Timbre 2018: Timbre is a many splendored thing

Tanna Schulich Recital Hall
Elizabeth Wirth Music Building
Schulich School of Music
McGill University
Montréal, Québec

4 – 7 July 2018

Program & Abstracts

SSHRC CRSH
Canada

McGill
Schulich School of Music
École de musique Schulich

CIRMMT

Published by:
The Schulich School of Music
McGill University
555 rue Sherbrooke ouest
Montréal, QC
Canada H3A 1E3

This conference was made possible with the generous support of the Social Sciences and Humanities Research Council of Canada, the Schulich School of Music of McGill University, and the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT).

Citation:
In E. Thoret, M. Goodchild & S. McAdams (Eds.), *Timbre 2018: Timbre Is a Many-Splendored Thing*, Montreal, QC: McGill University

ISBN 978-1-77247-013-0

Organizing Committee

Stephen McAdams	*Chair*
Meghan Goodchild	*Co-chair*
Etienne Thoret	*Paper chair*
Moe Touizrar	*Musical event chair*
Kit Soden	*Tutorial/Round table chair*
Hannah Darroch	*Logistics and managerial support*
Aliza Thibodeau	*Logistics and managerial support*
Manda Fischer	*Graphic design*

International Evaluation Committee

Martha de Francisco	*McGill University (Sound recording)*
Philippe Depalle	*McGill University (Signal processing)*
Emily Dolan	*Harvard University (Musicology)*
Mounya Elhilali	*Johns Hopkins University (Neuroscience)*
Philippe Esling	*IRCAM/Université Pierre et Marie Curie (Computer science)*
Cornelia Fales	*Indiana University (Ethnomusicology)*
Robert Hasegawa	*McGill University (Music theory)*
Stephen McAdams	*McGill University (Music psychology)*
Luis Naón	*Haute école de musique de Genève (Composition)*
John Rea	*McGill University (Composition)*
Friedemann Sallis	*University of Calgary (Musicology)*
Zachary Wallmark	*Southern Methodist University (Popular music studies, Music psychology)*
Stefan Weinzierl	*Technische Universität Berlin (Acoustics)*

WELCOME FROM THE DEAN OF THE SCHULICH SCHOOL OF MUSIC

It is my great pleasure to welcome you to Montreal, Quebec, and to the Schulich School of Music of McGill University for the 2018 conference, *Timbre Is a Many-Splendored Thing*. As a hub of cultural activity, especially in the summer months, Montreal offers the ideal environment to bring together accomplished researchers from across the globe to discuss, discover, and examine the place and power of timbre in music.

The Schulich School of Music is an international leader in professional music training and research, with critically acclaimed programs in classical performance, opera, jazz, early music, and contemporary music. The School is also a pioneer in the humanities, sound recording, and music technology, and is committed to studying musical language as part of a shared social and cultural fabric. An international conference dedicated to timbre enables a collaborative investigation of the distinct character of this intriguing musical element. To host such a gathering is our privilege.

The Schulich School of Music is proud of the work Dr. Stephen McAdams has done in bringing together experts from the areas of music theory and analysis, ethnomusicology, psychology, technology, composition, and digital humanities. This event offers the opportunity for us to transcend disciplines as we explore the significance of timbre, and enhance awareness of its role in music composition, scholarship, science, and performance. My hope is that you all leave here having gained new insights, made new connections, and inspired to continue to shape the future of this important field of study.

Brenda Ravenscroft, PhD
Dean, Schulich School of Music of McGill University

McGill   Schulich School of Music
École de musique Schulich

SCHEDULE

WEDNESDAY | 4 JULY 2018

Library Seminar Room A-510
enter library on 3rd floor of Elizabeth Wirth Music Building

TUTORIALS AND ROUND TABLES

09:30 Tutorial on Computer-Aided Orchestral Simulation: **DENYS BOULIANE & FÉLIX BARIL** — OrchPlay (OrchSim): The orchestra in your classroom and research center! A new software for pedagogy and research

11:30 **BREAK**

12:00 Tutorial on Computer-Aided Orchestration: **PHILIPPE ESLING, CARMINE EMANUELE CELLA, LÉOPOLD CRESTEL, AXEL CHEMLA-ROMEU-SANTOS & MATHIEU PRANG** — The computer as an orchestra

14:00 **BREAK**

15:30 Round table: **ROBERT FINK, CORNELIA FALES, MELINDA LATOUR, CATHERINE PROVENZANO, PAUL THÉBERGE, STEVE WAKSMAN, ZACHARY WALLMARK, SIMON ZAGORSKI-THOMAS** — *The Relentless Pursuit of Tone*: A round table on timbre, technology, history, and genre in popular music

17:00 **BREAK**

17:30 Round table: **JOHN REA, VICTOR CORDERO, FABIEN LÉVY, LAURIE RADFORD, LASSE THORESEN** — Orchestration pedagogy in the 21st century

19:00 **END**

REGISTRATION AND OPENING RECEPTION @ ELIZABETH WIRTH MUSIC BUILDING LOBBY

18:00 **REGISTRATION**

19:00 **OPENING RECEPTION**

21:00 **END**

THURSDAY MORNING | 5 JULY 2018

Tanna Schulich Recital Hall, ground floor of Elizabeth Wirth Music Building

07:30 **COFFEE AND REGISTRATION**

08:00 **OPENING ADDRESSES**

LLOYD WHITESELL, Assoc. Dean, Schulich School of Music, McGill University – Welcoming address

STEPHEN MCADAMS, McGill University – 'What is timbre?' vs. 'What can we do with timbre?' Picking the right questions

RESEARCH KEYNOTE 1 – MUSICOLOGY

08:45 **EMILY DOLAN**, Harvard University – Timbre eternal

TALK SESSION 1: CHAIR LLOYD WHITESELL

09:45 **HEATHER BUFFINGTON-ANDERSON** – Black power voices: Examining timbre and polyvocality in Oscar Brown, Jr. and Nina Simone

10:10 **REBECCA FLORE –** The social life of timbre: Discussing sound color in online guitar communities

10:35 **ISABELLA VAN ELFEREN –** Timbre: Aesthetics of vibration

11:00 **COFFEE BREAK**

TALK SESSION 2: CHAIR LLOYD WHITESELL

11:30 **FEDERICA DI GASBARRO** – Timbre and formal functions in Edgard Varèse's *Amériques*

11:55 **PASCAL DECROUPET** – Phrasing through timbre: Timbral variety, figural coherence and formal evidence in Stockhausen's *Kontakte* and Ferneyhough's *Time and Motion Study II*

12:20 **DAVID BLAKE & ALESSANDRO BRATUS –** Timbral syntax in Björk's *Utopia*

12:55 **POSTER SESSION 1 & LUNCH**

CHRISTOPH REUTER, ISABELLA CZEDIK-EYSENBERG, SALEH SIDDIQ & MICHAEL OEHLER – The closer the better: The role of formant positions in timbre similarity perception and timbre classification

MATTHEW ZELLER – Timbral function in Klangflächetechnik

SONG HUI CHON & SUNGYOUNG KIM – Understanding the perception of acoustic space from the perspectives of timbre

SIMON WALOSCHEK, ARISTOTELIS HADJAKOS & AXEL BERNDT – Score and sound-based interface for exploring music performances

SERGIO GIRALDO, RAFAEL RAMIREZ, GEORGE WADDELL & AARON WILLIAMON – Computational modeling of new timbre dimensions for automatic violin tone quality assessment

ETIENNE RICHAN & JEAN ROUAT – Towards translating perceptual distance from hearing to vision using neural networks

MOE TOUIZRAR – Sound signifying light: The sunrise topic as a transmodal perceptual metaphor

MARGARET SCHEDEL & ERIKA HONISCH – New wor(l)ds for old sounds

LAURENT POTTIER – Local sound signatures for music recommendations

SALEH SIDDIQ, CHRISTOPH REUTER, ISABELLA CZEDIK-EYSENBERG & DENIS KNAUF – The (obvious) influence of pitch and dynamics on the perceptual dimensions of the timbre of musical instruments

MATTSON OGG & L. ROBERT SLEVC – An investigation of the acoustic features supporting the perception of timbre and other auditory objects

ROBERT NORMANDEAU – Timbre spatialisation

JACK KELLY, JULIAN NERI, DIEGO QUIROZ & ALEJANDRO ASPINWALL – Communication through microphone choice and placement

Thursday Afternoon | 5 July 2018

Musical Event 1

14:45 Bruno Deschênes – The tone-color melodies of the Japanese shakuhachi

Talk Session 3: Chair Philippe Esling

15:25 Kai Siedenburg – Testing relative perception and context-sensitivity of timbral brightness

15:50 Charalampos Saitis & Kai Siedenburg – Exploring the role of source-cause categories in timbral brightness perception

16:15 Coffee break

Talk Session 4: Chair Philippe Esling

16:45 Etienne Thoret, Baptiste Caramiaux, Philippe Depalle & Stephen McAdams – A computational meta-analysis of human dissimilarity ratings of musical instrument timbre

17:10 Sven-Amin Lembke – Measuring temporal interdependence between duet performers along timbre, dynamics, and pitch

17:35 Jonathan Berger, Talya Berger, Eoin Callery, Elliot Kermit-Canfield & Jonathan Abel – Timbre, texture, space and musical style: The interplay of architecture and music in Rome's Chiesa di Sant'Aniceto

18:25 Continuation of Poster Session 1

Musical Keynote 1 – Koto

20:00 ANDO Masateru, Kisoukai Koto Association — Perception and cognition of playing the koto

20:20 ANDO Masateru & ANDO Tamaki — The world of the koto

21:20 End

Friday Morning | 6 July 2018

Tanna Schulich Recital Hall, ground floor of Elizabeth Wirth Music Building

08:00 Coffee

Research Keynote 2 – Cognitive neuroscience

08:45 Vinoo Alluri, International Institute of Information Technology, Hyderabad — Timbre in the brain

Talk Session 5: Chair Zachary Wallmark

09:45 James O'Callaghan – Rupture of timbre and source-bonding as musical aesthetic

10:10 Stephen Spencer – Arthur Lange and the spectrotone system of orchestration

10:35 Laurie Radford – Locating timbre: Challenges of timbral design in multichannel electroacoustic music

11:00 Coffee break

Talk Session 6: Chair Zachary Wallmark

11:30 **Carmine Emanuele Cella & Philippe Esling** – Open-source modular toolbox for computer-aided orchestration

11:55 **Léopold Crestel, Philippe Esling, Daniele Ghisi & Robin Meier** – Generating orchestral music by conditioning SampleRNN

12:20 **Tristan Carsault, Axel Chemla-Romeu-Santos & Philippe Esling** – Learning spectral transforms to improve timbre analysis

12:55 **Poster Session 2 & Lunch**

- **François-Xavier Féron, Catherine Guastavino & Benjamin Carat** – Acoustical analyses of extended playing techniques in *Pression* by Helmut Lachenmann
- **Nathalie Hérold** – Studying timbral structures in 19th-century piano music: Overview and challenges
- **Victor Cordero, Kit Soden & Éric Daubresse** – New perspectives in orchestration pedagogy
- **John Sheinbaum** – Timbre, early cinema, and the perception of time in Mahler's Symphony No. 5
- **Amit Gur** – Timbre and texture: Overlapping phenomena
- **Adrien Bitton, Axel Chemla-Romeu-Santos & Philippe Esling** – Timbre transfer between orchestral instruments with semi-supervised learning
- **Aurélien Antoine, Eduardo Miranda, Jean-Michaël Celerier & Myriam Desainte-Catherine** – Generating orchestral sequences with timbral descriptors
- **Frederic Le Bel** – From timbre decomposition to music composition
- **Megan Lavengood** – A musicological approach to the analysis of timbre
- **Sarah Schoonhoven** – Gender, timbre and metaphor in the music of Wendy Carlos
- **Flora Henderson** – Cross-cultural gestures: Timbral juxtapositions in Takemitsu's *November Steps* (1967) for shakuhachi, biwa and orchestra
- **Pierre Michel** – Instrumental timbre, groupings and texture: Analysis of form and expression in pre-spectral European ensemble and orchestral music
- **Lindsey Reymore & David Huron** – Identifying the perceptual dimensions of musical instrument timbre
- **Thiago Roque & Rafael Mendes** – Towards timbre solfège from sound feature manipulation

Friday Afternoon | 6 July 2018

Musical Event 2

14:45 **Miller Puckette & Kerry Hagan** – *Who was that timbre I saw you with*?

Talk Session 7: Chair Stefan Weinzierl

15:25 **Robert Hasegawa** – Timbral hybrids in Philippe Hurel's *Leçon de choses*

15:50 **Lasse Thoresen** – Timbre-as-heard: Spectromorphological considerations

16:15 **Coffee break**

Talk Session 8: Chair Stefan Weinzierl

16:45 **Julie Anne Nord** – Wagner's associative orchestration in the Orchesterskizze for *Tristan und Isolde*

17:10 **Jason Noble, Max Henry, Etienne Thoret & Stephen McAdams** – Timbre and semantics in sound mass music

17:35 **Asterios Zacharakis & Konstantinos Pastiadis** – Examining the influence of tone inharmonicity on felt tension and timbral semantics

18:00 **Continuation of Poster Session 2**

Musical Keynote 2 – Tabla

20:00 **Shawn Mativetsky**, Tabla, McGill University, with **Pankaj Mishra**, sarangi — Rivers: A rhythmic journey to Varanasi

21:20 **End**

Saturday Morning | 7 July 2018

Tanna Schulich Recital Hall, ground floor of Elizabeth Wirth Music Building

08:00 **Coffee**

Research Keynote 3 – Ethnomusicology

08:45 **Cornelia Fales**, Indiana University — Playing with timbre

Talk Session 9: Chair Vinoo Alluri

09:45 **Simon Zagorski Thomas** – See me, feel me: Timbre as a multi-modal experience

10:10 **Paul Théberge** – Timbre, genre, and the reframing of analog processes in digital audio production

10:35 **Zachary Wallmark** – Effects of cross-modal Stroop interference on timbre perception

11:00 **Coffee break**

Talk Session 10: Chair Vinoo Alluri

11:30 **Marcelo Caetano** – The sound morphing toolbox

11:55 **Ben Luce & James W. Beauchamp** – David A. Luce's research on musical instrument timbre and its subsequent application to analog synthesizer design and digital analysis/synthesis

12:20 **Charlie Sdraulig** – The effect of loudness on the perceptual representation of voiceless vowel and fricative timbres

12:55 **Poster Session 3 & Lunch**

Jennifer Beavers – A case for Ravel: Timbre, neurological decline, and Ravel's last compositions

Danilo Rossetti & Jônatas Manzolli – The emergent timbre in live-electronic music: Volume and spectral liveness estimation from audio descriptors

Noah Kahrs – Consonant and dissonant timbres in two works of Sofia Gubaidulina

Nora Engebretsen – Acousmatic listening, chimeric percepts, and auditory hallucinations: Conceptualizing "perceptualized" timbre through the "predictive processing" model

Axel Chemla-Romeu-Santos, Adrien Bitton, Philippe Esling & Goffredo Haus – Unsupervised timbre spaces through perceptually regularized variational learning

Gregory Lee Newsome – Visualizing orchestration with Orcheil

Luca Guidarini – Drowning into synthesis: The role of timbre in Fausto Romitelli's compositional techniques

Lauro Pecktor de Oliveira – The sound of colour in R. Murray Schafer's (1933) Seventh String Quartet (1998)

Maria Perevedentseva – Timbre as 'structuring structure' in underground electronic dance music

Ivan Simurra & Marcelo Queiroz – Comparative analysis between verbal attributes of timbre perception and acoustic correlates in contemporary music excerpts

Tanor Bonin – 箏の唱歌: Phonetic encoding of musical contour in the traditional Ikuta scores of Japan

Michael Mandel & Song Hui Chon – Using bubble noise to identify time-frequency regions that are important for timbre recognition

Amy V. Beeston, Alinka Greasley & Harriet Crook – Implicit timbral assessments of recorded and live music with hearing aids

Saturday Afternoon | 7 July 2018

Musical Event 3

14:45 **Anthony Tan & Noam Bierstone** – *Pose IV: In situ* (2016) for solo percussionist on amplified piano and electronics

15:05 **Robert Normandeau** – *StrinGDberg* (2001-03)

Talk Session 11: Chair Cornelia Fales

15:25 **Eric Maestri –** Timbre is a technomorphic thing: A comparative analysis of three case studies

15:50 **Shen Li & Renee Timmers –** The communication of timbral intentions between pianists and listeners and its dependence on audio-visual listening conditions

16:15 **Coffee break**

Talk Session 12: Chair Cornelia Fales

16:45 **Jason Winikoff** – The mikakaji: Timbre in Zambian luvale percussion

17:10 **Hazel Burns & Gabriel Ferreyra** – A comprehensive guide to recording the Kichwa instruments of Ecuador

17:35 **Conclusions**

18:05 **Continuation of Poster Session 3 / Closing Reception**

20:00 **Musical Keynote 3: Electroacoustic music**

Jean-Baptiste Barrière, Image Auditive, composition/video, with **Camilla Hoitenga**, flute — Timbre and synthesis, the promises of freedom

21:20 **End**

TUTORIALS AND ROUND TABLES

OrchPlay (OrchSim)**: The Orchestra in your Classroom and Research Center! A New Software for Pedagogy and Research**

Denys Bouliane†, Félix Baril[1],
[1] Music Research, Schulich School of Music, McGill University, Montréal, Québec, Canada
OrchPlayMusic, Montréal, Québec, Canada
† Corresponding author: denys.bouliane@orchplaymusic.com

Aims/goals

Orchestration is a *skill* to be acquired (the "Art of Orchestration"). It supposes solid technical knowledge, extensive analysis of reference scores, and years of practice. In this respect, it might not be all too different from learning traditional counterpoint or harmony where "models" are presented, analyzed and typical "rules of conduct" are deducted and categorized according to epoch, style, genre, and instrumentarium.

One exciting extra challenge for learning orchestration is to develop an intimate connection between the written score and the aural result. Musicianship relies on dictation and solfège ("informed, active listening") to achieve this goal, but orchestration requires a special form of ear training that includes timbral recognition of a wide array of specific instrumental colors that are constantly evolving according to tessitura and dynamics. Furthermore, it also requires the capacity to recognize ("timbral dictation") and to imagine ("timbral solfège") endless possibilities of combinations (orchestral techniques). This plethora of perceived orchestral effects, such as blend, timbral emergence, segregation, stratification, contrasts and many others, are based on the psychology of perception and are currently at the core of the research done by the Orchestration and Perception Project at McGill over the past several years.

Studying to become an orchestrator or verifying and scrutinizing the different perceptual categories based on the auditory grouping processes (McAdams & Bregman, 1979) require a direct contact to scores and their acoustical rendering by an orchestra.

In this respect, *the reference* is the *live orchestra*. In the best of all worlds, one would have unlimited access to an orchestra which would be happy to perform any excerpt from the repertoire in one's prescribed instrumental combinations for study (segmenting in different layers and instrumental groupings) and research (verifying different perceptual hypothesis). But this ideal situation might well remain a mere fantasy… unless a multi-billionaire patron becomes interested in the idea!

Many young musicians, students and researchers tend to rely today on different types of music notation software or digital sequencers with "play" capabilities. Despite some commendable efforts, their acoustical rendition of the instrumental world is so far from reality that their pedagogical value is very limited, bordering at times on the counter-productive and amounting to a form of ear "de-training."

Then, what about providing a more serious learning tool? A tool that could adequately model the complex acoustical reality of a symphonic orchestra and allow the precise isolation of each instrument and/or instrumental combination of each parameter of an instrumental performance in a normalized environment. In other words, short of having an orchestra at our disposal in each classroom and research center, could *the best possible substitute* be conceived and made available?

Background information

The OrchSim system is constantly developing and expanding since 2010. The OrchPlay software, developed in 2015, has been made available to the public in the summer of 2016.

Methodology

ORCHSIM

OrchSim ties together in a coherent whole some of the best virtual instruments to propose a hyper-realistic rendition of symphonic works. The OrchSim team has developed this system over the last 7 years, testing each step regularly with focus groups of listeners (music students at various levels, professional musicians, as well as sound recording specialists). Several steps were necessary:

I- Integration, creation and harmonization of all virtual instruments, creating a large symphonic orchestra from the best available sample banks. Each virtual instrument had to "behave" like the acoustic instrument on which it is modelled. We constructed a comprehensive taxonomy of all modes of playing incorporating their interactions.

II- Establishment of a link between virtual instruments and symbolic notation.

OrchSim assigns a specific command to our sample engine for every symbol (text of graphic) that a musician can encounter in a score. OrchSim was then enriched with a new series of symbols (non-printable) to allow—as orchestral musicians would do—the modulation of most parameters of the "primary" information transmitted by the score into a real instrumental performance.

III- Rendering of music scores and validation of results.

A series of renderings of representative works of the repertoire were produced. Each excerpt was modelled on a selection of reference recordings and each individual part on the performance of a professional performer. As OrchSim developed, these renderings were tested, gradually reworked and improved in search of higher quality and more convincing realism. Over the years, we submitted these renderings to many professional colleagues (composers, researchers, performers, sound engineers) for critical appreciation and collected their comments. All OrchSim recordings are produced in true multi-track format, enabling the selection of each instrument individually that can be heard from their specific location on stage and in the acoustics of the hall. The recordings are made in 24-bit audio format and then converted to the .opl proprietary format that can be played by the OrchPlay software. Any instrumental combination can be selected, and the balance can be modulated at will.

ORCHPLAY

The OrchPlay software plays the multitrack OrchSim files and provides access to all individual instrumental tracks of the full orchestra thus enabling subsets of instruments involved in a particular orchestral effect to be heard in isolation or within the full musical context. Bookmarks and Bookmarks Lists (referring to any compilation of pieces) can also be exported and shared with all other users. Standard stereo files can be imported as OPL files. Other related documents can be linked as well. Several new functions (score follower and automatic reformatting) are in development in conjunction with the new OrchView software.

Integration with ORCHVIEW and Orchard (ORCHestration Analysis and Research Database)

OrchPlay is an integral part of the large international research projects MAKIMONO (Multimodal Analysis and Knowledge Inference for Musical Orchestration) and ACTOR (Analysis, Creation and Teaching of Orchestration) as the audio hub to exchange and study stereo and multichannel recordings in a normalized environment.

Presentation at *Timbre Is a Many-Splendored Thing*

The workshop will briefly present and demonstrate OrchSim but will mainly focus on the use of OrchPlay for pedagogy and research. It will also present its many ties to the annotation software OrchView.

ORCHPLAY https://orchplaymusic.com/

The Computer as an Orchestra

Philippe Esling[†1], Carmine-Emanuele Cella[2], Lèopold Crestel[1], Axel Chemla-Romeu-Santos[1] and Mathieu Prang[1]
[1] IRCAM, Paris, France
[2] HEM, Genève, Swiss
[†] esling@ircam.fr

Aims/goals

This workshop will present state-of-the-art results regarding assisted orchestration from different standpoints. After a theoretical introduction, a practical session will incorporate specifically developed tools.

Background information

Musical orchestration is the subtle art of writing musical pieces for orchestra, by combining the spectral properties specific to each instrument in order to achieve a particular sonic goal. Indeed, this question puts forward one of the most complex, mysterious, and dazzling aspects of music, which is the use of timbre to shape musical structures in order to impart emotional impact. Timbre is the complex set of auditory qualities (usually referred as the sound color) that distinguish sounds emanating from different instruments. Intuitively, an auditory object is defined by several properties, which evolve in time. Decades of simultaneous research in signal processing and timbre perception have provided evidence for a rational description of audio understanding. We address these questions by relying on both solid perceptual principles and experiments on known empirical orchestration examples, and by developing novel learning and mining algorithms from multivariate time series that can cope with the various time scales that are inherent in musical perception.

In this quest, we seek tools for the *automatic* creation of musical content, a better understanding of perceptual principles and higher-level cognitive functions of the auditory cortex, but also generic learning and analysis techniques for data mining of multivariate time series broadly applicable to other scientific research fields.

In this workshop, we will introduce several researches and tools that have been developed around computer-aided orchestration:

1 – The first automatic orchestration system called Orchids/Orchidea, which turns any sound into an orchestral score by attempting to reproduce the timbre of this audio target.
2 – The current work on the Live Orchestral Piano targets correlations existing in the work of prominent composers. The system provides a way to play on a traditional piano, while the different notes are projected into a full classical orchestra in real-time.
3 – Our current research focused on automatic inference through variational learning to allow an automatic deciphering of these dimensions.

Methodology

The workshop will be provided in a mixed format of short scientific presentations with hands-on tutorials and demonstrations

The Relentless Pursuit of Tone: A Roundtable on Timbre, Technology, History, and Genre in Popular Music

Robert Fink[1†], Cornelia Fales[2], Melinda Latour[3], Catherine Provenzano[4], Paul Théberge[5], Steve Waksman[6], Zachary Wallmark[7] and Simon Zagorski-Thomas[8]
[1] Department of Musicology, UCLA, Los Angeles, CA, USA; [2] Department of Folklore and Ethnomusicology, Indiana University, Bloomington, IN, USA; [3]Department of Music, Tufts University, Boston, MA, USA; [4]Department of Music, New York University, New York, NY, USA; [5]Department of Music, Carleton University, Ottawa, Canada; [6]Department of Music, Smith College, Northampton, MA, USA; [7]Department of Music, Southern Methodist University, Dallas, TX, USA; [8]London College of Music, University of West London, London, UK
[†]rfink@humnet.ucla.edu

Aims/goals

"The Relentless Pursuit of Tone" is the masthead motto of *Premier Guitar,* a trade magazine aimed at professional guitarists and those who follow them. It is also the title of a new collection of essays (Oxford, 2018) that seeks to apply diverse perspectives from musicology, ethnomusicology, popular music studies, neuroscience, and musical production studies to the timbral study of popular music.

Participants will begin with very short (5 minute) position statements based on their own research, then respond, first to each other and eventually to members of the audience. We hope both to provide a "state of the art" for the timbral analysis of popular music, and to explore general issues of technology, affordances, sonic cognition, and generic/cultural politics.

Background information

The participants in this roundtable are all contributors to *The Relentless Pursuit of Tone:* their essays cover a wide range of topics, from extreme metal (Wallmark) to perceptual learning in rave culture (Fales) and the subwoofer as an instrument (Fink); to the techno-politics of digital pitch correction (Provenzano) and the guitar-synth (Waksman); through more general issues of aural space (Théberge) and cognitive schemata (Zagorski-Thomas) in studio recording.

Methodology

The methodologies represented in the roundtable vary widely, from ethnographic field work (Fales, Provenzano) to cognitive schemata (Wallmark, Zagorski-Thomas), history of technology and acoustics (Fink, Waksman) and musicological analysis (Théberge, Zagorski-Thomas)

Results

The various participants will discuss their research results.

Conclusions

Anyone making any kind of popular music participates in an open-ended negotiation among the diverse constituents of a complex signal chain, a network of interacting agents which includes organic (human hands, ears, brains), mechanical (saddles and strings), analog (pickups, amps, microphones), and digital (pitch correction, looping, DSP) components. The pursuit of tone can lead to new definitions of genre, subjectivity, authenticity, even challenge fixed perceptions of race and gender.

References

Fink, R., Latour, M. & Wallmark, Z. (Eds.), *The Relentless Pursuit of Tone: Timbre in Popular Music.* New York, NY: Oxford University Press, 2018.

Roundtable on Orchestration Pedagogy

Roundtable moderator: John Rea
Invited guests: Lasse Thoresen, Laurie Radford, Fabien Lévy, Victor Cordero

John Rea: *Orchestration Pedagogy in the 21st century*

Since the 18th century, orchestration and its attendant pedagogies have evolved further perhaps than the institutions for which they were originally designed to serve, symphony orchestras. Today, pedagogies abound.

Web-based information – beyond books, as it were, roars across cyberspace. And numerous orchestration study outposts, like so many cottage industries, promote their expert services.

One can teach oneself; one can access virtual playback orchestras; one may choose automated/computational approaches; one can be provided with heuristic techniques for composing, arranging, and fashioning timbres, etc. Whether an artistic task is real, virtual, or augmented, these pedagogies promise the acquisition of music industry-ready professional skills.

Whither orchestration pedagogy, what is its likely future?

Lasse Thoresen

With so much explicit information on orchestration available for students of orchestration in books, on the web, and in computer programs, the question of the future of orchestration pedagogy will by necessity have to be directed towards information not provided by these means. One field that easily goes missing is our understanding of any performer's technique on his instrument: what is awkward, what is practical, and how complex figures may be shared to facilitate individual playing. The spatiality of a symphony orchestra needs attention. An aspect of composition clearly diminished by the existence of score-writing computer programs is that of inner hearing which is essential to the creative process. Moreover, following the introduction of microtonality into contemporary composition, composers have required knowledge of the capacity of individual instruments to produce such sounds in different registers and velocities. Knowledge of extended playing techniques and their notation is also needed for contemporary music.

Orchestration pedagogy must find a new balance between the auditory effects of orchestration—a study that may be facilitated by digital means preferably enhanced by spectromorphological terminologies—and hands-on experience with musicians, including that with orchestra and stage. The traditional study of orchestration needs also to be supplemented by information about new extended playing techniques and notation.

Laurie Radford

There are certainly at present an abundance of pedagogical facilities, websites, materials, references and mentors to guide the would-be orchestrator or scholar of orchestration practices. These do not address, in most cases, several recent developments in musicking at large: the introduction of a new instrumentarium that either exceeds the affordances of conventional acoustic instruments or offers an entirely different sound world with idiosyncratic ways for the human performer and creator to access expressive means; and the development of new modes of listening that have arisen from the introduction of new spaces and places for audition as well as new roles and functions for the listening act. Related to both of these developments is the "sonic turn" that evolved during the 20th century and now occupies a decisive and growing place in both practice and scholarship. New tools, new skillsets and a new conceptual framework require radically new approaches to orchestration pedagogy and, perhaps, a redefinition of the "orchestra" that is the object of study.

Fabien Lévy: *"Functional orchestration": a research field in its Prehistory*

Orchestration has been mainly taught in a quite intuitive and non-verbalized musical way. Treatises in orchestration are often treatises on instruments. At the National Conservatory (Paris), composer and orchestration professor Marc-André Dalbavie (1961-) along with a few of his former students have been developing since the 90's some general and verbalized principles for orchestration under the name "functional orchestration," mostly based on acoustics, psychoacoustics and musical common practice. There is, however, a big need for this field to be further developed in collaboration with scientists, musicologists and theoreticians.

Victor Cordero

The transmission of know-how in the field of orchestration relies historically on three pillars: orchestration treatises, teaching in conservatories or universities, and symphonic literature itself.

In the present era of globalization, where communities of all kinds emerge daily on the web, online courses, and tutorials on orchestration have multiplied. It could be said that the subject is attracting increasing interest from both professional and amateur musicians, unearthing orchestration from its niche in order to democratize it, and to make it seemingly accessible to everyone.

On the other hand, it is clear that there is currently a convergence of common interests with regard to orchestration; many national and international research projects—each one more ambitious than the next—are good proof of this. Such a wide-ranging pedagogical community allows for exchanging views and for synthesizing them in a context of dynamic and positive sharing. Interdisciplinarity has been essential in this process in order to respond to such a complex subject as orchestration. Thus, the combination of complementary skills from various fields (psychoacoustics, signal processing, acoustics, etc.) can enable us to take a decisive step in the pedagogy of orchestration.

Opening and Keynote Addresses

‘What is Timbre?’ versus ‘What Can We Do with Timbre?’ Picking the Right Questions

Stephen McAdams
Department of Music Research, Schulich School of Music, McGill University, Montreal, QC, Canada
stephen.mcadams@mcgill.ca

A glance through these proceedings testifies to a rising interest in musical timbre across a multitude of fields, as well as to its under-explored richness, at least until very recently. It also makes evident that timbre is many things to many people and has many functions. Understanding this elusive, because multifarious, concept will require a harmonization and interconnection across disciplines of concepts, vocabulary, definitions, and analytical approaches. It will also require us to identify the right questions to ask. Here are two sets of them that seem important to me.

What is timbre?

This question has generated unending, and in many cases unfruitful, debate in many fields and across fields. Aside from the negative definition, which only tells us what timbre isn't, many other conceptions abound (see McAdams, 2013; McAdams & Goodchild, 2017).

Some people consider timbre as a sounding object, such as the timbre of a clarinet. This conception presumes some *thing* that is common to all clarinets played at any pitch or dynamic with any articulation and embouchure configuration by any clarinetist. Furthermore, this *thing* is different for every kind of instrument. But then why do people writing or talking about the clarinet give different names to different registers in reference to the perceived sound qualities: chalumeau, throat tones, clarion, altissimo? If a clarinet has *a* timbre, how can it have *different* timbres? And how far does this generalized instrument timbre (or macrotimbre) apply to the whole clarinet family from piccolo to contrabass? And why then is it so difficult for people to recognize that sounds in different registers of a given instruments or voices are coming from the same sound source?

Other people consider timbre as a physical sound event: “listeners were presented with a timbre.” This conception of timbre has led to all kinds of misconceptions and given much fodder for circular musings among philosophers and epistemologists of sound, and even the occasional music theorist, musicologist or perceptual psychologist. In particular, it leads to difficulties in traversing the inevitable gap between the realms of the physical or material on the one hand, and the perceptual or conceptual on the other. If it's a physical event, then what is in the mind of the listener?

Another conception of timbre is as a perceptual quality that derives from certain acoustic properties of sound events. This conception has led to the spontaneous generation of multitudes of quantitative “timbre” descriptors in psychoacoustics and music information retrieval (Siedenburg, Fujinaga & McAdams, 2016). Such audio descriptors are derived from different acoustic representations of a sound signal, which one attempts to correlate with various measurable perceptual properties, such as dissimilarities among sounds. But if the pure sensory information were enough to explain perceptual ratings, why are these ratings so often difficult to model satisfactorily? Extensions of this approach accept that the knowledge a listener has accumulated about categories of mechanical or electroacoustic devices that produce sounds can also affect listeners’ judgments. But to what extent do the results of these kinds of studies, often conducted on isolated sounds, generalize to performed musical sequences and various combinations of instruments?

Timbre may also be characterized as a set of qualia related to “perceptual” dimensions that can be characterized with verbal descriptions. This approach presumes a one-to-one or at least very tight relation between perception and language. But are we always able to find the right words for all the wonderfully subtle qualities that we perceive in musical sounds and their blended combinations? And if not, what are the limits of these linguistic characterizations?

Another recent notion concerns timbre as an emergent property of many sounding objects, referred to as “polyphonic timbre.” This notion is thought to characterize the sound of groups of instruments, particularly in pop music, and contributes to the recognition of a specific band or musical genre. But is this really “timbre” per se or does it stand in relation to timbre as a perceptual construct like harmony stands in relation to pitch?

What can we do with timbre?

From the point of view of the musician and composer, how we characterize timbre is not as important as how it can be used musically (see McAdams & Goodchild, 2017; Goodchild & McAdams, 2018).

Timbre as a perceptual property arises from the spectral fusion of acoustic information into a perceived event. How can new timbres be constructed or transformed in the mind of a listener through instrumental blends, sound synthesis and processing, or mixing and editing?

Timbre affects auditory stream formation and the shape of and memory for musical patterns. What role do similarities and differences in timbre play in the organization of sequences of musical events into one or more auditory streams or in the formation of foreground and background layers in orchestral music and other sonic textures? What timbral properties make some sounds more salient than others and thus attract a listener's attention to them so they occupy the perceptual foreground? In what ways can timbral variations across a musical sequence be used to create rhythmic patterns, making some events seem more accented and, for example, giving the overall pattern a sense of groove in pop music? Do the perception and memory of timbre allow us to perceive timbre intervals and contours and to recognize re-presentations of them with different timbres as preserving the timbral relations?

Sudden changes in timbre can structure music into spans of similar timbres set off by these changes. How can timbral contrasts be employed to segment sequences at different hierarchical levels in order to structure music through changes in instrumentation or sound processing? Can orchestration—as the choice, combination and juxtaposition of timbral qualities—reinforce or contradict structures formed of pitch and duration? Under what conditions do timbre-based structures dominate pitch-based structures? What are the possibilities of sculpting smaller-scale and larger-scale spectral shapes and timbral trajectories that evolve through time, all the while maintaining a perceptual coherence and imparting a dynamic form to the music?

And timbre has other functions as well. What timbral properties contribute to the establishment of a sonic identity (of a performer, a band, a composer, a musical genre, a gender, a political or social status)? How can timbre be used to create emotional tone and fashion an emotional trajectory for a listener?

Whither timbre research?

Many questions. Much to do. Time to get started!

Acknowledgments

This work is supported by the Quebec FRQSC, the Canadian NSERC and SSHRC, and Canada Research Chairs programs.

References

Goodchild, M., & McAdams, S. (2018). Perceptual processes in orchestration. In E. Dolan & A. Rehding (Eds.), *The Oxford handbook of timbre*. New York, NY: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780190637224.013.10

McAdams, S. (2013). Musical timbre perception. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 35-67): New York, NY: Academic Press.

McAdams, S., & Goodchild, M. (2017). Musical structure: Sound and timbre. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 129-139). New York, NY: Routledge.

Siedenburg, K., Fujinaga, I., & McAdams, S. (2016). A comparison of approaches to timbre descriptors in music information retrieval and music psychology. *Journal of New Music Research*, *45*(1), 27-41.

Timbre Eternal

Emily I. Dolan
Department of Music, Harvard University, Cambridge, MA, USA
edolan@fas.harvard.edu

Timbre today

Timbre studies have exploded in the twenty-first century, which is attested to by this conference and by a profusion of recent and forthcoming publications: Isabella van Elferen recently edited a special issue of *Contemporary Music Review* devoted to timbre; later this year, Robert Fink, Melinda Latour, and Zachary Wallmark will publish their volume on timbre in popular music, *The Relentless Pursuit of Tone*; Alexander Rehding and I are co-editing the Oxford Handbook of Timbre, which will include over twenty essays by scholars in a wide range of fields (Musicology, Ethnomusicology, Theory, Organology, Perception and Cognition, Classics, Anthropology, History of Science). The notion that timbre is a marginalized, overlooked, and understudied parameter—that commonplace opening gambit—is fast becoming untenable. What we face now is a profusion of studies, cutting across genres, methodologies, and disciplines.

Timbre everywhere

In this talk, I first reflect on the "timbral litany" in today's scholarship: timbre has no standardized language; it lacks a systematic theory; timbre is defined negatively, and so forth. In particular, I focus on the tension between the many claims of timbre's central important to musical experience, on the one hand, with the reality that we often, on the other hand, talk over and past timbre, abstracting music from timbre's specificities. This tension, I argue, can ultimately be understood historically. In this talk, I use the example of Gluck, his orchestration, and its changing reception. This example is useful for several reasons: first, Gluck's operatic reforms hinged, in part, on a particular deployment of his orchestral forces, one that demanded a particular kind of aural attention to his orchestra as a kind of emotional subtext. Without a robust vocabulary for discussing orchestral timbre and orchestration, contemporary accounts point more vaguely to the power of Gluck's "harmony." Not all listeners, however, could apprehend Gluck's orchestration, and I trace those listeners who struggled as they listened. By the nineteenth century, however, Gluck's innovations became commonplace and canonized, held up as examples of orchestral excellence. Wagner famously both celebrated Gluck as an innovator while also finding himself underwhelmed with Gluck's music when he encountered his music in performance.

Conclusions

The story of Gluck is not isolated: the outlandish, experimental, and challenging orchestral effects of one generation often become, for the next generation, the baseline and standard. I trace what we could call the timbral cycle, whereby the very centrality of timbre leads directly to its disappearance from the mainstream of musical discourse. In addition to recovering timbre—the noble goal of so much contemporary scholarship—we might also attend to what it means to listen past timbre, and what is at stake when we let timbre go.

References

Dolan, E. & Rehding, A. (Eds.). (Forthcoming) *The Oxford Handbook of Timbre*. Oxford and New York. Oxford University Press.

Fink, R., Latour, M., & Wallmark, Z. (Eds.). (2018) *The Relentless Pursuit of Tone: Timbre in Popular Music.* Oxford and New York. Oxford University Press.

van Elferen, I. (Ed.). (2017) *Contemporary Music Review*, 36 (6).

The Perception and Cognition of Playing the Koto

ANDO Masateru
Kisoukai Koto Association, Tokyo, Japan
ando.masateru@kisoukai.org

About the koto as an instrument

The koto, a representative traditional instrument of Japan, is 1.8 m long and has a body hollowed out from paulownia wood. The back is glued to this body, making a cavern which creates a resonating effect. The koto has 13 strings; many people use polyester strings today, but the instrument is strung traditionally with silk strings.. Silk strings lack durability but the timbre, in particular the beauty of the resonance and the easiness of the performance, is superior to polyester. Each string has a bridge and a scale is made from the position of bridges based on the Pythagorean temperament (Ando, 1989).

The tuning of the strings is decided by the piece performed; therefore, the performer needs to make the necessary alterations beforehand. The performer uses and fits together five types of intervals, the same pitch, a semitone (minor 2nd), fourth, fifth and octave. During the performance, the method of playing and changes in temperature and humidity can cause the strings to be out of tune from time to time. It is necessary, therefore, to have a technique, forged by experience and training, in order to perform while adjusting the position of the 13 bridges.
The performer plays by fixing ivory plectra to the thumb, index finger and middle finger of the right hand.

The history of the Koto

During the Yayoi period (300 BC–300 AD), a zither called a wagon was used, but during the Nara period (710–794) koto came to Japan from China as gagaku (imperial court) musical instruments. In the Tale of Genji, koto and their families appear as instruments for both male and female aristocracy. It is in the Edo (1603-1867) period that the father of the modern koto, YATSUHASHI Kengyō (1614-1685), is said to have established the fundamentals for koto music. In the Showa period (1926-1989), MIYAGI Michio (1894-1956) was well known as the “Father of Modern Japanese Music.” Left totally blind by the age of 8, Miyagi was apprenticed to become a koto musician. Miyagi is known not only for his more than 350 extant compositions, but also as an essayist, educator, and inventor of instruments.

The perception and cognition in playing the Koto

Cognition of high and low pitch

Recordings of the actual state of tuning the strings showed that, for the relationship between the G on the second string and its octave on the 7th string, when there was deviation of 10 cents from their ideal frequency difference (in this case, the sounds will clearly be perceived as being out of tune), a beat could be detected when the test sounds’ duration reached 650 ms. However, the player was able to judge at 250 ms whether a sound was in tune or whether the pitch was higher or lower. At 400 ms, the performer began to move the bridges. As a result, we conducted an auditory experiment using test sounds of 7 durations between 50 and 500 ms whose interval discrepancies were varied in 9 steps from -20 to +20 cents. The results showed that experienced performers could distinguish a 10-cent discrepancy (including the direction of discrepancy) at 200 ms. Perceptual performance depended on the degree of discrepancy, and even when the discrepancy was the same, differences in the direction of the discrepancy produced different results. From these findings, it can be concluded that the memory of judging the timbre helps in the judgment of intervals. Along these lines, a correlation has been found between the interval judgment and the technical ability of performers (Ando, 2007).

Cognition in connection with differences in sound production timing

The jushichigen (17-string koto) is a large, lower register instrument which was proposed by MIYAGI Michio and appeared in 1921. Performers of the koto exchange the instrument to perform. It is often noted that there is a delay in performance timing when there is group performance with the koto. I noticed the effort by the performer to solve this time delay and, by using electromyography (EMG), I found that the movement for the production of the sound was brought forward by 26 ms. Summarizing the time series, playing the jushichigen koto takes 26 ms longer than playing the koto. In an experimental study, it was possible to discern a discrepancy of 4 ms between two koto sounds. As there are other reports of detecting discrepancies of a few milliseconds, differences of 26 ms should be clearly perceived. Even though the koto and the jushichigen koto appear to look similar, they are different instruments. The performer of the jushichigen koto is aware of the fact that there is a delay and because of this it is possible to adjust and give a good performance when playing in an ensemble. It is expected that these findings would help in performances and performance education (Ando, 1982).

Summary

Because the koto is an instrument where the strings are tuned with movable bridges, it is necessary for the performers to keep paying attention to slight changes in the intervals, by tuning the strings before performance, and by modulation and correction during performance. In addition, in an ensemble performance, it is necessary to consider the special quality of the instrument in the adjustment of the timing in the plucking of the instrument.

References

Ando, M. (1989). Koto scales and tuning. *Journal of the Acoustical Society of Japan (E)*, *10*(5), 279-287.

Ando, T. (2007). *Salient factors of pitch perception on the Koto tuning—Influence of timbre instead of beats*. Doctoral Dissertation, Tokyo University of the Arts. (in Japanese).

Ando, M. (1982). *An analysis of the time elements in the plucking movements of the 13- and 17-stringed koto*. Doctoral Dissertation, Tokyo University of the Arts. (in Japanese).

The World of the Koto

ANDO Masateru† and ANDO Tamaki
Tokyo, Japan
†ando.masateru@kisoukai.org

Program and Program Notes

1. ROKUDAN NO SHIRABE（六段の調）------------------------- YATSUHASHI Kengyo 17c mid
(Six Steps Tune)

Koto: ANDO Masateru

Rokudan no shirabe or *Rokudan*, by Yatsuhashi Kengyo (1614-1685), is the earliest extant work for koto. Prior to this time, the koto was used as a member of the gagaku (court music) ensemble, but it was not in charge of melody. *Rokudan* is a piece in six sections or steps played without pause. Each section has the same number of beats, but the beginning starts slowly, the tempo becomes faster as the step goes forward, and the last section will also slowly conclude the piece. Most of the traditional techniques used in playing the koto can be found in this work. In the same year of Yatsuhashi's death, J.S.Bach was born. For about 400 years, this piece has been handed down without a musical score.

2. GODAN GINUTA（五段砧）--- MITSUZAKI Kengyo 19c mid
(Kinuta in Five Steps)

Koto high part: ANDO Tamaki
Koto low part: ANDO Masateru

"Go" is five, "Dan" is steps or sections, and "Kinuta" is the tool name in Japanese. In old days, people beat the cloth with a stick at the water side to soften the cloth. The tools of this is called Kinuta.
As the sounds of beating cloth were so rhythmical, the rhythmic pattern was taken into koto music.
Mitsuzaki Kengyo(? –1853) composed this piece in the middle of the 19th century during the Edo period.
The piece is a koto duet, played by low-pitched and high-pitched kotos, which are the same instrument. This piece was the first one to include two different parts that were composed by one person in Japanese music.

3. MIZU NO HENTAI（水の変態）-- MIYAGI Michio 1909
(Transformations of Water)

Koto basic part: ANDO Masateru
Koto counter part: ANDO Tamaki

Miyagi composed this work, which was his first work, at the age of 14 while he was living in South Korea. The lyrics are Waka describing the mysteries of nature and are taken from the textbook which Miyagi heard his younger brother read aloud. The form of this tegoto-mono (work with a tegoto, or interlude) is as follows: introduction – first song – tegoto – middle song – tegoto – final song. Although the form and playing techniques are conventional, listeners are left with the impression of something totally new and different from older works. This is due to the importance Miyagi placed on individual creativity, something which is also found in modern Western music. Mizu no Hentai is rich in original melodies and tone colours, which depict transformations of water and this results in the work which is rich in individuality.

〔Translation according to content〕

＊霧 小山田の霧の中道踏み分けて 人来と見しは案山子なりけり
Fog Cutting a path through the fog at a small rice field among the hills
The person who appears is merely a scarecrow

*雲 明け渡る高嶺の雲にたなびかれ 光消えゆく弓張の月
Clouds Swayed by the clouds over the lofty peak
The light of the half-moon is extinguished by the clouds

*雨 今日の雨に萩も尾花もうなだれて 憂ひ顔なる秋の夕暮れ
Rain Today's rain presses down upon the bush clover and pampas grass
The autumn sunset has a lonely face

*雪 更くる夜の軒の雫の絶えゆくは 雨もや雪に降り変はるらん
Snow In the deepening night the trickle from the eaves is fading
The rain turns to snow

*霰 むら雲の絶え間に星は見えながら 夜行く袖に散る霰かな
Hail Stars appear between the parting clouds
Toward evening hail scatters on the sleeve

*露 白玉の秋の木の葉に宿れりと 見ゆるは露の計るなりけり
Dew What appears to be white marbles resting on the autumn leaves
Is only dew

*霜 朝日さすかたへは消えて 軒高き家かげに残る霜の寒けさ
Frost The morning sun melts the frost on the eaves
But in the shadow of the house the remaining frost is cold

4. KAZOEUTA HENSouKYOKU（数え唄変奏曲）--------------------------- MIYAGI Michio 1940
(Variations on Counting Song)

Koto: ANDO Masateru

"Kazoeuta" (Counting Song), which was a well-known melody at that time when this work was composed, serves as the theme for this set of variations. This is the fourth theme and variations composed by Miyagi, and his only theme and variations for solo koto. Kazoeuta Hensoukyoku is the first work to utilize complicated techniques brought in from Western music, and the variations unfold continuously with both hands playing fast arpeggios, tremolos, harmonics, pizzicatos, glissandos, and other techniques. It is written in such a way that, although it is performed by only one player, the listener perceives it as a duet.

5. SEOTO（瀬音）-- MIYAGI Michio 1923
(The Sound of Rapids)

Koto: ANDO Masateru
Jyushichigen: ANDO Tamaki

Composed by Miyagi in 1923 as a showcase for his newly-invented jyushichigen (17-string koto), Seoto is based on his impressions of the upper rapids of the Tone River in Gunma Prefecture. Written for 13-string Koto and Jyushichigen, this work fully exploits the capabilities of both instruments.
Miyagi chose to use the three-part A-B-A form. The first part depicts the sound of the rapids, followed by a slower section representing the song of the boatman and the quickening flow of the river as it carries a raft, and then ends with the return of the sound of the rapids.
In this duet piece, you may find a time delay between the koto and jyushichigen, or the performer's skill may cover the delay.

Timbre in the Brain

Vinoo Alluri
Cognitive Science Lab, International Institute of Information Technology, Hyderabad, Telangana, India
vinoo.alluri@iiit.ac.in

Background

Timbre, in its very nature, is abstract. The brain, as we all know, is the most complex system that exists. So, one can only imagine the challenges that can be encountered while investigating timbre processing in the brain. The central focus of timbre research from a neuroscientific point of view, until recently, has been typically on brain responses to changes in acoustic structure or related to sound source categorization (ex: violin vs piano) (Giordano et al., 2014; Halpern, Zatorre, Bouffard, & Johnson, 2004; Overath, Kumar, von Kriegstein, & Griffiths, 2008). While these studies have successfully identified brain regions participating in processing of individual feature manipulations or categorization, they have relied on the controlled auditory paradigm, wherein the sounds are typically presented in isolation and/or synthesized and manipulated artificially. Thus, as a result, such settings paint but only an incomplete picture of the neural underpinnings of the perceptual processes involved in timbre processing, as the results do not necessarily translate to how the human brain processes, in parallel, the multitude of musical features when listening to, for example, a recording of real orchestral music. The aim of the keynote is to present recent neural studies that have attempted to address timbre perception in more ecological settings and touch upon the reliability and validity of the results, and also to highlight the challenges encountered in such settings and future directions.

Timbre in the naturalistic paradigm

Studying music listening as a continuous process using naturalistic stimuli could provide more accurate accounts of the processing of individual musical features in the brain. The novel naturalistic paradigm is one such setting that is more ecological than hitherto encountered settings, both stimulus-wise and task-wise, as participants freely listen to real music without performing any other task while their brain activity is being recorded. This approach of combining functional Magnetic Resonance Imaging (fMRI), Music Information Retrieval (MIR), and multivariate analysis, was first introduced by Alluri et al., (2012) and has recently been acknowledged as a new trend in the field of neuromusicology (Neuhaus, 2017). The results revealed that timbre processing, in addition to auditory regions, was associated with large-scale brain networks in musicians. These regions particularly encompassed cognitive regions previously linked with processing cognitive and attentional load. Additionally, the recruitment of the somatomotor areas of the brain also indicates that listening to natural timbres triggers action-perception mechanisms in musicians' brains. These results reflect the dual representative nature of timbre in the brain, as acoustic structure being processed in the auditory cortices, and as sound-source identification (as reflected by mirror neuron regions and those related to attention).

A fundamental problem that plagues the field of neuroscience is replicability and validity of the results, more so when novel paradigms and methods are introduced. A subsequent replication study performed by the same group revealed high replicability of the neural correlates of timbre pertaining to auditory areas (Burunat et al., 2016). Interestingly, increasing sample size of participants revealed more similar large-scale patterns in both musicians and non-musicians (Alluri et al., in prep). To test the generalizability of the results, a similar approach was used to build encoding models, wherein timbral features were used to predict brain activity and the robustness of the created models were tested across different stimuli (Alluri et al., 2013). As a result, activations in brain regions that were associated with timbre processing were predicted to a significant degree with different stimulus sets and varying participant pools. Decoding studies, which aim to predict stimulus features from brain responses, also revealed that timbral features could be decoded more accurately than the rhythmic and tonal ones (Toiviainen, Alluri, Brattico, Wallentin, & Vuust, 2014). However, several challenges remain concerning the underlying perceptual dimensions of timbre, and their acoustic correlates in such naturalistic settings, let alone methodological challenges in the novel paradigm. The talk is aimed at addressing some of these issues and open avenues of looking at timbre in a holistic sense.

Conclusions

Timbre processing, in the context of music, takes a different identity than hitherto put forth definitions in previous experimental settings. Timbre contributes to the putative emotional content of music (Brattico, Brattico, & Vuust, 2017), which is one of the fundamental driving forces behind the field of cognitive neurosciences of music. Hence, now is the time, due to technological and methodological advances in understanding human perception, to revisit timbre with rigour.

Acknowledgments

I would like to thank Prof. Stephen McAdams for setting the standard in research that one aims at achieving and his encouragement. I extend my sincere gratitude to Prof. Elvira Brattico for showing me the door to neuroscience. And, last but not the least, Prof. Petri Toiviainen, my beacon, without whom this wonderful journey would not have been possible, and for providing never-ending support both professionally and personally.

References

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage*, *59*(4), 3677–3689. http://doi.org/10.1016/j.neuroimage.2011.11.019

Alluri, V., Toiviainen, P., Lund, T. E., Wallentin, M., Vuust, P., Nandi, A. K., … Brattico, E. (2013). From Vivaldi to Beatles and back: predicting lateralized brain responses to music. *NeuroImage*, *83*, 627–36. http://doi.org/10.1016/j.neuroimage.2013.06.064

Brattico, P., Brattico, E., & Vuust, P. (2017). Global sensory qualities and aesthetic experience in music. *Frontiers in Neuroscience*, *11*(APR), 1–13. http://doi.org/10.3389/fnins.2017.00159

Burunat, I., Toiviainen, P., Alluri, V., Bogert, B., Ristaniemi, T., Sams, M., & Brattico, E. (2016). The reliability of continuous brain responses during naturalistic listening to music. *NeuroImage*, *124*, 224–231. http://doi.org/10.1016/j.neuroimage.2015.09.005

Giordano, B. L., Pernet, C., Charest, I., Belizaire, G., Zatorre, R. J., & Belin, P. (2014). Automatic domain-general processing of sound source identity in the left posterior middle frontal gyrus. *Cortex*, *58*, 170–185. http://doi.org/10.1016/j.cortex.2014.06.005

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, *42*(9), 1281–1292. http://doi.org/10.1016/j.neuropsychologia.2003.12.017

Neuhaus, C. (2017). Methods in neuromusicology: Principles, trends, examples and the pros and cons. In A. Schneider (Ed.), *Studies in Musical Acoustics and Psychoacoustics* (pp. 341-374). Cham, Switzerland: Springer Nature. http://doi.org/10.1007/978-3-319-47292-8

Overath, T., Kumar, S., von Kriegstein, K., & Griffiths, T. D. (2008). Encoding of Spectral Correlation over Time in Auditory Cortex. *Journal of Neuroscience*, *28*(49), 13268–13273. http://doi.org/10.1523/JNEUROSCI.4596-08.2008

Toiviainen, P., Alluri, V., Brattico, E., Wallentin, M., & Vuust, P. (2014). Capturing the musical brain with Lasso: Dynamic decoding of musical features from fMRI data. *NeuroImage*, *88*, 170–180.

Rivers: A Rhythmic Journey to Varanasi

Shawn Mativetsky† (tabla), Schulich School of Music, McGill University, Montreal, QC, Canada
Pankaj Mishra (sarangi), Kolkata, West Bengal, India
†shawn.mativetsky@mcgill.ca

Program

Tabla solo in vilambit and madhya teental (traditional, arr. Shawn Mativetsky)

Program notes

Many people think of the tabla primarily as an accompaniment instrument. While it is true that the tabla plays a vital role as *taal*-keeper for vocalists and instrumentalists and is prominently featured in the accompaniment of *kathak* dance, it also has a rich solo tradition. Dating back to the late 18th century, the Benares *gharana* of tabla, founded by legendary musician Pandit Ram Sahai (1780-1826), has a long history of tabla artists who excel at solo tabla performance. Pandit Ram Sahai himself is said to have performed tabla solo for seven consecutive nights, in the court of Nawab Wazir Ali Khan in Lucknow. After this magnificent solo performance, Pandit Ram Sahai then returned to Varanasi (while musicians continue to use the name Benares, the city is now officially known as Varanasi), where he went into seclusion, making changes to the techniques, creating new forms, and composing new repertoire; these innovations resulted in what we now know as the Benares *gharana* of tabla.

Of course, performances of tabla solo are not relegated to performers of the Benares style alone; there are six distinct lineages, or *gharanas*, of tabla playing, each with its own variation on technique, repertoire and interpretation. The first tabla *gharana* to have developed is that of Delhi, followed by offshoots Ajrada and Lucknow. The Farukhabad and Benares *gharanas* connect back to the Lucknow *gharana*. The Punjab *gharana* is said to have developed independently and was until relatively recently a pakhawaj *gharana*.

Shawn Mativetsky's Guru, Pandit Sharda Sahai (1935-2011), was renowned for his solo performances. He was the torchbearer of the Benares *gharana*, a direct descendant of Pandit Ram Sahai. His tabla solos are still referred to as THE authoritative example of the Benares style of tabla playing, and he was highly respected by musicians both in and outside of the Indian music tradition.

The art of the tabla solo in the Benares *gharana* is highly developed and complex. A full-length tabla solo typically lasts upwards of one hour and demonstrates the full breadth of the repertoire. In a solo performance, the sonic nuances of the tabla are explored, showcasing the spontaneous creativity and virtuosic technique of the artist. The solo itself is made up of numerous compositions, strung together in a specific way, according to the performance practices of the tradition. Some compositions are fixed and are to be played as they have been passed down through the generations, while others form seeds for theme-and-variation improvisation. Some compositions are cyclical, fitting comfortably into the *taal*, while others are cadential, causing friction with the underlying rhythmic cycle, creating an effect of tension and release. One can view the tabla solo form like any good story, with an introduction, development, and conclusion. Whether the solo lasts five minutes or five hours, it should follow this overall structure in some way.

A tabla solo can be set to any rhythmic cycle, but the principal *taal* for tabla solos is most definitely *Teentaal*, a 16-beat cycle. This is because most repertoire for tabla has been initially composed (and taught) in *Teentaal*, and so a solo in this cycle will showcase the compositions in their original form. Additionally, the cycle's subdivision into four *vibhags* (subdivisions of the *taal)* of four *matras* (beats) each is intuitive for the listener to follow. Of course, other *taals* are sometimes used, such as *Rupaktaal* (seven beats) or *Jhaptaal* (ten beats), in addition to some less commonly used *taals*. However, these tend to require modification of the original materials in order to fit the structure of the given *taal*. This is why most tabla soloists tend to prefer *Teentaal* for their longer solo performances.

A tabla solo is typically accompanied by a repeating melody, called *lehra*, which serves to provide an aural outline of the underlying cycle. This melody loops throughout the entire solo, providing the performer and audience with a constant reference point. The melody will typically be the same length as the *taal* being performed, so for *Teentaal*, the melody is 16 beats long. For tonight's performance, the lehra will be expertly performed by Pankaj Mishra on sarangi, which is considered to be the ideal instrument for accompaniment of tabla solos.

The tabla solo opens with *bhumika*, which literally translates as 'introduction'. This allows the listener to settle in to the *taal*, while also allowing the soloist to warm up. During the *bhumika*, which is fully improvised, the *taal* is not alluded to in any obvious fashion, with phrases floating over the barline, so to speak. One will hear resonant sounds of the *dahina,* but no resonant sounds of the *baya*. Once we do hear the bass tones of the *baya*, this signifies the transition to the *uthaan*, which translates as 'rising up'. The *uthaan*, which is also improvised, allows the tabla player to explore various subdivisions of the beat, starting slowly, and moving up through several subdivisions. The *uthaan* ends with a *tihai*, a rhythmic cadence made up of a phrase repeated three times, where the last note of the last phrase lands on *sam* (beat one of the cycle). Beat one is vitally important, as it is the anticipation of 'one' that allows the tabla player to play with expectations and to create tension and release through rhythmic dissonance. The repeated *tihai* phrase moves around within the cycle, creating tension, which eventually is resolved when it finally lands on beat one.

The next phase of the solo, the development section, features improvisation of variations on fixed themes. These themes, and the rules and strategies for improvisation, are passed down through the generations. The principal theme-and-variation form is called *kaida*, which literally translates as 'rule'. In *kaida* improvisation, the tabla player must maintain the structure of the *kaida*, which is akin to poetry: drum poetry! The improvised variations can use only the thematic material found within the source *kaida*, and the improvisations must follow in a gradual sequence. The second variation is based on the first; the third is based on the second, and so on. This highlights the spontaneous creativity of the performer. The Benares tabla solo development typically begins with a special theme-and-variation composition called *Benarsi Theka*, which is a three-part improvisation alluding to the *alaap, jor,* and *jhala* heard at the opening of a *raga* performance. This will be followed by any number of other theme-and-variation forms, such as *kaida, rela, bant,* and *angrusthana.* The goal is to explore each composition to the fullest, pushing one's limits, and taking the audience on a rhythmic journey. Each theme-and-variation improvisation is concluded with a *tihai*. *Theka* is played between all items of the solo, as a return to home base, and functioning as a musical palette cleanser, so to speak.

Towards the end of the development section, fixed compositions are introduced. These must be performed as they have been passed down through the generations, with the utmost accuracy. Fixed compositions are typically played at at least two speeds. The composition is introduced slowly, at single speed, where each sound can be individually appreciated. This is followed by a performance of the exact same composition at double speed, which completely changes the perception of the musical material, as new textures and contours become apparent, and the beauty of the composition is revealed. Any number of these fixed compositions, known as *gat*, *fard*, or *gat-fard,* can be performed.

While the introduction and development sections typically take place at a slow tempo, known as *vilambit laya*, the concluding section is normally at a moderate tempo, called *madhya laya*. Here, fixed compositions are featured, and are often recited before being played. Any number of *tukra*, *chakradar tukra*, *paran*, or *chakradar paran* can be performed, depending on the length of the solo and the mood of the performer. One will usually play *theka*, then speak the composition, followed by a performance of the same, and then return to the *theka*, and so on. Speaking the composition allows one to hear the poetry of the *bols* before the composition is played on the tabla – the same composition, performed both orally and percussively!

One will often see audience members keeping *taal* at performances of Indian classical music and reacting audibly to *tihais* and other enjoyable musical moments, either with applause or vocal exclamations such as "*kya baat hai"* or "*Wah! Wah!"* This audience interaction is important, as this positive feedback lets the performers know that the audience is understanding the musical language and enjoying the performance. Just like at a jazz performance, if you hear something that you enjoy, you don't have to wait to the end to applaud.

Playing with Timbre

Cornelia Fales
Department of Folklore & Ethnomusicology, Indiana University, Bloomington, IN, USA
cfales@indiana.edu

Introduction

This paper is in ternary A-B-A form; that is, it has two themes and three sections. The A-theme consists of thoughts on the treatment of timbre in ethnomusicology, historically, at present, and in an ideal future. It reviews the infamous problems timbre presents to all scholars of music, and then discusses the special problems of timbre for ethnomusicologists. The B-theme attempts to demonstrate some of the points made in the previous section with examples from my own work on timbre in music from various parts of the world. I will describe some rough-and-ready solutions devised to work around timbre-related difficulties encountered in the field, and later in analyses of the music and interviews I collected. Many of these "solutions" were temporary and far from satisfactory, and so I return to the A-Theme to discuss the future of timbre studies in ethnomusicology. The paper ends with a call for interdisciplinary efforts to demystify a parameter of profound significance to the experience of music.

A-theme: Timbre in ethnomusicology

It will not surprise anyone that for most of the twentieth century, the parameter of timbre was as often overlooked in ethnomusicology as it was in the fields of Western music history and theory. Unique to ethnomusicology, however, are certain historically recurring patterns of reference to timbre that point to the special irony of its omission from studies of music where it is often the most salient parameter. This section will also briefly trace the history of timbre awareness in ethnomusicology as closely entwined with the development of audio technology for fieldwork and analysis.

The last several decades have seen increasingly sophisticated approaches to timbre research in the various disciplines concerned with music. Here I present a taxonomy of current directions in timbre research in ethnomusicology and related fields, ranging from experimental to theoretical to analytical to purely descriptive. The structure and details of this taxonomy demand reflection as to whether we even need more or different tactics in our analysis and theorizing. A comparison of current investigations of musical timbre to the long evolution of approaches to pitch in Western music leads me to suggest that timbre study is still in its infancy. This section concludes with a discussion of what timbre research still needs in order to begin accumulating data to support paradigmatic approaches equipped with a common vocabulary, a body of established premises, and a developed methodology.

B-theme: Projects in timbre research

In 2008, David Huron published an opinion essay in the journal Nature, warning that the vast diversity of world music was disappearing due to the rampant imperialism of Western music. I agree with Huron that part of the tragedy of music extinction is the concurrent loss of idiosyncratic modes of perception responding to idiosyncratic music. But there are still cultures in corners of the world whose music traditions feature ingenious manipulation of timbre and whose listeners seem to have developed special perceptual proclivities according to the requirements of their music. Experience with music cultures such as these has left me in a state of permanent humility in encounters with unfamiliar music, and a hardline strategic mandate to resist the assumption that what I am hearing is the same as what indigenous listeners are hearing.

In this section, I claim that timbre is one of the most perceptually malleable of musical parameters, and thus one of the most definitive in distinctions between traditional musical forms. Among my particular interests in musical timbre is the phenomenon of spectral fusion. The features that qualify individual concurrent acoustic components for admission into a vertical group are captured by the grouping principles enshrined in the theory of Auditory Scene Analysis (Bregman, 1994). Less well understood is the process of spectral fusion that glues individual components into a single sound. It turns out that there are degrees of fusion. Tones produced on a hand-made instrument by a skillful musician may loosen the glue of a timbral unit, releasing one or more harmonics to express their individuality in the form of pitch. Or, the same tone may subsume alien components that meet the requirements for membership only very imprecisely. A "universal" feature of the world's music cultures is some form of drone-based music in which the drone is calculated to fade into and out of auditory attention as it is cyclically captured into, then released by, the timbre of another tone.

Alert fieldworkers also discover that spectral fusion is subject to perceptual learning with the realization that timbral elements to which they may not be particularly sensitive, are nevertheless highly salient to indigenous listeners. The extraordinary flexibility of auditory perception has worked to shape another of my research interests: the development of techniques to elicit from indigenous listeners descriptions of what exactly they are hearing at any point in time. These techniques require some inventiveness, since many of the languages I deal with have little or no vocabulary regarding perceived musical sound, and since the information I want involves a degree of metacognition that is uncomfortable for listeners unaccustomed to the practice. This section of the paper will include other examples of timbral manipulation, as well as categories of timbre types implicated in traditional practices.

A-theme reprised: Conclusions

In a general sense, the future of timbre studies in nonempirical fields depends on a sharing of information across disciplines. Ethnomusicologists who intend to focus on specific uses of timbre need to understand the nuances of timbre production and perception; they need to keep up with developing research on timbre and emotion; they need proficiency in some sort of acoustic analysis and methods to relate the results of that analysis to the perception of the music reported by indigenous listeners. The fields of auditory science, on the other hand, could profit from the breadth of view offered by ethnomusicologists well and widely exposed to timbral practices of the world's musicians. Examples of auditory principles—both demonstrated and sometimes contradicted—are plentiful in the collective experience of field workers. Many of these same researchers have discovered aspects of timbre perception of importance to listeners about which little work has been done. (Why are the mechanics of perceptual fusion still mysterious, or worse, conflated with the operations of feature binding?) From a practical standpoint, scientists who want to do true cross-cultural research could benefit from the expertise of ethnomusicologists in the often-tricky business of dealing with indigenous musicians and listeners on their own turf. Finally, if the disciplines engaged in subjects relevant to timbre want to explore the territory separating learned and innate perceptual proclivities, I submit that there are few populations more suited to the project than ethnomusicologists.

References

Huron, D. (2008). Lost in music, *Nature*, 453, 456-457.
Bregman, A. (1994). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Timbre and Synthesis, the Promises of Freedom

Jean-Baptiste Barrière [1†] and Camilla Hoitenga[2]
[1] Image Auditive, Paris, France
[2] Cologne, Germany
[†]jb@barriere.org

Program

I – Talk by J.-B. Barrière: Timbre and Synthesis, the Promises of Freedom
II – Visual Concert, with Camilla Hoitenga, flutes & J.-B. Barrière, electronics & video
    J.-B. Barrière — *Hybris* (1988), for computer-generated sounds, 10’
    Kaija Saariaho — *NoaNoa* (1992), for flute, electronics & video, 10’
    J.-B. Barrière — *Crossing the Blind Forest* (2011), for bass flute, piccolo, electronics & video, 10’
    J.-B. Barrière — *Contemplation* (2016), for alto flute, electronics & video, 10’

The legacy of Computer Music pioneers, in particular the research and musical works of John Chowning and Jean-Claude Risset, has allowed, through analysis and synthesis, to open wide the field of all possible sounds to composition, with a freedom of imagination and possible levels of control never available before. However, synthesis remains today largely a utopia, since “ready made” sounds, based mostly on sampling and sound fluxes of granular techniques, are largely dominating production, in commercial music as in contemporary music. More challenging techniques, like physical modeling, additive synthesis and others proceeding from analysis, are indeed often considered too tedious, complex, and demanding.

Building one’s own sounds/instruments, should be among the first challenges of a composer, starting early during formative years. In this search, understanding and mastering timbre construction, perception and cognition, is a major step. Hence, the idea of continuity between the instrumental and synthesis domains, which provides a territory to explore starting from the existing map of the instrumental and vocal worlds, sedimented by thousands of years, in order to further develop knowledge, concepts and tools, and conceive sounds to create new expressive music. Through analysis, processing, and synthesis, timbre becomes a powerful metaphor to organize a path toward the intricate complexity of music.

During the eighties, I imagined a method for building timbre interpolations based on the following structural derivation scheme: modeling (where the model comes from the analysis of traditional acoustic instruments or arbitrary construction), hybridization (where two models or more are mixed together statically), interpolation or transition (where two models or more are mixed together dynamically), extrapolation (transformation preserving the recognition of the reference), and abstraction (transformation alienating the recognition of the reference). These different steps were realized with an original method of analysis, processing, and synthesis by computer, named “Models of Resonance,” that I developed with Yves Potard and Pierre-François Baisnée at IRCAM.[1] Through this approach, I was trying to set up the foundations of an esthetics of *hypermorphosis*, a morphogenesis of sonic materials that could be the ground for genetic mutations inside sound synthesis.

This approach is illustrated in the computer piece *Hybris* (1983). It found its origin in studies of the specific mode of playing wind instruments, where the musician sings or speaks and plays at the same time. The resulting sound is three-fold: voice filtered through the instrument’s resonances, normal sound of the instrument, and amplitude modulation between the two previous sources (which, depending of their respective pitches, can produce more or less inharmonic textures). I imagined inverting the process: having the sound of the instrument filtered by the vocal tract. This became a metaphor for the idea to filter any sound by any other sound, and we developed an analysis technique to generalize the process and built a large database of instrumental models.

[1] Barrière, J.-B., Potard, Y. & Baisnée, P.-F. (1985). Models of continuity between synthesis and processing for the elaboration and control of timbre structures. In *Proceedings of 1985 International Computer Music Conference,* ICMA, pp. 193-198. See also: Potard, Y., Baisnée, P.-F. & Barrière, J.-B. (1991). Méthodologie de synthèse du timbre : l’exemple des modèles de résonance. In Barrière, J.-B. (Ed.), *Le Timbre, métaphore pour la composition*, pp. 135-163. Paris: Christian Bourgois.

This was first implemented at IRCAM in 1985 in non-real time with the Chant program, then in 1986 in real time on the 4X synthesizer; then ported in 1988 on the Reson8[2], a specially designed processing card as peripheral of a Macintosh, controlled with Max; then during the 90s, this code was adapted by Adrian Freed into Max/MSP in the object resonators at CNMAT. I continue to develop its applications with Thomas Goepfer. The analysis part was made available as the application ResAn, part of Diphone Studio.

The processing/synthesis part of the Models of Resonance consists mainly of a bank of filters controllable in frequency, amplitude and bandwidth, which can be excited by concrete or instrumental sound (or e.g., with physical modeling in Modalys) and by noises sources (e.g., extracted with SVP/Audiosculpt), thereby allowing one to merge processing and synthesis. One interest of Models of Resonance is precisely to allow one to conceive of the continuity between processing and synthesis, and therefore to advance at the same time the continuity between instrumental and electronic writing.

Models of Resonance have been used widely for thirty years by composers around the world, and from the beginning by Kaija Saariaho, who worked with them in nearly all the electronic parts of her pieces, including *Noa Noa* (1992), for flute and electronics,[3] *Io*, *Amers*, etc., or her opera *L'Amour de loin*. She built models from the analysis of complex natural or instrumental sounds (multiphonics, speech, etc.), and compositionally (as chords directly in Finale, or generated and manipulated in OpenMusic, and with the help of perceptual models such as Terhardt's masking algorithms[4]), emphasizing ambiguities between timbre and harmony, working on interpolations between "instrumental" and "compositional" models.[5]

I have used Models of Resonance in a number of live mixed pieces, such as *Crossing the Blind Forest* (2011) and *Contemplation* (2016), as well as in many virtual reality or interactive musical and visual installations, or in multimedia shows, such as *100 Objects to Represent the World* (1997) or *The Art of Change Opera* and *38th Parallel* (both coming in 2019), where either the instrumentalist or the interactor explores underlying timbre maps, designed with various analysis descriptors.

Since the end of the 90s, in search of bridges between music and visuals to conceive "visual concerts," I also use all sorts of filters in the visual domain, working as metaphors and extensions of filters in music: the musicians' faces and gestures are filmed with cameras and are filtered by various image sources, like the voice being filtered by the flute in the instrumental part, or the flute being filtered by the voice in the electronics.[6] In this way, timbre proliferates into color and forms, and music resonates into images.

---

[2] Barrière, J.-B., Baisnée, P.-F., Baudot, M/-D., & Freed, AA. (1989). A multi-DSP system and its musical applications. In *Proceedings of 1989 International Computer Music Conference, Columbus, Ohio,* Computer Music Association, Berkeley.

[3] Barrière, J.-B., Chabot, X. & Saariaho, K. (1993). On the realization of *NoaNoa* and *Près*, two pieces for solo instruments and IRCAM signal-processing workstation. In *Proceedings of the 1993 International Computer Music Conference*, ICMA, pp. 210-213.

[4] Terhardt, E., Stoll, G. & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *The Journal of the Acoustical Society of America*, 71/(3), 679-688. And its implementation in Iana; Todoroff, T., Daubresse, É. & Fineberg, J. (1995). IANA: A real-time environment for analysis and extraction of frequency components of complex orchestral sounds and its application within a musical realization. In *Proceedings of the 1995 International Computer Music Conference*, ICMA, pp. 292-294.

[5] Saariaho, K. (1987). Timbre and harmony: Interpolations of timbral structures. *Contemporary Music Review* 2(1), 93-133.

[6] Barrière, J.-B. & Barrière, A. (2016). When music unfolds into image: Conceiving visual concerts for Kaija Saariaho's works. In Y. Kaduri (Ed.)., *The Oxford Handbook of Sound and Image in Western Art*, Oxford University Press.

Concert program notes

Jean-Baptiste Barrière — *Hybris* (1988), for computer-generated sounds

Hybris (from the Greek "hybrid," "mixity," but also "loss of all measure") is a piece on timbre interpolation processes as they are only realizable with computer sound synthesis, and in particular with the Models of Resonance. *Hybris* explores a structural derivation scheme from modeling to abstraction inside an arbitrarily defined timbre space, and more specifically between models derived from double bass and tam tam sounds. These models were analyzed for their content in terms of both timbre and harmony and were then manipulated in the computer-aided composition environment Esquisse (an ancestor of Patchwork/PWGL and OpenMusic), so as to serve as a basis for the construction of the different processes contributing to the elaboration of the form of the piece. The piece evolves paradoxically between "program music" and a form of constructivism: it can be perceived as the story of two very rich sonic objects that elaborate together complex relationships, and cross a variety of musical situations, reconsidering permanently the limits between the real and the artificial; or as well as the systematic exploration of relations between timbre and harmony. When composing the piece, I was often thinking of the last chapter of the book by Primo Levi, *The Periodic System*, where the writer tries to imagine a novel based only upon the adventures of an atom of carbon, its multiple transformations and becomings. In a similar fashion, I would like to think of *Hybris,* as a—rather dramatic—story, on the emancipation of two resonances.

Kaija Saariaho — *NoaNoa* (1992), for flute, electronics & video

In NoaNoa (meaning 'Fragrant' in Tahitian), Kaija Saariaho wanted "to write down, exaggerate, even abuse certain flute mannerisms that had been haunting me for some years, and thus force myself to move onto something new. The title refers to a wood cut by Paul Gauguin called NoaNoa. It also refers to a travel diary of the same name, written by Gauguin during his visit to Tahiti in 1891-93. The fragments of phrases selected for the voice part in the piece come from this book." The flute part covers a wide range of techniques, including breath, speaking, and singing in the instrument (Hoitenga, 2011). The electronic part was originally realized with Max on the IRCAM workstation on a Next machine, as a pedagogical piece to demonstrate the possibilities of real-time electronics, and includes a rather large palette, with delays, harmonizers, sampling, Models of Resonance, spatialization, as well as score-following (abandoned in later versions). It was premiered at the Darmstadt Summer courses in 1992 and has since become one of the most played contemporary flute pieces. The video part was premiered in 1998 at the Contemporary Art Museum of Zurich, as part of the Visual Concerts series I then started to develop. It includes visual materials developed from Gauguin etchings and especially shot footage, processed through a variety of techniques, in particular various types of filtering (often called "masks") meant to work as "analogons" and prolongations to the ones in the instrumental part (voice being filtered by the flute), as well as the electronics part (flute and voice being filtered by the Models of Resonance).

Jean-Baptiste Barrière — *Crossing the Blind Forest* (2011), for bass flute, piccolo, electronics & video

Crossing is a piece composed especially for and dedicated to Camilla Hoitenga, who gave the world première in New York in 2011. It is a kind of evocation of Maurice Maeterlinck's play *Les Aveugles* (*The Blind*), itself based loosely on the similarly titled painting by Pieter Bruegel the Elder. It revisits and develops the flute materials I composed for the multimedia show *Deux Songes de Maeterlinck d'après Bruegel* (*Two Dreams of Maeterlinck after Bruegel*), premiered by Hoitenga in Marseilles in 2007. In this piece, the flutist is in a certain way playing the character of the blind people of Maeterlinck's play, lost in the forest, therefore bound to death. She is lost in an unknown world and must heighten all her sensations and skills in order to try to survive the dangers all around her. The virtuoso flute playing is challenged by electronic transformations (various types of delays and filtering, including spectral delays and Models of Resonance) in an uncertain conflict, one whose outcome may be left open, undecided. Images, mixing cross-transformations of the live performance of the flutist with treated images of forests devastated by storm, are meant to represent and accompany this quest undertaken by means of the senses.

Jean-Baptiste Barrière — *Contemplation* (2016), for alto flute, electronics & video

Contemplation, presented in a version especially arranged for this concert, is inspired by a famous poem of Victor Hugo. The French version of the text is "orchestrated" in the electronics part by exploring a complex timbre space, while the flutist delivers the English version through the flute playing, processed through spectral delays, Models of Resonance, and spatialization routines. The visual part is inspired by the less known but nevertheless fantastic drawings of Hugo, mostly showing mysterious landscapes of ruins and desolation, which are confronted and blended with images of the migrant crisis in Europe and the destructions of Syrian cities. This poem is considered in this piece as a departure point for an interrogation about the status of mankind in the real-time information age: submitted to a constant data storm, we are drowning in information, and are just left to contemplate. But what can we actually do, faced with realities which escape us?

Victor Hugo — "Les contemplations," Livre 6, *Au bord de l'infini*, XIV (translated/adapted by J.-B. Barrière)

O gouffre ! l'âme plonge et rapporte le doute.
Nous entendons sur nous les heures, goutte à goutte,
Tomber comme l'eau sur les plombs ;
L'homme est brumeux, le monde est noir, le ciel est sombre ;
Les formes de la nuit vont et viennent dans l'ombre;
Et nous, pâles, nous contemplons.

Nous contemplons l'obscur, l'inconnu, l'invisible.
Nous sondons le réel, l'idéal, le possible,
L'être, spectre toujours présent.
Nous regardons trembler l'ombre indéterminée.
Nous sommes accoudés sur notre destinée,
L'oeil fixe et l'esprit frémissant.

Nous épions des bruits dans ces vides funèbres ;
Nous écoutons le souffle, errant dans les ténèbres,
Dont frissonne l'obscurité ;
Et, par moment, perdus dans les nuits insondables,
Nous voyons s'éclairer de lueurs formidables
La vitre de l'éternité.

O abyss, the soul plunges and brings back doubt.
We hear the hours, drop by drop,
Fall upon us like water on lead;
Man is foggy, the world is black, the sky is dark;
The forms of the night go back and forth in the shadow;
And us, pale, we contemplate.

We contemplate the obscure, the unknown, the invisible,
We probe the real, the ideal, the possible,
The being, always present spectre.
We watch the undetermined shadow trembling,
We lean on our destiny,
The eye fixed, and the mind vibrant.

We spy on noises in these mournful voids;
We listen the breath, wandering in the darkness,
Whose obscurity shivers;
And, by moments, lost in the fathomless nights,
We see illuminated by a formidable flare
The window of eternity.

Hoitenga, C. (2011). *The flute music of Kaija Saariaho. Some notes on the musical language.* http://hoitenga.com/site/flute-music-of-kaija-saariaho

ORAL AND POSTER PRESENTATIONS

Black Power Voices: Examining Timbre and Polyvocality in Oscar Brown, Jr. and Nina Simone

Heather Buffington-Anderson
Music, Claflin University, Orangeburg, South Carolina, USA
hbuffingtonanderson@claflin.edu

Aims/goals

This paper examines the ways in which artist-activists Oscar Brown, Jr. and Nina Simone manipulate vocal timbre as a strategy of resistance within the context of the Black Power and Black Arts Movements. Through analyses of Brown, Jr.'s "Rags and Old Iron," and Simone's "Four Women," this paper aims to highlight how these vocalists present multiple timbres, or "voices." In hearing these as polyvocal performances, we may understand timbre as a site for resisting and disassembling expectations of Black vocality as well as an activist strategy for reimagining Black identity.

Background information

Challenging the notion that timbre is an "unmanipulable attribute," Nina Eidsheim's research demonstrates that the voice can emit many timbres. Eidsheim explains that because the voice arises from inside the body, vocal sounds are commonly understood as natural expressions, which becomes a metaphor for unified subjectivity. Timbral character, Eidsheim explains is heard "according to schemas of racialized, gendered or otherwise categorized bodies in accordance with the values of the given society" (2008, 178). Her fieldwork on vocal morphology demonstrates that while is no distinction between voices of different races, and that such notions are rooted in biological essentialism, critics continue to locate Blackness in timbre as a way to name racial difference.

Evie Shockley explains that gender norms and hierarchies figured heavily in constructions of Blackness within the Black Power era. Black music was at the center of the Black Power Movement and theorizations of a Black Aesthetic. The voice and vocality were central to the aesthetics of the Black Arts Movement as poets often performed their work shifting "from casual talk to heightened speech and sometimes to musical chant," while free jazz musicians sounded the extremes of vocal capacity evoking sounds of screams, cries, and wails (Keyes 2004, 34). However, the discourse of the music of this era often figures male instrumentalists' sounding of the human voice – particularly John Coltrane's scream – as not only innovative, but also as a powerful reference for many Black Arts poets and theorists. As such, vocalists are absent in discussions of radical black aesthetics and politics. Focusing on the voice, and particularly timbre, provides an opportunity to rethink radical Black politics as hypermasculine, and may bring into relief narratives of Black vocalists who were critical to developing complex and perhaps competing formulations of radical Black identities.

Methodology

Within this paper, transcriptions of recorded and live performances of "Rags and Old Iron" and "Four Women" are presented in notational and spectrographic formats and analyzed for timbral variations. These analyses are read in dialogue with a reception study of Brown, Jr.'s and Simone's critical reception from both mainstream and Black press publications in an effort to understand the ways in which various audiences hear and interpret the vocality and identities of the performers. The musical analyses and reception studies are interpreted through Evie Shockley's notion of polyvocality in black women's poetry. Shockley's theoretical framework builds upon Mae Gwendolyn Henderson's trope of "speaking in tongues," which describes the ability black women develop to "speak in a plurality of voices as well as in a multiplicity of discourses" (Henderson 1989, 22). Speaking in tongues – which involves both speaking and hearing – allows for a simultaneity of languages and received meanings. Henderson argues that black women have developed such a heteroglossia to be heard by others simultaneously as homogenous and heterogeneous to their audiences. Shockley's model allows for diverging polyvocal projects with varying purposes and objectives; such a framework avoids essentializing black women's voices, and responds to the voice which changes tone, timbre, and language over time and space. My modulated approach of polyvocality is centered around vocal timbre and I suggest both Brown, Jr. and Simone present multiple multiplicity of voices in various forms that are spoken, chanted, sung, and texted. Through the lens of polyvocality, it is possible to understand how artists-activists reformulate racial and gender hierarchies within the Black Power Movement.

Results

Both Oscar Brown, Jr. and Nina Simone accesses the notion that timbre is heard and seen as a construction of identity within their performances; however, each vocalist disrupts the notion of timbre as something that is essentialized – as something fixed – by quickly donning and shedding various vocal guises. By shifting timbral attributes including articulation, diction, tone, vibrato, and register, Brown and Simone present characters in rapid succession. Both vocalists theatrically sound fragmented musical codes to reference the class, gender, race, age, or region specific of their characters. The use of sonic caricature, which is often combined with humor and irony, challenges negative racialized stereotypes. By fragmenting and reassembling these codes, Brown and Simone collapse sonic expectations of Black vocality. Consideration of the vocalists' reception shows that while these polyvocal performances were challenging to mainstream critics – many of whom referred to Brown and Simone as artists who "defy categorization" – many Black critics and intellectuals recognized their use of timbre and genre as aesthetic strategies to reimagine Black identity and radical Black politics.

Conclusions

Timbral variation is central to both Brown and Simone in their process of reimagining Black manhood and womanhood respectively within the Black Power era. Using timbre to achieve mutability and multiplicity in the voice, Simone deconstructs four stereotypes of black women in her composition, including the Mammy, the Tragic Mulatta, Jezebel, and the Sapphire. Similarly, Brown, Jr. alternates between various vocal timbres within "Rags and Old Iron" to effectively challenge singular notions of Black masculinity within the Black Power Era. Timbral analysis and theorizations of timbre provide a more nuanced understanding of the ways in which musicians resist and protest gendered and racialized constructions of the voice.

References

Eidsheim, N. S. (2008). *Voice as a Technology of Selfhood: Towards an Analysis of Racialized Timbre and Vocal Performance* (Ph.D.). University of California, San Diego. Retrieved from http://search.proquest.com.ezproxy.lib.utexas.edu/docview/304655431/abstract/197368F59B8413EPQ/1

Eidsheim, N. (2012). Voice as action: Towards a model for analyzing the dynamic construction of racialized voice. *Current Musicology*, *93*. Retrieved from http://currentmusicology.columbia.edu/authors/nina-eidsheim

Eidsheim, N. S. (2011). Marian Anderson and "sonic blackness" in American opera. *American Quarterly*, 63(3), 641–671. https://doi.org/10.1353/aq.2011.0045

Keyes, Cheryl Lynette. 2004. *Rap Music and Street Consciousness.* University of Illinois Press.

Shockley, E. (2011). *Renegade Poetics: Black Aesthetics and Formal Innovation in African American Poetry*. Iowa City: University of Iowa Press.

The Social Life of Timbre: Discussing Sound Color in Online Guitar Communities

Rebecca Flore
Department of Music, University of Chicago, Chicago, Illinois, United States
rflore@uchicago.edu

Aims/goals

My presentation seeks to answer two questions: what language do electric guitar enthusiasts use to discuss timbre, and how are these timbral descriptors codified, learned, and spread? To answer these questions, I will focus my attention on online communities of these guitarist-audiophiles. In doing so, I hope to outline some of the ways in which knowledge flows in a modern, socially networked context.

Background information

Timbre is a slippery, imprecise, ineffable phenomenon, elusive for musicians and non-musicians alike. We typically resort to comparisons, especially defining timbre as the difference between sound A and sound B. For those with a strong background in acoustics—for instance, composers of electronic music—timbre can be explained and quantified in terms of perceptual and acoustical elements. For dedicated consumers of boutique guitar effects, however, timbre is qualitative. Such enthusiasts are often both musicians, usually playing guitar or bass, and audiophiles, meaning that they take a great interest in sound quality and high fidelity. These guitarist-audiophiles collect not only gear, but simultaneously create a substantial subjective vocabulary to describe the timbres produced by these technologies.

Methodology

In order to compile a vocabulary of tone descriptors, I surveyed language used on a variety of online havens for guitarist-audiophiles: threads on online guitar forums such as The Gear Page, Off-Set Guitars, and Rig-Talk; listings on online gear marketplace Reverb; interviews with guitarists on YouTube; and the websites of effects manufacturers like Z.Vex and Earthquaker Devices. Additionally, I use these online interactions to understand how guitarist-audiophiles can translate terminology into technique. As Leslie Gay notes, rock musicians often develop their skills by "listening to and imitating rock recordings, acquiring an initial repertory and sense of what constitutes a good rock sound" (1998, p. 84). The internet has created new avenues for this kind of attentive, obsessive listening. I explore two such avenues: audio-visual demonstrations done by authorities such as famous musicians and effects manufacturers, and fan-made gear demo videos.

Results

From my survey, I compiled a list of terms and which pedals they were used to describe tone and found that the specialized terminology surrounding tone falls into three categories: metaphorical language about timbre, language that describes the changes to the electrical signal that produce the sound, or the acoustical features of the sound itself. The most interesting of these categories is the metaphorical language because these metaphors require a relational frame (Zbikowski, 2015, p. 175), one that is acquired by matching up the heard quality of a sound with its common description in words. Using both authoritative and fan-made gear demo videos, a viewer may learn by listening, transferring whatever descriptive understanding they have of timbral effects onto the actual sound being produced. Whereas the authoritative demo videos foreground certain settings so that viewers may replicate the tones of well-known musicians, fan-made demos are meant to showcase the variety of sounds available in each piece of gear, to provide expert users with ideas for formulating unique tone colors.

Conclusions

First and foremost, the timbral vocabularies of these communities of guitarist-audiophiles provide ample opportunity to fortify the theory that metaphor is absolutely crucial to production and dissemination of timbral knowledge. Furthermore, fully thinking through the use of timbre metaphors can help us tune our attention on what it means to repurpose language from the visual or tactile modalities to describe sound.

Second, tracing these vocabularies provides an existing corpus of timbral descriptors *in use by musicians*. As with many aspects of sound and music it is difficult to translate them into words, but guitarist-audiophiles have developed a system of signifiers for sound that can be learned. While these communities are centered around the musical practice of electric guitar and bass playing generally, the online aspect of these communities allows them to exist beyond genre divisions. Musicians whose tastes range from metal to punk to jazz to Christian rock to noise and experimental music can exchange information and learn from one another. By studying the vocabularies that guitarist-audiophiles have built around the concept of tone, we get a rich view into a culture that has gained a qualitative, implicit, practice-based understanding of timbre.

References

Baym, N. K. (2012). Fans or Friends? Seeing Social Media Audiences as Musicians Do. J*ournal of Audience & Reception Studies* 9(2), 286–316.

Burns, A.-M., Bel, S., & Traube, C. (2017). Learning to Play the Guitar at the Age of Interactive and Collaborative Web Technologies. Presented at the Sound and Music Computing Conference, Espoo, Finland.

Carfoot, G. (2006). Acoustic, Electric and Virtual Noise: The Cultural Identity of the Guitar. *Leonardo Music Journal* 16, 35–9.

Fales, C. (2002). The Paradox of Timbre. *Ethnomusicology* 46(1), 56–95.

Fales, C. & Berger, H. (2005). 'Heaviness' in the Perception of Heavy Metal Guitar Timbres: The Match of Perceptual and Acoustic Features Over Time. In *Wired for Sound: Engineering and Technologies in Sonic Cultures* (pp. 181–97). Middletown, Conn: Wesleyan University Press.

Gay, L. C. (1998). Acting Up, Talking Tech: New York Rock Musicians and Their Metaphors of Technology. *Ethnomusicology* 42(1), 81–98.

McAdams, S. (2013.) Musical Timbre Perception. In D. Deutsch (Ed.), *The Psychology of Music*, 3rd ed. (pp. 35–67). San Diego: Elsevier.

McAdams, S., Douglas, C., & Vempala, N. N. (2017). Perception and Modeling of Affective Qualities of Musical Instrument Sounds across Pitch Registers. *Frontiers in Psychology* 8(153), 1–19.

Scotto, C. (2017). The Structural Role of Distortion in Hard Rock and Heavy Metal. *Music Theory Spectrum* 38(2), 178–99.

Traube, C., & Lavoie, M. (2008). The Guitar as an Extension of the Voice - Phonetic Gestures Underlying Guitar Timbre Perception and Description. *The Journal of the Acoustical Society of America*, 123(5), 3657–3657.

Zbikowski, L. M. (2002). *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. New York: Oxford University Press.

Zbikowski, L. M. (2015). Musical Semiotics and Analogical References. In *Music, Analysis, Experience: New Perspectives in Musical Semiotics* (pp. 167–83). Leuven: Leuven University Press.

Timbre: Aesthetics of Vibration

Isabella van Elferen
Department of Performing Arts, Kingston University London, UK
i.vanelferen@kingston.ac.uk

Aims/goals

Aiming to formulate a theory of timbre, this music-philosophical paper seeks to connect the two existing, but opposed, idealist and materialist schools of timbral reflection (see Background Information). My theorisation of timbre is conceived with the paradoxes of timbre as its main premise: as tone colour unites the material and the less material, so timbral theory should unite materialism and idealism. I argue that tone colour is both a materialist object *and* an idealist Thing: tone colour involves both the physical and the ephemeral, and it inhabits the gap between those opposites. This paradoxical position is theorised through a combination of Michel Chion's acoulogy of sonic vibration and Jane Bennett's vital materialism. Timbre, I propose, is a meta-acoustic vibration of sound waves crossing over from instruments to listening bodies to minds and imagination, which are connected in a vibrational continuum. The sonic energy carried by timbre's vibrating sound waves are an affectively powerful vitality interacting with musical assemblages of humans and nonhumans. In aesthetic experience, timbre's vibrant vitality engenders "tone pleasure" (Herder) in a "*sentiment sonore*" (Lyotard) that includes material practice, identity and politics as well as affective lines of flight and sublime aporia. The paper will focus on timbre in Western music since the Enlightenment.

Background information

The heterogeneous aesthetic event of timbre is hard to define, and theorisations of timbre are scarce. There are a number of studies on discrete aspects of timbre. Research on timbre production, timbral envelope, and harmonic spectrum represent established strands of organology, psychoacoustics, and creative practice (Parker, Schaeffer, Stockhausen). Other studies focus on timbre's relation to corporeal identity, to genre, or to the meanings attributed to certain tone colours (Barthes, Dolar, Moore). Finally, a number of composers and philosophers have reflected on the sublime ineffability of timbral perception (Jankélévitch, Lyotard, Nancy, Schoenberg). Timbral theory, as a result, is split into two schools of thinking: a materialist approach studying the circumstances of sound production on the one hand, and, on the other, an idealist approach studying the effects it has on the listening experience while leaving timbre "in-itself" at a distance. While each of these approaches has yielded valuable results, their practical and intellectual exclusivity has led to an undesirable musicological dead end: the two elements that form the entwined basis of the timbral event have been irreconcilably separated in timbral theory.

Methodology

Timbre: Aesthetics of Vibration seeks to explore not just one of the most enigmatic aspects of music—tone colour—but, through that, also one of the most often-debated questions of music epistemology: the simultaneity of the material and the immaterial in musical aesthetics. In order to achieve an overview of the existing field of timbral research, assess it with critical precision, and offer a new theorisation of this complex subject, the paper combines musicological and philosophical theories.

Results

Timbre can only be assessed with critical precision when we acknowledge, and operationalize, its paradoxical qualities. Timbre must be conceived as an aesthetic moment (Dolan), a "distribution of the sensible" (Rancière) affecting the material and the less material at the same time

Conclusions

In its paradoxical inclusivity of seemingly opposed characteristics, timbre is representative of a great deal of ongoing debate in music philosophy. Timbre is not the only aspect of music that is characterised by a simultaneity of material and immaterial characteristics, and of realist and idealist evaluations. Lacking the disciplining structures of discursive embedding, tone colour's material/immaterial vitality and materialist/idealist elusiveness are more highlighted than, say, melody's; but, arguably, any separate building block of music is similarly paradoxical and unknowable. Musicology has not reached consensus on the question of whether music should be considered material or immaterial, and whether it wants to think music in realist or idealist terms (Dodd, Hamilton). Overcoming timbral binarism, therefore, demands an engagement with one of the major unsolved problems of music epistemology, and that is what the paper sets out to do. Timbral aesthetics, I argue, affords interaction with musical epistemology itself.

References

Barthes, R. (1977) "The Grain of the Voice", 179-189 in *Image Music Text: Essays selected and translated by Stephen Heath*. London: Fontana Press.

Bennett, J. (2010) *Vibrant Matter: A Political Ecology of Things*. Durham and London: Duke University Press.

Boulez, Pierre (1987) "Timbre and Composition – Timbre and Language". *Contemporary Music Review* 2: 161-171.

Chion, M. (2016) *Sound: An Acoulogical Treatise*. Duke University Press.

Dodd, J. (2007) *Works of Music: An Essay in Ontology.* Oxford: Oxford University Press.

Dolan, E. (2013) *The Orchestral Revolution: Haydn and the Technologies of Timbre*. Cambridge: Cambridge University Press.

Dolar, M. 2006. *A Voice and Nothing More*. Cambridge, MA: MIT Press.

Hamilton, Andy (2007) *Aesthetics & Music*. London: Continuum.

Herder, J. G. ([1769] 1987) *Viertes Wäldchen,* in *Werke*, vol. II: *Herder und die Anthropologie der Aufklärung*, ed. W. Pross. Munich: Carl Hanser Verlag.

Jankélévitch, V. ([1983] 2003) *Music and the Ineffable*. Transl. Carolyn Abbate. Princeton: Princeton University Press.

Lyotard, J-F. (1991) *The Inhuman*. Cambridge: Polity Press.

Moore, A.F. (2012). *Song Means: Analysing and Interpreting Recorded Popular Song*. Farnham: Ashgate.

Nancy, J.-L. (2007). *Listening*. New York: Fordham University Press.

Parker, Barry (2009). *Good Vibrations: The Physics of Music.* Baltimore: The Johns Hopkins University Press

Rancière, J. (2009) *Aesthetics and its Discontents*. Cambridge: Polity Press.

Schaeffer, P. (1952) *A la recherche d'une musique concrete*. Paris: Éditions du Seuil.

Schaeffer, P. (1966) *Traité des objet musicaux*. Paris: Éditions du Seuil.

Schoenberg, A. (1922) *Harmonielehre. II. vermehrte und verbesserte Auflage.* Vienna: Universal-Edition.

Stockhausen, K. (1963-1984) *Texte zur Musik*. Cologne: Verlag M. DuMont.

Timbre and Formal Functions in Edgard Varèse's *Amériques*

Federica Di Gasbarro
Swiss National Science Foundation (SNF) Postdoc at Université Côte d'Azur, Nice, France
f.digasbarro@gmail.com

Aims/goals

Starting in the 1930s, Edgard Varèse expresses a new conception of timbre, "changed from being incidental, anecdotal, sensual or picturesque, it would become an agent of delineation [...] and an integral part of form" (Varèse, 1966, 11-12). The present contribution aims to show how the first score written after his arrival in New York, *Amériques* (1918-21), while bearing witness to maturity reached by his harmonic thinking, has to be considered an ideal test bed for a new figural and sonic understanding of music that was significantly developed in his later scores, including the revision of this *opus 1* (1927-29). Indeed, in order to clarify its "thematic" architecture, Varèse here calls into play new means as proper form-building factors, first of all orchestration and timbre. Hence, challenging the structural role of the thematic relationships in favor of 'figural' ones, the composer launches a fundamental reconsideration of the notion of timbre as one of the strongest shaping forces replacing structural harmonic functions.

Background information

Although former studies have noted motivic peculiarities within *Amériques* (Nanz 2003), the dominating discourse concerning the piece's overall form remains the so-called "block form." A combined analysis of harmonic-, contour- and coloring-factors, however, reveals how Varèse takes advantage of the concept of thematic development as a primary means to connect materials and delineate sections and parts according to a teleological conception of musical time and form. On the other hand, Varèse's timbral innovations have until now been primarily investigated in his percussion ensemble score, *Ionisation* (1931), considering both the qualities of the percussion sounds and their constructive potential (Wen-chung 1979, François 1991, Youatt 2012), while an interesting a multi-parametric analytic perspective can be found in Lalitte's considerations on an enlarged concept of "harmonie-timbre" in Varèse's chamber works (Lalitte 2003).

Methodology

Due to the lack of sketches for *Amériques* and of an explicit compositional theory, score analysis and comparison of the genetic stages of the text were chosen as a primary approach. The main sources for this investigation are not only the manuscript and the different printed scores, but also the annotations contained in the instrumental parts, on Stokowski's conducting score and on Varèse's "reference copies.". The common post-tonal analytical approaches have been enlarged, including a wider set of parameters responsible for the sound unit perceived in its integrity (Tenney 2015, 36-60).

Results

The form of *Amériques* proved to be structured in four parts, whereby the 11 main variants added in the different revision phases clearly contribute to strengthen precisely this structural characteristic. Part I is divided into three sections according to sonata-form functions, the first phrase (mm. 1-8) exposing the essential thematic ideas of the whole piece. This material supports the developmental process through its changes in contour and timbre, acting as shaping signal for the formal articulations. Not only in the revision Varèse altered this first part by cutting the less homogeneous material and inserting new occurrences of the main themes. Toward its end (RN 23-24), through an alteration in the orchestration— by mixing and thickening the layers of a texture now perceived as much denser—he also succeeds in replacing a weak post-cadential prolongation with a forward driving force that prepares the dense texture at the beginning of the next part. Alternating thematic and transitional sections, part II assumes the role of a supplementary development. Yet, while the former motivic materials reoccur in vertical densifications weakening the perception of discernable contoured entities, the transitional passages reinforce precisely the amorphous aspect and the timbral relevance of the textures, lessening the feeling of a temporal trajectory. Part III gives a glimpse of the function assigned to the orchestration as a means of delineation of two "volumes of sound," each characterized by a homogeneous, stable timbre. The independence between the formal articulation points in the woodwind melody (high register) and the continuous ostinato-accompaniment in lower instruments generates a feeling of instability, conferring on this Scherzo-part the function of tension relaxation and preparation for the finale. The last part (IV), almost untouched in the revision, features a mobile arrangement of sound layers changing in density and timbral attributions. The continuous layer of the brasses acts as stable "recto tono" that rears up in figural densifications; on it, tutti chordal explosions interact with irregularly reemerging "fragments of memory" derived from earlier material but restated here without any further development. Analyzing the layers according to color and density one can distinguish formal units and phraseological functions and perceive a directional motion toward the final saturation.

Conclusions

Amériques, with its prolonged genesis and successive revisions, represents a fortunate case study for the investigation of the development and emergence of Varèse's peculiar mature language. In the four-stage itinerary developed in the score he not only stages the very transition from a thematic to a figural musical conception, but also gradually discloses the fundamental principles that will be established as the implicit elements of his style. Such transformations act at the levels of both the form, from clearly hierarchized structures towards less immediately intelligible ones, and of the sonic substance, which loses its nature of "sound idea" and generates rather "figural entities," potentially containing infinite "musical actualities". In this context, the non-harmonic parameters and especially the timbral qualities become primary factors in the definition of temporal and spatial densities, thus replacing the harmonic-factor as the most important shaping-factor. From *Amériques* onwards, specific figural types crystallized as prominent building material in Varèse's lexicon, falling under three categories of sound entities: *points* (temporally-spatially concise), *lines* (temporally extended) and *planes* (sudden or gradual spatial extended sound entities). Samples taken from two "steps onwards" just before and after *Amériques*' revision (*Intégrales* and *Ionisation*) show how Varèse further develops this figural thinking towards a "composing out" of even single sounds. Finally, the migration of nearly same figural actualities from one piece to another becomes proof of a new role entrusted to the figure: no longer a distinctive thematic element of a single score because of its inherent contour features, but rather a pretext to look into the sonic results of the interaction between multiple sound parameters. This way of "composing-analyzing" a musical substance permitted Varèse to lay the foundations for a new musical thinking that affected generations of composers.

References

François, J.-Ch. (1991). Organization of scattered timbral qualities: A look at Edgard Varèse's Ionisation. *Perspectives of New Music*, *29(1)*, 48-79.

Lalitte, Ph. (2003). L'architecture du timbre chez Varèse : la médiation de l'acoustique pour produire du son organisé. *Analyse musicale*, *47*, 34-43.

Nanz, D. (2003). *Edgard Varèse. Die Orchesterwerke*. Berlin: Lukas Verlag.

Tenney, J. (2015). *From Scratch. Writings in Music Theory*. Urbana: University of Illinois Press.

Varèse, E. (1966), The Liberation of Sound. *Perspectives of New Music*, *5(1)*, 11-19.

Wen-chung, Chou (1979). Ionisation: the function of timbre in its formal and temporal organization. In S. Van Solkema (Ed.), *The New Worlds of Edgard Varèse* (26-74). New York: ISAM.

Youatt, A. P. (2012). *Analysing Edgard Varèse's Ionisation using digital spectral analysis*. MA Thesis University of Arizona.

Phrasing through Timbre: Timbral Variety, Figural Coherence and Formal Evidence in Stockhausen's *Kontakte* and Ferneyhough's *Time and Motion Study II*

Pascal Decroupet
Arts Department (Music), Université Côte d'Azur, Nice, France
pascal_decroupet@hotmail.com

Aims/goals

The aim of the present contribution is to compare the ways in which Karlheinz Stockhausen in *Kontakte* (for piano, percussion and tape, 1960) and Brian Ferneyhough in *Time and Motion Study II* (for cellist and live-electronics, 1976) have developed compositional strategies to generate figures synthesizing different timbral means within one musical gesture. If both composers use permutational devices for distributing these enriched timbres to ensure a certain aesthetically intended irregularity and imprevisibility, they both have limited themselves to a reduced number of figure types to generate coherence at different levels of the musical organization. Furthermore, the distribution of specific variants of their figure types as well as their adaptation to the respective contexts within the overall process reveal formal preoccupations at both a microscopic and macroscopic level. These combinatorial means of pre-organization help not only to understand the compositional process but serve also to perform these pieces by providing categories to elaborate a context-sensitive phrasing of timbrally composite figures.

Background information

In *Kontakte* the arsenal of timbres in the instrumental domain is enlarged through electronic sounds; in *Time and Motion Study II* the timbral possibilities are extended through new techniques of sound production on the cello. In the literature, the formal construction through timbral distribution has not been adressed so far. Indeed, for *Kontakte*, most authors (among them Kirchmeyer 1963, Heikinheimo 1972, Mowitz 2002 and Mowitz 2010) have discussed the composition on the basis of Stockhausen's 1961 article on "Momentform" (Stockhausen 1963, pp. 189-210), even if this concept was not yet formulated when *Kontakte* was undertaken. For *Time and Motion Study II*, Fitch has transcribed an important sketch from Ferneyhough's preparatory materials (Fitch 2013, p. 210) without deepening its incidences on the proper compositional level, while Iddon considers mainly the role and consequences of the live-electronic set-up (Iddon 2006). A very welcome exception is an article by Dack (1998) who analyses the beginning of *Kontakte* (precisely the two "non serial" sections) with the help of Schaeffer's tools of typomorphological analysis (Schaeffer 1966).

Methodology

The present research combines two different methodologies with the aim to relate both ends of the musical communication chain, i.e., sketch studies for the poietic aspect (according to Molino/Nattiez in Nattiez 1987) and Schaeffer's "reduced listening" ("écoute réduite") to describe qualitatively the sonic result from an aesthesic perspective (Schaeffer 1966).

Results

The close sketch study for both pieces allows to clarify the compositional building elements as well as the means of phraseological and formal articulation. Throughout the twentieth century, from Varèse's structural use of percussion instruments to Stockhausen's attemps to build the final section of *Gesang der Jünglinge* primarily by way of timbre repartitioning, the status of timbre within the composer's counsciousness changed in a decisive way. Whereas in an instrumental score like *Gruppen*, specific instrumental families helped to colour the respective sections of a serially derived time-structure, in *Kontakte*, the overall-plan (which remained partly unrealized) highlights in a countinuous timbre-transformation from low metallic noises to a mix of high wood and skin noises. The different steps in this process consist in either a concentration on one "fixed" timbral family (now defined as a field of resulting sounds disregarding the sound source), or the illustration of a strategic timbre-change (possible only with the electronics). This choice of the timbral family is furthermore combined with specific morphological types (point, group, tremolo). In section III all the serial determinations are realized by the sole instruments (and the added electronic sounds need to be considered in relation to this structural background) while in section IV the combination with the tape shows different possibilities of interaction and of exchange. The extended techniques used in *Time and Motion Study II* allow to enrich the sonic possibilities of the cello by including also explicitly "noisy" sound results. This is what the interaction of the two formal levels ("structures" *vs* "sequences") is about at the beginning of the score. Indeed, the first "Sequence" that interrupts the initial "Structure" is separated from the context by spectromorphological oppositions (short fixed tonic sounds against longer mobile and complex glissando structures). Nevertheless, Ferneyhough transgresses parts of his pre-compositional determinations to add subcutane continuities depending from the actual sonic environment (one of the groups of short tonic sounds is prolonged by a sustain resorting precisely the tremolos and glissandos from the previous "structure").

The description of the chosen sections with the support of Schaeffer's tool of typomorphologie (Schaeffer 1966, readapted by Thoresen within the more general context of spectromorphological analysis, Thoresen 2015) reveals to which amount the pre-compositional decisions sketched by both composers anticipate the perception of the musical results and which supplementary rules (present or not in the composers' sketches) are necessary to complete an aural understanding of these pieces.

Conclusions

The combination of sketch studies and aural description of the sonic result through reduced listening allows to identify a limited number of sonic categories as well as of phraseological and formal connection rules that organize these compositions. These rules sum up to a kind of "solfege" particularly adequate to this kind of sound-oriented music. Such a "solfege" is useful both for the aural understanding and for the conception of an informed performance practice of these compositions.

References

Dack, J. (1998). Strategies in the analysis of Karlheinz Stockhausen's Kontakte für elektronische Klänge, Klavier und Schlagzeug. *Journal of New Music Research*, *27* (1-2), 84-119.

Ferneyhough, B. (1995). *Collected writings*. J. Boros and R. Toop (Eds.). Amsterdam: Harwood Academic Publishers.

Fitch, L. (2013). *Brian Ferneyhough*. Bristol: Intellect.

Heikinheimo, S. (1972). *The electronic music of Karlheinz Stockhausen. Studies on the esthetical and formal problems of its first phase*. Helsinki: Acta Musicologica Fennica 6.

Iddon, M. (2006). On the entropy circuit: Brian Ferneyhough's Time and Motion Study II. *Contemporary Music Review*, *25* (1-2), 93-105.

Kirchmeyer, H (1963). *Zur Entstehungs- und Problemgeschichte der 'Kontakte' von Karlheinz Stockhausen*. Grosskönigsdorf: Wergo [record Wergo 60009].

Mowitz, M. (2002). *Die Form der Unendlichkeit. Aspekte der Momentform und der seriellen Struktur in Karlheinz Stockhausens Kontakte*. Essen: Die Blaue Eule.

Mowitz, M. (2010). *Die Form der Klänge. Stockhausens Konzeption einer einheitlichen musikalischen Zeit am Beispiel der Komposition "Kontakte"*. Universität Osnabrück: Electronic Publishing Osnabrück.

Nattiez, J.-J. (1987). *Musicologie générale et sémiologie*. Paris: Bourgois.

Schaeffer, P. (1966). *Traité des objets musicaux*. Paris: Seuil.

Stockhausen, K (1963). *Texte zur elektronischen und instrumentalen Musik*. Cologne: DuMont.

Thoresen, L. (2015). *Emergent musical forms: Aural explorations*. London (CA): University of Western Ontario.

Timbral Syntax in Björk's *Utopia*

David Blake[1†], Alessandro Bratus[1]
[1] Crane School of Music, SUNY-Potsdam, Potsdam, NY, USA
[2] Department of Musicology and Cultural Heritage, University of Pavia, Pavia, Italy
[†] blakdak@potsdam.edu

Aims/goals

This paper proposes an analytical approach to timbral syntax in popular music. We argue for the importance of contrasting timbral groupings for both formal and extramusical meaning. It demonstrates this theory through analyzing Björk's 2017 album *Utopia*.

Background information

Popular music theory has largely adopted the nomenclature of primary (e.g. pitch, rhythm, harmony) and secondary (e.g. timbre, texture, volume, dynamics) analytical parameters from Western art music theory. This terminology stems from the belief that primary parameters are syntactical—that is, they constitute a song's identity, phrasal grammar, and formal structure. By contrast, secondary parameters are perceived to be coloristic—important for perception, but not structurally integral. For over two decades, popular music scholars have theorized how musicians and record producers selectively choose and manipulate timbre for meaningful purposes, but the residues of primary/secondary divide have tended to deflect timbral analysis away from form toward either individual moments or general characteristics. We argue that related timbres are purposefully arranged in popular music, and that these arrangements are essential for a work's structure and meaning.

Methodology

Our theory of timbral syntax is predicated on two fundamental aspects of the parameter's cognitive perception: multidimensionality, the ability to transmit sonic information along innumerable material and phenomenal axes; and perceptual fusion, the tendency to comprehend soundworlds by synthesizing sounds with shared characteristics. We argue that together these aspects have enabled popular musicians to use a wide palette of what we will call "contrasting timbral groupings" to delineate musical syntax at multiple levels, from phrases and sections to songs and albums.

This paper focuses on three contrasting timbral groupings common in popular music that are found on Björk's *Utopia*. The first, clean and noisy, focuses on the clarity of source and pitch identification. The second, acoustic and electronic, distinguishes between the complex, aperiodic waveforms of acoustic sources and the regularized, periodized waveforms produced via electronic manipulation. The third, centered and decentered, indexes the spatial distribution of a sound as a result of record production. We contend that the production, repetition, and transformation of vocal and instrumental timbres along these three axes delineate song form, contribute to hermeneutic interpretation, and shape a specific ecology of listening experience. To account for both the formal and extramusical implications of timbral syntax, we utilize a holistic analytical approach that blends traditional analysis of form, pitch, and rhythm; spectrography; audio file analysis; and consideration of album visuals and Björk's stage personae.

Results

Björk's *Utopia* offer rich examples for analyzing how timbral relationships delineate form and create sonic environments through record production. The qualitative dimensions of sound defined by the oppositions clean/noisy and acoustic/electronic are constantly reinforced by the interaction of three elements: Björk's voice, produced both without overt production and with heavy distortion; shakuhachi sounds which both focus on pitch clarity and the noisiness of breath and material grain; and the synthesized keyboards and drum machines, which have varying degrees of tangible material reference. Each instrument sound pairs natural modes of sound production with artificial timbral components, enabling the complex interplay of utopia and dystopia arise in the relationship presented between the human and the technological in the album's artistic production.

Along with considering the relationships between instruments, we also focus specifically on the manipulation of Björk's recorded voice. The use of different production and distortion techniques, the frequent deployment of polyphonic voice tracking, and the constant displacement of sung and spoken fragments within the stereophonic space all fragment the sung text, leading to multiple potential layers of signification. On the whole, the syntactical exploration of the timbral dimension seems a way to reclaim a dominant position for a feminine recorded persona, a position amplified by Björk's own visual and performative persona. The distinct productions of her voice articulate groove-based repetitive forms that align her sounds with disco-oriented dance rather than the verse/chorus form of rock. Her recorded tracks overwhelm the listener with aural stimuli that combine to detach the listener from the natural world, reflecting the utopia of the song title and the sanctuaries of disco clubs.

Conclusions

Our project has three main conclusions. First, we contend that timbre has an integral syntactical function in popular music, suggesting that timbral analysis needs greater attention to relationships between and transformations of timbre. Second, we argue that considering timbre's syntactical function can open up new models for connecting formal meaning and extramusical meaning in recent recorded popular music repertoires. Finally, we conclude by suggesting that the possibility of a syntactical theory of timbre questions the appropriateness of any sort of parametrical hierarchization in popular music theory, suggesting a more holistic and dehierarchized analytic approach to popular music analysis.

References

Blake, D. K. (2012). "Timbre as Differentiation in Indie Music." *Music Theory Online* 18, no. 2. http://www.mtosmt.org/issues/mto.12.18.2/mto.12.18.2.blake.html.

Cogan, R. (1984). *New Images of Musical Sound*. Cambridge, MA: Harvard University Press.

Heidemann, K. (2016). "A System for Describing Vocal Timbre in Popular Song." *Music Theory Online* 22(1). http://www.mtosmt.org/issues/mto.16.22.1/mto.16.22.1.heidemann.html.

Leydon, R. (2012). "Clean as a Whistle: Timbral Trajectories and the Modern Musical Sublime." *Music Theory Online* 18(2). http://www.mtosmt.org/issues/mto.12.18.2/mto.12.18.2.leydon.html.

Meintjes, L. (2003). *Sound of Africa! Making Music Zulu in a South African Studio*. Durham, NC: Duke University Press.

Mirka, D. (2001). "To Cut the Gordian Knot: The Timbre System of Krzyzstof Penderecki." *Journal of Music Theory* 45(2): 435–56.

Moore, A. F. (2012). *Song Means: Analysing and Interpreting Recorded Popular Song*. Burlington, VT: Ashgate.

Osborn, B. (2016). *Everything in its Right Place: Analyzing Radiohead*. New York: Oxford University Press.

Walser, R. (1993). *Running with the Devil: Power, Gender, and Madness in Heavy Metal Music*. Middletown, CT: Wesleyan University Press.

Zagorski-Thomas, S. (2014). *The Musicology of Record Production*. New York: Cambridge University Press.

The Closer the Better: The Role of Formant Positions in Timbre Similarity Perception and Timbre Classification

Christoph Reuter[1†], Isabella Czedik-Eysenberg[1], Saleh Siddiq[1] and Michael Oehler[2]
[1] Musicological Department, University of Vienna, Vienna, Austria
[2] Musicological Department, University of Osnabrück, Osnabrück, Germany
[†] christoph.reuter@univie.ac.at

Aims/goals

The aim of this study is to test the advantage and suitability of the formant paradigm for musical timbre similarity prediction as well as timbre classification by investigating two questions:
Do shorter distances between the formant positions of two sounds correlate with the impression of greater timbre similarity (and vice versa)?
Is the calculated vector of formant positions an adequate means for a precise computational classification of the involved instruments?

Background information

There are only a few approaches to a comprehensive description of musical instrument timbres. Beside MFCCs (Loughran, Walker, O'Neill, & O'Farrell 2008) and MPS (Elliott, Hamilton, & Theunissen 2013; Thoret, Depalle, & McAdams 2017), the formant paradigm has proven successful so far especially for wind instruments: Formants turned out to be a useful predictor for timbre blending and separation as well as for timbre similarity (Reuter 1996; Lembke 2014; Lembke & McAdams 2015). Plotting the formants of 586 wind instrument sounds of all reachable pitches in *ff* and *pp* into a X/Y scatter diagram (X-axis: first formant, Y-axis: second formant), it is possible to distinguish musical instruments and their registers by their formant positions in this depiction (Reuter, Czedik-Eysenberg, Siddiq, & Oehler 2017).

Methodology

In a listening test, 22 participants rated the (dis)similarity of 40 loudness-adjusted timbre combinations including wind instruments like flute, oboe, clarinet, bassoon, trumpet, trombone, French horn, tuba on a scale between 1 and 8 (8= maximum dissimilarity). Half of the stimuli contained sounds with extremely close/overlapping formant regions while the other half contained sounds with very distant formant regions. Instrumental sounds were taken from the Vienna Symphonic Library (VSL) and adjusted to a matching ANSI-loudness level via Matlab and Genesis loudness toolbox.
With the help of support vector machines (SVM), k-nearest-neighbor (KNN), and further methods, we tested the precision of the instruments' classification by means of their formants using five-fold cross validation. Furthermore, we examined a set of 89 additional timbre descriptors to improve the classification.

Results

Between the perceived timbre (dis)similarity and the Euclidean distance of the formant positions a remarkable correlation ($r=0.759$; $p<0.001$) could be shown: Timbres with very close formant positions were perceived as very similar while timbres with very distant formant positions were deemed very dissimilar.

When matching the formant positions of the examined timbres with their respective musical instruments we achieved a precision of 46.1% (cubic KNN). By adding further timbre descriptors, we could increase the classification precision to 84.6% (quadratic SVM). The resulting confusion matrix corresponded very well with human mismatching of musical instrument timbres: e.g., flute, clarinet and oboe were confused relatively often as well as trumpet and trombone or bassoon and tuba sounds.

Conclusions

The perception of timbre similarity was shown to be closely connected with the sounds' formant positions. Regarding computational musical instrument classification, formant positions alone do not seem to be a sufficient criterion compared to other more powerful feature combinations, but regarding the evaluation of perceived timbre similarity, the distance between formant positions can serve as a helpful explanatory tool.

Acknowledgments

This project is supported by funds of the Österreichische Nationalbank (OeNB, Anniversary Fund, project number: 16473). Special thanks to the Vienna Symphonic Library for supporting our work with two free "Vienna Super Package" licenses.

References

Elliott, T.M., Hamilton, L.S., & Theunissen, F.E. (2013). Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones. *Journal of the Acoustical Society of America*, 133(1), 389-404.

Lembke, S.-A. (2014). *When timbre blends musically: Perception and acoustics underlying orchestration and performance*. (Doctoral dissertation). Retrieved from http://digitool.library.mcgill.ca/webclient/DeliveryManager?pid=130507

Lembke, S.-A. & McAdams, S. (2015). The role of spectral-envelope characteristics in perceptual blending of wind-instrument sounds. *Acta Acustica united with Acustica* 101, 1039-1051.

Loughran, R., Walker, J., O'Neill, M., & O'Farrell, M. (2008). The use of mel-frequency cepstral coefficients in musical instrument identification. Proceedings of the 2008 International Computer Music Conference, ICMC 2008, Belfast, Ireland. https://quod.lib.umich.edu/i/icmc/bbp2372.2008.083/1

Reuter, C. (1996). *Die auditive Diskrimination von Orchesterinstrumenten*. Frankfurt: Lang.

Reuter, C., Czedik-Eysenberg, I., Siddiq, S., & Oehler, M. (2017). Formanten als hilfreiche Timbre-Deskriptoren. Deutsche Gesellschaft für Akustik (Ed.), *Proceedings of the 43. DAGA, Fortschritte der Akustik* (pp. 190-193).

Thoret E., Depalle P., & McAdams S. (2017). Perceptually Salient Regions of the Modulation Power Spectrum for Musical Instrument Identification. *Frontiers in Psychology,* 8:587. doi: 10.3389/fpsyg.2017.00587

Timbral Function in *Klangflächetechnik*

Matthew Zeller
Department of Music, Duke University, Durham, North Carolina, U.S.A.
matthew.zeller@duke.edu

Abstract

The *Klangfläche*, or sound sheet, is a potent and rich compositional device that facilitates new ways of thinking about and analyzing musical timbre. Rarely discussed in English-language scholarship, it is traditionally considered from a harmonic standpoint. Monika Lichtenfeld, who coined the term in her 1970 essay, "Zur technik der Klangflächenkomposition bei Wagner," as well as Carl Dahlhaus, argue for a negatively framed understanding of the technique based primarily on non-chord tones and pitches extrinsic to the underlying harmony. Timbral function, however, can be decisive for interpreting this music and can force us to reconsider our harmonic assumptions. Building on Lichtenfeld and Dahlhaus's formulations of *Klangflächen*, I reformulate the technique through the lens of timbre, and show that timbre often challenges harmonic considerations.

Klangfläche technique evolved in two ways during the late nineteenth and early twentieth centuries, both with timbre at their core. First, *vis a vis* common practice tonality, I consider the music of Gustav Mahler and Richard Strauss. The *Klangflächen* of Mahler and Strauss illustrate that compositional parameters such as timbre, registral placement, and tempo, may take precedence over the traditional parameters of pitch and harmony in determining what is part of the underlying harmonic progression or teleological progress of the work. Rather than frame the extrinsic sonorities negatively as exempted from organic development, I formulate them positively as independent, non-teleological entities within a framework of the teleologically composed world. Second, I show that Arnold Schoenberg's seminal *Fünf Orchesterstücke* Op. 16 No. 3, "Farben," can be interpreted as a *Klangfläche*. Through this reading, pitch classes, musical gestures, and timbral events in "Farben" can be explained as harmonic/static/teleological and non-harmonic/active/non-teleological musical events. Schoenberg's use of *Klangfläche* technique became a foundational moment in post-tonal twentieth century music, providing the impetus for numerous works and compositional trajectories.

Understanding the Perception of Acoustic Space from the Perspectives of Timbre

Song Hui Chon† and Sungyoung Kim
Department of Electrical, Computer, and Telecommunications Engineering Technology, Rochester Institute of Technology, Rochester, NY, U.S.A.
† songhui.chon@gmail.com

Aims/goals

In the current study we seek to better understand the perception of acoustic spaces by musicians and listeners from the perspectives of timbre.

Background information

Sounds cannot be heard without a space, therefore an acoustic space may be considered as a hyper instrument that could affect the perception of music. In this paper, we apply findings from timbre perception research to better understand the perception of acoustic space by performers and listeners.

Methodology

Analyses were performed on 27 short binaural recordings of nine string quartet groups performing one excerpt in three acoustic conditions, as well as the data from perceptual experiments reported in an earlier study (Chon et al., 2015). The three acoustic conditions include one natural and two virtually enhanced rooms with similar but different spatial profiles. Performers completed the questionnaire after performing in each of the three conditions (randomized in order). Listeners answered a different questionnaire to evaluate the spatial characters of the acoustic space in the recordings. Timbre descriptors were computed using the Timbre Toolbox (Peeters et al., 2011) and correlated to find the best correlates of each question.

Results

Table 1. Timbre correlates to nine questions asked to performers after performing in each acoustic space. The answers were marked on a 5-point Likert scale. Two questions, quality of reverberation and enjoyment of playing, were not correlated with timbre descriptors due to lack of relevance. * indicates significant results at alpha of 0.05; ** at 0.01; and *** at 0.001.

PC	Question	r	*p*	Descriptor Group	Descriptor
1	Amount of reverberation	.36	.09	ERB FFT	Frame Energy
	Envelopment	.27	.17	ERB FFT	Frame Energy
	Height Sensation	.73	< .05*	ERB FFT	Spectral Decrease
2	Ease of hearing own instrument	−.63	< .001***	ERB Gammatone	Frame Energy
	Ease of hearing others	.39	< .05*	ERB Gammatone	Skew
	Ease of maintaining tempo	.44	< .05*	ERB Gammatone	Frame Energy
	Ease of hearing dynamics	−.52	< .01**	ERB Gammatone	Frame Energy
3	Clarity	−.60	< .001***	Harmonic	Spectral Centroid
	Tonal Balance	−.27	.10	Harmonic	Spectral Centroid

Three principal components (PCs) were previously found from performers' responses. Table 1 summarizes the nine questions in their three PCs, along with their timbre correlates. The first PC, *Spatial Attributes*, and the second PC, *Stage Support*, show highest correlations with auditory-model descriptors. It is notable, although perhaps not surprising, that the frame energy is the correlate for more than half of the questions in the first two PCs. This may suggest that performers are most sensitive to the acoustic energy that *they hear* in the performance space. The timbre correlate of the third PC, *Tonal Quality*, is spectral centroid in the harmonic domain. In sum, performers appeared to be most sensitive to the energy level of sounds they hear and the harmonic center of gravity of the acoustic space where they perform.

Table 2. Timbre correlates to 12 questions asked to listeners, who would choose one excerpt out of three that would best answer each question. * indicates a significant correlation at alpha of 0.05. + indicates a strong trend that approaches 0.05.

Question		ρ	Descriptor Group	Descriptor
Naturalness	Most	−.29	Harmonic	Spectral Skew
	Least	−.17	Harmonic	Spectral Decrease
Source Distance	Farthest	−.38+	ERB FFT	Spectral Decrease
	Closest	.27	Harmonic	Kurtosis
Room Size	Largest	.20	ERB FFT	Spectral Decrease
	Smallest	−.30	ERB FFT	Spectral Decrease
Clarity	Most	−.21	Harmonic	Spectral Decrease
	Least	−.29	Harmonic	Spectral Decrease
Loudness	Most	−.25	Harmonic	Spectral Decrease
	Least	.40*	Harmonic	Odd-Even Harmonic Ratio
Preference	Most	.25	ERB FFT	Spectral Decrease
	Least	−.28	ERB FFT	Spectral Decrease

Table 2 lists the timbre correlates to 12 questions to listeners. Spearman's rank order correlations were calculated, because the listeners' response data were probability data (of each acoustic condition to be chosen for each question). The degrees of correlations are mild at best, and only one shows a significant correlation at alpha of 0.05. The meaning of this significant correlation is difficult to understand, but perhaps it simply means that the detailed spectral pattern in the harmonic domain affected the loudness rankings of acoustic conditions. Note that both the harmonic and the ERB FFT domains are also important in listeners' judgments. But the mild correlations in this table may imply a different experience—that listeners evaluated an indirect experience of the space from binaural recordings whereas performers evaluated a first-hand experience of performing in and responding to the acoustic spaces.

Conclusions

We presented an analysis of evaluations of acoustic spaces by performers and listeners using the Timbre Toolbox. This current analysis presented a scientific way to understand the spatial perception from a timbre perspective. It turns out that the timbre descriptors based on the auditory model and harmonic analysis seem more relevant to explain the responses from the performers as well as the listeners (although not as significant). The descriptors in these two groups, especially those listed in the tables above, may be useful for future assessment of acoustic spaces for music performance.

References

Chon, S.H., Ko, D., & Kim, S. (2015). Listeners' responses to string quartet performances recorded in virtual acoustics. *In the Proceedings of AES 59th Conference, Montreal, Canada.*

Peeters, G., Susini, P., Misdariis, N., Giordano, B. L., & McAdams, S. (2011). The timbre toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, 130(5), 2902–2916.

Score and Sound-Based Interface for Exploring Music Performances

Simon Waloschek†, Aristotelis Hadjakos, Axel Berndt
Center of Music and Film Informatics, Detmold University of Music, Germany
† s.waloschek@cemfi.de

Aims/goals

Musical interpretations and timbre can be examined using sound analysis. Such interfaces provide possibilities for spectrographic analysis and manual annotation. However, in order to present the results of the musical analysis gained by such an approach, it is often necessary to refer to the musical score. Currently used interfaces could benefit from including synchronized scores, ideally the scores used in the recording with hand-written annotations, markings and scribbles by the performer.

Background information

Current sound analysis applications do not have support for music scores. Sonic visualiser (Cannam 2010) is one of the most advanced tools of this kind. It offers various possibilities for spectral analysis and plugins can provide additional audio analysis methods. Annotation layers can be used to mark time instances, to show images or annotate text. But the musical score, which is often the basis for the performers' interpretation, remains absent. On the other side, audio-to-score alignment applications, such as Prätzlich et al.'s (2015) FreiDi:syncPlayer, have not been taken into account as visual analytics tools to interface music performances and visualize the relation of score annotations and recorded audio. With our tool MuPEx (Music Performance Explorer) we present a first attempt to introducing scores to such analysis contexts.

Methodology

The synchronization of scores with audio recordings requires a common data basis for both representations. Data formats such as MIDI and MusicXML are often utilized as an intermediate link between these two modalities but lack the ability of representing actual images and image annotations. We therefore opted for MEI, a file format primarily used to describe musical notation and the underlying graphical layout. The actual alignment of audio and MEI was done using dynamic time warping with chroma features (Müller 2007). In order to link back to the scans of original scores used by the performers, we developed Vertaktoid, an Android app that lets the user annotate measures in a score (Mexin et al. 2017). This allows us to link from an aligned audio via the MEI measure ID to pixel coordinates in the original scores.

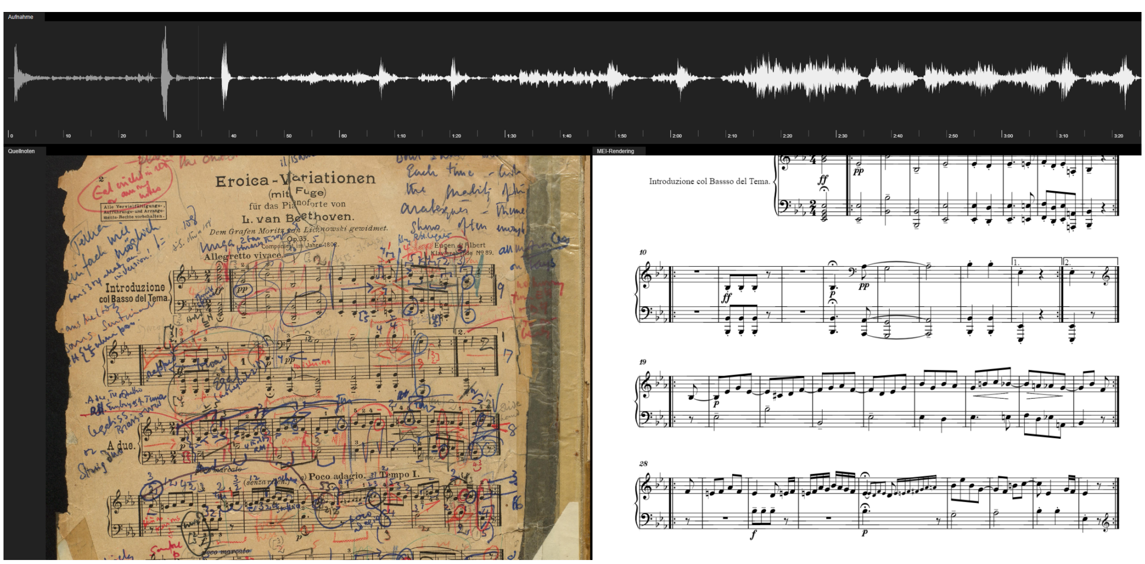


Fig. 1: The MuPEx interface with three different sources (audio, scan and score rendering).

Results

We developed a web-based interface (see fig. 1) that combines music recordings with their corresponding score. By calculating the alignment between audio files, scans of the scores used during production and clean renderings of the scores, the interface allows for transmedial navigation and annotation. Clicking in either one of the three views moves the playback cursor in all other views to the corresponding positions. Annotations can therefore be made in the view that has the strongest relation to its content, e.g., timbre-related annotations can be made directly in the waveform while the (written) reason for the specific change in timbre is highlighted in the scan.

Conclusions

The development of MuPEx was motivated by musicologists in the field of performance research who were struggling for a long time with a media break between scores, audio recordings and audio analysis visualizations. Audio-to-score alignment marks the first step towards a more comprehensive conflation of various music representation modalities, each providing a different perspective to the music. Beyond score and waveform visualization we consider several future additions, including frequency and phase spectrograms, video integration, and plottings of results from third-party analysis tools such as Sonic Visualizer.

Potential application scenarios of MuPEx are not limited to performance research where it enables a visual analytics-based methodology that is new to this field. We see further interesting applications also in music teaching and as part of digital music editions.

References

Cannam, C., Landone, C., & Sandler, M. (2010). Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the 18th ACM international conference on Multimedia* (pp. 1467-1468). ACM.

T. Prätzlich, M. Müller, B. W. Bohl, and J. Veit (2015). Freischütz Digital: Demos of audio-related contributions. In *Demos and Late Breaking News of the International Society for Music Information Retrieval Conference (ISMIR)*.

Roland, P. (2002). The Music Encoding Initiative (MEI). In *Proceedings of the International Conference on Musical Applications Using XML* (pp. 55–59).

ViFE (Virtueller Forschungsverbund Edirom) (2016). Edirom Editor 1.1.26. https://github.com/Edirom/Edirom-Editor, accessed: 2017-20-02.

Y. Mexin, A. Hadjakos, A. Berndt, S. Waloschek, A. Wawilow, G. Szwillus (2017). Tools for Annotating Musical Measures in Digital Music Editions. In *14th Sound and Music Computing Conf. (SMC-17).*

Müller, M. (2007). *Information retrieval for music and motion* (Vol. 2). Heidelberg: Springer.

Computational Modelling of Timbre Dimensions for Automatic Violin Tone Quality Assessment

Sergio Giraldo[1†], Rafael Ramirez[1], George Waddell[2] and Aaron Williamon[2]
[1] Music Technology Group, Pompeu Fabra University, Barcelona, Spain
[2] Centre of Performance Science, Royal College of Music, London, United Kingdom
[†] sergio.giraldo@upf.edu

Aims/goals

Automatic assessment of music performance is an open research area widely studied in the past. A vast number of systems aiming to enhance the learning process of a musical instrument are being developed in recent years. However, most of the systems focus on the assessment of pitch and onset accuracy, and very few pay attention to tone quality. This is particularly true in violin music education, where although a consensus exists on what is a good or a bad tone quality, there is not a formal definition due to its subjectivity. We present a machine learning approach for the automatic assessment of violin tone quality. We depart from our previous work on the preliminary modelling of several dimensions involving tone quality. Based on recorded examples of tones with different qualities defined and recorded by a professional violinist, we applied machine learning techniques to learn computational models able to evaluate tone quality from extracted audio features. The tone quality models were implemented into a real-time-visual-feedback system.

Background information

The quality of a performed sound is assumed to be a contribution of several parameters of sound such as pitch, loudness, and timbre. Eerola et al. (2012) identify 26 acoustic parameters of timbre among several instrument groups, which might reflect instrument performance techniques and expressive intentions. Automatic characterization of dynamics and articulation from low-level audio features has been studied by Maestre & Gómez (2005) in the context of expressive music performance. Knight et al. (2011) study the automatic assessment of tone quality in trumpet sounds using machine learning techniques. Romani Picas et al. (2015) make use of machine learning techniques to identify good- and poor-quality notes of the trumpet, clarinet and flute, given training data consisting of low and high level audio features extracted from performed musical sounds with each instrument. Giraldo et al. (2017a, 2017b) conducted perceptual tests on the quality of performed musical notes to study the correlations among previously defined terms used in the literature (i.e., pitch-dynamic-timbre stability, richness, and attack clarity) and a list of tone qualities provided by music experts in violin education.

Tone Qualities	
Dark	Bright
Cold	Warm
Harsh	Sweet
Dry	Resonant
Light	Heavy
Grainy	Pure
Coarse	Smooth
Closed	Open
Restricted	Free
Narrow	Broad

Table 1. Proposed list of tone quality by music experts.

Methodology

We obtained recorded examples of each of the tone qualities listed in Table 1, by a professional violinist. Eight notes were recorded for each tonal label using different combinations of fingerings, strings, and bow directions (i.e. up and down), for a total of 80 notes. Low- and high-level descriptors were extracted from the audio signal using the Essentia library (Bogdanov et al. 2013). Automatic feature selection methods (i.e., wrapper with genetic search) were used to obtain a subset of low-level frame based, spectral descriptors. After feature selection, machine learning techniques were applied to obtain a model for each tone quality pair in Table 1 (e.g., bright-dark, cold-warm), thus obtaining a total of 10 models. For each model, three machine learning schemes were compared, i.e., Linear Regression (LR), Artificial Neural Networks (ANN) with one hidden layer, and Support Vector Machines (SVM) with linear kernel.

Results

The accuracy of each model was calculated based on the percentage of correctly classified instances. The average accuracy of each of the methods compared, across the 10 models of tonal quality pairs was 79% (5%std), 89% (5%std), 85% (5%std) for LR, ANN, and SVM, respectively. The ANN models were then chosen to be implemented on a real-time system for tone quality assessment, where the probability of each binary class was mapped into a visual feedback system, using a spider chart scheme.

Conclusions

A computational approach to automatically assess the quality of performed violin sounds is proposed. We recorded examples of expert defined tone qualities. High and low-level descriptors were extracted from the audio signal and machine learning models were obtained to predict the different quality dimensions. The accuracy obtained indicate that the selected features contain sufficient information to predict the studied tone qualities.

Acknowledgments

This work has been partly sponsored by the Spanish TIN project TIMUL (TIN2013-48152-C2-2-R), the European Union Horizon 2020 research and innovation programme under grant agreement No. 688269 (TELMI project), and the Spanish Ministry of Economy and Competitiveness under the Maria de Maeztu Units of Excellence Programme (MDM-2015-0502).

References

Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., Roma, G., Salamon, J., Zapata, J. R., Serra, X. et al. (2013), Essentia: An audio analysis library for music information retrieval., in 'ISMIR', pp. 493–498.

Eerola, T., Ferrer, R. & Alluri, V. (2012), 'Timbre and affect dimensions: evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds', *Music Perception: An Interdisciplinary Journal,* 30(1), 49–70.

Knight, T., Upham, F. & Fujinaga, I. (2011), The potential for automatic assessment of trumpet tone quality., in 'ISMIR', pp. 573–578.

Maestre, E. & Gómez, E. (2005), Automatic characterization of dynamics and articulation of expressive monophonic recordings, in 'In *Proceedings of the 118th Audio Engineering Society Convention*, Citeseer.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N. & McAdams, S. (2011), 'The timbre toolbox: Extracting audio descriptors from musical signals', The Journal of the Acoustical Society of America 130(5), 2902–2916.

Romani Picas, O., Parra Rodriguez, H., Dabiri, D., Tokuda, H., Hariya, W., Oishi, K. & Serra, X. (2015), A real-time system for measuring sound goodness in instrumental sounds, in 'Audio Engineering Society Convention 138', Audio Engineering Society.

Giraldo, S., Ramírez, R., Waddell, G. and Williamon, A. (2017a). "A Computational Approach for Measuring Performance Quality in Violin Tones ." In *International Symposium in Performance Science*, 29 August 2017, Reikiavik, Iceland.

Giraldo, S., Ramírez, R., Waddell, G. and Williamon, A. (2017b). "A Real-time Feedback Learning Tool to Visualize Sound Quality in Violin Performances." In *Proceedings of the 10th International Workshop on Machine Learning and Music (MML 2017)*, 6 October 2015, Barcelona, Spain.

Towards Translating Perceptual Distance from Hearing to Vision Using Neural Networks

Etienne Richan[†], Jean Rouat
NECOTIS, Département de génie électrique et génie informatique, Université de Sherbrooke, Sherbrooke, QC, Canada
[†] etienne.richan@usherbrooke.ca

Aims/goals

The goal of this research project is to develop a system mapping sounds to images in a manner that respects perceptual distances in both modalities. In other words, the degree of perceived difference between two sounds should be reflected in the dissimilarity of the images generated from those sounds. This is not a trivial problem, and there does not seem to be much work of reference on the translation of perceptual distances between modalities. In our approach, timbral features are used to measure perceived auditory distance and a neural network for style transfer is used to synthesize textural images. Our software project allows users to select sound-image pairings that are meaningful to them, which can then be extrapolated to other sounds. The generated images aim to assist the task of distinguishing between different musical sound samples without needing to hear them.

Background information

It has been shown that features such as spectral centroid and log attack time are highly correlated with the perceptual classification of musical timbre (McAdams et al., 1995). It is of interest to study what visual metaphors would be most effective for representing these auditory dimensions. Research on audio-visual correlations has generally focused on relatively simple auditory and visual features such as loudness, pitch, position, shape and colour (Walker, 1987). Recent studies have explored the more complex correlation between timbral features and 2D and 3D shapes (Adeli et al., 2014; Soraghan, 2014). Other studies have investigated the correspondences between timbral and textural properties such as coarseness, granularity and regularity (Giannakis, 2006). While these studies have discovered some strong correlations between auditory features and visual attributes, they are limited in the manner that they map a single auditory parameter to a single visual parameter.

Neural networks provide a general framework for extracting features from complex distributions in a hierarchical manner. The hierarchical representation of visual features by neural image classification networks is comparable to that of the visual cortex (Bogdanov et al., 2013) (at least in the initial layers). This allows us to suppose that features extracted by these networks can be used as a measure for visual perceptual distance.

Style transfer networks are a recent offshoot of research on neural networks for image classification. The covariance matrices between feature maps at different layers of these deep convolutional networks are representative of the stylistic (or textural) structure of images (Gatys et al., 2016). Initial solutions allowed for the transfer of artistic style from one image to another by iterative optimization of an input image through backpropagation (Gatys et al., 2016). More recent approaches (Li et al., 2017) have greatly sped up the process by directly altering the feature maps of an input image to match those of the desired style using a signal whitening and colouring approach. Stylized images can be reconstructed from this latent representation using a pretrained decoder network. This algorithm provides a novel approach to controlling image space; instead of directly manipulating coordinates, colour schemes or geometry, we can use the network to synthesize and interpolate between different textures.

Methodology

The research project is divided into three phases. First, finding a suitable space for audio samples based on timbral features where proximity is representative of their perceptual similarity. The Essentia library (Bogdanov et al., 2013) is used to extract a set of low-level timbral features from monophonic musical samples. The nSynth dataset (Engel et al., 2017) provides testing and training samples. It contains ~300000 four-second notes from ~1000 acoustic, electric and synthetic instruments, played across their respective ranges. The dataset also provides semantic quality annotations for each note (e.g., bright, dark, distorted).

We train a shallow neural network to learn a mapping from timbral features to a space where notes with common qualities are found closer together. We focus on 2D and 3D spaces, as these are more appropriate for user interaction and evaluation. The learned spaces are compared to those found by dimensionality reduction techniques. In the second phase, we investigate the expressive potential of style transfer networks by testing a variety of input patterns, methods for controlling the flow of information through the network and interpolating between latent representations. Finally, the first two phases are integrated into a software project allowing for the definition, visualization and extrapolation of audiovisual pairings. By selecting a subset of samples in timbral space to be paired with textural images, images can be generated for all other samples based on their proximity to the reference samples by interpolating between latent representations in the style transfer network.

Results

Interpolating between textures allows us to generate images for a large set of sounds based on a much smaller set of sound-image pairings. For good results, this approach requires that pairs be well distributed throughout the sample space. The poster will mainly focus on results from the first two phases of research: the comparison of timbral spaces and controlling texture synthesis.

Conclusions

Neural networks allow us to approach image synthesis in a radically different way than we are used to. They also provide a mechanism for learning highly nonlinear transformations across modalities. The difficult task of comparing distances within different modalities can be assisted by the embedding learned by a neural network. Our method for defining and generating audio-visual pairings opens new avenues for research as well as artistic endeavors. This tool can assist in the work of scholars and artists who must navigate large libraries of samples. We also believe it can help extend the study of audiovisual correlations and the search for effective visualizations of timbral features.

References

Adeli, M., Rouat, J., & Molotchnikoff, S. (2014). Audiovisual correspondence between musical timbre and visual shapes. *Frontiers in Human Neuroscience*, *8*.

Bogdanov, D., Serra, X., Wack, N., Gomez, E., Gulati, S., Herrera, P., … Zapata, J. (2013). ESSENTIA: an open-source library for sound and music analysis (pp. 855–858). ACM Press.

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., & Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Scientific Reports*, *6*.

Engel, J., Resnick, C., Roberts, A., Dieleman, S., Eck, D., Simonyan, K., & Norouzi, M. (2017). Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders. *arXiv:1704.01279 [Cs]*.

Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on* (pp. 2414–2423). IEEE.

Giannakis, K. (2006). A comparative evaluation of auditory-visual mappings for sound visualisation; Auditory-visual mappings for sound visualisation; Kostas Giannakis. *Organised Sound; Cambridge*, *11*(3), 297–307.

Li, Y., Fang, C., Yang, J., Wang, Z., Lu, X., & Yang, M.-H. (2017). Universal Style Transfer via Feature Transforms. *Advances in Neural Information Processing Systems 30*, 386–396.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychological Research*, *58*(3), 177–192.

Soraghan, S. (2014). Animating Timbre-A User Study. In ICMC. Retrieved from http://www.smc-conference.net/smc-icmc-2014/images/proceedings/OS8-B04-AnimatingTimbre.pdf

Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics*, *42*(5), 491–502.

Sound Signifying Light: The Sunrise Topic as a Transmodal Perceptual Metaphor

Moe Touizrar
Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT),
Schulich School of Music, McGill University, Montreal, Quebec, Canada
moe.touizrar@mcgill.ca

Aims/goals

In recent literature, musicologists have identified the sun—and its light—as a proto-topical category in Western classical music. Unlike most other musical topics, light and its modulations exist as conspicuously visual phenomena, and this raises an important question about how topical metaphors are perceived when listening to music: what does it mean to hear a sunrise? And, why would so many composers, each ostensibly independently of one another, develop and adapt similar orchestral strategies and forms in order to convey an experience devoid of sound? In this paper, I aim to shed some light on this issue by drawing on recent developments in the perception and cognition of music to delineate the transmodal and metaphoric nature of the sunrise topic in orchestral music.

Background information

Although several scholars have raised the subject of timbral depictions of light in orchestral music (Loughridge, 2016; Sisman, 2013; Tarasti, 2012; Greenwald, 2010; Kramer, 2009), it was Sisman (2014) who first proposed that the orchestral sunrise might in fact constitute a discrete topic. And yet, despite these recent, albeit brief, scholarly mentions, none have expounded or populated the topic beyond a few isolated cases, let alone offered speculation as to how the compelling timbral-temporal metaphor might operate, either in the hands of the composer or in the mind of the listener.

Methodology

This poster presentation highlights the history and musical features of the sunlight topic in orchestral music. First, I enumerate instances of sunlight depiction in music in a timeline that runs from 1761-1913. Second, using the opening of Haydn's Symphony No. 6 (*Le matin*), I discuss ways in which orchestration helps create the energetic semblance of the sunrise, which is re-presented as a perceptible transmodal musical metaphor (Larson, 2012; Zbikowski, 2008; Bruhn, 2000; Hatten, 1995). Lastly, I combine evidence from studies in the perception of musical timbre (McAdams and Goodchild, 2017**)** with a listener-oriented phenomenology to show how the sonic phenomenon of the musical sunrise can be analyzed (Goodchild and McAdams, forthcoming).

Results

My speculations with regard to metaphorical perceptions of music demonstrate possible ways in which a purely visual phenomenon can be understood aurally. Correspondingly, I enumerate key orchestrational and temporal characteristics of musical sunrises, demonstrating that their selective assemblage and sequencing forms a powerful gestalt that composers from Haydn to Schoenberg have used to signify both the motion and the dynamic intensity of the solar orb.

Conclusions

The results of this research point not only to the establishment of the sunrise as a valid topic of musical discourse—a topic for which instrumental timbre and orchestration are essential; but, in addition, my preliminary conclusions concerning perceptual resemblance between visual and auditory scenes may also help to promote and facilitate new explanations for the perceptual semantics and pragmatics of transmodal metaphors in music.

References

Barsalou, L.W, Simmons, W.K., Barbey, A.K., and Wilson, C.D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Science*, 17:2, 84-91.

Bruhn, S. (2000). *Musical ekphrasis: Composers responding to poetry and painting*. Hillsdale, NY: Pendragon Press.

El Rafaie, E. (2015). Reconsidering "Image Metaphor" in light of perceptual simulation theory. *Metaphor and Symbol*, 30, 63-76.

Forceville, C. (2008). Metaphor in pictures and multimodal representations. In R.W. Gibbs, Jr. (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 462-482). Cambridge, MA: Cambridge University Press.

Goodchild, M., & McAdams, S. (forthcoming). Timbre perception and orchestration. In A. Rehding & E. Dolan (Eds.), *Oxford Handbook of Timbre*, New York, NY: Oxford University Press.

Greenwald, H. M. (2010). Son et lumière: Verdi, *Attila*, and the sunrise over the lagoon. *Cambridge Opera Journal*, 21(3), 267-277. DOI: 10.1017/S0954586710000170

Hatten, R. S. (1995). Metaphor in music. In E. Tarasti (Ed.), *Musical signification: Essays in the semiotic theory and analysis of music*. New York, NY: De Gruyter.

Kramer, L. (2009). Recalling the sublime: The logic of creation in Haydn's *Creation*. *Eighteenth-Century Music*, 6(1), 41-57.

Larson, S. (2012). *Musical forces: Motion, metaphor, and meaning in music*. Bloomington, IN: Indiana University Press.

Lakoff, G. & Johnson, M. (1980*). Metaphors we live by*. Chicago, IL: University of Chicago Press.

Loughridge, D. (2016) *Haydn's sunrise, Beethoven's shadow*. Chicago, IL: University of Chicago Press.

McAdams, S. & Goodchild, M. (2017). Musical structure: Sound and timbre. In R. Ashley & R. Timmers (Eds.), *The Routledge Companion to Music Cognition*, pp. 129-139, New York, NY: Routledge.

Sisman, E. (2014). Symphonies and the public display of topics. In D. Mirka (Ed.), *The Oxford Handbook of Topic Theory* (pp. 90-117). New York, NY: Oxford University Press.

_______. (2013). Haydn's solar poetics: The *Tageszeiten* symphonies and enlightenment knowledge. *Journal of the American Musicological Society*, 66(1), 5-102. DOI:10.1525/jams.2013.66.1.5.

Tarasti, E. (2012). The semiosis of light in music: From synaesthesias to narratives. *Semiotics of classical music: How Mozart, Brahms and Wagner talk to us* (Chapter 12, pp. 301-334). Boston, MA: De Gruyter.

Zbikowski, L. M. (2008). Music and metaphor. In R.W. Gibbs, Jr. (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 502-524). Cambridge, MA: Cambridge University Press.

New Wor(l)ds for Old Sounds

Margaret Schedel†, Erika Honisch
Music Department, Stony Brook University, Stony Brook, NY USA
† margaret.schedel@stonybrook.edu

Aims/goals

Evidence for the "sonic turn" in and beyond the humanities is everywhere: in the calls for papers of recent interdisciplinary conferences, in the popularity of sound-oriented blogs, in the formation of sound studies interest groups in academic professional societies, and in the collaborations of electroacoustic composers with social scientists. It is less evident—given the general emphasis in sound studies on contemporary sonic cultures and practices—that a significant line of inquiry focuses on the richly sonic past.

Background information

Studies exemplifying this historicist impulse draw attention to the acoustic properties of ancient and early modern spaces, and those of more recent built environments (Blesser and Salter, 2007; Fisher, 2014); they search archival documents for the sounds of colonial encounter (Rath 2005) and the hubbub of England in the Victorian period and earlier (Picker 2003; Cockayne, 2007); they find traces of the noisy mediaeval city in manuscript illuminations (Dillon 2012); they document sound and its silencing to trace shifting urban identities and values (Bjisterveld 2008; Thompson, 2002); they investigate the properties of instruments and technologies, from monochords to metronomes, developed to chart interval space and measure musical time (Grant, 2014); they consider the collision of early recording technology with traditional Western musical aesthetics (Rehding, 2005). Collaborative digital projects recreate past sound worlds, embedding reconstructed sounds in 3D virtual space, as in Mylène Pardoen's The Sounds of Eighteenth-Century Paris, or situating records (both aural and textual) of sound in specific locations, as with the ever-expanding London Sound Survey.

Methodology

The interest in timbre, changing technologies, and acoustics that animates these projects also drives the work of practitioners and historians of electroacoustic music. Indeed, the vocabulary and methodologies developed by electroacoustic musicians to build a sonic lexicon, research the sounds of the past, and contextualize the impact of technology on sonic creativity are ideally suited to historically oriented sound studies.

The purpose of this paper is to explore the many points of resonance between the questions raised by electroacoustic specialists and those taken up by scholars who work on the sounds of the pre-electric past. How can we build bridges between these two exciting fields? With this in mind, "New Wor(l)ds for Old Sounds" applies the insights afforded by electroacoustic technologies, practices and vocabularies to sounds and spaces before the widespread adoption of electric sound in North America and Europe, roughly 1925. By its very etymology "electroacoustic" implicates the electric; we have chosen to be deliberately provocative and to project electroacoustic music studies to pre-electric objects and, conversely, test methodologies as well as the relevance/applicability of historical knowledge to the field.

More specifically we probe how electroacoustic language can be fruitfully used to discuss technologies, compositions, and listening practices before the advent of recording and electronically generated sound. What kinds of sounds emerge when we examine textual documents or historical musical instruments using a vocabulary of timbre informed by electroacoustic music? What do the re-creative possibilities of electroacoustic technology tell us about the obsolete or imaginary musical instruments described in music theory treatises (Athanasius Kircher, Musurgia Universalis, 1650); the utopian sound-houses described by Francis Bacon (The New Atlantis, 1624); the "invisible music" channeled into the palace of Christian IV of Denmark (Spohr, 2012); the acoustic properties of the cavernous Salle des Machines in Berlioz's arias?

And on the other hand, how do pre-electric practices and technologies continue to inform current electroacoustic practices?

Results

We rethink of the relationship between past and present conceptions of timbre, space, and sonic ecology, and the history of sound-based listening.

Conclusions

Using the language of electro-acoustic music, we can more accurately describe the sound world of the pre-electric past.

Acknowledgments

This paper grows out of the editorial and a call for works for the journal Organised Sound. Special thanks to Leigh Landy for his support.

References

Bjisterveld, K. (2008). *Mechanical Sound: Technology, Culture and Public Problems of Noise in the Twentieth Century*. Cambridge, Mass.: MIT Press.

Blesser, B. and L.-R. Salter. (2007). *Spaces speak, are you listening? Experiencing aural architecture*. Cambridge, Mass.: MIT Press.

Cockayne, E. (2007). *Hubbub: Filth, Noise & Stench in England 1600-1770*. New Haven: Yale University Press.

Dillon, E. (2012). The Sense of Sound: Musical Meaning in France, 1260–1330. Oxford, New York: Oxford University Press.

Fisher, A. (2014). *Music, Piety, and Propaganda: The Soundscapes of Counter-Reformation Bavaria*. Oxford, New York: Oxford University Press.

Grant, R. (2014). *Beating Time and Measuring Music in the Early Modern Era*. Oxford, New York: Oxford University Press.

Picker, J. (2003). *Victorian Soundscapes*. Oxford, New York: Oxford University Press.

Rath, R. C. (2005). H*ow Early America Sounded*. Ithaca, NY: Cornell University Press.

Rehding, A. (2005). Wax Cylinder Revolutions. *Musical Quarterly*, *88* (2005): 123–160.

Spohr, A. (2012). 'This Charming Invention Created by the King' – Christian IV and His Invisible Music. *Danish Yearbook of Musicology*, 39: 13-33.

Thompson, E. (2002). *The Soundscape of Modernity: Architectural Acoustics and the Culture of Listening in America, 1900–1933*. Cambridge, Mass: MIT Press.

Local Sound Signatures for Music Recommendations

Laurent Pottier
CIEREC, UJM (Université Jean Monnet), Saint-Etienne-Lyon, France
laurent.pottier@univ-st-etienne.fr

Aims/goals

We lack tools to represent and analyze music (without using scores). This poses a problem for analyzing electroacoustic music, but also for music on which the score does not provide all the useful information for the analysis. This text describes new developments in the automated analysis of short sound samples (ca. 2 seconds). Our research focuses on identifying and characterizing the "sound qualities of music," seeking to determine what makes it possible for us to recognize an artist by listening to a musical excerpt of only one or two seconds of duration. Our goal is to try to find out how to characterize this global signature, a mixture of sounds and rhythms, to be able to describe it, to compare pieces of music and classify them.

Background information

In traditional western instrumental music, the study of timbre is usually reduced to that of orchestration. With electrified music, this does not adequately describe the nature of the sound, the nature of timbre currently depends more on many sound processing tools (especially on the guitar) and mixing techniques used in the studio. The case is even more complex with electroacoustic music for which any sound source can find musical applications.

In recent years, however, works on sound analysis has been carried out, particularly around the MIR (Music Information Retrieval) studies, but if we look at the archives of the latest ISMIR conferences, the main researches concern the characterization of the singing voice, the recognition of the harmonies, the instruments, the tempo, but very few of them are related to a global characterization of the timbre. Tools like the MIR Toolbox (Lartillot & Toiviainen, 2007), Sonic Vizualizer (Cook & Leech-Wilkinson, 2009) or Zsa.Descriptors (Malt & Jourdan - 2008) offer audio descriptors that allow, each of them, to measure one of the characteristics of the timbre, but these descriptors are much too reductive, even in being combined, to give a good description of the timbre. Furthermore, the development of audio streaming sites on the Internet has changed our consumer behavior in the music field. Faced with the overabundance of data, streaming audio companies develop recommendations tools that can help consumers to discover new music. A company like Niland (Tardieu, 2016) uses algorithms that calculate distances between pieces of music, based on spectral analyzes and deep learning algorithms. But these technologies are mainly intended to determine the genre of the music, the instrumentation, the tempo, even the mood it produces on the listener and are not adapted to no-mainstream music, such as electroacoustic music for example.

Methodology

Our sound analysis technique characterizes the sound qualities perceived in sound samples of short duration (~ 2 seconds), that corresponds to the durations for which our perception can recognize the nature, the origin, and even the composer of the music. This analysis is at the same time a spectral and a temporal analysis. The technique consists of splitting the sound into 27 channels (by 1/3 octave) and then measuring the RMS energy of each band at various sampling rates (more and more fine when reaching high frequencies). This step is realized inside the Max/MSP program. From these tables of values, which represent a "local signature of the music," we have extracted, using the python language, various spectral and rhythmical audio descriptors: single values (centroid, standard low and high deviation, skewness, kurtosis, rhythmical flux) and one-dimensional vectors (amplitude envelope, overall spectral envelope, multiband zero-crossing energy derivates). With these descriptors, we can calculate

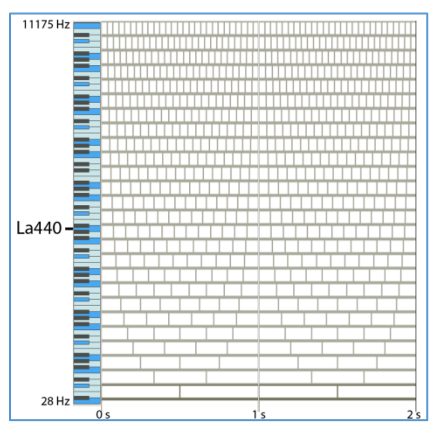


Figure 1: structure of the table for the analysis

various distances and similarities between signatures and proceed to unsupervised machine learning applications on a large set of musical pieces of various categories.

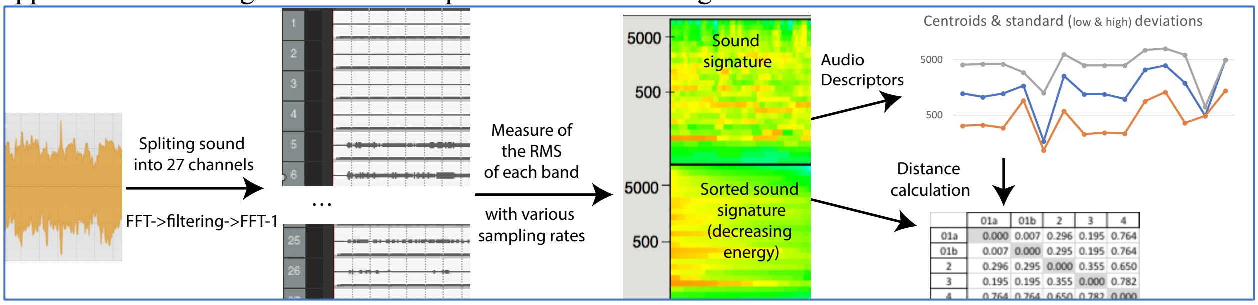


Figure 2: the successive steps of the analysis

Results

The analysis of the signature of the music was conducted on a set of varied music, including test sounds (sinus, pulse train, pink noise), commercial music (rock'n roll, progressive rock, noise music), contemporary music, electroacoustic music and a selection of indie music of 1D touch, a streaming platform provided by 1D Lab company. We built a database to store all the data (in MySQL). Over the next few weeks, we will run a machine learning application to compare the sound signatures of our database. The next step will be to offer musical recommendations, based on the sound signature, on larger music databases provided by streaming companies such as 1DLab (Diefenbach, 2016).

Conclusions

The method produces a table that constitutes an acoustic signature of a short musical extract. From this table, and some audio descriptors associated, it is possible to measure distances and estimate similarities between different pieces of music, or within the same piece. Our collaboration with the Vibrations Acoustique laboratory should allow to perform perceptive tests to assess the relevance of different calculations of distances and proximity between musical pieces and define categories of acoustic qualities of these extracts. Our collaboration with the Connected Intelligence team of the Hubert Curien laboratory and the 1DLab audio streaming company aims to enrich their methods of music recommendation systems.

Acknowledgments

We would like to thank the UJM Foundation, which funds this project, and our collaborators: 1DLab company, Lab. Vibrations Acoustic (INSA-Lyon), Hubert Curien laboratory (UJM), GRAME (Lyon).

References

Bonardi, A., Bossis, B., Couprie, P., Dufeu, F., Malt, M., Pottier, L. (2013). “Outils pour l'analyse de la musique électroacoustique”. *Proceedings of the Journées d'Informatique Musicale (JIM 2013),* 199-200.

Cook, N. & Leech-Wilkinson, D. (2009). A musicologist's guide to Sonic Visualiser. King's College London. Retrieved from http://www.charm.rhul.ac.uk/analysing/p9_1.html

Diefenbach, D., Lhérisson, P.-R., Muhlenbach, F., Maret P. (2016). “Computing the Semantic Relatedness of Music Genre using Semantic Web Data”. SEMANTiCS 2016, Retrieved from http://ceur-ws.org/Vol-1695/paper23.pdf

Lartillot, O. & P. Toiviainen. “A Matlab Toolbox for Musical Feature Extraction from Audio”, *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, 2007.

Malt, M. & Jourdan, E. (2008). “Zsa.Descriptors: a library for real-time descriptors analysis”. *Proceedings of the 5th Sound and Music Computing Conference*, Berlin, p.134-137.

Tardieu, D. (2016). “L'intelligence artificielle va changer la façon dont vous écoutez de la musique”. 18/04/2016, HUFFPOST, FR edition, Retrieved from http://www.huffingtonpost.fr

The (Obvious) Influence of Pitch and Dynamics on the Perceptual Dimensions of the Timbre of Musical Instruments

Saleh Siddiq[1†], Christoph Reuter[1], Isabella Czedik-Eysenberg[1] and Denis Knauf[2]
[1] Institute of Musicology, University of Vienna, Vienna, Austria
[2] Vienna University of Technology, Vienna, Austria
[†]Corresponding author email: saleh.siddiq@univie.ac.at

Aims/goals

The main objectives of this study are: (1) Identifying the perceptual dimensions of the timbre of musical instruments together with their respective physical correlates. (2) Factoring in pitch and musical dynamics in order to provide a reasonable basis for the investigation of musical instruments—in distinction from the timbre of single tones. (3) Adopting the methodical basis of classic Timbre Space (TS) studies and complementing it with the modern possibilities of sound analysis provided by Music Information Retrieval (MIR) as well as elaborated quantitative statistical evaluation.

Background information

According to the standard definition, timbre "is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar." (ANSI, 1960). Although this definition was heavily criticized (e.g. Shouten in Plomp, 1070; Bregman, 1990; Patterson, Walters, Monaghan, & Gourdrain, 2010), it effectively took pitch and dynamics (here: "loudness") out of the equation. Subsequently, researchers most often reduced musical instruments to one single note and thus compared them at the same pitch and loudness level (e.g. Wedin & Goude 1972, Grey 1975, Krumhansl 1989, McAdams, Winsberg, Donnadieu, de Soete & Krimphoff, 1995; Lakatos, 2000). Such studies, in rough accordance, established a spectral scale (mostly brightness) and a temporal scale (mainly onset-related) as the correlates of the first two perceptual dimensions. Dimension 3, if existent, doesn't appear to have such a commonly established correlate (Caclin, McAdams, Smith & Winsberg 2005, Reuter & Siddiq, 2017). On the downside, these results, as consistent and solid as they are, only apply to what was effectively investigated: isolated single sounds. Unfortunately, the results were taken as representative of the entire timbre spectrum of the instruments these sounds stemmed from. It is commonly known that any given musical instrument is capable of producing a vast spectrum of sounds (e.g. Houtsma, 1997) that vary in timbre, dependent of other factors like pitch, dynamics (here: force of excitation), and articulation. Studies with several tones per instrument suggest the global dimensions of timbre perception change, once pitch and/or dynamics are taken into account (e.g. Marozeau, Cheveigné, McAdams, & Winsberg, 2003; Handel & Erickson, 2004). Hence, the questions are: What actually, are the perceptual dimensions of the timbre of musical instruments and what are their salient underlying physical properties?

Methodology

In this experiment, the perceptual relations of five non-percussive musical instruments (bassoon, cello, clarinet, flute, trombone) were tested. All tested sounds were actually recorded sounds, taken from the Vienna Symphonic Library (VSL). Pitch and dynamics were incorporated by sampling three different notes in three different dynamic levels per instrument, leading to a total of 45 sounds.

In a hearing experiment (pairwise comparison), participants (n = 38) were asked to subjectively rate the dissimilarities of the 45 musical instrument sounds. Based on these ratings and by means of a multidimensional scaling (MDS), a low-dimensional perceptual space was calculated. This, so far, is basically an adoption of the methods of the classical TS studies (McAdams, 1999).

On top of that, hierarchical clustering was performed to explore the underlying relation of the point cloud in the MDS-configuration. With the help of modern MIR, several features of the sounds were extracted. Correlations between the feature values and ranks along the spatial dimensions were calculated for every sound in order to identify the most suitable physical features to explain these dimensions. Using machine learning methods, regression models were trained for each of the dimensions and implemented in the form of a software tool to predict and later evaluate positions of new sounds in the virtual space.

Results

The results reveal some interesting interdependent findings. In short, the main results are: (1) musical instruments expectedly strew all across the perceptual space. (2) Thus, musical instruments are rather heterogeneous tonal entities than singular one-timbre-objects. (3) Primarily pitch, but also dynamics, is revealed as a salient factor and an inseparable part of timbre.

Conclusions

According to these findings, it becomes clear that musical instruments cannot be appropriately tested based on a single tone. Moreover, and in contradiction to the standard definition, it is impossible to separate "the timbre" of a musical instrument from pitch. Increasing pitch necessarily means at least (i.e. ignoring more complex matters like registers) increasing brightness and decreasing roughness at the same time. Hence, the timbre of musical instruments has to be addressed on another layer than the timbre of single sounds. As a consequence, the global perceptual dimensions of "single-sound" TS must not be applied to entire musical instruments.

Acknowledgments

Supported by funds of the Österreichische Nationalbank (OeNB, Anniversary Fund, project number: 16473). Special thanks to the Vienna Symphonic Library for supporting our work with two free "Vienna Super Package" licenses.

References

ANSI (1960). *American standard acoustical terminology (including mechanical shock and vibration)*. New York, NY: American Standard Association.

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America*, 118(1), 471–482.

Grey, J. M. (1975). *An exploration of musical timbre using computer-based techniques for analysis, synthesis and perceptual scaling*. PhD dissertation, University of Stanford.

Handel, S., & Erickson, M. (2004). Sound source identification: The possible role of timbre Transformations. *Music Perception*, 21(4), 587–610.

Houtsma, A. (1997). Pitch and timbre: Definition, meaning and use. *Journal of New Music Research*, 26(2), 104–115.

Krumhansl, C. (1989). Why is musical timbre so hard to understand? In S. Nielzen & O. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43–53). Amsterdam, NL: Elsevier.

Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics*, 62(7), 1426–1439.

Marozeau, J., Cheveigné, A., McAdams, S., & Gourdrain, E. (2003). The dependency of timbre on fundamental frequency. *Journal of the Acoustical Society of America*, 114(5), 2946–2957.

McAdams, S. (1999): Perspectives on the contribution of timbre to musical structure. *Computer Music Journal*, 23(3), 85–102.

McAdams, S., Winsberg, S., Donnadieu, S., de Soete, G., & Krimphoff, J. (1995): Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58(3), 177–192.

Patterson, R., Walters, T., Monaghan, J., & Gourdrain, E. (2010). Reviewing the definition of timbre as it pertains to the perception of speech and musical sounds. In Lopez-Poveda, E., Palmer, A., & Meddis, R. (Eds.), *The Neurophysiological bases of auditory perception* (pp. 223–233), New York, NY: Springer

Reuter, C., & Siddiq, S. (2017). The colourful life of timbre spaces. Timbre concepts from early ideas to meta-timbre space and beyond. In Wöllner, C. (Ed.), *Body, sound and space in music and beyond:*

An Investigation of the Acoustic Features Supporting the Perception of Timbre and Other Auditory Objects

Mattson Ogg[1,2†] and L. Robert Slevc[1,2]
[1] Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD, USA
[2] Department of Psychology, University of Maryland, College Park, MD, USA
[†] mogg@umd.edu

Aims/goals

The primary aim of this research is to understand whether the perception of non-musical auditory objects be characterized along the same acoustic dimensions as musical timbre.

Background information

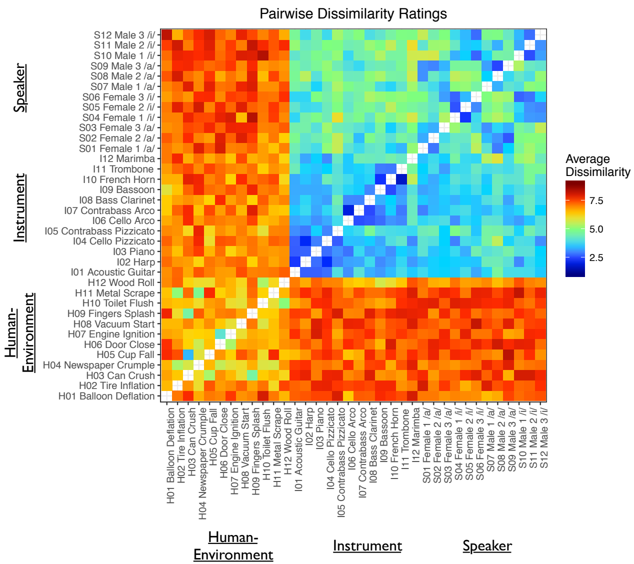
Timbre is a multi-dimensional attribute used to distinguish sound sources in music (McAdams & Giordano, 2009). Our understanding of timbre suggests clear cognitive and perceptual parallels to other domains of auditory object identification such as speaker identification (Baumann & Belin, 2010), speech perception (Piazza, Iordan & Lew-Williams, 2017), and environmental sound identification (Gygi, Kidd & Watson, 2007). Specifically, the identification of sounds across domains involves 1) a multidimensional set of cues, 2) matching stored representations of objects to acoustic cues that pertain to the features of a sound source, and 3) using this information to help organize auditory scenes. These similarities invite a closer examination of how these different domains of acoustic research (that are often investigated in isolation) might together provide a clearer understanding of the general acoustic properties that humans use to identify objects in their environment (Ogg, Slevc & Idsardi, 2017). Furthermore, the ubiquity of music in audiovisual works, such as film, and the increasing use of sample and field-recording based compositional techniques make this endeavor potentially valuable for creative and aesthetic pursuits.

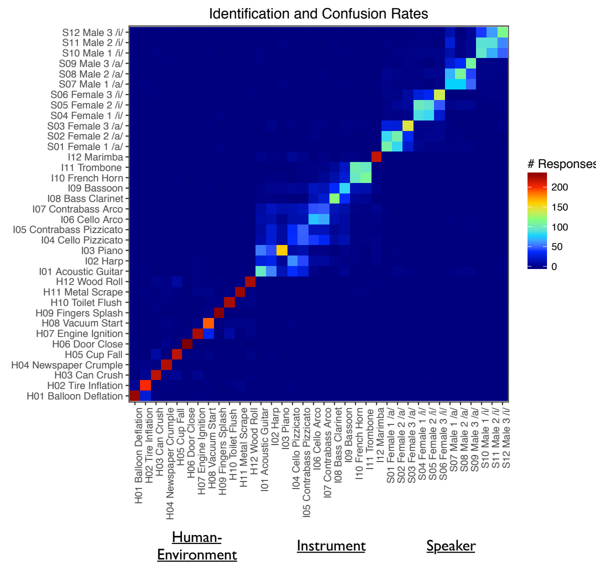
Methodology

To better understand the acoustic features listeners use to identify behaviorally relevant auditory objects we conducted a study of dissimilarity ratings (61 listeners each making 630 pairwise dissimilarity judgments) as well as identification and confusion rates (24 listeners each making 360 identification judgments of ten repetitions per sound source) for a large set of natural sound tokens (each 250 milliseconds in duration). Stimuli included 12 human-environmental sounds from everyday objects and materials, 12 instrument timbres, and 12 speech sounds (2 vowels from 6 different speakers: 3 male and 3 female). The pitches of the instrument sounds were matched to the pitches of the speech tokens. Responses were analyzed via multidimensional scaling and analyses of trial-level difference ratings or confusion rates as a function of the sounds' acoustic properties including noisiness (De Cheveigné & Kawahara, 2002), spectrotemporal variability (Elliot & Theunissen, 2009) and aspects of their temporal and spectral envelopes (Peeters, Giordano, Susini, Misdariis & McAdams 2011).

Results

Dissimilarity ratings and confusion rates indicated high levels of similarity among sounds within the same category and greater dissimilarity among sounds from different categories. Acoustic dimensions relevant to the study of musical timbre were reliably associated with these behavioral data, such as attack-time and features of the spectral envelopes as well as changes in the spectral envelope over time. Dissimilarity ratings and identification rates were also correlated with one another, suggesting that these provided converging evidence regarding timbre and sound source perception.





Conclusions

These results suggest that the acoustic dimensions relevant to timbre perception do, in fact, generalize to other domains of auditory object perception. Additionally, variability in spectral envelopes, which is sometimes inconsistently associated with timbre ratings, played a role in explaining the dissimilarity and confusion rates among this diverse set of sound sources. These results suggest a common perceptual and acoustic framework in which timbre and other sound sources are related to one another.

References

Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: common dimensions for different vowels and speakers. *Psychological Research*, *74*(1), 110-120.

De Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, *111*(4), 1917-1930.

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Computational Biology*, *5*(3), e1000302.

Gygi, B., Kidd, G. R., & Watson, C. S. (2007). Similarity and categorization of environmental sounds. *Attention, Perception, & Psychophysics, 69*(6), 839-855.

McAdams, S., & Giordano, B. L. (2009). The perception of musical timbre. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology*, (pp. 72– 80). New York, NY: Oxford University Press.

Ogg, M., Slevc, L. R., & Idsardi, W. J. (2017). The time course of sound category identification: Insights from acoustic features. *The Journal of the Acoustical Society of America*, *142*(6), 3459-3473.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The timbre toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, *130*(5), 2902-2916.

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology*, *27*(20), 3162-3167.

Timbre Spatialisation

Robert Normandeau
Faculté de musique, Université de Montréal, Montréal, Québec, Canada
robert.normandeau@umontreal.ca

Aims/goals

In this text, I will present *space* as a musical parameter that can now be considered as important in acousmatic music composition as the other musical parameters in instrumental music. Some aspects of the sound spatialisation could be considered as exclusive to the acousmatic language. The spatialisation of sound presented here is an immersive one, where the listeners are placed in a situation where speakers surround them. I will present my own cycle or works within which I have developed a new way to compose the space that I call *Timbre Spatialisation*.

With instrumental music in the ‘60s, composers explored spatialisation, creating works that assigned performers to different locations in the concert hall (Stockhausen’s *Gruppen* or *Carré*). However, these works are limited to the timbre of the instruments: the violin on the left side will always sound like a violin on the left. The sound and the projection source are linked together. What is specific to the acousmatic medium is its virtuality: the sound and the projection source are not linked.

A speaker can project any kind of timbre. Furthermore, today, with the appropriate software, all these sounds can be located at any point between any group of speakers. What is unique in electroacoustic music is the possibility to fragment sound spectra amongst a network of speakers. When a violin is played, the entire spectrum of the instrument sounds, whereas with multichannel electroacoustic music, timbre can be distributed over all virtual points available in the defined space.

This is what I call *timbre spatialisation*: the entire spectrum of a sound is recombined only virtually in the space of the concert hall. Each point represents only a part of the ensemble. It is not a conception of space that is added at the end of the composition process—a frequently seen approach, especially today with multitrack software—but a real composed spatialisation. *Timbre Spatialisation* is a musical parameter that is exclusive to acousmatic music.

Background information

A cycle of works started in 2001, that includes *StrinGDberg* (2001-03) 18', *Éden* (2003) 18', *Palindrome* (2006-07) 16', *Kuppel* (2006) 18', *Pluies noires* (2008) 18', *Anadliad* (2010) 14’, *La part des anges* (2012) 14’, *L'engloutissement* (2017) 17'.

Methodology

At the beginning this research was essentially intuitive. In the movement of *Clair de terre* (1999) entitled *Couleurs primaires et secondaires* (Primary and Secondary Colors), I had the idea to divide the timbre of a group of transformed sounds of a Balinese gamelan into different registers and to send these different registers to different speakers. It was only a short movement (2'54") of a large work (36'), but I had the feeling that this way of spatializing the timbre was quite novel at the time. I decided then to push my music a little bit further in that direction.

The first piece in which I systematically explore this concept is *StrinGDberg* (2001-03). The piece is constructed using different superimposed layers. The first layer is composed of a single recording of an improvisation on a hurdy-gurdy that lasts about a minute. Stretched, filtered, and layered, the sound of the hurdy-gurdy, distributed in a multiphonic space, is revealed, layer-by-layer, over the length of the piece. Out of the middle part of this improvisation, I kept twelve “consonants”—the attacks of the notes—and twelve “vowels”—the sustained parts between the attacks. Both the consonants and vowels were then “frozen.” The entire 24 were filtered by four dynamic band pass filters, the parameters of which changed over sections and time. The opening of each filter increased over the duration of the work and the center frequency is constantly changing. That means that the sound is globally filtered at the beginning of the work and it ends up at a point where the entire spectrum is unfiltered.

In *StrinGDberg*, the form, duration, rhythms and proportions were derived from the original improvised melody. All the sound files for the work were created and organized with multichannel diffusion in mind. This is another defining characteristic of what I call *Timbre Spatialisation*. The different filtered bands are assigned to different loudspeakers, sixteen in the original version, thirty-eight in the final version. The final mix is then formed in the concert hall and in different ways for every listener.

Results

As mentioned before, the research led to a cycle of works. But there is one aspect that occurred very rapidly at the beginning of the process: the lack of tools to easily work with the sound material. Since 2008, I have founded a research group called Groupe de recherche en immersion spatiale (GRIS) involved in the development of electroacoustic tools for sound Spatialisation in 2D and 3D. These tools are *Zirkaloy* (2009), *OctoGRIS* (2010-16), *ZirkOSC* (2012-15), *SpatGRIS* (2016-…) and *ServerGRIS* (2017). Essentially, these tools are integrated into the work flow of the composers to help them to experiment and explore this idea of *Timbre Spatialisation*.

Conclusions

We do not know very much about the perception of the space in a musical context. Of course, psycho-acousticians have explored the perception of distance and localization, but most of the time they have done so by using very primitive sounds, such as static, that don't address how musical gestures in space are perceived. As Robert Sazdov and others have mentioned recently, "Composers of electroacoustic music have engaged with elevated loudspeaker configurations since the first performances of these works in the 1950s. Currently, the majority of electroacoustic compositions continue to be presented with a horizontal loudspeaker configuration. Although human auditory perception is three-dimensional, music composition has not adequately exploited the creative possibilities of the elevated dimension" (Sazdov, Payne, Stevens 2007).

We still have a great deal to explore about space perception and I think that this research should be made in parallel with music composition. This research could include experiments on how we perceive sound movement in 3D, how we perceive the same musical gesture presented at the front of the audience compared to the back, how we perceive low frequency content from different locations, what the perceptual threshold regarding location angle is and thus, how many speakers we need to feel completely immersed by the sound. In other words, what is relevant in terms of musical perception in an immersive environment?

Acknowledgments

Thanks to Ludger Brümmer and the team at the ZKM (Karlsruhe, Germany) and to the team at the GRIS. Specail thanks to the actual programmer at the GRIS, Olivier Bélanger.

References

Normandeau, R., « Timbre spatialisation : the medium is the space », Organised Sound, vol. 14, no. 3, 2009.

Pulkki, V. 1997. Virtual sound source positioning using vector base amplitude panning. *Journal Audio Engineering Society,* 45(6): 456-466.

Sazdov, R., Paine, G. and SY, K., Perceptual Investigation into Envelopment, Spatial Clarity and Engulfment in 3D Reproduced Multi-channel Loudspeaker Configurations. Electroacoustic Music Studies 2007. Leicester: EMS.

H. Vaggione, « Décorrelation microtemporelle, morphologies et figurations spatiales du son musical », dans Es- paces sonores, p. 17–29, Paris : Editions musicales transatlantiques, 2002. http://jim.afim-asso.org/jim2002/articles/ L27_Vaggione.pdf.

Verron, C., Gauthier, P-A., Langlois, J, & Guastavino, C. (2013). Spectral and spatial multichannel analysis/synthesis of interior aircraft sounds. *IEEE transactions on Audio, Speech and Language Processing*, 21(7), 1317-1329.

Communication through Microphone Choice and Placement

Jack Kelly[1†], Julian Neri[2], Diego Quiroz[1], and Alejandro Aspinwall[1]
[1] Sound Recording Area, McGill University, Montréal, Québec, Canada
[2] Music Technology Area, McGill University, Montréal, Québec, Canada
[†] jack.kelly@mail.mcgill.ca

Aims/goals

The goal of the following experiment is to determine whether or not recording engineers are able to communicate specific intentions through microphone choice and placement. We aim to establish to what extent these intentions are successfully communicated to the listener. This research was undertaken as a seminar project, through which we formulated a pilot study. This paper explains the study and motivates further research into the engineer's influence on a sound's timbre and capacity for communication.

Background information

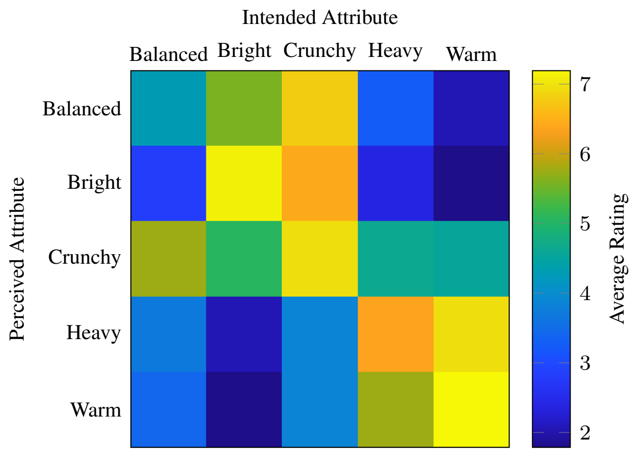
Electroacoustic transducers have an undeniable influence on the timbre of a recorded sound source. It has been established that loudspeakers apply a unique timbral filter to recorded sounds. Listeners are able to distinguish one type of loudspeaker from another, when timbre is the only variable. On the sound capture side, microphone choice and placement can alter the temporal response and frequency content of the source. The decisions made by sound engineers influence the resulting attack time, spectral centroid, and spectral flux of a sound source. Given the engineer's ability to manipulate the timbre of the sound source through recording decisions, we hypothesize that the recording engineer can convey specific perceptual intentions (a sound that is "crunchy," "bright," or "warm," for instance) to the listener.

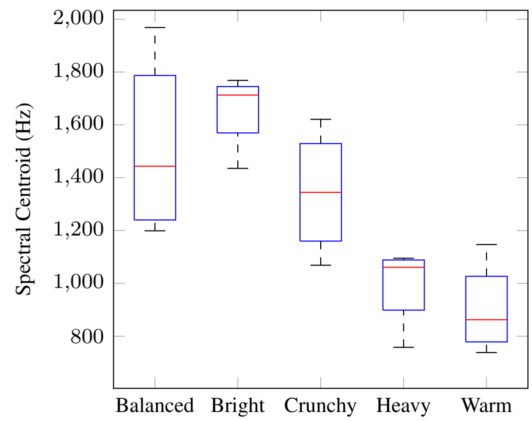
Methodology

We applied Juslin's standard paradigm to explore whether the recording engineer can communicate their intentions through expressive cues to the listener, consisting of an encoding and decoding experiment. For the encoding portion, 4 recording engineers were presented with an electric guitar amplifier that had a pre-recorded electric guitar track looping through it. The engineers were given a selection of microphones of various brands, polar patterns, and operation principles, and a list of 5 specific intentions: balanced, bright, crunchy, heavy, and warm. They were asked to record the amplifier using one microphone and placement strategy per intention. For the decoding portion, a listening test was run in which 8 listeners were asked to rate the level of each possible intention on a scale of 0 to 10 for each recording. The audio samples were loudness matched and randomized. The set of 20 ratings from each listener was normalized.

Results

The data were formatted into confusion matrices, showing how well the engineer communicated an intent to the listener and whether some intents were confused for others. The columns correspond to the engineer's intended attribute while the rows correspond to the listener's perceived attribute. Each cell displays the average rating given by the listeners for an encoding-decoding pair. Ideally, the matrix has ratings of 10 along the diagonal and 0 elsewhere. The confusion matrix shown below includes ratings averaged over all engineers and listeners. Overall, the listeners successfully classified bright, crunchy, heavy, and warm. Individually, two engineers successfully communicated balanced while two did not (individual matrices not shown). The spectral centroid for each of the 20 recordings was computed to compare recording decisions. The variance of spectral centroid for bright, heavy, and warm is relatively small. Listeners tended to confuse balanced and crunchy recordings, which have relatively large variances and close spectral centroids.





Conclusions

Our pilot study shows that we are indeed able to measure the effectiveness of intent-based communication from engineer to listener via timbral filtration. This lays the groundwork for future studies on the communication between recording engineer and listener, and the effects of specific recording techniques on the captured sound's timbre. Future research will extend this study by considering a variety of instrument types and larger set of sound attributes to encode and decode.

Acknowledgments

We wish to acknowledge the advice and guidance of Dr. Joe Plazak, the course instructor who guided our preliminary research. This research will be conducted with the support of CIRMMT and the Department of Music Research at the Schulich School of Music, McGill University.

References

B. Bartlett, Tonal effects of close microphone placement (1981). *J. Audio Eng. Soc*., vol. 29, no. 10, pp. 726-738.

B. Bartlett, Microphones: Small is Beautiful (1986). in *AES 81st Convention*, Los Angeles, USA.

B. Bartlett, Choosing the Right Microphone by Understanding Design Tradeoffs (1987) *Journal of the Audio Engineering Society*, Vol 35, No 11, pp. 924 - 944.

J. Eargle (2004) *The Microphone Book*. New York: Van Nostrand Reihold Company. Ch. 1 – 8

M. Hibbing (2001). Design of Studio Microphones with Extended High-Frequency Response, in *AES 111th Convention*, New York, USA, pp. 1-3.

D. Josephson (1999). A Brief Tutorial on Proximity Effect, in *AES 107th Convention*, New York, USA, pp. 1-13.

P. Juslin (1997). Emotional communication in music performance: a functionalist perspective and some data, *Music Perception: An Interdisciplinary Journal*, vol. 14, no. 4, pp. 383-418.

M. Lavandier, P. Herzog, and S. Meunier (2004). The restitution of timbre by loudspeakers in a listening room: perceptual and physical measurements, *AES 117th Convention*, San Francisco, CA.

The Tone-Colour Melodies of the Japanese Shakuhachi

Bruno Deschênes, independent ethnomusicologist, shakuhachi player
(has obtained his *shihan*, Japanese master title, in 2016)
Email: bruno@musis.ca

Aims/goals

The pieces of the solo repertoire for *shakuhachi*, a Japanese bamboo flute originally played by the monks of a Zen Buddhist sect during Japan's Edo period (1603-1868), are considered by Japanese *shakuhachi* masters to be tone-colour melodies, not melodies of tones. This musical event will give a succinct overview of how these melodic lines are based on tone-colours.

Description of event

The event will be in 2 parts. The first part will show how the phrases and the tones of the solo pieces for *shakuhachi* are based on tone-colour. Examples will be performed on the *shakuhachi*, while their notated versions will be projected on a screen, both in Japanese and Western notations. The traditional notation for *shakuhachi* uses Japanese characters (from the *katakana* syllabary); they do not refer to pitches but to fingerings. Some tones can be produced with different fingerings or with head movements, thus producing distinctive tone-colours. For example, in a phrase, a pitch can be heard 2 or 3 times with different fingerings, thus producing distinct tone-colours for each of them. Others can be produced with partial opening of holes. The standard notes produced while covering the holes are clear, while the ones with partial openings are softer. The difference of intensities between tones has become a characteristic of the tone-colour aesthetic of this music. Few other stylistic characteristics of the playing of the *shakuhachi* will be presented as well.

In the second part, a piece (or 2, if time permits) will be performed.

The shakuhachi in context

The Japanese *shakuhachi* masters I studied with told me on numerous occasions that being in pitch is secondary in playing the traditional solo repertoire for *shakuhachi*. The way this five-hole flute is made makes it nearly impossible to be perfectly in tune according to the Western tempered scale. When performing, *shakuhachi* players must put the emphasis not on the pitches, but on the tone-colour of each tone produced. These pieces were not producing harmonious melodic lines. For these monks, the *shakuhachi* was not a musical instrument, but a spiritual tool. Moreover, the flute was not tuned according to a scale or a mode. At the end of the 19th century *shakuhachi* makers started to pay attention to tuning, so that it could play with the *koto* and the *shamisen*, as well as compete with Western instruments. The music these monks composed has no common measure with folk music or the music from the entertainment world of the Edo era.

Historically, the melodies of all genres of traditional Japanese music are composed based on tone-colour, as indicated by Tokumaru (1991, 2000) for the *shamisen*, a three-string lute, and the *tsuzumi*, a Japanese percussion; and de Ferranti (2000) for the *biwa*, the Japanese pear-shaped lute. Surprisingly, there are no extensive studies of tone-colour on the *shakuhachi*.

Today, the emphasis on tone-colour on the *shakuhachi* does not concern only the tones, but as well the choice of flutes of different lengths for any pieces being performed. This is a 20th-century development. There are 2 main types of *shakuhachi* in Japan: the traditional one in which the nodes inside the bamboo are drilled and the bore is left as is; and a modern one in which the bore is filled with plaster and lacquer to make it even, as well as making it louder and more in tune with Western instruments. A musician will choose among his many flutes the one he considers has the best overall tone-color for a particular piece. Such choices are not made based on tuning, tonality or pitch, but on the distinctive tone-colour of a flute.

Interestingly, two flutes made by a *shakuhachi* maker will differ in tone-colour, but more so between makers, as well as between flutes of different lengths, and between the modern flute and the traditional one.

In learning the *shakuhachi*, the student must learn to decondition herself or himself from the Western beliefs that a melody is solely based on tones or pitches, and pay attention to tone-colours, which can be arduous for non-Japanese musicians trained in Western music.

References

Deschênes, Bruno (2017a). *Le* shakuhachi *japonais, Une tradition réinventée*, Paris: L'Harmattan.

Deschênes, Bruno. (2017b). "A preliminary approach to the analysis of *honkyoku*, the solo repertoire of the Japanese *shakuhachi.*" *NEMO-Online* 4-6:123-143. http://nemo-online.org/archives/1674.

de Ferranti, Hugh (2000). *Japanese Musical Instruments*, Oxford: Oxford University Press.

Tokita, Alison McQueen (1996). "Mode and Scale, Modulation and Tuning in Japanese Shamisen Music: The Case of the *Kyomoto* Narrative," *Ethnomusicology* 40-1:1-33.

Tokumaru, Yoshihiko (1991). "Le timbre dans la musique japonaise," in *Le timbre, Métaphore pour la composition*. Paris: Ircam-Centre Pompidou/Christian Bourgeois Éditions, p. 90-97.

Tokumaru, Yoshihiko (2000). *L'aspect mélodique de la musique de* syamisen, Paris: Éditions Peeters.

Testing Relative Perception and Context-Sensitivity of Timbral Brightness

Kai Siedenburg
Department of Medical Physics and Acoustics, University of Oldenburg, Oldenburg, Germany
kai.siedenburg@uni-oldenburg.de

Aims/goals

The goal of this study is to better understand the relative perception of timbral brightness. Brightness perception is a central component of timbre and an important parameter of the musical discourse. In this presentation, I report on two experiments that revealed novel effects of ambiguity and context sensitivity of brightness shift judgements.

Background information

A critical issue for cognitive music theorists, composers, and psychologists alike is the question whether pitch and aspects of timbre share principles of perceptual organization. From a theoretical perspective, Arnold Schoenberg once speculated that "pitch is nothing but tone color measured in one direction" (1911/1978, p. 421). In the psychological literature, however, comparative approaches to the study of both attributes are relatively rare (although see, McDermott & Oxenham, 2008).

One important aspect concerns the phenomenon of relative perception, which is a hallmark feature of pitch: melodies can generally be transposed along frequency without losing their identity. For timbre, the literature suggests that the contour of single-dimensional shifts is well perceivable (Wood, 2015; McDermott & Oxenham, 2008), whereas listeners have difficulties with recognizing transpositions of intervals in multi-dimensional timbre space (Ehresman & Wessel, 1978; McAdams & Cunible, 1992). Another question concerns the aspect of context sensitivity, where Chambers et al. (2017) demonstrated rapid and long-lasting context effects for shift judgments, using so-called "Shepard tones". These tones possess octave-space partials and a bell-shaped spectral envelope. Shifting partials along (log-)frequency while keeping the envelope fixed generates the well-known Shepard (-Risset) illusion—a sequence of tones with continually ascending or descending pitch, which has been used in numerous musical compositions. Another important feature of Shepard tones concerns the aspect of perceptual ambiguity. As for half-octave pitch shifts, listeners are on average equally likely to report upward or downward pitch shifts (Shepard, 1964). Chambers et al. capitalized on this ambiguity and demonstrated that in the presence of a prior context, participants tended to report shifts (up/down) that best encompassed the frequency components of the context.

In this presentation, I will present two listening experiments that were based on a transfer of the classic (pitch-based) Shepard illusion to timbre. The basic idea was to generate stimuli which allowed for cyclic shifts of brightness while pitch was kept fixed. The experiments tested whether the direction of brightness shifts was perceived in a systematic manner and whether the reported context effects of Chambers et al. would generalize to timbral brightness perception.

Methodology

Stimuli were digitally synthesized tones and allowed to vary the spectral envelope, and hence manipulate brightness instead of pitch. In contrast to the octave-spaced partials of classic Shepard tones, here sounds were characterized by harmonically-spaced partials that ensured a fixed pitch percept at the fundamental frequency f0 (additionally amplified to ensure a robust pitch percept at f0). The partials' amplitudes were determined by a spectral envelope that was a combination of a global bell-shaped curve and a more fine-grained cyclic envelope component that affected brightness if shifted along frequency. Note that the resulting cyclic nature of fine-grained envelope shifts creates a timbral analogy to the classic pitch-based Shepard illusion: sound sequences with fixed pitch but continually increasing or decreasing brightness.

The rest of the experiments were modeled after Exps. 1 and 2 of Chambers et al. (2017). Exp. 1 presented isolated pairs of tones (T1 and T2) and listeners judged whether they perceived T1 or T2 as brighter. The tones T1 and T2 differed by a shift of the fine-grained components of the envelopes by 1/6, 1/3, 1/2, 2/3, or 5/6 octaves. In a second experiment, zero to four tones preceded the target pair. Context tones had fine-grained envelopes with shifts that were randomly selected from the range 1/12-5/12 octaves or the other half of the octave (7/12-11/12).

Results

Effects unfolded analogously to the original pitch experiments (Chambers et al., 2017): Exp. 1 demonstrated that shifts of the cyclic envelope component were perceived as systematic changes in brightness. Furthermore, half-octave shifts of the fine-grained spectral envelope yielded perceptual ambiguity in the sense that participants were roughly as likely to report upward as downward shifts. Results from Exp. 2 demonstrated that brightness judgments for ambiguous shifts were strongly biased towards the direction of the prior context.

Conclusions

These findings constitute the first systematic effect of ambiguity in the relative perception of timbre and establish an effect of context sensitivity of timbral brightness judgements. Moreover, the striking similarities to the findings for pitch shifts (Chambers et al., 2017) suggest the existence of analogous principles of perceptual processing of pitch and timbral brightness and invite further psychological and musical exploration.

Acknowledgments

This work was supported by a Carl von Ossietzky Young Researchers' Fellowship from the University of Oldenburg.

References

Chambers, C., Akram, S., Adam, V., Pelofi, C., Sahani, M., Shamma, S., and Pressnitzer, D. (2017). Prior context in audition informs binding and shapes simple features. *Nature Communications*, 8:15027.

Ehresman, D. and Wessel, D. L. (1978). Perception of timbral analogies, vol. 13 of *Rapports de l'IRCAM.* Paris, France: IRCAM Centre Pompidou.

McAdams, S. and Cunible, J.-C. (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society: Biological Sciences*, 336:383-389.

McDermott, J. H., Lehr, A. J., and Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science*, 19(12):1263-1271.

Schoenberg, A. (1911/1978). Theory of Harmony [Harmonielehre]. University of California Press (R. E. Carter, Trans. from original German edition, 1911), Berkeley, CA.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, 36(12):2346-2353.

Wood, C. (2015). Perception of timbre intervals. Master's thesis, McGill University Montréal, Canada.

Exploring the Tole of Source-Cause Categories in Timbral Brightness Perception

Charalampos Saitis[1†] and Kai Siedenburg[2]
[1] Audio Communication Group, Technical University of Berlin, Berlin, Germany
[2] Department of Medical Physics and Acoustics, University of Oldenburg, Oldenburg, Germany
[†]charalampos.saitis@campus.tu-berlin.de

Aims/goals

This paper reports an experiment investigating the role of sound source categorical information in the semantic description of timbre. Ratings of dissimilarity based on brightness were compared with ratings of general dissimilarity to examine whether the former rely more on acoustic cues or are subject to source-cause categories as the latter (Siedenburg *et al.*, 2016). An additional goal was to address the question whether direct brightness dissimilarities would yield a similar ordering of sounds compared to the dimension of the general dissimilarity space usually associated with brightness (e.g., Zacharakis *et al.*, 2015). Approaching these questions constitutes an important step towards consolidating our understanding of musical brightness—one of the perceptually most salient attributes of timbre which plays important roles in music composition and the perception and induction of musical emotions (McAdams *et al.*, 2017; Goodchild, 2016).

Background information

Much previous work on timbre semantics has considered only recorded notes from acoustic musical instruments or their synthetic emulations (see Zacharakis *et al.*, 2014, for a review and new findings). In a recent study, Siedenburg and colleagues (2016) have shown that long-term familiarity with and knowledge about sound source categories (e.g., bowed or plucked strings, blown woodwinds, struck timpani) influence the perception of timbre as manifested in dissimilarity ratings. Here we hypothesized that categorical cues may further exert an effect on the semantic description of musical timbre because of the strong link between linguistic and perceptual representations (Zacharakis *et al.*, 2015; Samoylenko *et al.*, 1996).

Methodology

Participants

Thirty musicians took part in the experiment (15 females, 15 males; average age = 24 years, SD = 5 years, range = 18–39 years). They had at least 12 years of musical experience (average years of instruction and performance = 20 years, SD = 7 years, range = 12–33 years) and were remunerated for their participation.

Stimuli

Stimuli consisted of the same 14 recordings of single tones from common musical instruments used by Siedenburg et al. (2015, Sec. 2.1.2). All sounds had a fundamental frequency of 311 Hz and a duration of 500 ms, were equalized in perceived loudness, and only left channels were used.

Procedure

Participants first completed a standard pure-tone audiogram to ensure normal hearing with hearing thresholds of 20 dB HL or higher. Subsequently, they listened to all sounds in pseudorandom order to familiarize with the different timbres available in the set. The main experiment comprised three tasks carried out in the same order for all participants. In the first task, two stimuli were presented successively with an interstimulus interval of 300 ms and participants were asked to rate how dissimilar the two sounds were. Ratings were provided through a continuous scale with marks between 1 (= identical) and 9 (= very dissimilar) at the extremes. Each stimulus pair was presented once in one order (AB or BA for sounds A and B) and the order of presentation was counterbalanced across individuals. Pairs of identical stimuli were included, yielding 105 comparisons in total. Participants could listen to each pair as many times as desired but were encouraged to move at a reasonable pace. Four example trials were given. The second task was identical to the first, but this time participants were instructed to rate how dissimilar the two sounds were in terms of brightness. In the third task, all sounds were presented together, and participants were asked to rate each on brightness. The whole experimental session lasted approximately 70 min.

Analysis

A modeling framework analogous to the one used by Siedenburg et al. (2016) was employed: Brightness dissimilarity ratings (second task) were predicted using a partial least-squares regression model that takes audio descriptors from the Timbre Toolbox (Peeters *et al*., 2011) as regressors. It was then tested whether adding predictors derived from sound source-cause categories significantly improves the model fit. General dissimilarity ratings (first task) were analyzed using multidimensional scaling (MDS). The best fitting space was subsequently explored regarding a candidate dimension that maps onto differences in brightness, and dissimilarities along this dimension were then correlated to those from the brightness ratings. Finally, simultaneous brightness ratings (third task) were used to obtain a direct scaling of the attribute across the tested sounds.

Results

We hypothesize that previous findings regarding the influence of source-cause categories in dissimilarity ratings (Siedenburg *et al*., 2016) are also to be found in the semantic scenario of brightness ratings. We further hypothesize that the brightness dimension obtained from MDS analysis of general dissimilarity ratings converges on the same orderings of sounds as for direct brightness ratings.

Conclusions

Evidence of categorical effects in the semantic description of timbre corroborates its hybrid nature, where the audio-induced sensory representation available to the listener further induces cognitive-categorical information stored in long-term memory. Furthermore, considering the affective mediation of musical semantics (Koelsch, 2011), a better understanding of timbral brightness perception can help improve timbre blending and orchestration strategies for conveying emotional intention in music.

Acknowledgments

CS wishes to thank the Alexander von Humboldt Foundation for support through a Humboldt Research Fellowship. KS is supported through a Carl von Ossietzky Young Researchers' Award.

References

Goodchild, M. (2016) *Orchestral gestures: Music-theoretical perspectives and emotional responses*. PhD thesis, Dept. of Music Research, McGill University, Montreal, Canada.

Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Frontiers in Psychology*, 2:110.

McAdams, S., Douglas, C., and Vempala, N. N. (2017). Perception and modeling of affective qualities of musical instrument sounds across pitch registers. *Frontiers in Psychology*, 8:153.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., and McAdams, S. (2011). The Timbre Toolbox: Extracting audio descriptors from musical signals. *J. Acoust. Soc. Am*., 130(5):2902–2916.

Samoylenko, E., McAdams, S., and Nosulenko, V. (1996). Systematic analysis of verbalizations produced in comparing musical timbres. *Int. J. Psych.,* 31(6):255–278.

Siedenburg, K., Jones-Mollerup, K., and McAdams, S. (2016). Acoustic and categorical dissimilarity of musical timbre: Evidence from asymmetries between acoustic and chimeric sounds. *Frontiers in Psychology*, 6:1977.

Zacharakis, A., Pastiadis, K., and Reiss, J. D. (2014). An interlanguage study of musical timbre semantic dimensions and their acoustic correlates. *Music Percept.,* 31(4):339–358.

Zacharakis, A., Pastiadis, K., and Reiss, J. D. (2015). An interlanguage unification of musical timbre: Bridging semantic, perceptual, and acoustic dimensions. *Music Percept.,* 32(4):394–412.

A Computational Meta-analysis of Human Dissimilarity Ratings of Musical Instrument Timbre

Etienne Thoret[1†], Baptiste Caramiaux[2], Philippe Depalle[1], Stephen McAdams[1]
[1]Schulich School of Music, McGill University, Montreal, QC, Canada
[2] LRI, Univ. Paris-Sud, CNRS, Inria, Université Paris-Saclay
[†] etienne.thoret@mcgill.ca

Aims/goals

In this study, we modeled the perceptual dissimilarity ratings of former studies on timbre spaces with a computational framework. The first goal is to shed light on the relevant acoustic cues that have likely been used by listeners to make these dissimilarity ratings.

Background information

Musical instrument timbre has been intensively investigated through dissimilarity rating tasks (Grey, 1977, Grey & Gordon, 1978, McAdams et al., 1995, Iverson & Krumhansl, 1993; Lakatos, 2000; Krumhansl, 1989). It is now well known that audio descriptors such as attack time and spectral centroid, among others, account well for the dimensions of the timbre spaces underlying these dissimilarity ratings. Nevertheless, it remains very difficult to reproduce these perceptual judgments from distances computed on acoustical representations such as the waveform or the spectrogram. Biologically inspired representations based on spectrotemporal modulation spectra such as spectrotemporal receptive fields (STRF) have been shown to be well suited to reproduce human dissimilarity ratings (Patil et al., 2012). Moreover, recent studies have shown that Modulation Power Spectra (MPS), a similar kind of representation, are well suited to represent musical instrument timbre (Elliott et al., 2013; Thoret et al., 2016, 2017).

Methodology

Here we propose a meta-analysis of six former studies on timbre spaces in light of these recently developed representations (Grey, 1977, Grey & Gordon, 1978, McAdams et al., 1995, Iverson & Krumhansl, 1993; Lakatos, 2000; Patil et al., 2012). We implemented a computational framework that optimizes the correlation between the perceptual results and distances obtained from a set of different acoustic representations. This framework is inspired by the one proposed by Patil et al. (2012). It applies a dimension reduction (PCA) to a representation when its dimensionality is too high, and then optimizes a Gaussian kernel in order to maximize the correlation between perceptual distances and objective distances computed between reduced representations. Here we distinguished two different kinds of representations, those inspired by the physiology of the auditory system and those that are computed with traditional series of Fourier transforms. For each kind of representation, we considered four different representations: spectrum, spectrogram, MPS and STRF. In the case of the auditory model and the STRF representation, this study replicates Patil et al. (2012). We then extend this study to seven other representations and five other studies. In addition to this analysis, we also analyzed how the optimization process, i.e., the Gaussian Kernel, filters the original representations in order to maximize the correlation.

Results

We observed that distances computed from the STRF provide the best correlation with the perceptual results across the 12 timbre spaces, whereas other representations such as the MPS or the Fourier spectrum provided lower correlations. This is globally coherent with the results of Patil et al. (2012). Nevertheless, we observed that the correlation depends on the timbre space being modeled. This result strongly suggests a context effect of the timbre space sound set on the human dissimilarity ratings. Finally, for the studies that provided a fairly high correlation between perceptual dissimilarity ratings and mathematical distances, we highlighted the parts of the representations contributing the most to the correlation, suggesting new insights into the underlying perceptual metrics. For example, in the case of the Fourier spectrogram, we observed that the beginning of sounds and specific spectral shapes were emphasized depending on the timbre space. Finally, we showed that the optimized Gaussian kernel generalizes fairly well across the different timbre spaces for STRF representation, although it remains fairly low for lower level representations such as the spectrum or spectrogram. This result is quite promising as we may expect to determine a general perceptual metric to model human dissimilarity ratings based on the STRF representation. This expectation is also supported by the fact that Gaussian kernel generalize the best for STRF representations with an average Spearman correlation of .6 between the optimized Gaussian kernels for the 12 different timbre spaces.

Conclusions

These results complement the conclusions drawn by Patil et al. (2012) who suggested that the STRF contained the relevant information to reproduce the perceptual dissimilarity ratings of musical instrument timbres. Nevertheless, we here observed a context effect that depends on the timbre space being modeled.

Acknowledgments

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada awarded to SM (RGPIN-2015-05208, RGPAS-478121-15) and to PD (RGPIN-262808-2012) as well as a Canada Research Chair awarded to SM and an EU Marie Skłodowska-Curie fellowship to BC (Project MIM, H2020-MSCA-IF-2014, GA no. 659).

References

Elliott, T. M., Hamilton, L. S., & Theunissen, F. E. (2013). Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones. *The Journal of the Acoustical Society of America*, 133(1), 389-404.

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbrea. *The Journal of the Acoustical Society of America*, 94(5), 2595-2603.

Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In S. Nielzén & O. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43-53). Amsterdam, The Netherlands: Excerpta Medica.

Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics*, 62(7), 1426-1439.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *The Journal of the Acoustical Society of America*, 61(5), 1270-1277.

Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *The Journal of the Acoustical Society of America*, 63(5), 1493-1500.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58(3), 177-192.

Patil, K., Pressnitzer, D., Shamma, S., & Elhilali, M. (2012). Music in our ears: the biological bases of musical timbre perception. *PLoS Computational Biology*, 8(11), e1002759.

Thoret, E., Depalle, P., & McAdams, S. (2016). Perceptually salient spectrotemporal modulations for recognition of sustained musical instruments. *The Journal of the Acoustical Society of America*, 140(6), EL478-EL483.

Thoret, E., Depalle, P., & McAdams, S. (2017). Perceptually salient regions of the modulation power spectrum for musical instrument identification. *Frontiers in Psychology*, 8, 587.

Measuring Temporal Interdependence between Duet Performers along Timbre, Dynamics, and Pitch

Sven-Amin Lembke
Music, Technology and Innovation Research Centre, De Montfort University, Leicester, United Kingdom
sven-amin.lembke@dmu.ac.uk

Aims/goals

Several computational techniques were explored to identify and characterize the interdependence between musicians in duet performances, both in its magnitude and orientation. The focus lay on detecting temporal patterns of interdependence for acoustic measures related to timbre, dynamics, and pitch.

Background information

Timbre has received little attention in the study of musical performance, especially concerning the coordination of musicians in ensemble playing. Lembke et al. (2017) showed that the role of leadership affects timbral coordination of musicians in that accompanying performers (followers) adjust toward a darker timbre, accompanied by slightly softer dynamics, compared to when the same musicians lead in performance. This was quantified through corresponding decreases in spectral frequency (e.g., centroid) and sound level, respectively. Furthermore, this coordination between musicians increases in the course of a performance, as measured through coefficients of variation and cross correlation. Although these last two measures can assess the degree of coordination, the orientation of the influence, commonly known as the directionality or causality, still remained unknown. Dedicated measures can be used for the purpose of identifying the directionality of interdependence (Papiotis et al., 2014), based on autoregressive or non-linear techniques.

Method

Three measures were compared in their utility to identify and describe directional interdependencies between performers, namely, one measure based on Granger causality (Barnett & Seth, 2014) and two measures related to non-linear coupling (Chicharro & Andrzejak, 2009; Sugihara et al., 2012). Pairwise interdependence coefficients were computed for time series of the acoustic measures spectral centroid, relative sound level, and fundamental frequency that accounted for coordination of timbre, dynamics, and pitch, respectively.

Interdependence was evaluated at each of 16 measures of performances between one horn and one bassoon player in unison, based on an excerpt from Mendelssohn-Bartholdy's *A Midsummer Night's Dream* (see Lembke et al., 2017). Given that leadership was known to have affected timbral adjustments, separate analyses were considered for each role assignment (e.g., leading bassoon with following horn and vice versa). In addition, a further distinction into two eight-measure phrases was made, since the later phrase also exhibited higher coordination between performers. Each interdependence analysis comprised a total of 64 performances across eight different pairs of performers.

Results

Among the three tested measures, Granger causality revealed the least clear patterns of time-series interdependence, while also showing a tendency to fail obtaining autoregressive models for up to 15% of the time-series data. By contrast, both tested non-linear coupling techniques were more robust and always yielded results; the non-linear coupling coefficient L (Chicharro & Andrzejak, 2009) revealed the clearest patterns of interdependence, in agreement with its previous evaluation (Papiotis et al., 2014).

Different patterns of interdependence emerged across the acoustic measures. For spectral centroid (timbre) and sound level (dynamics), a clear tendency for bassoon players to adjust toward horn players became apparent, although this only manifested itself in the second phrase. As illustrated for the second phrase in Fig. 1 (right panel), bassoonists (circles) exhibited higher L coefficients than hornists (diamonds) in most measures, which signifies a directional dependence of bassoonists on hornists. By contrast, for the coordination of fundamental frequency (pitch), horn players showed a tendency to adjust towards bassoonists, moreover, to about the same degree throughout both phrases.

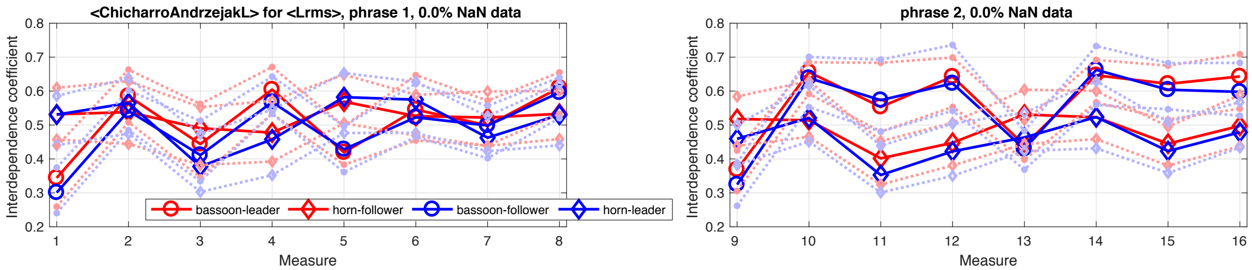


Fig. 1: Interdependence coefficient L (Chicharro & Andrzejak, 2009) for sound level (L_{rms}) per measure of a 16-measure excerpt, separated into two phrases. Solid lines and large symbols represent medians whereas dotted lines and small symbols describe quartiles across performances (N = 64). Pairs of coefficients of the same color concern a duet performance; the coefficient pairs of other color concern duets of the opposite role assignment. Interdependence between instrumentalists corresponds to a clear separation between a pair of coefficients; the directionality of the dependence always points from higher to lower value.

Conclusions

The results revealed characteristic patterns of interdependencies for different acoustic measures. Whereas bassoonists exhibited dependencies toward hornists for both timbral and dynamics adjustments, two parameters that seem interrelated (Lembke et al., 2017), the inverse dependency applied to the coordination of pitch or intonation. In addition, coordination of spectral centroid and sound level again became more apparent in the later phrase (Lembke et al., 2017). Interdependence across measures appeared limited to a relationship between instruments and not performer roles, which suggests that (within-subjects) differences in leadership may affect performances as a whole and not vary over time.

The use of interdependence measures to characterize musical performance are at an early stage of development; the current results should therefore be considered descriptive in nature. Suitable methods of statistical inference still need to be incorporated, which requires determining appropriate coefficient estimates for the null-hypothesis case (Papiotis et al., 2014). Also, these measures require further validation on controlled performances, for instance, ones including known one-way dependencies. Overall, such measures are valuable to elucidate interdependencies among even three or more musicians, characterizing how they adjust to each other's timbre, dynamics, and pitch and providing insight into which parameter may be of greater importance at a given time and for a particular musical context.

Acknowledgments

The analyzed data were acquired at the Music Perception and Cognition Lab (director: Stephen McAdams) and the Centre for Interdisciplinary Research on Music Media and Technology (CIRMMT), McGill University.

References

Barnett, L., & Seth, A. K. (2014). The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods*, *223*, 50–68.

Chicharro, D., & Andrzejak, R. G. (2009). Reliable detection of directional couplings using rank statistics. *Physical Review E*, *80*(2), 26217.

Lembke, S.-A., Levine, S., & McAdams, S. (2017). Blending between bassoon and horn players: An analysis of timbral adjustments during musical performance. *Music Perception*, *35*(2), 144–164.

Papiotis, P., Marchini, M., Perez-Carrillo, A., & Maestre, E. (2014). Measuring ensemble interdependence in a string quartet through analysis of multidimensional performance data. *Frontiers in Psychology*, *5*(September), 963.

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500.

Timbre, Texture, Space and Musical Style: The Interplay of Architecture and Music in Rome's Chiesa di Sant'Aniceto.

Jonathan Berger[1†], Talya Berger[2], Eoin Callery[1], Elliot Kermit-Canfield[1], Jonathan Abel[1]
[1] Center for Computer Research in Music and Acoustics, Stanford University, Stanford, California, USA
[2]Department of Music, Stanford University, Stanford, California, USA
[†] brg@ccrma.stanford.edu

Aims/goals

The spaces in which music is performed and experienced typically, and often profoundly, affect the resultant timbre. However, relatively little research exists on the considerable effect of architectural acoustics on musical timbre (Halmrast, 2000). Still fewer studies consider the relationship between musical style, performance practices, and architecture (Boren et al., 2013).

In this paper, we consider the role of architectural acoustics in the evolution of musical timbre and style by considering a particular church in Rome—the Chiesa Sant'Aniceto in Palazzo Altemps, an early seventeenth century church with rich and complex acoustical qualities—and the music written specifically for that church—a codex that has recently come to light after a long period of absence. Working with this unique corpus of source materials and computational models of the acoustics of that church, we demonstrate how the music's timbre and textural characteristics capitalize on the acoustical properties of the space (Huang et al., 2008). Using these analyses, we address the broader question of the interrelationship between compositional practice and architectural design beyond this single case study. Using an acoustic model of Chiesa dei Gesu in Rome, we demonstrate how distinctly different the music composed for Sant'Aniceto would sound in a larger structure with significantly longer decay times and resonant eigenfrequencies.
We conclude with hypotheses regarding timbre and texture as affected by the resonant and reverberant properties of the architectural space for which the music was meant to be performed.

Background information

The Chiesa di Sant'Aniceto in Rome's Palazzo Altemps was constructed in 1604 under the auspices of Duke Giovanni Angelo Altemps. Duke Altemps was a collector and composer of music. A large codex of works composed for Sant'Aniceto was compiled in 1604 with contributions by significant composers of the period, including Palestrina, Felice Anerio and others, as well as a few works by Duke Altemps himself. The codex was moved from the Altemps to the Colegio Romano, from where it was stolen and disappeared until 1996 when it was purchased by Museo Nazionale Romano and returned to the museum at Palazzo Altemps. The authors had the unique opportunity to perform acoustic tests in the church and to study the music contained in the codex. The existence of a body of works created specifically for performance in a church that was constructed in the same period provides an extraordinary opportunity to study the relationship between musical timbre and architecture. Furthermore, the architecture of Chiesa di Sant'Aniceto has undergone remarkably little change since its construction, thus affording a unique condition for studying timbre, texture and musical acoustics as they inform historical performance practices.

Methodology

Impulse responses of two churches in Rome, Chiesa di Sant'Aniceto and Chiesa dei Gesu, were measured using balloon pops (Abel et al., 2010) and sine sweeps as stimuli. Along with photogrammetric studies, the data was analyzed and used to create convolution reverberation models. A codex of choral and instrumental works composed specifically for performance in one of the modeled churches was studied and selected works were transcribed. A professional early music ensemble recorded the works in a near-anechoic space. The acoustic model of both churches were applied to the recordings. The resultant recordings were verified in subsequent experiments in the source spaces with both signal analysis and subjects' comparative evaluations. The musical scores were analyzed in terms of timbre and texture, as were the pre- and post-processed recordings.

Results

We note an apparent correspondence between the acoustical properties of Sant'Aniceto and aspects of tessitura, density, texture, timbre and rate of change in the music. In general, we propose that the church was created with particular awareness of the acoustical challenges of performing contrapuntal music in the style of the early Baroque period.

Conclusions

Reverberation is a fundamental aspect of music, with particular impact on timbre (Halmrast, 2000). Across time and cultures, music of sacred rituals has often been performed in spaces with significantly long decay times (Pentcheva et al., 2011). Some studies have proposed theoretical frameworks in which T30 exceeding ten seconds provides a highly reverberant feel that influenced melody and temporality of monophony and drone [7]. Here we explore the intentional attenuation of decay time in order to preserve the musical characteristics of polyphony that evolved in the 16th and 17th centuries. Musical timbre of this period is tightly bound with acoustical clarity which is a primary factor in architectural acoustics.

References

Halmrast, T (2000). Orchestral Timbre: Comb-Filter Coloration from Reflections. *Journal of Sound and Vibration*, 232(1), 53-69

Boren, B, et al. (2013). Acoustic Simulations of Renaissance Venetian Churches" Acoustics in Practice, 1(2), 17-27.

Huang, P. et al. (2008). Reverberation Echo Density Psychoacoustics". Audio Engineering Society Convention Paper.

Abel, J, et al. (2010). Estimating Room Impulse Responses from Recorded Balloon Pops. Audio Engineering Society 129th Convention.

Pentcheva, B. (2011). Hagia Sophia and Multisensory Aesthetics. *Gesta,* 50(2), 93-111.

Rupture of Timbre and Source-Bonding as Musical Aesthetic

James O'Callaghan
Independent Researcher, Montréal, Québec, Canada
jamesocallaghan3@gmail.com

Aims/goals

The proliferation of multimedia in contemporary music has opened the doors to the separation of timbre from physical source. Composers have capitalized on the ability to manipulate familiar timbres and disassociate them from their original sources, and this has become an important aspect of the aesthetic of much new music. Through a diverse survey of contemporary works, I aim to identify the manipulation of source-bonding (or association of a sound with a physical cause) as a significant compositional focus that contributes to a sophisticated conception of timbre that resists source-identification as part of the definition.

Background information

Traditional orchestration manuals and introductory texts often closely pair the concept of timbre with instrumentation. Definitions often appeal to identification of the source of a sound, describing timbre, for example, as something "from which [a sound's] source or manner of production can be inferred" (Dictionary.com). This identification, known as "source-bonding" (Smalley 1997), is not included in all definitions of timbre and some make efforts to explicitly separate it. In spite of such an exclusion, it is difficult to "parameterize" timbre, its meaning holistic enough to become a "multidimensional waste-basket category for everything that cannot be labeled pitch or loudness" (McAdams & Bregman 1979, p. 34). Composer François-Bernard Mâche suggests that "neutrality of [material becomes] fictive as soon as the choice of an instrument or a means of articulation is allowed to become more important than the choice of the notes" (Mâche 1983). In much contemporary music, efforts to assert or obscure sound's material origin have become central to musical discourse.

Methodology

The paper proceeds with a perceptual analysis of selected works to highlight different means of separating timbre from source-bonding in musical contexts. The analysis will be based on my own observations of the works, consulting their scores, programme notes, and surrounding theoretical texts, as well as auto-ethnographic analysis of my own compositional process.

Results

At least since the advent of electroacoustic music, composers have explored the disassociation of sound from physical source, most explicitly in acousmatic music, which is defined precisely by this condition (Schaeffer 1966, p. 91). While the concealment of sound-source is an important aesthetic value of much acousmatic music, the electronic medium nonetheless brings sound-source identification to the fore of the listening experience. While in traditional instrumental music "what is causing the sound" quickly becomes a redundant question, music in the age of mechanical reproduction allows it to become an enticing mystery. For example, the gradual revelation of sound source governs form in music by Parmerud and Gubaidulina.

Mixed music, involving conventional instruments and electronics, often appeals to an ambiguous blending of sound-source as an ideal, situating the electronics as an extension of instrumental timbres. The precondition of this ambiguity has been a source of inspiration in instrumental composition, as in the concealment of instrumental sources in works by Waller, or the reification of instrumental sound-production as musical discourse through so-called "extended techniques" in the work of Lachenmann, whose "musique concrète instrumentale" aims to reassert the significance of source-bonding in instrumental music upon the advent of it having been brought into relief through the contradistinction of Schaeffer's musique concrète (Ryan & Lachenmann 1999).

On the other hand, composers have repurposed musical instruments to connote sources other than themselves, especially through the imitation of environmental sound sources in Mâche's "phonography"

(Mâche 1983) and Einbond's "musique instrumentale concrète" (Einbond 2016). The contemporary technique of "mimetic instrumental resynthesis" (O'Callaghan 2015a) affords the use of spectrographic analysis and software-assisted orchestration to potentially create false source-bondings. Finally, composers of mixed music also are increasingly using novel spatialization strategies, especially by treating instruments as sites of electronic sound through the application of transducer speakers, to blur source identification and create ambiguous disconnects between physical sound production and sounding result. This is a significant element in music by Steen-Andersen, Alessandrini, Moroz, and myself.

Conclusions

A mutual concern for the materiality of sound, and the phenomenon of distinguishing timbre from physical source, runs as a current through practices in contemporary music that seem otherwise aesthetically distant from another. Electroacoustics, whether as an element of the music-as-heard, a part of the compositional process, or simply a source of inspiration, have been instrumental in the means that composers have re-imagined the role of the "material" in understanding timbre in music. The manipulation of timbres in this music can significantly influence source-bonding, such that the reduction of timbre to physical source becomes a meaningless metric.

References

Alessandrini, P. (2014). *Bodied Chambers*. Programme note. Retrieved from http://patriciaalessandrini.com/bodiedchambers

Benjamin, W. (1936/1969). The Work of Art in the Age of Mechanical Reproduction. In H. Arendt (Ed.) *Illuminations* (pp. 217-251). New York: Schocken.

Donin, N. (2015). Sonic Imprints: Instrumental Resynthesis in Contemporary Composition. In G. Borio (Ed.) *Musical Listening in the Age of Technological Reproducibility.* Farnham: Ashgate.

Einbond, A. (2016.) Musique instrumentale concrète: Timbral transcription in *What the Blind See* and *Without Words*. In J. Bresson (Ed.) *The OM Composer's Book, Vol. 3*. Paris: Delatour France.

Gubaidulina, S. (1993). *String Quartet No. 4*. Programme note. Retrieved from http://www.musicsalesclassical.com/composer/work/24115

Lachenmann, H. (1969). *Pression.* Wiesbaden/Leipzig/Paris: Breitkopf & Härtel.

Mâche, F. (1983/1992). *Music, Myth, and Nature: or, The Dolphins of Arion*. S. Delaney (Trans.) Switzerland: Harwood Academic Publishers.

McAdams, S., Bregman, A. (1979). Hearing Musical Streams. *Computer Music Journal,* 3(4).

McAdams, S. (1993). Recognition of Sound Sources and Events. In S. McAdams & E. Bigand (Eds.) *Thinking in Sound (*pp. 146-198). Oxford: Clarendon Press.

Nonnenmann, R. (2005). Music with Images—The Development of Helmut Lachenmann's Sound Composition Between Concretion and Transcendence. *Contemporary Music Review,* 24(1).

O'Callaghan, J. (2015a). Mimetic instrumental resynthesis. *Organised Sound,* 20(2).

O'Callaghan, J. (2015b). Resounding but not sounding: diffusing into interior spaces of instruments. *eContact!* 17(3). [http://econtact.ca/17_3/ocallaghan_resounding.html]

Parmerud, Å. (1992). *Les objets obscurs*. Programme note. Retrieved from https://www.electrocd.com/en/piste/ina_c_1030-4.2

Ryan, D., Lachenmann, H. (1999). Composer in Interview: Helmut Lachenmann. *Tempo*, 210.

Schaeffer, P. (1966). *Traité des Objets Musicaux.* Paris: Éditions de Seuil.

Smalley, D. (1997). Spectromorphology: Explaining Sound-Shapes. *Organised Sound,* 2(2).

Steen-Andersen, S. (2007). *Chambered Music*. Frederiksberg: Edition·S.

Timbre [Def. 1] (n.d.) Dictionary.com. Retrieved from http://www.dictionary.com/browse/timbre

Waller, J. F. (1997). *Eye-Masking.* Programme note. Retrieved from http://felipewaller.com/compositions/eye-masking

Arthur Lange and the Spectrotone System of Orchestration

Stephen A. Spencer
Music Department, The Graduate Center (CUNY), New York, NY, U.S.A
sspencer1@gradcenter.cuny.edu

Aims/Goals

As a popular music arranger, Arthur Lange (1889-1956) was well-known for his orchestral effects. He was particularly famous for his unique blends and stark contrasts, which he eventually systematized in his *Spectrotone System of Orchestration* (Lange, 1943), a complex technicolor chart of timbre correspondences. Through a detailed reading of the *Spectrotone System*, citing brief examples from all periods of Lange's career—from his early dance numbers and film scores to his late concert music—I aim to explore the ways in which Lange's orchestration theory schematizes the principles that account for the integration, segregation, and segmentation of auditory streams.

Background Information

There is a distinct lack of scholarship about Arthur Lange and his tone-colour chart, but primary sources by Lange himself are plentiful. His first book, *Arranging for the Modern Dance Orchestra* (Lange, 1926), shows him grappling with a systematic approach to instrumental blend and contrast. A series of newspaper articles from the 1920s and a collection of unpublished autobiographical material also suggest a career-long preoccupation with the systematization of timbre relationships, which he eventually realized in 1943 with the *Spectrotone* chart. Orchestration treatises that discuss tone-colour registers and provide examples of blend and contrast had existed long before Lange's chart (e.g. Berlioz & Strauss 1948; Gevaert, 1912; Koechlin, 1954; Rimsky-Korsakov, 1964), but the *Spectrotone System* is a radically different arrangement that relies on visual colour to facilitate an understanding of tone-colour relationships. Recent scholarship has attempted to substantiate and unify the underlying principles of orchestration using empirical means (Goodchild & McAdams, forthcoming; Lembke, 2015; McAdams & Goodchild, 2017); the *Spectrotone System*—owing to the succinctness of its layout, its unique colour-scheme, and its position within the history of popular music—would be fertile ground for further research of this kind.

Methodology

First, I provide a brief but thorough overview of Lange's *Spectrotone* chart, highlighting in particular his ten tone-colour categories, his distinction between *tone-colour* and *timbre,* and the relationship between the chart and more contemporary multidimensional models of timbre (Grey, 1977; Grey & Gordon, 1978; Iverson & Krumhansl, 1993; Wessel, 1979). Next, I discuss Lange's four categories of blend—*perfect*, *close*, *complementary*, and *remote*—and demonstrate how these relate to the principles of auditory scene analysis (McAdams & Bregman, 1979; Bregman, 1990) and other conceptions of timbral integration (Sandell, 1991). Lastly, I discuss several brief musical examples from Lange's career as a dance-band arranger, film orchestrator, and concert music composer, and use the *Spectrotone System* to illustrate visually some salient examples of blend and contrast.

Results

By examining the *Spectrotone System* and using it as an analytical tool, I reveal how Lange's visual abstraction of tone-colour facilitates a practical understanding of the principles of auditory scene analysis. I suggest that the chart can be enhanced by the inclusion of recent findings in music perception and cognition, but also demonstrate that future research in these fields could be enriched by the chart and its graphical representation of orchestrational principles.

Conclusion

This research points simultaneously toward the history of orchestration and to its possible future: an understanding of the *Spectrotone System* can tell us a great deal about the history of popular music arranging, the aesthetic standards of Hollywood orchestration, and the rise of systematic approaches to timbre during the Interwar period. Moreover, Lange's method of representing visually the principles of segmentation, integration and segregation points the way forward for orchestration pedagogy, orchestral analysis, computer-assisted orchestration, and composition.

References

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Berlioz, H. and Strauss, R. (1948). *Treatise on instrumentation.* (T. Front, Trans.). New York, NY: Kalmus.

Gevaert, F-A. (1906). *Nouveau traité d'instrumentation.* (E. Suddard, Trans.). Paris, France: H. Lemoine.

Goodchild, M. & McAdams, S. (forthcoming). Timbre perception and orchestration. In A. Rehding & E. Dolan (Eds.), *Oxford Handbook of Timbre*, New York, NY: Oxford University Press.

Grey, J.M. (1977). Multidimensional perceptual scaling of musical timbre. *Journal of the Acoustical Society of America, 61,* 1270-1277.

Grey, J.M. & Gordon, J.W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America, 63,* 1493-1500.

Iverson, P. & Krumhansl, C.L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America, 94,* 2595-2603

Koechlin, C. (1959). *Traité de l'orchestration: en quatre volumes.* Paris, France: M. Eschig.

Lange, A. (1926). *Arranging for the modern dance orchestra.* New York, NY: Robbins Music Corporation.

Lange, A. (1943). *Arthur Lange's Spectrotone system of orchestration: A color-graphic exposition of tone-color combinations and balance as practiced in modern orchestration.* Los Angeles, CA: Co-Art

Lembke, S-A. (2015). *When timbre blends musically: Perception and acoustics underlying orchestration and performance* (Doctoral dissertation). McGill University, Montreal, Canada.

McAdams, S. & Bregman, A. (1979). Hearing musical streams. *Computer Music Journal, 3*(4), 26-43.

McAdams, S. & Goodchild, M. (2017). Musical structure: Sound and timbre. In R. Ashley & R. Timmers (Eds.), *The Routledge Companion to Music Cognition*, pp. 129-139, New York, NY: Routledge.

Rimsky-Korsakov, N. (1964). *Principles of orchestration*. (E. Agate, Trans.). New York, NY: Dover.

Sandell, G. J. (1991). *Concurrent timbres in orchestration: a perceptual study of factors determining blend* (Doctoral dissertation). Northwestern University, Evanston, IL.

Locating Timbre: Challenges of Timbral Design in Multichannel Electroacoustic Music

Laurie Radford
Music Division, School of Creative and Performing Arts, University of Calgary, Calgary, Alberta, Canada
lradford@ucalgary.ca

Aims/goals

This paper summarizes the role of timbre in electroacoustic music and situates it within the ever-evolving spatial designs of this immersive sound art. The frontal and binaural facts of human listening are considered in the context of multichannel electroacoustic music where signals arrive at a listener's position at optimal and less than optimal angles. The ramifications of these discrepancies in localized aural apprehension for the compositional objectives of timbral design (the spectromorphological foundation of much electroacoustic music) are discussed. Changes to listening practices and listening entrainment that is particular to this sound discipline are also discussed and several methods for negotiating timbre and spatial design for timbral fidelity and predictability under multichannel audio conditions are reviewed.

Background information

The bio-neurological mechanisms employed by humans to locate sound in the world around them are well known: distance, interaural time (ITD) and amplitude differences (IAD), the perception of reflected sound, the engagement of memory, and the unconscious and conscious shifting of the head and thus the angle of audio reception to better localize particular sounds combine to afford reasonably precise localization of 3D audio sources under a variety of conditions. In addition, given the primarily frontal orientation in the listening act and the masking function of the head, the inevitable result of this system of localization is that the timbral characteristics of certain sounds are modified dependent on the source and on listener orientation. This situation presents challenges for accurately and predictably achieving some of the fundamental objectives of multichannel electroacoustic music: the creation and control of an immersive and animated audio stage in which the emanation of sound sources from any and all directions is compositionally meaningful (Normandeau, 2009; Kim-Boyle, 2008; Nyström, 2011), in conjunction with the refinement and fidelity of spectromorphological profiles that are fundamental to the language and syntax of the art form (Smalley, 1994; Smalley, 1997). There would appear to be a conflict between the desire for absolute timbral definition, often painstakingly achieved in the studio environment (Hong and Beilharz, 2007; Lynch and Sazdov, 2017), and the distribution of these sound materials on immersive audio systems (from point-source high-density loudspeaker arrays (HDLA) to higher-order ambisonic (HOA) and wave-field synthesis (WFS) systems) given the spectral modifications that are effectuated by the human localization system. Changes to listening practices and listening entrainment that have developed under increasingly refined acousmatic circumstances are also considered. Several methods for approaching a more satisfactory balance between timbre and spatial design are discussed and examined.

Methodology

The research was conducted through: 1) a review of previously published research in the area; 2) a consideration of scientific knowledge about hearing and sound perception in conjunction with compositional objectives and technical affordances of current multichannel audio systems; 3) outline of several proposed methods for negotiating timbre and spatial design in multichannel electroacoustic music.

Results

The research will result in a succinct overview of the relationship between timbral and multichannel spatial design in electroacoustic music, and a report on suggested methodologies for compositional control and negotiation between these two important parameters.

Conclusions

The conclusions of the research will address the consideration that methodological adjustments to timbral design are necessary in multichannel electroacoustic music composition in order to achieve and ensure a fundamental objective of the art form: spectromorphological (timbral) clarity and fidelity.

References

Einbond, Aaron & Schwarz, Diemo. Spatializing timbre with corpus-based concatenative synthesis. *Proceedings of the International Computer Music Conference* (ICMC) June 1–5, 2010, New York.

Kim-Boyle, David. Spectral Spatialization – An overview. *International Computer Music Association*, *Vol. 2008* (August).

McGee, R. Spatial modulation synthesis. In *Proceedings of the International Computer Music Conference 2015*, 246–249.

Lynch, Hugh & Sazdov, Robert. A perceptual investigation into spatialization techniques used in multichannel electroacoustic music for envelopment and engulfment. *Computer Music Journal, 41(1)*, pp. 13–33. MIT Press 2017.

Normandeau, Robert. Timbre spatialisation: The medium is the space. *Organised Sound*, *14(3)*, 277-285. Cambridge University Press, 2009

Nyström, Eric. Textons and the propagation of space in acousmatic music. *Organised* Sound, *16(1)*, 14–26. Cambridge University Press, 2011.

Ramakrishan, Chandrasekhar. Zirkonium: Non-invasive software for sound spatialisation. *Organised Sound, 14(3)*, 268-276. Cambridge University Press, 2009.

Hong Jun Song & Beilharz, Kirsty. Spatialization and timbre for effective auditory graphing. *Proceedingsof the 8th WSEAS International Conference on Acoustics & Music: Theory & Applications*, Vancouver, Canada, June 19-21, 2007.

Smalley, Denis. Defining timbre – refining timbre. *Contemporary Music Review*, *Jan 1994*.

Smalley, Denis. Spectromorphology: Explaining sound-shapes. *Organised Sound*, 2(2), 107-126. Cambridge University Press, 1997.

Smalley, Denis. Space-form and the acousmatic image. *Organised Sound*, 12(1), 35-58. Cambridge University Press, 2007.

Open-Source Modular Toolbox for Computer-Aided Orchestration

Carmine Emanuele Cella† and Philippe Esling
IRCAM CNRS – UMR 9912, 1 Place Igor Stravinsky, F-75004 Paris, France
† cella@ircam.fr

Aims/goals

Computer-aided orchestration is defined as the search for relevant sound combinations that could match a given sound target. Several computational approaches for musical orchestration have been proposed in the literature. However, most approaches were bounded to *time-blind* features, by the use of averaged descriptors over the whole spectrum, which do not account for the temporal structures that are crucial to musical timbre perception. Furthermore, most research proposed in the past was related to proprietary source code, which hampers the development of a research community around the topic of orchestration. In this paper, we introduce a novel open-source and modular toolbox for computer-aided orchestration. We rely on an evolutionary orchestration algorithm that allows a constrained multiobjective search for a target timbre, in which several perceptual dimensions are jointly optimized while accounting for their micro-temporal structures. We introduce the overall modular architecture of the resulting toolbox, which allows knowledge extensibility, multiple and abstract targets and efficient time series querying.

Background information

Tackling the problem of orchestration from a scientific standpoint involves the use of empirical knowledge and unveils numerous facets of complexity. Current knowledge of orchestration is summarized in orchestration treatises that mostly collect examples across the classical repertoire. The challenges of orchestration can be organised into three major areas of complexity: *combinatorial*, *temporal,* and *timbral* complexities. In this context, orchestration consists in finding a mixture of instrument samples that jointly minimizes different perceptual distances to a given target. (Psenicka, 2003) first developed a system called SPORCH (SPectral ORCHestration) where the search is based on an iterative matching pursuit algorithm. Later, (Rose, 2009) proposed a tool to analyze existing orchestrations to create new ones, by modelling the instrumental knowledge with the average harmonic spectra. Finally, an interesting approach was proposed by (Carpentier, 2008). Although the formulation of the problem remains the same, the system is based on multiobjective genetic exploration, which can propose a *set* of optimal solutions, rather than a single instrumental combination. All the systems described above are inherently limited by relying on *time-blind* static features, making them therefore unable to cope with time-evolving sounds. This only addresses static timbres, which is in contradiction with the basic cognitive principles of music perception. Recently, we introduced a novel approach with an instrumental model representing both temporal and spectral properties (Esling, 2012). This model allows capturing relevant properties of instrumental timbres while handling their temporal structures. The system architecture has been designed to fully separate instrumental knowledge from search algorithms, working with virtually any kind of timbre modeling. Hence, this approach addresses explicitly the three different facets of complexity (timbre multidimensionality, combinatorial explosion of instrument mixtures and micro-temporal structures variability) by the joint use of a multicriterion optimization and evolutionary algorithms augmented with efficient time series processing techniques.

Methodology

Our approach is illustrated in Figure 1 (left). The goal is to find the mixtures of instruments that best match a given timbre, by relying on a set of pre-computed instrumental features. The user specifies the instruments he would like to use (constraints) and the characteristics of the sound to be produced (target) through a sound file which defines the audio features to reproduce. Then, an orchestration engine searches among the instrumental features database to suggest instrumental combinations that should sound close to the target. The problem is, therefore, to find an efficient way to converge towards closely sounding elements (circumventing the underlying combinatorial complexity of the problem). In order to reach a wider community, we decided to implement the proposed approach by creating an open-source library in C++. The main aim of this toolbox is to deal with all previously underlined complexities inherent to assisted orchestration and leave the users able to focus on specific orchestration requests. The library we introduce here, called *liborchidea*, is based on the existing software *Orchids* (Esling, 2012). One of the advantages of this new implementation is its modularity. The toolbox is made of modules that deal with separate problems of analysis, search, representation, synthesis and so on, as shown in Figure 1 (right).

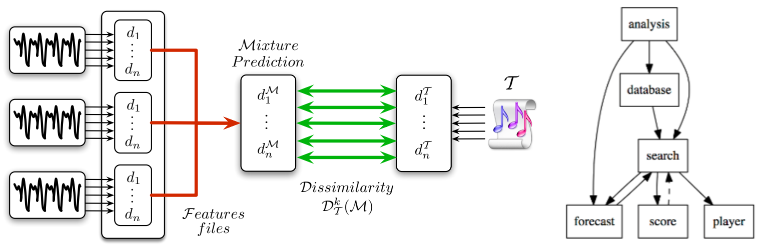


Figure 3. (Left) Reifying musical orchestration by using a target sound, which should be reconstructed by using a set of features from pre-analyzed instruments. (Right) Structure of the open-source modular liborchidea toolbox proposed in this paper.

A major outcome of this layout is the possibility of creating specific workflows for orchestration adapted to the needs of each user. Moreover, an additional benefit of this architecture is that the library can become the core of different applications such as components of larger systems or standalone software.

Results

The current implementation has been tested against old implementations and independently. The testing frameworks have been implemented in Matlab and are available with the library.
Analysis accuracy. Extensive testing has been conducted on the features computed by our version, which shows a very high level of accuracy as compared to previous versions.
Forecast accuracy. During the search, our algorithm builds forecast features of the instrumental mixtures generated (predicting their spectral features). We performed extensive testing of these forecasts with real features computed from the mixture waveforms and obtain satisfactory results.
Retrieval of known mixtures: in order to verify the reliability of our approach, we generated known mixtures of sounds present inside the instrumental dataset and used them as target. The goal is to assess the ability of the algorithm to retrieve instrumental mixtures of increasing complexity.
The preliminary test phase has been encouraging, providing good accuracy for all three categories. Moreover, the library provides specific use cases for separate modules of analysis, forecast and search. Hence, users can combine the components in the way that fits best their needs for assisted orchestration.

Conclusions

In this paper, we introduced our approach for solving assisted orchestration queries. This entails the creation of a novel C++ open-source library called *liborchidea* aimed at this purpose. The library builds on previous implementations of the described approach (Esling, 2012) and has been extensively tested. It is currently freely available for interested researchers and composers. This library will be the core of several projects around assisted orchestration in the next future. We believe that its modular approach will let researchers advance in this subject and will permit significant improvements in the field.

Acknowledgments

This work benefited from the funding of the MAKIMOno 17-CE38-0015-01 project of the French ANR.

References

Carpentier, G. (2008). Approche computationnelle de l'orchestration musicale. PhD thesis, UPMC-P6.
Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing Surveys (CSUR)*, *45*(1), 12.
Esling, P. (2012). Multiobjective time series matching and classification. PhD thesis, UPMC-P6.
Psenicka, D. (2003). SPORCH: An Algorithm for Orchestration Based on Spectral Analyses. In *ICMC*.
Rose, F., & Hetrick, J. E. (2009). Enhancing orchestration technique via spectrally based linear algebra methods. *Computer Music Journal*, *33*(1), 32-41.

Generating Orchestral Music by Conditioning SampleRNN

Léopold Crestel†, Philippe Esling, Daniele Ghisi and Robin Meier
IRCAM CNRS – UMR 9912, 1 Place Igor Stravinsky, F-75004 Paris, France
† leopold.crestel@ircam.fr

Aims/goals

Recent breakthroughs in statistical learning have opened up the possibility to generate audio signals by learning directly from datasets of complex waveforms. However, most of the studies in this field analyse the success of these systems to generate speech signals or single instrument music such as piano recordings (Mehri, 2016; Oord, 2016). Here, our objective is to design an audio waveform synthesis system that is able to produce orchestral music by allowing the model to extrapolate from orchestral pieces observed in a large dataset. Furthermore, whereas most of the previous learning-based approaches do not allow for user control at generation time, we implement a mechanism to address this deficit. First, we introduce a conditioning mechanism to the recently proposed SampleRNN (Mehri, 2016) and evaluate the success of our model to generate different types of single instruments. We show how our conditioning can allow for switching between different instruments. Then, we evaluate the same system on its capacity to generate signals for full orchestral music by training the model on a large set of orchestral recordings. Finally, we show that our conditioning mechanism can be used as a new sound synthesis method that can allow for stimulating creative possibilities.

Background information

The question of generating musical signals has been extensively studied over the past decades. Most of the previous approaches on this topic were defined in the spectral domain as a complex set of relatively simple manipulation (subtractive, additive or modulation synthesis). However, several recent breakthroughs in audio waveform generative models based on neural networks (Oord, 2016) have obtained astounding results in speech synthesis both qualitatively and quantitatively (Mehri, 2016). These systems rely on learning the structure of audio waveforms directly from a given set of audio files, without computing a spectral transform and in an unsupervised manner. Furthermore, whereas the modelling capacity of most audio synthesis systems is limited to micro-temporal evolutions, SampleRNN is able to generate sequences with macro-temporal structure. Because statistical learning essentially combines the properties of existing audio files, it can be compared to some extent to granular synthesis. However, granular synthesis cannot extrapolate from observed samples and is limited to their concatenation.

Both WaveNet (Oord, 2016) and SampleRNN (Mehri, 2016) were essentially designed and optimised to generate speech. Hence, when applied to music, their results are less satisfying. Furthermore, because the range of sounds generated by the model is strongly determined by the type of music contained in the training set, we observed that if there is no consistency in term of style or timbre between the training files, the model may fail to grasp the waveform structure and produce noisy sounds. However, we believe that maintaining a heterogeneous database could lead to a better model, as the total number of seen examples is increased. We believe that if we manage to train the model with a segmented database while being able to condition the model over given labels, we could obtain both a more robust and versatile model along with a mechanism for generating cross-stylistic examples. Here, we implement a mechanism to address this issue by conditioning SampleRNN and show how our conditioning can switch between different styles of instruments and even to generate signals for full orchestral music.

Methodology

Database. In order to train our models, we rely on a dataset of several different orchestral recordings to provide sufficient diversity. This dataset is composed of 1502 recordings taken from 57 composers. Regarding the instruments, we use a set of 282 recordings that are pieces written for 11 different instruments. Both datasets correspond to 364.2 hours of raw audio signals.

Training. Learning a model to generate audio waveforms can be done by training it on a *sample-by-sample* predictive task. Given a waveform segment between time indices $t - N$ and t, the model tries to predict the

sample at time $t + 1$. This predicted sample is compared with the original one through a negative log-likelihood loss, and gradient descent over the model's parameters is performed.
Architecture. Our contribution is to extend the SampleRNN model with a conditioning mechanism, which conditions the generation process with high-level label information. This mechanism is implemented by multiplying the weights of the original SampleRNN network by a supervised value depending on a given set of labels (Taylor, 2009). In practice, this labelling mechanism consists of splitting the dataset in disjoint subsets and assigning a label to each of this subset. The subsets can be either explicitly created by the user by organising the files in several folders, or automatically extracted from the files (splitting the database between low and high brilliance, or any other computable feature). At generation time, the user can define a time series as labelling values which will shape the evolution of the generation process. Here, we target orchestral music as the most fertile framework for our model and provide the possibility to create subsets with the specifics sonic identities of composers (used as labels).
The qualitative performances of the model are assessed using the same negative log-likelihood loss function as for the training criterion on a one-step prediction. While this criterion gives no guarantee over the long-range behaviour of the generated waveform, it is sufficient for comparing and discarding models. To assess the model's ability to generate music with long-range structures, we relied on several human listening aimed at qualitatively evaluating the capacity of a model, where authors and external participants of the laboratory ranked the models according to their sound quality, creative interest.

Results

Generating multiple instruments. We can produce interesting acoustic generation for single instruments separately and assess the capacity of the model to generate various orchestral instruments. The model trained with our proposed conditioning on a dataset is able to successfully switch between different instruments at generation time given the conditioning value. However, artefacts are still present at the switching point which shows instability that should be addressed with more complex datasets.
Generating full orchestral music. The model trained on full orchestral music is able to generate interesting orchestral textures with long-term structures that denote an understanding of macro-temporal evolutions inside orchestral music. Although the overall audio quality is good, it must be noted that large segments of the generation are textures rather than clear-cut notes and sequences, which seems to be coming from the system trying to oscillate between large ranges of possibilities and producing a mean.
Providing control over complex orchestral music. The proposed conditioning allows to alleviate the previously mentioned texture effects. However, the generated music can still not be clearly tied to the specific composer. However, the user control steers the generation in a novel and creative way.

Conclusions

We proposed a model for generating raw audio signals for full orchestral music. Based on SampleRNN, our model implements a label conditioning mechanism which offers the possibility for the model to efficiently learn on larger databases (using explicit structuring information of the dataset), and offers user control at generation time (by modifying the value of the label along time). While the quantitative evaluation mainly provided us with a selection criterion, the qualitative analysis showed that our model is able to generate creative stimulating sounds, with a good audio quality.

Acknowledgments

This work benefited from the funding of the MAKIMOno 17-CE38-0015-01 project of the French ANR.

References

Mehri, S., Kumar, K., Gulrajani, I., Kumar, R., Jain, S., Sotelo, J., ... & Bengio, Y. (2016). SampleRNN: An unconditional end-to-end neural audio generation model. arXiv preprint arXiv:1612.07837.

Oord, A. V. D., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., ... & Kavukcuoglu, K. (2016). Wavenet: A generative model for raw audio. arXiv preprint arXiv:1609.03499.

Taylor, G. W. and Hinton, G. E. (2009). Factored conditional restricted boltzmann machines for modeling motion style. In Proceedings of the 26th ICML, pp. 1025–1032. ACM.

Learning Spectral Transforms to Improve Timbre Analysis

Tristan Carsault†, Axel Chemla--Romeu-Santos and Philippe Esling
IRCAM CNRS – UMR 9912, 1 Place Igor Stravinsky, F-75004 Paris, France
† tristan.carsault@ircam.fr

Aims/goals

Here, we argue that one of the key aspects for understanding timbre derives directly from the spectral transforms used to analyse its characteristics. Almost all timbre studies rely on the computation of a Fourier transform. However, the success of further analyses is highly dependent on the type of spectral representation. Here, we show that transforms can be learned depending on the specific timbre of each instrument by relying on recent advances in *statistical learning*. We propose different approaches to perform relevant feature extraction by relying directly on the signal waveform information. The transforms are learned through an instrument-specific objective (presence/absence in an acoustical mixture) adapting the computation of the transform to the timbre of each instrument. We compare the various transforms obtained and evaluate their properties as a novel proxy to timbre analysis.

Background information

The complexity of musical signal results from the acoustical mixture of different instruments having their own timbre characteristics. Almost all works in timbre analysis research relies on a time-frequency representation (Burgoyne, 2007) of the sound. However, the Fourier Transform might not always be the optimal representation for all analyses. This remark appears more clearly in other research fields, such as machine learning, where classification problems use preferentially the Constant-Q transforms (CQT).

Recently, several breakthroughs in *unsupervised generative models* based on neural networks (Oord, 2016) have obtained astounding results in speech synthesis both qualitatively and quantitatively (Mehri, 2016), while learning directly from the raw audio waveform information. The large dimensionality of the inputs is handled by a hierarchical organization within the network. In the case of *WaveNet* (Oord, 2016), the complexity of the waveform signal is handled by first computing an organized hierarchy of *dilated convolutions* at different scales. Regarding the *SampleRNN* model (Mehri, 2016), the idea is to use as a first layer an auto-regressive process to address the specificity of waveforms and, then, to learn on the coefficients of this process. Both models are able to learn an intermediary representation that generates speech that is perceptually very close to human production.

Here, we hypothesize that if those models are able to obtain such results, there might inherently learn a more adapted transform as an intermediary step of the learning process. Therefore, we introduce a model that is comprised of two complementary learning systems. The first part trains a deep dilated convolution network akin to *WaveNet* to learn a given spectral transforms in a supervised way. The second part then builds a classifier on top of this learned transform and learns a supplementary supervised instrument recognition task. However, the computation of the original transform (first part) is still allowed to be modified. Hence, we aim to make the system learn its own spectral representation that is built specifically for a given task and the different timbre of the inputs. By analysing how the original transform layers are adapted depending on different orchestral instruments, we study how we could use the transform as a proxy for timbre analysis. To do so, we require the model that we learned to modify its weights towards an extremely specific transform by learning the instrument recognition task with a *very low capacity* classifier above the transform network. Therefore, we ensure that the discrimination has to be made by the spectral transform part rather than by the classifier layer on top. We study how the original transform layers are adapted depending on different orchestral contents by analysing how the representation accomplishes this particular task. Finally, we study how each transform could impact further timbre analysis by comparing the saliency and weights of the different learned transforms.

Methodology

We rely on *Studio On Line* (SOL) to obtain a dataset for 12 instruments (Piano, Flute, Clarinet, Trombone, English-Horn, French-Horn, Oboe, Saxophone, Trumpet, Violin, Violoncello, Bassoon), with 10 playing styles for each. We use samples annotated with the same intensity (to remove effects from the pitch and loudness and incentivize the networks to focus on the timbre). Then, we extract all 12 ms windows of signals with a Blackman window from each sample. For each window, we compute the CQT that will be used as a target for the pre-training of the networks. We normalize the transforms so that they have zero mean and unit variance.

Regarding the models, we use a specifically designed architecture. First, we use two layers of hierarchical dilated convolutions as defined in the WaveNet system (Oord, 2016). Then, 2 layers of convolutions are used to increase the capacity of the model and reduce the dimensionality of the data. Finally, we add one fully-connected layer, to map to the dimensionality of the CQT.

As our goal is to learn more adapted spectral transforms, we start by learning a common model on the whole dataset of orchestral instrument samples. This model learns by providing the CQT as the supervised target that the network must approximate. The model learns with the SGD optimization algorithm with a learning rate $\eta = 10^{-3}$. Once the model converges, we add the low-capacity classifier composed by a single fully-connected layer in order to learn the instrument discrimination task.

For these complete models, we learn one different supervised task for each instrument, in order to study how the transform learned by the first layers may vary. To do so, we create mixtures with potential appearance of a given instrument to obtain a labelled dataset that associates every frame with the presence of specific instruments. Thus, the whole network must learn to identify the timbre of that instrument.

Results

Comparing different transforms and their weightings. After learning to approximate the CQT, the global network converges to a very high accuracy (mean L_2 error of 0.071), which confirms that we are able to learn the transform. After training on instrument recognition, the very high variance of the weights and final representation obtained confirms that the transform adapts to the timbre of each instrument.

Comparing the efficiency in musical instrument recognition. By applying the transform learned for one instrument to recognize another, we obtain very significant drops in accuracy that depends on the classified instrument (e.g. $\Delta acc(Bassoon) = 11 \pm 2\ \%; \Delta acc(Violin) = 32 \pm 1\ \%$), which confirms that the transforms learned are highly specific to each instrument.

Using transforms as a proxy to timbre analysis. When using the saliency to determine the specificities of each transform, the variability and complexity of the results precluded a straightforward result. However, specific tools should be developed for using these transforms as a new way of performing timbre analysis.

Conclusions

We have shown that we could learn spectral transforms that are specific to the timbre of each orchestral instrument. Various analyses confirmed that these were highly variable and specific, which indicates that they have to adapt to instrumental timbre (shown through variance analysis and classification accuracies). However, the very high variability and complexity of the first layers precluded a direct understanding of their inner mechanisms, calling for the development of more specific analysis tools. These timbre-specific transforms still open very intriguing avenues of research as a novel proxy for timbre understanding.

Acknowledgments

This work benefited from funding of the MAKIMOno 17-CE38-0015-01 project of the French ANR.

References

Burgoyne, J. A., & McAdams, S. (2007). A meta-analysis of timbre perception using nonlinear extensions to CLASCAL. In *International Symposium on Computer Music* (pp. 181-202). Springer, Berlin.

Mehri, S., Kumar, K., Gulrajani, I., Kumar, R., Jain, S., Sotelo, J., ... & Bengio, Y. (2016). SampleRNN: An unconditional end-to-end neural audio generation model. *arXiv preprint arXiv:1612.07837*.

Oord, A. V. D., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., ... & Kavukcuoglu, K. (2016). Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*.

Acoustical Analyses of Extended Playing Techniques in *Pression* by Helmut Lachenmann

François-Xavier Féron[1,3†], Catherine Guastavino[2,3], Benjamin Carat[4]
[1] CNRS, LaBRI-SCRIME UMR 5800, Université de Bordeaux, France
[2] School of Information Studies and Schulich School of Music, McGill University, Montreal, Canada
[3] Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Canada
[4] Conservatoire à Rayonnement Régional de Bordeaux, PESMD Bordeaux Aquitaine
[†] francois-xavier.feron@u-bordeaux.fr

Aims/goals

Since the beginning of the 20th century, composers have relentlessly been developing new instrumental techniques in order to reveal and explore the wide timbral palette of acoustical instruments. Articulating a musical grammar based on the production of new sounds is an essential aspect of Helmut Lachenmann's musical language. In *Pression* (1969-70/2010) for solo cellist, the composer radically redefined the way of playing the cello, paving the way for what he called "instrumental *musique concrete.*" The scoring systems used in this score aim to indicate the actions the cellist has to perform onto various parts of the cello (strings, soundboard, bridge, and tailpiece, as well as the bow). Despite this unusual notational system and extended playing techniques, this piece has become an integral part of solo cello repertoire. The first version of the score written in 1969-70 and published by Hans Gerig in 1972 contains some textual explanations (in German) and some graphical representations that are not very precise. At the time, performers worked in close collaboration with the composer who would explain precisely how to perform these new playing techniques and the resulting sounds. Forty years later, stimulated and helped by cellist Lucas Fels, Lachenmann produced a new bilingual version of the score, much more intelligible and precise (Féron, 2015). But his notational system still emphasises the performers' gestures rather than the resulting sounds. What are the timbral specificities of sounds produced by these different playing techniques? Could acoustical analyses be useful for performers?

Background information

Pression has received research attention from different perspectives. The structure of the piece was analyzed by Jahn (1988) and Kaltenecker (2001) in terms of actions carried out by the performer. Mosh (2006) compared different audio recording to discuss variations in performance. Orning (2012) discussed her experience of the piece as a performer with the two versions of the score. Other performance studies were recently conducted using motion capture at the Institute for Computer Music and Sound Technology of the Zurich University of the Arts (Schacher et al., 2016). Because of its crucial visual dimension, *Pression* has often been video recorded and many performances are available on internet video platforms or DVDs. But the singularity of actions asked by the composer might make it difficult for performers to assimilate all the extended playing. In a DVD-rom recently produced by Hermann and Walczak (2013), all of Lachenmann's extended playing techniques are grouped by instrumental family and described via textual explanation, score excerpts and video examples. As a logical complement, we recorded individually all *Pression*'s playing techniques and characterized them through acoustical analyses in order to describe the sound results and reveal their main characteristics.

Methodology

On December 5, 2017, a recording session was organized at the SCRIME (Studio de Création et de Recherche en Informatique et Musiques Expérimentales) in Bordeaux (France) with the cellist Benjamin Carat. He was asked to play individually each playing technique, recorded simultaneously with a stereo pair of SCHOEPS CMC6--U (at one meter at ear height), a contact microphone Schertler BASIK attached near the tuning pegs, and a AKG C414B close the cello at the end of the tailpiece. Two full HD cameras were used for the video recording. The first (Sony HDCAM HDW750P) was positioned in front of the cellist giving an overview of him. The second (Sony XDCAM PMWEX1) serves to zoom on his hands. Acoustic analyses were conducted from the mono signal of AKG C414B. We resorted to Audiosculpt software for spectrograms.

Results

We illustrate the analysis with an example shown in Figure 1. In this score excerpt, the performer is holding the bow with his two hands and rubbing the strings vertically, moving toward of the scroll (descending line) or toward the bridge (ascending line). According to the score notice, “A dry, brilliant, quasi-perforated rattling sound will result.” The left hand grabs the hair and changes position in order to modify the length of the hair. This results in a pitch variation inside the noisy rattling sound as shown in the spectrograms [0-2800 Hz]. The left pane describes to a vertical rubbing on string III with a 1/4 hair length: the fundamental frequency around 155 Hz corresponds to the open string (Db3) and stable formants appear around 1000 Hz. The right pane describes to the same gesture but with a hair length varying from ¼ to ½: the formant gradually goes down while the fundamental frequency remains stable.

Figure. 1 – Impact of the hair bow’s length on “perforated rattling” sound.

Conclusions

With *Pression*, Lachenmann builds the foundation for “intrumental musique concrète,,” which consists in exploring the sound potential of instruments by disregarding well-established aesthetic canons. This masterpiece reveals a new sound vocabulary where clear pitches are almost completely absent, which calls for a study of timbre, as emancipated from pitch. Spectrogram analyses provide new insights for characterizing the subtleties of this new vocabulary. Additional analysis will be conducted using the Matlab Timbre Toolbox for audio descriptors (Siedenburg *et al.* 2016). Documenting the various timbres in that way can provide guidance for performers trying to acquire the corresponding playing techniques.

Acknowledgments

This research was begun as a part of the ANR GEMME (2013-2017) coordinated by Nicolas Donin at Ircam and was pursued at LaBRI-SCRIME. We are very grateful to Thibaud Keller and Christian Faurens who were respectively in charge of audio and video recordings.

References

Féron, F-X. (2015). Enjeux et évolution du système de notation dans *Pression* (1969-70/2010) pour un(e) violoncelliste de Helmut Lachenmann, *Circuit – Musiques contemporaines*, 25 (1), 55-66.

Hermann, M. & Walczak, M. (2013). Erweiterte Spieltechniken in der Musik von Helmut Lachenmann / Extended Techniques in the Music of Helmut Lachenmann, CD-ROM Breitkopf & Härtel,

Jahn, H-P. (1988). *Pression*. Einige Bemerkungen zur Komposition Helmut Lachenmanns und zu den interpretationstechnischen Bedingungen, Musik-Konzepte, 61-62 ‘Helmut Lachenmann’, 40-61.

Kaltenecker, M. (2001). Avec Helmut Lachenmann, Paris, Edition Van Dieren.

Mosh, U. (2006). Das Unberührte berühren – Anmerkungen zur Interpretation von Helmut Lachenmanns Werken Pression und Allegro sostenuto, in Jörn Peter Hiekel (ed.), *Musik inszeniert. Präsentation und Vermittlung zeitgenössischer Musik heute, Veröffentlichungen des Instituts für Neue Musik und Musikerziehung Darmstadt*, (46), Mainz, Schott, .25-46.

Orning, T. (2012). *Pression* – a performance study, Music Performance Research, (5), 12-31.

Schacher, J-C., Bisig, D. & Neff, P. (2016). Exploring Gesturality in Music Performance, Proceedings of the 13th Sound and Music Computing, 17.

Siedenburg, K., Fujinaga, I., & McAdams, S. (2016). A Comparison of Approaches to Timbre Descriptors in Music Information Retrieval and Music Psychology, *Journal of New Music Research*, 45(1), 27-41.

Studying Timbral Structures in 19th-Century Piano Music: Overview and Challenges

Nathalie Hérold
Labex GRÉAM, University of Strasbourg, Strasbourg, France
nathalieherold@hotmail.com

Aims/goals

My general research project aims at understanding the structural role of timbre in 19th- and 20th-century western art music. It takes the point of view of analytical and theoretical musicology, including some necessary incursions into the domains of acoustics, psychology, signal processing, and also musical practice. To what extent and under what conditions can timbre be considered as a form-bearing dimension within musical discourse? With a particular focus on piano music of the first half of the 19th century, this contribution presents an overview of the main issues, and a synthesis of the key challenges pertaining to this question that has been so far too little discussed by analytical and theoretical musicology.

Background information

The consideration of timbre in the frame of music analysis and theory has not been largely developed yet, contrary to other musical parameters such as pitch – and, to a certain extent, duration – which have given rise to a great number of analytical theories (Schenkerian, Neo-Riemannian, Generative Theory of Tonal Music, etc.). In spite of the invaluable contributions of authors such as Erickson (1975), Cogan (1984), Lerdahl (1987), or Delalande (2001), neither a musical and explicit theory of timbre, nor a real "musicology of timbre" (McAdams 2015), has been developed. Moreover, the analytical and theoretical discourse on timbre has highly focused on 20th-century music (Barrière 1991), in particular on music after Debussy (Guigue 2009). As regards repertoires before Debussy, a few authors such as Dolan (2013) and Hérold (2011, 2014, 2015) nevertheless brought some insight on timbre in respectively Haydn's orchestral music and piano music of the first half of the 19th century.

Timbre is also as a relatively problematic – but no less promising – concept within musicology, and particularly music analysis and theory, due to its different kinds of terminological uses. The distinction between "identity timbre" and "qualitative timbre" (Castellengo 2015) underlines the basic distinction between, on the one hand, timbre in the sense of an instrumental differentiation and, on the other, a global timbre that can be produced by a group of instruments, as well as the different timbres that can be produced by one single instrument—for example the piano, which can give rise to a specific pianistic "orchestration." Following Hugues Dufourt, timbre can be considered not only as a multidimensional parameter, but also as a musical "metaparameter" that includes all musical parameters.

Methodology

In music analysis and theory, and particularly as regards timbre, methodological issues are an important matter of discussion, namely because methods have to be adapted as much as possible to the musical specificities of the works under study. Therefore, my contribution discusses, on the one hand, some epistemological aspects concerning the different ways it is possible to approach timbre in piano repertoires from the past. The relevancy of various types of sources (published scores, sketches, commercial and experimental recordings, period and contemporary instruments, discourses of musicians, etc.) is examined, as well as the conditions for pursuing an empirical research on timbre in collaboration with musicians (Clarke and Cook 2004).

My contribution presents, on the other hand, some propositions for a timbre analysis method, with applications to piano pieces by composers such as Beethoven, Chopin, Schumann, and Liszt. This method of timbre analysis is based on both score and sound analysis. As regards score analysis, several textural and pianistic timbral factors are examined, among others spacing between simultaneous notes, horizontal density, pedals, registers, articulations, dynamics, etc. These timbral factors lead to the identification of different degrees of fused and segregated timbral configurations from the score. As regards sound analysis, timbre is approached through computational tools such as sonograms and audio features—namely the spectral centroid, particularly relevant for the study of timbral brightness, as well as audio features integrating perceptive models. These tools are used in close relation to musical listening—as allowed by software like *Acousmographe*, *AudioSculpt*, or *Sonic Visualiser*—and are particularly useful to examine large-scale timbral configurations on the level of an entire piano piece.

One main challenge of this type of analytical method is to highlight constants of musical timbre that are robust enough to withstand performance changes and can thus be considered as characteristic of the timbral composition of a musical piece.

Results

The results take the form of annotated scores and sonograms, graphs, and formal diagrams, and concern short-scale as well as large-scale musical structures. New structures of musical pieces—in terms of segmentation, articulation points, climaxes, and cadences—can be highlighted on the basis of timbre analysis. Theoretically, these structures can be formalized as strings of concatenated timbral units on different formal levels and give rise to timbral hierarchies and spaces within a given piano piece. These timbral structures are likely to converge or diverge with the musical structures deduced from pitch-based analytical theories, and it is thus essential to discuss them in light of their musical relevancy—including the consideration of compositional and listening processes.

Conclusions

On the basis of one specific corpus centered on 19th-century piano music, this contribution gives an overview on how to study timbral structures, and discusses the main challenges posed by timbre analysis. It draws some general elements for a theory of timbre to be applied to the analysis of the piano repertoire of the first half of the 19th century and, by highlighting the function of timbre in the evolution of musical forms, it suggests reconsidering the theory of musical forms from the late-classical and early-romantic period. Open to interdisciplinary dialog and to collaborations with musicians, it raises many scientific and artistic challenges, due to the particular situation of timbre at the very heart of the musical phenomenon.

Acknowledgments

This research project is supported by a 10-year grant allocated by the ANR (French National Agency for Research) to the Excellence Laboratory GRÉAM (Experimental Research Group on Music in Act).

References

Barrière, J.-B. (Ed.) (1991). *Le timbre, métaphore pour la composition*. Paris: IRCAM-Bourgois.

Castellengo, M. (2015). *Écoute musicale et acoustique : Avec 420 sons et leurs sonagrammes décryptés*. Paris: Eyrolles.

Clarke, E., and Cook, N. (Eds.). (2004). *Empirical musicology: Aims, methods, prospects*. New York: Oxford University Press.

Cogan, R. (1984). *New Images of Musical Sound*. Cambridge, MA: Publication Contact International.

Delalande, F. (2001). *Le son des musiques : Entre technologie et esthétique*. Paris: INA-Buchet-Chastel.

Dolan, E. (2013). *The Orchestral Revolution: Haydn and the Technologies of Timbre*. Cambridge: Cambridge University Press.

Erickson, R. (1975). *Sound Structure in Music*. Berkeley: University of California Press.

Guigue, D. (2009). *Esthétique de la sonorité : L'héritage debussyste dans la musique pour piano du XX^e^ siècle*. Paris: L'Harmattan.

Hérold, N. (2011). *Timbre et forme : La dimension timbrique de la forme dans la musique pour piano de la première moitié du dix-neuvième siècle*. Doctoral dissertation, University of Strasbourg.

Hérold, N. (2014). Timbre, melody, and spacing in Chopin's Nocturne Op. 9 n° 2: From modeling to performance. *Sonus*, *35*(1), 1-16.

Hérold, N. (2015). La modélisation du timbre et de son organisation formelle dans les œuvres pour piano de la première moitié du dix-neuvième siècle. In X. Hascher, M. Ayari, and J.-M. Bardez (Eds.), *L'analyse musicale aujourd'hui/Music Analysis Today* (pp. 93-110). Sampzon: Delatour.

Lerdahl, F. (1987). Timbre hierarchies. *Contemporary Music Review*, *2*(1), 135-160.

McAdams, S. (2015). *Perception et cognition de la musique*. Paris: Vrin.

New Perspectives in Orchestration Pedagogy

Victor Cordero[1†], Kit Soden[1,2,3], Eric Daubresse[1,4]
[1] Music, Haute école de musique de Genève/Neuchâtel, Geneva, Switzerland
[2] Schulich School of Music, McGill University, Montreal, Quebec, Canada
[3] CIRMMT, McGill University, Montreal, Quebec, Canada
[4] IRCAM, Paris, France
[†] victor-manuel.cordero-charles@hesge.ch

Aims/goals

The purpose of our contribution is to give an overview of the transformations of pedagogical practices at the Haute école de musique (HEM) de Genéve/Neuchâtel, Switzerland, in the field of orchestration following the *Orveo* and *e-Orch* research projects. This research has been conducted at the HEM jointly with McGill University (CA) and IRCAM (FR).

There are currently four axes of educational innovation:

1. Development of new pedagogical approaches to orchestration based on a systematic analysis of scores and perception-based research, in order to develop new pedagogical methods and material, including the creation of thesaurus of orchestration.
2. Introduction of science-based notions into orchestration pedagogy.
3. Creation of teaching modules for orchestration in custom built software.
4. Formalization of the relationship between orchestration and the perception of musical form.

Background information

Axis 1: The current pedagogical approach to orchestration is exemplar-based. As Piston (1955) writes in his book on orchestration:

> The composer or orchestrator who scores chords with consistent success from the standpoint of balance of tone, smooth blend of instrumental color, and appropriate dynamic effect, does so because of a reliable memory of the sound made by each instrument at any given pitch and dynamic level. He does not do so because of formulas and devices learned from books. (p. 444)

An optimum pedagogy would ideally straddle these two approaches, with the end goal being a unified theory of orchestration that encompasses timbre (the set of auditory qualities that distinguish sounds emanating from different instruments or their blended combinations), as well as pitch, and other temporal and technical elements.

Axis 2: The physical phenomena involved in the formation of the orchestra's sound have been partially neglected in the treatises. It seems to us essential to introduce the student in the understanding of concepts such as: the directivity of sound sources (modes of radiation of instruments), the additive intensity of sound sources, the effects of masking, etc. These have a major influence on perceptual processes such as timbral blend mixing, stratification, segregation-integration, etc.

Axis 3: From the analysis of a sound source, the application Orchis/Orchidée (Carpentier et al, 2012) proposes a map of instrumental combinations approaching perceptively this source of origin by combining the instrumental sounds of the orchestra. The tool does not orchestrate in the place of the student, but it helps to broaden its conception on the timbre and stimulates its creativity.

Another interface that was developed, Pleione, makes it possible to navigate in a virtual orchestra from symbolic data (conventional notation) and from a series of timbre descriptors of the signal selected audio.

Axis 4: The interaction between the formal articulations and the different levels of contrasts in the orchestration is a fundamental element that determines to a great extent the overall perception of form in the listener and is organized in a hierarchical way to create different structural levels.

Methodology

• Development of software to help orchestration (*Orchis-Pleione*).
• Analysis of multitrack recordings of a real symphony orchestra (symphony orchestra of the HEM).
• Analysis of orchestration treatises for verification and study.
• Development of a thesaurus of orchestration terms, extracted from treatises and books on orchestration.
• Analysis and discussion of results from the new orchestration software.

Results

The influence of research projects has had an important impact on the evolution of teaching practices at the HEM in Geneva. The two consecutive interdisciplinary projects (*Orveo* and *e-Orch*) laid the foundations for the development of an innovative orchestration pedagogy based on these interdisciplinary methods. The use of innovative software to give new ideas and help with orchestration, the use of real recordings of orchestral examples, and the development the formalization of a theory of orchestration based on technical and perceptual processes, are just some of the concrete examples of the transformation that the teaching of orchestration is experiencing at the moment.

Conclusions

Orchestration can be approached in several ways: it is mainly seen as the art of writing, from symbolic data, musical works for instrumental numbers of variable size, combining them with each other in order to produce orchestral effects. If the art of orchestration can be summed up in a few words, it hides in fact a very great complexity: indeed, because of the multidimensional character of the timbre, it becomes necessary, as soon as one wants to study it in depth, to use skills in very different fields. This research was carried out within the framework of an international project, one of whose objectives is to better understand the fundamental aspects of the orchestration, to model them, then to put them at the service of the creation and the pedagogy of orchestration.

Acknowledgments

We would like to acknowledge our colleagues at the HEM Luis Naón and Michael Jarrell, and our partners in this international collaboration: Stephen McAdams, Philippe Esling, Meghan Goodchild, and Carmine Cella.

References

Adler, S. (2000). *Study of Orchestration*, 3rd ed. New York, NY: W.W. Norton.
Barrett, W.A. (1879). *An Introduction to Form and Instrumentation*. London, UK: Rivingtons.
Berlioz, H. (1844; 1905; 1948). *Treatise on Instrumentation*. Trans. T. Front. New York, NY: Kalmus.
Brant, H. (2009). *Textures and Timbres: An Orchestrator's Handbook*. New York, NY: C. Fischer.
Carpentier, G., Cordero, V. & Daubresse, E. (2012). Cartographier le timbre. Projet Orchis/Pléione à la Haute Ecole de Musique de Genève. *Dissonances*, 119, 48-63.
Carpentier, G., Daubresse, E., Garcia Victoria, M., Sakei, K. & Villanueva Carratero, F. (2012). Automatic orchestration in practice. *Comput. Mus. J.*, 36, 24-42.
Carse, A. (1925). 1964. *The History of Orchestration*. New York, NY: Dover Publications.
Goodchild, M. & McAdams, S. (2017). Perceptual processes in orchestration. In E. Dolan & A. Rehding (Eds.), *The Oxford Handbook of Timbre*, Oxford University Press, New York, NY. (forthcoming)
McAdams, S. (2013). *Musical timbre perception*. In Deutsch, D., ed., Psychology of music, 3rd edition, chap. 2, pp. 35–67. San Diego, CA: Academic Press.
McAdams, S., & Bregman, A. (1979). Hearing musical streams. *Computer Music Journal*, 3(4), 26-43.
McKay, G.F. (1963). *Creative Orchestration*. Boston, MA: Ally and Bacon.
Piston, W. (1955). *Orchestration*. New York, NY: W.W. Norton and Co.
Rimsky-Korsakov, Nicholas. (1873; 1912;1964). *Principles of Orchestration: With Musical Examples Drawn From His Own Works*. Edited by M. Steinberg. Translated by E. Agate. New York, NY: Dover.

Timbre, Early Cinema, and the Perception of Time in Mahler's Symphony No. 5

John J. Sheinbaum
Lamont School of Music, University of Denver, Denver, CO, USA
jsheinba@du.edu

Aims/goals

This paper explores Gustav Mahler's processes of timbral continuity and discontinuity in his Fifth Symphony (1904) and proposes contextually relevant frameworks for understanding how Mahler's use of timbre shapes a listener's perception of time.

Background information

Mahler's oeuvre is filled with music about other music, passages that point to time frames beyond the immediate present. These interconnections are often signaled through notable uses of timbre. Mahler's engagement of multiple time frames is often understood through precursors from 19th-century symphonic, operatic, and popular traditions, with which he was intimately familiar (de La Grange, 1997; Freeze, 2017). Less explored are the deep affinities Mahler's music shares with the burgeoning science of experimental psychology, and with the art form of film at the turn of the 20th century, as asserted in Theodor W. Adorno's seminal book on the composer (1960/1992).

Methodology

This research was conducted primarily through score study and close listening to recordings of Mahler's Fifth Symphony, as well as viewing examples of early film and interpreting psychological texts from the turn of the twentieth century.

Results

In the Fifth Symphony, distinctive opening gestures and themes simultaneously evoke past, present, and future. Numerous later passages explicitly reference earlier movements. Most notably, a famously surprising "breakthrough" in the second movement looks forward all the way to the climax of the fifth-movement finale. The temporally provocative passages in the Fifth Symphony communicate long-range connections through some of the most audible and palpable elements of music, such as timbre, mode, rhythm, and melodic contour, rather than exact connections of absolute key area or precisely matching thematic material most likely to be uncovered through untimed study of the score.

Numerous techniques first used in the 1890s by film pioneers such as Georges Méliès, George Albert Smith, and the Edison Manufacturing Company similarly created seemingly impossible juxtaposed images or illusions of continuity across disjunct images, suggesting new sources of metaphor for the spectacle and narrative flow of Mahler's music.

These visual and sonic modes of cultural production additionally resonate with models of temporal perception posited by experimental psychologists such as Wilhelm Wundt (1832-1920) and William James (1842-1910). Their empirical approaches to the introspective construction of time and memory, in addition to the clinical psychoanalytic method developed by Freud, laid the foundations for modernist approaches to consciousness and the self. James's landmark text *The Principles of Psychology* (1890), for example, posits a "stream of consciousness" that "mixes" past, present, and future, and conceives of memory as "a network of attachments" formed by "exciting diverse brain-tracts."

Conclusions

Mahler employs orchestral color not as mere fleshing out of the "real" music, or a disturbing distraction from it, but rather as an essential key, which listeners use to forge connections across large spans of time and recognize processes of continuity, discontinuity, and culmination. These techniques, while somewhat distinctive to Mahler's oeuvre, nevertheless can be understood more deeply by seeking connections to other forms of cultural production and theories about the human mind during the same generation.

References

Adorno, T. W. (1992). *Mahler: A musical physiognomy* (E. Jephcott, Trans.). London: University of Chicago Press. (Original work published 1960).

Freeze, T. (2017). Popular music and the colloquial tone in the posthorn solos of Mahler's Third Symphony. In J. Barham (Ed.), *Rethinking Mahler* (pp. 183-201). New York: Oxford University Press.

James, W. (1890). *The principles of psychology*. New York: Henry Holt.

La Grange, H. de. (1997). Music about music in Mahler: Reminiscences, allusions, or quotations? In S. E. Hefling (Ed.), *Mahler studies* (pp. 122-168). Cambridge: Cambridge University Press.

Timbre and Texture: Overlapping Phenomena

Amit Gur
Antwerpen Research Institute for the Arts | University of Antwerp, Belgium
amittgur@gmail.com

Aims / Goals

The principal aim of this paper is to compare notions of *timbre* and *musical texture*, to explore the overlap in their meaning and to demonstrate how a definition of musical texture can contribute towards better understanding of timbre. Timbre is often defined by negation: "everything that cannot be labeled pitch or loudness" (McAdams & Bregman, 1979, p. 34). I will attempt to characterize the aesthetic quality of timbre in a more affirmative way. To achieve that I will present a theoretical model for texture that incorporates the notion of timbre: a perceived unity, characterized by different levels of homogeneity. In this context I will not address texture and timbre as different phenomena but approach timbre as a borderline case of texture, characterized by a *stronger* homogeneity at the expense of a *lower* perception of plurality.[1]

Background information

The term texture in music, although used commonly, appeared very late in music dictionaries,[2] and a consensual definition has not yet been formed. Even in the domain of visual perception, where texture has been studied for decades, researchers have great difficulty in arriving at a definition that covers all its manifestations. In my ongoing PhD research, I explore three mutually related attributes of texture (both visual and auditory): 1) there are too many elements to be perceived simultaneously or to be held in working memory, therefore the quantitative measurement is lost; 2) because we cannot perceive all elements simultaneously, we tend to rely on a common quality possessed by all of them; 3) because texture is defined by a common quality, the number of elements constituting it is irrelevant.

These three attributes are manifested in a unique feature of texture: it is experienced between boundaries (over space or time) but is not characterized by them. I call this feature *boundlessness nature*. In the following example you can see three different objects, defined by different boundaries, that share the same texture. In the auditory domain the medium is time, thus a change in the number of elements of texture affects its temporal boundaries.

Timbre and texture

In musical discourse, timbre and texture are sometimes used interchangeably. Timbre is often referred to as "a texture of a single sound" (for example, in a popular source like Wikipedia). Both are perceptual qualities characterized by unity, but the unities of timbre and texture are different. The philosophical idea of *emergence* is a good way to demonstrate a difference between them: an emerging property occurs through the aggregation of several components. It does not exist in individual components and becomes apparent only when they are combined. Christian von Ehrenfels[3] distinguishes between two kinds of emerging properties in perception: *preservative* and *eliminative*. In the former case, the property does not replace the individuality of its components; Ehrenfels offers melody as an example. In the latter case, the emerging property replaces the perception of its components, for example in color. Timbre is an *eliminative* emerging property since it replaces the individual perception of overtones,[4] while texture is *preservative*, since its

[1] Dubnov S. 1996 discusses the perception of a timbre with a high level of temporal divisibility as a borderline case of texture.

[2] Dunsby J. 1989, discussed at length the evolution of the notion of texture among musicians.

[3] Ehrenfels mentions the idea in his essay 'On Gestalt-qualities' (1988).

[4] Kubovy, M., & Van Valkenburg, D. (2001) suggest that timbre is an *eliminative emerging property*.

components, the textons,[1] are evident in perception. Similarly to texture, timbre shares the unique feature of *boundlessness nature;* timbre, like color,[2] is experienced between boundaries but is not characterized by them. In other words, texture and timbre are both *vertical* qualities in music, which means they are perceived as stationary qualities over the time domain.

Methodology

- I start from the assumption that timbre is a perceptual phenomenon that relates to the notion of texture.
- Through conceptual analysis, I get a grip of the differences and similarities between the two phenomena: I use the concept of *emergence* to demonstrate a fundamental difference between the two, showing how texture preserves its components in the auditory scene, while timbre replaces them. I relate to a feature that exists in texture and timbre—they both exist between temporal boundaries but are not defined by them. I call this feature "boundlessness nature."
- I discuss different attributes that are similar to texture and timbre: number of constituting elements, relationship to temporal boundaries, attentive role - figure and background and more.
- I show how by increasing temporal density, through granular synthesis, a texture will lose its *preservative* quality and become *eliminative*: in this case texture will be interpreted as timbre.

Conclusions

Texture and timbre are perceptual phenomena which are similar in nature: both are characterized by a relative unity over time and emerge from the combination of numerous elements. However, while elements of texture are perceptually evident, in timbre they are not. Texture and timbre can be placed on a continuum: an increasing temporal density of elements in auditory texture, causes greater homogeneity. In this context, we can regard timbre as a borderline case of texture that has a very high level of homogeneity at the expense of a weaker perception of plurality.

References

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, 80, 97–126.

Dubnov S. (1996). *Polyspectra Analysis of Music Texture*, PhD dissertation, Hebrew University, Jerusalem.

Cohen, D. & Dubnov S. (1996). Gestalt phenomena in musical texture. In M. Leman (Ed.), *Music, Gestalt, and Computing* (pp. 386-405). Springer, Berlin, Heidelberg.

Dunsby J. (1989), Consideration of Textures. *Music & Letters*, Vol. 70, No. 1 (Feb., 1989), pp. 46-57

Ehrenfels, C. von (1988), On `gestalt qualities'. In B. Smith (Ed.), *Foundations of gestalt theory*; Munich, Germany: Philosophia Verlag, pp. 82 ± 117. (Original work published 1890)

McAdams, S. & Bregman, A. S. (1979), Hearing musical streams. *Computer Musical Journal*, 3(4): 26-43.

McAdams S. (1984), *Spectral Fusion, Spectral parsing and the Formulation of Auditory Images*, Ph.D. dissertation, Stanford University, CCRMA report no. STAN-M-22, Stanford, CA.

Helmholtz, H. 1954 (Original German Edition in 1863, English translation in 1954). *On the Sensations of Tone* (translation by Alexander John Ellis). Dover.

[1] This term was coined by Bela Julesz, it refers to the elements constituting a texture.

[2] Helmholtz 1954 discussed at length the analogy between timbre and color

Timbre Transfer between Orchestral Instruments with Semi-Supervised Learning

Adrien Bitton[1,†], Axel Chemla--Romeu-Santos[1,2], and Philippe Esling[1]
[1] IRCAM CNRS – UMR 9912, 1 Place Igor Stravinsky, F-75004 Paris, France
[2] Laboratorio d'Informatica Musicale, Università Degli Studie di Milano, Via Festa del Perdono, 20122 Milano, Italy
[†] adrien.bitton@ircam.fr

Aims/goals

Understanding the timbre qualities of orchestral instruments is a central question in music perception, usually studied through signal descriptors. Recent advances in *unsupervised generative models* provide an alternative way to obtain higher-level features to study timbre by constructing *latent spaces*, learned on sets of sounds through hierarchical models. Once learned, these models allow *weight sharing* with others to provide a common knowledge space. Hence, transferring latent spaces allows *domain translation* so that data are encoded in one architecture, and generated by another one. Alternatively, *semi-supervised models* allow conditioning the latent space structure according to additional information such as the instrument or note labels, in order to explicitly interact with these. We propose to use and compare such architectures to achieve what we call *timbre transfer*: using instrument-specific generative systems for analyzing timbre organizations relative to each one, but also to build a common timbre space of higher abstraction that can be used for analysis purposes and sound translation from one instrument to the other.

Background information

Observations of complex data are caused by several underlying generative factors with specific perceptual features. However, timbre factors are usually entangled and cannot be extracted directly from the audio data. Here, we address the problem of modeling continuous timbre representations with a hierarchical approach, based on multi-layered inference generative systems. Recent advances in *unsupervised* and *variational learning* allow the inference of high-level features by directly generating from the space of factors of variations and compare to the original data. In the framework of *Deep Latent Gaussian Models* (DLGM), given hierarchical continuous spaces such that data can be generated recursively from the top layer with a generative model (called *decoder*)

$$p_\theta(\mathbf{z}_L, ..., \mathbf{z}_1, \mathbf{x}) = p_\theta(\mathbf{z_L}), p_\theta(\mathbf{z}_{L-1}|\mathbf{z}_L)...p(\mathbf{x}|\mathbf{z}_1)$$

where the densities are parametrized by neural networks. We approximate the intractable inference model with a distribution mirroring a simpler generative process

$$q_\phi(\mathbf{z}_1, ..., \mathbf{z}_L|\mathbf{x}) = q_\phi(\mathbf{z}_L|\mathbf{x})...q_\phi(\mathbf{z}_1|\mathbf{x})$$

DLGMs are trained with a variational procedure, by maximizing the following lower-bound

$$\mathcal{L}_{\theta,\phi,\mathbf{x},\mathbf{z}} = \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}\left[\log p_\theta(\mathbf{x} \mid \mathbf{z})\right] - \beta \sum_{l=1}^{L} \left[\mathcal{D}_{KL}\left(q_\phi(\mathbf{z}_l \mid \mathbf{x}) \| p(\mathbf{z}_l)\right)\right]$$

which can be explained as the optimization of a likelihood (*reconstruction*) term and a regularization term that forces the system to learn an encoding that corresponds to a *prior* distribution.

Recently, the idea of *style transfer* tries to use properties of deep networks to apply the style of an input to another (Gatys, 2016). An interesting method of *image translation* (Liu, 2017) was proposed based on the idea of a *shared latent space*, where variations between different data are matched in a common projection space. Here adapted to orchestral instruments in order to perform *timbre transfer* by training the first layers for each instrument separately (instrument relative features), while higher latent layers will be shared by all instrument-specific architectures to represent common higher-level timbre space. This architecture allows us to perform timbre transfer by encoding the spectral content from a first instrument and decoding this latent variable with the decoder of another instrument against an adversarial criterion.

On the other hand, semi-supervised architectures allow the building of a conditioned latent space which may explicitly be organized according to instrument and note labels. Within each sub-space the algorithm aim at modeling the separate instrument timbres. Given that they are consistently organized, timbre transfer would then be performed by switching the instrument condition before the decoding stage. By conditioning a single auto-encoder network together with the several instrument classes, we make use of the full database at once and alleviate the limitation of simultaneously processing two domains.

Methodology

Datasets. We use the *Studio On Line* (SOL) recordings of 12 instruments (in Winds, Strings, Keyboards, Brass) with 10 playing styles for each in every pitches and velocities available on which are computed the Non-Stationary Gabor Transform (NSGT) on the Mel Scale for its beneficial properties in terms of input representation (logarithmic scale, quality/dimensionality trade-off) and invertibility for generation.
Evaluation. First, we train 3-layered DLGMs separately on every instrument and perform classification tests with Multilayer perceptron (MLP) classifiers to evaluate how the obtained representations are instrument-dependent and how many latent dimensions are required. Then, we jointly train couples of models for various pairs of instruments with the procedures described in the previous section. We evaluate the quality of within-instrument and cross-instrument generation and observe the projection of notes and instruments in the shared latent space with respect to acoustical descriptors and symbolic properties such as conditioning labels. We test our timbre transfer assumption and compare different architectures and conditioning results in likelihood on separate test data and aural evaluation.

Results

Hierarchical timbre spaces. By comparing latent spaces obtained from models trained on each instrument separately or across all available instruments, we obtain clear organized features, both instrument-dependent and global, confirming the shared latent space hypothesis.
Timbre transfer. We study how timbre information is abstracted and transferable by performing instrument translation, quantitatively and qualitatively assessed in its quality and realism.
Analyzing latent space traversal. For qualitative evaluation of the generative factors in the encoded representation, we keep all the latent units fixed to a given set of values and only vary one dimension which is traversed. The samples generated appear to express the range of variation within the corresponding high-level feature. Hence, we can separately modify such features and control the properties of resynthesized sounds directly from the unsupervised representation. Alternatively, we perform interpolations in between notes or instruments to experiment with morphings from the latent features.

Conclusions

We studied the idea of *timbre transfer* by learning variational models adapted to each orchestral instrument. Different latent spaces provide insights on the spectral organization for each instrument through non-linear dimensionality reduction. By matching the high-level latent spaces learned for each instrument, we could transfer the decoder of one instrument to the encoder of another. We used this idea to perform innovative generation processes. By evaluating the traversal of latent space and the quality of timbre transfer, we provide insights on the transferability of timbre among families of instruments.
As a continuation of this timbre study, the intricate relationship between timbre and playing style may be studied according to its correlation to the sensitivity of certain latent units across different subsets of playing styles from our dataset, possibly enabling its control while performing *timbre transfer*.

Acknowledgments

This work benefited from funding of the MAKIMOno 17-CE38-0015-01 project of the French ANR.

References

Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision* (pp. 2414-2423).
Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
Liu, MY, Breuel, T. & Kautz, J. (2017). Unsupervised Image-to-Image Translation Networks, NIPS 2017

Generating Orchestral Sequences with Timbral Descriptors

Aurélien Antoine[1†], Eduardo R. Miranda[1], Jean-Michaël Celerier[2], Myriam Desainte-Catherine[2]
[1] Interdisciplinary Centre for Computer Music Research (ICCMR), Plymouth University, Plymouth, United Kingdom
[2] Laboratoire Bordelais de Recherche en Informatique (LaBRI), University of Bordeaux, Talence, France
[†] aurelien.antoine@postgrad.plymouth.ac.uk

Aims/goals

This project aimed to develop a computing system capable of generating sequences of orchestral instrument combinations matching specific perceptual qualities represented by timbre properties.

Background information

One of the characteristics of orchestration is the ability to create unique sounds or textures, an outcome appreciated by composers and utilized to produce specific perceptual effects. This compositional practice involves writing for several instruments, which means a large number of potential combinations. The significant search space is challenging when developing computing system designed to aid in orchestral composition. The presented research combines the work on analysis and classification of timbre within orchestral audio and machine learning developed at Plymouth University's ICCMR (Antoine et al. 2016; Antoine and Miranda 2017a; Antoine and Miranda 2017b) and the research into methods and tools to write and execute interactive scenarios developed at University of Bordeaux's LaBRI, conducted under the OSSIA (Open Scenario System for Interactive Applications) project (Celerier et al. 2015; De la Hogue et al. 2014).

Methodology

The presented approach uses machine learning methods for classification and regression tasks. First, supervised learning algorithms are utilized to harness and learn specific timbre properties from orchestral audio sources. Here, Artificial Neural Networks (ANNs) are trained with a training corpus manually labeled by the authors in order to create classification models to automatically estimate and evaluate the perceptual qualities of instrument combinations. The second set of supervised learning methods provides regression models to predict the timbre properties of note combinations without the need to perform an acoustical and psychoacoustical analysis. Once again, ANNs algorithms are utilized to create regression models, which have been trained using sets of instrument combination examples along with their calculated timbral values. Notes are then combined following predefined interval rules and are output only if they match the desired perceptual qualities. Therefore, the search algorithm is guided by the potential perception of the instrument mixture. The different timbre classification and regression models have been incorporated into the OSSIA *Score* software[1], which offers an interface to design scenario of orchestration sequences. Users can select different groups of orchestral instruments and specify the desired perceptual qualities represented with timbral descriptors such as *brightness* or *roughness*.

Results

The regression and classification models created by supervised learning methods have been able to predict timbral characteristics of note combinations, which can be used to identify perceptual features directly from instruments information. Such techniques, when integrated into a generative system, allowed for the instant perceptual characteristics estimation of instrument combinations, and thus, to generate orchestrations based on their potential perceptual qualities. Their integration in the OSSIA *Score* software (Fig. 1) has enabled a method to generate sequences of instrument combination guided by lists of timbral descriptors describing their potential perceptual qualities. Moreover, *Score* incorporates

[1] https://ossia.io/

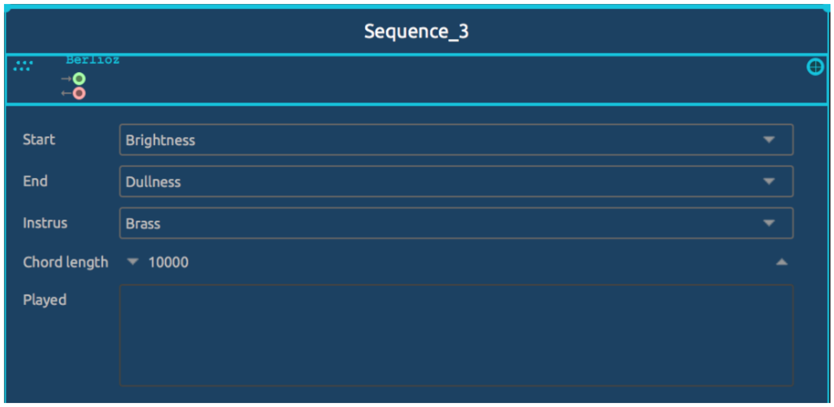


Figure 4. The generative process embedded in a OSSIA Score object.

conditional processes, which offers the design of multiple scenarios. This can be used as an exploratory technique to test different orchestration ideas.

Conclusions

This paper presented an approach designed to generate instrument combinations based on their potential perceptual qualities. The use of supervised machine learning methods has enabled the automatic prediction of timbre characteristics representing the perception of sonic properties, without the need to perform acoustical and psychoacoustical analysis on an audio source. The integration of the different classification and regression models in the OSSIA *Score* software has offered a method to design interactive and conditional scenarios representing sequences of orchestral instrument combinations to be generated following timbral descriptors. Such an approach could be utilized to refine the significant instrument search space faced in computer-aided orchestration systems.

References

Antoine, A., Williams, D., Miranda, E.R. (2016). Towards a Timbral Classification System for Musical Excerpts, *Proceedings of the 2nd AES Workshop on Intelligent Music Production*. London, UK, 2016.

Antoine, A. and Miranda, E.R. (2017a). Musical Acoustics, Timbre, and Computer–Aided Orchestration Challenges, *Proceedings of the 2017 International Symposium on Musical Acoustics (ISMA)*. Montréal, Canada.

Antoine, A. and Miranda, E.R. (2017b). Computer Generated Orchestration: Towards Using Musical Timbre in Composition, *Proceedings of the 9th European Music Analysis Conference (EuroMAC 9 – IX CEAM)*. Strasbourg, France.

Celerier, J-M., Baltazar, P., Bossut, C., Vuaille, N., Couturier, J-M. (2015). OSSIA: Towards a unified interface for scoring time and interaction, *Proceedings of TENOR 2015 Conference*. Paris, France.

De la Hogue, T., Baltazar, T., Desainte-Catherine, M., Chao, J., Bossut, C. (2014). OSSIA: Open Scenario System for Interactive Applications, *Proceedings of Journées de l'Informatique Musicale*. Bourges, France.

From Timbre Decomposition to Music Composition

Frédéric Le Bel
Music Representation Team, IRCAM (UMR 9912 STMS), Paris, France;
Centre de recherche en informatique musicale, MUSIDANCE, Université Paris 8, Paris, France.
frdric.lebel@gmail.com | http://repmus.ircam.fr/lebel/

Introduction

This abstract proposes to overview a work in progress that aims to develop a computer-aided-composition (CAC) approach to structuring music, by means of audio clustering and graph search algorithms. Although parts of this idea have been investigated in order to achieve different tasks such as corpus-based concatenative synthesis, musical genre recognition or computer-aided orchestration to name a few, the challenge remains to find a way of integrating these techniques into the composition process itself, not to generate material, but to explore, to analyse and to understand the full potential of a given sound corpus (sound file database) prior to scoring a musical piece; being instrumental, acousmatic or mixed. As opposed to mainstream CAC tools, mostly focusing on generative methods, the following one proposes an analytical approach to structuring music based on auditory attributes and their physical correlates. Basically, the idea is to use unsupervised machine-learning to extract hidden structures from sound corpuses' features space and to translate them into musical structures at different stages of the composition; from the micro-form to the macro-form. From another perspective, the idea is to elaborate musical works based, not on patterns proliferation or similar techniques, but rather on relationships that bring together different sound entities. Consequently, the goal here is to unfold the algorithmic structure through a prototype software and to reveal how it was used to compose a recent piece of mine: *Il ne s'agit pas de le recomposer*, for augmented string quartet and fixed media, in order to examine the methodology, to discuss a few important co-lateral problematics, to expose the limitations of such an approach and finally to share ideas for future developments.

Structural overview

Based on a typical unsupervised machine-learning architecture, the following algorithmic structure may be divided into three distinct stages. The first one focuses on data extraction (audio features extraction), the second one focuses on data analysis (features selection, clustering algorithm and evaluation criterion), and the third one focuses on data sorting (graph search algorithm). Essentially, the task is to deduce a musical structure from a sound file database (audio and not symbolic).

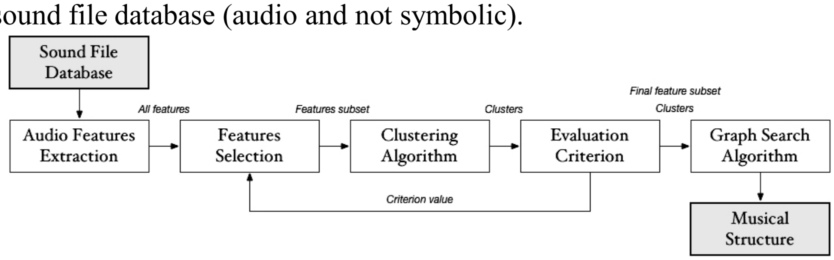


Figure 5. Algorithmic structure

From timbre decomposition… (data extraction)

In the frame of this work, the audio features extraction consists of decomposing each sound file from the database by mapping the signal's short-term Fourier transform (STFT) magnitude into a lower-dimensional domain that more clearly reveals the signal characteristics. Assumed to represent specific auditory attributes, these features inform different aspects of the temporal structure, the energy envelope, the spectral morphology and the harmonic content of sounds. From these low-level data models, the sound files are then compared, taken pairwise, on a dissimilarity/distance basis in order to generate what could be seen as an n-dimensional timbre-space upon which the clustering algorithm can be later applied. Briefly, the evaluation criterion for the features selection aims at maximizing the inter-clusters distances and at minimizing the intra-clusters' ones. In this particular case, the timbre-space metaphor should be considered as an exploratory structure of dissimilarity data rather than a comprehensive perceptual coordinate system of timbre.

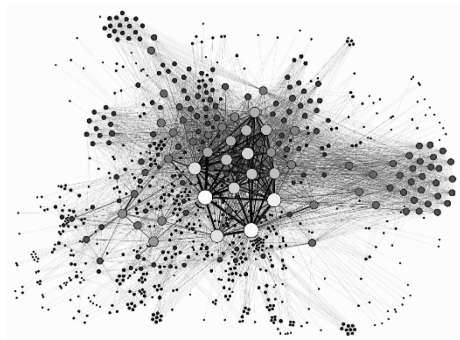


Figure 2. 2D timbre-space network. The nodes represent the sounds and the edges the distances between them.

…To musical structure (data sorting)

Following an extensive data analysis stage (the features selection and the clustering algorithm), the graph search consists of finding path(s) within the timbre-space that could suggest way(s) of sequencing the different clusters of sounds and their components in a chronological order. Considering that the resulting network may be seen as a complete, undirected and weighted graph, multiple approaches could be used but spanning trees were investigated first. More precisely, the minimum spanning tree (MST) seemed to provide an interesting premise accordingly with the timbre-space's underlying dissimilarity data structure. A MST is a subset of the edges of an undirected and weighted graph that connects all the nodes together with the minimum possible total edge weight without any cycles. In other words, it is a tree (not a path) whose sum of edge weight is as small as possible. Hence, the MST does not represent an ordered sequence of clusters but rather some sort of an optimized road map through the timbre-space from which multiple paths may be drawn. However, the clusters and their components remain connected in a way that the global similarity is maximized. The Kruskal algorithm was used in the frame of this work in order to obtain structures such as the following one.

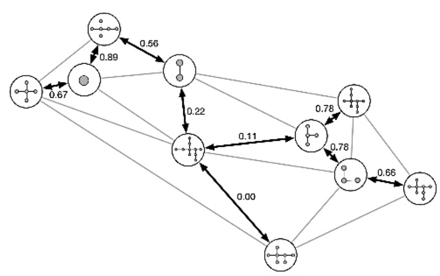


Figure 3. Two-layer MST extracted from the timbre-space, where the big nodes represent the clusters of sounds, the small nodes represent the sounds and the edges on both layer represent the transition/emission probabilities.

The previous two-layer MST already suggests a rather clear musical structure, or a particular definition of it, but yet remains to be expressed in the time domain. For that, the model is translated into a hidden markov model (HMM) for which the edges, originally distance based, are converted into transition and emission probabilities respectively for each layer (global structure and local structures) according to the following principle: $1\text{-}d_i/\text{argmax}(d)$. In other words, the closer or the more similar are two sounds, or two clusters of sounds, higher is the probability to transit from one to the other and vice-versa. Finally, a customized polyphonic step sequencer is used to articulate the resulting probabilistic model on a timeline and let the musical structure be heard. My last piece: '*Il ne s'agit pas de le recomposer*', for augmented string quartet and fixed media, is entirely based on these principles and was composed as a proof of concept in order to assess, as objectively as possible, the artistic potential of such an approach to composition.

For more details: http://repmus.ircam.fr/lebel/from-timbre-decomposition-to-music-composition

A Musicological Approach to the Analysis of Timbre

Megan Lavengood
Music department, George Mason University, Fairfax, VA, USA
meganlavengood@gmail.com

Aims/goals

This paper presents a musicological methodology for the analysis of timbre. The methodology combines acoustics (via the study of spectrograms) with cultural criticism, which produces data that can then be used to generate effective, timbre-based music analyses. Case studies are briefly referenced, which demonstrate the practical utility of this approach.

Background information

Music theory is a field that focuses on developing methodologies for close analysis of specific pieces of music. Traditionally, music analysis of this sort is done with a printed musical score; therefore, music theories have tended to favor musical domains captured in that medium, like rhythm, pitch, and structure. Timbre is only an abstract idea in the musical score, and thus historically, timbre has typically been neglected in music analysis. Yet timbre is one of the most immediate aspects of our musical experience, so many contemporary music theorists have recently become interested in timbre analysis.

In music theory, two distinct approaches to timbre analysis exist, with complementary strengths and limitations. First, music theorists from the 1980s adopt a positivist mindset and look for ways to quantify timbral phenomena, often using spectrograms, avoiding any cultural dimensions of their work (e.g., Cogan 1984). Second, writings of the past five years focus on the cultural aspects of timbre but make no use of spectrograms (e.g., Heidemann 2016).

Methodology

The methodology presented in this paper synthesizes the above approaches by situating spectrogram analysis within a broad cultural context, taking direct account of listener experience, i.e., "perceptualization" (Fales 2002), through the notions of markedness (Hatten 1994). A vocabulary for timbre analysis is established, which relies on a system of binary oppositions (e.g., *percussive/soft, rich/sparse,* etc.) that are defined partly via spectrogram analysis. The binaries are further defined as either marked (+) or unmarked (–). In any linguistic binary-oppositional pair, the terms are semiotically asymmetrical; a marked term linguistically contains some kind of additional conceptual information that an unmarked term does not have (Hatten 1994). In the present study, the unmarked term within each opposition is the one that is most timbrally normative within a given musical context. The marked term is in some way non-typical. To interpret the data gathered through spectrogram analysis, timbre analysis is connected with textural function, narrative, and ethnography to interpret the acoustic data of spectrograms.

To this end, the methodology is applied in a case study that uses 1980s popular music as a focal repertoire.

Results

Sounds used in a given track are categorized as belonging to one of three groups, or *instrumentational categories*: a) *core sounds*, which articulate structural aspects of pitch and rhythm of the song, b) *melody sounds,* which are the voice and any instrument replacing the voice, or c) *novelty sounds*, used primarily for coloristic effects. This paper focuses on 1980s popular music; categorization therefore was determined by analysis of the instrumentation of many 1980s singles. A correlation arises between the timbral characteristics of these instruments and their instrumentational category: the *core* and *melody* sounds share unmarked timbral properties. *Novelty* sounds are intrinsically difficult to generalize but tend to feature marked timbral characteristics.

Conclusions

This work shows that effective musicological analyses of timbre can be produced by combining spectrogram study with broader cultural contextualizations. While much musicological work has moved away from spectrograms as a reliable source of information on timbre, this methodology demonstrates that spectrograms can still play a role in culturally-sensitive timbre analyses. Understanding certain timbral features as being unmarked indicates that those features allow a texture to blend into the groove's fabric, rather than demanding attention, while marked features are useful as hooks or signature sounds. Furthermore, instances of subversion of these timbral norms enable the analyst to locate musical meaning created through the manipulation of timbres.

This paper further demonstrates the value and necessity of the inclusion of timbre in popular music analyses. Because popular music is primarily distributed through recordings, timbre takes on particularly immediate significance in popular music. Developing approachable methodologies such as this one is an essential step in advocating for greater attention to timbre in analysis of popular music, and music of other genres.

References

Cateforis, T. (2011). *Are We Not New Wave? Modern Pop at the Turn of the 1980s*. Ann Arbor: University of Michigan Press.

Cogan, R. (1984). *New Images of Musical Sound*. Cambridge, MA: Harvard University Press.

Fales, C. (2002). The Paradox of Timbre. *Ethnomusicology*, *46*(1), 56. https://doi.org/10.2307/852808

Hatten, R. S. (1994). *Musical Meaning in Beethoven: Markedness, Correlation, and Interpretation*. Bloomington: Indiana University Press.

Heidemann, K. (2016). A System for Describing Vocal Timbre in Popular Song. *Music Theory Online*, *22*(1). Retrieved from http://www.mtosmt.org/issues/mto.16.22.1/mto.16.22.1.heidemann.html

Malloch, S. (2000). Timbre and Technology: An Analytical Partnership. *Contemporary Music Review*, *19*(2), 155–72.

McAdams, S., & Giordano, B. L. (2009). The Perception of Musical Timbre. *The Oxford Handbook of Music Psychology*, 72–80.

Saldanha, E., & Corso, J. F. (1964). Timbre Cues and the Identification of Musical Instruments. *Journal of the Acoustical Society of America*, *130*(5), 2021–6.

Wallmark, Z. T. (2014). *Appraising Timbre: Embodiment and Affect at the Threshold of Music and Noise* (Dissertation). University of California Los Angeles.

Gender, Timbre, and Metaphor in the Music of Wendy Carlos

Sarah Schoonhoven
Butler School of Music, University of Texas, Austin, Texas, USA
sarah.schoonhoven@gmail.com

Aims/goals

This paper draws on critical and scholarly reception of synthesized works by Carlos ranging from the 1969 release of "Switched-On Bach" through the present day to explore and unpack the gendered implications of the timbral metaphors used to describe her works. My analysis builds on the work of Tara Rodgers surrounding timbral metaphor and its use in electronic musics. Similarities in the metaphors used to describe her music over the past half century provide insight into not only the "tone color" of the sound but also the music's cultural implications and the audience's perceptions of the music's creator. This paper will examine, compare, and analyze metaphors to characterize Carlos's "March from Clockwork Orange" in order to analyze the ways that perceptions of Carlos and her perceived gender identity have fed into interpretations of her music.

Background information

Tara Rodgers suggests in her dissertation, "Synthesizing Sound: Metaphor in Audio-Technical Discourse and Synthesis History," that a feminist analysis of a musical work would reject the "fundamental parameters of sound" such as pitch, meter, and harmony, and examine timbre through metaphor. Using metaphor to describe music is something that theorists have already been doing, according to her, whether or not the field has been completely aware of such. Consciously studying timbral metaphor, then, involves an awareness of the metaphoric framework, bringing to light the relations between a musical example and the power structures surrounding it.

Methodology

This work is done by examining critical reception and scholarship surrounding Carlos's works, both those composed for *A Clockwork Orange* and those created as independent works, before examining the metaphors Carlos uses to describe her own sounds. The paper then analyzes Carlos's "March from A Clockwork Orange," an electronic realization of Beethoven's "Turkish March," using Carlos's own metaphors and metaphorical categories from her writings.

Results

Carlos's "March from A Clockwork Orange" uses synthesized timbres to draw connections between sections and bring out voices that might have otherwise been hidden in the fuller orchestration of the original Beethoven transcription. Her metaphors tap into the "human" element of the performance, both in its allusions to acoustic instruments and its motion-focused metaphors, which draw the listener into a special and temporal sound event. At a larger level, however, Carlos utilizes the timbres available on her synths and vocoder to exaggerate the troping of the "Turkish march" and "grand chorale" topics beyond the capabilities of normative Western classical instruments. The grand becomes grander, and the markedly different becomes outright ridiculous and absurd.

Conclusions

Wendy Carlos's music has been incredibly influential to electronic music genres, and whether her lack of current popularity is due to gendered perceptions of her music or simple changes of taste, analysis of her music is still a valuable tool for understanding twentieth century electronic music. Electronic timbres have been largely overlooked in discussions of twentieth century artists, and an analysis of timbral metaphors can give the scholar insight into the cultural assumptions and social implications surrounding a work.

References

Babaoğlu Balkiş, L. (2010). Defining the Turk: Constructing meaning in operatic Orientalism. *International Review of the Aesthetics and Sociology of Music,* 41(2), Croatian Musicological Society.

Bell, A. (1979, May 17). Wendy/Walter Carlos: a Candid conversation with the "Switched-On Bach" composer who, for the first time, reveals her sex-change operation and her secret life as a woman. *Playboy*.

Cateforis, T. (2014). *Are we not new wave?* Ann Arbor: University of Michigan Press.

Hatten, R. (2004). *Interpreting musical gestures, topics, and tropes: Mozart, Beethoven, Schubert.* Indianapolis: Indiana University Press.

Peraino, J. (2015). Synthesizing difference: the Queer circuits of early synthpop. In *Rethinking difference in music scholarship,* O. Bloechl, M. Lowe, and J. Kallberg (Ed.), 287-314. Cambridge: Cambridge University Press.

Pinch, T., & Trocco, F. (2004). *Analog days: the Invention and impact of the Moog synthesizer*. Cambridge, MA: Harvard University Press.

Rodgers, T. (2010). Synthesizing sound: Metaphor in audio-technical discourse and synthesis history. PhD diss., McGill University.

Wajcman, J. (2006). the Feminization of work in the Information Age. *Women, gender, and technology.* Champaign: University of Illinois Press.

Cross-Cultural Gestures: Timbral Juxtapositions in Takemitsu's *November Steps* (1967) for Shakuhachi, Biwa and Orchestra

Flora Henderson
Independent Scholar, UK
arolf1187@icloud.com

Aims/goals

To show how inharmonic frequencies – rough timbres – are privileged musical attributes in Japanese music and how this inharmonicity has been used by composers, and to demonstrate a framework for discussing these privileged timbres of Japanese music in musical analysis. This contributes to discussion of how timbral inharmonicity is used in non-western musical traditions.

Background information

Timbre has equal status with pitch in many genres of East Asia. Within that paradigm in Japan, rough timbres – inharmonic frequencies – are privileged in many music traditions, for example the traditions of the shakuhachi flute and the biwa lute. This privileging of rough timbres is expressed through instrumental organography, defined performance techniques, and as an element of musical structure.

Composers have been attracted to these rough timbres, and the large range of timbral techniques available, and have explored these sounds and techniques in cross-cultural works, notably Toru Takemitsu's (1930–1996) *November Steps* (1967) for shakuhachi, biwa and western orchestra. In this work, Takemitsu juxtaposed the Japanese and western instruments through a series of exchanges between the two traditions, culminating in a lengthy cadenza for the shakuhachi and biwa and followed by a short orchestral finale. Timbral range is prominent in the structure of the work and when writing for the Japanese instruments, Takemitsu explores their timbral techniques to articulate the privileged roughness that he labelled "beautiful noise" (Takemitsu 1995: 64–65).

Despite numerous analyses of this work (Smalldone 1989, de Ferranti and Narazaki 2002), Takemitsu's use of timbre in the shakuhachi-biwa exchanges with the orchestra (excluding the cadenza) has been underexplored. Timbral analysis has little scope in western musical analysis, while in ethnomusicology, models of timbral analysis tend to be specific to the music tradition they represent (Feld 1982, Stobart 1996). In acoustics and cognition, useful frameworks for timbral analysis have been proposed (Grey 1977, Toiviainen et al. 1995, Caclin et al. 2006), although findings have often been undermined by an overreliance on western musical forms or by technical limitations (Serman and Griffiths 2002l in modeling timbre with approaches that can be applied to cross-cultural musics, contributing to discussion of the use of timbre as a cross-cultural musical vehicle.

Methodology

Timbre can be defined as a gesture (Halmrast et al. 2010). Using a model of gestural analysis (Henderson 2016) based on research in acoustics, cognition and gesture (Ben-Tal 2012; Tsang 2002, and McAdams 2004), individual timbral gestures can be identified and the use of these gestures at phrasal level and in an overall timbral trajectory can be discussed. Individual timbral gestures are defined according to the parameters of the musical work and knowledge of the music tradition of that instrument. This enables a focus on the shakuhachi and biwa exchanges with the orchestra, in particular Takemitsu's use of the rough timbres of these Japanese instruments in this work, and how this "beautiful noise" is a cross-cultural vehicle.

Results

For the Japanese instrument sections, Takemitsu uses both traditional and non-traditional gestures of timbral inharmonicity, often in unconventional phrasal structures, to create a timbral trajectory consonant with the timbral trajectory he explores through the orchestral voices. He uses timbral heterogeneity to juxtapose cultural differences between Japanese and western voices, however he also uses timbral homogeneity between the Japanese and western instruments, often with pitch movement, to create a gestural synthesis in which the two traditions meet.

Conclusions

Takemitsu successfully uses the intentional inharmonicity of Japanese instrument timbres in *November Steps* to create both cross-cultural juxtapositions and cross-cultural synthesis. This analysis also allows us to see how these timbres can be used to construct and develop musical form in the same way that melody and rhythm are used, how we can frame intentional timbral inharmonicity as a medium of cross-cultural musical engagement and provides a model applicable to timbral analysis in other music traditions.

References

Ben-Tal, O. (2012). Characterising musical gestures. *Musicae Scientiae, 16*(3), 247–261.

de Ferranti, H., Narazaki, Y. (Ed.). (2002). *A Way a Lone: Writings on Toru Takemitsu*. Tokyo: Academia Music Limited.

Godøy, R. I., Leman, M. (Ed.). (2010). *Musical Gestures: Sound, Movement, and Meaning*. New York: Abingdon, UK: Routledge.

Gritten, A., King E. (Ed.). (2006). *Music and Gesture*. Aldershot, UK; Burlington VT, USA: Ashgate Publishing Ltd.

Gritten, A., King, E. (Ed.). (2011). *New Perspectives on Music and Gesture*. Aldershot, UK: Ashgate Publishing Ltd.

Halmrast, T., Guettler, K., Bader R., Godøy, R. I. (2010). Gesture and Timbre. In R. I. Godøy, Leman, M. (Ed.), *Musical Gestures: Sound, Movement and Meaning* (pp. 183–211). New York; Abingdon, Oxon: Routledge.

Henderson, F. (2016). *Contemporary composition for shakuhachi and western instruments: timbral gesturality in the analysis of cross--cultural music.* (PhD), SOAS, University of London.

McAdams, S., Depalle, P., Clarke, E. (2004). Analysing Musical Sound. In E. C. Clarke, N (Ed.), *Empirical Musicology*. Oxford: OUP.

Takemitsu, T. (1995). *Confronting Silence: Selected Writings of Toru Takemitsu* (Y. Kakudo, Glascow, G., Trans.). Berkeley, CA.: Fallen Leaf Press.

Takemitsu, Tōru. (1967). November Steps. New York, London: C. F. Peters Corporation.

Tōru Takemitsu performed by the Tokyo Metropolitan Symphony Orchestra and Hiroshi Wakasugi. (1994). Track 2 – November Steps, on the album "Takemitsu: Visions, November Steps". SLG, LLC. Denon CO 79441.

Tsang, L. (2002). Towards a theory of timbre for musical analysis. *Miusicae Scientiae, 6*(1), 23–47.

Instrumental Timbre, Groupings and Texture: Analysis of Form and Expression in Pre-Spectral European Ensemble and Orchestral Music.

Pierre Michel
Department of Music, Strasbourg University 1, Strasbourg, France
pmichel@unistra.fr

Aims

In this paper I aim to show the validity of considering instrumental groupings, and the specific timbre created in the textures as one of the major analytic tools to clarify what emerges for the listener when first experiencing an experimental musical work. Because this kind of music can rarely be understood directly through its pitch-logic (harmony, melody), a part of the understanding and emotional connection comes from the timbre—as an important part of the musical texture—which enables the listener to distinguish different "moments" or sound-qualities of the piece, and then to delve deeper inside each "moment." In this paper, texture and instrumental timbre are treated in a coherent way to analyse and to perceive musical form and morphology in orchestral and ensemble works of the 60s and 70s. Works used to illustrate this approach include the "post-serial" music of Bernd Alois Zimmermann (Cello *Concerto "en forme de pas de trois"* and Gilbert Amy (*D'un espace déployé*), as well as works by Goffredo Petrassi (*Estri)* and György Ligeti (*Chamber Concerto*).

Background information

My research into this topic began when I met György Ligeti in 1979 and studied his music for a series of collected essays. I was fascinated by his orchestral and ensemble music and I translated some of his essays into French. I later went on to work on chamber, ensemble and orchestral pieces by many other composers such as Luigi Dallapiccola, Goffredo Petrassi, B.A. Zimmermann, Klaus Huber, Franco Donatoni, Ivan Fedele, Tristan Murail, and Michèle Reverdy. As an instrumentalist (saxophone player), I have also made special analyses of performance aspects and have made research into specific new instrumental techniques.

Methodology

This research is the result of long experience in analyzing contemporary music. The musical works I discuss in this paper do not readily lend themselves to methods such as set theory or other kinds of methods used for serial music. I focus rather on the sound, the timbre and the energy of the music as a way of interpreting its form. Nobody can really hear and control all the details of these pieces, which is why a more realistic texture-based approach (as defined by the composers György Ligeti, and Tristan Murail, and by the scholars Wallace Berry and Joel Lester) that involves the segmentation of the piece under observation into different sections via its timbre (sound) and dynamic activity (including rhythm, dynamics, etc.) is the method chosen here. This approach focuses on "the changes in degrees of density and textural diversity" (Berry, 1976), taking the instrument groupings— also meaning timbre or tone color—as one of the main factors, together with those defined by Joel Lester: spacing, register, rhythm, but also with other common elements such as melodic writing—with or without figures—heterophony, and polyphonic features. The analysis of "functional successions of textural events" (Berry, 1976) and the observation of different types of musical texture are related to the listening of the pieces and graphic representations. In this sense the point of departure is to consider listening to this kind of music at a global level. This approach could be likened to ethnomusicology or the analysis of electroacoustic music, or even to the so-called "Taxonomic listening": "When we adopt a taxonomic listening intention, we recognize and subtract parts of the music, we compare it to other parts and we look for an overall shape or logical form" (Delalande). But as these pieces of music I have chosen do not fall within the scope of a more traditional musical education, it is necessary that "the music fulfils certain criteria" (Delalande), or—and this is my goal—that in the case of complex music, the musicologist should draw out certain perceptive criteria for the listener to focus on. This kind of musical "mediation" should initially be based by the musicologist on a sort of "back and forth" between the score and the recording. The feedback of the listener, who listens without a score, would then confirm or reject the criteria selected by the musicologist from the point of view of timbres and textures.

Results

I have used these methods for a long time in different kinds of contemporary music pieces, and I know they were close to the ideas of many composers following the major works of Stockhausen, Xenakis and Ligeti. But it seems to be a useful way of understanding general elements of musical form in complex ensemble or orchestral music of other composers too. Offering listening approaches to this kind of music can be one of the basic results, with graphic representation or tables with the main characteristics of each moment of the form, in order to render the formal dynamic of the piece understandable, and how meaning is conveyed through "some textural functions in delineation of form" (Berry) and whether or not listening to the instrumental timbre can reinforce the perception of musical form.

Conclusions

This approach seems to be one possible way among others to understand and appreciate orchestral or ensemble music of this period. Whilst seemingly close to the older "orchestration" approach, it is in fact well adapted to the new ways of composing music after the 1950s and focuses on the trends that emerged before the advent of French "spectral music" (i.e. mainly the two decades up to 1975).

References

Barrière, J.-B. (1991). *Le timbre, métaphore pour la composition*, Paris, Ircam/Christian Bourgois Editeur.

Berry, W. (1976-1987). *Structural functions in music*, New York, Dover Publications.

Boulez, P. (1991). Le timbre et l'écriture. Le timbre et le langage. In J.-B. Barrière (Ed.), *Le timbre, métaphore pour la composition*, Paris, Ircam/Christian Bourgois Editeur.

Chemillier, M. (2001). György Ligeti et la logique des textures, Analyse Musicale vol. 38, pp. 75-85. Delalande, F. (2013). *Analyser la musique, pourquoi, comment ?* Paris, Ina Editions.

Dunsby, J. (2004). Texture. In J.-J. Nattiez (Ed.), *Musiques – Une encyclopédie pour le XXIe siècle. 2 Savoirs musicaux*, Paris, Actes Sud/Cité de la musique.

Grisey, G. (1991). Structuration des timbres dans la musique instrumentale. In J.-B. Barrière (Ed.), *Le timbre, métaphore pour la composition*, Paris, Ircam/Christian Bourgois Editeur.

Grisey, G. (1987). *Tempus ex Machina*: A composer's reflections on musical time, Contemporary Music Review, 1987, Vol. 2 pp. 239-275.

Lester, J. (1989). *Analytic Approaches to Twentieth-Century Music,* New York, Norton.

Ligeti, G. (2001). *Neuf essais sur la musique*. Geneva, Contrechamps.

Ligeti, G. (1960). Metamorphoses of Musical Form, in Die Reihe 7 (Vienna).

Michel, P. (1995). *György Ligeti*, Paris, Minerve.

Michel, P. (1999). Ligeti : *Concerto de chambre* : l'écriture musicale au service d'une dynamique de la forme. In Analyse musicale vol. 34, pp. 31-51.

Michel, P. (2006). *Musiques contemporaines – Approches analytiques 1950-1985*, Paris, Minerve.

Murail, T. (2005). *Models & Artifice – The Collected Writings of Tristan Murail*, Contemporary Music Review, Vol. 24, Parts 2+3, 2005, Issue Editors: Joshua Fineberg and Pierre Michel.

Identifying the Perceptual Dimensions of Musical Instrument Timbre

Lindsey Reymore† and David Huron
School of Music, The Ohio State University, Columbus, Ohio, The United States
† reymore.1@osu.edu

Aims/goals

We aim to develop a model of the perceptual dimensions of musical instrument timbre, that is, musical instrument qualia. This model is then used in an experiment designed to generate perceptual profiles of various musical instruments. Our goal in constructing these perceptual profiles is not only to understand better how people experience timbre and characterize these experiences, but also to create a resource for future studies in timbre and orchestration.

Background information

Different musical instruments do not simply exhibit different musical timbres; they also evoke different associations and phenomenological experiences, or what we call in this study "musical instrument qualia." Informal descriptions of instrument timbres seem to employ stereotypical characterizations, such as "airy" for the flute or "heavy" for the tuba. Previous research has sought to identify dimensions of timbre through listener judgments of paired comparisons (Grey, 1977; Kendall, Carterette, & Hajda, 1999). However, by mapping perceptual dimensions to acoustic rather than auditory categories, the paired comparisons approach does not attend directly to how phenomenological experiences differ for the listener. Investigations into the semantic dimensions of timbre demonstrate that a number of descriptive terms map consistently onto particular acoustic correlates of timbre (e.g. Zacharkakis, Pastiadis, & Reiss, 2014). In a quantitative corpus analysis of texts on orchestration and instrumentation, Zachary Wallmark (2017) found descriptive terms for instrument timbre could be sorted into relatively few distinct categories and demonstrated latent semantic groupings among these categories.

Methodology

Interviews were conducted with 23 professional musicians who were asked to describe 20 different musical instrument sounds. These responses were parsed into component ideas for content analysis. The component ideas were pile sorted independently by the two authors, a process which yielded an initial list of perceptual categories or dimensions of musical instrument timbre. To test the reliability and usefulness of the proposed dimensions, participants are asked to rate musical instrument sounds on these refined dimensions. Principal components analysis is applied to these results to suggest ways in which categories with significant overlapping variance can be collapsed to reduce the number of dimensions, producing a final version of the model. In the last stage of the study, a new pool of participants rates a variety of instrument sounds according to each of the dimensions in the refined model; these results generate perceptual profiles for each of the rated musical instruments.

Results

460 interview responses from 23 professional musicians were parsed into 2,487 unique component ideas, 502 of which occurred multiple times. The researchers independently performed pile sorts of these 502 ideas, producing 59 and 70 categories; through mutual discussion and review, these were reconciled to 77 categories. Results for the remaining stages of the study are forthcoming.

Conclusions

In this project, we derive a model of the dimensions of musical instrument qualia. This model is then used to generate perceptual profiles for musical instrument sounds. The profiles will reveal common ways in which musical instrument timbres are perceived and characterized and will serve as a resource for future studies of orchestration, particularly of the role of the perception and experience of different timbres in orchestration.

References

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America,* 61(5), 1270-177.

Kendall, R.A., Carterette, E.C., & Hajda, J.M. (1999). Perceptual and acoustical features of natural and synthetic orchestral instrument tones. *Music Perception,* 16(3), 327-363.

McAdams, S. (2013). Timbre as a structuring force in music. *Proceedings of Meetings on Acoustics, 19, 035050.*

Wallmark, Z. (2017). Timbre semantics in orchestration: A corpus-linguistic study. *Conference of the Society for Music Perception and Cognition*, San Diego, July 30-August 3.

Zacharakis, A., Pastiadis, K., & Reiss, J. D. (2014). An interlanguage study of musical timbre semantic dimensions and their acoustic correlates. *Music Perception*, 31(4), 339-358.

Towards Timbre Solfège from Sound Feature Manipulation

Thiago Rossi Roque[1†], Rafael Santos Mendes[2]
[1] IA, UNICAMP, Campinas, São Paulo, Brazil
[2] DCA/FEEC, UNICAMP, Campinas, São Paulo, Brazil
[†] t072515@dac.unicamp.br

Aims/goals

This work presents the initial results from a research on a new sound synthesis technique, aiming to sort out the timbre of sound instruments into a set of unidimensional, predictable and recognizable parameters by using sound features. At this point of the research, an analytical analysis/synthesis system for the extraction and manipulation of perceptual spectral sound features have been developed using the Fractal Additive Synthesis technique as basis framework.

Traditional sound synthesis techniques lack from a good mapping from its control parameters to the final resulting sound (Hunt & Kirk, 2000). By using feature manipulation techniques, we believe this problem could be overcome by defining a set of manipulable, orthogonal and unidimensional sound features.
Pierre Schaeffer, in his seminal work "Traité des objets musicaux," proposed an outlook of timbre, in the context of reduced listening, by sorting out this phenomenon using a set of morphological criteria (Chion, 2009). In this work, an analogy between Schaeffer's morphological criteria and the concept of sound features is proposed in a way that the timbre phenomenon could be decomposed by abstractable "pertinent traits," similar to the traditional notion of pitch, duration and intensity.

Background information

Sound features are an important tool of the science called Music Information Retrieval (MIR). Instead of using this technique just for the extraction of information about each sound sample, this project implements a feature modulation system to enable timbre manipulation by the establishment of a link between each sound feature and a perceptual element.

Fractal Additive Synthesis (FAS) is a spectral modeling analysis/synthesis technique similar to Xavier Serra's Spectral Modeling Synthesis (SMS) (Serra, 1989). However, unlike SMS, FAS does not codify each harmonic partial as a perfect periodic ideal signal. It takes into account the natural pseudo-periodicity of tonal signals by performing a fractal, 1/f profile analysis of the sidebands of each harmonic partial (Polotti, 2003). By this approach, FAS is capable to synthesize sounds with far better organicity, giving more emphasis on codifying the stochastic aspect of sound, even using the same amount of stored data.

By its low set of parameters needed to codify periodic/pseudo-periodic signal, FAS has also proven itself a good candidate to be used as the basis framework for the extraction of sound features from pitched sounds, justifying its choice for this project.

Methodology

In this initial stage, just a small set of spectral sound features has been chosen for this research, these would be: Spectral Centroid (SC), Spectral Spread (SS), Even to Odd Harmonic Ratio (EOR), Tristimulus (TR), Mean Harmonic Band Hurst Exponent (MHE) and Harmonic Band Correlation Coefficient (HBCC)[1].

To study the possible understanding of sound features as pertinent traits, we seek to analyze a possible recognizability and previsibility on each feature. This study was carried out in three stages.

[1] Mean Harmonic Band Hurst Coefficient and Harmonic Band Correlation Coefficient are two stochastic features developed specifically for FAS; for detailed definition of these features, see Roque (2017).

In the first one, each feature was extracted from samples of four different instruments (cello, trumpet, oboe and clarinet) playing close pitches. The second stage consists on the analysis of the variation of each feature for a same instrument but along different pitches. These first two stages aim study how the value of some features varies for different samples with different timbre characteristics and along different pitches. The third stage consists on the modulation of the features followed by a careful listening of the synthesized modulated sound. For this last stage, only SC, EOR, MHE and HBCC were used. The main goal of this stage was to seek for a monotonic relationship between the amount of modulation and its related perceptual timbristic aspect.

Results

From the results obtained in the first two stages of this study, many perceptual characteristics could be observed by the extracted features values. As expected, the SC presented an easier to be noticed relation with a perceptual characteristic, strongly related to the notion of brightness. The MHE also presented an easily perceived perceptual characteristic, but, unlike expected, presented a greater relation with the background noise than with the notion of pseudo-periodicity. The TR has been shown to be an interesting feature when associated to the SC for varying pitches.

The third stage of the analysis showed interesting results. All features presented a monotonic relation between the amount of modulation and its related feature value. Although highly non-linear, this monotonic relation expresses a good indication of a possible previsibility. On the other hand, except for the SC, it was possible to obtain sounds from different instruments showing a same feature value but with diverging perceptual characteristic associated with this feature. A detailed presentation of these results can be found in Roque (2017).

Conclusions

This first set of results on the sort out of the timbre phenomenon into a set of unidimensional, predictable and recognizable parameters shows a promising possibility for this usage of sound features. Among studied features, SC presented itself as the best candidate for an abstractable pertinent trait. There is still work to be done, mainly by expanding this analysis/synthesis system for other perceptual features, including temporal ones.

Following recent researches on sound features allied to deep learning and autoencoders, this machine learning approach seems promising for the search of new features towards this morphological criteria analogy. Nevertheless, there is still an important semantic gap to be overcome about the real relation between each sound feature and its associated perceptual aspect. Only massive listening tests with a large number of listeners might lead to final conclusions.

References

Chion, M. (2009). *Guide to sound objects: Pierre Schaeffer and musical research*. Trans. J. Dack and C. North.

Hunt, A., & Kirk, R. (2000). Mapping strategies for musical performance. In *Trends in gestural control of music* (pp. 231-258). Paris: Ircam.

Polotti, P. (2003). *Fractal additive synthesis: spectral modeling of sound for low rate coding of quality audio* (Doctoral dissertation). École Polytechnique Fédérale de Lausanne, Lausanne, Swiss.

Roque, T. R. (2017). *Extração e Modulação de Descritores Sonoros a Partir da Síntese Aditiva Fractal* (Master dissertation). State University of Campinas, Campinas, Brazil.

Who Was That Timbre I Saw You With?

Miller Puckette[1†], Kerry Hagan[2]
[1] Music Department, University of California, San Diego, California, USA
[2] Department of Computer Science and Information Systems, University of Limerick, Limerick, Ireland
[†] msp@ucsd.edu

Program

Who Was That Timbre I Saw You With? 2017 8-10 minutes

Program Notes

The idea of a timbre space has long captivated computer music researchers (Grey, 1977; Wessel, 1979; McAdams, 1995). These researchers essentially propose to project the space of "all possible musical timbres" onto a low-dimensional Euclidean space. Here we propose another, more exploratory approach, where we attempt to span an interesting subset of possible timbres using a highly non-Euclidean source space. Rather than mapping multiple possibilities into a low-dimensional space, we start from the unruly space of possible hand gestures, and we map these gestures back into timbres. Despite limitations, we might be able to span an interesting subset of timbres – the larger and the unrulier, the better.

Using LEAP controllers, we obtain 57-tuples of joint locations to millimeter accuracy in real time. The complexity of timbre requires a greater dimensionality than our system. However, our model is much richer, and the system can be traversed more nimbly, than an array of low-dimensional controllers such as a slider bank.

We want our system to: be highly responsive, generate many static timbres; produce an effective timbre regardless of input; and respond commensurately to changes in the input. Here we propose a two-dimensional dynamical synthesis method that can produce a wide variety of possible timbres (harmonic/inharmonic; smooth/rough; variability of spectral envelope) depending on nine independent parameters.

We program the LEAP controller to report specific spatial relationships of the hand that determine the nine parameters. A reliable set of timbres and gestures can be recreated, but there is more room for serendipity and surprise than planning. As performers, we find this similar to learning a new instrument without the guarantee that it can be mastered.

On the other hand, this uncertain situation makes for more interesting musical potential. So, in order to further reduce the possibility of planning and mastery, we use two performers, one hand apiece, swapping a digit or two so that neither can completely control any one sound.

Acknowledgments

In Memoriam: Robert Wood (1792-1847)

Timbral Hybrids in Philippe Hurel's *Leçon de choses*

Robert Hasegawa
Department of Music Research, Schulich School of Music of McGill University, Montreal, QC, Canada
robert.hasegawa@mcgill.ca

Aims/goals

Philippe Hurel (1955–) has been described as a "post-spectral" composer, building on the achievements of the L'Itineraire group (Gérard Grisey, Tristan Murail, Hugues Dufourt) while developing a unique personal style that combines classic spectral procedures with a wide range of external influences (Pousset 2000). His music deploys diverse and vividly colored harmonic materials, with a density of rhythmic activity and polyphonic development comparable to the most complex serial works of the 1950s and 60s. Hurel has observed that composers of his generation, faced with the legacy of spectral music, have sought to "build coherence with many different or contradictory elements," a dialectical approach that "goes beyond the original spirit of spectralism" (Hurel 2015). Through a detailed analytical study of the 1993 work *Leçon de choses,* for thirteen musicians and electronics, my paper will explore how Hurel's use of "timbral hybrids" allows the reconciliation of drastically contrasting elements: recorded concrete sounds, computer-manipulated audio files, artificial frequency modulation spectra, MIDI-controlled samples of acoustic instruments, "instrumental synthesis" chords, and abstract intervallic series and patterns.

Background information

Titled after a 1975 novel by Claude Simon, *Leçon de choses* not only invokes the "object lesson" (*leçon de choses*), a pedagogical approach first developed in the nineteenth century using concrete objects to introduce more abstract ideas, but also suggests through wordplay "the sound of things" (*le son de choses*). Hurel's work uses seven "sound objects" as its primary musical material (Example 1): these include the recorded sounds of a tart pan and a tape casing struck with a percussion mallet, the sampled sound of a double bass, and a variety of synthesized sounds, some combining instrumental sounds with complex frequency modulation spectra. The objects are differentiated by their gestural profiles ("objet scalaire," "avalanche") as well as their timbres (represented as frequency spectra in Example 2). Most often, the objects are presented simultaneously by the electronics and the acoustic ensemble, with notes assigned to specific instruments through the spectral composition technique of instrumental synthesis. As Catherine Tognan (1994) describes the work's evolution, these initially presented objects are "eroded" into a wash of undifferentiated polyphony in which the live ensemble is mirrored by a MIDI "fake orchestra," then gradually "reconstituted" into their original form.

Methodology

My research on Hurel draws on existing literature (Tognan 1994), sketch and manuscript materials published by Hurel (Hurel and Lelong 1994a), independent score and audio analysis, and interviews with Hurel carried out in 2015 and 2016. I propose the idea of "timbral hybrids" to describe several different intermediate states between clearly defined and recognizable timbres. The idea has close parallels with Albert Bregman's description of perceptual "chimeras," Cornelia Fales's "timbral anomalies," and Gérard Grisey's description of timbre-chords as "something hybrid for our perception, no longer really a timbre, without yet being a true chord, a sort of mutant of contemporary music." Timbral hybrids emerge from the "crossing" of (1) single-source real-world timbres and their imitation in the ensemble through instrumental synthesis, (2) electronically synthesized and acoustic sounds, (3) real instruments of the ensemble and their counterparts in the "fake orchestra," (4) sonorities fusing simultaneous events into a single percept and polyphonic "clouds," and (5) abstract FM synthesis spectra and spectra of real objects and instruments. All of these "crossings" can produce intermediary states including aspects of both of the "parent" phenomena.

Results

A close analysis of *Leçon de choses* reveals the important role of timbral hybrids at all levels of musical design, from the smallest details to the large-scale determinants of form. Hurel's use of hybrids allows a variety of "trompe l'oreille" effects and illusions affecting the listener's analysis of auditory streams. Tracing the creation and development of the various types of hybrids in the work offers a clearer understanding of formal design in post-spectral music, particularly the complex relationship of spectral compositional techniques with perceptual outcomes.

Conclusions

This project suggests the value of new analytical concepts such as the "timbral hybrid" in discussions of spectral and post-spectral music. Particularly characteristic of Hurel's approach are (1) a mix of "instrumental synthesis" techniques with other harmonies derived by frequency modulation, (2) the use of interpolation between radically different materials to produce intermediary hybrids, (3) an emphasis on recognizability, the ability of the ear to identify particular musical objects. The form designed by Hurel makes the recognizability of objects particularly striking, as the seven musical objects are presented alongside various distorted and "fake" forms (in both electronics and instruments) that blur the boundaries of perceived aural reality. Hurel's music is quite unique in the post-spectral school in its frequent use of strict, self-imposed compositional constraints (on pitch, rhythm, formal durations, patterns of recurrence) to control structure and stimulate inventiveness. More broadly, this project seeks to develop and demonstrate new conceptual tools for the analysis of post-spectral music, particularly mixed acoustic/electronic works.

Acknowledgments

This research is made possible by support from the Fonds de recherche du Québec – Société et culture (FRQSC) and the Social Sciences and Humanities Research Council (SSHRC).

References

Hirsbrunner, T. (2013) "Autour du spectralisme : prolongements, critiques, voies parallèles." In Nicolas Donin and Laurent Feneyrou, *Théories de la composition musicale au 20e siècle* (pp. 1603–1625). Lyon: Symétrie.

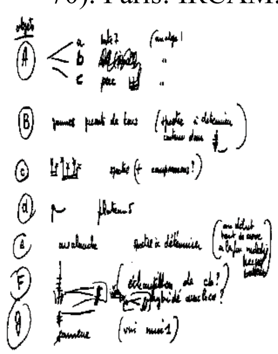
Hurel, P. (2015) "La musique spectrale…à terme !" Retrieved from http://www.philippe-hurel.fr/musique_spectrale.html

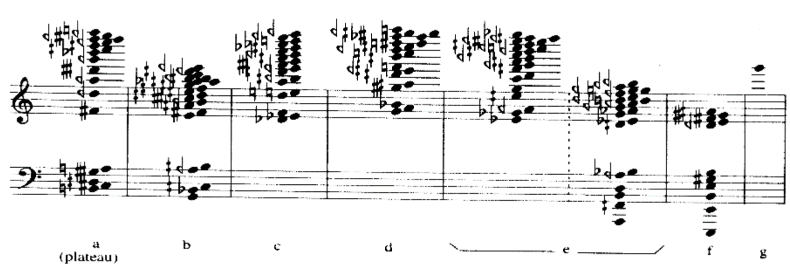
Hurel, P. and Lelong, G. (1994a) "*Leçon de choses* : Étapes." In *Philippe Hurel,* Cahiers de l'IRCAM 5 (pp. 71–86). Paris: IRCAM.

Hurel, P. and Lelong, G. (1994b) "Entretien avec Philippe Hurel." In *Philippe Hurel,* Cahiers de l'IRCAM 5 (pp. 7–41). Paris: IRCAM.

Pousset, Damien. (2000) "The Works of Kaija Saariaho, Philippe Hurel and Marc-André Dalbavie—Stile Concertato, Stile Concitato, Stile Rappresentativo." *Contemporary Music Review, 19/3,* 67–110.

Tognan, Catherine. (1994) "*Leçon de choses* : analyse." In *Philippe Hurel,* Cahiers de l'IRCAM 5 (pp. 43–70). Paris: IRCAM.





Example 1 (Hurel & Lelong 1994a) Example 2 (Tognan 1994)

Timbre-as-Heard: Spectromorphological Considerations.

Lasse Thoresen
The Norwegian Academy of Music
Lasse.thoresen@nmh.no

Aims/goals

The eventual aim of my project within the larger ACTOR project is to throw light on the *exosemantic* implications of timbre-as-heard (*H-timbre*) – exosemantics being the field of extra-musical meanings associated with the sound and the music by the listener. I will develop further a phenomenological approach that uncovers the listeners' *semioses* involved in the act of attributing a meaning to H-timbre – semiosis being the specific act by which the mind establishes a correlation between a percept (a signifier) and its meaning (a signified). Eventually tacit knowledge relevant for the understanding of H-timbre will be made explicit and may thus be shared with others in the contexts of artistic and musicological research. However, a number of initial specifications and definitions must first be made clear, and the purpose of the present paper is to prepare the ground for the semiotic study at a later stage.
The study of timbre as a pure acoustic phenomenon (i.e., described through spectral analyses, spectrograms etc.), though relevant for the composer and the scientist, is not of direct relevance for the understanding the *musical* implications of timbre, the latter being the primary purpose of my study. My approach will serve to contextualize isolated findings in acoustics and perception psychology into a verbal discourse that refers directly to a music-as-heard.

Background information

This project is a continuation of the Aural Sonology Project (www.auralsonology.com).

Methodology

The basic method employed is observation (including explicit phenomenological reflection on the intrinsic *intentionalities* involved in musical listening and interpretation) combined with structural correlation of observed entities that eventually will be the subject of semiotic interpretation. Music is observed on four ontological levels: Articulation level one: *The sound-object* (approached through the intentionality reductive listening. The study of sounds-objects is the main field of spectromorphology); Articulation level two: *Compound sound patterns* (taxonomic listening level two); Articulation level three: *form-building patterns* (taxonomic listening level three); *Semantic level* (interpretations based on the *semioses comparison, causal inference, connotation,* and *recognition*, and their concatenation into *semiotic chains*). By accessing the anatomy of the sound-object (Articulation level one) through reductive listening one can learn to appreciate its spectral and energetic aspects that often form the basis of H-timbre. H-timbre is related through extension to texture-as-heard (Articulation level two), and will be the subject of a later paper.

Results

The result will be a precise and transparent study of the multiple aspects of H-timbre in its sounding locus, its relationship to musical structure, and its exosemantic interpretations.

Conclusions

The concept of timbre as used to describe elements in traditional, interval-based music, cannot be defined as an isolated percept. In order to realize the full implications of the concept of timbre, it has to be studied in terms of timbre-as-heard in the context of a sounding, musical discourse. H-timbre then enters into relationship with its defining opposite: structural values. Timbre understood in this way will be more or less equivalent with what P. Schaeffer in his *Traité* used to call "character." *Structural values* vs. *character* may parallel the relationship between verbally communicated messages vs. meta-communicated indices (voice sonority, intonation, accompanying gestures). Moreover, in sound-based (as opposed to interval-based) music characters and values will be shown not have the same musical and structural functions.

References

Chion, M. (1983). *Guide des objets sonores. Pierre Schaeffer et la recherche musicale*. Paris. Buchet / Chastel. Institut national de la communication audiovisuelle. (For translation, see Dack and North 1999.)

Chion, M. (1988). Du son à la chose. Hypothèse sur l'objet sonore. *Analyse musicale* 2.

Schaeffer, P. (1966). *Traité des objets musicaux*. Paris. Éditions du Seuil (2nd edition, 1977).

Smalley, D. 1994. Defining Timbre, Refining Timbre. *Contemporary Music Review*, 10(2): 35–48.

Thoresen, L. (2007). Spectromorphological Analysis of Sound Objects: An adaptation of Pierre Schaeffer's Typomorphology. *Organised Sound,* 12(2): 129-141

Thoresen, L. (2009). Sound-objects, Values, and Characters in Åke Parmerud's Les Objets Obscurs. With the assistance of Andreas Hedman. *Organised Sound,* 14(3). Cambridge University Press.

Thoresen, L. (2010). Form-building gestalts and metaphorical meaning. With the assistance of Andreas Hedman. *Organised Sound,* 15(2). Cambridge University Press.

Thoresen, L. (2012). Exosemantic Analysis of Music-As-Heard. *Proceedings of the Electroacoustic Music Studies Conference Meaning and Meaningfulness in Electroacoustic Music*, Stockholm, June 2012. www.ems-network.org

Thoresen, L. (2015). Emergent Musical Forms. Aural Explorations. *Studies in Music from the University of Western Ontario. Vol 24.* London, Ontario.

Wagner's Associative Orchestration in the *Orchesterskizze* for *Tristan und Isolde*

Julie Anne Nord
Don Wright Faculty of Music, University of Western Ontario, London, Ontario, Canada
jheikel@uwo.ca

Aims/goals

In *Oper und Drama* (1851), Richard Wagner portrays the voice of the orchestra as an integral part of the music's role in the drama, calling it the "pure organ of the Feeling, [which] speaks out the very thing which Word-speech [i.e., the libretto] in itself can *not* speak out,… *the Unspeakable*,… [which] is shewn plainly enough by the Instruments of the orchestra themselves, whereof each for itself, and infinitely more richly in its changeful union with other instruments, speaks out quite clearly and intelligibly" (Wagner, 1893, p. 316). Though Wagner avoided engaging in specific analysis of any of his works, his confidence in the orchestra's expressive capacity suggests that orchestration and timbre should not be as overlooked as it has been in recent scholarship. My research examines what I term Wagner's "associative orchestration"—the use of orchestral colour for dramatic means—particularly in works or revisions he completed after the theoretical writings from his time of exile in Zurich (1849–1858). Specifically, this paper uses Wagner's own marginalia in his orchestral sketches for *Tristan und Isolde* (1857–1859) to determine the moments in the drama for which Wagner had particular orchestral timbres in mind even before he completed his drafts for full orchestra.

Background information

Many of the individual musical features of Wagner's works have been the focus of research into how his music portrays the drama, especially those of the leitmotif, key structures, and musical form. Wagner's use of leitmotifs has driven the predominance of motivic-based analysis in studies by Thomas S. Grey (1996, 2008), F. E. Kirby (2004), and Matthew Bribitzer-Stull (2007, 2015), among others. Similarly, Wagner's use of form and symbolic use of key structures has been a popular topic of discussion (Bailey, 1977; Grey, 1988; Lorenz, 1966; Newcomb, 1981; Petty, 2005). Despite the growth of research in these related areas, Wagner's orchestration has remained largely unexplored. In light of this scholarship, I argue that Wagner uses timbre expressively in his mature works, which contributes to music's role in the drama and his concept of *Gesamtkuntswerk*.

Methodology

My analysis of Wagner's use of timbre in *Tristan* draws on my study of Wagner's theoretical writings from his Zurich period where he discusses the role of the orchestra in his dramatic works; orchestral treatises from the time—particularly Berlioz's 1844 *Grand traité d'instrumentation et d'orchestration modernes*, which Wagner held in his library; and Wagner's sketches housed at the Richard Wagner Nationalarchiv und Forschungsstätte in Bayreuth, Germany.

Robert Bailey (1970) identified three stages in Wagner's compositional process for the music of *Tristan*: a *Kompositionskizze* in which the composer drafted the vocal settings of his libretto; an *Orchesterskizze* which shows his initial elaboration of the orchestral accompaniment on 2–4 staves; and finally, a complete draft. My interest is in the middle stage, the *Orchesterskizze*, which contains not only the full vocal parts with the orchestra represented in reduction, but also marginalia where Wagner indicates his timbral intentions more fully. The nearly 130-page draft contains only 25 such marginalia, most of which coincide with major dramatic moment or important leitmotifs. My analysis considers these moments and leitmotifs within the larger drama, studying where Wagner retained these timbral colours in his final scores, how they evolve as the leitmotifs are repeated in the course of the drama, and how they function rhetorically or expressively.

Results

My research positions timbre as a musical element that functions associatively in a similar way to Robert Bailey description of associative tonality (1977) or Matthew Bribitzer-Stull's characterization of the leitmotif as associative (2015). In *Tristan*, timbre joins these oft-studied elements to enable a deeper reading of Wagner's drama and its characters.

Conclusions

My analysis of Wagner's orchestration for *Tristan* is part of my larger study of timbre in Wagner's mature works. Just as Bailey notes a shift in Wagner's use of tonality from expressive to associative, I anticipate a similar change in his orchestral techniques, where instruments are used not only in mimetic or topical ways, but also associatively—where timbre calls to mind events or dramatic themes through association over the course of the drama. Furthermore, my work on timbre has the potential to inform further studies on timbre as a rhetorical and expressive element in music, especially in media, like film music, which are already known to draw upon Wagner's orchestral language.

Acknowledgments

I wish to acknowledge the support of the Richard Wagner Nationalarchiv und Forschungsstätte in Bayreuth, Germany and their archival librarian, Kristina Unger. *My research is also supported by the Social Sciences and Humanities Research Council of Canada.*

References

Bailey, R. (1970). *The genesis of "Tristan und Isolde" and a study of Wagner's sketches and drafts for the first act*. Princeton: Princeton University.

Bailey, R. (1977). The structure of the *Ring* and its evolution. *19th-Century Music, 1,* 48–61.

Bailey, R. (Ed.). (1985). *Richard Wagner: Prelude and transfiguration from "Tristan and Isolde*." New York: W. W. Norton.

Berlioz, H., & Strauss, R. (1948). *Treatise on instrumentation, enl. and rev*. (T. Front, Trans.). New York: E. F. Kalmus.

Bribitzer-Stull, M. (2007). Naming Wagner's themes. In M. Bribitzer-Stull, A. Lubet & G. Wagner (Eds.), *Richard Wagner for the new millennium* (pp. 91–110). New York: Palgrave MacMillan.

Bribitzer-Stull, M. (2015). *Understanding the leitmotif: From Wagner to Hollywood film music*. Cambridge: Cambridge University Press.

Dahlhaus, C. (1979). *Richard Wagner's music dramas*. (M. Whittall, Trans.). Cambridge: Cambridge University Press.

Dolan, E. I. (2013). *The orchestral revolution: Haydn and the technologies of timbre*. Cambridge: Cambridge University Press.

Grey, T. S. (1988). Wagner, the overture, and the aesthetics of musical form. *19th-Century Music*, *12*, 3–22.

Grey, T. S. (1995). *Wagner's musical prose: Texts and contexts*. Cambridge: Cambridge University Press.

Grey, T. S. (1996). *wie ein rother Faden*: On the origins of *Leitmotif* as critical construct and musical practice. In I. Bent (Ed.), *Music theory in the age of romanticism* (pp. 187–210). Cambridge: Cambridge University Press.

Grey, T. S. (2008). Leitmotif, temporality, and musical design in the *Ring*. In T. S. Grey (Ed.), *The Cambridge companion to Wagner* (pp. 85–114). Cambridge: Cambridge University Press.

Kirby, F. E. (2004). *Wagner's themes: A study in musical expression*. Warren, Michigan: Harmonie Park Press.

Lorenz, A. (1966). *Das Geheimnis der Form bei Richard Wagner.* Tutzing, Germany: H. Schneider.

Newcomb, A. (1981). The birth of music out of the spirit of drama: An essay in Wagnerian formal analysis. *19th-Century Music*, *5*, 38–66.

Petty, J. C. (2005). *Wagner's lexical tonality*. Lewiston: The Edwin Mellen Press.

Wagner, R. (1893). *Oper und Drama*. In *Richard Wagner's prose works.* (pp. 1–376). (W. A. Ellis, Trans.). New York: Broude Bros.

Wessel, D.L. (1979). Timbre space as a musical control structure. *Computer Music Journal, 3*(2), 45-52.

Timbre and Semantics in Sound Mass Music

Jason Noble†, Max Henry, Etienne Thoret, Stephen McAdams
Centre for Interdisciplinary Research in Music Media and Technology, Schulich School of Music, McGill University, Montreal, QC, Canada
† jason.noble@mail.mcgill.ca

Aims/goals

Listeners frequently associate musical experience with extramusical images, ideas, concepts, and yet not only has this phenomenon been little studied, it has been *actively suppressed* by formalist music theory and the related ideology of "absolute" music. A possible explanation is that the musical parameters traditionally considered the most hierarchically important—stable pitch and synchronous rhythm—are idiosyncratic to music and rarely encountered in nonmusical auditory experience. But if we expand our understanding of what is essentially "musical" in musical sound to include parameters that have traditionally been considered incidental, secondary, or decorative—especially timbre—the basis for meaningful association with extramusical domains may become clearer. "Sound mass," a branch of late 20th and early 21st century music that diminishes the perceptibility of individual sound events in order to focus attention on global properties of the massed totality, provides ideal material for the study of relations between timbre and semantic associations. Sound mass reorients erstwhile hierarchies of musical parameters, such that timbre is a primary shaping force of musical gesture and identity while pitch and rhythm are perceptually attenuated. In this study, we examine relations between timbral properties of excerpts of sound mass music, revealed through audio descriptor analysis, and semantic association, revealed through a pair of empirical experiments in which participants rated the same excerpts along batteries of semantic scales.

Background information

Semantic associations with music are notoriously variable, but mounting evidence suggests they are not necessarily arbitrary. In an unpublished pilot study, Noble and Kazazis (2017) found that the semantic associations of *brightness*, *noisiness*, and *density* are significantly correlated with distributions of spectral energy. Wallmark (in press) claims that "semantic practices constitute a vital component in the transmission of discursive knowledge about timbre," and suggests that "we cannot help but to perceive and cognize musical timbres through the mediating filter" of culturally and historically situated semantic associations. He found that timbres were consistently described in orchestration treatises using semantic categories for *affect*, *matter*, *metaphor*, *mimesis*, *action*, *acoustics*, and *onomatopoeia*, and noted that "historical convention has designated some descriptors as better fits than others." Huovinen and Kaila (2015) point out that "listeners' making sense of music by extrinsic means does not necessarily require ... 'isomorphic' relations." They provisionally accept Nicholas Cook's (2001) idea of attribute selection as a basis for cross-domain mappings between musical experience and extramusical associations but emphasize that the bases for such mappings are not always homologous: they may also be *topical*. Listeners may relate musical experience to extramusical domains on the basis of qualitative similarities between music and other domains (e.g., "bright" timbres may be associated with bright light), or they may do so with learned, culturally contingent musical *topics*. Topics relate musical styles, instruments, and prototypical gestures to situations in which they are usually produced or experienced (e.g. "horn fifths" are a sign of the pasture or the hunt; the distorted electric guitar is a sign of rock and roll and its associated imagery). Margulis (in press) demonstrated that topical associations between orchestral music and film imagery are pervasive and often surprisingly consistent between listeners. Unfamiliar styles of contemporary music such as sound mass are frequently paired in movies with horrifying, fantastical, and futuristic tropes, so we may expect to find topical associations along those lines in listeners' interpretations of such music.

Methodology

Thirty-eight participants listened to 40 excerpts (average duration: 15 s) of postwar music featuring sound mass, rating each along three blocks of semantic scales with terms used by composers and theorists to describe sound mass. Categories in the first block reflected definitions of sound mass: *fusion, density, complexity, homogeneity* (see Douglas, Noble, and McAdams, 2016). Categories in the second block were in adjectival form: *volatile, atmospheric, busy, static, formless, impenetrable, voluminous, kaleidoscopic*. Categories in the third block were in nominal form: *gas, liquid, solid, clouds, wind, water, webs, galaxies, crystals, machinery, herds/crowds/swarms*. Participants also had the option of selecting "questionable relevance" for each category, and to add their own categories if they wished to report an association not represented in the list. Later, 20 participants completed a follow-up experiment with the same experimental design and stimuli but with the grammatical forms of the categories in Blocks 2 and 3 reversed (e.g., *volatile* was replaced by *volatility*; *gas* was replaced by *gaseous*). Stimuli were initially analyzed subjectively (i.e., their properties of rhythm, pitch, dynamics, timbre, and gesture were described by the researchers based on subjective listening). These analyses were used to interpret the data collected in the experiments. A more acoustically precise description of the stimuli using audio descriptors will enable semantic associations reported by listeners in the experiments to be linked more systematically to timbral properties.

Results

Participants were significantly consistent in their semantic ratings of these stimuli, and the ratings could often be plausibly explained in terms of musical properties including timbre (e.g., *volatile* was associated with noisy and unstable timbres, *atmospheric* was associated with vocal timbre). Principle Component Analysis revealed five underlying principle components which also seem to be timbrally driven (e.g., PC1 involved heterogeneous timbres and granular textures, PC4 involved loud dynamics and spectral saturation over a broad compass). In order to further understand semantic association with timbre an acoustic analysis has been done based on classical audio descriptors (Peeters et al., 2011). Preliminary tests suggest that spectral flux correlates moderately with 'formless' and 'machinery'; further tests will focus on quantifying timbral heterogeneity in complex musical textures and on evaluating correlations with semantic categories. Participants' self-reported categories suggested topical associations from science fiction, horror, fantasy, and fairy tale movies, including affects (e.g., *creepy*), objects (e.g., *lasers*), and very specific tropes (e.g., *Cinderella's evil stepmother*).

Conclusions

The findings of this study support our hypotheses that extramusical associations are not arbitrary and can be explained in terms of relations between musical attributes (especially timbre) and attributes of other domains. Those relations may be based in similarity/homology/isomorphism between musical and extramusical domains, or they may be based in culturally determined topical associations rooted in contexts of production or reception (such as movies). This study contributes to the growing body of evidence that underscores the pervasiveness of extramusical associations across many styles of music, and implicitly challenges formalist hierarchies and doctrines that have historically minimized the role of timbre and excluded extramusical associations from music theory and analysis.

References

Cook, N. (2001). Theorizing musical meaning. *Music Theory Spectrum, 23* (2), 170-95.

Douglas, C., Noble, J., & McAdams, S. (2016). Auditory scene analysis and the perception of sound mass in Ligeti's *Continuum*. *Music Perception, 33* (3), 287-305.

Huovinen, E. & Kaila, A-K. (2015). The semantics of musical topoi: An empirical approach. *Music Perception, 33* (2), 217-43.

Margulis, E. H. (in press). An exploratory study of narrative experiences of music. *Music Perception*.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The Timbre Toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, 130(5), 2902-2916.

Wallmark, Z. (2018). A corpus analysis of timbre semantics in orchestration treatises. *Psychology of Music*, http://journals.sagepub.com/doi/10.1177/0305735618768102.

Examining the Influence of Tone Inharmonicity on Felt Tension and Timbral Semantics

Asterios Zacharakis†, Konstantinos Pastiadis
School of Music Studies, Aristotle University of Thessaloniki, Greece
† aszachar@mus.auth.gr

Aims/goals

This piece of research aimed to investigate the relationships between tone inharmonicity, timbral semantics and perceived musical tension. A connection between the luminance-texture-mass (LTM) model of timbral semantics (Zacharakis et al., 2014; Zacharakis & Pastiadis, 2016) and the concept of musical tension—which is not often studied under the prism of timbre—was pursued.

Background information

Recent research has presented evidence that inharmonicity is a timbre attribute that contributes to inducing musical tension (Farbood & Price, 2017). The link between tension and inharmonicity in this study was examined using static tones with fixed amounts of inharmonicity. Zacharakis et al. (2014) also found a relationship between inharmonicity and timbral semantics. However, the results on which of the identified semantic dimensions (namely luminance, texture and mass) was more closely related to inharmonicity were inconclusive. In addition, evidence concerning how time-varying inharmonicity can relate to both tension and the LTM model of timbral semantics is lacking.

Methodology

Thirty-five musically trained listeners (12 male, mean age: 23, average years of musical practice: 13.5) were presented (through a pair of PreSonus HD7 headphones) with two synthesized tones (a sawtooth wave consisting of 30 harmonic partials and a square wave consisting of 15 partials) of two different pitches (220 & 440 Hz) and of time-varying inharmonicity. The inharmonicity was varied either through a random positive/negative displacement of each partial as a percentage of its harmonic frequency (with a maximum of 4%) or through a positive exponential displacement that affected higher partials more heavily according to the stiff piano string equation (Fletcher, 1964). Inharmonicity linearly increased towards the specified maximum inharmonincity level and then decreased back to zero within a total duration of 30 seconds.[1] During repeated stimulus presentations, participants were asked to provide time-varying tension, luminance, texture and mass profiles for each tone (8 tones x 4 qualities = 32 profiles per participant) by horizontal hand movements on a Kensington Expert wireless trackball mouse and a custom designed LabVIEW GUI that sampled the trackball's horizontal axis coordinate every 5 milliseconds and offered participants a real-time visualisation of the profiles they were creating.

Analysis

Raw responses were subsampled by calculating the mean value over adjacent non-overlapping rectangular time windows (.5 secs = 10 samples). The resulting 60-sample time series were subjected to first-order differentiation and replacing positive/negative values with 1 and -1, respectively. Each participant's data were then normalized within each quality by his/her maximum rating on this particular quality. The time series were next integrated and smoothed using a cubic spline interpolant. Finally, linear trends were removed from each individual participant's time series to ensure stationarity (Dean & Bailes, 2010). The processed time series were averaged over every stimulus and each of the four qualities under study. Figure 1 presents the mean profiles along with their 95% confidence intervals.

[1] The stimuli can be found online at: http://ccm.web.auth.gr/timbreandtension.html.

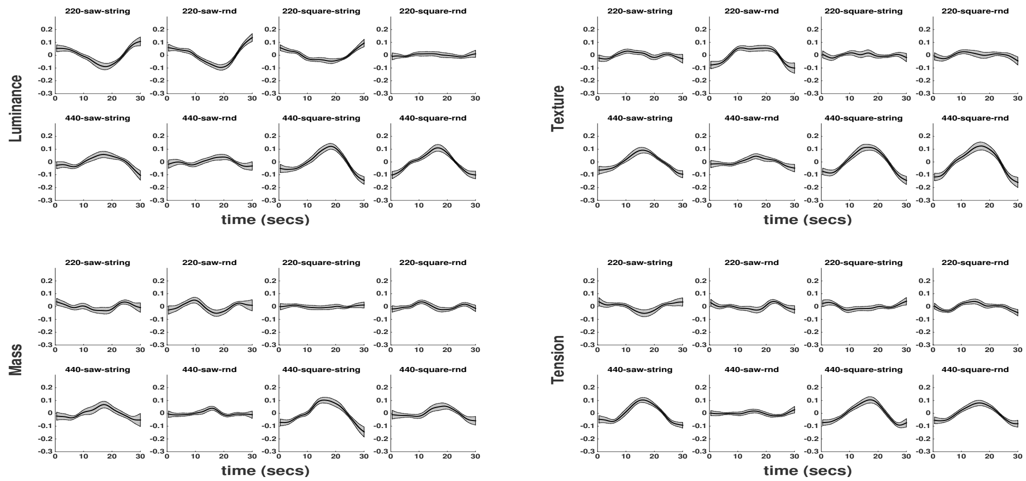


Figure 1. Mean temporal profiles of luminance, texture, mass and tension. Grey fills represent the 95% CI. Top rows: 220 Hz, bottom rows: 440Hz.

Conclusions

- The participants' agreement as represented by the Cronbach's Alpha was high for luminance, texture and tension (.87, .86, .82) but lower for mass (.66) in accordance with past evidence (Zacharakis et al., 2014; Zacharakis & Pastiadis, 2016).
- Statistically significant changes in tension can be induced through continuous manipulation of inharmonicity with a strong positive relation for higher F0s, thus expanding existing evidence for static sounds (Farbood & Price, 2017).
- Stepwise multiple regressions showed that stronger increases in tension (evident in stimuli 4, 5, 7 & 8) were more strongly predicted by texture (roughness). For the remaining four stimuli (1, 2, 3 and 6), tension changes were predicted through a synergy of texture with either luminance or mass, or through a synergy of all three.

Acknowledgments

This study was supported by a post-doctoral scholarship issued by the Greek State Scholarships Foundation (grant title: "Subsidy for post-doctoral researchers," contract number: 2016-050-050-3-8116), which was co-funded by the European Social Fund and the Greek State.

References

Dean, R. T., & Bailes, F. (2010). Time series analysis as a method to examine acoustical influences on real-time perception of music. *Empirical Musicology Review, 5(4),* 152-175.

Farbood, M. M., & Price, K. C. (2017). The contribution of timbre attributes to musical tension. *The Journal of the Acoustical Society of America*, *141*(1), 419-427.

Fletcher, H. (1964). Normal vibration frequencies of a stiff piano string. *The Journal of the Acoustical Society of America*, *36*(1), 203-209.

Zacharakis, A., & Pastiadis, K. (2016). Revisiting the Luminance-Texture-Mass Model for Musical Timbre Semantics: A Confirmatory Approach and Perspectives of Extension. *Journal of the Audio Engineering Society*, *64*(9), 636-645.

Zacharakis, A., Pastiadis, K., & Reiss, J. D. (2014). An interlanguage study of musical timbre semantic dimensions and their acoustic correlates. *Music Perception*, *31*(4), 339-358.

See Me, Feel Me: Timbre as a Multi-modal Experience

Simon Zagorski-Thomas
London College of Music, University of West London, London, UK
simon.zagorski-thomas@uwl.ac.uk

Aims/goals

The aim of the project is to explore the experience of timbre in popular music through the analytical potential of a theoretical framework about the embodied and multi-modal nature of perception, cognition and action in the case study of The Who's 1969 album *Tommy.* As the framework focuses on the perception of timbre as a dynamic, activity-based process, a further aim is to explore the potential for creating visual and/or interactive animations rather than text as the research output of the analytical process.

Background information

Alex Case has summarized the definition of timbre as everything in an audio signal "not covered by loudness and pitch" (Case, 2012, p. 111). This idea of sound and music as being made up of a series of discrete features (of which timbre is one) is a manifestation of the human ability to categorize. Trying to reconcile various aspects of the ecological approach to perception (Clarke, 2005; Gibson, 1979; Moore, 2012; Smalley, 2007; Zagorski-Thomas, 2018) and embodied cognition (Feldman, 2008; Iacoboni et al., 2005; Lakoff, 1990; Lakoff & Johnson, 2003) with the multi-modal nature of the inextricably entwined world of perception and action (McGurk & MacDonald, 1976) led me to the notion that multiple modes of perception were a logical pre-requisite for categorical thought. The idea that what I see and what I hear, although they are fundamentally different sensations, are aspects of "the same" phenomenon requires some cognitive mechanism for marking the different sensations as "the same." This, along with the idea that knowledge and understanding emerge from the experience of doing (Ingold, 2013; Noë, 2004), suggest that categorization—the notion of "things" and characteristics—is extracted from these basic building blocks of activity-based understanding. Under this model timbre is experienced and understood as the sound of something happening somewhere and that the basis of this and musical meaning in general is either directly empathic or based on a metaphorical connection with previous experience. Either we hear the sound of some activity that we recognize as something we have engaged in and we understand its meaning based on our direct previous experience (empathy) or we have to find some features of it that are similar to our previous experiences and use these experiences to hypothesize an interpretation (metaphor). This combination of the ecological and embodied approaches has parallels with Middleton's three categories of musical meaning: gesture, connotation and argument (Middleton, 1990).

Methodology

Given this theoretical model, the research problem was to develop a methodology for its application to a case study: The Who's 1969 album *Tommy*. Using a series of audio and audio-visual extracts from *Tommy,* the empathic and metaphorical characteristics of various timbres will be explored from mono-modal (audio or visual only), bi-modal (audio-visual) and multi-modal (participatory) perspectives. This involves a series of practical experiments about the gestural and connotative features of particular timbres that demonstrates both the near-universal 'meaning' of certain sounds and the idiosyncratic and unique interpretations that we nevertheless produce. Techniques of hypothetical (or real) substitution (Tagg, 2000) re-enactment (Meynell, 2017) and spectral analysis will also be used where appropriate.

Results

The presentation will use case study examples to explore a range of guitar, bass, drum and vocal timbres through the multi-modal nature of our experience of musical sound. Through a process of demonstration, comparison and schematic representation, the presentation of 'results' will aim to expose the cognitive mechanisms through which timbre is understood as having both broad, species-wide features as well as narrower community, group and individual ones.

Conclusions

The presentation demonstrates the usefulness of this method as an analytical tool and provides a window on an embodied approach to the analysis of timbre. It also suggests a potential for visual analytical animations of music as sound that may have creative / artistic potential as well as an instructional / educational one.

References

Case, A. (2012). *Mix Smart: professional techniques for the home studio*. Taylor & Francis.

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

Feldman, J. A. (2008). *From Molecule to Metaphor: A Neural Theory of Language* (1st MIT Press Paperback Ed). MIT Press.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Psychology Press.

Iacoboni, M., Milnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizolatti, G. (2005). Grasping the Intentions of Others with One's Own Mirror Neuron System. *PLOS Biology*, *3*(3). https://doi.org/10.1371/journal.pbio.0030079

Ingold, T. (2013). *Making: Anthropology, Archaeology, Art and Architecture*. London and New York: Routledge.

Lakoff, G. (1990). *Women, Fire, and dangerous Things: What Categories Reveal about the Mind*. Chicago: University Of Chicago Press.

Lakoff, G., & Johnson, M. (2003). *Metaphors We Live By* (2nd ed.). University Of Chicago Press.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

Meynell, A. (2017). *How Recording Studios Used Technology To Invoke The Psychedelic Experience: the difference in staging techniques in British and American recordings in the late 1960s* (PhD). University of West London, London.

Middleton, R. (1990). *Studying Popular Music* (illustrated edition). Open University Press.

Moore, A. F. (2012). *Song Means: Analysing and Interpreting Recorded Popular Song*. Farnham: Ashgate Pub Ltd.

Noë, A. (2004). *Action in Perception*. MIT Press.

Smalley, D. (2007). Space-form and the acousmatic image. *Organised Sound*, *12*(1), 35–58.

Tagg, P. (2000). *Kojak--fifty Seconds of Television Music: Toward the Analysis of Affect in Popular Music*. Mass Media Music Scholars' Press.

Zagorski-Thomas, S. (2018). The Spectromorphology Of Recorded Popular Music: the shaping of sonic cartoons through record production. In R. Fink, M. L. O'Brien, & Z. Wallmark (Eds.), *The Relentless Pursuit Of Tone: Timbre In Popular Music*. New York: Oxford University Press, USA.

Timbre, Genre, and the Reframing of Analog Processes in Digital Audio Production

Paul Théberge,
Institute for Comparative Studies in Literature, Art and Culture, Carleton University, Ottawa, Ontario, Canada
paul.theberge@carleton.ca

Aims/goals

The aim of this research is to understand the manner in which changes in technology have become central to the modern understanding of timbre and its role in contemporary music. From analog synthesis to spectral analysis, from multi-track recording consoles to digital audio workstations, our ability to conceptualize, visualize, and manipulate the timbral characteristics of music have been enabled and facilitated by technology; at the same time, the effects of these technologies and technical processes have become central to our understanding of the sonic characteristics of various musical genres. While digital technologies have expanded the range, scope and depth of signal processing operations, they have also been used to emulate earlier, analog signal processors and to automate engineering practices associated with the overall sonic shaping of musical productions. Digital technology has thus become an important vehicle for the reification of timbre as a definable and reproducible feature of musical genres.

Background information

As Jonathan Sterne has argued, our understanding of analog technologies has been largely defined in relation to the emergence of digital technology during the 1980s and 90s, and the resulting polarization of analog versus digital has led to a variety of romantic discourses and conceptual misunderstandings. The remediation of analog processes within digital algorithms plays a curious role in this redefinition of the analog: it suggests that analog technologies have a fixed architecture and a relatively predictable set of sonic characteristics—a "sound," or timbre—that can be analyzed and emulated. Furthermore, because analog technologies and signal processing practices are implicated in the production and recording of musical genres of the past, there is the assumption that the sonic characteristics of musical genres are also relatively predictable and can be similarly analyzed and emulated (e.g., through genre-specific mastering algorithms). The use of spectral analysis and other visualization techniques have been critical in the establishment and justification of these assumptions.

Methodology

This research uses methodologies derived from science and technology studies, acoustics, media archeology, discourse analysis, and the musicology of sound recording to explore the relationships between technology and musical culture. Two case studies, dealing with the extreme ends of the recording process will be explored: the first addresses new designs in microphone technologies, and the other, the use of algorithms in the mastering process; digital emulation is central to both case studies.

Results

A better understanding of the relationship between old and new recording processes in a post-digital era.

Conclusions

While timbre in music of the past was primarily associated with orchestration and arranging, on the one hand, and the more or less direct manipulation of instrumental sound through performance techniques, on the other, contemporary notions of timbre have been influenced by ideas drawn from physics and acoustics and, in the practical realm, on the impact of sound recording and signal processing technologies. Because music is ultimately a practical art, changes in technology have become critical not only for our understanding of how timbre is analyzed, manipulated and reproduced in audio production but, also, in how we understand the role of timbre in the definition and reproduction of musical genres.

References

Bolter, J. & Grusin, R. (1998). *Remediation: Understanding New Media*. Cambridge, MA: MIT Press.

Hartmann, W. (2017). *Modern Acoustics and Signal Processing*. Berlin: Springer.

Sterne, J. (2016). Analog. In Peters, B. (Ed.), *Digital Keywords: A Vocabulary of Information Society and Culture* (pp. 31-44). Princeton: Princeton University Press.

Effects of Cross-Modal Stroop Interference on Timbre Perception

Zachary Wallmark
Meadows School of the Arts, Southern Methodist University, Dallas, Texas, USA
zwallmark@smu.edu

Aims/goals

The present study explored the interrelationship between semantic and timbre processing through the use of a modified Stroop task. In investigating "cross-talk" between perceptual modalities, the study posits a mechanistic link between descriptive routines and musical timbre perception. Results suggest that categorical and continuous descriptions of timbre may modulate processing speed.

Background information

Although the semantic structure of timbre has long been debated, there is little consensus concerning the cognitive mechanisms undergirding common descriptive practices. Lexical activation to timbre is typically conceived as *categorical* ("that's a trumpet") or *continuous* ("that sound is bright") (McAdams, 1993), and methods such as semantic differential scales have been used to quantify these relationships. Many of these techniques require explicit verbal evaluation. Here a modified Stroop test was employed in two experiments to explore cross-modal interactions between timbre semantics and perception.

Methodology

In Experiment 1, the effect of categorical word stimuli on timbre perception was explored using a semantic-auditory Stroop task. Participants ($N = 46$, musicians and non-musicians) were presented concurrently with auditory and written word stimuli in congruent, incongruent, and control conditions: clarinet and trumpet signals paired with the words CLARINET or TRUMPET, and violin signal paired with XXXX-type control. Reaction time responses were measured in two trials: attend-sound (in which participants were asked to choose the instrument they heard while ignoring the word on the screen), and attend-word (choose the correct word while disregarding the sound).

Experiment 2 extended this paradigm to test for an effect of semantic-auditory interference in the perception of timbres paired with continuous, cross-modal descriptive adjectives. Participants ($N = 54$) were presented with paired auditory and word stimuli for a reaction time task. Auditory stimuli consisted of 32 natural and synthesizer timbres (1.5s, D#4 pitch) that received the strongest and most consistent ratings on bipolar *luminance* ("dark–bright") and *texture* ("smooth–rough") scales (Zacharakis, Pastiadis, & Reiss, 2014) in a pilot experiment testing 93 signals. Word stimuli consisted of these four adjectives plus a XXXX control in an attend-word only trial.

Results

Experiment 1: Reaction time (log) and accuracy data were subjected to mixed repeated-measures ANOVAs, which revealed significant main effects of trial (attend-word vs. attend-sound) and condition (congruent, incongruent, control). Across both trials, participants took significantly longer (47 ms) to respond to incongruent pairings compared to congruent. Moreover, participants were faster in the attend-word trial (with lower dispersion), suggesting that responding to words was attended by greater cognitive ease than the timbre identification task. The error rate was low across all conditions (5%) and showed a significant association with reaction time. Incongruent pairings produced the highest error rate (9%).

Experiment 2: Mixed ANOVAs on (log) reaction time and error rates revealed that responses to the control condition were significantly shorter than the other two, indicating Stroop interference in response to the experimental task. The difference between congruent and incongruent was non-significant, as were all other main effects. However, significant interactions between condition/stimuli type (natural vs. synthesizer) and condition/group (music-majors vs. non-majors) were observed. Similarly, analysis of error rate revealed a significant interaction between descriptive sense modality and condition, with errors rising for the control between luminance and texture and falling for the other two conditions.

Conclusion

This study used a cross-modal Stroop task to assess the effects of categorical and continuous timbre descriptors on timbre processing speed, and conversely, of timbral stimuli on semantic processing. Experiment 1 found that congruent vs. incongruent pairings of instrument timbres and written names caused significant Stroop interference, indicating reciprocal interference between semantic and auditory modalities. This means that the most common categorical mode of musical timbre discrimination—identifying a timbre's source by name—may influence perception. Experiment 2 used bipolar cross-modal timbre descriptors (*dark-bright, smooth-rough*) with validated timbral stimuli in a semantic-auditory Stroop-type task. There was no observable main effect of congruent/incongruency on processing speed, suggesting that timbres typically described as *bright*, *dark*, *smooth*, or *rough* did not by themselves influence the processing of these luminance and texture adjectives. However, congruency of timbre/word pairs exhibited two-way interactions with stimuli type (natural vs. synthesizer), group (music-majors vs. non-majors), and descriptive modality. In the first interaction, this suggests that incongruent pairings of adjectives and natural timbres led to "cross-talk" between semantic and timbre processing systems, slowing reaction time, while synthesizers largely did not. Taken together, while congruency alone did not substantially interfere with semantic processing, it interacted with stimuli type, participant musical background, and modality.

In sum, this study demonstrates that semantic and timbre processing mechanisms may be interrelated. This effect is most pronounced in source identification, i.e., when timbre percepts conform to discrete categories. Timbral stimuli, however, do not seem to as strongly influence responses to common cross-modal descriptive adjectives, though they interact significantly with other variables. This suggests that categorical and continuous features of timbre perception, while dissociable in certain aspects, are closely coupled when we reconcile sensory and semantic frames of reference (Siedenburg, Jones-Mollerup, & McAdams, 2016). Implications for our understanding of the timbre-language connection are discussed in conclusion, along with an analysis of the acoustic correlates of cross-modal timbre associations.

References

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition*. Oxford: Clarendon.

Siedenburg, K., Jones-Mollerup, K., & McAdams, S. (2016). Acoustic and categorical dissimilarity of musical timbre: Evidence from asymmetries between acoustic and chimeric sounds. *Frontiers in Psychology, 6*(1977). doi:10.3389/fpsyg.2015.01977

Zacharakis, A., Pastiadis, K., & Reiss, J. D. (2014). An interlanguage study of musical timbre semantic dimensions and their acoustic correlates. *Music Perception, 31*(4), 339–358.

The Sound Morphing Toolbox

Marcelo Caetano
INESC TEC, Sound and Music Computing group, Porto, Portugal
mcaetano@inesctec.pt

Aims/goals

Sound morphing is a transformation that gradually blurs the categorical distinction between the source and target sounds by blending sensory attributes (Caetano & Rodet, 2013). The focus and ultimate challenge of most sound morphing techniques is to interpolate across dimensions of timbre perception to achieve the desired result. There are several sound morphing proposals in the literature with few open-source implementations freely available, making it difficult to reproduce the results, compare models, or simply use them in other applications such as music composition, sound design, and timbre research. This work presents the sound morphing toolbox (SMT), which contains Matlab implementations of two sound morphing algorithms. The SMT is freely available and open-source.

Background information

In music composition, sound morphing allows the exploration of the sonic continuum (Wishart, 1996). Notable examples include Jonathan Harvey's Mortuos Plango, Vivos Voco (Harvey, 1981), Michael McNabb's Dreamsong (McNabb, 1981), and Trevor Wishart's Red Bird (Wishart, 1996). These morphs were achieved by hand. Ideally, given the input sounds, sound morphing should allow automatically setting input parameters that control the morph to achieve the desired result (Slaney et al., 1996).

Sound morphing is also used in audio processing, sound synthesis, and sound design. Tellman, Haken and Holloway (1995) proposed a sound morphing technique based on sinusoidal modeling that is intended to improve the performance of a sample-based synthesizer by morphing between sounds of the same instrument to obtain intermediate pitches, dynamics, and other effects. Kyma (http://kyma.symbolicsound.com/) is a commercial sound design environment based on sinusoidal analysis with Loris (Fitzet, Haken, Lefvert, Champion & Odonnell, 2003) that also allows morphing.

Signal processing is commonly used to study timbre perception. McAdams, Winsberg, Donnadieu, Soete and Krimphoff (1995) and later Caclin, McAdams, Smith and Winsberg (2005) built timbre spaces with synthetic musical instrument sounds to investigate the effects of controlled manipulations. Sound morphing techniques have been used to investigate different aspects of timbre perception. Grey and Gordon (1978) investigated the perceptual effect of exchanging the shape of the spectral energy distribution between pairs of musical instrument sounds. More recently, Carral (2011) used spectral morphing to determine the just noticeable difference in timbre for trombone sounds. Siedenburg et al. (2016) investigated the acoustic and categorical dissimilarity of musical timbre with morphing.

Methodology

The SMT has implementations of two sound models and two sound morphing techniques in Matlab R2015a. The sinusoidal model (SM) (McAulay & Quatieri, 1986) and the Source-Filter Model (SFM) (Caetano & Rodet, 2013) were implemented, along with the morphing techniques that use respectively the SM (Tellman et al., 1995; Fitz et al., 2003) and the SFM (Caetano & Rodet, 2013). There is also an implementation of the Phase Vocoder (PV) (Dolson, 1986) which is used to time-scale the input sounds to the same duration previously to morphing for both models (SM and SFM.) The SMT is open-source and freely available at https://github.com/marcelo-caetano/sound-morphing/.

Results

The SMT automatically morphs between two input sounds by setting the morphing parameter α. Only the source sound S is heard when $\alpha = 0$, whereas only the target sound T is heard when $\alpha = 1$. Intermediate values of α correspond to morphed sounds with different combinations of S and T. The SMT performs stationary, dynamic, or cyclostationary morphs (Caetano & Rodet, 2013).

Conclusions

The SM model can produce striking morphs for a broad palette of sounds (Fitz et al., 2003; Tellman et al., 1995). However, SM morphs are generally not perceptually linear with respect to α (Caetano & Rodet, 2013). The SFM model and morph algorithm were developed to obtain morphs that are as perceptually linear as possible when α varies linearly. The morph algorithms can be used in music composition, sound synthesis and design, and notably to investigate timbre perception. When morphing musical instrument sounds, the result can create the auditory illusion of interpolated timbres (Fitz et al., 2003; Hikichi & Osaka, 2001) or hybrid musical instrument sounds (Caetano & Rodet, 2013). Morphed musical instrument sounds can be used to investigate cutting-edge issues in timbre perception (Carral, 2011; Siedenburg et al., 2016) or to explore the sonic continuum in music composition (Wishart, 1996).

Acknowledgments

FCT (SFRH/BPD/115685/2016) and FourEyes Research Line within project "TEC4Growth - Pervasive Intelligence, Enhancers and Proofs of Concept with Industrial Impact/NORTE-01-0145-FEDER-000020" financed by the North Portugal Regional Operational Program (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF).

References

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America*, 118(1), 471-482.

Caetano, M., & Rodet, X. (2013). Musical Instrument Sound Morphing Guided by Perceptually Motivated Features. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(8), 1666-1675.

Carral, S. (2011). Determining the Just Noticeable Difference in Timbre Through Spectral Morphing: A Trombone Example. *Acta Acustica united with Acustica*, 97(3), 466-476.

Dolson, M. (1986). The Phase Vocoder: A Tutorial. *Computer Music Journal*, 10(4), 14-27.

Fitz, K., Haken, L., Lefvert, S., Champion, C., & Odonnell, M. (2003). Cell-Utes and Flutter-Tongued Cats: Sound Morphing Using Loris and the Reassigned Bandwidth-Enhanced Model. *Computer Music Journal,* 27(3), 44-65.

Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *The Journal of the Acoustical Society of America*, 63(5), 1493-1500.

Harvey, J. (1981). "Mortuos Plango, Vivos Voco": A Realization at IRCAM. *Computer Music Journal*, 5(4), 22-24.

Hikichi, T., & Osaka, N. (2001). Sound timbre interpolation based on physical modeling. *Acoustical Science and Technology*, 22(2), 101-111.

McAdams, S., Winsberg, S., Donnadieu, S., Soete, G. D., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58(3), 177-192.

McAulay, R., & Quatieri, T. (1986). Speech analysis/Synthesis based on a sinusoidal representation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(4), 744-754.

McNabb, M. (1981). "Dreamsong": The Composition. *Computer Music Journal*, 5(4), 36-53.

Siedenburg, K., Jones-Mollerup, K., & McAdams, S. (2016). Acoustic and Categorical Dissimilarity of Musical Timbre: Evidence from Asymmetries Between Acoustic and Chimeric Sounds. *Frontiers in Psychology*, 6.

Slaney, M., Covell, M., & Lassiter, B. (1996). Automatic audio morphing. *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*.

Tellman, E., Haken, L., & Holloway, B. (1995). Timbre Morphing of Sounds with Unequal Numbers of Features. *Journal of the Audio Engineering Society*, 43(9), 678-689.

Wishart, T. (1996). *On sonic art*. London: Routledge.

David A. Luce's Research on Musical Instrument Timbre and its Subsequent Application to Analog Synthesizer Design and Digital Analysis/Synthesis

Benjamin P. Luce[1†], James W. Beauchamp[2]
[1] Natural Sciences Department, Northern Vermont University, Lyndonville, Vermont, USA
[2] Dept. of Electrical & Computer Engineering and School of Music, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA
[†] benluce137@gmail.com

Aims/goals

The aim of this research is to revisit, trace the impact of, and extend, the timbre research results of the 1960 Era Musical Acoustics Team at MIT, and those of its core member David A. Luce during the period 1959-1975. These results included first demonstrations of the importance of attack and decay transients for instrument identification, detailed analyses of harmonic amplitude attack transients, and Luce's discovery of "pitch-invariant spectral envelopes" that exist for many instruments. We also trace the subsequent application of this knowledge in Luce's development of some of the first analog synthesizers, both monophonic and polyphonic, that could produce reasonable simulations of conventional instruments.

Background information

David A. Luce and his MIT team carried out their initial research during the period 1959-1963 under the supervision of Professor Melville Clark Jr. Luce was the first to use computer analysis of musical tones to perform a thorough analysis of timbre, for 14 conventional instruments. Subsequent to this work he developed an analyzer/synthesizer to extend this research. During 1968-1971 he created a series of analog synthesizers (not commercialized) that could mimic many conventional instruments remarkably well, and then developed commercial synthesizers for Moog Music, most famously the Polymoog. Following his passing in Spring of 2017, we began re-evaluating his work in detail and tracing its influence forward. Our analysis has also given rise to some new conjectures which we are investigating.

Methodology

Our present research is being pursued via computer analysis of musical tones, including both existing tone sets as well as new sets we are generating via real instruments and synthesizers, including the Polymoog, the Taurus Bass Pedals, and other instruments designed or influenced by Luce.

Results

We have recreated and evaluated the accuracy of the original pitch-synchronous analysis technique used by David Luce for obtaining harmonic evolution data, and found it to be sufficiently accurate for most purposes. We have also shed light on the "frequency normalization" approach he developed for the averaging of steady-state spectra. We have recovered some samples of recordings of his original non-commercialized instruments and are in the process of analyzing the timbral characteristics of these. We have also been able to show direct connections between his original research and the use of formant filters and other special features of the Polymoog. We have found this instrument to be remarkably capable for acoustic instrument simulation as well as a useful generator of novel musical timbres. Similar comments apply to the Taurus Pedals. We will present some relevant demonstrations of these in our talk.

Our results in retracing the influence of Luce's research are generally too extensive to summarize here, but to give some indication, Beauchamp's own work in instrument analysis/synthesis at the University of Illinois at Urbana-Champaign has been impacted by Luce's research results in several significant ways. This included the incorporation of harmonic attack delays in Beauchamp's Harmonic Tone Generator (HTG) in 1964, based on Luce's observation of these in his computer analyses; the creation of an analysis/synthesis program based on Luce's approach starting with a 1966 Audio Engineering Society preprint and a chapter in the 1969 book *Music by Computers*; Beauchamp's development of nonlinear synthesis models that utilized correlations between spectral centroid and spectral envelopes in the 1980s;

and the development by Beauchamp and Andrew Horner of a more accurate instrument synthesis method based on a family of spectral envelopes indexed by their spectral centroids in the 1990s.

Conclusions

We find that David Luce's contributions have had a wide, although perhaps somewhat uneven, impact on the field of musical timbre research, and have not been very well understood or appreciated up to this point. This is especially true of his work on instrument spectral envelopes, which he obtained with the above-mentioned frequency normalization approach. These results suggest a more general connection between timbre and perceived dynamic level that we are currently investigating.

References

Beauchamp, J. W. (1965). *Electronic Instrumentation for the Synthesis, Control, and Analysis of Harmonic Musical Tones* (Doctoral dissertation, University of Illinois), pp. 1-143

Beauchamp, J. W. (1966). Additive Synthesis of Harmonic Musical Tones. *Journal of the Audio Engineering Society*, 14(4), 332-342

Beauchamp, J. W. (1966). Transient Analysis of Harmonic Musical Tones by Digital Computer. *Audio Engineering Society Preprint* No. 479, pp. 1-18.

Beauchamp, J. W. (1969). A Computer System for Time-Variant Harmonic Analysis and Synthesis of Musical Tones. *Music by Computers* (H. F. von Foerster and J. W. Beauchamp, eds), pp. 19-62.

Beauchamp, J. W. (1982). Synthesis by Spectral Amplitude and "Brightness" Matching of Analyzed Musical Instrument Tones. *Journal of the Audio Engineering Society*, 30(6), 396-406.

Beauchamp, J. W. (2007), "Analysis and Synthesis of Musical Instrument Sounds", in *Analysis, Synthesis, and Perception of Musical Sounds: Sound of Music*, J. W. Beauchamp, Ed., Springer, N. Y., pp. 1 – 89.

Beauchamp, J. W. & Horner, A. B. (1995). Wavetable Interpolation Synthesis Based on Time-Variant Spectral Analysis of Musical Sounds. *Audio Engineering Society Preprint* No. 3960, pp. 1-17.

Clark Jr, M., & Luce, D. A. (1982). *U.S. Patent No. 4,365,533*. Washington, DC: U.S. Patent and Trademark Office.

Clark Jr, M., Luce, D., Abrams, R., Schlossberg, H., & Rome, J. (1963). Preliminary experiments on the aural significance of parts of tones of orchestral instruments and on choral tones. Journal of the Audio Engineering Society, 11(1), 45-54.

Clark Jr, M., Robertson, P. T., & Luce, D. (1964). A preliminary experiment on the perceptual basis for musical instrument families. *Journal of the Audio Engineering Society*, *12*(3), 199-203.

Luce, D. A. (1963). Physical correlates of nonpercussive musical instrument tones (Doctoral dissertation, Massachusetts Institute of Technology).

Luce, D. A. (1975). Dynamic spectrum changes of orchestral instruments. *Journal of the Audio Engineering Society*, *23*(7), 565-568.

Luce, D. A. (1978). *U.S. Patent No. 4,099,439*. Washington, DC: U.S. Patent and Trademark Office.

Luce, D., & Clark Jr, M. (1965). Duration of attack transients of nonpercussive orchestral instruments. *Journal of the Audio Engineering Society*, *13*(3), 194-199.

Luce, D., & Clark Jr, M. (1967). Physical Correlates of Brass-Instrument Tones. *The Journal of the Acoustical Society of America*, *42*(6), 1232-1243.

The Effect of Loudness on the Perceptual Representation of Voiceless Vowel and Fricative Timbres

Charlie Sdraulig
CCRMA, Department of Music, Stanford University, Stanford, California, USA
sdraulig@ccrma.stanford.edu

Aims/goals

How do changes in loudness effect the perceptual representation of timbre in a corpus of recorded voiceless vowels and fricatives? I sought to uncover the acoustic cues informing listeners' perception of timbre in this corpus, as well as the extent to which these cues might vary at different loudness levels. I normalized the perceived loudness of my corpus to two levels (ITU-R, 2015), creating two versions—one loud and the other soft. Via listeners' dissimilarity ratings and Multidimensional Scaling (MDS), I aimed to compare the perceptual representations of each version of the corpus. I hypothesized that the stimuli in the soft condition would be judged as more similar in timbre to one another, as any relevant acoustic descriptors would provide less prominent cues for dissimilarity judgements at softer levels. The choice of stimuli in this study was motivated by its future applications in soft noise-based vocal compositions using timbre dissimilarity as a formal device.

Background information

MDS methods have been employed for harmonic and percussive (i.e. noisy) instrumental timbres (Lakatos, 2000), vocal timbre (Erikson, 2008), and speech perception (Baumann & Belin, 2010). Though interactions between timbre and loudness have been demonstrated, I am not aware of any studies that have explicitly explored the effect of changes in loudness on the perceptual representation of timbre itself (McAdams, 2013).

Methodology

15 music students at Stanford University took part in this experiment. Their average age was 26.1 years (SD = 4.76). 6 were female. All spoke English as their first language and reported normal hearing. The corpus consisted of eight voiceless stimuli. Stimuli 1 to 3 were whispered English vowels [ɑ], [ʌ], and [i]. Stimulus 4 was a whispered English [ɹ̠] (a postalveolar approximant). Stimuli 5 to 8 were English fricatives [s], a diffuse [ʃ], a focused [ʃ], and [x]. Each stimulus was 1-second in duration, with controlled 50-ms ramped onsets and offsets. All stimuli were recorded by one male vocalist. Lip to microphone distance was 12 inches. The perceived loudness of each stimulus was normalized to two levels (ITU-R, 2015). The loud condition level was -23.0 Loudness K-weighted, relative to Full Scale (LUFS). The soft condition level -54.1 LUFS the lowest analyzed Integrated Loudness for the eight samples recorded. Each version of the corpus was presented diotically over headphones in a separate block, in a classical pairwise comparison task. Paired stimuli were separated with 500-ms of silence. The loud corpus was presented at approximately 60 dB SPL; the soft corpus was presented at approximately 40 dB SPL. Listeners recorded their dissimilarity ratings on a horizontal scale from 1 (very similar) to 9 (very dissimilar).

Results

A two-way ANOVA on listeners' ratings yielded significant main effects for perceived loudness, $F(1, 784) = 6.50$, $p = .011$, $\eta_p^2 < .01$ as well as for stimuli pair, $F(27, 784) = 19.09$, $p < .001$, $\eta_p^2 = .40$, but no significant interaction between these factors, $F(27, 784) = 1.16$, $p = 0.267$, $\eta_p^2 = .04$. 3 dimensional INDSCAL (Carroll and Chang, 1970) solutions were found for the loud (Kruskal's stress-1 = 0.171) and soft (Stress-1 = 0.175) conditions. Overall, on average, listeners weighted dimension 3 above 2, and 2 above 1. Audio descriptors were calculated in the Timbre Toolbox (Peeters, Giordano, Susini, Misdariis, & McAdams, 2011). For the loud condition: dimension 3 strongly correlated with *spectral spread (interquartile range)* ($r(6) = 0.91$, $p = .002$) and *spectral rolloff (median)* ($r(6) = 0.88$, $p = .004$); dimension 2 strongly correlated with *spectral kurtosis (IQR)* ($r(6) = -0.93$, $p < .001$) and *spectro-temporal variation (IQR)* ($r(6) = -0.88$, $p = .003$); dimension 1 strongly correlated with *noise energy (med)* ($r(6) = -0.90$, $p = .002$). For the soft condition: dimension 3 strongly correlated with *spectral spread (IQR)* ($r(6) = -0.90$, $p = .002$) and *spectral rolloff (med)* ($r(6) = -0.85$, $p = .007$); dimension 2 strongly correlated with *spectro-*

temporal variation (IQR) ($r(6) = -0.96$, $p < .001$) and *spectral kurtosis (IQR)* ($r(6) = -0.87$, $p = .005$); dimension 1 strongly correlated with *noise energy (med)* ($r(6) = -0.90$, $p = .002$)

Conclusions

As expected, stimuli in the soft condition were perceived as slightly more similar to one another than in the loud condition—indicated by the significant main effect for perceived loudness from the ANOVA, as well as the overall mean ratings for the loud ($M = 5.64$, $SD = 1.55$) and the soft ($M = 5.32$, $SD = 1.42$) conditions respectively. With the caveat that a modest number of listeners rated a relatively small corpus, I found that the audio descriptor correlates for both conditions' MDS solutions were consistent, despite the approximately 20 dB SPL difference in presentation level. Wide bands of spectral energy and statistical descriptions of spectral distribution were highly correlated with dimension 3 i.e. spectral spread and rolloff. Dimension 2 was highly correlated with spectro-temporal variation i.e. spectral flux. Dimension 1 was parsed categorically. This was confirmed by correlating dimension 1 with a binary code where 0 was assigned to stimuli 1 to 4, and 1 was assigned to stimuli 5 to 8 (loud condition ($r(6) = 0.96$, $p < .001$); soft condition ($r(6) = 0.97$, $p < .001$)). This dimension likely reflects whether a given stimuli was a fricative or not, possibly cued by the presence or absence of the turbulent air flow generated during the production of fricatives. For musical composition, possible implications of these findings are as follows: to delineate formal landmarks (e.g. the beginnings and ends of sections) via timbre dissimilarity at softer levels, larger changes in wide bands of spectral energy and spectral flux may be required, potentially coupled with categorical changes in vocal production for further contrast. In addition, the audio descriptor correlates found here suggest a perceptual space that may inform the classification of these voiceless sounds in phonetics.

Acknowledgements

Thank you to Jonathan Berger, Takako Fujioka, Madeline Huberth, and the anonymous reviewers for their valuable comments.

References

Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: common dimensions for different vowels and speakers. *Psychological Research, 74*, 110-120.

Carroll, D., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika, 35*, 283-319.

Erickson, M. L. (2008). Dissimilarity and the Classification of Male Singing Voices. *Journal of Voice, 22(3),* 290-299.

International Telecommunication Union-Radiocommunication Sector (2015) Recommendation ITU-R BS.1770-4 Algorithms to measure audio programme loudness and true-peak audio level. Retrieved from https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1770-4-201510-I!!PDF-E.pdf

Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics, 62(7)*, 1426-1439.

McAdams, S. (2013). 2 - Musical Timbre Perception. In Diana Deutsch (Ed.), *The Psychology of Music Third Edition* (pp. 35-67). USA: Academic Press.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The Timbre Toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America, 130(5)*, 2902-2916.

A Case for Ravel: Timbre, Neurological Decline, and Ravel's Last Compositions

Jennifer Beavers
Department of Music, University of Texas at San Antonio, San Antonio, Texas, United States
jennifer.beavers@utsa.edu

Aims/goals

The aim of this talk is to examine the correlation between Ravel's cognitive decline towards the end of his life and its potential impact on his compositional practice, most notably observed in the structural role timbre plays within the musical form.

Background information

As early as 1926, Ravel began complaining about exhaustion, anemia, and memory loss. Following a 1932 taxi accident, Ravel experienced a rapid neurological decline. By 1933, his symptoms worsened to include memory gaps, neurasthenia, aphasia, and apraxia and he could no longer compose. A right-sided craniotomy was performed by his surgeon, Clovis Vincent, in December of 1937. There was no mention of a subdural hematoma in his operation notes, ruling out head trauma from the accident, but it was noted that his brain was collapsed. As such, many neuroscientists interpret the accident as a 2nd major factor that revealed a pre-existing brain disease, which would likely be diagnosed today as fronto-temporal dementia. Ravel slipped into a semi-coma following the surgery and died eight days later. The cognitive struggles Ravel experienced at the end of his life have fascinated researchers in the medical profession since his physician's first publication in 1948 (Alajouanine 1948). To date, over 40 publications in neurological and scientific journals show how doctors have attempted to unravel his neurological symptoms (for example, Baeck 1996; Cytowic 1976; Dalessio 1984; Otte et al. 2003) and make connections between his condition and its effect on his art (Amaducci et al. 2002; Baeck 2002; Cybulska 1997; Seeley et al. 2008; Sellal 2008, Sergent 1993). As no post-mortem was granted, the case for Ravel ultimately remains speculative.

Ravel's cognitive decline has also captivated writers, filmmakers, and radio programmers (Abumrad 2012; Eschenoz 2007; Weinstein 2001). Some discussions are positioned as conjecture, while others openly assert the obviousness the Ravel's neurological disease on his compositions, in which *Boléro* is, for many, an obvious example of Ravel's heightened attention to timbre. More, Ravel's personal correspondence points to significant struggles he faced with regard to creativity as well as physical and mental challenges at the end of life (Orenstein 1990). Yet, music analysts fail to connect these neurological symptoms to changes in Ravel's compositional strategy. As Straus states, "[music] critics have preferred to avert their gaze" (2011, 42). In so doing, we are perhaps depriving ourselves of a vital piece of the puzzle when it comes to understanding Ravel's late style of composition.

Traditional forms, most notably, the sonata form, appear throughout Ravel's oeuvre. In compositions after the 1928, however, a change is observed: structurally-important moments, typically reserved for intricate manipulations of melodic and harmonic expectations, are replaced by timbrally-marked moments. In works such as *Boléro* (1928), *Chansons madécasses* (1925-26), the Piano Concerto for the Left Hand (1929-30), and the Piano Concerto in G (1929-31), timbre is assigned a structural role that shows a marked change in compositional priority. This paper examines Ravel's troublesome neurological history alongside his heightened focus on timbre in his late years.

Methodology

This is primarily an analytical paper in which I take a multi-modal approach to contextualizing timbre within Ravel's late compositions. My approach is informed by medical and historical research, Ravel's personal correspondence, and musical observations. I have analyzed formal and harmonic characteristics of pre- and post- World War I compositions to establish stylistic tendencies of each period.

Results

Compared to early compositional strategies, pieces composed after 1928 showcase timbre in remarkable ways. With *Boléro*, timbre replaces musical development. In the Piano Concerto for the Left Hand, melodic development and form generation is intimately connected to two contrasting timbre and contour motives. In this way, *Boléro* and the left-hand concerto can be classified as timbrally-generated forms. The Piano Concerto in G, utilizes timbral manipulations to conceal and transform the piano and orchestra, in which instruments are figuratively and musically transformed into a jazz piano, a Spanish guitar, a harp, and a musical saw or theremin; in this way, the G-major concerto can be categorized as a timbrally-marked form. The results of harmonic and formal analysis of these late works, thus, reveal a marked use of timbre, which by and large function as structural markers that transcend the rather conventional forms he employed.

Conclusions

I conclude that Ravel's enhanced focus on timbre around 1928 aligns with documentation that he was experiencing cognitive decline. This is not to say that his deteriorating health condition creates a disability reading in which timbre replaces the sophisticated confluences of melody and harmony of earlier works. Rather, I posit that this shift towards timbre might be motivated, at least in part, by influences outside of his conscious awareness. His shift towards timbre could perhaps dually indicate the changing internal landscape of his brain, or what some neuroscientists refer to as right-hemisphere generated processes, while also reflecting how the expert orchestrator leveraged his abilities to reflect the modern soundscape of his world—a world that heard initial performances of jazz and early electric instruments such as the theremin. In this way, Ravel's attention to timbre anticipates the radical sonic experimentations of the following decades.

References

Abumrad, J. (Host). (2012, June 18). *Unraveling Boléro* [Audio podcast]. Retrieved from http://www.radiolab.org/story/217340-unraveling-bolero/

Alajouanine Th. (1948). Aphasia and artistic realization. *Brain, 71*, 229-41.

Amaducci, L., Grassi, E., & Boller, F. (2002). Maurice Ravel and right-hemisphere musical creativity: influence of disease on his last musical works? *European Journal of Neurology, 9*(1), 75-82.

Baeck, E. (1996). Was Maurice Ravel's illness a corticobasal degeneration? *Clinical Neurology and Neurosurgery, 98*, 57-61.

Baeck, E. (2002). Maurice Ravel and right hemisphere creativity. *European Journal of Neurology, 9*(1), 321.

Cybulska, E. (1997). Boléro unraveled. A case of musical perseveration. *Psychiatric Bulletin, 21*, 576-77.

Cytowic, R.E. (1976). Aphasia in Maurice Ravel. *Bulletin of the Los Angeles Neurological Society*, *41*(3), 109-14.

Dalessio, D.J. (1984). Maurice Ravel and Alzheimer's disease. *Journal of the American Medical Association, 252*, 3412-3.

Eschenoz, J. (2007). *Ravel: A novel*. New York: New Press.

Orenstein, A. (1990). *A Ravel Reader*. New York: Columbia University Press.

Otte, A., De Bondt, P., Van De Wiele, C., Audenaert, K. & Dierckx, R. (2003). The exceptional brain of Maurice Ravel. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research, 9*(6), 154-59.

Sellal, F. (2008). A few comments on Ravel's disease. *Brain, 131*, e98.

Seeley, W., Matthews, B., Crawford, R., Gorno-Tempini, M., Foti, D., Mackenzie, I., & Miller, B. (2008). Unravelling Boléro: progressive aphasia, transmodal creativity and the right posterior neocortex. *Brain, 131*, 39-49.

Seeley, W., Matthews, B., & Miller, B. (2008). A few comments on Ravel's diseases. *Brain, 131*, e99.

Sergent, J. (1993). Music, the brain and Ravel. *Trends in Neurosciences, 16*, 168-72.

Weinstein, L. (2001). *Ravel's Brain* [Motion picture]. Finland: Rhombus Media.

The Emergent Timbre in Live-Electronic Music: Volume and Spectral Liveness Estimation from Audio Descriptors

Danilo Rossetti[†], Jônatas Manzolli
Interdisciplinary Nucleus for Sound Communication, University of Campinas, Campinas, São Paulo, Brazil
[†] danilo.rossetti@nics.unicamp.br

Aims/goals

The general aim of this research is to analyze the emergent timbre (McAdams, 2013; Rossetti, Teixeira & Manzolli, 2017; Rossetti & Manzolli, 2017) in live-electronic music performance. We believe that the emergent characteristics of the timbre result from the interactions between the different overlapped sound layers, such as the tape sounds, the dynamic part of real-time processing, and the instrumental sounds which can be or not expressed by extended techniques. The specific goal is to understand the temporal and spectral activities of these emergent sound structures in terms of Volume and Spectral Liveness. The Volume can help us to perform a formal segmentation of the analyzed pieces, and the Spectral Liveness determines the flux qualities of the achieved emergent sonorities.

Background information

Our analysis approach is based on the idea that audio descriptors (Peeters, 2004) can help us to understand some of the questions about live-electronic pieces raised above. Our viewpoint is that the interaction between instrumental and electroacoustic sounds (under fixed support or real-time processed) produces a considerable sound flux, which can be analyzed and correlated to the spectral activity of the musical timbre. The paradigm shift related to timbre was introduced by Truax (1992), who sustained that it should be approached by a non-linear, inter-related and interactive background among its microstructures. Truax proposed the concept of "Volume" as a complex percept or the perceived magnitude of sound. Later, Malt & Jourdan (2009) proposed a different notion of "Volume," named BStD (Brightness Standard Deviation), which is a graphical representation. It is a three-dimensional graphic based on the evolution of the Spectral Centroid (*C*), Spectral Standard Deviation (*SD*) and Loudness (*L*) descriptors. The distance between *C* and *SD* curves in addition to the color associated with the intensity perception, in dB, determines this representation. Complementary to this notion, we propose the analysis of the orthogonality between the values of the Spectral Flux (*Flu*)—a measure of how quickly the spectrum amplitude changes—and the Spectral Flatness (*Fla*)—the quantity of noise found in the spectrum, in opposition to a harmonic configuration—descriptors. We name this feature Spectral Liveness, and the orthogonality is represented by a two-dimensional phase space plot.

Methodology

Three live-electronic compositions where chosen for the analysis: two from Rossetti (*Desdobramentos do contínuo*, for violoncello, and *Proceratophrys boiei*, for tenor saxophone) and one from Manzolli (*Ressonâncias*, for piano). We assume that these compositions have emergent timbres formed by the proposed three-layer combination (instrumental, tape and real-time processed sounds) in space and time. A segmentation of the pieces is performed in the Volume graphic with the *Flu* and *Fla* curves. The applied descriptors belong to the PDescriptors library, developed in Pd by Monteiro (2012). The Spectral Liveness graphic shows the correlation between *Flu* and *Fla* values (a scale between 0 and 1). These values represent low or high spectral activity (*Flu*), and harmonic or noisy spectral configuration (*Fla*) of the timbre. In Fig. 1 an example of a Volume graphic with Spectral Flux and Spectral Flatness curves is shown, representing the analysis of *Proceratophrys boiei*. A possibility of a formal segmentation of this piece is performed. In the Spectral Liveness graphic (on the right) pair values of *Flu* (X-axis) and *Fla* (Y-axis) are plotted. It is important to emphasize that the Volume graphic is a temporal representation, while the Spectral Liveness graphic is an out-of-time representation. The detailed graphics of the three analyzed pieces will be included in the Poster Presentation of the Conference.

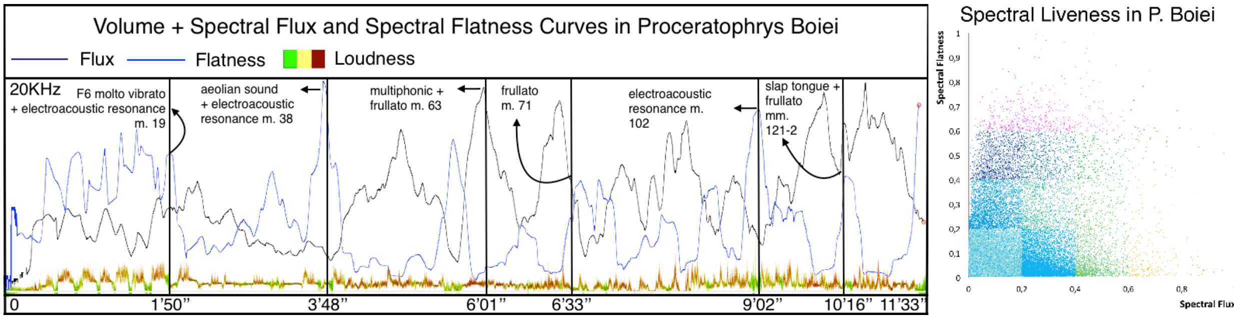


Fig. 1: Volume (formal segmentation) with *Flu* and *Fla* curves, and Spectral Liveness graphics

Results & Conclusion

In the temporal representation of the Volume graphic with *Flu* and *Fla* curves, a possible segmentation of *Proceratophrys boiei* is presented based on the given information of these descriptors. The time duration, measures of the score, and sound features of the points of segmentation are indicated. These points were decided from the peaks of both *Flu* and *Fla* curves (3'48'', 6'01'', 9'02''), and important increasing or decreasing movements (1'50'', 6'33'', 10'16''), always having in consideration the score information and a critical listening. The Spectral Liveness plot correlates *Flu* and *Fla* values and, in this piece, it indicates a concentration of points in values between 0 to 0,6. Few higher values of *Flu* and *Fla* are found (0.6 to 1), mostly when the complementary variable has a low value (0 to 0.4).

The musical analysis model for live-electronic music proposed in this article intends to be complementary to traditional analysis techniques. The gains coming from this model are the possibility of capture and represent timbre variations which cannot be represented from traditional analyzes or musical notations. The Volume graphic can be useful to perform a formal segmentation of electroacoustic pieces based on the extracted information from *BStD, C*, *SD*, *L*, *Flu* and *Fla* descriptors. This information can indicate timbre subtle variations that the ear does not perceive without a very concentrate listening. The Spectral Liveness plot gives us a figure that represents the timbre behavior on the entire piece. High *Flu* values indicate the presence of huge amplitude variations, while low *Flu* values indicate few variations of amplitude. On the other hand, high *Fla* values describe a spectral noisy configuration, while low *Fla* values describe a more harmonic spectral configuration.

Acknowledgments

Rossetti is supported by FAPESP (2016/23433-8) and Manzolli by the CNPq (305065/2014-9).

References

Malt, M., & Jourdan, E. (2009). Le 'BStD' – Une représentation graphique de la brillance et de l'écart type spectral comme possible représentation de l'évolution du timbre sonore. In *l'Analyse Musicale Aujourd'hui, Crise ou (R)evolution*. Strasbourg, France.

McAdams, S. (2013). Musical Timbre Perception. In D. Deutsch (Ed.), *The Psychology of Music* (Third Edition) (pp. 35-67). San Diego: Academic Press.

Monteiro, A. (2012). *Criação e performance musical no contexto dos instrumentos musicais digitais* (Master dissertation). University of Campinas, Campinas.

Peeters, G. (2004). *A Large Set of Audio Features for Sound Description (Similarity and Classification) in the CUIDADO Project* (Cuidado Project Report). Paris: Ircam.

Rossetti, D., Teixeira, W., & Manzolli, J. (2017). Emergent Timbre and Extended Techniques in Live-Electronic Music: An Analysis of *Desdobramentos do Contínuo* Performed by Audio Descriptors. *Musica Hodie 18(1)*. In press.

Rossetti, D., & Manzolli, J. (2017). Uma análise de *Ressonâncias* para piano e sons eletrônicos. *Revista Opus 23(3)*, (pp. 193-221).

Truax, B. (1992). Musical Creativity and Complexity at the Threshold of the 21st Century. *Interface 21*, (pp. 29-42).

Consonant and Dissonant Timbres in Two Works of Sofia Gubaidulina

Noah Kahrs
Composition Department, Eastman School of Music, Rochester, NY, USA
† nkahrs@u.rochester.edu

Aims/goals

Sofia Gubaidulina's music is often noted for its timbral inventiveness and dramatic approach to form, but these are difficult to relate to one another analytically. Most prior discussion of her music focuses on how either of these techniques relate to her religious philosophies, not on how to make sense of the sounds that result. However, her published interviews on her compositional preplanning allow us to identify consonance and dissonance among certain timbres in her work. This paper seeks to relate Gubaidulina's structural approach to consonance and dissonance to the timbres that result, to see how timbre functions in her music and which features thereof contribute to consonance and dissonance.

Background information

For analyzing Gubaidulina's music, Valentina Kholopova and Philip Ewell (2014) have presented a "parameter complex" in which consonance and dissonance are identified among different musical parameters, including rhythm, pitch, and phrase continuity, as well as timbre. Rebecca Leydon (2012) has also discussed "timbral trajectories," in which Gubaidulina's music operates along a pitch-noise axis. Vera Lukomsky (1999) and Jennifer Milne (2007) have discussed the religious poetics of Gubaidulina's formal planning, but not necessarily in relation to timbral distinctions. Farbood and Price (2017), Pressnitzer, McAdams, Winsberg, and Fineberg (2002), Kendall and Carterette (1993), and Lerdahl (1987) have all discussed the applicability of quantitative timbre attributes to musical tension.

Methodology

Two Gubaidulina works, *Meditation on the Bach Chorale "Vor deinen Thron tret ich hiermit"* and *Am Rande des Abgrunds*, were analyzed with standard music-theoretical methods as well as those of previous Gubaidulina research. *Meditation* was also analyzed in terms of the sketches in the Paul Sacher Stiftung in Basel, Switzerland, and the contents of the sketches were compared against published interviews describing Gubaidulina's precompositional process and philosophy. Certain timbres were identified as consonant and dissonant, and representative samples of those timbres were analyzed with Matlab, so that several timbre attributes (i.e. spectral centroid) could be calculated and compared to other musical factors.

Results

The sketches for *Meditation* show a wide range of structural divisions, which overlap in the middle section of the piece and correspond to changes of playing techniques and thus of timbre. Figure 1 shows major structural divisions and corresponding timbres in the middle section of the piece; numbers along the top correspond to rehearsal marks in the score, and numbers in the next row show relative proportions. Because Gubaidulina's published interviews state that dissonance occurs between conflicting structural divisions, middle-section techniques such as tremolo and circular bowing can be considered dissonant, whereas more traditional playing techniques are consonant. Likewise in *Am Rande*, pizzicato and tremolo are dissonant, whereas arco harmonics are consonant. In both pieces, dissonant timbres have consistently higher spectral centroid and spectral flatness, and thus similarity to white noise, than consonant timbres.

Conclusions

In Sofia Gubaidulina's music, timbre explicitly articulates large scale structural divisions. By comparing Gubaidulina's published interviews to her sketches, specific timbres in her work can be considered consonant and dissonant. Dissonant timbres generally have higher spectral centroid and spectral flatness, suggesting that Gubaidulina's consonance-dissonance axis may lie along the pitch-noise axis. Gubaidulina's treatment of consonant and dissonant timbres is similar to traditional functional harmonic treatment of consonance and dissonance, suggesting parallels between timbral similarity to white noise in twentieth-century music and harmonic tension in earlier music.

Acknowledgments

My travel to the Paul Sacher Stiftung in Basel, Switzerland, was supported by a Summer International Travel Grant from the University of Chicago College Study Abroad Office. My access to Gubaidulina's sketch material was made possible by the staff of the Paul Sacher Stiftung, particularly by librarian Evelyne Diendorf and musicologist Felix Meyer.

References

Ewell, P. A. (2014). The Parameter Complex in the Music of Sofia Gubaidulina. *Music Theory Online*, 20(3).

Farbood, M. M., & Price, K. C. (2017). The contribution of timbre attributes to musical tension. *The Journal of the Acoustical Society of America*, 141(1), 419-427.

Kendall, R. A., & Carterette, E. C. (1993). Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck's adjectives. *Music Perception: An Interdisciplinary Journal*, *10*(4), 445-467.

Lerdahl, F. (1987). Timbral hierarchies. *Contemporary Music Review*, *2*(1), 135-160.

Leydon, R. (2012). Clean as a Whistle: Timbral Trajectories and the Modern Musical Sublime. *Music Theory Online*, 18(2).

Lukomsky, V. (1999). 'Hearing the Subconscious': Interview with Sofia Gubaidulina. *Tempo*, (209), 27-31.

Milne, J. D. (2007). The rhythm of form: Compositional processes in the music of Sofia Gubaidulina. ProQuest.

Pressnitzer, D., McAdams, S., Winsberg, S., & Fineberg, J. (2000). Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Perception & psychophysics*, *62*(1), 66-80.

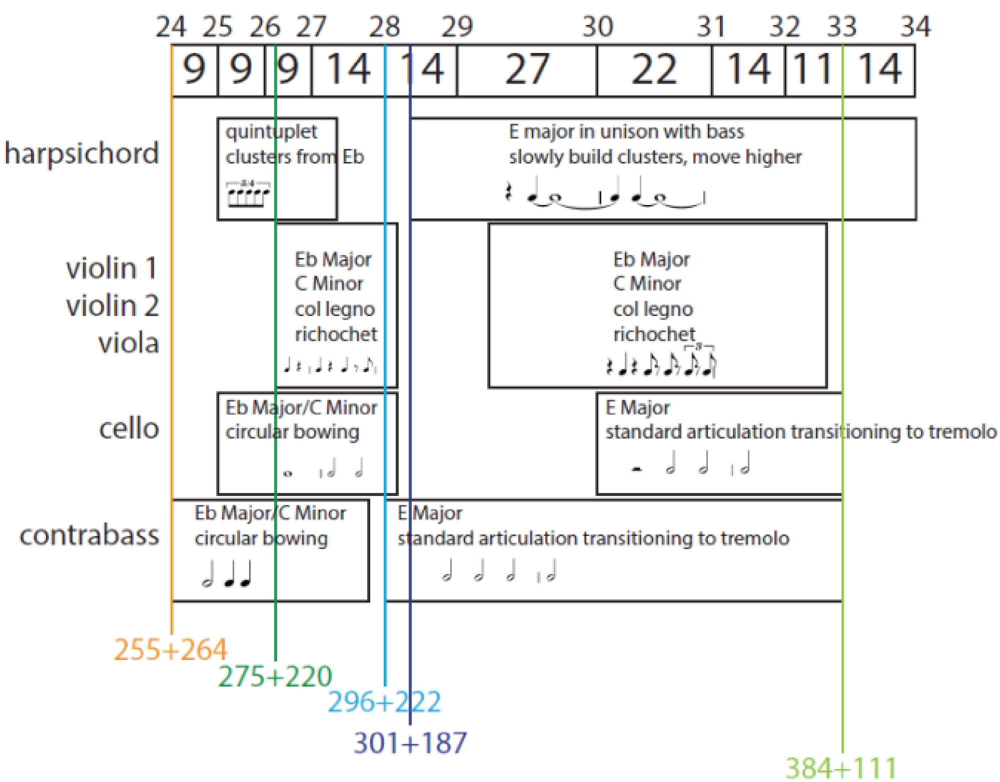


Figure 1: Structural divisions by timbre versus by sketch study in Gubaidulina's *Meditation*

Acousmatic Listening, Chimeric Percepts, and Auditory Hallucinations: Conceptualizing "Perceptualized" Timbre through the "Predictive Processing" Model

Nora Engebretsen
Department of Musicology/Composition/Theory, Bowling Green State University, Bowling Green, OH, USA
norae@bgsu.edu

Aims/goals

This poster presentation addresses issues related to the perception and analysis of "chimeric" timbres (i.e., those without readily identifiable physical sources) in acoustic and electroacoustic repertoires. Responding to Brian Kane's (2014) characterization of acousmatic sounds as anxiety inducing—due to the listener's inability to bridge the gap between these sounds' sonic effects and their physical sources/causes—the present study applies Andy Clark's (2016) "predictive processing" model to explore ways in which the imagination might deal with the ambiguity of such sounds. This approach is used to better understand the relationships between listeners and their environments and, specifically, between subjective interpretations and the physical world.

Background information

This work builds on increasingly "perceptualized" music-analytic engagements of timbre, both in a general a sense of involving a phenomenological orientation and in a more specific sense of involving an understanding of timbre as a product of cognitive operations on acoustic inputs. Several recent works, including Leydon 2012, have been strongly influenced by ethnomusicologist Cornelia Fales's 2002 article, "The Paradox of Timbre." Fales's title plays on the often-cited paradox of timbre as "readily perceptible, but difficult to conceptualize" but also on a related paradox involving the relationship between perceptualized timbre (the listener's subjective sonic world) and the physical, acoustic world, which Fales leverages to account for timbre's power as a carrier of extra-musical meaning. Leydon adapts Fales's concept of perceptualization—and particularly her conception of timbre's ability to signify states of both embodiment and, though fission of previously fused timbres, metaphysical transcendence—to create an interpretive framework for timbre in a range of twentieth-century works.

The notion of timbre as the product of a source and cause in the physical world is crucial for both Fales's and Leydon's theories, but the existence of an identifiable sound source is not a universal given. Indeed, acousmatic sound, as defined by Pierre Schaeffer and further conceptualized by Dennis Smalley (1997) and Brian Kane (2014), is explicitly untethered from a physical source and from attending expectations. As listeners, we are confronted with and asked to make sense of such sounds. This cognitive task aligns well with Clark's (2016) model of "predictive processing," which defines our engagement with the world as a constant process of "guessing" about the incoming sensory data, evaluating goodness of fit of the data against our conceptual models, and revising and acting on our guesses accordingly.

Methodology

The project centers on interpretive accounts, informed by the concept of predictive processing, of several short passages drawn from works that grant the listener varying degrees of accessibility, via timbral information, to the physical sound source. Works discussed include Edgar Guzmán's *∞¿?* for amplified bassoon and electronics (featuring shifting timbral relationships between the bassoon and electronic sounds), Adrian Moore's *Study in Ink* (in which digital manipulations distance sound from its physical source), and Stockhausen's *Kontakte* (Nr.12, for electronic sounds alone). These analytical accounts, like Clark's own, do not specify details of cognition, in terms of neural processing, but instead aim to "provide functional explanations of the empirical data" (Clark 2016, 2)—here assumed to be explanations of the sound sources—arrived upon through processes of guessing and testing. The possibility of a purely acousmatic listening experience is not considered in this context.

Results

The examples suggest that if context activates what Clark terms an "apt set of priors…poised to impact processing without further delay" (42), we cleave to those priors. Once we've locked in to a bassoon timbre, for example, we continue to hear bassoon timbre even in the face of mounting evidence to the contrary. Conversely, the absence of clear contextual cues related to sound source leaves us to navigate the sensory data with a lower degree of certainty about our predictions, but even so these predictions remain grounded in experience of the world. A more interesting result emerges as the examples illustrate the applicability of the predictive element of Clark's model, through what Clark describes as the "perceptual strangeness of unexpected sensation"—here "prediction errors" and surprises "induced by a mismatch between the sensory signals encountered and those predicted" (79). These include timbral slippages or misattributions, and a crossing of a parametric boundary between timbre and rhythm. These moments raise a substantial issue to be considered, if only briefly: the nature and musical potential of prediction and expectation (see, for example, Huron and Margulis 2010) in relation to a domain generally held to be non-syntactical.

Conclusions

The use of "predictive processing" as a tool for modeling listeners' experiences of chimeric timbres resonates with Kane's characterization of the anxious listener devising "frantic hypotheses" about sound sources, in an attempt to disacousmatize sounds. The predictive processing model is also compatible with analytical narratives of timbral transcendence, though the notion that we perceive a world of our own construction, separate from the physical data is one Clark rejects. The predictive model is shaped by and shapes our interaction with the world, and the priors we bring to that interaction include not only physical bodies, but also a range of sociocultural structures and practices. Perhaps the most promising line of inquiry suggested by the study, however, is the relationship between "prediction errors" and musical expectations grounded in our perceptualizations of timbre.

References

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. New York, NY: Oxford University Press.

Fales, C. (2002). The paradox of timbre. *Ethnomusicology*, 46(1), 56–95.

Huron, D. & Margulis, E. (2010). Musical expectancy and thrills. In Juslin, P. & Sloboda, J. (Eds.), *Handbook of music and emotion: Theory, research, applications*. New York, NY: Oxford University Press.

Kane, B. (2014). *Sound unseen: Acousmatic sound in theory and practice*. New York, NY: Oxford University Press.

Leydon, R. (2012). Clean as a whistle: Timbral trajectories and the modern musical sublime. *Music Theory Online, 18(2)*.

Smalley, D. (1997). Spectromorphology: Explaining sound-shapes. *Organised Sound,* 2(2), 107–126.

Unsupervised timbre spaces through perceptually-regularized variational learning

Axel Chemla--Romeu-Santos[1†], Adrien Bitton[1], Goffredo Haus[2], Philippe Esling[1]
[1]IRCAM CNRS – UMR 9912, 1 Place Igor Stravinsky, F-75004 Paris, France
[2]Laboratorio d'Informatica Musicale, Università Degli Studie di Milano, Via Festa del Perdono, 20122 Milano, Italy
† chemla@ircam.fr

Aims/goals

Timbre spaces have been used to study the relationships between different instrumental timbres based on perceptual ratings. These spaces are usually constructed with multidimensional scaling and then analyzed with signal descriptors. However, this approach is highly limited as it computes fixed spaces with limited interpretability and no generative capabilities. Here, we show that unsupervised *variational auto-encoders* (VAE) can alleviate all of these limitations. We perform the first unsupervised construction of an abstract timbre space by relying on VAEs to organize a target latent space. We introduce a specific regularization of this space with perceptual constraints, allowing to directly enforce a given perceptual metric or human similarity ratings onto these spaces. We compare our approach to existing timbre spaces and study how the usual acoustic descriptors behave along our non-linear latent dimensions. Finally, we exhibit the generative capabilities of the VAE by generating sounds with smooth evolution of timbre perception.

Background information

Music perception research has demonstrated that timbre could be partially described through acoustical descriptors of different sounds, by studying multidimensional timbre spaces (Grey, 1975). These timbre spaces have been explored based on the idea of using multidimensional scaling to organize perceptual ratings and correlate spectral descriptors as predictive dimensions of timbre perception (McAdams, 1995). Although these spaces provide interesting avenues of analysis, they are inherently limited by:

(1) ordination techniques (MDS) produce fixed spaces that must be recomputed for any new point:
(2) interpretation is bounded by descriptors rather than descriptions derived directly from the analysis;
(3) these spaces are unable to generalize on novel points or generate data from these spaces.

Deep generative models challenge these constraints and allow us to specify any distribution we want to match for both inference and generation. In partiucular, we use here *Variational Auto-Encoders* (VAEs) (Kingma, 2013) that jointly learns an inference distribution $q_{\phi}(\boldsymbol{z} \mid \boldsymbol{x})$ (inferring hidden factors of variation $\boldsymbol{z}$ based on a given input $\boldsymbol{x}$) and a *generative* process $p_{\theta}(\boldsymbol{x} \mid \boldsymbol{z})$ (regenerating input $\boldsymbol{x}$ based on factors $\boldsymbol{z}$). Parameters of both distributions are computed with deep neural networks paremeterized respectively with parameters θ and ϕ, learned through the *evidence lower-bound* (ELBO)

$$\mathcal{L}_{\theta,\phi,\mathbf{x},\mathbf{z}} = \mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})}\left[\log p_{\theta}(\mathbf{x} \mid \mathbf{z})\right] - \beta \mathcal{D}_{KL}\left(q_{\phi}(\mathbf{z} \mid \mathbf{x}) \| p(\mathbf{z})\right)$$

This criterion optimizes jointly the reconstruction error of the original data and the divergence between a specified prior $p(\boldsymbol{z})$ and the inference distribution $q_{\phi}(\boldsymbol{z} \mid \boldsymbol{x})$ learnt during training. The β parameter allows for a balance between the reconstruction and the regularization criterion, so the learnt representation is more likely to capture the inner factors of variation of the data. With this simple ELBO, the organization of the *latent space* ($\boldsymbol{z}$) is ruled by inherent mutual-information constraints between the data and latent spaces. This yields drastically compressed while smooth representations that allows to sample realistic data anywhere from the latent space. Furthermore, adding specific regularization $R_{\phi,\theta}(\boldsymbol{x},\boldsymbol{z})$ to the above objective allows us to enforce the latent space to exhibit specific properties

$$\mathcal{L}_{\theta,\phi,\mathbf{x},\mathbf{z},\beta} = \mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})}\left[\log p_{\theta}(\mathbf{x} \mid \mathbf{z})\right] - \beta \mathcal{D}_{KL}\left(q_{\phi}(\mathbf{z} \mid \mathbf{x}) \| p(\mathbf{z})\right) + \mathcal{R}_{\phi,\theta}(\mathbf{x},\mathbf{z})$$

Hence, we can drive each axis of this latent space to maximize the derivative of user-defined *attribute functions* in the data space, hence building a latent space with arbitrary metric properties.

Methodology

Datasets. We rely on both the *Studio On Line* (SOL) database and data coming from previous timbre spaces studies. We selected 7,200 samples to represent 12 different instruments (Piano, Cello, Violin, Flute, Clarinet, Trombone, Horn, Oboe, Saxophone, Trumpet, Tuba), with 10 playing styles for each. We normalized the range of notes used by taking common tessitura, and samples annotated with the same intensities (to remove effects from the pitch and loudness). Each instrument is represented by 600 samples equally spread across styles. Regarding the data from timbre studies, we rely on samples and perceptual ratings from (Grey, 1975; McAdams, 1995). All datasets are detailed in corresponding articles.

Learning. VAEs are trained on normalized Modulation Spectrum (MPS) of both datasets that are then fed into a multilinear-PCA, with corpus-wide normalization to preserve relative intensities of samples. First, VAEs are learned on subsets grouped by instrumental families, in order to see if the corresponding latent spaces preserve local timbral properties and show a perceptual coherent behavior among their axis. Then, we proceed on the full dataset and check if the representation obtained still fulfills these properties.

Regularization. We regularize the latent space by injecting the perceptual ratings, through the term $R_{\phi,\theta}(\boldsymbol{x}, \boldsymbol{z})$ that imposes that the distance in latent space is proportional to similarity ratings. As we do not have this information for all data, we add this term only for the corresponding samples. The learning still gains from learning on examples without ratings as it helps the regularization of the latent space.

Evaluation. We analyze the number of relevant latent dimensions, and perform manifold reduction techniques to recover an exploitable representation. We evaluate the quality of the representation by:

(1) Analyzing the repartition of the data projections with respect to acoustical descriptors;
(2) Performing a classification test based on either the original spectrum or the latent representation, in order to evaluate which yields better classification performance (indicating higher disentanglement);
(3) Comparing state-of-the-art timbre spaces and our unsupervised spaces;
(4) Evaluating the reconstruction quality and auditory consistency among the latent axes.

Results

Comparing unsupervised spaces from different family of instruments. We observe that the latent spaces obtained for each instrument separately, and on the whole dataset show high consistency between each other but also with previous human studies, even without performing regularization. This confirms our hypothesis that timbre spaces can be constructed in an unsupervised fashion.

Perceptually-regularized spaces between all instruments. The further timbre spaces that are perceptually-regularized provide almost identical relations as those coming from the state-of-art timbre studies. However, as underlined these spaces can now be used to analyze novel instruments without ratings.

Analyzing the evolution of signal descriptors across the latent space. By visualizing how the traditional descriptors behave when we perform a traversal of the unsupervised latent spaces, we observe that they have a non-linear behavior in our spaces. This tends to confirm our hypothesis that they are insufficient.

Generating from the timbre space. From the latent spaces, we directly obtain smooth morphing between the timbres of different instruments while remaining perceptually relevant, which allows to turn our analytical spaces into generating synthesizers that could potentially feed further perceptual studies.

Conclusions

Unsupervised variational learning provides a space that is relevant to timbral information with respect to different criteria. The capacity of these models to generalize to different instrument families and the ability to analyze the variance with respect to a whole learnt corpus allows for the extrapolation of perceptual results to new sounds and instruments without the need to collect new measurements. The organization of these latent spaces given a human perceptual metric proves the their flexibility and provides a perceptually regularized latent space from which generation of audio content is straightforward.

Acknowledgments

This work benefited from funding of the MAKIMOno 17-CE38-0015-01 project of the French ANR.

References

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *JASA, 61*(5), 1270-1277.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres. *Psychological Research*, *58*(3), 177-192.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Visualizing Orchestration with Orcheil

Gregory Lee Newsome
Faculty of Music, University of Toronto, Toronto, Ontario, Canada
greg.newsome@utoronto.ca

Aims/goals

Orcheil is a web application for the visualization of orchestration. The goal of the application is to provide a clear and immediate representation of when an instrument is or is not sounding in a piece of music. This can reveal, for instance, the instrumentation of a composite timbre, or the role of orchestration in articulating musical form. Orcheil has a straightforward interface, and is driven by symbolic music notation data, specifically a MusicXML file.

Why might this type of visualization be useful? The study of orchestration is a time-consuming endeavour whose methodology is largely static, consisting of critical listening, score study, and the historic work of Berlioz, Rimsky-Korsakov, Forsyth, Piston, and Adler. While the effectiveness of this methodology is proven, there is I believe an opportunity to complement it with data visualization, and it is for this reason that I have built Orcheil.

Background information

Orcheil can be thought of as a generalized Dolan Graph generator, and I am therefore indebted to the work of Dr. Emily Dolan. A Dolan Graph consists of multiple sets of line segments, with each horizontal set representing an instrument part. The x-axis denotes time, and the y-axis the work's instrumentation. Over time, the line segments for each part may start and stop—when they are present, the instrument is sounding, when they are not, the instrument is resting. There is no pitch information, rhythmic detail, or performance direction present in the graph.

I have also drawn upon Dr. Meghan Goodchild's doctoral dissertation, which contains a wealth of visualization methodology from a music perception perspective. With respect to existing software in this area, there is a music visualization application, the MoshViz package, focusing on harmonic and melodic pattern recognition. However, while basic orchestration information is present in MoshViz, it is secondary to the visualization of pitch.

Methodology

The methodology of creating Orcheil was essentially the software development process—deciding upon platform support, committing to a feature set, designing the data structures and user interface, reviewing and testing various frameworks and libraries, and implementing, testing, and deploying the finished application.

I decided to make Orcheil a web application to simplify user interaction with it, and to ensure cross-platform compatibility. The musicology toolkit music21 was chosen for inclusion because it has a mature symbolic music notation parser for MusicXML. While music21 can also plot a Dolan Graph, Orcheil does not use this functionality, instead drawing the visualization directly in the web browser using Scalable Vector Graphics (SVG). The web framework Flask was selected for its simplicity and versatility, and like music21, it is written in Python, making for a consistent application codebase.

Results

The product of research is the web application Orcheil, which may be found at <http://orcheil.ca/>, though in the strictest sense the result is the application codebase, which is freely available under the BSD 2-Clause License at <https://github.com/GLN/orcheil>.

Figure 1: An Orcheil visualization of Gustav Mahler's *Kindertotenlieder I*

Conclusions

An Orcheil visualization can offer a more immediate understanding of a work's orchestration, whether for analysis or pedagogy, than gleaning it from a score or a recording. This is especially true for complex music with many instruments. The role of timbre is at the forefront of an Orcheil visualization—it is an orchestration-first representation of music.

There is also potential for corpus analysis using Orcheil, for instance to assess the use of a specific vertical relationship in the complete output of a given composer, or to examine brightening-darkening technique in a certain historical period.

Acknowledgments

The Orcheil web application is based upon an original concept by Dr. Emily Dolan, from a conversation between us at the ACTOR Partnership Meeting at McGill University in Montreal, August 3-4, 2017.

Orcheil is built with the programming language Python, the web framework Flask, and the musicology toolkit music21.

References

Cantareira, G. D., Nonato, L. G., & Paulovich, F. V. (2016). MoshViz: a detail+overview approach to visualize music elements. *IEEE Transactions on Multimedia, vol. 18, no. 11*, pp. 2238-2246.

Dolan, E. (2013). *The Orchestral Revolution: Haydn and the Technologies of Timbre*. Cambridge: Cambridge University Press.

Goodchild, M. (2016). *Orchestral Gestures: Music-Theoretical Perspectives and Emotional Responses* (doctoral dissertation). McGill University, Montréal.

Drowning into Synthesis: The Role of Timbre in Fausto Romitelli's Compositional Techniques

Luca Guidarini
Dipartimento di Musicologia e Beni Culturali, University of Pavia, Cremona, Italy
luca.guidarini@gmail.com

Aims/goals

My presentation focuses on Fausto Romitelli's (1963-2004) compositional techniques, which I have investigated for my dissertation, analyzing the case of *An Index of Metals* (2003)*, Hellucination I – Drowningirl*. This section of the video-opera presents a situation in which the morphological quality of the timbre configures the piece structurally: the so-called instrumental synthesis realized by the ensemble imitates a specific case of FM synthesis in which the modulator frequency comes progressively closer to the carrier frequency. In so doing the simulation process, together with its corresponding electronics, assumes the architectural principle of the composition.

Background information

Since the first collection of essays devoted to his work (Arbo 2003), Fausto Romitelli's compositional output has been investigated aesthetically and philosophically. His compositional language includes elements from rock and electronic popular music, but in a more explicit way than, for example, Murail, who writes, "The spectacular development of synthesizers, of electronic sound, owes considerably more to Pink Floyd than to Stockhausen" (Murail 1980, p. 79). Like Pink Floyd's Richard Wright, Romitelli invested heavily in the unique sound properties of synthesizers, which served as the compositional starting point for *An Index of Metals, Hellucination I – Drowningirl.* Alessandro Arbo, in his article, *Fausto Romitelli: Six Keywords* (2014) considers "sound" as the first word that could describe Fausto Romitelli's compositional style. Arbo remarks that, "He [Romitelli] drew inspiration from the about-turn effected by the composers of the *Itinéraire*, in the wake of other important 20th-century composers. Much more than 'compose with sounds', what was at issue, for him, was to 'compose sound'" (2014, p. 42), in a different vein when compared to the so-called *Polish school*, his work seems rather more influenced by some experimental rock bands from the 60s and 70s or by the style of popular electronic music producers of the 90s, such as *The Jimi Hendrix Experience,* Robert Fripp, or *Aphex Twin*. The relationship between the spectral techniques and sonic-oriented popular music defines the opening moments of *An index of Metals*: the first sound heard is a sample from Pink Floyd's *Shine on You Crazy Diamond*, which then becomes a structural element in the composition. On a deeper level, Romitelli's use of electronic and *synthetic-sculpted sound* shape his compositional techniques, creating a personal approach to instrumental synthesis.

Methodology

As in the cases of Luigi Nono and Bruno Maderna, who "did not publish their theoretical procedures, a question arises of how to define the technical procedures, the study of sketches becomes essential to grasp the technical issues to which the works from time to time gave an answer" (Borio, 1999, p. 7). Contrary to these most known figures, Romitelli's premature death means his sketches becomes indispensable to anyone seeking a deeper appreciation of his music and compositional techniques.

By examining his archive, currently preserved by the Giorgio Cini Foundation in Venice, I was able to examine the process of *listening – analyzing – composing* the timbre that shaped *Hellucination I – Drowningirl*, in order to analyze the musical results on the score, revealing Romitelli's conception of timbre in the compositional process.

Results

Roy Lichtenstein's *Drowning Girl* inspired Kenka Lekovich to write *Metalsushi*, from which the poetic text for *An Index of Metals* was extracted. The image and the sensation of drowning, of sliding into something that is "an initiatory celebration of the metamorphosis and fusion of matter, a light show, in which an extension of the perception of the self beyond the physical limits of the body is provoked by means of techniques of transference and fusion in an alien material. It is a path towards perceptive saturation and hypnosis, one of total alteration of the habitual sensorial parameters" (Milanaccio, n.d.) is explicit in both the text and in the music of *Hellucination I – Drowningirl.*

Since the text is a linguistic evocation of the sensation of sinking into this hypersensitive experience the musical material representative of this phenomenon is hard to describe analytically. The study and the analysis of the sketches enabled me to clarify the musical and synthesis techniques that points towards these expressive results and connects the compositional process of this section with the composer's interest for the frequency modulation synthesis, that enables him to create a complex world of harmonic and inharmonic sounds. Unlike Murail's *Gondwana*, in which FM synthesis was only a kind of "pitch generator machine," in *Hellucination I – Drowningirl* has structural functions, since the ensemble is seen as an imaginary synthesizer with FM. Starting from a fixed and implicit carrier frequency, this section of *An Index of Metals* simulates the process by which the modulator frequency becomes progressively closer to the carrier one, creating a long, structural, and lively glissando.

Conclusions

The timbre, which in this specific example of Romitelli's music, is not only a qualitative and multidimensional element of the music, but also a process of its gradual definition over time is the main formal dimension for the development of an entire section of *An Index of Metals*. New strategies to control gestures, pitches, dynamics and density are used, in order to generate and handle the timbre in unprecedented and original ways.

References

Arbo, Alessandro (Ed.) (2003). *Il corpo elettrico. Viaggio nel suono di Fausto Romitelli*, Monfalcone, Il teatro

Arbo, Alessandro (2014). Fausto Romitelli's Six Keywords. *On the Page*, Universal Music Publishing Classical, 42-45

Borio, Gianmario (1999). Sull'interazione fra lo studio degli schizzi e l'analisi dell'opera. In G. Borio, G. Morelli and V. Rizzardi (Eds.), *La nuova ricerca sull'opera di Luigi Nono*, Archivio Luigi Nono (Studi I), Firenze: Olschki, 1-21

Milanaccio, Roberta (trad. by Ward-Perkins, Hugh), Fausto Romitelli. Universal Music Publishing Classical, <http://www.umpgclassical.com/en-GB/Composers/R/Romitelli-Fausto.aspx> (accessed 12-13-2017)

Pustijanac, Ingrid (2009). La Ricezione del pensiero spettrale negli ultimi due decenni in Italia. *I*n A. Collisani, G. Garilli & G. Mercadante (Eds.), *Italia – Francia. Musica e cultura nella seconda metà del XX secolo,* L'epos, pp. 347-372

Rose, Francois (1996). Introduction to the Pitch Organization of French Spectral Music. *Perspectives of New Music* 34(2), 6-39.

Murail, Tristan (2005). The Revolution of Complex Sounds. *Contemporary Music Review*, 24(2-3), 121-135. (Ed. orig. Murail, Tristan (1980) La revolution des sons complexes. *Darmstddter Beitrdge zur Neuen Musik*, Mainz, London, New York, B. Schott's Sohne, 77-92).

Truax, Barry (1977). Organizational Techniques for C:M Ratios in Frequency Modulation. *Computer Music Journal,* Vol. 1, No. 4, 39-45.

The Sound of Colour in R. Murray Schafer's (1933-) Seventh String Quartet (1998)

Lauro Pecktor de Oliveira
School of Creative and Performing Arts, University of Calgary, Calgary, Alberta, Canada
lauro.pecktordeolive@ucalgary.ca

Aims/goals

I will examine how R. Murray Schafer (1933-) dealt with sound colour in the creative process of his Seventh String Quartet (1998), notably the semantic relationship between the use of visual colours in Schafer's working documents and the timbre of the resulting work.

Background information

This research is inserted in the array of studies dedicated to the exploring of the creative process of musical works. Even though genetic criticism originated in literature, there has been a steady and growing movement expanding its use towards other creative areas. In music, several scholars worldwide have been contributing to this expansion and adaptation of genetic criticism (Chaves, 2012; Donin, 2012; Katz, 2012; Kinderman, 2012; Sallis, 2015). An example of this expanding movement can be observed in the international biennial conference *Tracking the Creative Process in Music*, organized by Nicolas Donin, which, in every edition, counts with an ever-growing number of international participants. It is my understanding that this research that I am presenting is the first genetic study on R. Murray Schafer's Seventh String Quartet.

Methodology

The Seventh String Quartet's working documents were firstly approached by categorizing the documents, and secondly by creating hypotheses of the creative process within the documents. The categorization is straightforward and usually demands an attention to the nature of the documents; roughly, this stage involves establishing a chronological order to the documents and assessing their degrees of completeness. It is in this first stage that a working document is defined as a sketch, a draft, a fair-copy, or whatever kind of suitable definition in accordance to its nature. The first stage helps giving an overall view of the documents and sets the ground for the elaboration of hypotheses of the creative path. In this sense, and in its expansion to music, this research resounds with questions elaborated by Nicolas Donin (2009) on how a genetic approach to a musical piece should be:

- What chains of compositional operations gave rise to the work as we know it today?
- What type of singular musical logic can be reconstructed on this basis, without simply presenting the chronology of the creative process or analyses oriented toward structural listening?
- How can our knowledge of the genesis of the work modify our perception to it?

Results

The working documents of Schafer's Seventh String Quartet reveal a series of dialogues between the composer and his intertextualities (such as the myth of *Tapio*, the colours of Guido Molinari's paintings, the text by a mental hospital patient, etc.), between the documents and their building-blocks (pitches, harmonies, rhythms, gestures, etc.), between external constraints by commissioners (for example, the addition of the soprano, which was not part of the original idea for the piece). Within these dialogues perhaps the most appealing to the eye when studying the working documents is the use of colour; it seems to be the composer's way to extrapolate the standard notation to a more free and independently creative exploration of sound, and to, in a personal way, relate with another artist and another art form: Guido Molinari (1933-2004) and his paintings. This approach not only gives an alternative notation to timbre, but also enables each performer to discover ways to signify *visual* colour into *sound* colour. Schafer's interest in the use of colour in his Seventh Quartet was so great by the end of its creative process that he composer decided to ask the performers of the Molinari Quartet (who commissioned and premiered the piece) to dress in different colours: the first violinist in red, second violinist in blue, violist in yellow, cellist in green, and soprano in white.

Conclusions

The use of genetic criticism in this research was crucial for the organization and understanding of these working documents. There is still much to be unveiled from these documents not only related to the Seventh String Quartet, but also to R. Murray Schafer's oeuvre in general.

Acknowledgments

I would like to acknowledge Regina Landwehr and Allison Wagner from the Archives and Special Collections of the University of Calgary for their assistance in this research. I would also like to acknowledge Dr. Friedemann Sallis for all the help and support in this endeavour.

References

Chaves, C. G. L. (2012). Processo Criativo e Composição Musical: Proposta para uma Crítica Genética em Música. *Congresso Internacional da Associação de Pesquisadores em Crítica Genética, X Edição*. 237-245.

Donin, N. (2012). Empirical and Historical Musicologies of Compositional Processes: Towards a Cross-fertilisation. In Dave Collins (Ed.) *The Act of Musical Composition: Studies in the creative process*. 1-26. New York: Routledge.

Donin, N. (2009). Genetic Criticism and Cognitive Anthropology: A Reconstruction of Philippe Leroux's Compositional Process for Voi(rex). In William Kinderman and Joseph E. Jones (Ed.) *Genetic Criticism and the Creative Process*. 192-215. Rochester: University of Rochester Press.

Ferrer, D. (2009). Variant and Variation: Toward a Freudo-bathmologico-Bakhtino- Goodmanian Genetic Model?. In William Kinderman and Joseph E. Jones (Ed.) *Genetic Criticism and the Creative Process*. 35-50. Rochester: University of Rochester Press.

Katz, S. L. (2012). The Influence of the Extra-musical on the Composing Process. In Dave Collins (Ed.) *The Act of Musical Composition: Studies in the Creative Process*. 171-186. New York: Routledge.

Kinderman, W. (2012). *The Creative Process in Music from Mozart to Kurtág*. Champaign: University of Illinois Press.

Lind, S. (2009). Interpreting Gesture as Motive: A Transformational Perspective on Replication in R. Murray Schafer's Seventh String Quartet. *Intersections* 28 (2): 51-71

Sallis, F. (2015). *Music Sketches*. Cambridge Introductions to Music. Cambridge: Cambridge University Press.

Sallis, F. (2004). Coming to terms with the composer's working manuscripts. In Friedemann Sallis and Patricia Hall (Ed.) *A Handbook to Twentieth-Century Manuscripts*. 43-58. Cambridge: Cambridge University Press.

Schafer, R. M. (2014). *My Life on Earth & Elsewhere.* Ontario: Porcupine's Quill.

Schafer, R. M. (2000). *String Quartets 1-7*, Interpreted by the Quatuor Molinari and Marie-Danielle Parent, soprano, CD, ATMA Classique ACD 2 2188/89 (double album).

Schafer, R. M. (1998). *Seventh String Quartet*. Ontario: Arcana Editions.

Schafer, R. M. (1998). Seventh String Quartet, with soprano obligato. *University of Calgary, Archives and Special Collections, The Richard Johnston Canadian Music Archives Collection, The R. Murray Schafer Fonds, Accession No.716/02.3, 2.1*.

Timbre as 'Structuring Structure' in Underground Electronic Dance Music

Maria Perevedentseva
Department of Music, Goldsmiths, University of London, UK
mpere001@gold.ac.uk

Aims/goals

This poster offers a multi-modal formal and semiotic analysis of timbre in Objekt's 2017 single "Theme from Q" in order to investigate how timbre in underground electronic dance music (UEDM) functions as a sign, and what kinds of musical, conceptual and embodied knowledge it activates. I argue that timbre occupies a central meaning-making site in UEDM, acting as a distilled signifier of genre and encoding notions of value, authenticity, public, place and time. My principal aims, therefore, are to explore what can be gained from attending to timbre in UEDM and to identify an effective means of accessing and interpreting its affordances.

Background information

With very few exceptions (e.g., Fales, 2005), literature on dance music has been curiously silent on the issue of timbre and as recently as 2014, Mark Butler wrote that "methodologies and theoretical frameworks for discussing [UEDM's] sonic dimensions simply [do] not exist" (p.15). Yet, outside the academy, notions of timbre permeate UEDM discourse, and have been shown – in rock music culture at least – to perform important genre- and self-identifying functions for listeners (Van Elferen, 2018). This suggests that timbre plays a significant role in UEDM fans' appreciation of their music of choice and warrants further investigation. The largely non-source-oriented nature of synthesised sound has repeatedly stalled attempts to formalise a semantics of electronic music. However, cognition of obscure timbres has, in the electroacoustic sphere, been successfully theorised as relying equally on the physical properties of the sound, environmental factors, and on listeners' willingness to recall their past sonic experiences in order to fill in any perceptual gaps (Windsor, 2000). Framed in this way, timbre presents a unique opportunity to reappraise "from within" the dynamic musical and social structures that underpin UEDM culture.

Methodology

In this poster, I use Ben Curry's (2017) adaptation of Peirce's concept of simultaneous bonds to analyse selected timbral gestures in "Theme from Q." The track is notable – though not unusual – for its stylistic hybridity, featuring Deep House organs, slap bass licks, sci-fi pads, breakbeats and Speed Garage basslines. Curry (ibid.) writes that for a musical sign to generate meaning, it must simultaneously signify as an icon (formally), an index (actually) and a symbol (conventionally) – a claim I follow in my own analysis. I use spectrograms and close listening to explore the pure physical characteristics of a given timbre. At the indexical level, I consider how the timbre interacts with the musical ecology (the virtual environment created by the intrinsic musical structure of the track) of "Theme from Q" as a whole. Lastly, at the symbolic level, I look at how its deployment here relates to past conventional uses of that timbre, both within UEDM and in other electronic musics. My longitudinal approach adapts Denis Smalley's (1997) classic study of sound-shape perception and transformation to identify the social and technological catalysts for timbral change over time—a kind of historical spectromorphology.

Results

My results bring into sharper relief the rich tapestry of timbral signifiers in "Theme from Q." Spectral analyses comparing individual timbres in this track to their historical uses reveal how Objekt's sound design digresses from and subverts historical models, relying on these "canonic" gestures being recognised as such and perceptually filled in by his listeners. A bassline gesture which appears twice in the track is especially interesting, as the spectromorphological transformation it undergoes in both of its entries appears to condense into only a few seconds' duration a longer historical shift from the timbral composition of basslines in mid-1990s House and early Garage into those found in Dark Garage and Dubstep in the following decade. This dilation from a closed, hollow timbre to a wide, squelchy spectrum represents one of the most primal gestures in electronic dance music—or indeed any music created using instruments with rotary filters—and is a defining characteristic of Acid House, a North-American genre widely regarded as having kick-started the global rave movement. In this way, this single timbral gesture conceptually unites the disparate stylistic threads running throughout "Theme from Q," as well as situating the track at a point in UEDM's historical continuum that allows for this kind of creative self-reflection and reconstruction.

Conclusions

The idiosyncratic musical and social make-up of underground electronic dance music culture enables timbre to activate certain known and knowable meanings and affordances, and my methodology in this poster has suggested some ways in which these meaning structures can be accessed and articulated. The tightly-woven self-referentiality of UEDM allows me to bypass the original obstacle to cognition in acousmatic sound (that of source-bonding to "real" instruments and environments) and instead penetrate the, by now, highly codified meanings attached to timbres in this scene. I hope to have shown that, in attending to timbre in UEDM, the listener and analyst can uncover multiple layers of signification which convey vast amounts of information, expand and contract temporal planes and mine the listeners' own musical and affective knowledge to create a holistic and perceptually stable listening experience. Due to space constraints, I have not been able to interrogate in any depth the issues surrounding the language used by analysts and listeners to describe timbre, but this remains an important avenue for future research.

References

Butler, M. (2014). *Playing with something that runs. Technology, improvisation and composition in DJ and laptop performance*. New York, NY: Oxford University Press.

Curry, B. (2017). 'Valency-actuality-meaning. A Peircean semiotic approach to music', *Journal of the Royal Musical Association*, 142(2), 401-443.

Fales, C. (2005). 'Short-circuiting perceptual systems. Timbre in ambient and techno music' in P. Greene & T. Porcello (Eds.), *Wired for sound. Engineering and technologies in sonic cultures* (pp.156-180). Middletown, CT: Wesleyan University Press.

Objekt (2017), 'Theme from Q' on *Objekt #4* [Vinyl]. Berlin, Germany: Objekt.

Smalley, D. (1997). 'Spectromorphology. Explaining sound shapes', *Organised Sound*, 2(2), 107-126.

Van Elferen, I. (2018). 'Dark timbre: The aesthetics of tone colour in goth music', *Popular Music*, 37(1), 22-39.

Windsor, L. (2000). 'Through and around the acousmatic: the interpretation of electroacoustic sounds' in S. Emmerson (Ed.), *Music, electronic media and culture* (pp.7-35). Aldershot: Ashgate Publishing Limited.

Comparative Analysis Between Verbal Attributes of Timbre Perception and Acoustic Correlates in Contemporary Music Excerpts

Ivan Eiji Simurra†, Marcelo Gomes de Queiroz
Department of Computer Science/Sonology Research Center - University of São Paulo, São Paulo, Brazil
† ieysimurra@gmail.com

Aims/goals

The aim of the experiment is to identify significant acoustic correlates associated with verbal attributes classification of orchestral timbres which are based on technical aspects of contemporary music.

Background information

We present a further study of an ongoing research project focus on timbre characteristics of orchestral instruments by the agency of the relationship between verbal attributes and acoustic features. Our study is based on the literature related to the analysis of acoustic correlates of verbal attributes classification of music timbres. The study performed by Alluri and Toiviainen (2012) aimed to verify acoustic correlates for the polyphonic timbre perception in a cross-cultural context between Indian and Western Music. Zacharakis, Pastiadis and Reiss (2014) investigated the underlying structure of musical timbre semantics through an analysis of verbal descriptions of different timbres and a correlation analysis between extracted acoustic features and semantic dimensions indicated the prominent acoustic correlates. Wu, Horner and Lee (2014) attempted to identify the most perceptually relevant timbral attributes relating verbal emotion categories with acoustic features like spectral centroid and attack time. Siedenburg and Müllensiefen (2017) explored to what degree similarity judgments of short music clips can be modeled with low-level acoustic features related to timbre.

Methodology

The sound database was designed to address some aspects pertaining to contemporary music techniques mainly to create new sounds and textures (Barrière, 1991). The whole framework was based on music composed by Ravel, Debussy, Stravinsky, Messiaen, Schoenberg, Ligeti, Grisey, Scelsi, Lachenmann, Xenakis and Sciarrino. In total there are 33 sound stimuli each with a duration of 5.0 seconds. We divided all 33 sound stimuli into three different groupings in which: a) eleven stimuli based on the original transcription of the musical scores; b) eleven stimuli based on the variation of the musical instruments techniques maintaining the same set of musical instrumentation design; c) eleven stimuli based on the variation of the set of musical instrumentation design. This approach aims to cover timbral variations through both by modifying musical techniques and/or musical instrumentation even while maintaining some musical characteristics such as pitch and rhythm.

Results

In previous work (Simurra & Queiroz, 2017), we analyzed all auditory stimuli through Verbal Attributes Classification by Semantic Differential method (Osgood, 1964). Using Factor Analysis, we were able to identify major perceptual categories associated with tactile and visual properties such as Mass, Brightness, Color and Scattering. The method of the acoustic feature extraction in the present study was the common timbre- related computational features which are based on the literature such as Peeters (2004), Malt & Jourdan (2009) and Bogdanov et al (2013). A total of sixteen acoustic features were extracted using Libxtract (Bullock, 2007) and Librosa (McFee et. al., 2015). We calculated the correlation coefficient among all acoustic features. To calculate the correlation between multidimensional features such as MFCC, Contrast and Tristimulus concurrently with all the unidimensional acoustic features we calculated the RV Coefficient. Next, we applied Principal Component Analysis to reduce the multidimensional features. For each stimulus we extract the first PCA for MFCC, Contrast and Tristimulus. Then we calculate the Factor Analysis (FA) for all acoustic features to cluster them (Factor 1 = Centroid/Roll-Off/Standard Deviation; Factor 2 = Noisiness/Irregularity/RMS; Factor 3 = Tristimulus/Zero Crossing Rate; Factor 4 = Flux; Factor 5 = MFCC). Finally, we calculated the correlation coefficients between Factor Loadings from verbal attributes classification and all the acoustic correlates. In the following table, we present the correlation between these factors related to the all 33 auditory stimuli.

	Centroid/Roll-Off/Std	Noisiness/Irregularity/RMS	Tristimulus/Zero Crossing	Flux	MFCC
Mass	-0.19251337	-0.7476604	0.06116310	-0.38703209	-0.33856952
Brightness	0.52139037	0.2155749	0.090909	-0.10828877	0.09692513
Color	-0.20220588	0.3181818	-0.19284759	-0.01537433	0.41544118
Scattering	-0.05280749	-0.1052807	-0.31149733	0.05514706	-0.24498663

Conclusions

We presented a study of an exploratory experiment linking verbal attributes with acoustic correlates of a set of orchestral timbres based on contemporary music repertoire. Our first findings demonstrated that some acoustic features were related to the verbal attributes. The set consisting of Centroid/ Roll-Off/ Standard Deviation is related to Sound Brightness. Noisiness/Irregularity/RMS are close to Sound Mass. MFCC are related to the Sound Color and Tristimulus/Zero Crossing Rate are related to Sound Scattering. The main purpose of the investigation on this stage of the research focus on the assessment of the acoustic correlates of the vast sound universe quaintness belonged to modern and contemporary music associated with verbal attribute descriptions.

Acknowledgments

The authors would like to thank the funding agencies FAPESP and CNPq for their support.

References

Alluri, V., & Toiviainen, P. (2012). Effect of enculturation on the semantic and acoustic correlates of polyphonic timbre. *Music Perception: An Interdisciplinary Journal*, 297-310.

Barrière, J.-B. (1991) Le Timbre: métaphore pour la composition, Christian Bourgois.

Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O. & Serra, X. (2013). Essentia: an open-source library for sound and music analysis. In *21st ACM Conference* (pp. 855-858).

Bullock, J. (2007). Libxtract: a Lightweight Library for audio Feature Extraction. In *ICMC*.

Chau, C.-j., Wu, B., and Horner, A. (2015). The Effects of Early-Release on Emotion Characteristics and Timbre in Non-Sustaining Musical Instrument Tones. In ICMC.

Malt, M., & Jourdan, E. (2009). Real-time uses of low level sound descriptors as event detection functions using the max/msp zsa. descriptors library. *12th Brazilian Smposium on Computer Music*.

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). librosa: Audio and music signal analysis in python. In *14th python in science conference* (pp. 18-25).

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1964). *The measurement of meaning*. Peeters, G. (2004). A large set of audio features for sound description in the CUIDADO project.

Siedenburg, K., & Müllensiefen, D. (2017). Modeling timbre similarity of short music clips. *Frontiers*

Simurra, I. E. & Queiroz, M. (2017). Pilot Experiment on Verbal Attributes Classification of Orchestral Timbres. In *Audio Engineering Society Convention 143*. Audio Engineering Society.

Wu, B., Horner, A., & Lee, C. (2014). Musical timbre and emotion: The identification of salient timbral features in sustained musical instrument tones equalized in attack time and spectral centroid. In *ICMC*.

Zacharakis, A., Pastiadis, K., and Reiss, J. D. (2014). An interlanguage study of musical timbre semantic dimensions and their acoustic correlates. Music Perception: An Interdisciplinary Journal, pp. 339–358.

箏の唱歌: Phonetic Encoding of Musical Contour in the Traditional Ikuta Scores of Japan

Tanor Bonin
Music Research, McGill University, Montreal, QC, Canada
† tanor.bonin@mail.mcgill.ca

Aims/goals

In addition to the tonal and metrical information codified in standard Western music notation, the traditional scores of the Ikuta school of Japan also exhibit phonetic transcriptions of *timbral* structure and development within the music. This timbral transcription is known as the *shouga* (唱歌). The focus of this research was to analyze the structural relations between the timbral, tonal, and temporal dimensions of a traditional koto piece, Rokudan, by the master YATSUHASHI Kengyo (1614–1685), as conveyed in the Ikuta scores.

Background information

In the summer of 2017 I had the honour and privilege of studying traditional Japanese koto performance with sensei ANDO Masateru. The koto (箏) is the national instrument of Japan, a 13-string zither first introduced from China in the early Nara period (8th century). ANDO-sensei generously offered me one of his instruments and two lessons each week for the duration of a three-month fellowship in Japan. As the lessons were conducted in Japanese and I possess only limited elementary Japanese speech skills, my learning the instrument and Japanese musical style derived primarily from a sensitivity (感性) to the sonic, postural, and aesthetic-emotional patterns of the music. I thus present all intellectual descriptions of the musical experience, musical structure and theory, and performance techniques from this point of reference.

Each of the strings of the koto is tuned with a moveable bridge called a *ji* and is activated with three ivory plectra on the right thumb, index, and middle fingers, called *tsume*. Intimate control over the pitch and timbre of the performance is afforded by a formal lexicon of articulation techniques for the right hand and the complementary pitch modifying and tuning techniques executed by the left hand. Performance errors in the "timbral" or "tuning" dimensions of my playing were not addressed with differentiated concepts; understanding of the correct sound was to be achieved only through proper holistic execution. In time, it occurred to me that this perspective was thoroughly reflected in the notation of the traditional Ikuta scores.

The Ikuta school (*ryu*) was founded by Ikuta Kengyo (1666–1716) and concerned the merging of the koto and shamisen music of the Osaka region. The scores of the Ikuta school include phonetic notation of the music's timbral form (called *shouga*) in addition to the tonal and metrical notations commonly found in standard Western notation. I became fascinated with the intuitive structural correspondence between the shouga and melodic phrases within the music, particularly in the work titled *Rokudan* (六段の調; 'Six steps' [of being]) by Yatsuhashi Kengyo. Though I was unable to complete all six movements of this piece before my departure from Japan, I felt compelled to further explore this work to whatever extent remained possible by the means available to me. I asked for ANDO-sensei's blessing and decided to turn to a structural analysis of the music's form.

Methodology

Each note-event in Rokudan was represented in a 3-dimensional melodic x metrical x timbral (phonetic) space. Using IDyOM (Pearce, 2005), "a framework for constructing multiple-viewpoint variable-order Markov models for predictive statistical modelling of musical structure," a conditional probability distribution represented the estimated likelihood of each note-event on the basis of Shannon entropy metrics derived from preceding notes and note-sequences. Auto-correlative and coincident relations between the piece's melodic, metrical, and timbral dimensions were then summarized in terms of functional and structural kernels within the musical form, allowing for statistical descriptions of melodic, metrical, and timbral covariance throughout the musical piece.

Results

Final results are forthcoming in the Spring of 2018. Preliminary results indicate that the "timbral arcs" codified in the *shouga* of the musical score are directly related to the phase of both the melodic contour and the metrical boundaries of the musical form.

Conclusions

This work indicates that the timbral, tonal, and temporal structures of traditional koto music are intimately connected, and that these interrelations are effectively relayed in the notation conventions of the traditional Ikuta scores. Future work may extend this analysis to a larger corpus of traditional Japanese music, and may also consider whether similar structural relations exist in other musical genres, such as those of the West, despite the fact that they are not explicitly represented in codified formulations within the score. Such a possibility may bear out in the orchestration of classical works, the tone shaping and phrasing of popular electric guitar- and synthesizer-based music, or the generative constraints of experimental and improvisational musics.

Acknowledgments

This research was supported by NSERC CGS-D and JSPS Doctoral Fellowships awarded to TB.

References

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition* (Doctoral dissertation, City University London).

Ando, M. (1986). *Koto of the Ikuta School* (生田流の箏曲). Ed. Yoshikawa Hisashi.

Using Bubble Noise to Identify Time-Frequency Regions that are Important for Timbre Recognition

Michael I. Mandel[†1] and Song Hui Chon[2]
[1] Computer and Information Science, Brooklyn College, CUNY, NY, NY, USA
[2] Electrical, Computer & Telecommunications Engineering Technology,
Rochester Institute of Technology, Rochester, NY, USA
[†] mim@sci.brooklyn.cuny.edu

Aims/goals

It has traditionally been difficult to identify the perceptual correlates of musical timbre, which seem to vary according to the experimental context. This work investigates these correlates using a listening test and analysis paradigm that we have recently introduced for identifying time-frequency regions of a sound that are important for its correct identification (Mandel et al., 2016). By applying this technique to an instrument recognition task, we are able to characterize the regions of individual recorded notes that allow a listener to correctly identify the instrument playing it, i.e., the regions of important timbral content.

Background information

Our approach is based on the visual "bubbles" technique (Gosselin & Schyns, 2001) which aims to study the cues used by biological visual systems. It measures the ability of subjects to perform a visual identification task when an image is occluded except for a small set of randomly placed "bubbles." Mandel et al. (2016) first described using bubble noise to analyze speech perception, identifying the time-frequency regions important to phoneme recognition in vowel-consonant-vowel nonsense words. They showed that the cues identified correspond well with those identified in general by acoustic phonetics and that different listeners exhibit similar important regions. Thoret et al. (2016) utilize a somewhat related bubble technique for studying timbre. They apply a single bubble in the modulation spectrum domain per stimulus. Instead of mixing a stimulus with bubble noise, they apply a specific modulation filter to it and find that musical instrument timbres may be characterized by specific spectrotemporal modulations.

Methodology

The bubble noise technique described by Mandel et al. (2016) measures the "intelligibility" of a sound when mixed with many different instances of "bubble noise." Bubble noise is designed to provide random combinations of time-frequency "glimpses" of a note of interest through a loud noise field. The noise is constructed by attenuating elliptical bubbles in a uniform noise field with a spectrum matching that of the clean stimuli. The locations of these bubbles are selected uniformly at random in time and ERB-scale frequency and their duration and bandwidth (in ERB) are fixed as well. The target sounds are isolated notes from the Vienna Symphonic Library and have been normalized to be the same loudness and duration. All notes are either C4 or E4, but come from one of four instruments: clarinet, French horn, oboe, and trumpet. Using two different pitches allows us to determine whether the important regions follow the harmonics, which change with the fundamental frequency, or remain fixed and thereby "follow" the fixed resonances of the instrument. The listener is presented with one note mixed with bubble noise and must select the instrument that played it. The four instrument sounds for a specific listener were all on one pitch—either C4 or E4 but not both. To analyze the results, we measure the correlation across many different noise instances between whether a given mixture was correctly identified and the audibility at each time-frequency point. If a given time-frequency point is important for identifying the timbre of a given note, then it should be audible in mixtures in which the instrument was correctly identified and inaudible in those that were not correctly identified. We quantify this relationship using the correlation coefficient and its statistical significance according to a one-sided ANOVA with two levels. The experiment was carried out with 12 undergraduate students enrolled in the acoustics course at RIT. Each listener heard 200 mixtures of bubble noise with four notes, either all of the instruments playing C4 or E4, taking approximately 75 minutes to listen and respond to all 800 mixtures. Due to variability in students' musical backgrounds, a large range of accuracies was observed (average 32% to 69%). After excluding six low accuracy students, we analyzed the results from the six remaining students (two on C4, four on E4).

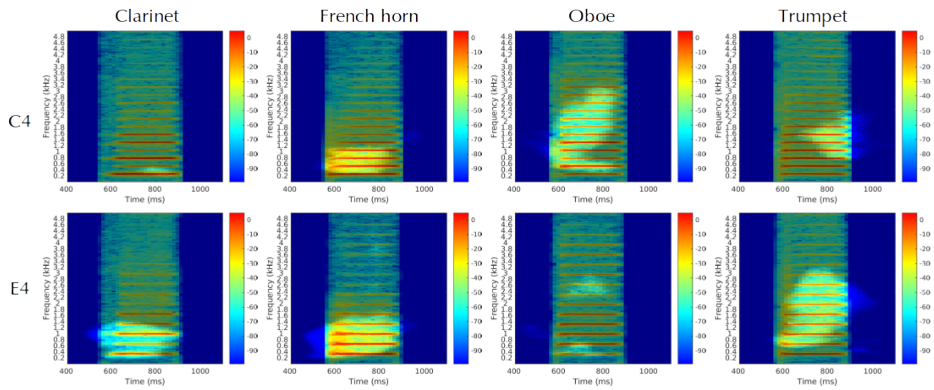


Figure 6: Preliminary results of the experiment with six subjects. Lightened regions exhibit significant correlations between correct identification of the instrument in a mixture and audibility of that region.

Results

Figure 1 shows preliminary results from this experiment. While these results are not conclusive, they suggest that importance follows the harmonics for clarinet, French horn, and oboe, whereas it remains fixed for the trumpet. This makes sense for the woodwinds as the clarinet has a distinctive pattern of alternating strong and weak harmonics and the oboe has a distinctively weak fundamental, both of which relate to the pitch. For the brass instruments, C4 and E4 use the same fingering (no keys pressed), and so represent different excitations of the same tube. They should therefore have fixed resonances with different harmonics, and it appears that the listener utilizes this feature in the case of the trumpet, but perhaps not for the French horn. Further analysis and statistical testing of these results should clarify this issue.

Conclusions

These results show that the bubble noise technique can provide useful insights into the specific time-frequency regions of individual recordings of notes that listeners utilize to identify the instrument performing that note. Thus, it can be used to localize important timbral features of notes with few *a priori* assumptions about their nature in the context of a natural instrument-identification task where non-timbral qualities are normalized out. In order to generalize to an instrument's entire range, several notes from the instrument can be analyzed and conclusions extrapolated.

References

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research*, 41(17), 2261-2271.

Mandel, M. I., Yoho, S. E., & Healy, E. W. (2016). Measuring time-frequency importance functions of speech with bubble noise. *Journal of the Acoustical Society of America*, 140(4), 2542-2553.

Thoret, E., Depalle, P., & McAdams, S. (2016). Perceptually salient spectrotemporal modulations for recognition of sustained musical instruments. *Journal of the Acoustical Society of America*, 140(6), EL478-EL483.

Implicit Timbral Assessments of Recorded and Live Music with Hearing Aids

Amy V. Beeston[1†], Alinka Greasley[1] and Harriet Crook[2]
[1] School of Music, University of Leeds, Leeds, UK
[2] Department of Neurotology, Sheffield Teaching Hospitals NHS Trust, Sheffield, UK
[†] avbees@gmail.com

Aims/goals

This paper reports findings from a project investigating the impact of hearing loss and hearing aid technologies on the experience of listening to music. We take a social-psychological approach and examine a snapshot of almost 1000 responses to an in-depth online survey on music listening. We report here on a subset of the data relating implicitly to timbral aspects of the sound, querying whether respondents experience any imbalance between high and low frequencies while listening to recorded and live music with hearing aids.

Background information

Given the importance of music in society, and the prediction that 20% of the population will have some degree of hearing loss by 2031 (Hill et al., 2015), it is perhaps surprising that so little research has addressed music listening – and hence, timbre perception – with hearing aids. Music presents acoustic signals with numerous key differences from the speech signals that a digital hearing aid is primarily designed to enhance (Chasin & Russo, 2004). As a result, music listening with hearing aids is typically challenging and unsatisfactory (Chasin & Hockley, 2013; Madsen & Moore, 2014). Recent work is also beginning to assess the enormous variation between individuals in terms of both their hearing loss and their use of hearing aid technology in the context of music (Greasley et al., 2018).

Methodology

We examined a snapshot of N=981 responses to an online survey, and analysed responses to a subset of questions about experiencing difficulties with "too much treble" and "too much bass" whilst listening to recorded and live music with hearing aids. The survey questions were designed with data from almost 200 hearing aid users in pilot studies, comprising 179 responses to a waiting-room survey and 22 in-depth interviews (Greasley et al., 2015). The resulting survey investigated hearing level, use of hearing aid technology, music listening in live and recorded settings, and music-related discussions with audiologists. It contained around 70 mainly non-obligatory questions, required around 45 minutes to complete, and was advertised widely via UK and international networks.

Results

Around 95% of survey participants responded to a question asking them to remember and report the category of hearing loss that their audiologist had diagnosed them with (N=928, mild 5.2%, moderate 40.2%, severe 34.4%, profound 13.3%, don't know 7.0%. After removing participants that responded "don't know" to this question, the remaining respondents were grouped according to their self-reported level of hearing loss (N=863, mild 5.6%, moderate 43.2%, severe 37.0%, profound 14.3%).

Participants were asked how often they experienced difficulties with "too much treble" or "too much bass" when listening to recorded (N=690) and live (N=592) music with hearing aids (selecting one of six response alternatives: never; occasionally; sometimes; frequently; all the time; don't know). Around one in eight people responded "don't know" in each of these questions (recorded/treble 13.9%; recorded/bass 13.2%; live/treble 11.8%; live/bass 10.6%), perhaps highlighting the relative obscurity of musical terms such as "treble" and "bass" in everyday language, and revealing the inherent difficulty of discussing sound quality and timbre perception with non-specialist audiences. These participants were discounted from further analysis, and the remaining response categories were transformed into binary variables reflecting whether participants "never" or "at least occasionally" experienced the difficulty in question (i.e., combining responses across occasionally, sometimes, often, and all the time).

When listening to recorded music with hearing aids, approximately two-thirds of people reported difficulties – at least occasionally – of experiencing "too much treble" (N=594, overall 71.9%). There was no significant association, however, between people reporting difficulty with excessive treble experiences and the severity of their reported hearing loss (N=594, X^2 (3) = 0.71, p = .871). Problems experiencing "too much bass" when listening to recorded music with hearing aids, on the other hand, were significantly associated with the self-reported level of hearing loss (N=599, X^2 (3) = 18.83, p < .001). The worse the category of hearing loss, the greater the proportion of participants in the group reported having difficulties with excessive bass experiences (N=599, mild 44.1%, moderate 54.1%, severe 64.8%, profound 76.9%).

A very similar pattern of results was observed for reports of listening to live music with hearing aids. Though the experience of "too much treble" was again prevalent in live settings (N=522, overall 72.8%), no significant association was found between this experience and the level of hearing loss reported (N=522, X^2 (3) = 1.59, p = .661). As with recorded music listening, however, there was a significant association between hearing loss category and the experience of "too much bass" in live settings (N=529, X^2 (3) = 12.44, p = .006), with larger numbers of people again reporting this problem in the more severe hearing loss categories (N=529, mild 54.5%, moderate 62.9%, severe 72.4%, profound 80.6%).

Conclusions

This paper examines experienced imbalance between the amplification of high and low frequencies while listening to music with hearing aids. Patterns in the data revealed that around two-thirds of people experienced difficulties on occasion with too much treble in recorded and live music, regardless of their level of hearing loss. The experience of too much bass, however, varied with hearing loss level: those with more severe hearing losses reported problems more often in both recorded and live settings.

Though not directly revealed in the current results, both individual and contextual factors may contribute to this difference. Firstly, individual factors such as levels of engagement, training, preferences and personality will influence how people perceive music and tolerate discomfort. Further, people may have different understandings of the terms bass and treble, depending on their fluency with musical vocabulary. Secondly, differences exist among participants classified in the same audiological category in terms of both their hearing loss itself and their adoption, fitting and use of hearing aid technologies. It will be important in the future to tease out the relative contribution of these different factors.

Acknowledgments

Funded by the Arts and Humanities Research Council, UK (Grant number: AH/M00368X/1).

References

Chasin, M., & Hockley, N. S. (2014). Some characteristics of amplified music through hearing aids. *Hearing Research*, 308, 2-12.

Chasin, M., & Russo, F. A. (2004). Hearing aids and music. *Trends in Amplification*, 8(2), 35-47.

Greasley, A. E., Crook, H. and Fulford, R. (2015). *Hearing aids for music: Exploring initial findings from questionnaire and interview studies.* In: 9th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Manchester, UK, 17-22 Aug.

Greasley, A. E., Crook, H., Beeston, A. V., Salter, J. and Fulford, R. (2018). *Outcomes of an online survey examining challenges and strategies for music listening with hearing aids*. In: 15th International Conference on Music Perception and Cognition (ICMPC) and 10th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Graz, Austria, 23-28 July.

Hill, S., Holton, K., and Regan, C. (2015). *Action plan on hearing loss*. Official document of the Department of Health and NHS England.

Madsen, S. M., & Moore, B. C. (2014). Music and hearing aids. *Trends in Hearing*, 18, 2331216514558271.

Pose IV: in situ (2016) - for Solo Percussionist on Amplified Piano and Electronics

Anthony Tan[1†], Noam Bierstone[2]
[1] Visual and Performing Arts (Music), University of Colorado, Colorado Springs, Colorado, USA
[2] Schulich School of Music, McGill University, Montréal, Quebec, Canada
[†] atan2@uccs.edu

Aims/goals

This event will present a mixed work for solo percussionist playing amplified acoustic piano with live electronics. The piano timbre and its contrapuntal manipulation, from a timbre perspective, represent the theoretical/compositional approach to this work.

This was a collaborative project between Noam Bierstone (Percussionist) and Anthony Tan (composer). Our goals were two-fold:

1) Explore how a percussionist could approach the piano as a solo instrument.
2) Examine the 'corporeality' of the piano timbre.

Description of event

Duration: 17 min.
Recording: https://soundcloud.com/anthonytan-2/poseiv-in-situ-for-piano-and-electronics-live-recording
Video: https://youtu.be/iWeErHLHjIM

The piano represents a culturally "heavy" musical instrument, which makes it very difficult to divorce its timbre from historical associations. By writing a piano piece that only a percussionist could play, we explored new performative and timbre approaches to this instrument. This work uses the entire physical body of the piano, not just the timbre produced by playing the keys, as a timbral and musical source. Through the use of mallets, we explored the interaction created between the physical contact of a prosthetic object against the piano. The percussionist acts as an exciter, probing and exploring this sound structure *in situ*, or "in position." Here, catalyzed formants and electronic feedback processes take place within situations of calming intensity. Furthermore, the electronics metaphorically serve as a microscope expanding sounds contained within. More specifically, using spectral filtering and pitch shifters, we explored creating pitch orientated contrapuntal lines within the spectral structure of the piano timbre. Further, timbre-orientated music requires longer durations in order for the listener to enter into a "timbre-listening" situation. Therefore, another exploration in this work involves the act of doing something slowly and deliberately for long periods of time, serving as a physical meditation through repetitive movements. This allows both the performer and the listener to enter into a listening situation whereby micro-timbral variations can be perceived. Part of a larger series of works entitled *Poses* whereby each 'pose' examines the relationship between music and an extra musical topic, this work imagined timbre itself as a physical body, contained within the shell of a piano.

StrinGDberg (2001-03)

Robert Normandeau
Faculté de musique, Université de Montréal, Montréal, Québec, Canada
robert.normandeau@umontreal.ca

Aims/goals

StrinGDberg (2001-03) was the first piece of my cycle *Timbre Spatialization*. The cycle is the object of a paper presentation.

Description of event

The piece lasts for 18' and illustrates the concept of *Timbre Spatialization*. The piece is made for an immersive environment, which usually comprises between 16 and 64 speakers. For the current *2018 Timbre* conference, the 5.1 version will be played.

Acknowledgments

Thanks to Jean-François Denis, owner of the empreintes DIGITALes label.

References

Normandeau, R. 2005. StrinGDberg (2001-03). On Puzzles. Montréal: Empreintes Digitales, IMED 7505. Reprint IMED 1775.

Timbre is a Technomorphic Thing: A Comparative Analysis of three Case Studies

Eric Maestri
Music Department, University Paris 8 Vincennes-Saint Denis, MUSIDANSE/CICM, Saint-Denis, France
eric.maestri@gmail.com

Aims/goals

Focusing on timbre as a *technomorphic thing*, this paper investigates its use by confronting three musical works: Tristan Murail's *Les nuages de Magellan* (1973), for e-guitar, two ondes Martenot and percussions, Pierre Alexandre Tremblay's *asinglewordisnotenough III* (2015), for cello and electronics, and a *live coding* improvisation for cello and computer by Thor Magnuson and Miguel Mira (2015) (Maestri & Mori, 2017). Adopting Caporael's definition, *technomorphism* is considered as the attribution of technological characteristics to humans (Caporael, 1986). On this basis, researches have proven that computer users professionals—like engineers or programmers—technomorphize their behaviors and their relationships with human beings (Lum, Shelstad, Harris, & White, 2014). The paper argues that the utilization of electronic tools technomorphizes the sound and that the timbre is the means of its implementation. Thus, instead of being a neutral concept, towards analysis, synthesis and transformations, it biases sound accustoming listeners, musicians and musicologists.

Background information

Even if timbre was always part of the compositional practice, it gained importance in 19th century thanks to the development of orchestration techniques, conceived as the melting of the instruments' timbre, (Berlioz, 1844) and increased its centrality in the last century's music, when it became a parameter of musical composition (Solomos, 2013). Thanks to computer assisted composition and mixing techniques, from "art music" to EDM, sound occupies a crucial role in musical composition. It marks the contemporary compositional means and determines the appearance of specific musical morphologies (Roads, 2015). Sound is technologically conceived and manipulated towards its parameters as timbre. Timbre grounds techniques and compositional procedures on new cognitive basis (McAdams, 1989); in music it is a specific morphophoric element (Duchez, 1989); it implies musical forms that enhance human-machine interaction in music as well (Maestri, 2017). Through the evolution of this notion, sound as timbre has become an operational category based on the manipulation of its parameters, allowing the implementation of techniques such as computer assisted orchestration (Crestel & Esling, 2017) and musical querying (Leman, 2007). Timbre is a musical thing that results from a massive technological exploitation which transforms the previous empirical use (i.e. the instrumental melting in Berlioz) in a powerful tool for musical making. For this reason, it is worth to approach this question under the perspective of *technomorphism*.

Methodology

The methodology aims at showing technomorphic sound morphologies and musical forms in concrete cases. This analysis will follow a *workflow* in three parts: 1) spectromorphological aural segmentation (functional structures, textural movement and spectral typology) (Smalley, 1997), 2) spectral and morphological analysis and 3) comparison of the morphologies considered as the distribution of the sound elements of the musical events. Three works are chosen because of their substantial differences in instrumentation and use of technologies in sound generation and transformations: 1) Murail's piece uses an e-guitar, Ondes Martenot and percussions, 2) Tremblay's piece is for cello and electronics based on DSP processing, 3) Magnusson and Mira's improvisation is a mixed piece for cello and computer generating the electronics. These pieces express three kinds of sound compositions: 1) orchestration of and with electric instruments, 2) transformation and mixing of instrumental and electronic sounds with digital tools, 3) generation of sustained sound morphologies that accompany and solicit the improvisation of the cello.

Results

This paper will show the following aspects: 1) the definition of sound morphologies as *technomorphic*, questioning the concept of timbre as biased by technology; 2) the use of sound as timbre in three divergent musical works; 3) the pattern of their morphological common and/or different features from the point of view of *technomorphism*: is there a *technomorphic* sound morphology?

Conclusions

In conclusion, the author will argue that the analyzed musical morphologies are "techno-morphologies." The paper will discuss timbre as a notion biased by technology evoking the difference in sound morphologies between electronic and non-electronic musical works. The author will invite to consider the results of the paper as a starting point of a larger research project that aims at considering timbre from the perspective offered by the notion of *technomorphism*. This notion, that proposes a cultural perspective about electronic musical practices, is currently undervalued and has a strong heuristic power that deserves to be deployed in music.

References

Berlioz, H. (1844). *Grand traité d'instrumentation et d'orchestration modernes*. Paris: Schonenberger.

Caporael, L. R. (1986). Anthropomorphism and Mechanomorphism: Two Faces of the Human Machine. *Computer in Human Behavior, 2*, 215–234.

Crestel, L., & Esling, P. (2017). Live Orchestral Piano, a system for real-time orchestral music generation. In *Proceedings of the 14th Sound and Music Computing Conference 2017*. Espoo, Finland.

Duchez, M.-E. (1989). An historical and epistemological approach to the musical notion of "form-bearing" element. *Contemporary Music Review*, *4*(1), 199–212.

Leman, M. (2007). *Embodied Music Cognition and Mediation Technology*. Cambridge: MIT Press.

Lum, H. C., Shelstad, W. J., Harris, M. H., & White, M. M. (2014). Human or Superhuman: Individual Differences in the Perception of Technomorphism. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 1424–1428).

Maestri, E. (2017). A Typo-Morphological Approach to Human-Machine Interaction Analysis in Music. *Organised Sound*, *22*(3), 315–323.

Maestri, E., & Mori, G. (2017). Live Coding: Four case studies' sound morphologies. In *GATM 2017*. Rimini, Italy. Retrieved from https://halshs.archives-ouvertes.fr/halshs-01644118

McAdams, S. (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, *4*(1).

Roads, C. (2015). *Composing Electronic Music. A New Aesthetic*. Oxford New York: Oxford University Press.

Smalley, D. (1997). Spectromorphology: explaining sound-shapes. *Organised Sound*, *2*(2), 107–126.

Solomos, M. (2013). *De la musique au son. L'émergence du son dans la musique des XXème-XXIème siècles*. Rennes: Presses universitaires de Rennes.

The communication of timbral intentions between pianists and listeners and its dependence on audio-visual listening conditions

Shen Li[†], Renee Timmers
Department of Music, University of Sheffield, Sheffield, United Kingdom
[†] sli37@sheffield.ac.uk

Aims/goals

A listening experiment was conducted to examine the accuracy of communication of timbral intentions to listeners and its dependence on visual and aural component of musical performance.

Background information

Studies on the audio-visual presentation of musical performance have investigated its impact on perceived expressiveness (Davidson, 2005) and musical appreciation/evaluation (Platz & Kopiez, 2012). Little is known about whether a precise timbral intention by a pianist can be communicated to listeners. An interview study (Li & Timmers, 2017) on the conceptualization of piano timbre revealed pianists' extensive utilization of timbral intentions in piano performance; the findings also suggested that the timbre concept of a pianist is enriched by embodied experience such as bodily preparations, indicating the relevance of visual cues.

Methodology

Thirty musical excerpts that were played by three pianists using ten timbres (bright/dark, heavy/light, relax/tense, round/sharp, velvety/dry) were recorded both aurally and visually. Twenty-one undergraduate music students took part in the listening experiment using a within-subject design, and were presented all the musical performances either in audio only, visual only, or both audio and visual. They were required to rate to what extent each ten timbres were expressed in each music excerpt on 1-9 Likert scales. The dependent variable was the ratings of each timbre evaluation; while independent variables were the factor of performer/piece and the mode of audio-visual presentation.

Results

We first calculated the percentage of correct of listeners' responses for each timbre by re-coding data with either 1 (correct) or 0 (incorrect). The response was correct when the rating of target timbre is higher than ratings of the rest of nine timbres for the same music excerpt; the response could also be coded as correct when the rating of target timbre is relatively higher for the target instruction excerpt compared to the other nine instructions. The preliminary results showed that all the ten timbres were communicated with an accuracy above chance level, although the average percentage correct was considerably higher for some timbres (sharp, relaxed, light) than for others. Dark and round were the most difficult ones to communicate.

Three-way repeated measures ANOVA was conducted to examine the effect of performer / piece, listening mode and instruction on perceived timbre. First, significant main effects of instruction were found for almost all timbres except for darkness. It means, excerpt for the darkness of timbre, the three performers were able to make effective contrast between instructions in their performances and communicate successfully to the listeners. Secondly, two-way interaction between instruction and AV condition revealed that particular mode of audio-visual presentation affects the communication of several timbres. For instance, relaxation and tension were mainly communicated by visual information while the velvety timbre is communicated relying on sound. Finally, three-way interaction effects suggest that the three pianists showed different capacity to communicate timbral intentions to listeners and relied differently on audio-visual presentation modes. For example, heaviness was communicated better by piece/performer 1 aurally while better by piece/performer 3 with visual information; sharpness was communicated more successfully by piece/pianist 1 presented with visual only while better by piece/pianist 3 aurally.

Conclusions

This study demonstrated that it is possible for pianists to reliably communicate abstract timbral intentions to (musically trained) listeners. It also showed that several but not all timbres can be communicated effectively in the audio only condition; both visual and auditory information contribute to the communication process. This research suggests that performers can potentially learn to communicate the music using both the sound AND the body.

Acknowledgments

This research is funded by Chinese Scholarship Council.

References

Davidson, J. W. (2005). Bodily communication in musical performance, in D. Miell, R. Macdonald, & D. Hargreaves (Eds.), *Musical communication*, pp. 215-237.

Platz, F., & Kopiez, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception*, 30(1), 71-83.

Li, S., & Timmers, R. (2017, 31 July – 4 August). Exploring pianists' embodied concepts of piano timbre: An interview study. In V. Dyck (Ed.) *Proceedings of the 25th Anniversary Conference of the European Society for the Cognitive Sciences of Music,* pp 112-117, Ghent, Belgium.

The Mikakaji: Timbre in Zambian Luvale Percussion

Jason Winikoff
Department of Music, Tufts University, Medford, MA, USA
jason.winikoff@tufts.edu

Aims/goals

By examining performance technique of the Zambian *mikakaji*, I aim to provide a case study of how Luvale drummers interact with timbre. This analysis serves as the ethnographic justification for my close consideration of technique, rhythm, and the instrument itself. Beyond my discussion of the timbral qualities of a part, I explain how rhythmic variation in a common *mikakaji* break results in changes in density. This yields a phenomenon I term a "timbral shift": when the timbral profile of the ensemble's gestalt changes. Ultimately, I argue that to move towards an ethnographically informed understanding of Zambian Luvale percussion music, we must consider sound quality as vital as rhythmic competency.

Background information

The Luvale are a Bantu group of people primarily located in present-day northwestern Zambia and eastern Angola. To the academic community (Jordan 2000, 2006; Kubik 1981; Wele 1993) their rich culture seems to revolve around initiation camps (*Mukanda* and *Wali*) and the *makishi* (ancestral spirits manifesting in the living world as masked dancers). Occasionally (Tsukada 1988) research is dedicated to the songs of these ceremonial functions. The drumming tradition that accompanies these cultural staples is almost always overlooked, though. In this paper I focus on one instrument of the Luvale percussion ensemble: the *mikakaji*. The *mikakaji* are two sticks struck against the body of a drum or a metallic idiophone affixed to the lead drum.

Methodology

This research is based off of several fieldwork trips to various parts of Zambia. Research was primarily conducted in Lusaka and Zambezi village. Participant-observation and a quest towards bi-musicality defined my modes of ethnomusicological inquiry. I worked closely with three culture groups – conducting interviews, taking lessons from, performing alongside, and befriending members of Lenga Navo, Tulizangenu (Chota), and Likumbi lya Mize Chibolya. In this paper I valorize the ear, mind, and body's ability to organize and draw meaning from sound; consequentially, human subjective response to music often constitutes the "data" for this study.

Results

The emic explanation for sticking choices revolves around the notion that each hand, stick, part of the stick, and surface struck will yield a different timbre. The repetition of the fastest pulses (Koetting 1986) on the mikakaji is not timbrally consistent; this is part of the basic musical idea of this instrument. This thus problematizes the notion of repetition. Furthermore, musical excitement is often created through changes in the timbral profile that result from alterations to both simultaneous and successive density.

Conclusions

Luvale musicians are not just sensitive to timbre, they actively engage with it in performance. Non-timbral musical elements (such as rhythm, technique, density, and the physical instrument) affect timbre. Rhythmically defined and manipulated "parts" (Locke 1998) only offer a partial explanation of a drum's basic musical idea and effect. It may often make more sense to conceive of percussion parts in terms of the variation of the timbres of the pitched and "pitchless" (Fales 1995) sounds produced. In this way I conceive of rhythm as a means of controlling, presenting, and varying timbre. I conclude that rhythm can be an expressive means of interacting with the fundamental aspect: timbre. My conclusions may help explain choices in technique, instrument construction, and ensemble layering.

Acknowledgments

I first and foremost must thank the Luvale people for granting me a window into their beautiful culture. The impact of their friendship, patience, knowledge, and talent is invaluable. I extend a special thank you to my primary teachers: Kapalu Lizambo, William Vunda, and Samuel Samutala. My friends in Nomakanjani have also had a profound influence on my love and understanding of percussion, dance, and song in Zambia. I next must extend my gratitude towards my mentor and thesis advisor, David Locke. This work would not have been possible without funding from the Granoff Special Projects and Research Grant as well as the Graduate Student Research Competition award. I thank Tufts University for the stimulating academic environment and my family and friends for the love and support.

References

Dreier, James (2001). . Kiko's Samba. *Percussive Notes* 39(1): 18-19.

Euba, Akin (1990). *Yoruba Drumming: The Dundun Tradition*. Bareuth: Bayreuth University.

Fales, Cornelia (1995). Acoustic Intuition of Complex Auditory Phenomena By "Whispered Inanga" Musicians of Burundi. *The World of Music* 37(3): 3-35.

Heidmann, Kate (2016). A System for Describing Vocal Timbre in Popular Song. *Music Theory Online* 22(1).

Koetting, James, & Roderic Knight (1986). What Do We Know About African Rhythm? *Ethnomusicology* 30(1): 58-63.

Lavengood, Megan (2017). *A New Approach to the Analysis of Timbre*. The City University of New York.

Locke, David (1998). *Drum Gahu: An Introduction to African Rhythm*. Tempe, AZ: White Cliffs Media.

Parillo, Michael (2008). Zakir Hussain Global Drummer. *Modern Drummer* 32(1): 72-76, 78, 80, 82, 84.

Serwadda, Moses, & Hewitt Pantaleoni (1968). A Possible Notation for African Drumming. *African Music* 4(2): 47-52.

Wele, Patrick (1993). *Likumbi Lya Mize and Other Luvale Traditional Ceremonies*. Lusaka, Zambia: Zambian Educational Publishing House.

Whitman, David Robert (2011). Two Lecture Recitals of Percussion Music: "Notation for Stroke-Type Analyses" and "Tabla and Indian Musical Concepts in Modern Drum Set Performance". University of Southern Mississippi.

Documentation and Acoustic Analysis of the Kichwa Instruments of Ecuador

Hazel Burns[†], Gabriel Ferreyra
Colegio de Música, Universidad San Francisco de Quito, Quito, Pichincha, Ecuador
[†] hburns@usfq.edu.ec

Aims/goals

The folkloric music of Ecuador's Andes region shares a similar cultural heritage with that of Peru and Bolivia. Consequently, much of the information that exists about traditional musical instruments from this area emphasizes a primarily Peruvian, and to a lesser extent, Bolivian point of view, with little documentation existing on instruments exclusively used in Ecuadorian music. Presently, we are undertaking a project to document and acoustically analyze the folkloric instruments of Ecuador's Kichwa people, with specific emphasis given to instruments endemic to Ecuador. Our acoustic analysis includes measurements of each instrument's radiation patterns and timbral characteristics, as well as suggested recording approaches, and is based on similar studies done with classical orchestral instruments. The goal is to give representation to the Kichwa folkloric instruments of Ecuador through the compilation of a concise and comprehensive guide that includes information on each instrument's timbral characteristics and radiation properties, as well as musical function and desired cultural aesthetics. This will aid sound engineers and ethnomusicologists in accurately recording them.

Background information

Mainland Ecuador consists of three principal regions: the Andes mountain range (Sierra), the Coast, and the Amazon jungle (Oriente). Each geographical region has its own distinct indigenous cultures and musical traditions. This study focuses exclusively on the traditional instruments of the Kichwa people of the Sierra region, who constitute Ecuador's largest indigenous population and are an ethnic and linguistic division of the broader Quechua culture found throughout Ecuador, Peru, Bolivia, Colombia, and Argentina.

The traditional music of Ecuador was first documented in 1925 by French ethnomusicologists Raoul and Marguerite d'Harcourt, who published fifty-four transcriptions of Ecuadorian music in their book, *La musique des Incas et ses survivances*. Other notable publications of Ecuadorian music have been Luis Segundo Moreno's *La musica en el Ecuador*, Gilbert Chase's *A Guide to the Music of Latin America*, and Paulo de Carvalho Neto's *Diccionario ecuatoriano de folklore*. However, these texts focus largely on the preservation and analysis of indigenous melodies and rhythms and contain little information about the instruments that were used. Robert Stevenson's book *Music in Aztec & Inca Territory* mentions this lack of archaeological documentation, stating that many studies choose to focus on the instruments of Peru and overlook the nearby and culturally similar countries of Bolivia and Ecuador.

Kichwa folkloric music and its traditional instruments are an important part of the national cultural identity of modern Ecuador and heavily influence Ecuador's popular music scene. This is evident in the success of groups such as traditional folkloric ensemble Ñanda Mañachi, who was awarded "best classical music album of the year" by the *Académie Charles-Cros* in France; emerging Andean hip-hop group Los Nin, who incorporate indigenous melodies and instruments and rap in both Spanish and Kichwa; and DJ Nicola Cruz, who fuses traditional Ecuadorian music and instruments with electronic sounds. Nonetheless, there is little documentation and currently no standardized approach to recording many of these instruments.

Basing our analysis of the Kichwa instruments' timbral properties and radiation patterns on Jürgen Meyers' *Acoustics and the Performance of Music* and Michael Dickreiter's *Tonmeister Technology*, the aim of our research is to fill this gap and add a further layer of understanding to these instruments from an acoustical point of view. In addition to serving as a reference guide for ethnomusicologists and modern recording engineers, this kind of formal study, which includes the contribution of prominent Kichwa musicians, will be the first to specifically represent these instruments within an Ecuadorian cultural setting.

Methodology

In order to accurately represent traditional Kichwa instruments in recorded form, it is important they be analyzed from both a musical and sonic point of view and that the knowledge, aesthetics, and opinions of Kichwa musicians and luthiers be given proper representation. For this purpose, information is being gathered using the following methods:

1. Investigation of previous ethnomusicological research pertaining to Andean folkloric music, with specific focus on the instruments' construction, history, and musical function.
2. Interviews with prominent Kichwa musicians and luthiers to gain insight into each instrument's history, musical function, and acoustic design.
3. Measurements and recordings of each instrument's radiation patterns and sonic properties using a 25-microphone array in a controlled environment and spectral analysis software.
4. Experimentation with microphone position and technique in collaboration with Kichwa musicians to create a recording reference guide.

Results

The end result of this research project is to represent the folkloric instruments of the Kichwa people of Ecuador through a formal analysis and the compilation of a concise reference guide including a library of recorded samples. This guide will serve as a basis for understanding each instrument's musical function, radiation properties, and desired timbral characteristics, and will aid ethnomusicologists and recording engineers in accurately capturing these instruments in recorded form.

Conclusions

Due to the popularity and prevalence of traditional Andean instruments in Ecuadorian music today, and the underrepresentation of these instruments from a specifically Ecuadorian perspective, there is a need for their formal analysis. This project will add to the canon of ethnomusicological research of Ecuador and give representation to the knowledge of the Kichwa people. Specific emphasis on each instrument's sonic characteristics, musical function, and desired timbre in recorded representation will further aid sound engineers and ethnomusicologists in understanding the instruments' desired properties and accurately recording them. Our presentation will outline the methodologies used to analyze and document these instruments, and discuss the limitations faced in undertaking such research in Ecuador.

References

Chase, G. (1962). *A Guide to the Music of Latin America* (2nd ed.). Buffalo, NY: Pan American Union and the Library of Congress.

Dickreiter, M. (1989). *Tonmeister Technology: Recording Environments, Sound Sources and Microphone Techniques*. (S. F. Temmer, Trans.). New York, NY: Temmer Enterprises, Inc. (Original work published 1984)

List, G. (1966). Ethnomusicology in Ecuador. *Ethnomusicology, 10*(1), 83-86. Chicago, IL: University of Chicago Press.

Meyer, J. (2010). *Acoustics and the Performance of Music: Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instrument Makers* (5th ed.). New York, NY: Springer.

Olsen, D. A., & Sheehy, D. E. (2014). *The Garland Handbook of Latin American Music* (2nd ed.). New York, NY: Routledge.

Stevenson, R. (1968). *Music in Aztec & Inca Territory*. Berkeley and Los Angeles, CA: University of California Press.

Stimulus Timbre Modulates Pupil Dilation During an Absolute-Pitch Test

Patricia Vanzella[1,2†], Francisco J. Fraga[1,4], Guilherme Alves Delmolin de Oliveira[3] , João Ricardo Sato[1,2]
[1] Interdisciplinary Unit for Applied Neuroscience, Universidade Federal do ABC, São Bernardo do Campo, Brazil
[2] Department of Music, Universidade de Brasilia, Brasília, Brazil
[3] Center for Mathematics, Computation and Cognition, Universidade Federal do ABC, S. Bernardo do Campo, Brazil
[4] Center for Engineering, Modeling and Applied Social Sciences, Universidade Federal do ABC, Sto André, Brazil
[†] p.vanzella@ufabc.edu.br

Aims/goals

We investigate whether pitches in different timbres influence accuracy, reaction times, and pupillary responses of musicians while performing an absolute pitch test. Our results help to understand how vocal and instrumental tones are processed and may elucidate how musicians with varying degrees of absolute pitch resort to different strategies to identify musical stimuli in different timbres.

Background information

Absolute Pitch (AP) is the ability to effortlessly identify or sing pitches without reference (Parncutt & Levitin, 2001; Ward, 1999; Takeuchi & Hulse, 1993). It has been shown, however, that pitches in certain timbres are more difficult for AP possessors to identify, e.g., piano tones are easier than sung tones (Vanzella & Schellenberg, 2010). Recent studies have revealed that pupillary responses are a sensitive and reliable indicator of the extent of central nervous system processing allocated to a task. Increases in pupil diameter correspond to increases in the amount of information processing (Granholm & Steinhauer, 2004). If AP possessors find vocal tones more difficult to label, we hypothesized that the cognitive load would be higher when identifying them and hence there would be greater pupil dilation when compared to the identification of piano tones. Since AP possessors usually label tones without effort, we also hypothesized that the higher the accuracy in the AP test, the smaller would be the reaction times and pupil dilation.

Methodology

We studied pitch identification performance on piano and vocal timbres by analyzing hit rates, reaction times and pupillary responses of 18 undergraduate and graduate music students with varying degrees of AP. We used a pitch discrimination task while recording their pupillary responses with an eye tracker (Mobile Eye 5, 60 Hz, Monocular (right) – ASL). The task was designed in a block paradigm. Each block consisted of six 1-second-duration tones, with 5-second interstimulus interval. There were twelve blocks in all: six for the vocal timbre and six for the piano timbre. The blocks were presented in a pseudo-randomized order, and each block was separated by a 30-second interval of silence. The tones presented included all pitches of the chromatic scale between A3 and Gsharp5. The reaction time was measured as the elapsed time between the onset of the played tone and the onset of the participant's vocal response using a Voice Activity Detection algorithm. Hit rates were measured by auditory recognition of each trial response. Pupillary responses were matched with each participant's reaction time in the task.

Results

As expected, behavioral results revealed significant longer reaction times for vocal tones compared to piano tones. There was no difference in accuracy for piano and vocal tones. However, there was a significant negative correlation between reaction time and accuracy for both piano and vocal tones. On the psychophysiological level, we observed pupil dilation in the whole sample of participants during the pitch labeling task. Moreover, as hypothesized, pupillary responses for instrumental and vocal timbres were distinct, with larger pupil dilation for the latter. We found no correlation between accuracy and pupil dilation and no correlation between reaction time and pupil dilation.

Conclusions

We investigated the amount of resource allocation needed during an AP test by analyzing changes in pupil diameter as well as behavioral data (in terms of accuracy and reaction times). We observed larger pupil dilation for vocal tones. AP possessors identified vocal tones just as accurately as piano tones, albeit with slower reaction times. This may indicate that either these tones required more effort to be labeled or that they were more salient than piano tones. Such salience could be explained by the fact that the sound of human voice is probably the most familiar sound for our species. Vocal sounds are usually loaded with various levels of information (linguistic and paralinguistic, including identity and affective cues) that simultaneously activate distinct neural pathways (Belin, Zatorre, & Ahad, 2002). Although evidence in the literature is mixed, pupillary responses have also been associated to emotional processing (Mudd, Conway, & Schindler, 1990; Steinhauer, Boller, Zubin, & Pearlman, 1983; Libby, Lacey, & Lacey, 1973). Since vocal sounds are frequently loaded with emotional information, it might be the case that the larger pupil dilations for vocal stimuli observed in our study would be better explained in terms of emotional arousal. Further investigation is needed to evaluate this possibility.

Acknowledgements: We thank Jackson Cionek from BrainSupport-Brazil for the technological support.

References

Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Cognitive Brain Research*, *13*(1), 17-26.

Granholm, E., & Steinhauer, S.R. (2004). Introduction: Pupillometric measures of cognitive and emotional processing. *International Journal of Psychophysiology*, 52, 1–6.

Libby, W.L. Jr., Lacey, B.C., & Lacey, J.I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology*, 10, 270–294.

Mudd, S., Conway, C.G., & Schindler, D.E. (1990). The eye as music critic: Pupil response and verbal preferences. *Studia Psychologica*, 32, 23–30.

Parncutt, R., Levitin, D.J. (2001) Absolute pitch. In S. Sadie (Ed.), *New Grove dictionary of music and musicians* (pp. 37–39). New York: St. Martins Press.

Steinhauer, S.R., Boller, F., Zubin, J., & Pearlman, S. (1983). Pupillary dilation to emotional visual stimuli revisited. *Psychophysiology*, 20, 472.

Takeuchi, A.H., & Hulse, S.H. (1993). Absolute pitch. *Psychol Bull* 113, 345–361.

Vanzella, P., & Schellenberg, E.G. (2010). Absolute pitch: Effects of timbre on note-naming ability, *PLoS ONE, 5*, e15449. doi:10.1371/ journal.pone.0015449.

Ward, W.D. (1999). Absolute pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 265–298). San Diego: Academic Press.

AUTHOR INDEX

Abel 85
Alluri 27
Ando, M. 23, 25
Ando, T. 25
Antoine 111
Aspinwall 73
Baril 13
Barriere 33
Beauchamp 146
Beavers 150
Beeston 174
Berger, J. 85
Berger, T. 85
Berndt 55
Bierstone 176
Bitton 109, 158
Blake 48
Bonin 170
Bouliane 13
Bratus 48
Buffington-Anderson 38
Burns 184
Caetano 144
Callery 85
Caramiaux 81
Carat 99
Carsault 97
Celerier 111
Cella 15, 93
Chemla-Romeu-Santos 15, 97, 109, 158
Chon 53, 172
Cordero 17, 103
Crestel 15, 95
Crook 174
Czedik-Eysenberg 50, 67
Daubresse 103
Decroupet 46
Depalle 81
Desainte-Catherine 111
Deschenes 75
Di Gasbarro 44
Dolan 22
Engebretsen 156
Esling 15, 93, 95, 97, 109, 158
Fales 16, 31
Feron 99
Ferreyra 184
Fink 16
Flore 40
Ghisi 95
Giraldo 57
Greasley 174
Guastavino 99
Guidarini 162
Gur 107
Hadjakos 55
Hagan 127
Hasegawa 128
Haus 158
Henderson 119
Henry 134
Herold 101
Hoitenga 33
Honisch 63
Huron 123
Kahrs 154
Kelly 73
Kermit-Canfield 85
Kim 53
Knauf 67
Levy 17
Latour 16
Lavengood 115
Le Bel 113
Lee Newsome 160
Lembke 83
Li 180
Luce 146
Maestri 178
Mandel 172

Manzolli 152
Mativetsky 29
McAdams 20, 81, 134
Meier 95
Mendes 125
Michel 121
Miranda 111
Neri 73
Noble 134
Nord 132
Normandeau 71, 177
O'Callaghan 87
Oehler 50
Ogg 69
Pastiadis 136
Pecktor de Oliveira 164
Perevedentseva 166
Pottier 65
Prang 15
Provenzano 16
Puckette 127
Queiroz 168
Quiroz 73
Radford 17, 91
Ramirez 57
Rea 17
Reuter 50, 67
Reymore 123
Richan 59
Roque 125
Rossetti 152
Rouat 59
Saitis 79
Schedel 63
Schoonhoven 117
Sdraulig 148
Sheinbaum 105
Siddiq 50, 67
Siedenburg 77, 79
Simurra 168
Slevc 69
Soden 103
Spencer 89
Tan 176
Theberge 16, 140
Thoresen 17, 130
Thoret 81, 134
Timmers 180
Touizrar 61
van Elferen 42
Vanzella 186
Waddel 57
Waksman 16
Wallmark 16, 142
Waloschek 55
Whitesell 6
Williamon 57
Winikoff 182
Zacharakis 136
Zagorski-Thomas 16, 138
Zeller 52