Sound perception and design in multimodal environments

Thesis · December 2015

DOI: 10.13140/RG.2.1.3849.6728

1 author:

Permagnus Lindborg
City University of Hong Kong

Sound perception and design in multimodal environments

PERMAGNUS LINDBORG

DOCTORAL THESIS IN SPEECH AND MUSIC COMMUNICATION

Stockholm, Sweden 2015

TRITA-CSC-A-2015:17
ISSN-1653-5723
ISRN-KTH/CSC/A-15/17-SE
ISBN 978-91-7595-771-5

KTH Royal Institute of Technology
School of Computer Science and Communication
Department of Speech, Music, and Hearing (TMH)
Lindstedtsvägen 24, SE-100 44 Stockholm, Sweden

Akademisk avhandling som med tillstånd av Kungl. Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen fredagen den 11 december 2015 klockan 10:00 i Kollegiesalen, KTH, Brinellvägen 8, Stockholm.

Tryck: Universitetsservice US-AB

To Joyce

Abstract

This dissertation is about sound in context. Since sensory processing is inherently multimodal, research in sound is necessarily multidisciplinary. The present work has been guided by principles of systematicity, ecological validity, complementarity of methods, and integration of science and art. The main tools to investigate the mediating relationship of people and environment through sound have been empiricism and psychophysics.

Four papers focus on perception. In paper A, urban soundscapes were reproduced in a 3D installation. Analysis of results from an experiment revealed correlations between acoustic features and physiological indicators of stress and relaxation. Paper B evaluated soundscapes of different type. Perceived quality was predicted not only by psychoacoustic descriptors but also personality traits. Sound reproduction quality was manipulated in paper D, causing two effects on source localisation which were explained by spatial and semantic crossmodal correspondences. Crossmodal correspondence was central in paper C, a study of colour association with music. A response interface employing *CIE Lab* colour space, a novelty in music emotion research, was developed. A mixed method approach supported an emotion mediation hypothesis, evidenced in regression models and participant interviews.

Three papers focus on design. Field surveys and acoustic measurements were carried out in restaurants. Paper E charted relations between acoustic, physical, and perceptual features, focussing on designable elements and materials. This investigation was pursued in Paper F where a taxonomy of sound sources was developed. Analysis of questionnaire data revealed perceptual and crossmodal effects. Lastly, paper G discussed how crossmodal correspondences facilitated creation of meaning in music by infusing ecologically founded sonification parameters with visual and spatial metaphors.

The seven papers constitute an investigation into how sound affects us, and what sound means to us.

Keywords: sound, perception, design, multimodal, environment, soundscape, music, listening, crossmodal, psychoacoustics, psychophysiology, personality trait, emotion, appraisal, aesthetics, colour, correlation, regression, classification, sonification

Sammanfattning

Denna doktorsavhandling handlar om ljud i sammanhang. Eftersom informationsbehandling genom sinnena alltid är multimodal så kräver ljudforskning en tvärvetenskaplig forskningsansats. Arbetet i denna avhandling har vägletts av principer såsom systematik, ekologisk validitet, samspel mellan metoder, och integration av vetenskap och konst. De viktigaste redskapen för att undersöka den ömsesidiga påverkan mellan människa och miljö genom ljud har varit empiri och psykofysik.

Fyra artiklar handlar om perception. I artikel A återskapades urbana ljudlandskap i en 3D-ljudinstallation. Analys av experimentresultat avslöjade samband mellan akustiska mått och fysiologiska markörer av stress och avslappning. Artikel B utvärderade olika typer av ljudlandskap. Upplevd kvalitet kunde prediceras inte bara av psykoakustiska mått utan även av personlighetsdrag. Ljudåtergivningskvalitet manipulerades i artikel D och orsakade två effekter på lokalisering av en ljudkälla vilka förklarades av rumslig och semantisk korsmodala kopplingar. Korsmodalitet var huvudpunkten i artikel C, en studie av färgassociation till musik. Ett användargränssnitt utvecklades som använder färgrymden *CIE Lab*, en nyhet i forskningfältet musik och känslor. En abduktiv metod stödde hypotesen att känslouttryck medierar korsmodala kopplingar, vilket framgick av regressionsmodeller och intervjuer med försökspersonerna.

Tre artiklar handlar om design. Fältundersökningar och ljudmätningar utfördes i restauranger. Artikel E kartlade samband mellan akustiska, fysiska och perceptuella särdrag, med fokus på formbara element och material. Detta arbete fortsattes i artikel F varigenom en taxonomi av ljudkällor utvecklades. Analys av enkätdata avslöjade perceptuella och korsmodala effekter. Slutligen, artikel G diskuterade hur korsmodala kopplingar främjade meningsskapande i musik genom att ekologiskt motiverade sonifikationsparametrar samverkade med visuella och spatiala uttryck.

De sju artiklarna utgör landmärken i avhandlingens utforskande av hur ljud påverkar oss, och vad ljud betyder för oss.

Acknowledgements

Many people have influenced the work in this thesis. I have been helped along by someone showing an interest, sparking an idea, posing a question, giving an answer. I am grateful for every one of numerous moments of joyful insight.

A few individuals have been profoundly important to me. Without their active contribution and sustained support, I cannot see how these pages, dear reader, could now be before your eyes.

I thank my supervisor Anders Friberg, for piloting my traverse between the Scylla of scope and Charybdis of detail;

I am indebted to Ser Wee, Vibeke Sorensen, and Louis-Philippe Demers who made it possible for me to pursue doctoral studies while working at ADM/NTU;

I thank Adriaan Du Toit, Yong Rong Zhao, Vladimir Todorovic, Martin Reiser, and Lonce Wyse, for being great friends and colleagues;

My thoughts go to morfar Olov Alvfeldt and Leon Milo whose inspiration was fundamental;

Thanks to my Singaporean family, Mama, Dajie, Tom, everyone;

Tack Mamma och Pappa, ni är två solvarma klippor, Katarina, Anders, Ami, Jon, Jonas, hela släkten;

Tack Älvi Weiyu, för meningen;

Thank you Joyce, for just about everything.

PerMagnus Lindborg

Singapore & Stockholm, November 2015

www.permagnus.org

Papers included in the thesis

This thesis is the original work of the candidate except for commonly understood and accepted ideas or where references have been made. It consists of seven papers, and an introduction ('kappa'). The principal contribution to each papers was made by the candidate. Coauthors' contributions are stated below.

Paper A

Lindborg PM (2013, June). "Physiological measures regress onto acoustic and perceptual features of soundscapes". *Proceedings of the 3rd International Conference on Music Emotion (ICME3)*. Jyväskylä, Finland, 11th - 15th June 2013. Geoff Luck & Olivier Brabant (Eds.). University of Jyväskylä, Department of Music. ISBN 978-951-39-5250-1. URL https://jyx.jyu.fi/dspace/handle/123456789/41614 (9 Oct. 2015).

Paper B

Lindborg PM & Friberg AK (2015, submitted). "Personality traits influence perception of soundscape quality". *Environment and Behavior*.

Anders Friberg contributed to manuscript authoring.

Paper C

Lindborg PM & Friberg AK (2015, accepted). "Colour Association with Music is Mediated by Emotion: Evidence from an Experiment using a CIE Lab Interface and Interviews". *PLOSone*.

Anders Friberg contributed to data analysis and manuscript authoring.

Paper D

Lindborg PM & Kwan NAK (2015, May). "Audio Quality Moderates Localisation Accuracy: Two Distinct Perceptual Effects?". *Proc. 138th Convention of the Audio Engineering Society*, May 2015, Warsaw, Poland. AES #9313. DOI: 10.13140/RG.2.1.3587.8565. URL http://www.aes.org/e-lib/browse.cfm?elib=17737.

Nicholas Kwan contributed to experiment design, conducted the experiment, and contributed to manuscript authoring.

Paper E

Lindborg PM (2015, March).“Psychoacoustic, Physical, and Perceptual Features of Restaurants: A Field Survey in Singapore”. *Applied Acoustic*s 92, 47-60. DOI: 10.1016/j.apacoust.2015.01.002. URL http://www.sciencedirect.com/science/article/pii/S0003682X15000031.

Paper F

Lindborg PM (2015, submitted). “A taxonomy of sound sources in restaurants”. *Applied Acoustics*.

Paper G

Lindborg PM (2015, accepted). “Interactive Sonification of Weather Data for The Locust Wrath, a Multimedia Dance Performance”. *Leonardo*, MIT Press.

Other works

Other publications and artworks that have informed the research.

Papers (selection)

Lindborg PM & Styles S (2015, forthcoming). “Editorial". *Proc. 2nd International Symposium on Sound & Interactivity (Si15). ICMA Array.*

Lindborg PM & Liu DY (2015, July). “Locust Wrath: an iOS Audience Participatory Auditory Display”. *Proc. 21th International Conference on Auditory Display (ICAD).* Vogt K, Andreopoulou A & Goudarzi V (eds.). Institute of Electronic Music and Acoustics (IEM), University of Music and Performing Arts Graz (KUG), 8-10 July, Graz, Austria. ISBN 978-3-902949-01-1. URL http://iem.kug.ac.at/fileadmin/media/institut-17/proceedings150707_3.pdf, p. 125-132.

Lindborg PM (2014, Oct.). “Editorial - Special Issue on Sound Art and Interactivity in Singapore: SI13 and More”. *eContact! 16.2*. URL http://cec.sonus.ca/econtact/16_2/.

Lindborg PM (2014, Oct.). “Sound Art Singapore: Conversation with Pete Kellock, Zul Mahmod and Mark Wong”. *eContact! 16.2* (Lindborg, ed.). URL http://cec.sonus.ca/econtact/16_2/lindborg_singapore.html .

Lindborg PM & Lim, M.J.Y. (2013, Jul.). "Design of an Interactive Earphone Simulator and Results from a Perceptual Experiment". *Proc. Sound and Music Computing Conference* 2013 *(SMC)*, Stockholm, Sweden. July 2013, ISBN 978-91-7501-831-7. Logos Verlag, Berlin. URL http://www.speech.kth.se/smac-smc-2013/, p. 74-79.

Aw M, Lim CS & Lindborg PM (2013, Jul.). “SmartDJ, An Interactive Music Player for Music Discovery by Similarity Comparison”. *Proc. Sound and Music Computing Conference (SMC)*. Stockholm, Sweden. July 2013. URL http://www.speech.kth.se/smac-smc-2013/, p. 776-781.

Lindborg PM (2012). “Correlations Between Acoustic Features, Personality Traits and Perception of Soundscapes”. *Proc. Joint Conference 12th International Conference on Music Perception and Cognition & 8th Triannial Conference of the European Society for the Cognitive Sciences of Music (ICMPC-ESCOM)*, Thessaloniki, Greece. URL http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/594_Proc.pdf

Lindborg PM (2012). “How fun is this? A pilot questionnaire study to investigate visitors’ experience of an interactive sound installation” (poster). *Proc. Joint Conference 12th International Conference on Music Perception and Cognition & 8th Triannial Conference of the European Society for the Cognitive Sciences of Music (ICMPC-ESCOM),* Thessaloniki, Greece.

Lindborg PM & Koh JBT (2011, Jul.). “Multidimensional spatial sound design for ‘On the String’”. *Proc. International Conference of Computer Music (ICMC)*. University of Huddersfield, UK. July 2011. URL http://hdl.handle.net/2027/spo.bbp2372.2011.014.

Lindborg PM (2010. Dec.). "Singapore Voices: an interactive installation about languages to (re)(dis)cover the intergenerational distance". *IM: Interactive Media*. Issue 6: Special issue on ‘Performance’. National Academy of Screen and Sound (NASS), Australia. ISSN 1833-0533. URL http://wwwmcc.murdoch.edu.au/nass/issue6/pdf/nass-im6-performance-article-04-lindborg.pdf.

Lindborg PM (2010, Aug.). "Perception of emotion portrayal in cartoons by visually and aurally oriented people". *Proc. 11th International Conference on Music Perception and Cognition (ICMPC11)*. Seattle, Washington, USA. (Demorest, Morrison, Campbell eds.). ISBN 1-876346-62-0.

Artworks (selection)

Sonification sound installation and physical sculpture

Lindborg PM (2015, Mar.). *LW24, sculptural 24-channel auditory display. Mixed materials*. In *Beyond the Horizon* (Sabapathy & Lim, curators). ADM, NTU, Singapore. 19 March - 15 April 2015. Also exhibited at the National Gallery of Singapore, 1-10 December 2015.

Lindborg PM & Liu DY (2015, Jan.). “Locust Wrath #3.” In Liong, Koh, Lindborg et al. *Make it New: Future Feed*. National Design Centre, Singapore. 7-8 February 2015.

Lindborg PM (2014, Sep.). *Locust Wrath #2. Six-channel frontal auditory display with sonification of weather data 2015-55 covering the Mediterranean Basin*. Music program of the joint International Computer Music and Sound & Music Computing Conferences (ICMC-SMC). Onassis Cultural Centre, Athens, Greece. 21-26 September 2014.

Lindborg PM (2013, Sep.). “Locust Wrath. Spatialised sonifications of meteorological data.” In Liong A, Koh JBT, Lindborg PM & ArtsFission Dance Company. *The Locust Wrath*. HeloTrans Gallery, Singapore. 27-28 September 2013. (http://www.dansing.org.sg/2013/09/12/the-locust-wrath-by-the-arts-fission-company/)

Interactive /Generative sound installation and physical sculpture

Koh JBT, Lindborg PM, Pertout A, Shin S, Fasciani S & Stromberg D (2015, Aug.). *When We Collide. Generative quadraphonic and collaborative sound installation*. Soundislands Festival (Si15), The ArtsHouse and ArtScience Museum, 18-23 August 2015, Singapore.

Lindborg PM, Koh JBT & Yong RZ. (2013, Sep.) *The Canopy v2. Performances and installation with interactive sound scultpure with 3D*

spatial diffusion. *Mixed materials*. World Stage Design Festival (WSD), Cardiff, 6-9 September 2013. (http://www.wsd2013.com/).

Lindborg PM (2013, Aug.). *Graviton Semantic*. *Interactive electroacoustic soundscape composition with 3D spatial diffusion*. Music program of the International Computer Music Conference (ICMC). Perth, Australia. 11-17 August 2013. (http://icmc2013.com.au/program/installations-2/).

Lindborg PM, Koh JBT & Yong RZ. (2011, Jul.) *The Canopy*. Interactive electroacoustic sound sculpture with 3D spatial diffusion. *Mixed materials*. Music program of the International Computer Music Conference (ICMC), University of Hudderfield, UK, 31 July - 5 August 2011.

Koh JBT, Lindborg PM, Fu M & Tan F (2011). *On the String, DVD*. Commissioned by National Library, Singapore.

Lindborg PM (2010, Jun.). *Graviton Dance*. *Generative electroacoustic soundscape composition with 3D spatial diffusion*. In Koh JBT, Lindborg PM, Chian KH et al. (2010). *On the String, theatre of music*. Multimedia performance commissioned by National Arts Council for Singapore Arts Festival. Esplanade Theatre, Singapore, 4-5 June 2010.

Site-specific collaborative sound installation

Hausswolff CM, Lindborg PM, Grönlund T, Nisunen P, Harding M, The Sons of God, Kirkegaard J, LaBelle B, Nilsen BJ, Petursson F, Pomassl F, Ceeh A, Ödlund C, Thirwell J, Urstad M & Winderen J. (2015, Feb.). *freq_out 1.2 ∞, LP*. Ash International 12.0. London, UK.

Hausswolff CM, Lindborg PM et al. (2014, Nov.). *Freq-Out 1.2 ∞*. *Permanent site-specific sound installation*. Skandion Clinic, Uppsala, Sweden. Since 11 Nov. 2014. (http://freq-out.org/infinity/).

Hausswolff CM, Lindborg PM et al. (2013, Feb.). *Freq-out 9*. *Site-specific collaborative sound installation*. Stedeljik Museum, Amsterdam, the Netherlands. Sonic Acts Festival / Dark Universe, 23–25 Feb. 2013. (http://freq-out.org/9/).

Hausswolff CM, Lindborg PM et al. (2012, Feb.). *Freq-out 8*. *Site-specific collaborative sound installation*. Museum of Modern Art, Stockholm, Sweden. 13-19 Feb. 2012. (http://www.modernamuseet.se/en/Stockholm/Exhibitions/2012/freq_out8, http://freq-out.org/8/).

Introduction

Content

Overview

The thesis Introduction, structured in six chapters, aims to situate the papers included in the thesis within their broader context. Several research themes, concepts, and definitions relevant to the papers are discussed. The chapters have been written from the 'outside in' and are not based on material already included in the papers, each of which has a background or review section pertaining to the topic. The style is more general in character than in the papers and relies to a higher degree on citations and referenced reformulations of authority. The goal of the Introduction is to synthesise relevant aspects of psychoacoustics with research in music and soundscape perception.

- Chapter 1 presents the central research questions and discusses the guiding principles of the work.
- Chapter 2 gives a brief introduction to each of the seven included papers. Highlights, a summary, and main contributions are stated.
- Chapter 3 reviews aspects of sound perception which have direct relevance to five of the papers. Papers A and B investigate physiological responses and individual differences in soundscape perception. Paper C deals with crossmodal correspondences between music and visual colour. While paper D focusses on source localisation, paper F is about source appraisal.
- Chapter 4 discusses sound design, in particular the requirements of practice-oriented acoustic design. Papers E and F investigate restaurants from this aspect, presenting a typology of acoustic design elements and a taxonomy of sound sources. Paper G describes an original artwork with a sound design based on principles of crossmodal correspondences between music, imagery, and spatiality.
- Chapter 5 considers a way to classify multimodal environments based on acoustic communication type, leading to a framework with two conceptual dimensions, *Plasticity* and sensorial *Complexity*. Multimodal environments can be interpreted with the help of these constructs.
- Chapter 6 concludes the Introduction and outlines future work.

Introduction

Effective soundscape research scarcely begins before the scientific community acquires firm answers to questions like the following: What are the fundamental entities of which the sonic environment is composed? How do these interact with each other and with the senses? What questions may legitimately be asked about sound and what techniques employed in seeking solutions? (Kuhn 1962, p. 4-5, paraphrased)

1. Background

The sonic environment is in constant flux. People, machines, animals, and nature itself: they all make noises, sometimes in order to communicate, and often as a by-product of some other activity, some other need. As each one of us moves through everyday environments in daily life, our listening perspective changes, and thereby our expectation and appraisal of the soundscape. With little or no effort, we can recognise and recall places we have visited, sometimes years ago, triggered by sound alone. What we think of as a 'place' always has a sound attached to it.

This doctoral thesis connects areas of understanding related to sound in context, specifically: sound perception and design in multimodal environments. Sound perception refers to ways in which we take in, listen to, and reflect upon sound. Sound design refers to ways in which we describe, classify, and shape sound. Multimodal environments are all the places that we can visit; and they all have sound.

Research questions

The fundamental question addressed in this thesis is how people relate to sound in context. The answers have been sought by investigating sound as a designable material and people's responses to sound in nature, servicescapes, and in performing arts. Each of the seven papers in the thesis centres around a question, which when taken together form the backbone of an inquiry into what sound *means* to us.

- Paper A: *How is the body affected by the sonic environment?*
- Paper B: *Can personality explain how people perceive soundscapes?*
- Paper C: *How does perception of music relate to colour?*
- Paper D: *Does audio quality affect the perceived auditory space?*

- Paper E: *Which design features are important for restaurant soundscape quality?*
- Paper F: *Which sounds are liked or disliked in restaurants?*
- Paper G: *Can sound be designed so that its meaning is understood intuitively?*

Similar questions are being voiced by many researchers and artists. Some have been investigated for a long time, yet in some cases, finding satisfactory answers have proven elusive. Researchers and artists might place their focus on music, soundscapes, or fundamental audition; sometimes with specific probes, and at other times with general inquiries. The questions addressed in this thesis sample all these categories.

The driving force behind my research has been an urge to understand what makes people *tick*. Why do we sometimes experience a sense of place when interacting with an environment? How can we judge the aesthetic qualities of soundscape, or of music? What is the essence of sound: how and why does it affect us?

Science and art

The motto of KTH – *Vetenskap och konst* – might be interpreted as *Science and Art* in different ways. The Swedish word *veta* means 'to know' and *konst* means 'art'. However, the latter stems from *kunna*, which means 'to be able to' (Hellquist 1922, p. 341). Considering that KTH is a technical university, an interpretation of the motto would be "to know and know-how".

The contributions of this thesis intersect science and art. Science has an idealistic side, which might be thought of as 'philosophy' or 'blue skies research' (depending on perspective), and a utilitarian side, which might imply 'development' or 'entrepreneurship' (again, depending on who is speaking). Likewise, art has an idealistic side, as in 'art for art's sake', and a utilitarian side, i.e. 'craft' or 'entertainment'. The situation is not straightforward as the perspectives are intertwined and interdependent; in reality, they represent the extremes encompassing a continuum. Fig. 1 illustrates their relationship as the vertices of a tetrahedron.

The papers included in the thesis traverse the faces, edges, and interior of this tetrahedron, drawing from various fields of knowledge. From science: psychoacoustics, acoustics, acoustic ecology, and music emotion. From art: sonification, sound installation, and multimedia performance. Transfer of knowledge between the fields was enabled by sharing know-how of technologies and methods, including practical technologies such as audio engineering, recording, and synthesis (to generate material for experiments and artworks), and analytical methods such as statistics and stochastics (to understand and create structured experiences).

The multifaceted topic, which among other things covers a wide range of physical environments, from nature and urban soundscape via servicescape to multimedia performance and installation, has necessitated a variety of work. The papers include studies on perception of sound through physiological responses, quality appraisal, sound characterisation, and crossmodal association. In addition, a substantial part of engineering work and knowledge has been required for the experiments, including design of sound through classification of sources, acoustic elements, and interior design materials, combined with application of 'know-how' of sound through analysis, interface design, programming, recording, and diffusion.

Guiding principles

Four principles of 'know and know-how' that link science and art can be discerned. These principles have guided the work presented in this thesis.

Knowledge grows from systematic observation

Scientific knowledge is traditionally characterised by three qualities: theory, truthfulness, and systematicity (Bech & Zacharov 2006, p. 18-20). Science does not solve practical problems per se, but it may originate from practice, or be used for practical purposes. It does not rely on belief, and demands thorough documentation so as to be reproducible. Theory is a scaffold for scientific work. Truthfulness demands an active interest in phenomena of the world. The value of empirical observation is measured against systemised previous knowledge. Systematicity relies on the labour of making observations and conducting experiments.

But these are not principles reserved for scientists; they apply to any agent, from earthworm to composer. In turn, they depend on more fundamental behavioural principles, such as sensory identification of phenomena, memory, and expectancy. Truthfulness and systematicity enable growing knowledge by the method of induction. Induction is the logical reasoning that takes us from the particular to the general, and eventually, to theory. It is the basis of empiricism.

Ecological validity promotes relevance

Much psychoacoustic research uses perceptually basic stimuli such as sinusoidal waves and noise bursts (Moore 2012; Pulkki & Karjalainen 2015). This applies to paper D as well. Similar sounds are the concrete material in the artwork presented in paper G and in other works (e.g. Lindborg, Koh & Yong 2013; Lindborg 2013; Hausswolff, Lindborg et al. 2012).

A very different class of sounds, namely real-world soundscapes, have been employed as stimuli in a series of experiments and field surveys in papers A, B, C, and D (also Lindborg & Lim 2013) for investigations of how people relate to sound in multimodal environments. Maintaining a high degree of ecological validity has been important in the design of these empirical perceptual experiments (cf. Gaver 1993; Friberg 2012).

Methods are complementary

Research methods are not “either-or”. They are designable like clay and take on numerous guises: systematic procedures within a continuum of possible actions (for an overview, see Pulkki & Karjalainen 2015 ch. 17 and 8). By tradition, methods are referred to as ‘objective’ when they rely primarily on machines, and ‘subjective’ when they mainly depend on humans.[1]

In this project, knowledge about sound perception was gained through studying the interaction between hearing and other sensory modalities, and through comparing subjective and objective measures.

‘Machine’ measurement of sound pressure level (Kinsler et al. 1999, ch. 12 and 13) was fundamental to six papers in this thesis, paper A through F. Model-based computational extraction of audio features, i.e. psycho-acoustic descriptors (Fastl & Zwicker 2007; Cabrera 2014; Nilsson 2007b) and music features (Friberg et al. 2014; Lartillot et al. 2008, 2013), was

[1] The deeper implications of these labels, if they were to be taken literally, would be far-reaching. An ontology of 'machine' was set forth by Leroi-Gourhan (1964). See Lindborg (2003) for a discussion of its implications for musical creativity.

employed in papers A, B, C, and E. In papers D and G, audio features played a role as design factors.

‘Human’ measurement has been equally fundamental to the project. Response data in four different modalities have been employed: physiological, semantic, visual, and motoric. Responses captured by psychophysiological reactions (e.g. electrodes on the skin, as in paper A) give evidence of autonomic nervous processing which is primarily involuntary though constantly interacting with the central nervous system (Kreibig 2010, p. 396). Responses via semantic methods rely on introspection and are thus essentially cognitive, yet may provide insight into lower-level processing, e.g. pre-cognitive sound perception, through crossmodal correspondences (Spence 2011). For example, a way to gauge a person’s relationship to a sonic phenomenon, be it a single sound, music, or the soundscape as a whole, might be to ask “What do you hear?”. The response is a function of multiple translations, triggered by the original sensation and processed through classification, appraisal, and memory, resulting in an expression, possibly in a different modality. Semantic

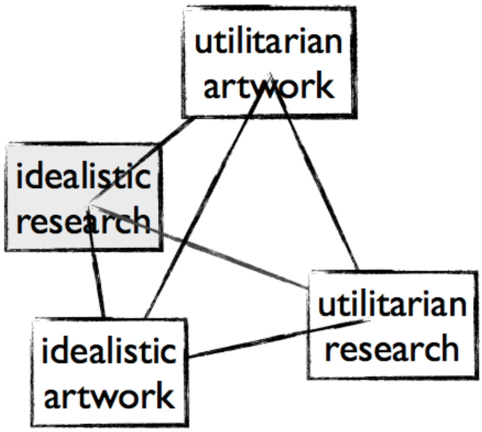

Figure 1. Art and science in idealistic and utilitarian modes.

responses can be in various formats, from unconstrained recorded speech (used in pilot experiments) to constrained written text (Bauer MW & Gaskel 2000; paper F). Visual response methods include sorting of objects or symbols (used in pilots) and colour association (paper C). Labelled graphic rating scales and to some extent Likert scales combine semantic and visual response modalities (papers B, E, F). Finally, motor response actions can take numerous forms, from the basic pressing of a button to indicate spatial direction (paper D) to the highly elaborate schemes a professional dancer develops in counterpoint with aesthetic sonification (Vickers & Hogg 2006; Hermann et al. 2011; paper G). Fig. 2 illustrates how responses to sound can be captured through various channels of measurement and eventually modelled as features (cf. Friberg et al. 2014)

The data generated through objective measurement and many of the subjective methods lend themselves well to quantitative analysis, involving statistics (Kabacoff 2011; Everitt & Hothorn 2010; Howell 2007). A special kind of response is the free-form semantic modality (spoken or written). In this project, such data were gathered through focus interviews (paper C) and open questionnaire items (paper F), and analysed with qualitative analysis methods. Furthermore, some objective data such as observations of typical menu items and interior design materials (paper E) also called for qualitative analysis.

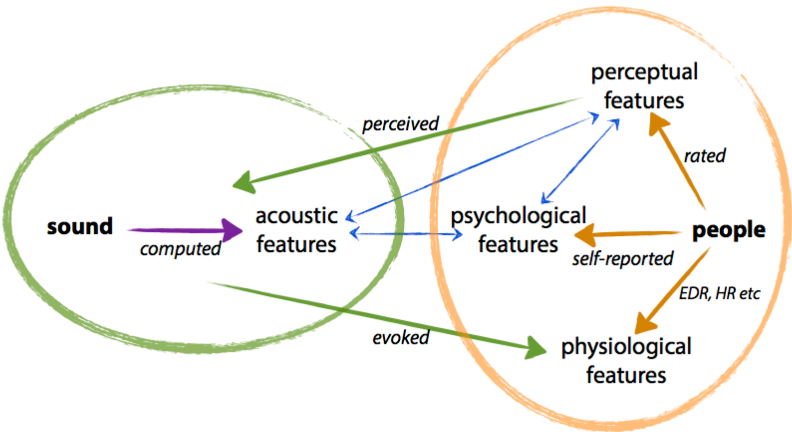

Figure 2. Overview of relationship between sound and people described as features in different modalities.

Qualitative methods might yield structured and detailed data that can be analysed statistically in meaningful ways. The results from quantitative methods might produce models or schemes of classification that feed into qualitative methods. In each of the papers included in this thesis, more than one method was employed. Attempts were made to compare and harmonise knowledge gained from the different sources.

Art is a laboratory for sensory research

As discussed above, *art* is the know-how required to design a sensory experience. This is the sense in which the word is used in the thesis. For example, an experimental procedure that involves an original design is in itself an artwork.

To conduct perceptual experiments in laboratory, software interfaces to present audio and video stimuli to participants were designed. The participants could, for example, respond via semantic scales, visual sorting of symbols, or free-form text. The design based on established praxes for graphical user interfaces (GUI) and enabled papers B and D (also Lindborg 2010; Lindborg & Lim 2013). A more adventurous undertaking was the design from scratch of the *CIE Lab* colour response interface (described in paper C). Experiments with multichannel sound installation design (in particular Lindborg & Koh 2011) made paper A possible through the physical construction of a 3D sound diffusion rig for the realistic reproduction of Ambisonic soundscape recordings. The same sound system was then used in the development of interactive installations and performances (Lindborg, Koh Yong 2011, 2013; Lindborg 2015, Mar.). Paper G focusses on the design of an interactive sonification system which set in motion the “Locust Wrath” series of artwork (Lindborg 2013, 2014, 2015; Lindborg PM & Liu DY 2015).

Introduction

Well then, Älvi, I'll explain how it works. A single gingerbread man might go to pot during the baking so badly that one can barely make it out. But after having seen twenty or thirty gingerbread men – more or less shapely – I can tell with certainty what the cookie cutter looks like. This I can do, even if I haven't seen it. (Gaardner 1991, p. 92, paraphrased)

2. Summaries

The papers in this thesis report work that revolves around four main themes; perception, design, modality, environment (see Fig. 3). Each of the seven publications included in the thesis contributes to knowledge and know-how in some way; either by replicating findings that strengthen existing theory, by proposing novel theory and methods that facilitate future undertakings, or by creating original designs that aim to stimulate a listener's curiosity and satisfaction. The next pages present the highlights, aims and method, and contributions of each paper.

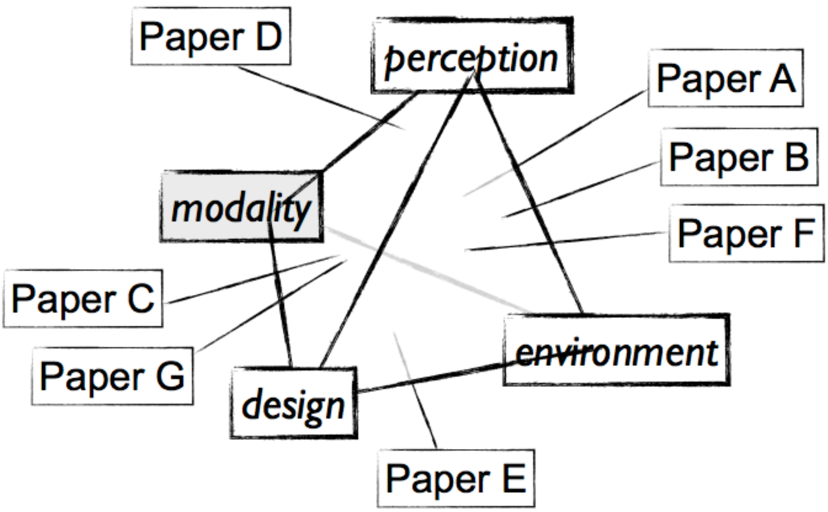

Figure 3. Overview of papers in relation to the main themes of the thesis.

Paper A

Physiological measures regress onto acoustic and perceptual features of soundscapes

Key question

- *How is the body affected by the sonic environment?*

Highlights

- Physiological responses are fundamental to understanding perception.
- The autonomic nervous system is associated with emotions.
- We conducted a perceptual experiment in an innovative 3D sound installation.
- Multichannel responses were captured while participants were immersed in the soundscape.
- Analysis revealed relationships between audio features and ANS activation.

Figure 4. Photo from the experiment in paper A.

Summary

There is no exact model for the relationship between the autonomic nervous system and emotion. Psychophysiological response methods have been used to investigate music perception, but have been more rare in soundscape research. Since it is known that health is influenced by the sonic environment, the study here presented aimed to investigate the nature and strength of relationships between soundscape features and physiological responses linked to relaxation and stress.

In a controlled experiment, seventeen healthy volunteers individually moved freely inside a 3D audio installation while blindfolded. They were immersed in a randomised sequence of twelve soundscapes (nature, urban parks, eateries, shops) reproduced with Ambisonic techniques (see Fig. 4). Five psychophysiological responses were captured: *Skin Conductance, Peripheral Temperature, Blood Volume Pulse Amplitude, Heart Rate,* and *Thoracic Respiration Amplitude.*

Response envelopes were de-trended over the whole session, and residual fluctuation was assumed to reflect physiological reactions to the soundscape. Means and trends were calculated and regressed onto orthogonal acoustic and perceptual stimuli features that had been previously determined.

Results showed that *Peripheral Temperature* was negatively correlated with *SoundMass* (≈ loudness) and with *Calm-to-Chaotic*. In calm soundscapes, such as rural parks, skin temperature was higher, indicating reduced system activation which typically leads to emotional relaxation. In chaotic soundscapes such as shops, the opposite response indicated stress. *Heart Rate* was negatively correlated with *VariabilityFocus* (≈ timbral brightness and fluctuation strength). Heart rate increase in soundscapes with low-frequency fluctuation, such as a street environments, indicated stress, while in soundscapes with the opposite character, such as a waterscape, heart rate decrease indicated emotional relaxation.

Contributions

Paper A contributes to soundscape research by extending findings from other domains to the context of soundscape perception. The experimental design was original, systematic, and ecologically valid. Results underline the importance of psychophysiological methods for understanding basic sound perception.

Paper B

Perception of soundscape quality is influenced by personality traits

Key question

- *Can personality explain how people perceive soundscapes?*

Highlights

- Personality is a stable individual difference that affects perception.
- Two listening experiments using binaural soundscape recordings are described.
- Analysis revealed the influence of personality traits on quality ratings.
- Results showed that the effect size of personality traits was up to one-tenth of psychoacoustic descriptors in one experiment.
- Broad personality dimensions predicted narrow noise sensitivity.

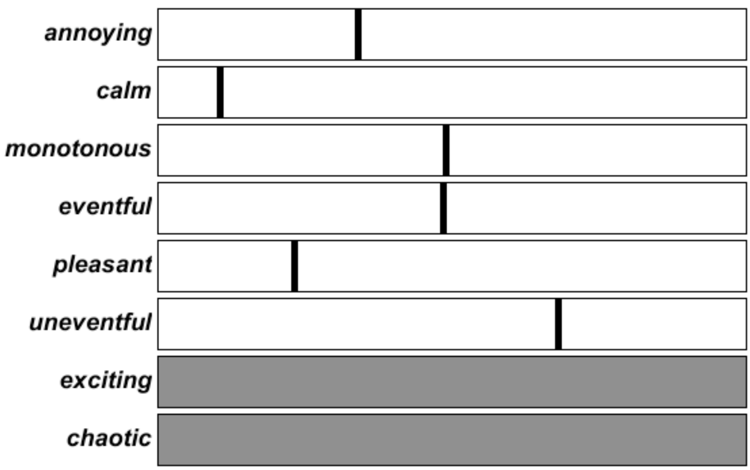

Figure 5. Part of the SSQP response interface for experiments in paper B.

Summary

Empirical soundscape research aims to link the human perception of sonic environments, for example evidenced in quality ratings, with objective measures, for example psychoacoustic descriptors. Individual personality factors might contribute to explaining part of the variation in responses. While noise sensitivity has long been part of investigations, the influence of broad personality dimensions, such the *Big Five*, on perceived soundscape quality has only recently started to be considered.

We conducted two consecutive listening experiments (n = 43, n = 42), where recordings of soundscapes of predefined type (rural parks, urban parks, shops, and restaurants) in which *Pleasantness* and *Eventfulness* were estimated from ratings of the *Swedish Soundscape Quality Protocol*. Personality traits of the participants were measured using self-report instruments (see Fig. 5 for an example), and psychoacoustic descriptors were extracted from the Ambisonic recordings.

In experiment 1, analysis of variance showed that quality ratings were strongly dependent on soundscape type, supporting previous research. Ratings were weakly influenced by certain personality traits, including *Emotional stability* and *Openness*. In experiment 2, a multivariate regression model with selected psychoacoustic descriptors and personality traits explained 25% of the variance in *Pleasantness* and 30% in *Eventfulness* (cross-validated adjusted R^2). The contribution of personality traits, as indicated by the squared semi-partial correlations, reached approximately a tenth of that of psychoacoustic descriptors. Explorations of the data revealed interaction effects between traits and objective descriptors. Further analysis showed that *Noise sensitivity* could be predicted by *Big Five* dimensions, replicating recently published research.

Contributions

Paper B contributes to soundscape research by focussing on individual differences. The results lend empirical support to the hypothesis that personality traits have a significant, albeit small, influence on the perceived quality of sonic environments.

Paper C

Colour Association with Music is Mediated by Emotion: Evidence from an Experiment using a CIE Lab Interface and Interviews

Key question

- *How does perception of music relate to colour?*

Highlights

- Crossmodal correspondences might be mediated by emotion.
- The design of a novel colour response interface is described.
- First use of *CIE Lab* in crossmodal music emotion research.
- We conducted a perceptual experiment with film music stimuli.
- Analysis revealed that emotions contributed over and above audio features in explaining variation in colour responses.
- Mixed methodology approach with participant interviews strengthened emotion mediation hypothesis.

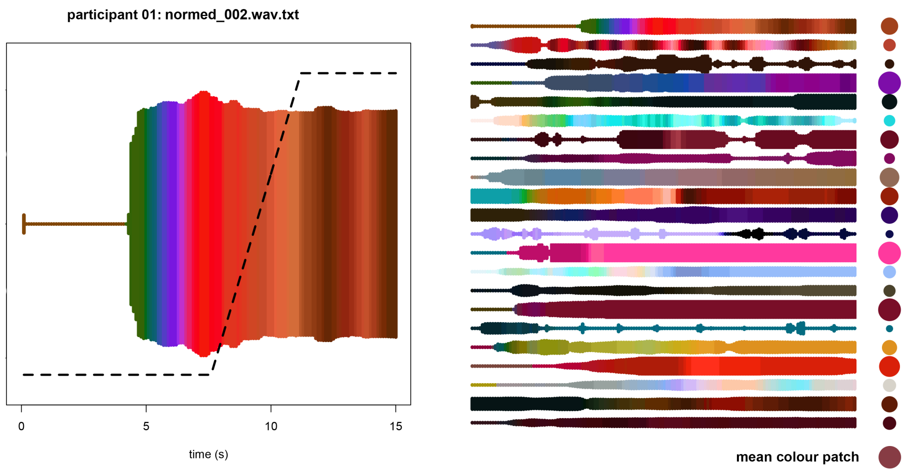

Figure 6. Illustration of how colour response data were weighted and averaged. For details, see paper C, Fig. 2.)

Summary

Crossmodal associations may arise at neurological, perceptual, cognitive, or emotional levels of brain processing. Higher-level modal correspondences between musical timbre and visual colour have been previously investigated, though with limited sets of colour. We developed a novel response interface that employs a tablet interface to navigate the *CIE Lab* colour space, which is based on a perceptual model.

The interface was used in an experiment where 27 film music excerpts were presented to participants (n = 22) who continuously manipulated the colour and size of an on-screen patch to match the music. Focus interviews were also conducted.

Analysis of the data revealed several patterns, e.g. that happy music was associated with yellow colours, music expressing anger with large red colour patches, and sad music with smaller patches towards dark blue. Correlation analysis suggested patterns of relationships between audio features and colour patch parameters. Using partial least squares regression, we compared models for predicting colour patch responses from audio features and ratings of perceived emotion in the music. Parsimonious models that included emotion outperformed the audio-only model and robustly explained between 60% and 75% of the variation in the colour patch parameters, as measured by cross-validated R^2.

Content analysis of structured spoken interviews with the participants provided further evidence of an emotion mediation mechanism, whereby people tended to match colour with the perceived emotion in the music.

Contributions

Paper C contributes to research in audiovisual crossmodal association by presenting a novel colour response method (see Fig. 6), by replicating and expanding upon previous finding, and by suggesting new relationships distilled through systematic analysis. The mixed-method approach of our study provides strong evidence that emotion can mediate crossmodal association between music and visual colour. The *CIE Lab* interface promises to be a useful tool in perceptual ratings of music and other sounds.

See also the thesis cover image which represents mean colour responses to the 27 stimuli plotted in 3D emotional space (Valence, Energy, Tension).

Paper D

Audio Quality Moderates Localisation Accuracy: Two Distinct Perceptual Effects?

Key question

- *Does audio quality affect the perceived auditory space?*

Highlights

- Localisation accuracy might depend on audio quality and other factors.
- We conducted a perceptual experiment with a factorial design.
- Results suggested a crossmodal effect of audio quality on spatial imaging.

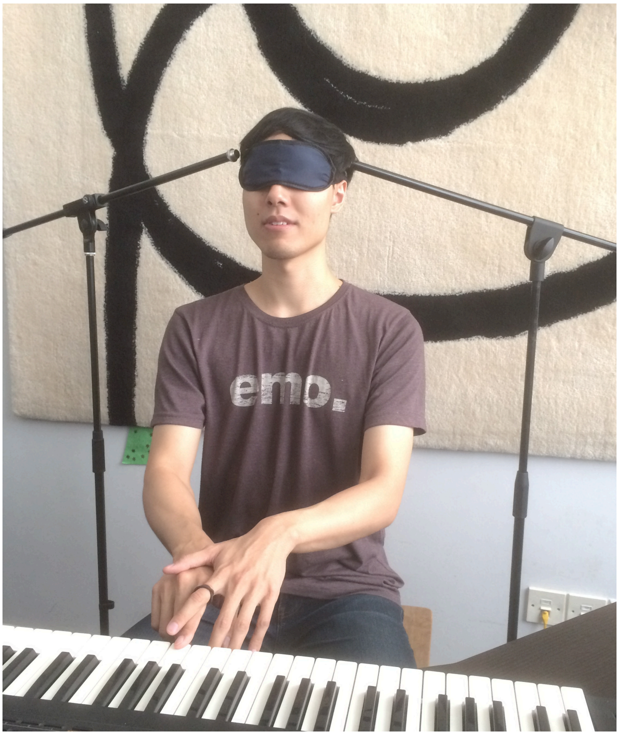

Figure 7. Photo from an experiment continuing the work described in paper D.

Summary

It is generally understood that higher audio quality enables better source localisation. Audio quality is known to cross-modally influence reaction speed, sense of presence, and visual quality. Previous accounts have anecdotally described subjective effects of lossy audio encoding on spatial imaging. Perceptual effects of audio quality will influence listener satisfaction.

We designed an experiment to test the effect of audio quality on the soundstage image, specifically, how compression rate affects source localisation in a situation with concurrent maskers. A realistic listening scenario was designed in which the listener was tasked to localise one instrument within a rhythmic pattern. As a proxy for audio quality, we generated stimuli with different rates of MP3 compression. To create a realistic listening situation, the stimuli were rhythmic patterns of drum samples typical of studio production music. Participants (n = 18) estimated the position of the snare drum target while compression rate, masker, and target position were systematically manipulated in a full-factorial repeated-measures experimental design.

Analysis of variance revealed that localisation accuracy was better in wide target positions than in narrow, with a medium effect size; and that the effect of target position was moderated by compression rate in different directions for wide and narrow targets. This suggested that there might be two perceptual effects at play: one, whereby increased audio quality causes a widening of the soundstage, possibly via a SMARC-like mechanism, and two, whereby it enables higher localisation accuracy. In the narrow target positions in this experiment, the two effects acted in opposite directions and largely cancelled each other out. In the wide target presentations, their effects were compounded and led to significant correlations between compression rate and localisation error.

Contributions

Paper D contributes to research by investigating a commonly held belief with a systematic and possibly new approach. Results suggest perceptual effects that might be linked to crossmodal correspondences. We have pursued this line of investigation more recently. Further data collection and a preliminary analysis have been made (see Fig. 7), and might be published at a later point.

Paper E

Psychoacoustic, Physical, and Perceptual Features of Restaurants: A Field Survey in Singapore

Key question

- *Which design features are important for restaurant soundscape quality?*

Highlights

- We advance a typology of acoustic design elements relevant to restaurants.
- Empirical results from a survey of 112 restaurants are reported.
- Correlation analysis revealed relationships between various features.
- Annotations of interior design materials supported classification by *Design Style* and *Food Style*.

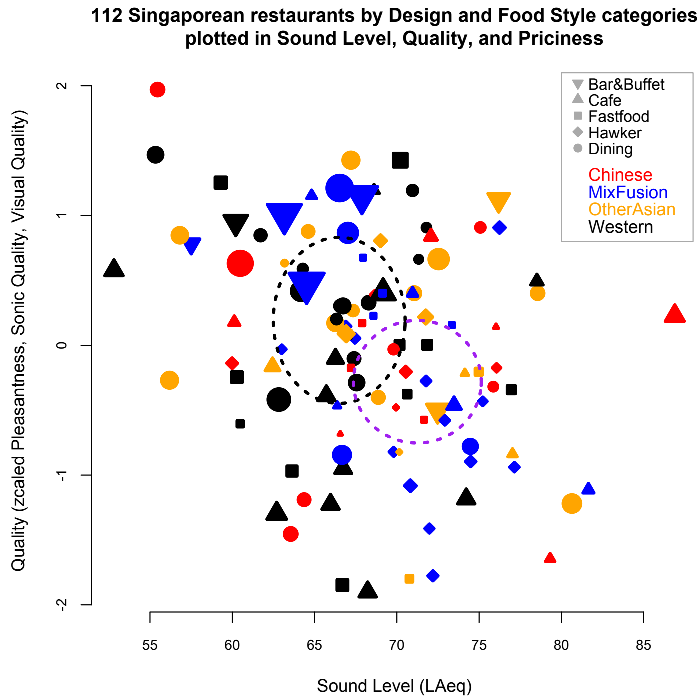

Figure 8. Plot of 112 Singaporean restaurants by selected features. For details, see paper E, Fig. 3.

Summary

Sound is a multi-faceted phenomenon and a critical modality in all kinds of sevicescapes, which have received less attention from researchers than the urban outdoor soundscape; there is a scarcity of published empirical data. At restaurants, our senses are intensively stimulated. They are social places that depend on acoustic design for their success. Contributing to theory building, this article proposes a typology of designed and non-designed acoustic elements in restaurants.

Few empirical studies on the psychoacoustics of restaurants have been published. To address this, we performed a survey of 112 restaurants. The collected data included sound level, audio recordings from which psychoacoustic descriptors such as *Loudness* and *Sharpness* were calculated, on-site perceptual ratings with the *Swedish Soundscape Quality Protocol*, and annotations of physical features such as *Occupancy*. We derived a measure of *Priciness* to compare menu cost levels between restaurants.

Correlation analysis revealed several patterns: for example, that *Priciness* was negatively correlated with Loudness. Analysis of annotations of interior design materials supported a classification of the restaurants in categories of *Design Style* and *Food Style*. These were investigated with MANOVA, revealing significant differences in psychoacoustic, physical, and perceptual features between categories among the surveyed restaurants (see Fig. 8). For example, less expensive places were noisier, the more pleasant restaurants had fewer sounds identified with people or machine sources, restaurants serving Chinese food had the highest prevalence of stone materials, and Western-menu places were the least loud.

Contributions

Paper C contributes to research in servicescape design by presenting a broad and detailed survey of restaurants. To our knowledge, there is no published account of a comparable survey of this size. The proposed typology of acoustic design elements, as well as the empirical findings, are valuable for managers, acoustic designers, and researchers.

Paper F

A taxonomy of sound sources in restaurants

Key question

- *Which sounds are liked or disliked in restaurants?*

Highlights

- We gathered perceptual data on sound and soundscape perception through a field questionnaire.
- Systematic classification of annotated characteristic sounds yielded a taxonomy.
- Reliability of a taxonomic clade with four levels was shown through internal and external validations.
- Analysis revealed perceptual and crossmodal effects.
- The taxonomy is applicable to field research and simulation.

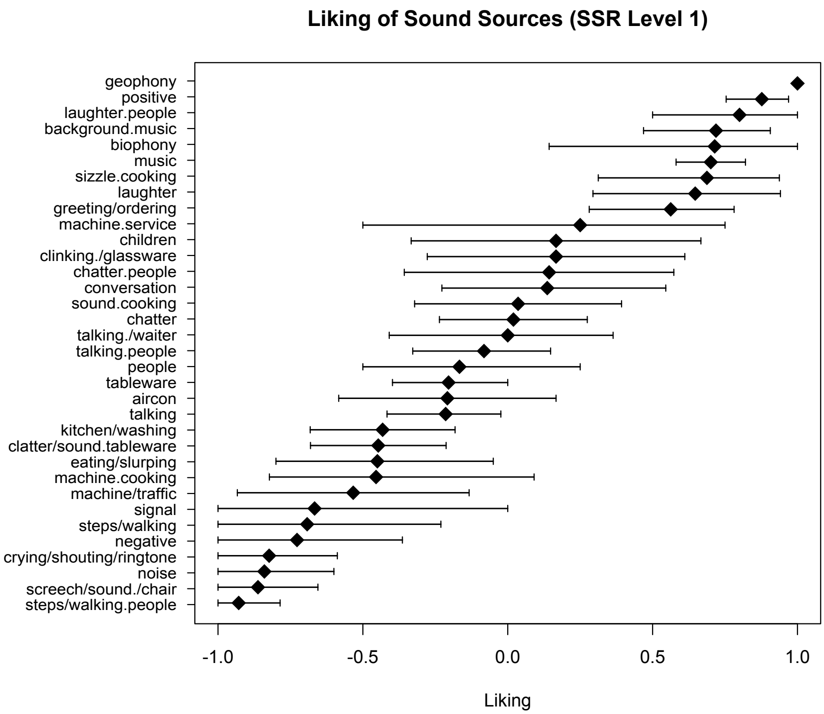

Figure 9. *Liking* of 34 sound sources in restaurants. For details, see paper F, Fig. 2.

Summary

Soundscape quality is understood in different ways depending on context. Restaurants are complex environments where all our senses are engaged. More or less designable sound sources such as background music, voices, and kitchen noises are believed to be important in relation to the overall perception of the soundscape.

We collected 1018 annotations of characteristic sounds from 393 customers in 40 restaurants. Annotations were analysed with a cladistic approach to develop a multi-level taxonomy of perceived sound sources in restaurants, SSR. Different classification taxa were evaluated by comparing *Liking* estimates (see Fig. 9) with *Pleasantness* ratings obtained on-site with the S*wedish Soundscape Quality Protocol*. A four-level clade was efficient and out-performed all the other alternatives. Internal validation of the *Pleasantness* construct was made by comparing it with an alternative measure derived through separate ratings (n = 7) of free-form descriptions of the environment as a whole. External validation was made with a data set obtained in a laboratory listening experiment where audio recordings of 15 restaurant environments were rated (n = 48). The validations demonstrated that the SSR taxonomy maintained good internal construct validity, was consistent across levels, and showed external robustness.

Exploring the data further, we found that voice-related annotations of characteristic sounds where 'people' was included as a specifier were more liked than those were it was not, which might be due to an emotional crossmodal association mechanism. We also found an order effect on *Liking* of characteristic sounds, evidenced by significant differences between first and last annotations, which might be due to an initially positive bias being countered by exposure to a task inducing a critical listening mode.

Contributions

The SSR taxonomy introduces a detailed classification of perceived sound sources in restaurants. It offers validated estimates of liked and disliked sound types structured in units and levels. Derived models will be useful for field research and design simulations of restaurant soundscapes. The empirical findings might inform theory, specifically research charting the perception of sound sources in multimodal environments.

Paper G

Interactive Sonification of Weather Data for *The Locust Wrath*, a Multimedia Dance Performance

Key question

- *Can sound be designed so that its meaning is understood intuitively?*

Highlights

- Sonification is a practice that bridges science and art.
- We report the process of creating the sound design for an intermedia artwork with dancers.
- A system for interactive sonification of climate data is described.
- Parameter mapping strategies were motivated by crossmodal correspondences.
- Meaning-creation through metaphors based on ecologically founded modes of listening are discussed.

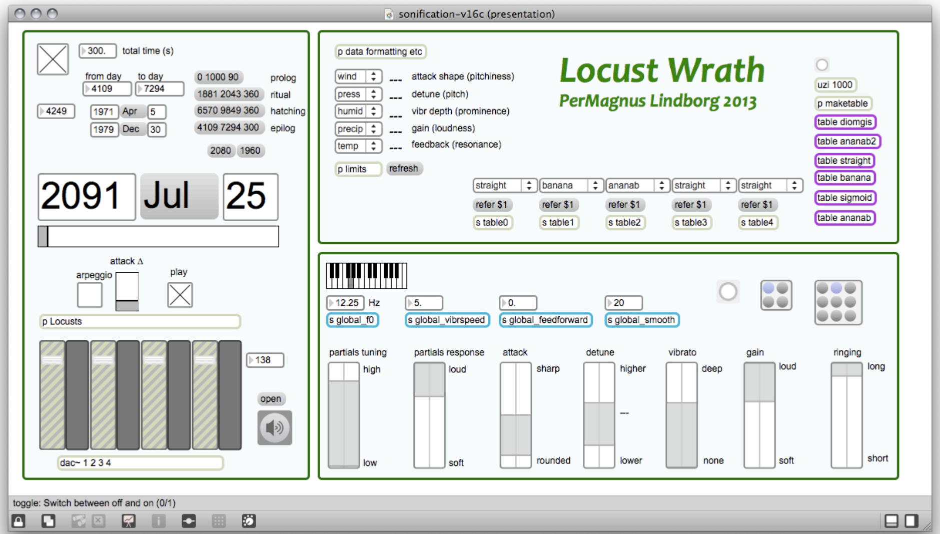

Figure 10. Locust Wrath GUI for interactive sonification (made in Max, Cycling ’74).

Summary

The two main approaches to data sonification, “aesthetic” and “scientific” sonification, are complementary. They have a substantial overlap in terms of methods and applications, but we argue that there is no conflict of interest between the two.

“The Locust Wrath” is a multimedia dance performance on the topic of climate change. To create the sound design, we developed a software for interactive sonification of meteorological data (see Fig. 10). An open-ended approach to parameter mapping allowed tweaking and improvisation during rehearsals, resulting in a large range of musical expression. The multichannel sonifications represented weather systems pushing through geography in complex patterns. Sonification compresses time; weeks of weather pass by in seconds. The climate was rendered as a piece of electroacoustic music, whose compositional form - gesture, timbre, intensity, harmony, spatiality - was determined by the data.

Paper G discusses aspects of aesthetic sonification, reports the process of developing the present work, and contextualises the design decisions within theories of crossmodal perception and modes of listening.

A companion paper (Lindborg & Liu 2015) presents a recent continuation of the project, where a novel spatialisation method for sonification was developed. Audience members participated to create a large, immersive sound design by connecting ad hoc iOS devices in a server-client network.

Contributions

The project described in paper G contributes to sonification research by bridging practices from science and art. The artwork is part of a long-term collaboration with dancers, designers for stage, and fellow sound artists. The subject matter of climate change is pressing and the public interest we have gained stimulates sound design research. The companion paper contributes by presenting a novel auditory display method.

Introduction

Brief Pause in the Organ Recital. The organ stops playing and the church becomes deathly quiet, but only for a couple of seconds. And the faint rumbling penetrates from the traffic outside, that greater organ. For we are surrounded by the murmurings of the traffic, it flows along the cathedral walls. The outer world glides like a transparent film and with shadows struggling pianissimo. And as if it were part of the street noise, I hear my pulse beating in the silence. I hear my blood circulating… I hear the truck that rumbles past and makes the six-hundred-year-old walls tremble. (Tranströmer 1984)

3. Sound perception

The latin word *percipere* means 'to receive'. This chapter first reviews aspects of how the sensory system processes information, in particular auditory. Then follows a section on how humans are affected by sound and how knowledge is gained about the environment through sound. Finally, fundamental concepts pertaining to soundscapes are discussed, starting with a holistic perspective and narrowing down towards the perception of single sound events.

Multimodal sensing

Sensory information processing is inherently multimodal. In normal situations, an organism perceives the environment using all its senses simultaneously. Crossmodal correspondence might take place at any stage of neural processing (Spence 2011). When auditive and visual stimuli are simultaneously presented, a crossmodal percept might reflexively be created. By this phenomenon, 'synchresis', a common cause is attributed to both. This might happen even if sensations were actually produced by different sources (Chion 2009 p. 492).

The senses have different range or specialisation; e.g. vision gives more accurate spatial information while audition gives more accurate temporal information. People are biologically compelled to interpret sensorial stimuli in terms of real-world elements, even when the elements they cognise (become aware of) are imaginary. For example, by default, perceivers attribute a specific physical source to any sonic stimulus. It is only through dedicated practice that people can de-learn this innate mental mechanism (cf. Schaeffer 1966 p. 95). Auditory perception normally involves one or the other form of crossmodal association as our hearing is conditioned (innate or acquired) on ecological situations where sounds events always appear as a part of a multimodal context (cf. Clarke 2005; Thibaud 2011; further discussed in papers C, F, and G).

Since this thesis is about sound the most important term might be *listening*. Philosophical perspectives on listening (e.g. Schaeffer 1960; Schafer 1994; Truax 2001; Juslin & Västfjäll 2008; Juslin 2013; Tuuri & Eerola 2012; Miller 2012) will be discussed after a review of how the brain processes auditory information.

Information processing stages

Three stages might be discerned in the complex sensory information processing of the human brain: sensation, perception, and cognition. The abbreviated perspective presented here is primarily ‘bottom-up’, but note that there are also ‘top-down’ mechanisms whereby conscious cognition modulates sensation and perception (for details see e.g. Västfjäll 2010; Schnupp et al. 2011; LeDoux 1998; Peretz 2010).

Sensation

Sensation is physio-neurological and mainly reflexive. An introduction can be found in Moore & Linthicum (2011). The brain stem is an ancient structure of the brain that subserves auditory and other sensorial perception. It controls or mediates fundamental psychological reactions such as attention reorientation and arousal, and modulates physiological functions such as heart rate and breathing (cf. paper A). However, “attention is difficult to define but easy to detect” (Mesulam 1998 p. 1036).

Sensing sound is crucial to survival. The biological task for the auditory system is to alert the organism to important changes in its environment. This principle is fundamental not only for individual earthworms in the compost heap, but also for relatively high-level activities such as music (see Huron 2006 in particular ch. 2). Audition is always active, and specialises in detecting sudden, extreme, or quickly approaching sounds that might necessitate action (Juslin & Västfjäll 2008, p. 564). Sound might act as a stressor in at least three ways: physiological stress on the hearing organ, in particular at high pressure levels (Moore 2012 ch. 4.8) but also at levels within legal limits of noise exposure (paper E); psychophysiological stress e.g. on heart rate and metabolism (paper A) especially with long exposure (Berglund & Nilsson 2007); and physiological stress to the voice in noisy environments (Rindel 2010; paper F).

The auditory system is tasked with scanning the entire spatial environment continuously, both when the organism is awake and asleep (cf. 'background listening', Truax 2001). It is has higher temporal resolution than vision, and higher spatial resolution than olfaction. Certain sound qualities are indicative of chance, such as sudden or extreme sounds, sounds that change very quickly, or sounds that are the result of strong force or large size (Juslin & Västfjäll 2008, p. 564; compare Truax p. 110-111). The autonomic nervous system (ANS) is part of an organism's system to control organs and body functions through neuronal (rapid, precise, differentiated) and hormonal (slower and more diffuse modification of metabolic functions) activity levels. Its task is to maintain homeostasis, i.e. to maintain a relatively constant inner environment for the organs to function properly, and assure the survival of the organism. It achieves this by effectuating adaptive responses to various environmental demands. For example, lesion studies indicate that the amygdala brain region, which processes information from the external environment, is central in ANS activation (Mesulam 1998 p. 1035). In fact, ANS might be involved in nervous activity at all levels (Kreibig 2010 p. 396). Depending on situation, it elicits quick response mobilisation, "fight-or-flight", via the sympathetic nervous system (SNS), or a "rest and digest" response via the parasympathetic nervous system (PNS). These two subsystems are continuously modulating bodily vital functions, usually in antagonistic fashion. Paper A provides further details and investigates how immersion in sound causes physiological changes.

Early processing encodes neuronal activity into multiple parameters of auditory stimuli, including frequency spectrum, loudness, time patterning, and spatial location (Västfjäll 2003; also paper D), to produce an integrated neural representation of the stimulus. This information is then passed on to the forebrain (Moore & Linthicum 2011 p. 1262).

Perception

Perception is neurological. It is the organisation, identification and interpretation of sensory information in order to represent and understand the environment. Perception mediates between, on the one hand, physiological sensation (Bech & Zacharov 2006; Pulkki &

Karjalainen 2015), and on the other hand, cognition (Mesulam 1998), which, per definition represents that which is available to the individual for introspection and awareness. Perception does not just filter and transmit sensory signals, but is shaped by signals from higher-level cognitive areas. For example, in a situation of danger, the organism goes on high alert and actively searches for a useful auditory signal. This is a case of intentional listening, what Truax (2001) would call 'listening-in-search'. Through different neurological pathways, information travels in 'sensory-fugal' or upstream direction as well as in 'sensory-petal' or downstream direction. The former refers to processing that feeds into higher-level, cognitive information, and the latter refers to cognitive actions such as learning, memory and expectation that direct perception. Perception depends on complex functions of the nervous system, but subjectively it is fairly effortless[2] as it takes place beneath the horizon of awareness.[3]

From the perspective of measurement, perception is non-linear, time-varying with both short-term and long-term effects, and depends on innumerable factors. Comparatively less is known about the auditory perceptual pathways than the corresponding pathways for visual processing. It is believed that early auditory processing areas encode more elementary features, such as frequency and pitch, whereas downstream areas encode more composite features related to e.g. localisation of sound sources (paper D; Asutay & Västfjäll 2015), categorisation of object-specific sounds (paper F), and the characterisation of individual voices (Mesulam 1998 p. 1023; Nahata 2015). Psychoacusticians model the relation between a physical sound event and a perceived auditory event as a probability function (Loy 2006 ch. 6). Existing psychoacoustic models are typically non-linear (Fastl & Zwicker 2007). The human auditory system has been half-jokingly described as "an unreliable measurement device for measuring an unpredictably behaving system" (Pulkki & Karjalainen 2015, p. 133). Yet this biological sensing device can be fascinatingly robust and reliable. Even though an auditory event is a perceptual construct and internal to the listener, aspects of lower-level

[2] Note that subconscious processing still taxes brain processing resources. Unheard noise is tiring because it activates the autonomic nervous system (cf. paper A)

[3] This leads to the concept of *intuition* which will not be pursued here.

processing might be accessed through introspection; for example, as a response to the question: ”what did you hear?” Note however that introspection is subjective and generally not acceptable as scientific evidence unless confounding factors are controlled and carefully contextualised. Statistic analysis of multiple observations improves reliability.

Cognition

Cognition is neuro-psychological. It is a faculty for the processing of information, applying knowledge, and changing preferences. Cognitive processes can be conscious or unconscious (Nahata 2015). It is the emergence of multiple parallel perceptual processing analyses, and manifests itself as memory, emotion, attention, language, thought and consciousness (Mesulam 1998; Huron 2006). What we call cognition is not *per se* neural. The *Language of Thought model* (LOT) describes cognition as a language operating on symbolic representations, "a computation in *mentalese*" (Fodor 1975 p. 62; discussed in Lindborg 2003). It is a set of functions that allow an individual to gain knowledge, also without prior access to exact sensorial information, through the mechanism of deductive inference. Aspects of cognition, such as awareness of self, were historically used to demarcate between humans and animals. It is however clear that some birds and mammals are able to plan their actions and have a sense of individual self, for example, by being able to recognise their appearance in a mirror (e.g. Rajala et al. 2010; cf. Premack 2007 for an opposing perspective). On the other hand, the behaviour of many animals can be machine-like ('sphexishness'; Hofstadter 1982 p. 532; see also Lindborg 2003).

Cognitive assessment of sound sources depends on hearing certain sounds as foreground events (‘synechdoche’; see below). The strength of reactions to foreground events depends on whether they are recognisable or not (Asutay, Västfjäll et al. 2012). Indeed, source identification might be an even stronger predictor of perceived quality than sound level (Nilsson 2007a). Auditory cognition is highly complex. The BRECVEMA theoretical framework (Juslin & Västfjäll 2008; Juslin 2013) goes beyond cognitive appraisal to describe eight mechanisms through which music listening may induce emotion: brain stem reflex, rhythmic entrainment, evaluative conditioning, [emotional] contagion, visual imagery, episodic memory, musical expectancy, and aesthetic judgement. The framework produces a number of hypotheses that allow for separate and focussed investigations (e.g. Juslin et al. 2010; Juslin, Gonçalo & Eerola 2015). Since this theory is based on ecological perception (Gaver 1993; Clarke 2005), most if not all of the eight mechanisms are likely to apply in soundscape listening as well. At the upstream (early sensation) end of the processing chain, there is no reason to assume that brain stem reflexes work any differently in the context of music from what they do in the context of everyday soundscapes. At the sensori-petal (high-level cognition) side of brain processing, the situation is not different. For example, precisely because the mechanism of musical expectancy depends on active and cognitive modulation (e.g. willpower, learning), it might be applicable to soundscape listening, if the attention is guided towards the variation and detail of specific sound sources in the environment, as well as their internal relationships and inherent communicative meaning, rather than their sheer number or the monotony of the soundscape as a whole that is produced when few and predictable sounds are allowed to dominate.

Affect

Affect is central to people’s life (Västfjäll 2010; Västfjäll et al. 2002). The word is an umbrella term that covers all evaluative states such as emotion, mood, and preference (Juslin & Västfjäll 2008, p. 561; Coan & Allen 2007). Affect is an essential part in people's experience with music (Huron 2006) and in social encounters (Lazarus & Folkman 1984), and it is just as relevant when attempting to understand feelings and reactions to everyday soundscapes. The concepts of emotion, appraisal, and individual differences will be reviewed here.

Emotion

Emotion is an affective reaction of relatively brief duration. Affective states are valenced, as in positive /negative, good /bad etc. (Osgood et al. 1957; Mehrabian & Russell 1974; Russell 1979), that may modulate lower-level processing stages in ‘top-down’ pathways (Mesulam 1998 p. 1032-1037). In other words, cognitive evaluation activates affect-congruent material in memory, and primes the individual towards judgements and actions in that same direction (Västfjäll 2010 p. 266). The distinction between induced and perceived emotion (which are related to recognition and experience) might in reality be blurred. Gabrielsson and Lindström (2010) suggested that the two alternatives could be seen as opposite extremes of a continuum. Empirical studies have found more similarities than differences between perceived and evoked emotion (Eerola & Vuoskoski 2011). Another important distinction in regards to emotion lies between models based on dimensional emotions (or continuous, such as arousal) and discrete emotion (or basic, such as happiness). See paper C where both were used.

Appraisal

Cognitive appraisal accounts for most emotions encountered in everyday listening (Juslin 2013; Juslin & Västfjäll 2008). The *Swedish Soundscape Quality Protocol* (SSQP; Axelsson, Nilsson & Berglund 2010, 2012; Axelsson 2011) was developed to measure the appraisal of soundscapes, in particular outdoor environments (e.g. Joo & Yin 2015), but also indoors (paper F) and in laboratory (papers A and B). Studies of the urban soundscape have shown that most people prefer natural over technological soundscapes (Schafer 1994; Axelsson et al. 2010; see papers B and F for discussions). Natural and human sounds show great variation. "Therefore", Truax writes, "it is not surprising that such sounds are preferred by most listeners, and that their particular qualities often serve as reference points in our evaluation of other sounds". (Truax 2001 p. 111).

It might be easier to identify what is annoying than what is pleasing (Cohen et al. 1987, p. 169). While isolated sounds might be appraised reliably as more or less pleasant, for complex sounds (perhaps also for simple sounds in context) it is possible to simultaneously hold positive and negative perceptions. Appraisal depends on expectations, preferences, mood, and current activities (Pulkki & Karjalainen, ch. 17 especially p. 351; discussed in papers B and F). Some sounds are almost universally undesirable, yet still appropriate in particular settings. For example, screeching sounds are normally unpleasant. Yet in a specific situation such sounds might represent something positive in the mind of individual perceivers. Schafer called "the scraping of the heavy metal chairs" at Parisian cafés "uncounterfeitable sound souvenirs... in need of protection" (Schafer 1994 p. 240).

The variation in evaluation of human sounds and sound generated by human facility presents a challenge to the taxonomist aiming to predict responses to prevalent sounds. Perceptual experiments have shown that soundscape evaluation depends on the activities that the listener imagines within it. For example, the ‘liking’ of sound categories might reflect context-dependent appraisal. Paper F showed that the perceived presence of ‘people’ was associated with a more positive appraisal of voice sounds (‘laughter’, ‘chatter’, and ‘talking’) than when it was absent. The observed effect might have been caused by an emotional crossmodal mechanism whereby the affective evaluation of a sound was influenced by the perceiver’s social relation to the source. This hypothesis might be tried with an Implicit Association Test (IAT; e.g. Västfjäll 2010 p.265). Paper F replicated previous findings that the “voices of others” are a major source of nuisance in restaurants. The problem is linked to the Lombard effect, which has been known for a century (see Brumm & Zolliger 2011 for historic and cross-species perspectives; also paper F). The effect is compounded with acoustically inferior design (see paper E).

Individual differences

As highlighted in the ISO definition, context and individual perception matter (BS/ISO 2014). There are infinitely many ways in which individuals can be said to differ. In the present context, a distinction is made whether sound perception applies to individuals, which is in the domains of physiology and psychology, or to groups of individuals and societies, which pertains to sociology or culture (Augoyard & Torgue 2006 p. 15). For soundscape perception, the most relevant physiological differences include gender and age (e.g. Yu & Kang 2010), and the most relevant psychological differences include broad personality traits (John & Srivastava 1999; Russell & Mehrabian 1977; Stansfeld 2000; paper B) and narrowly defined traits such as noise sensitivity (Shepherd, Heinonen-Guzejev et al. 2015; Belojevic et al. 2012; paper B). Accounting for these kinds of individual differences might be crucial in certain acoustic environments, such as schools, going beyond the attention paid to differences in terms of sociology or culture. People might individually adapt to partly inadequate or 'semi-designed' (see ch. 5) environments via habituation (Schafer 1994) or coping strategies (Lazarus & Folkman 1984), especially if empowered with choices (Andringa 2010; Payne 2013; Payne & Guastavino 2013) or if facilitated by technology (Lindborg & Lim 2013).

Sound as soundscape

The neologism 'soundscape' was introduced by Schafer in the 1960s. In the recent standard definition, it refers to a perceptual construct, while the term 'sonic environment' refers to a physical phenomenon (ISO 12913-1 2014, Part 1 §2:3; cf. Truax 2001 p. 50). Just as the term 'landscape' typically implies aesthetic appreciation, e.g. through painterly depiction, soundscape, as originally conceived in the World Soundscape Project (Truax 1999, 2001), was primarily concerned with the aesthetics of natural sonic environments, as well as their protection and preservation. The soundscape concept eventually shaped the field of acoustic ecology (Wrightson 2000). Over the past two decades, soundscape research further expanded to become an umbrella term for multiple practices, including sonic artwork in servicescapes (Hellström, Dyrssen et al. 2011), urban planning (Kang 2010), and public health (e.g. Berglund & Nilsson 2007).

Sonic effects in time and space

A design-oriented perspective on soundscape will be discussed in relation to the original conceptualisation. Augoyard, Amphoux, Thibaud, and collaborators (CRESSON) found Schafer's definition "too broad and blurred" for operationalisation in the context of architectural and urban planning (Augoyard & Torgue 2006 p. 6-7; see also Hellström 2003 p. 20-21). While much influenced by Schaeffer's work to capture the phenomenological aspects of the 'object sonore' (Schaeffer 1966 Livre V: morphology and typology), they sought a way to describe sound perception that was anchored in the real world and at the same time, more impactful and practicable than the soundscape concept. As a connector between these extreme poles they developed "the paradigm of a sonic effect" (Augoyard & Torgue 1995, 2006; Thibaud 2011; Hellström 2003) whereby the collective urban space is primarily understood as a "place of sociability" (Amphoux 2011, p. 6). The aim was to compare the physical characteristics of urban environments with accounts of perceptual effects the same places have on listeners (Augoyard & Torgue 2006, p. xii), linking the "interaction between the physical sound environment, the sound milieu of a socia-cultural community, and the 'internal soundscape' of every individual." (Augoyard & Torgue 2006 p. 9). Their (typological) classification contains 82 effects, which can be understood as categories of psychoacoustic perception, that are characterised in relation to domains, in particular ‘physical and applied acoustics’, and ‘psychology and physiology of perception’. Three major effects will be discussed here: ‘ubiquity’, ‘metabole’, and 'synechdoche'. See Hellström (2003) for an in-depth review.

Sonic environments extend in temporal and spatial dimensions. A listener in the world is constantly immersed in a soundscape that might be sparse or dense, thin or thick. Immersion in sound increases the strength of emotional responses (cf. Västfjäll 2003). This situation produces the first two perceptual effects.

Ubiquity

The latin word *ubique* means ‘everywhere’. Any background sound such as an urban traffic drone can be perceived as ubiquitous in the very literal sense that the acoustic propagation is highly dispersed and arrives at the listener from numerous if not all directions. Diffuse, unstable, omnidirectional sound presents an intrinsic tendency toward ubiquity, leading to a confusion between the locus of a sound and the locus of a source (Amphoux 1995 p. 142; see also paper F), and impedes upon the listener’s sense of orientation and location in space (Hellström 2003 p. 81). However, this aspect of background sound does not suffice to produce the effect. Beyond the mere acoustically produced difficulty or impossibility of locating a sound source, the ubiquity effect suggests a metaphysical dimension of sound perception. For the effect to occur, the sound “questions its location, and disavows, at least momentarily, the listener of its identification”[4] (Amphoux 1995 p. 142, my translation; cf. Amphoux 2006 p. 130-145; Hellström 2003, p. 81). Thus urban drones or roaring seas might be ubiquitous, but they do not cause the effect since the perceiving subject knows where they are. It is when the listener cannot locate a ubiquitous sound that it causes stress, especially if loud. “The uncertainty… establishes a power relationship between an invisible emitter and the worried receptor” (Amphoux 2006 p. 131). Music in servicescapes might produce the ubiquity effect not by being loud but by having neither centre nor circumference. “That”, Schafer declared, “is how

---

[4] “que le son interroge quant à sa provenance et qu’il tienne en échec, au moins quelques instants, celui qu’il interroge”

the medieval theologians defined God… the schizophonic voice, invisible but authoritative and omnipresent" (Schafer's foreword, in Augoyard & Torgue 2006 p. xv).

Ubuiquity is almost unequivocally negative. Sonic non-specificity drowns our ability to hear distant sounds (Augoyard & Torgue 2006 p. xv), and Schafer did not mince his words when referring to the traffic drone as a "sound sewer", suggesting that ugly soundscapes were "likely to result when a society trades its ears for its eyes, and… certain to result when this is accompanied by an impassioned devotion to machines." (Schafer 1977 p. 237). Reverberant spaces are prone to the ubiquity effect since reverberation reduces the possibility of orienting oneself by listening attention. It becomes difficult to perceive depth in sonic environments when distance cues drown in prolonged reverberation tails. This causes sound events to be perceived as lacking in temporal as well as spatial precision (cf Barron 2010). Such spaces, without "explicit, comprehensible sonic centre" (Hellström 2003 p. 139), might bring a feeling of discomfort[5] (Chelkoff 1995 p. 88; cf. Cohen et al 1987; Schafer 1994 p. 225). This is often but not always the case in semi-designed environments (see ch. 4 & 5). In designed contexts when reverberation is carefully controlled, the effect might obviously be entirely positive (e.g. in films; see Chion 2009 p. 242).

Metabole[6]

The metabolic effect is in time what ubiquity is in space[7] (Chelkoff 1995, p. 86-91). It refers to a soundscape where the relationship between its constituents parts (acoustic elements), which might be indistinguishable as individual events, is constantly shifting. The metabolic soundscape is unstable over time, yet the whole is perceived as a static entity, because the perception of the whole blurs clear distinction of detail. The effect is thus perceptual. It may also be structural, in the sense that the concept can be used as an attribute: an environment can be more or less metabolic. It might be an indicator of whether a *place* has been acoustically designed

---

[5] "les sons… dans l'effet d'ubiquité… sont anonymes (d'où un sentiment de malaise"

[6] The original term in Augoyard & Torgue (1995), 'métabole', was rendered as 'metamorphosis' in the 2006 translation, but 'metabole' is straightforward and is in fact more appropriate (see http://www.etymonline.com/index.php?l=m&p=26).

[7] "la métabole est au temps ce que l'ubiquité est à l'espace"

or should rather be considered as 'semi-designed' (see ch. 4–5). The uniformity of metabolic soundscapes tends to desensitise listeners cognitively (Truax 2001 p. 110-111), just as constant sound fatigues the hair cells of the inner ear (Moore 2012 p. 152; Schnupp et al. 2011 p. 67). Schafer found that cities in general did not display the deliberate and delicately attuned acoustic rhythms of natural soundscape, or even rural villages. Their principal feature was randomness rather than music-like structure: “The great profusion of rhythms cancel one another out.” (Schafer 1994 p. 233). Everyday listening (Tuuri & Eerola 2012) may be metabolic, or 'floating' (Aphoux 2003). The perspective of ubiquity and metabole give rise to the question of how sound events emerge in the mind of a listener, and how the perception of single events relate to the perception of the soundscape as a whole.

Sound as event

The recent ISO definition of soundscape emphasises that it is a perceptual construct originating in sound sources, distributed in space and time, in a physical environment. (BS/ISO 2014). The number and variation of sound sources in our environment is vast. Every organism has a limited capacity of processing auditory information that is generally lower than the amount of acoustic information typically projected upon it. In general, an organism’s learning can be seen as the process of updating expectancy (prior understanding of the environment) in light of perceptual evidence of the world (observations) enabling it to learn (posterior understanding). Perhaps due to innate ecological listening principles, we spontaneously attribute auditory phenomena to causal actions (Chion 2009 p. 471; see also Tuuri & Eerola 2012; papers C and F). However, it is not sound itself that pertains to things in the world, but rather, we understand sound as evidence of action. As Thibaud puts it, “sound gives access to what is happening” (Thibaud 2011).

The starting point of systematic conceptualisations of sound in context was given by Schafer. He made several typological divisions of the sonic realm. One classification was by referential aspects (Schafer 1994 p. 137-148, 268-270) i.e. by *source*[8]: 'natural sounds', 'human sounds', 'sounds and society', 'mechanical sounds', 'quiet and silence', and 'sounds

[8] Sources might be veridical or imagined.

as indicators'. This propos was further developed by Krause (2008) who defined three classes of sound source: ‘geophony’ (sounds of nature except biology), ‘biophony’ (biology except humans), and ‘anthrophony’ (human-generated sounds,[9]). Another classification Schafer made was by significance (Schafer 1994 p. 26, 152, 271-275): ‘keynote’ (background reference), ‘signal’ (foreground communication), and ‘soundmark’ (cultural symbol). The two classifiers differ in that the first refers to physical objects in the world, and the second to attributed purpose as understood by humans. This distinction inspired the taxonomic approach in paper F. For Schafer, categorisation according to associated meaning was perhaps the most important (cf. Schafer p. 131, p. 137; see also Hellström 2003, p. 51). His groundbreaking work influenced followers to categorise prevalent sounds, mostly in regards to outdoor urban soundscapes (e.g. Payne et al. 2007; Axelsson et al. 2010, 2012; for an overview, see Kang 2010) and occasionally in servicescapes (see paper F: Introduction and Table 10, for a review of relevant previous studies).n

Synechdoche

Synechdoche is one of the major effects in CRESSON's framework and is the basis of perceptive selection. (Thibaud 2006 p. 124). It is the ability to observe a sonic element as a perceptual entity within a general flux of acoustic information, to evaluate and select it[10] (Thibaud 1995 p. 143; also “selective listening” Augoyard & Torgue 2006 p. xiv). The effect is produced via directed attention (cf. “denotative mode” of listening, Tuuri & Eerola 2012 p. 148; “listening-in-search” Truax 2001 p. 22; Schaeffer 1966 Livre II in particular p. 112 ff.; Bosch & Andringa 2014). A ‘sound event’ is the smallest self-contained part of a soundscape. Like Schaeffer’s ‘sound object’, it is a phenomenological object to which semantic meaning might be attributed, but the sound event is not a “laboratory specimen” and rather a “nonabstractable point of reference, related to a whole of greater magnitude than itself” (Schafer 1994 p. 274).

The selection of one element necessitates the omission of others (*asyndète*; Augoyard & Torgue 1995 p. 27). As one sonic element is lifted out from the context, it becomes the foreground; what remains, becomes the

[9] 'This term is used by Krause (2008 p. 77), though it has occasionally been (erroneously?) cited as 'antropophony'.

[10] “la faculté de d’opérer une sélection valorisant l’un ou l’autre élément”

background. The perception of a sound as an event implies a range of high-level brain processing, such as source recognition, working memory cognitive association, and emotive appraisal. On the other hand, the background might evoke no conscious response at all, as it is experienced in a more holistic way (Botteldooren et al. 2009 p. 193). However, foreground and background perception are intricately linked. Paper F shows that the cognitive appraisal of characteristic sound events is significantly correlated with the rated quality of the soundscape as a whole.

The sound event is understood as evidence of a physical event attributed to a causal action, and in this sense, synechdoche is fundamental to the sonic effect model. This mechanism creates the gap between the external acoustic reference and the internalised 'auditory object' within the listener's perceptual system, accessible via introspection (Pulkki & Karjalainen 2015 p. 133-135). The faculty of a sound to stand out from the whole, and be understood as an event, emphasises the permanence of the attributed source; it is then more likely to be remembered. Synechdoche organises the perception of time and enables the experience of duration. (Thibaud 2006 p. 125). The perception of a brief sensory event can be prolonged by working memory, which links between reflexive association and long-term memory (and selects what to store). Working memory is a special type of attentional process which enables the temporary and conscious holding of information (Mesulam 1998 p. 1032). Thibaud argued that language learning can be explained by synechdoche (compare this with the hypothesis that learning happens through neural adaptation over multiple exposures to relatively strong, repeated stimuli; cf. Huron 2006 ch. 4; also Clarke 2005). A sound event only exists as an element within the temporal and spatial flux of a soundscape. Sound cannot be experienced outside time: time is its very nature (Amphoux 2003). The synechdoche effect might be triggered by qualities in the acoustic signal itself (cf. acoustic descriptors, Kinsler et al. 1999 ch. 12), qualities in the receiver (cf. listening intention, Schaeffer 1966 ch. 8 especially p. 154), or in the interaction between the two. Interaction qualities are those that are likely to trigger crossmodal correspondences. Receiver qualities might interact with different kinds of acoustic qualities; for example, steady-state or dynamic psychoacoustic characteristics (e.g. a loud sound, an approaching sound; cf. Fastl & Zwicker 2007); or semantic characteristics (e.g. a recognised sound; Tuuri & Eerola 2012); or emotional or affective characteristics (e.g. sounds inspiring laughter, awe, or chills; cf. Huron 2006). See papers C and G.

A human being is nothing but a question. And the question is: beast or angel? Beast or God? You – you're obsessed with alarms and everything. What do you think you're doing? Aren't you trying to lift yourself – and everyone else – up from the bestiality of this awful noise to the heavenly silence? (Gruska's pillow talk monologue, in Noise (Bean et al. 2007))

4. Sound design

The latin word *designare* means 'to mark out'. This chapter gives a background and an overview of acoustic design, i.e. a practice-oriented demarcation of tools, skills, and elements that are enabling the artful shaping of acoustic spaces. The review of sound perception theories highlights several challenges for applied design work. Can 'positive sound sources' be designed? What makes a soundscape 'vibrant'? Can we determine the degree of 'purposefulness' of a sonic environment?

Scope

It is well-known that knowledge of auditory perception (Moore 2012; Schnupp et al. 2011) in everyday and specialised environments (Mehrabian & Russell 1974; Cohen et al. 1987; Barron 2010) provides keys to acoustic design (Apfel 1998; Asutay, Västfjäll et al. 2012; Kinsler et al. 1999; Culling JF 2013; Hellström 2003; Hellström, Dyrssen et al. 2011; Hellström 2012; Navarro & Pimentel 2006; Novak et al. 2010; see also papers E and F). Since sound is a temporal phenomenon, acoustic design differs from most other modalities, such as visual, architectural, written, and commercial (as in pricing) aspects, all of which are essentially static. Since practically all utilitarian designs serve social and other functions that develop over time (exchange, transport, work, relaxation, play and so forth), temporal dynamisms is often implicit in these other design modalities, but it is only explicit in the acoustic modality. For example, while the architecture of a public urban space might be documented in the form of a photographic image or in a written text, its functional, hedonistic, and aesthetic success must be judged on whether its fully multisensorial realisation facilitates the programmed activity or not. Therefore, acoustic design matters.

Schafer's starting point was that human activity defines the soundscape and that humans are responsible for it. He wrote that it is “no accidental byproduct of society; rather it is a deliberate construction by its creators, a composition which may be as much distinguished for its beauty as for its ugliness.” (Schafer 1994 p. 237). The choice of individual physical activities may not always reflect an awareness of their acoustic importance. As Truax continued, “the design of an environment need not be intentional; it may be implicit, or indirect, as with most traffic and aircraft noise when it constitutes a prominent feature of a soundscape. Whether the soundscape is natural or artificial, and its elements intentional or simply by-products of other processes, its quality of design depends on how it functions.“ How to achieve good design can be learnt from observation of “positively functioning soundscapes” (Truax 2001, p. 109; see also Davies et al. 2013).

Schafer has imagined the professional scope of an acoustic designer before it existed as an established practice. He saw sound design as a set of principles to be employed in adjudicating and improving soundscapes. The task, he wrote, is to "show society what it is missing, and only by applying passion and talent could [the designer] hope to be successful". Moreover, he compared the designer’s function in society with that of an artist, namely, “to open out new modes of perception and to portray alternative life styles". Schafer continues to say that the artist-designer must not expect to win popularity easily and that a pragmatism is essential, such as engaging in "very practical preservation and repair work” (Schafer 1994 p. 239). The acoustic designer must identify imbalances in the soundscape and address them in a systematic way, in order to be able to “reintroduce strong and exhilarating sounds” (Schafer 1994 p. 243).

Methods

To achieve the scope of work thus outlined, Schafer recommended that acoustic designers develop a critical mindset and gain control of listening skills; he called it ‘clairaudience’. Three things are needed for clairaudience: a respect for the ear and voice; awareness of sound symbolism; and knowledge of the natural soundscape and its balancing mechanisms (Schafer 1994 p. 238). It is an analytical tool, whereby the soundscape as a whole is disintegrated into its components (Schafer 1994

p. 133). It instrumentalises the perceptual principle of synechodoche (see ch. 3). Clairaudience training entails that the listener makes conscious cognitive effort to redirect the attention towards specific mechanisms of auditory perception, so as to be able to relearn his or her framework for cognitive appraisal of the soundscape (cf. Tuuri & Eerola 2012; Juslin & Västfjäll 2008; see also Hellström 2003, p. 37-38).

An overarching idea in this thesis is that successful research in sound relies on making connections between knowledge and know-how. In regards to sound design and sound perception, several essential skills and methods can be identified. Knowledge is built from theory, contextualisation, typology, taxonomy, and vocabulary (see papers B, E, and F for examples). Know-how comes from clairaudience, observation, measurement, synthesis, modelling, and validation (relevant in papers A, D, F, and G). An illustration of how these concepts are connected is shown in Fig. 11.

Acoustic design in restaurants

What can sound design in physical environments achieve? The example that will be discussed in some detail at this point regards restaurants, a type of servicescape, and a sonic environment that the thesis work has focussed on. The limitations placed on acoustic design in restaurants is important in order to frame the concept of *Plasticity* that is introduced in the next chapter.

After the term 'soundscape' had been introduced, the idea of 'servicescape' (Bitner 1992) could follow. Servicescapes are social places, either by necessity (hospital, school) or choice (restaurant, airport), that involve transaction. To facilitate human exchanges, the servicescape is an environment where considerable attention has been paid to multimodal design. Servicescapes engage all the senses and thereby provide rich opportunities to study sound perception in context.

Restaurants are servicescapes characterised by "elaborate physical complexity" (Bitner 1992, p. 58). The complexity is evident in that ambient environmental conditions affect the senses not only through physical factors such as temperature, lighting, noise, music, and scent, but also through psychological factors such as memory, appraisal, and "imagery" (Cohen et al. 1987, p. 172). Restaurant design is subject to

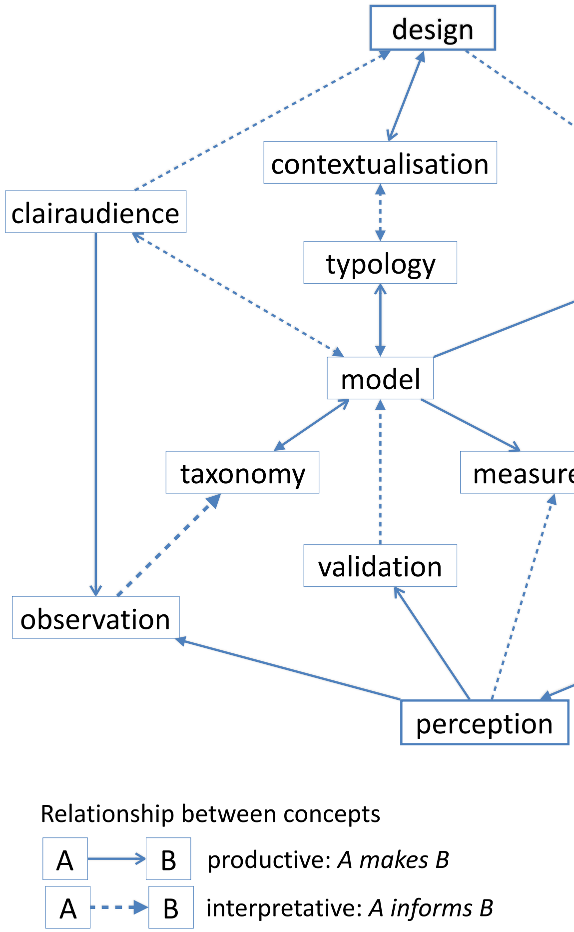

Figure 11. Concepts, methods, and tools underpinning sound design.

competing requirements. Factors such as visual aesthetics, ambience (including music), lighting, and service are measurable (Ryu & Jang 2008), and to varying degree designable.

Interior design can communicate cues to establish the character of a servicescape (Bitner 1992, p. 62; see also Frid 2013). In restaurants, a visual design might influence people's beliefs about the cuisine through colour schemes, light points, and artwork including printed menus. People can tell subconsciously whether a restaurant is ‘fast food’ or ‘diner’ just by looking at (and touching) interior design materials. Paper E showed that the proportion of stone, wood, plastic, metal, glass, and textile predicted both the 'design style' and 'food style' of restaurants. It is conceivable that a restaurant manager's pragmatic decision of which materials to use are heavily influenced by visual preferences. Yet such materials have very different acoustic performance, with implications for room acoustic. Therefore they have a secondary effect on soundscape quality which may or may not be intended. Another example is that sonic as well as visual factors affect people’s perception of the taste and flavour of food through crossmodal correspondences (Spence & Shankari 2010, Novak et al. 2010). Physical factors such as priciness and crowdedness are also associated with perceived quality and loudness (paper E).

Background music is probably the acoustic element that is the most obviously designed. It might be an effective way to project the “image” of a servicescape to prospective customers (Bitner 1992 p. 66; further developed in Ryu & Jang 2008). This effect relies on the psychological mechanism of evaluative misattribution (Västfjäll 2010 p. 26-269; paper F; see also Kahneman 2011 for examples). The relatively few systematic investigations of restaurant soundscapes have focussed on how music style (Wilson 2003; North & Hargreaves 1996) and sound level (Novak et al. 2010) influence customer behaviour and appraisal. The optimum stimulation level theory (Berlyne 1974; see also Axelsson 2011) links evaluative perception of a stimulus to its information rate (or complexity) via a ∩-shaped relationship.

Apart from the obvious background music, acoustic design elements generally influence the restaurant experience in subtle ways. For example, room acoustics influence speech intelligibility and vocal comfort (Rindel 2010). Paper F identified a range of specific sound sources, including glassware, screeching chairs, and the chatter of other customers, which were reliably appraised as positive ('liked') or negative ('disliked'). More than a thousand original annotations of characteristic sounds were analysed in order to create a multilevel taxonomy of sound sources in restaurants, that might be useful for soundscape simulation and field work. Paper E proposed a typology of acoustic design elements in restaurants. A primary distinction was made between ‘non-designed’ and ‘designed’ physical elements. Those in the first category were further classified according to origin (‘technological’, ‘human’, or ‘natural’; see ch. 3), and those in the latter were classified as either ‘active’ (i.e. music, masking noise) or ‘passive’ (enclosures, materials).

Acoustic designers must strive for variety and coherence, and cannot fixate on the "elimination" of certain sounds or promotion of "particularly interesting other ones", because design must "deal with the entire environment as a system of interactions between all elements” (Truax 2001 p. 113). Restaurant soundscapes are to a large extent determined by elements that can be classified and investigated, yet acoustic comfort is seldom prioritised. For example, being commercial spaces, restaurant attract customers by having large windows or open doors, preferably towards a busy street or shopping mall concourse (leaking in external noise). People generally react positively to an impression of cleanliness, so floors and tabletop materials are shiny (hard and acoustically reflective). Managers and waiters must observe multiple customers from a distance, so the plan design is open (no absorption of crowd noise and few lateral reflections from walls). Customers might want a ‘vibrant atmosphere’, so the music is selected to mix unobtrusively with the background din (a rationale for reverberant acoustics). Many restaurants are metabolic environments. Paper E showed that the sonic quality in the 112 surveyed restaurants was rated significantly worse than visual quality. It is a bleak picture, but the cloud is not without a silver lining. The contemporary situation represents an opportunity for acoustic designers. There is more room for improvement of the soundscape in restaurants than the corresponding visual environment. A reduction of ambient noise levels is likely to lead to higher overall quality ratings and might justify higher menu prices.

Any classification system or taxonomy is surrealistic...bringing together incongruous or anachronistic facts... The system used to organize such a vast number of designations will be arbitrary, for no sound has an objective meaning... The only framework inclusive enough to embrace all man's undertakings with equal objectivity is the garbage dump. (Schafer p. 133-137).

5. Multimodal environments

Environments might be understood by how they are perceived and how they are designed. The multimodal perception of an environment can be more or less complex. Complexity increases with the number of sensorial modalities containing uncorrelated information yet partaking in creating the whole; the degree of commonality and interaction between modalities. An environment can be more or less designable, or *plastic*[11]. It may communicate evidence of having been designed, or show promise of a potential to be reshaped.

Imagine a conceptual space mapped by dichotomies such as realistic–abstract, physical–virtual, pristine–built up, inviting–discouraging, place–non-place, and so forth. At one end are the 'non-designed' places found in natural wilderness, and at the other are the 'designed' soundscapes of immersive artworks. In the middle ground exist a multitude of complex environments that are 'semi-designed', such as urban parks and restaurants, places characterised by a mix of residual noise, metabolic sounds and specific place-characteristic sounds sources. They are characterised by a high degree of local unpredictability. Restaurants are semi-designed environments with particularly interesting soundscapes. Acoustic design elements in restaurants include but are not limited to music.

Complexity

Schafer wrote that "music is key to the utopian soundscape" (Schafer 1994 p. 237; see also ch. 4) and by this he meant that we should listen to environments as if they contained sounds that were intended, and organised as in a musical composition. Only with effort can listeners train themselves to hear city sounds as music, through conscious redirected attention and reshaping of their internalised appraisal framework. By modulating expectation, we can train ourselves to become constructively critical to the sonic environment. Yes, cities produce "dull murmurs,

[11] 'Plastic' in the sense the term *plasticity* is used in physics and sculptural arts

machine noise, the shifting and familiar racket created by people", but "every urban moment has a sound signature" (Augoyard & Torgue 2006 p. 4). The attribution of significance to sound events is key. It is the range and variation of sound events (cf. Truax 2001 p. 111) that listeners need to pay attention to, employing the appropriate listening strategy (Tuuri & Eerola 2012), and subjecting the event to appraisal. This goes back to Schaeffer's mode of objective and abstract listening which aims to 'grasp' the sound not only as an object but equally as a carrier of meaning (*comprendre*; Schaeffer 1966 p. 116), sometimes rendered as 'semantic listening'. A motivation for training oneself to "hear everything as a musical composition" (Vickers & Hogg 2006 p. 2015) is the reward: the transcendental freedom of being an active listener, rather than a passive 'hearer'. If the acoustic environment is a reservoir of sonic possibilities, then acoustic design elements constitute its instrumentarium (cf. Hellström 2003 p. 101). Every sound in the soundscape can be the object of analytic attention, care, and perhaps, passion. Urban listening goes beyond the reflexive 'like' or 'dislike'.

Sonic perception in everyday spaces need not be simplistic or defeatist. A contemporary urban listener wanting to promote herself from the naive experience (Augoyard & Torgue 2006 p. 13) of default background listening (Truax 2001) into the expert practise of willed switching between different modes of listening must first embrace the complexity of physical environments. While auditory perception is fundamentally one and the same in everyday and specialised listening, an understanding of sonic phenomena might happen through a "rediscovery of the pre-categorical approach to listening" (Augoyard & Torgue 2006 p. 13). This is in line with the embodied theory of meaning, which argues that creation of meaning is enabled by situating oneself in a flow of experiences (Tuuri & Eerola 2012 p. 144; see also Juslin 2013 on embodied cognition). For designers, it implies immersion in complex sonic environments and investigation of sound by consciously alternating between experiential, denotative, and reflective modes of listening. This perspective is discussed in papers F and G.

An individual immersed in an environment can evaluate the relationship between sensorial channels, by comparing information through all available channels, both sensorial stimulation in particular via auditory, visual, and tactile processing as well as internal processing such as memory (reflection) and imagination (projection). The relationship can take many forms. To use the metaphor of musical polyphony, it can be understood as a counterpoint of modalities and subjected to analysis. The degree to which modalities are intertwined and interdependent might be labelled 'sensorial *Complexity'*.

Plasticity

The sum total of sound events constituent of a soundscape may be perceived as more or less intended or purposeful for the sonically dependent functions, including in particular communication, restoration, and social activities that are taking place, or that a listener imagines could take place, in the environment. If the environment yields evidence of having been shaped with the explicit goal of facilitating these intents and purposes, we may refer to it as 'designed'. If the opposite is the case, we may refer to it as 'non-designed'. They are in reality the extreme ends of a continuum. The dimension along which a cognitive appraisal (specifically, aesthetic judgement) is made of perceived degree of intent (or purpose) might be labelled '*Plasticity'*. Three regions or levels can be identified.

- 'Designed' environments; techniques for soundscape design are deployed to full effect in an acoustically controlled situation. Sound is accorded a significant role to contribute to the experience as a whole, in terms of perceived quality and communicative intent. Examples include music, film, and multimedia performance (Wishart 1994, 1996; Chion 2009).
- 'Semi-designed' environments; evidence of a degree of attention to acoustic design. The soundscape as a whole contains both designed and non-designed elements. Examples include restaurants, other servicescapes, streetscapes, and urban parks (see ch. 4).
- 'Non-designed' environments; no clear intent or human communicative purpose behind sound events in the soundscape. This might happen either because there are simply no humans in the environment, or no effort has been made to shape the sounds so as to facilitate sonically dependent functions. Examples include pristine wilderness, underwater soundscapes, and certain extreme urban environments such as motorway underpasses. Note that these examples show that *Plasticity* as a

construct does not align with the "hi-fi vs. lo-fi" dichotomy introduced by Schafer (1994 p. 43).

Communication and meaning

While Schafer saw acoustic design as a tool for activism and practice, Truax (2001) focussed on a systematic investigation and demarcation of the parts and functions of soundscapes; in other words, acoustic design as the analysis of systems of communication. He identified three archetypal systems: *Soundscape, Music,* and *Speec*h. Each represents a specific way of organising sound, and each has its own purpose. The order in which he presented them is significant. Moving from left to right, he claimed that there is a decrease in the size of the acoustic repertoire (the number of different sound events); an increase in the strictness of the syntactical structure; and an increase in temporal density of information. To interpret it in another way: while all three archetypes of communication are meaningful in their respective context, meaning is expressed differently. The three might thus be ordered by a latent factor that could be conceptualised as the 'locus of meaningfulness', somewhere between detail and holistic scope. Truax claimed that "as we move towards the [soundscape side of] the continuum… meaning depends more and more on the relationship between elements, and between the elements and the whole…. The environmental sound signal (e.g. footstep, a bell, the wind, a car), whether foreground or background in perception, only acquires meaning through its context, that is, its complete relationship to the environment." (Truax 2001 p. 52-53). The spoken word, on the other hand, "can be stripped of its acoustic form in print, taken our of context, and still mean something, however incomplete."

The concept of 'communication between soundscape elements' is intriguing and demands a discussion. Communication is evident among birds and other animals; they can be attributed agency. But what about the flow of a river, or the flow of highway traffic? In what way can geophonies and other non-human environmental sounds be said to 'communicate'? Their sonic output is a by-product of some other process. Truax touches upon the question of whether a natural soundscape can be understood as designed (Truax 2001 p. 55), saying that it is a matter of hearing the soundscape as "a system of interrelated parts"; the coming together of an "acoustic community". The balance between variety and complexity of constituent sonic elements is the yardstick by which the naturalness of its design might be measured. Herein might be discerned a path to connecting the ideas of soundscape and communication. The degree to which any communication is intentional (this might very well be a necessary condition for us to speak of communication in the first place) aligns with the degree to which its acoustic ecology is designable, or *plastic,* as in possessing a potential of being shaped by human acoustic activity. At one extreme end is *Art*, which can be completely controlled and improved upon through design, and at the other is *Nature*, which cannot be controlled at all through design and cannot be improved upon [12].

In between the extremes of *Art* and *Nature*, there is *Servicescape*. In a conceptual continuum of environments along the dimension of *Plasticity*, it occupies the middle ground. Generally, the servicescape contains various elements, partly controlled and partly designed: *art* and *artifacts* (i.e. architecture, furniture, interior design, tools) as well as elements that are uncontrolled or non-designed: *nature* and *noise* (i.e. rain, machinery, movement of people). The mix of elements is characteristic of servicescapes, since they are environments for exchange, consumption, action, and production. When it comes to restaurant soundscapes, the situation is particularly complex, since both non-designed and partly designed acoustic elements come together. For example, art ≈ background music; artifacts ≈ table cloth and wall panels; nature ≈ birdsong through an open window; and broadband noise ≈ air conditioning.

Acoustic archetypes

Music, *Speech*, and *Soundscape* are systems of acoustic communication suggested by Truax, while *Nature*, *Servicescape*, and *Art* are archetypes of environments that can be interpreted along dimensions of *Plasticity* and sensorial *Complexity*. The six constituent concepts can be briefly outlined as follows:

- *Nature*: pristine wilderness, pelagic ocean, primal forest, benthic ocean, nature reserves, rural areas. Sound is present in all these environments, but they are only soundscapes when heard. Environments with human activity are more numerous than pristine wilderness. Nature is essentially multimodal yet extreme

[12] This is obviously stated in the spirit of Schafer.

environments might occasionally be perceived as having reduced modality, e.g. very silent environments (Antarctica, deserts, mountains, ocean).

- *Soundscape*: semi-designed urban spaces are either predominantly green (parks, mounds, boulevards) or grey (asphalt, concrete structures, building façades). Natural soundscapes and acoustic ecology go hand-in-hand.
- *Servicescape*: Semi-designed, predominantly social, multimodal design environment, including streetscapes, shops, lobbies, and restaurants. Sound is a critical modality that influences the perception of character, purpose, value, and activity of the servicescape as a whole. Servicescapes are multimodal.
- *Music*: musical instruments allow us to "refine the sounds of nature into a powerful form of human communication" (Truax 2001 p. 50-51). This includes both electroacoustic and acoustic composition. Soundscape composition is a form of electroacoustic art based on recordings of natural environments (Wishart 1994, 1996). It can be pure (selected, unaltered field recordings), mixed (including audio from other sources such as instruments), or synthetic (mimesis of natural phenomena). As music and soundscape composition calls for the sense of hearing alone, they essentially unimodal. [13]
- *Art*: temporal art creates highly designed environments, such as sound art, film, game, mobiles, dance, and music. Film and audiovisual performance are essentially bimodal expressions, fixed and interactive, respectively. Sound art often and additionally involves sculpture, thereby including the tactile sense to the auditive and visual.
- *Speech*: individual "voice soundmaking" is a highly designed mode of communication, by which a person's "concept of self and relationships to others, including the environment, are established" (Truax p. 34-35). Like music, it is primarily auditive and unimodal, though evidently it is closely connected to systems of visual signage i.e. written text and other visual symbols.

[13] Obviously, unimodality does not in the least imply simplicity.

How do environmental archetypes and communication systems relate? Since the dimensions of *meaning* (Truax' model) and *Plasticity* (as defined above) can be considered as parallel, the constituent parts of the acoustic communication systems and environmental archetypes can be placed along a single dimension, shifted so that they occupy intermediary positions. The complementarity of the two classification schemes is illustrated in Fig. 12.

The two schemes can be joined to form a single framework, a *Continuum of Acoustic Archetypes*. The framework expands upon Truax's communication model by interpreting the six concepts in a conceptual plane spanned by two dimensions, *Plasticity* and sensorial *Complexity*, which have been discussed earlier. Fig. 13 illustrates the continuum as a circle segment. Examples of actual environments are indicated at their approximate region within the continuous area within which all acoustic archetypes might be interpreted.

Discussion

In Fig. 12, an ellipse was drawn around *Music*, *Art*, and *Speech* to highlight the proximity of these concepts. Likewise, *Nature* and *Soundscape* are considered to be close concepts (while not ignoring their incongruencies). Servicescape straddles a middle ground, characterised by controlled as well as unintended sonic elements. Fig. 13 graphically indicates the scope of the seven papers included in this thesis within the conceptual continuum. The figure also illustrates subtypes within acoustical archetypes. For example, extreme wilderness and electroacoustic music, at either end of the *Plasticity* axis are considered to be of similar sensorial complexity, lower than that of restaurant servicescapes, since a description of the latter must necessarily include a large number of descriptors for complex social relations, including consumption, movement, and exchange.

It is through the process of contemplating the aims and results of the papers included in the thesis that the framework has gradually evolved. It is a greatly simplified model of actual, physical, and memorable sound in context. The framework has enabled comparison between diverse acoustic archetypes and has clarified their relationship.

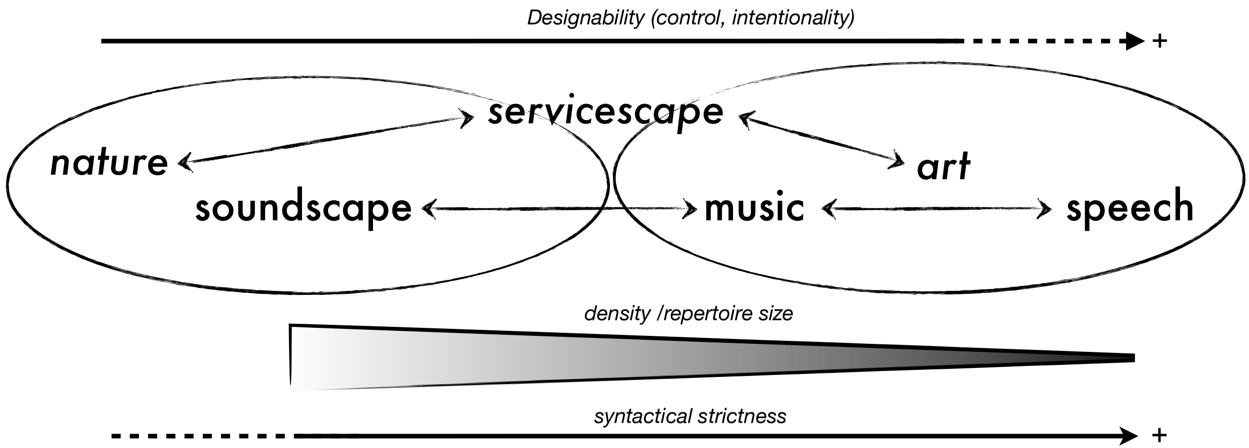

Figure 12. Illustration of Truax's three 'systems of Acoustic Communication' joined with three multimodal environments with different degree of *Plasticity*.

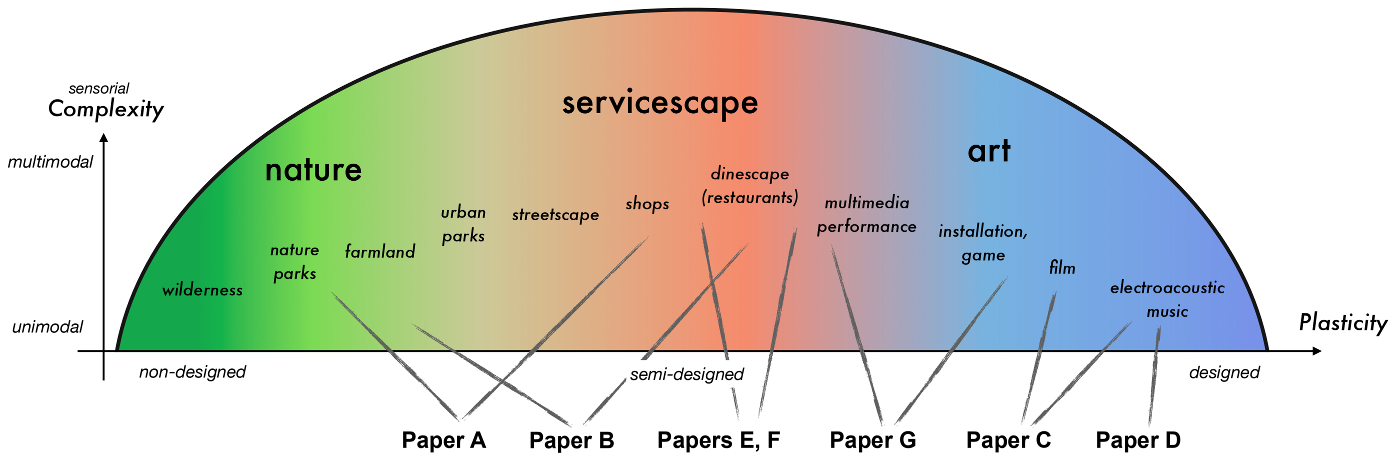

Figure 13. A continuum of multimodal environments ordered along dimensions of *Plasticity* and *Complexity*. The scopes of Papers A – G included in this thesis are indicated.

Introduction

The quarrel between Hume and Rousseau is symbolic: Rousseau was mad but influential, Hume was sane but had no followers.... The growth of unreason... is a natural sequel to Hume's destruction of empiricism [which] rests entirely upon his rejection of the principle of induction... To this extent, Hume has proved that pure empiricism is not a sufficient basis for science... [But what] these arguments prove... is, that induction is an independent logical principle... and that without this principle science is impossible. (Russel 1961 p. 646-647)

6. Conclusions and future work

The fundamental question addressed in this thesis was how people relate to sound. Answers have been sought by studying sound perception and sound design in context. The included papers synthesise conceptual and empirical research. The work has been carried out in a range of environments, from nature via restaurants to music, and with the use of different tools.

The three central chapters of the introduction brought together several wide-ranging concepts, enabling comparison and synthesis of otherwise isolated phenomena, practices, and theories. The outcome is a tentative theoretical framework describing a continuum of acoustic archetypes. The framework will need to be developed in the future and put to several tests, for example, as regards construct validity of the conceptual dimensions, parsimony of the model, and eventually, its usefulness for theory.

While the introductory section of the thesis aimed to be broad in scope, each of the seven papers that are included in the next section of the thesis has strived to be as detailed as possible. Each of the included papers investigated a demarcated research question, addressing different aspects of the central questioning of how people relate to sound. What they have in common, beyond the overall subject matter, are the four guiding principles of research, evidenced in each paper, namely that: knowledge grows from systematic observation; ecological validity promotes relevance; methods are complementary; and that art is a laboratory for sensory research.

The thesis has attempted to devote attention to both scope and detail. The challenge to achieve this within a single project is significant. Ultimately and inevitably, the task has become one of finding the right balance between the two approaches. There is a risk that in the end, neither attempt will have been fully satisfactory. Therefore, some of the many questionings that this thesis has brought up will be pursued in future work. In fact, since research is a process and not a goal in itself, some of the 'future work' is already well on its way.

By its very nature, research in multimodal perception necessitates a systematic approach where sensory modalities are compared two by two; this has to be iterated in each context and in several experimental declinations. The research in audiovisual crossmodal correspondences, specifically colour association with music (initiated with paper C), is ongoing. The goal of this research project is a robust model for colour-timbre association. The model might eventually be applicable in other contexts, such as the sonification project (see paper G) for concurrent audiovisual perceptualisation of data.

The investigation of restaurant environments and other semi-designed environments is being pursued. Visual material, already collected as part of earlier surveys and for example including photography, will be analysed and compared with published results which focussed on sound (e.g. papers E and F). The taxonomy of sound sources (presented in paper F) will be a crucial component in a 'restaurant soundscape simulator'. It might be extended with a visual counterpart, enabling sophisticated servicescape research in an interactive virtual audiovisual environment.

The question of why certain spaces conjure a sense of *place* calls for further probes in personality psychology (started in paper B). To understand how soundscape and music can be improved through acoustic design, both as efficiently well-functioning utilitarian environments and as idealistic aesthetic experiences, we may first have to improve our understanding of people. Only this way can we grasp the essence of sound and how it affects us.

Bibliography

Apfel RE (1998). *Deaf Architects & Blind Acousticians? A Guide to the Principles of Sound Design*. Apple Enterprises Press. New Haven, Conneticut.

Amphoux P, Moser G & Weiss K (2003). "L'identité sonore urbaine. Une approche méthodologique croisée". In Moser & Weiss (ed.) *Espaces de vie : aspects de la relation homme-environnement*, p. 201-220.

Amphoux P (2006) "Ubiquity". In Augoyard & Torgue (eds.) *Sonic Experience*. p. 130-145.

Amphoux P (1995). "Ubiquité". In Augoyard JF & Torgue H (eds.) *A l'écoute de l'environment. Répértoire des effets sonores*. p. 141-158.

Andringa TC (2010). "Soundscape and core affect regulation". *Proceedings of Interspeech*, Lisbon, Portugal, p. 1-10.

Asutay E, Västfjäll D, Tajadura-Jiménez A, Genell A, Bergman P & Kleiner M (2012). "Emoacoustics: A study of the psychoacoustical and psychological dimensions of emotional sound design." *Journal of the Audio Engineering Society* 60:1/2, 21-28.

Asutay E & Västfjäll D (2015, Jan.). "Attentional and Emotional Prioritization of the Sounds Occurring Outside the Visual Field". *Emotion*. URL http://dx.doi.org/10.1037/emo0000045. (31 Oct. 2015).

Augoyard JF & Torgue H (eds.) (1995). *A l'écoute de l'environment. Répértoire des effets sonores*. Editions paranthèses, France.

Augoyard JF & Torgue H (eds.) (2006). *Sonic Experience. A Guide to Everyday Sounds*. McGill-Queen's University Press, Quebec, Canada.

Axelsson Ö (2011). *Aesthetic Appreciation Explicated* [Thesis]. Stockholm University.

Axelsson Ö, Nilsson ME & Berglund B (2012). "The Swedish soundscape-quality protocol". *Journal of the Acoustical Society of America*, 131(4), 3476-3476.

Axelsson Ö, Nilsson ME & Berglund B (2010). "A principal components model of soundscape perception". *Journal of the Acoustical Society of America*, 128(5), 2836-2846.

Barron M (2010). *Auditorium Acoustics and Architectural Design*. 2nd edition. Spon Press.

Bauer MW & Gaskell G (eds.) (2000). *Qualitative Researching with Text, Image, and Sound*. Sage Publications, London.

Bean H, Hoffman P, Robbins T et al. (2007). *Noise*. [Film]. Seven Arts /Fuller Films. Anchor Bay Entertainment, USA.

Bech S & Zacharov N (2006). *Perceptual Audio Evaluation. Theory, Method, and Application*. John Wiley & Sons, Ltd.

Belojevic G, Jakovljevic B & Slepcevic V (2003). “Noise and mental performance : Personality attributes and noise sensitivity”. *Noise & Healt*h, 6:21, 77-89, http://www.noiseandhealth.org/text.asp?2003/6/21/77/31680 (31 Oct. 2015).

Berglund B & Nilsson ME (2007). Summary of the studies in soundscape perception, p. 14-23 in Spång, K.: *Soundscape Support to Health*. URL http://www.ljudlandskap.acoustics.nu/downloads/publikationer/ljudlandskap/slutrapporter/Summary+of+studies+and+results.pdf (31 October 2015).

Berlyne DE (1974). *Studies in the new experimental aesthetics: Steps toward an objective psychology of aesthetic appreciation*. Hemisphere.

Bitner MJ (1992). "Servicescapes: The Impact of Physical Surroundings on Customers and Employees". *Journal of Marketing* 56:2, 57-71.

Bosch KA & Andringa TC (2014). “The effect of sound sources on soundscape appraisal.” *Proc. 11th International Congress on Noise as a Public Health Problem* (ICBEN), Nara, Japan.

Botteldooren D, de Coensel B, van Renterghem T, Dekoninck L & Gillis D (2009). "The urban soundscape: a different perspective." Ghent University, Department of Information Technology, Acoustics Group, Ghent.

Brumm H & Zollinger SA (2011). "The evolution of the Lombard effect: 100 years of psychoacoustic research". *Behaviour*, 148(11-13), 1173-1198.

BS/ISO (2014). *Standard 12913-1 on Acoustics. Soundscape. Part 1: Definition and conceptual framework*. ICS Classification ISBN 978-0-580-78309-8.

Cabrera D (2014). *Psysound3* (Jan. 2014).URL http://www.densilcabrera.com/wordpress/psysound3/ (31 Oct. 2015).

Chelkoff G (2006). “Metamorphosis”. In Augoyard JF & Torgue H (eds.) *Sonic Experience*. p. 73-78.

Chelkoff G (1995). “Métabole”. In Augoyard JF & Torgue H (eds.) *A l’écoute de l’environment*. p. 86-91.

Chion M (2009). *Film, A Sound Art*. Columbia University Press.

Clarke EF (2005). *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford University Press.

Coan JA & Allen JJB (eds.) (2007). *Handbook of Emotion Elicitation and Assessment*. Oxford university press.

Cohen AJ (2010). “Music as a source of emotion in film”. Chapter 31 p. 879-908 in Juslin P & Sloboda J (eds.) *Handbook of Music and Emotion*. Oxford University Press.

Cohen AJ, Campanella A, Marshall L & Grant C (1987). “Perspectives on acoustics in environmental design.” *Journal of Architectural and Planning Research*, 4:2, 162-179.

CRESSON [Le Centre de recherche sur l'espace sonore et l'environnement urbain]. URL http://www.cresson.archi.fr/ (31 Oct. 2015).

Culling JF (2013). “ Energetic and Informational Masking in a Simulated Restaurant Environment”. In Moore BCJ (ed.) *Basic Aspects of Hearing, Advances in Experimental Medicine & Biology*. Springer Science+Business Media, New York.

Davies WJ & Adams, Bruce, Cain, Carlyle, Cusack, Hall, Hume, Irwin, Jennings, Marselle, Plack & Poxon (2013). “Perception of soundscapes: An interdisciplinary approach”. *Applied Acoustics* 74, 224–231.

Everitt BS & Hothorn T (2010). *A Handbook of Statistical Analyses Using R*. 2nd edition. CRC Press.

Eerola T & Vuoskoski J (2011). “A comparison of the discrete and dimensional models of emotion in music”. DOI: 10.1177/0305735610362821. *Psychology of Music* 39:1, 18–49.

Fastl H & Zwicker E (2007). *Psychoacoustics: Facts and Models*. 3rd ed. Springer-Verlag, Germany.

Friberg A, Schoonderwaldt E, & Hedblad A, Fabiani M & Elowsson A (2014). “Using listener-based perceptual features as intermediate representations in music information retrieval.” *Journal of the Acoustical Society of America*, 136(4), 1951-1963.

Friberg A (2012). "Music Listening from an Ecological Perspective". Poster, *Proc. ICMPC–ESCOM*, Thessaloniki, Greece.

Frid E (2013). *Restaurant Soundscapes in Stockholm*. (Project Report), KTH Royal Institute of Technology, Stockholm.

Fodor JA (1998). "The Language of Thought". In Bechtel & Graham (eds.), *A Companion to Cognitive Science*. Blackwell Publishers.

Gaarder J (1991/1997). *Sofies verden*. Aschehoug, Oslo. My translation. The cited passage has been paraphrased[14].

Gabrielsson A & Lindström E (2010). “The role of structure in the musical expression of emotions”. Chapter 14 p. 367-400 in Juslin PN & Sloboda JA (eds.) *Handbook of Music and Emotion*. Oxford University Press.

Gaver WW (1993). “How Do We Hear in the World? Explorations in Ecological Acoustics”. *Ecological Psychology*, 5:4, 285-313, DOI: 10.1207/s15326969eco0504_2.

[14] Original: “Jada, Sofie, jeg skal forklare med nærmere. En enkelt pepperkakemann kan være så misslykket etter all bakingen, hevingen og stekingen at det ikke er godt å si nøyaktig hva den forestiller. Men etter å ha sett 20-30 slike pepperkaker – som alltså kan være mer eller mindre perfekte – kan jeg med stor sikkerhet vite hvordan kakeformen ser ut. Dette kan jeg slutte meg til selv om jeg aldri har sett selve formen."

Hellquist E (1922). *Svensk etymologisk ordbok*. Gleerups förlag. URL http://runeberg.org/svetym/ (31 Oct. 2015).

Hellström B (2012). *Om kontorslandskapens akustik & arkitektur. Vad örat hör men ögat inte ser*. Arkus #68.

Hellström B (2011). "The Sonic Identity of European Cities". *Soundscape Studies and Methods*. The Finnish Society for Ethnomusicology Publ., no 9 (2002): 59-82. URL http://europeanacousticheritage.eu/wp-content/uploads/2011/09/The-Sonic-Identity-of-European-Cities1.pdf (31 Oct. 2015).

Hellström (2003). *Noise Design. Architectural Modelling and the esthetics of Urban Acoustic Space*. (Doctoral thesis). Bo Ejeby Förlag.

Hellström B, Dyrssen C, Hultqvist A, Mossenmark S & Sjösten P (2011). "Modelling the shopping soundscape". *Journal of Sonic Studies*, 1. URL http://journal.sonicstudies.org/vol01/nr01/a04 (31 Oct. 2015).

Hermann T, Hunt A & Neuhoff JG (ed.) (2011). *The Sonification Handbook*. COST Office and Logos Verlag Berlin GmbH. Available at http://sonification.de/handbook/. acc. 30 Oct. 2015.

Hofstadter DR (1982). *Metamagical Themas : Questing for the Essence of Mind and Pattern*. BasicBooks.

Howell DC (2010). *Statistical methods for psychology*. Intl. edition. Cengage Wadsworth.

Huron D (2006). *Sweet anticipation. Music and the Psychology of Expectation*. MIT Press.

John OP & Srivastava S (1999). "The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives". In Pervin & John (eds.) *Handbook of Personality. Theory and Research*. 2nd ed, Guilford Press, USA, pp. 102-138.

Joo YH & Jin YJ (2015). "Influence of urban contexts on soundscape perceptions: A structural equation modeling approach". *Landscape and Urban Planning* 141, 78–87.

Juslin PN, Gonçalo B & Eerola T (2015). "From Sound to Significance: Exploring the Mechanisms Underlying Emotional Reactions to Music." *The American journal of psychology*, 128.3, 281-304.

Juslin PN (2013)."From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions." *Physics of Life Reviews* 10.3, 235-266. URL http://www.sciencedirect.com/science/article/pii/S1571064513000638 (31 Oct. 2015).

Juslin PN, Liljeström S, Västfjäll D & Lundqvist LO (2010). "How does music evoke emotions? Exploring the underlying mechanisms". Chapter 22 p. 603-642 in Juslin P & Sloboda J (eds.) *Handbook of Music and Emotion*. Oxford University Press.

Juslin PN & Sloboda JA (eds.) (2010). *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford University Press.

Juslin PN & Västfjäll D (2008). "Emotional responses to music: The need to consider underlying mechanisms". *Behavioural and Brain Studies*, 31, 559 – 621. URL http://www.sciencedirect.com/science/article/pii/S1571064513000638 (a31 Oct. 2015).

Kahneman D (2011). *Thinking, Fast and slow*. Penguin Books.

Kang J (2010, Nov.). “From understanding to designing soundscapes”. *Frontiers of Architecture and Civil Engineering in China,* 4:4, 403-417.

Kinsler LE, Frey AR, Coppens AB & Sanders JV (1999). *Fundamentals of Acoustics*. 4th edition. Wiley.

Kreibig SD (2010). “Autonomic nervous system activity in emotion: A review”. *Biological Psychology*, 84, p. 394–421.

Kuhn T (1962). *The Structure of Scientific Revolutions*. 2nd edition. The University of Chicago. The citation was paraphrased[15].

Kabacoff R (2011). *R in action. Data analysis and graphics with R*. Manning.

Krause B (2008). "Anatomy of the Soundscape". *Journal of the Audio Engineering Society* 56:1/2, 73-80.

Lartillot O et al. (2015). *MIRtoolbox*. URL https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/ (31 Oct. 2015).

Lartillot O, Toiviainen P, & Eerola T (2008). "A matlab toolbox for music information retrieval". In *Data analysis, machine learning and applications*, p. 261-268. Springer Berlin Heidelberg.

LeDoux J (1998). *The Emotional Brain*. Weidenfeld & Nicolson. NY, USA.

Lazarus RS & Folkman S (1984). "Coping and adaptation". In Gentry (ed.) *Handbook of Behavioral Medicine,* Guilford Press, New York, pp. 282-325.

Leroi-Gourhan A (1964). *Le geste et la parole, vol. 1 & 2*. Editions Albin Michel.

Lindborg PM (2003). *Le dialogue musicien-machine : Aspects des systèmes d'interactivité musicale*. (Mémoire de DEA de musique et musicologie du XXe siècle). Marc Battier, dir.). Université de Paris IV Sorbonne. URL http://www.permagnus.org/pm/research/DEA/ (31 Oct. 2015).

Loy G (2006). *Musimathics, vol. 1 & 2*. MIT Press.

Mehrabian A & Russel JA (1974). *An Approach to Environmental Psychology*. MIT Press.

[15] Original: “Effective research scarcely begins before a scientific community thinks it has acquired firm answers to questions like the following: What are the fundamental entities of which the universe is composed? How do these interact with each other and with the senses? What questions may legitimately be asked about such entities and what techniques employed in seeking solutions?"

Mesulam MM (1998). “From sensation to cognition”. *Brain*, 121, 1013–1052.

Miller SS (2012). "Words on Sounds – Listening: Modes and Strategies". URL https://sightssoundswords.wordpress.com/2012/12/31/words-on-sounds-listening-modes-and-strategies/ (31 Oct. 2015).

Nahata A (2015). Entry in Q&A. URL https://www.researchgate.net/post/Exact_difference_and_relation_between_cognition_and_perception (31 Oct. 2015).

Navarro NPN & Pimentin RL (2006). “Speech interference in food courts of shopping centres”. *Applied Acoustics* 68, 364–375.

North AC & Hargreaves DJ (1996). “The effects of music on response to a dining area”. Journal of Environmental *Psychology*, 16, 55–64.

Novak CC, La Lopa J & Novak RE (2010). “Effects of Sound Pressure Levels and Sensitivity to Noise on Mood and Behavioral Intent in a Controlled Fine Dining Restaurant Environment.” *Journal of Culinary Science & Technology*. Taylor & Francis Group. 8:191-218.

Moore BCJ (2012). *An Introduction to the Psychology of Hearing*. 6th ed. Emerald Group Publishing, UK.

Moore JK & and Linthicum FH (2011). “Auditory System”. Chapter 34 in Mai JK & Paxinos G (eds.) *The human nervous system*. 2nd edition. Academic press.

Nilsson ME (2007b) "A-weighted sound pressure level as an indicator of short-term loudness or annoyance of road-traffic sound." *Journal of Sound and Vibration* 302.1, 197-207.

Nilsson ME (2007a). “Soundscape quality in urban open spaces”. *Proc. Inter-Noise*, Istanbul, Turkey. Inst. Noise Control Engineering, Poughkeepsie, New York, USA, pp. 1-7.

Osgood CE, George JS & Tannenbaum PH (1957). *The measurement of meaning*. Urbana, University of Illinois Press.

Payne S (2013). “The production of a Perceived Restorativeness Soundscape Scale”. *Applied Acoustics* 74, 255–263.

Payne SR & Guastavino C (2013, Sep.). “Measuring the perceived restorativeness of soundscapes: is it about the sounds, the person, or the environment?" *Proc. Inter-Noise*, Innsbruck, Austria.

Payne SR, Devine-Wright P & Irvine KN (2007, Aug.). "People’s perceptions and classifications of sounds heard in urban parks: semantics, affect and restoration". *Proc. Inter-Noise*, 36th International Congress and Exhibition on Noise Control Engineering, Istanbul, Turkey.

Peretz E (2010). “Towards a neurobiology of musical emotions”. Chapter 5 p. 99-126 in Juslin P & Sloboda J (eds.) *Handbook of Music and Emotion*. Oxford University Press.

Premack D (2007). "Human and animal cognition: Continuity and discontinuity". *Proceedings of the National Academy of Sciences*, 104(35), 13861-13867.

Pulkki V & Karjalainen M (2015). *Communication Acoustics*. John Wiley & Sons Inc.

Rajala AZ, Reininger KR, Lancaster KM & Populin LC (2010). "Rhesus monkeys (Macaca mulatta) do recognize themselves in the mirror: implications for the evolution of self-recognition". *PLoS One*, 5(9), e12865.

Rindel JH (2010). "Verbal communication and noise in eating establishments". *Applied Acoustics,* 71:12, 1156-1161.

Russell JA (1979). "Affective space is bipolar". *Journal of Personality & Social Psychology* 37(3), 345-356.

Russell JA & Mehrabian A (1977). "Evidence for a three-factor theory of emotions. *Journal of Research Personality,* 11(3), 273–294.

Ryu K & Jang S (2008). "DINESCAPE: a scale for customers' perception of dining environments". *Journal of Foodservice Business Research*, 11(1), 2-22.

Schafer RM (1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Destiny Books, USA, pp. 1-320.

Schaeffer P (1966). *Traité des objets musicaux. Essai interdisciplines*. Nouvelle édition. Seuil.

Schnupp J, Nelken I & King A (2011). A*uditory neuroscience: Making sense of sound*. MIT Press.

Shepherd D, Heinonen-Guzejev M, Hautus MJ & Heikkilä K (2015). "Elucidating the relationship between noise sensitivity and personality". *Noise & Health* 17, 168-171.

Spence C (2011). "Crossmodal correspondences: A tutorial review". *Atten. Perceptual Psychophysiology* 73, 971–995 DOI 10.3758/ s13414-010-0073-7.

Spence C & Shankari MU (2010). "The influence of auditory cues on the perception of, and responses to, food and drink". *Journal of Sensory Studies* 25, 406–430. DOI: 10.1111/j.1745-459X.2009.00267.x

Stansfeld SA, Haines MM, Burr M, Berry B & Lercher P (2000). "A review of environmental noise and mental health". *Noise Health*, 2:1-8. http://www.noiseandhealth.org/text.asp?2000/2/8/1/31756 (31 Oct. 2015).

Thibaud JP (2011, Oct.). "A sonic paradigm of urban ambiances". *Journal of Sonic Studies,* 1:1. URL http://journal.sonicstudies.org/vol01/nr01/a02 (31 Oct. 2015).

Thibaud JP (2006). "Synechdoche". In Augoyard & Torgue (eds.) *Sonic Experience*. p. 123-129.

Thibaud JP (1995). “Synechdoque”. in Augoyard JF & Torgue H (eds.) *A l’écoute de l’environment. Répértoire des effets sonores*. p. 134-140.

Tranströmer T (1984). "The Wild Market Square / Det vilda torget". In *The great enigma, new collected poems*. Translated by Robin Fulton. New Directions Books, New York 2006.

Truax B (1999). *Handbook for Acoustic Ecology*. URL http://www.sfu.ca/sonic-studio/handbook/ (acc. 31 Oct. 2015).

Truax B (2001). "Handbook of acoustic ecology (CD-ROM version)". *Computer Music Journal* 25, 93-94.

Truax B (2001). *Acoustic Communication*. Greenwood Publishing Group, USA. pp. 1-284.

Tuuri K & Eerola T (2012). “Formulating a Revised Taxonomy for Modes of Listening”. *Journal of New Music Research*, 41(2), 137-152.

Vickers P & Hogg B (2006), "Sonification abstraite/sonification concrete: An ‘Æsthetic perspective space’ for classifying auditory displays in the ars musica domain." *Proc. International Conference on Auditory Display*, p. 210-216.

Västfjäll D (2010). “The role of structure in the musical expression of emotions”. Chapter 10 p. 255-277 in Juslin P & Sloboda J (eds.) *Handbook of Music and Emotion*. Oxford University Press.

Västfjäll D.(2003). "The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments". *CyberPsychology & Behavior*, 6:2, 181-188.

Västfjäll D, Friman M, Gärling T & Kleiner M (2002). "The measurement of core affect: A Swedish self–report measure derived from the affect circumplex". *Scandinavian Journal of Psychology*, 43:1, 19-31.

Wilson S (2003). “The effect of music on perceived atmosphere and purchase intentions in a restaurant”. *Psychology of Music* 31, 93-112.

Wishart T (1996). *On Sonic Art*. Psychology Press.

Wishart T (1994). *Audible Design*. York, Orpheus the Pantomime.

World Soundscape Project [WSP]. URL http://www.sfu.ca/~truax/wsp.html (31 Oct. 2015).

Wrightson K (2000). "An introduction to acoustic ecology". *Soundscape: The journal of acoustic ecology*. 1:1, 10-13.

Yu L & Kang J (2010). "Factors influencing the sound preference in urban open spaces". *Applied Acoustics* 71 (2010) 622-633.

Papers