10 Conclusion

Sound, as we have seen, can be perceived in many ways by referring to many different objects, actions, processes or experiences. This poses a particular challenge for sonification, which most often addresses phenomena that have no sonic references. As such, sound merely offers abstract options for their representation. The nature of this challenge is best illustrated by reconsidering the vocal sound and quoting from Mladen Dolar's *A Voice and Nothing More*. Dolar (2006) Dolar opens the book by quoting Plutarch, who tells the story of a man who plucked a nightingale, and finding but little to eat exclaimed: "You are just a voice and nothing more." The nightingale appears to be a voice without substance. Similarly in vocal sonifications, the voice and not its message becomes the centre of attention. Plucking the nightingale leads to no insight, and in the same way, plucking a string can just make music and nothing more. We can similarly be carried away by rhythm instead of analytically listening to repetitions. The same applies to our fascination with technology, which can be so overwhelming that we forget to pay enough attention to the sound.

Another visual metaphor might be helpful to illustrate this thought. In astronomy, there is the concept of gravitational lenses. A light beam from a hidden star can be bent through the gravity of the star that hides it, so that the deviated light finds its way to the eyes of the observer. With the help of gravity, the observer perceives something that would normally be invisible. The visible star can, however, be a helpful medium and a trap at the same time. If the gravity becomes too strong, then the light is too attracted and cannot escape. The hidden star then remains inexistent for the observer.

All the different aspects of sonification mentioned above risk turning into an object of reference that traps the information carried in the sound. A sonification that works is therefore the successful struggle to create a message that points beyond the medium.

There is often a dispute about whether sonification is a scientific or an artistic practice. For science, one can say that sonification has to follow aesthetic considerations, as discussed above. But if we think of sonification only as art, we unnecessarily reduce its potential. Sonification can only succeed as a cutting-edge practice that transcends either discipline.

Acknowledgements Some historic aspects mentioned in this article have been published by the first author together with Theresa Schubert-Minski in the book see this sound Grond and Schubert-Minski (2009). The authors want to acknowledge further fruitful discussions with the participants of the *Science by Ear* symposium February 2010 in Graz. Finally, these thoughts were evaluated and compiled after integrating the important feedback from the sonification symposium March 2010 in Aix-en-Provence.

References

Baier G, Hermann T (2008) Temporal perspective from auditory perception. In: Vrobel S, Rössler OE, Marks-Tarlow T (eds) Simultaneity: temporal structures and observer perspectives, Chap. 21. World Scientific, Singapore, pp 348–363

Beller H (2011) Between the poles of mickey mousing and counterpoint. In: Daniels D, Naumann S (eds) Audiovisuology 2 essays. pp 102–119 Walther König, Köln

Bonner F, Devleschoward A (1995) AAEM minimonograph #45: the early development of electromyography. Muscle Nerve 18:825–833

Bovermann T (2009) Tangible auditory interfaces: combining auditory displays and tangible interfaces. PhD thesis, Faculty of Technology, Bielefeld University

Cassidy RJ, Berger J, Lee K, Maggioni M, Coifman RR (2004) Auditory display of hyperspectral colon tissue images using vocal synthesis models. In: Barrass S, Vickers P (eds) Proceedings of the 10th international conference on auditory display (ICAD2004), Sydney, Australia, http://www.icad.org/Proceedings/2004/CassidyBerger2004.pdf

Chion M (2008) Le phrase audio visuelle/die audiovisuelle phrasierung. In: Meyer PM (eds) Acoustic turn. Wilhelm Fink Verlag, München

d'Arsonval JA (1878) Telephone employe comme galvanoscope. Comptes rendus hebdomadaires des seances de l'Academie des sciences 86, Paris, pp 832–833

de Campo A (2007) Toward a data sonification design space map. In: Scavone GP (ed) Proceedings of the 13th international conference on auditory display, Schulich School of Music, McGill University, Schulich School of Music, McGill University, Montreal, Canada, pp 342–347, http://www.icad.org/Proceedings/2007/deCampo2007.pdf

Dolar M (2006) A voice and nothing more. MIT Press, Cambridge, MA

Dombois F (2002) Auditory seismology on free oscillations, focal mechanisms, explosions and synthetic seismograms. In: Nakatsu R, Kawahara H (eds) proceedings of the 8th international conference on auditory display, advanced telecommunications research institute (ATR), Kyoto, Japan, ICAD, Kyoto, Japan, http://www.icad.org/Proceedings/2002/Dombois2002.pdf

Dombois F (2008) Sonification, ein plädoyer, dem naturwissenschaftlichen verfahren eine kulturhistorische einschätzung zukommen zu lassen. In: Meyer PM (eds) Acoustic turn. Wilhelm Fink Verlag, München

Drake S (1980) Galileo. Oxford University Press, New York

Flückiger B (2001) Sound design die virtuelle Klangwelt des films. Schüren, Marburg

Garcia-Ruiz M, Gutierrez-Pulido J (2006) An overview of auditory display to assist comprehension of molecular information. Interact Comput 18(4):853–868

Gena P, Strom C (1995) Musical synthesis of DNA sequences. In: XI Colloquio di Informatica Musicale, Univeristà di Bologna, pp 203–204

Grond F, Schubert-Minski T (2009) Sonification, scientific method and artistic practice. In: Daniels D, Naumann S (eds) Audiovisuology compendium, Verlag der Buchhandlung Walter König, Köln, vol 1, pp 285–295, http://beta.see-this-sound.at/kompendium/abstract/70

Guttman S, Gilroy LA, Blake R (2005) Hearing what the eyes see. Psychol Sci 16(3):228–235

Hermann T (2002) Sonification for exploratory data analysis. PhD thesis, Bielefeld University, Bielefeld, Germany

Hermann T, Hunt A (2005) An introduction to interactive sonification (guest editors' introduction). IEEE MultiMed 12(2):20–24

Hermann T, Ritter H (2004) Sound and meaning in auditory data display. Proceedings of the IEEE (special issue on engineering and music–supervisory control and auditory communication) 92(4):730–741

Hermann T, Baier G, Stephani U, Ritter H (2006) Vocal sonification of pathologic EEG features. In: Stockman T (ed) Proceedings of the 12th international conference on auditory display. International Community for Auditory Display (ICAD), Department of Computer Science, Queen Mary, University of London UK, London, UK, pp 158–163

Hofstadter DR (1980) Gödel, escher, bach, an eternal golden braid. Vintage Books, New York

Kepler J (1967) Weltharmonik. Max Caspar, Munich

Kramer G (1994) An introduction to auditory display. In: Kramer G (eds) In auditory display. pp 1–79 Addison-Wesley, Boston, MA

Licht A (2007) Sound art, beyond music, between categories. Rizzoli International Publications, Inc., New York, NY

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nat Biotechnol 264:746–748

Rilke RM (1919) Ur-Geräusch, soglio, am Tage Mariae Himmelfahrt 1919

Vickers P (2005) Ars informatica-ars electronica: improving sonification aesthetics. In: HCI2005: workshop on understanding & designing for aesthetic experience

Vickers P, Hogg B (2006) Sonification abstraite/sonification concrète: an 'æsthetic perspective space' for classifying auditory displays in the ars musica domain. In: Proceedings of the 12th international conference on auditory display, London, UK, 20–23 June 2006

Walker BN, Nance A, Lindsay J (2006) Spearcons: speech-based earcons improve navigation performance in auditory menus. In: Proceedings of the international conference on auditory display (ICAD 2006), London, pp 63–68

Walton J (1952) The electromyogram in myopathy: analysis with the audio-frequency spectrometer. J Neurol Neurosurg Psychiatry 15:219–226

Wedenskii N (1883) Die telephonischen Wirkungen des erregten Nerven. Centralblatt für die medicinischen Wissenschaften 26:465–468

Williamson J, Smith RM, Hughes S (2007) Shoogle: excitatory multimodal interaction on mobile devices. In: CHI 2007 Proceedings. Shake, Rattle & Roll: New Forms of Input & Output, San Jose, CA, USA

Worrall D (2009) Sonification and information-concepts, instruments and techniques. PhD thesis, University of Canberra