Estonian Academy of Music and Theatre

Gerhard Bruno Erich Lock

Methodological Contributions to a Cognitive Analysis of Perceived Structural Musical Tension in Contemporary Post-Tonal Orchestral Music

Article-based doctoral dissertation in music theory and music psychology

Supervisor: Prof. Dr. Kerri Kotta

Consultant: Prof. Dr. Mauri Kaipainen (University of Helsinki)

Tallinn 2025

List of included publications

I – Lock 2010a (classificator 1.2; extended summary see 5.1, full text see A.1)
Lock, Gerhard 2010a. Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses „Oxymoron“ [On factors influencing musical tension and its perception in Erkki-Sven Tüür’s piece „Oxymoron“]. – *Res musica* 2: 63−74.
https://resmusica.ee/wp-content/uploads/2017/04/rm2_2010_63-74_Lock.pdf,
https://www.researchgate.net/publication/348008100

Ia – Lock & Kotta 2012 (classificator 3.4; extended summary see 5.2, full text see A.2)
Lock, Gerhard; Kotta, Kerri 2012. Musical tension as a response to musical form. – *Music Cognition: 12th International Conference on Music Perception and Cognition (ICMPC) / 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM).* 23.–28.07.2012. Thessaloniki: Aristotle University of Thessaloniki, 612−617. http://icmpc-escom2012.web.auth.gr/files/papers/612_Proc.pdf,
https://www.researchgate.net/publication/348007176

II – Lock 2014 (classificator 3.1; extended summary see 5.3, full text see A.3)
Lock, Gerhard 2014. Visualisierende Analyse [Visualizing analysis]. Erkki-Sven Tüür: *Oxymoron*. – *Musikalische Analyse. Begriffe, Geschichten, Methoden.* Hrsg. Felix Diergarten. Laaber: Laaber, 287−319. (Grundlagen der Musik 8, Hrsg. Felix Diergarten & Manuel Gervink).
https://www.researchgate.net/publication/284900168

III – Lock 2017a (classificator 1.2; extended summary see 5.4, full text see A.4)
Lock, Gerhard 2017a. Models and scientific modeling in music context. – *Problems in Music Pedagogy* 16 (2): 25−45.
http://pmp.du.lv/wp-content/uploads/2016/08/PMP_2017_Vol-16-2.pdf,
https://www.researchgate.net/publication/331132663

IV – Lock 2020 (classificator 1.1; extended summary see 5.5, full text see A.5)
Lock, Gerhard 2020. Salienz, Narrativität und die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung von post-tonaler Orchestermusik [Salience, Narrativity and the Role of Musical Parameters in the Analysis of Musical Tension in Post-Tonal Orchestral Music]. – *ZGMTH: Zeitschrift der Gesellschaft für Musiktheorie* 17 (2): 311–349.
https://doi.org/10.31751/1074, https://www.researchgate.net/publication/348003849

V – Lock 2024 (classificator 3.4; extended summary see 5.6, full text see A.6)
Lock, Gerhard 2024. Musical tension: Methodological dilemmas. – *Actes des Journées d'Informatique Musicale [Proceedings of the Musical Computer Days] JIM 2024.* 06.–08.05.2024. Marseille: PRISM (Perception, Representations, Image, Sound, Music) Laboratory, CNRS (The National Centre for Scientific Research), Aix-Marseille University, 303–312. https://cloud.prism.cnrs.fr/index.php/s/8aX4YaDXArst2xQ,
https://www.researchgate.net/publication/380726031

Methodological Contributions to a Cognitive Analysis of Perceived Structural Musical Tension in Contemporary Post-Tonal Orchestral Music

Metoodikat kognitiivselt analüüsimaks nüüdses posttonaalses orkestrimuusikas tajutud strukturaalset muusikalist pinget

Abstract

The basic goal of this dissertation is to intertwine music analysis and music psychology using a cognitive approach. The general aim is to advance analytical methods in order to enhance the understanding of the general object – complex musical structures and contemporary post-tonal orchestral music (CPTOM) – while observing/analyzing the special object of this dissertation – musical tension as a compound phenomenon – that is universally comprehensible but complicated to research systematically. In this dissertation music is understood as environment (publications III, IV). Furthermore, principles of modeling and analogies (publication III) as well as visualization and representation of music (publication II) are presented. The main objective is to detect, describe, make sense of and comprehend structural aspects that purportedly *trigger* the experience of musical tension as a temporal dynamic wave-like (real-world) phenomenon (TDWP) during *attentive listening* to CPTOM. The means to achieve this objective are defining and empirically analyzing musical tension in complex, cross-style and sound-centered post-tonal orchestral and symphonic music by the Estonian contemporary composer Erkki-Sven Tüür (b 1959): 4th Symphony/Percussion Concerto *Magma* (2002) and *Oxymoron* (2003) for large ensemble. The Introductory chapter presents in section 1 the basic goal and general object, the general aim and special object, the reasons to arrive at a cognitive approach and basic methodological aspects. Section 2 discusses music-theoretical and music-psychological thinking intertwined, introduces important methodological-philosophical matters (e.g. Kantian *analytic idealism*), temporal dynamic cognitive processes and *microgenesis*, psychophysical measurement principles, cognitive dynamic listening models, and *protonarratives* as a "pre-definition" of musical tension. In section 3 key concepts are introduced: music and tension, salience and attention, salience in complex sounds and contemporary music, musical parameters and narrativity. Musical tension has been defined music-theoretically as (perceived) structural musical tension (publications I, Ia, IV), later it is identified as Cognitive Musical Tension (CMT) (publication V) and a tension, cognition and narrativity joining cross-domain definition of ‘Perceived structural musical tension’ (PSMT) is presented (publication V, see 3.5). For this dissertation empirical experiments have been conducted with slider-controllers (publications I, Ia, II [N=7, N=6]; IV [N=26]) and the especially developed COSM: Cognitive Octagonal Slice Model (publication IV [N=14]) enabling to detect musical events as Impulses/“moments of change” and their ‘content’, musical parameters, via salience. The research methodology and design in section 4 uses triangulation and presents answers to the guiding questions provided by the especially developed Twelve Strategic Steps for modeling/analyzing of scientific models (TSSM, publication III) aiming at supporting the development of COSM. Section 5 presents extended summaries of and relations between the included publications. Section 6 arrives at the conclusions and implications of this dissertation. The appendices present the re-prints of the included publications (A.1–6); the technical steps to apply the research design of this dissertation (B); methodological questions that appeared after publications I and Ia (C.1–2); and some overall results of the COSM experiment. Data analysis has been conducted with the DBSCAN: Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise via cluster-finding and time series patterns. Specific results of the main study (publication IV) show that a hypothesized higher salience of “secondary” (SP) over “primary” musical parameters (PP) could be confirmed only partly in Tüür’s 4th Symphony/Percussion Concerto *Magma*. Overall results (see Appendix D) show that the musical parameters “dynamics” (SP), “rhythm” (PP) and “instrumentation/timbre” (SP) as so-called All salience level parameters (AL) are perceived as more salient than the remaining, so-called Basic salience level parameters (BL). This result is in line with general observations/expectations about musical tension, musical parameters as well as Tüür’s compositional tools and his postulated “integrated cognition” goal. The novelties and achievements of this dissertation are mainly methodological: the experiments (including COSM) can be understood as a “psychophysical measurement” procedure detecting “just noticeable difference” (JND) in Impulses/“moments of change” and musical parameters via salience. COSM represents an external, intersubjectively shared spatial standard-object to gain a meaningful audio-visual-salience based Kantian objective understanding of musical tension in CPTOM.

Keywords: perception, cognition, experience, attentive listening, psychophysical measurement, salience, perceived structural musical tension, post-tonal orchestral music.

Acknowledgments

I express my gratitude to all who have thought along with me during developing and realizing my dissertation. I apologise to list here only more salient academic supporters.

Supervisor: Prof. Dr. Kerri Kotta – for all intensive discussions and his valuable contributions to my thought process (e.g. the salient features' approach development – publications I, Ia), always supporting and trusting my ideas and solutions, for including me into two ETAg research grants under his leadership (2006–2008, 2010–2012) that enabled me to prepare as well as develop my dissertation topic; Prof. Dr. Jaan Ross – for professional advice during his supervision 2008–2012 and enabling me to publish publication I.

Consultant: Prof. Dr. Mauri Kaipainen (University of Helsinki, FI; Prof. emeritus at Södertörn University, SE) – for his interest in my COSM approach 2017–2018, for his expertise in the cognitive realm, supportive discussions and detailed feedback to publication V and the Introductory chapter 2022–2025.

Opponent: Prof. Dr. Anna Rita Addessi (University of Bologna, Società Italiana di Analisi e Teoria Musicale, GATM, IT) – for her valuable feedback, the confidence her research methods have given me for my own ideas and the questions at the pre-defense that helped me to reconcile my research focus and goals with a compositional perspective and to be more detailed in my description of Tüür's music.

English language editor: Dr. Mark Lawrence (City University of London, UK) – for his intensive editorial work on the long version of the Introductory chapter 2024–2025.

Supporters: Prof. Dr. Urve Lippus (1950–2015, *in memoriam*, former Head of the Musicology Department and doctoral studies at EAMT) – for her support and trust, and her proposed term *listening strategy*; Prof. Dr. Kristel Pappel (Head of the Centre for Doctoral Studies at EAMT) – for her trust in my ideas and my ability to succeed; Prof. Mart Humal – for the analytical skills and care in language and content editing I learned from his supervision of my Master's thesis (2002–2004), also for his proposed term *contrast-culmination* improving the salient features' approach (publications I, Ia); Hans-Gunter Lock (lecturer at EAMT, Estonian Academy of Arts, Head of the New Media department's studio, EAMT doctoral student) – for his technological and programming skills while building the slider-controller hard- and software (Max/MSP,

since 2007, TEDEA since 2011) and the COSM software (Max/MSP, since 2017); Maris Valk-Falk (1934–2016, *in memoriam*) – while collaborating with her (since 2005) I learned the basics of empirical research methods, she has supported my ideas always firmly; Dr. Marju Raju (EAMT) – for useful methodological and structural feedback (Introductory chapter); Prof. Dr. Allan Vurma – for suggesting to look at concepts like 'attention' and 'consciousness'; Dr. Paul Beaudoin (Brandeis University, US; music researcher, composer, visual artist and educator; Fulbright scholar at Tallinn University 2015, living and working in Estonia) – for his general support and profound English language editing of my earlier articles and papers; Prof. Dr. Felix Diergarten (DE, Grundlagen der Musik, publication II), Prof. Dr. Jelena Davidova (LV, PMP Journal, publication III) and Prof. Dr. Ariane Jeßulat (DE, ZGMTH Journal, publication IV) – for their valuable and supportive editorial work on my publications; Dr. Ilkka Kosunen (University of Helsinki, FI; Tallinn University Digital Technology Institute, DTI, EE) – for providing me with the DBSCAN algorithm (2017–2019); Mustafa Can Özdemir (doctoral student at Tallinn University DTI, TR) – for providing me with the Tableau visualization software; Dr. Aleksander Väljamäe (Tallinn University DTI) for connecting me to Dr. Kosunen, M. C. Özdemir and others, and enabling valuable insights into the realm of Human Computer Interaction (HCI) including outlooks to possible future research on my topic; Dr. Toby Gifford (Queensland Conservatorium, Griffith University, Brisbane, AU) – for programming me the (Max/MSP integrated) Java-based software (2012–2013) to automatize the tension curves reduction method (proposed in publication Ia); Dr. Peep Nemvalts (Tallinn University School of Humanities, Head of the Center for Academic Estonian) – for including me in his projects (since 2013) and the possibilities to present my dissertation topics in his organized Conferences on Academic Estonian (2017 on modeling principles in music and science; 2022 on the development of definitions of musical tension (2010–2022), 2023 on psychophysical measurement), for the intense editing of and related methodological discussions improving my Estonian summary; Prof. Dr. Mario Baroni (University of Bologna; Società Italiana di Analisi e Teoria Musicale, GATM, IT) – for his kind openness to my developing COSM model during GATM conferences in Rimini 2015 and 2019; Prof. Dr. Charles de Paiva Santana (BR, Aix-Marseille University; Perception, Representations, Image, Sound, Music: PRISM Laboratory, FR) – for all supportive discussions about music theory as well as empirical and computational methods since Rimini 2019, the invitation to his co-organized conference Journées d'Informatique Musicale (JIM, Musical Computer Days) 2024 in Marseille that enabled issuing of publication V; Dr. Joshua Mailman (Eastman School of Music University of Rochester, NY;

teaching e.g. at Columbia University NYC, US, music theorist, philosopher and researcher, musician, computing specialist and composer) – for many interesting discussions at conferences and his huge dissertation (2010, 690 pages text and 380 pages examples) on temporal dynamic form and emergent properties in music as well as numerous graphical realizations that ensured me to continue on my way in thinking via visualization/representation; Prof. Dr. Pia Tikka (Tallinn University Baltic Film, Media and Arts School, BFM, FI/EE) – for introducing me in 2017 to Mauri Kaipainen while having a shared interest in narrativity and cognition; Dr. Jaak Sikk (EAMT) – for inspiring and supportive discussions during our co-authored publication on improvisation and serendipity (2021–2022), he insisted on including the concept of ‘tension’ there, too; Andrus Kallastu (composer, organizer; doctoral student at EAMT) – for supportive discussions on our both dissertations in progress.

I thank all the participants of the empirical tests conducted since 2008. Without their active participation, patience and feedback my dissertation would not be in the state it is now. I also thank all my colleagues and students at Tallinn University (TLU) (since 2005 Institute of Fine Arts, since 2015 BFM) who have supported my development and have trusted my ideas and ability to succeed. Among these colleagues I mention Dr. Tiina Selke, Maris Kirme, Andres Avarand, Dr. Vaike Kiik-Salupere, Dr. Marit Mõistlik-Tamm, Raul Talmar, Krista Aren, Krista Simson, Katrin Saks, Dr. Birgit Vilgats, Dr. Iivi Zajedova and many others. Among my numerous students who applied under my supervision in their research projects slider-controllers as well as concepts related to e.g. (musical) tension, salience or narrativity either suggested by me or independently I mention Elisa Johanna Känd. She insisted on researching narrativity in her Audiovisual Media BA thesis (2016) to comprehend audiovisual narrativity via tension while applying slider-controllers – in my dissertation I, in turn, attempt to understand musical tension via narrativity.

I am thankful to all my friends, creative and research realms colleagues in Estonia and internationally for their support in all possible moments.

Finally, I particularly thank my family: especially my beloved parents Angelika and Hans-Jürgen Lock for their endless care, trust, patience and support in all possible and impossible situations!

Tallinn, 10.05.2025

Contents

Introductory chapter

> "A single musical event can become a battleground for competing worldviews, methodologies, and social stances."
>
> "Music, though a complex phenomenon, is often stripped of its complexity by the constraints of disciplinary boundaries and individual predispositions."
>
> (Almén 2006: 1)

1. Introduction

This doctoral dissertation observes and analyzes Cognitive Musical Tension (CMT) as 'Perceived structural musical tension' (PSMT) of contemporary post-tonal orchestral music (CPTOM) applying modeling and visualization to achieve a refined perspective on complex musical structures (the general object of this dissertation). Its basic goal is to combine music analysis and music psychology, its general aim is to advance analytical methods to enhance the comprehension of such music via musical tension (the special object). With its main objective this dissertation observes and analyzes the experience of musical tension via attentive listening focusing on structural aspects (musical parameters) in the music of the Estonian contemporary composer Erkki-Sven Tüür (b. 1959).

This dissertation comprises six publications in the following (generalized) methodological realms: speculative music analysis (salient features) and empirical experiments with slider-controllers (musical tension and musical form, N=7), combined visualization of such data (publication I – Lock 2010a) as well as visualization examples complementing publication I and a (visually inspired) reduction algorithm for data gained from empirical experiments (musical tension curves, N=7 & N=6, publication Ia – Lock & Kotta 2012); general principles of representation and visualization of music and their automatization processes as intuitive, systematic, system-based and automatized analysis methods,[1] examples e.g. based on data from publication I (publication II – Lock 2014); modeling principles in science, music, analysis, composition and education, Twelve Strategic Steps for modeling/analysis of models (TSSM) developed as guiding principles (publication III – Lock 2017a) for developing a new cognitive analysis model; empirical experiments with slider-controllers (musical tension, N=26, using TEDEA[2]) and first application of my new cognitive analysis model COSM[3] (N=14; combined narrativity-based analysis of TEDEA and COSM data) to observe the salience of Impulses and musical parameters in the context of narrativity while analyzing perceived structural musical tension (publication IV – Lock 2020, defined as PSMT in Lock 2022 and publication V – Lock 2024). The five scientific Dilemmas (publication V – Lock 2024) collocate and conclude the theoretical-methodological problems that have motivated and driven this dissertation. They are crucial to achieve an empirical analysis of musical tension (CMT) as a temporal dynamic wave-like (real-world) phenomenon (TDWP).

Tensional aspects in sound and music, musical tension and intensity of musical parameters have been observed structurally, psychologically (psychoacoustically) and cognitively in tonal (e.g. Lerdahl & Jackendoff 1983; Lerdahl 2001, Farbood 2006, 2008, 2012; Huron 2006; Lerdahl & Krumhansl 2007; Eitan & Granot 2007; Granot & Eitan 2011; Krumhansl & Lerdahl 2011; Herremans & Chew 2016; Herremans & Chuan 2017; Friberg *et al* 2019; Teo 2020; Liu *et al* 2020; Sun *et al* 2020) and nontonal/atonal/post-tonal/neo-tonal music (e.g. Addessi & Caterina 2000; Pressnitzer *et al* 2000; Kirschbaum 2001; Addessi 2010; Friberg *et al* 2019; Teo 2020). Musical tension is characterized, bordered and defined shortly in 1.2, in publications I, Ia, and IV, and more detailed in 3.1.

In section 1 the goal, the general aim, and the general (CPTOM) and special objects (PSMT) of this dissertation are introduced underpinned by the motivation to apply a cognitive approach based on visualization/representation. In section 2 the theoretical background (music-theoretical and music-psychological thinking, methodological-philosophical matters, temporal dynamic cognitive processes, *microgenesis*, psychophysical measurement) and a "pre-definition" of musical tension are introduced. In section 3 key concepts (tension, salience, musical parameters, narrativity, the five scientific Dilemmas that include the five Research

---

[1] First proposed in Lock 2006.

[2] Tension Design Experimental Apparatus consisting of hard- and software developed for this dissertation (Lock, G. & Lock, H.-G. 2011).

[3] Cognitive Octagonal Slice Model (Lock & Kotta 2015, Lock 2017b, Lock 2019, publication IV, Lock 2021) realized with software written by Hans-Gunter Lock.

Questions (RQs) and the five Statements (STs) as the methodological frame, and my cross-domain definition of musical tension (PSMT) are presented. Answers to the Dilemmas/RQs/STs can be found in publication V offering methodological solutions relevant for this dissertation. Finally, the main objective of this dissertation is presented in 3.7 underpinned by discussing crucial terms (*purportedly* and *trigger*) and the *why*-question.

The following subsections present the basic goal and the general object (complex musical structures analyzed/observed), the general aim and the special object (musical tension analyzed/observed), reasons and justifications of the proposed cognitive approach, and basic methodological aspects (introspection, intuition, perceptual experience, triangulation). These subsections prepare the reader for the upcoming theoretical-methodological section 2 on the theoretical background, and subsection 3 on the key concepts and a cross-domain definition of musical tension.

1.1 Basic goal and general object: complex musical structures

The basic goal of this doctoral dissertation is to intertwine music-analytical (theoretical) and music-psychological (empirical) thinking (e.g. Deutsch 1984; Deliège 1987, 1989; Clarke 1989; Cook 1994; Deliège 1996; Cross 1998a; Addessi & Caterina 2000; Deliège 2001; Cook & Clarke 2004; Pearce 2005; Addessi & Caterina 2005; Ockelford 2006, 2009; Deliège 2007; Addessi 2010, Friberg *et al* 2019) to enhance the comprehension of contemporary post-tonal orchestral music (CPTOM) – the general object – and musical tension – the special object of the observation and analysis of this dissertation.

Music analysis and music psychology are embedded in a broader field of interdisciplinary research that interconnects experts from psychology, music theory, composition, computer science, music technology, and engineering with experiences in musicology, cognition, psychology, signal processing etc.[4] The aim is "to better understand the structures that emerge in composition, performance, and listening, and how these structures interrelate" as an ongoing interest in understanding and applying computational music structure analysis "to segment or

---

[4] Dealing with both possibilities and constraints (Almén 2006: 1), see the second moto-quotation at the beginning of the Introductory chapter.

decompose music into patterns or units that possess some semantic relevance and then to group these units into musically meaningful categories" (Müller, Chew & Bello 2016: 148–150). The exploration of methods and comparison of results of music theoretical analysis with empirical psychological tests especially for post-tonal/atonal music has been explicitly conducted by members of the research group GATM (*Gruppo Analisi e Teoria Musicale*) led by the researcher of music analysis and systematic musicology Mario Baroni dealing e.g. with a listening-based and cognitive understanding of "macroform" (Baroni 2003, Addessi & Caterina 2005) and the perception of tension/relaxation (Addessi & Caterina 2000, Addessi 2010) of post-tonal/atonal music. More recently they have collaborated with researchers of tonal (and also pop-music) melodies to develop (as it is the tendency in psychological research, also in tension research) computational prediction models (Friberg *et al* 2019). While the above mentioned earlier research observed musical tension/relaxation in post-tonal/atonal string quartets (Addessi & Caterina 2000, Addessi 2010), which is a reasonable amount of voices and deals with manageable lengths' of parts or whole musical works, the later research (Friberg *et al* 2019) is reduced to single melodies but aims at a wider range of styles over several centuries (see more on several musical tension research approaches in 3.1). However, in the interest of my dissertation are rather complex musical structures of contemporary post-tonal orchestral music (CPTOM) (the general object) that are not easily manageable from the point of view of multiple and often dense voices and layers.

As follows, I shortly border here the terms *contemporary*, *post-tonal* and *orchestral* in the context of *symphonic* music. The terms *tension* and *musical tension* are bordered shortly below in the general aim of this dissertation (see 1.2). The term *contemporary* is used in this dissertation for music technically composed since around the 2000s and refers to its immediacy (being contemporary) to me during my master's studies (2002–2004 at the Estonian Academy of Music and Theatre) and preparing (2004–2007) to submit for doctoral studies (2008). The term *post-tonal* broadly refers to music that does not apply functional harmonic chord structures or its extensions (e.g. see Roig-Francolí 2008), this can be understood as the broad definition of *post-tonal* in contrast to the narrow definition of especially atonal and nontonal music. The contemporary post-tonal orchestral music (CPTOM) of the Estonian composer Erkki-Sven Tüür (b 1959) has been chosen based on the conditions of its contemporaneity to me and the broad definition of the term *post-tonal*. Furthermore, Tüür merges in his music several different styles. On a "mind map" branching into three styles or a Venn-diagram consisting of three areas covering atonal/nontonal/avant-garde,[5] "neoclassical"[6] and rock-music[7] (with which he started, before he became an avant-garde composer) Tüür's music (especially the 4th Symphony/percussion Concerto *Magma*) can be located between different Western music styles while more and more merging them. I have Tüür's music characterized as individual and universal, combining classical-romantic/traditional and modern/avant-garde compositional thinking reflecting sensibly and detailed our times but also as "timeless" while synthesizing many traditions universally and simultaneously (Lock 2019/2011: 158, Lock 2011b: 5). Tüür's music "includes, expresses and embodies time and space in many aspects and layers from the structural to the emotional dimensions, the constructed to the natural [organically growing], the philosophical and metaphorical to the semiotic, the chronological to the subjectively experienced [temporal dimension]" (Lock 2011b: 5–6, [Lock 2019/2011: 158/159]). In the context of universality, time and phenomenology (in the spirit of Merleau-Ponty 1945/2005) I see the listener in an encounter with Tüür's music as "the brooding presence moving to meet him, like a storm [tsunami or volcano] on the horizon" to be "touched, overwhelmed, astonished, exalted, even rendered speechless by the suggestiveness of its sounds, images and the memories it evokes" (Lock 2011b: 6, [Lock 2019/2011: 159]).

In Tüür's so-called first period until 2001 he describes his compositional approach as a "polystylistic metalanguage" (meta in the meaning of archetypes) – a combination of contrasts as structural blocks, an abstract drama with characters and dynamic event-chains that "continuously shift, widen and are compressed" (Rosin 2024: 15–16, referring to Tüür 2001: 3). In his so-called second, ongoing or vectorial period since 2002 Tüür has developed an original voice-leading- and interval-based "vectorial writing" or "vectorial" method, approach or technique (Tüür 2003/2007: 4, 6) of aggregating intervals to create a "multidimensional musical space" "based on a relatively short numeric combination/code" (Kotta 2011: 104, Kotta 2009/2010: 5, see also Kotta 2017, Rosin 2024: 17). The code/vector principle organizes rather the contrapuntal structure (vectorial voice-leading principle, Kotta 2009) of Tüür's music, less influencing its "rhetorical and formal [...] aspects" (Kotta 2011: 104). Therefore, the code/vector principle is not further investigated neither in Kotta (2011) nor in this dissertation. Although musical parameters are part of the methodological design, and pitch and rhythm constitute the counterpoint, Tüür's compositional principles to achieve this texture are not in

---

[5] Absence of functional tonality, experimentations with all musical parameters and the musical form.

[6] Extension of the functional harmonic traditions, the musical form stays rather classical.

[7] Including a rock music cadence in the 4th Symphony observed in the main experimental study, see publication IV.

the focus of this dissertation. Furthermore, musical tension cannot be comprehended without reference to musical form (publications I, Ia, IV). Kotta's analyzed rhetorical aspects are somehow akin to the narrative approach I later suggest in this dissertation (publication IV, see 3.5) but their relationship is shown in a comparatively manner mainly on the global level of musical form ("macroform," Baroni 2003, Addessi & Caterina 2005) in relation to the emergence of symphonic and sonata-form principles, not dealing with more specific "semantic" or deeper structural local levels.

Tüür's 4th Symphony/Percussion Concerto *Magma* (2002) is composed partly and *Oxymoron* (2003 for large ensemble) is the first work composed entirely with his "vectorial" method. These two works mark a change in Tüür's compositional method; *Magma* holds a transitional position, Tüür generates the basic musical material from constructed modes, not anymore from contrasting blocks and the role of organic development principle increases (Rosin 2024: 17–19, e.g. referring to Tüür 2012/2023). But it is a rather consecutive further development from the earlier contrastive to more extensive "transitional" or converging textures (see publication I, Appendix C.1) and organization of musical time (Kotta 2008, 2011).

Tüür also applies a special compound musical parameter he calls "tone" (in his words not simply "surface"-timbre)[8] as a form-, structure-, "potential energy/musical inertia,"[9] dramaturgy and therefore a narrative-constituting aspect that he also visualizes in the z- or depth-dimension during the planning phase of composing (Kotta 2009/2010: 4–5, Kotta 2009). However, his "own voice" he has achieved while merging the "vectorial" method with the "tone" principles since the 5th Symphony (2004) (Kotta 2009). So, the works analyzed in this dissertation exemplify an ongoing development of his musical language.

However, this dissertation neither analyzes the composer's technique nor is restricted to the observation of "primary" musical parameters (pitch, rhythm, harmony). It embraces both "primary" (PP) and "secondary" (SP) musical parameters (Meyer 1956, Snyder 2000, see 3.4).[10] PPs' leading role especially in functional harmonic music is extensively covered but for avant-garde and post-tonal music (also CPTOM) a stronger role for SPs has been suggested (Bauer

---

[8] Concerning *Magma* this is also described by Tüür as 'timbral harmony' in the context of colour as an important aspect in his music (mediated by Martin Anderson, see Tüür 2002/2007: 5, 8, 11).

[9] Kerri Kotta, has developed and applied principles of the qualitative state of musical energy (publication Ia) and on "energy conversion" in the music of Erkki-Sven Tüür (Kotta 2013).

[10] PP: pitch, rhythm, harmony; SP: dynamics, tempo, instrumentation/timbre, texture, effects. The secondary parameters are taken from Martin Kirschbaum (2001) while adding a combined category I call instrumentation/timbre – because, in this context, timbre is for me a feature of musical instruments, therefore expressed/detectable in instrumentation.

2001): see hypothesis 1 postulating the salience of SPs over PPs in Tüür's 4th Symphony/Percussion Concerto *Magma* in publication IV.

Tüür’s dense orchestral music is “rich” in details and intuitively “translucent” for listeners also without explicit music educational background (perceptually/cognitively and dramaturgically comprehensive on a narrative level, see 3.5, publication IV, see Lock 2011b: 5, Lock 2019/2011: 158) but rather “opaque” for the traditional analyst. What I mean with “translucent” is related to my notion of Tüür’s music as being “like an open book” for the experienced listener (particularly those also knowing Estonian and Tüür’s music) while “memories of meanings, styles, images, times and spaces from the distant and recent past of Western music – from ancient music to jazz and rock [...] – seemingly emerge from nowhere and fade back into sonically moving forms and the play of pure musical material” (Lock 2011b: 6, [Lock 2019/2011: 159–160]). I understand Tüür’s music as a kind of “perfection” of Hanslick’s (1854) idea of “sounding moving forms” (German *tönend bewegte Formen*) (Lock 2019/2011: 158, Lock 2011b: 5–6, see also 2.3). What I mean with “opaque” I have described mainly figuratively as follows:

> “For the less experienced listener or someone approaching it from a purely phenomenological standpoint, Tüür’s music is like a magic book. While pages turn, culminations appear, and natural phenomena like storms, volcano eruptions and frozen tsunamis roll over [us]; single notes and motives, gestures and abstract narratives, impulses and culminations [see Lock 2010a)], patterns and pulsations, contrasts and crucial changes in texture emerge rapidly or slowly, remain continuously or move smoothly from one state of the musical material to the next [(see Kotta 2008)]. Tüür’s titles and the concepts of his works often support and awaken these impressions, associations, analogies and synesthesia or natural imagery and phenomena.” (See titles like *Magma* and *Oxymoron*, Lock 2011b: 6–7, [Lock 2019/2011: 159–160].)

As an aspect of this “opaqueness” I also understand the fact that music-theoretical methods for post-tonal music are usually pitch-based and often do not cover neither the “tone,” the “transitional” merging textural/teleological and time-related developments (despite form-analysis covering also temporal dynamic aspects), the symphonic dramaturgy (in my view the narrativity) nor the wave-like tensional aspects of such music.

In the context of orchestral and especially symphonic music the main supervisor of this dissertation, Kerri Kotta (2009: 5), underlines that two at the first glance contrasting and “even antagonistic musical attitudes"[11] are merged in Tüür's music: the classical hierarchical (structural, linear, teleological), and a sound-centered multidimensional "tone"-based (spatial, non-linear, non-hierarchical) way of composing. Tüür's musical form incorporates musical archetypes of – but the composer does not deliberately follow – the classical symphonic tradition (Kotta 2009/2010: 4, Kotta 2009; Kotta 2011). However, the boundaries of its formal units are veiled through the wave-like dynamic development[12] of Tüür's music. Its texture spans from melodic-motific ("slow" time) to rather dense ("quick" time) and even sonoristic (soundfield-like "continuous" time) to rock-style percussion (rhythm-based) structures[13] (especially in the 4th Symphony/Percussion Concerto *Magma*). These features are the reasons why I have chosen Tüür's music to serve to exemplify the general object (CPTOM) of observation and analysis in this rather methodological dissertation. As follows I present the general aim and the special object – musical tension – of this dissertation.

1.2 General aim and special object: musical tension

The general aim of this dissertation is to advance analytical methods to enhance the understanding of complex, multilayered and dense contemporary post-tonal orchestral music (CPTOM) (partly also called sound-mass music) while observing and analyzing the special object of this dissertation – musical tension – as a compound phenomenon that can be experienced easily intuitively also by non-expert listeners but is complicated to research systematically (Farbood 2012, Lehne & Koelsch 2015), see 1.3, 2.2.[14]

Tension as a general human experiential trait is a complex and multidimensional phenomenon that is observable especially in the arts (Lehne & Koelsch 2015). *Tension* related to *sound* and

---

[11] See also the meaning of the title *Oxymoron* (2003): contradictory terms used in conjunction; incongruous, seemingly self-contradictory effect (Oxymoron DC [2025]).

[12] Rather deliberately composed tensional high-points and low-points define its form (see publication I, Appendix C.1). The wave-like dynamic feature is connected to musical tension as a temporal dynamic wave-like phenomenon (TDWP) (see publication V).

[13] The terms in quotation-marks have been developed and applied by Kotta (2008, 2011) and further developed by Kotta in Gerhard Lock, Charles de Paiva-Santana and Kerri Kotta (2021). Texture types that converge (merge) in Tüür's music: (1) sound-centered ["continuous" time, e.g. static chords, clusters], (2) melody- or line-centered ["slow" time], (3) rhythm-centered texture ["quick" time].

[14] On music-theoretical and music-psychological thinking intertwined see 2.1. See also the five scientific Dilemmas shortly presented later in 3.6 and explained in detail in publication V.

music can be understood as an emotional aspect, as something beyond emotions, as a semantic characteristic of music as well as on the structural level, e.g. as intensities and a process triggered by the interaction of several musical parameters (see 3.1). Musical tension has been crucial for composing, performing and understanding Western music for centuries. It has been often verbally described (also by non-music-background listeners) with verbalized narratives or graphical visualization (e.g. in the so-called energetics-tradition in the first half of the 20th century, see Kirschbaum 2001, Rothfarb 2002), see overview of music and narrative since 1900 (Mailman 2012/2013, Klein & Reyland 2012/2013).[15] Intensity contours of different musical parameters are perceivable as analogous and can be called musical "gestures"[16] independently of the specific parameter (Eitan & Granot 2007: 39). Musical tension manifests itself during the flux of music as temporal experience triggered by the sensation of musical parameters in interaction (Farbood 2012: 387). Musical tension is universally comprehensible (Farbood 2012, Lehne & Koelsch 2015), nonetheless, it has been less investigated using integrated formal analytical, structural and empirical approaches (see 3.1, 4.1).

In this dissertation I propose the following three approaches to enhance the comprehension of CPTOM: (1) I understand music as environment (publications III, IV, see 1.3). In an environment, therefore also in music as environment, sounds and musical events trigger cognition[17] – e.g. as a general metaphor (Reybrouck 2015), in the *acoustic* (or physical) and the *auditory* (or perceived) dimensions (Wiggins 2009, see 2.2 and publication V), as cognitive/knowledge-based dimension via principles of computation (Kaipainen 1994) or as "ecological approach" of "direct" perception in music (Clarke 2005).[18] (2) I also apply computational means as well as modeling (e.g. Dubnov & McAdams 2006, Dahan 2012, de Paiva Santana & Guigue 2020, Noble *et al* 2020, Antunes *et al* 2021/2023) (see publications III, IV, 4.1) based on principles of representation and visualization (see publication II) to achieve the general aim and the main objective of this dissertation. (3) My research has led me to consider also intuitive (musical) thinking alongside with scientific (musical) thinking (for music see Clifton 1975, for cognition and science in general see 2.2 and Kant 1998), that is mirrored in the juxtaposition of the so-called first-person *vs* third-person research methods.

---

[15] See also verbalizers' *vs* visualizers' cognitive learning styles (Mayer & Massa 2003, Chen & Sun 2012). See linguistic (language-based narratives) *vs* non-linguistic (including extra-linguistic structures called models, visualized representations) forms of scientific theories and their representations (Hendry & Psillos 2007).

[16] Compare the concept of 'gesture' (see 4.1) with *temporal dynamic forms* (Mailman 2010) and *dynamic forms of vitality* (Stern 2010), see 2.3.

[17] See the first moto-quotation (Almén 2006: 1) in the beginning of the Introductory chapter.

[18] This has implications for research methods and for contemporary post-tonal orchestral music, see section 4.

However, this dissertation does not expand its methodological background towards recently re-established phenomenological (first-person) approaches (Gallagher & Zahavi 2008/2012/2020; Valenzuela-Moguillansky & Demšar 2023) because phenomenology is not my expertise. Furthermore, intuitive musical thinking immediately evokes association with the concept of ‘speculation’[19] that generally has a rather negative connotation. However, speculation belongs to the methodological toolbox in science and also in music theory.[20]

According to the cognitive and neuroscience researcher in music Marcus Pearce (2005: 13–14) music theory is generally characterized as a speculative discipline where the analyst is applying “deduction from definitions of concepts, self-evident principles and generally accepted propositions” while feeling conviction. Empirical disciplines, in contrast, can be said to apply experimental and scientific methodologies while taking into account the epistemological status of scientific knowledge, the principle of falsification and its criticism on several levels (Pearce 2005: 13–16). Hence, music analysis – if it is not just statistically analyzing MIDI or other automatized notation/representation data (see publication II) –, is first of all conducted by humans based on music theoretical methods between art and science (Broman & Engebretsen 2007).

In this dissertation I prefer the concept of ‘intuitive’ over the concept of ‘speculative’ to characterize the listening process conducted by both the attentive listener and the researcher. ‘Intuition’ is often based on implicit pre-knowledge and plays an important role in an Kantian *analytic idealist* understanding of “inner sense” (Kant 1998, see 2.2). The etymological origin of the term *intuition* includes meanings of insight, direct or immediate cognition, and spiritual perception, the Latin term *intueri* means “look at, consider” (see tutor) (Intuition O-ED [2025]). In my view, intuition refers, broadly speaking, to the self-tutoring, self-teaching and self-learning[21] activity of the mind in dealing with (after Kant) the manifold of intuition via sensing while discovering “the one in many” via active unity construction: “In short, for the given

---

[19] The term *speculation* has an every-day pejorative meaning (guessing without certainty, Speculation CD [2025]) and a neutral or even academic meaning based on negative examples – the etymological origin of the term speculation means exploration and observation (Speculation DC [2025]).

[20] I generally do not use the dichotomy objective *vs* subjective in the usual meaning of scientific *vs* non-scientific or quantitative *vs* qualitative approaches, because these concepts are too complex and lead away from the scope of this dissertation. Only in the context of a Kantian understanding of (ap)perception via sensation and receptivity (objective in the meaning of objects in space and “outer” sense, Kant 1998) as an early psychophysical measurement approach I follow Kraus' (2013/2016) proposition to use an “outer” sense physical object to measure “inner” sense (see 2.2 and publication V).

[21] Self-learning as a principle is used also in artificial intelligence algorithms and fully accepted as a scientific approach.

manifold of intuition to become knowledge, it must be *synthesized*" (Wolff 1963: 149) (see in 2.2).

More concretely I propose in this dissertation a strategy for music analysis that spans from intuitive and systematic to system-based and automatized analysis methods (in the context of visualization of music, see publication II). These cover the aforementioned different methodological approaches using speculative (music theoretical) and my preferred intuitive (listening strategies and attentive listening) as well as empirical (perception experiments) (for both see publication I, Ia and IV), modeling (see publication III) and computational means (computation-supported data-capture and analysis see publication IV and section 4). As follows I present the reasons and justification why to arrive at a cognitive approach.

1.3 Towards a cognitive approach

The general and the special object of observation of this dissertation – music and musical tension – are both complicated to pin down analytically and scientifically (see the five scientific Dilemmas in 3.6 and publication V). Music's relation to science and comprehensibility via science, therefore strict and systematic analysis is problematic (Cross 1998b) (see 2.2). Strict analysis requires systematization to become scientific knowledge (Hoyningen-Huene 2013: 21 etc).

As mentioned in the subsection on the general aim (see 1.2) the understanding of music as environment (publications III, IV) affording and triggering cognition and intuitive musical thinking alongside with the opportunities of a broader field of interdisciplinary research (including computational means, modeling, representation and visualization) have guided me to enhance the comprehension of CPTOM and musical tension with a cross-domain cognitive approach. The latter mirrors the methodological actuality of today's research, e.g. in (embodied) cognitive or (not-yet-embodied) AI-based modeling approaches. As a methodological-philosophical basis for such a cognitive approach I have chosen *transcendental metaphysics* / *analytic idealism* proposed by the philosopher Immanuel Kant in the *Critique of Pure Reason* (first edition 1781, second edition 1787) (see Kant 1998). Recently *analytic idealism* has become again more prominent in a kind of renaissance (e.g. Kraus 2013/2016, Kastrup 2019).

Analytic idealism assumes that human perception and cognition is a solely mind-driven and active process while acknowledging immediacy and intuition in cognizing phenomena[22] in the world as a *synthetic unity* of *apperception* (Kant 1998: 282, see 2.2). Sound and musical events are such synthesized phenomena, music creates a musical world and is an agent of the world as environment (the Pythagorean-Platonic viewpoint, Mailman 2016: 12) open for humans' receptivity, sensitivity, thought and conceptualization.

More specifically, music understood as an abstract and cultural construct is describable by three representational domains: the *acoustic* (or physical), the *auditory* (or perceived) or the *graphemic* (or notated/also digitally encoded) domain (Wiggins 2009: 477 after Babbitt 1965, 2003: 204, see below and 2.2). The notated domain can be discriminated further into pre- and descriptive approaches (see publication II) that applies also to music analysis (DeBellis 2002) (see my enlarged visualization in Figure 1 in publication V).

In short, this dissertation proposes a cognitive approach because music and musical tension are both identifiable as cognitive. Cognitive Musical Tension (CMT) (see 2.2 and publication V) applies music theoretical (score-based listening) analysis as well as empirical (attentive listening) perception experiments (see publications I, IV) using (cross-domain) representation-visualization means (see publications I, Ia, II, IV) as well as modeling (see publications III, IV).

The justification of a cognitive approach is in line also with the following general goals the composer of the music observed and analyzed, Erkki-Sven Tüür, has expressed referring to sensation and integrated cognition in a modern acoustic world (Kotta 2009: 4):

> "I would like to unfold the entire genesis – from micro level to final composition – before the senses of the listener. I would like to let people follow the process of creation from the atomic elements to the birth of something big."
>
> "[...] it is my purpose to demonstrate the vitality of symphony as a genre that can absorb and integrate acoustic material so paradoxically taken from the same fragmentary world. I'd like to demonstrate the possibility of integrated cognition in [the] modern acoustic world."

Therefore, in my view it is appropriate to approach Tüür's music considering a cognitive approach as well as applying representation and visualization means, because the composer

---

[22] Appearances via sensitivity and receptivity as representations of the things-in-themselves (*noumena*) humans cannot cognize directly (Kant 1998: 8).

explicitly applies visualization as a tool in the composing phase (see 1.2), the musical score (as a form of symbolic visualization) generally represents the music visually (before/after it is performed), and the composer has expressed his integrated-cognition goals to cope with a modern acoustic world.

Representation and visualization are widespread in cognitive approaches in Western culture and science (e.g. Levitin 2002; Nanay 2007, 2009/2010; Shea 2018; Roy *et al* 2018). Modeling itself uses analogies and includes representations and visualizations. Models or families of models are non- or extra-linguistic conceptions of scientific theories (Hendry & Psillos 2007: 129). In my view models should be more complete (as a whole) than theories (that must stay open to confirmation/refutation), theories are the backbone of models (publication III). One can proceed from a model to a theory and *vice versa* using numerous types of visual representations/visualizations (Saldaña 2024). However, representations depend on enculturation (Hipólito 2022: 8) and are object of debate, e.g. rejection of the mind-computer analogy (Hipólito 2022: 8, 12; on the mind-computer metaphor see also Kaipainen & Tikka 2022), repudiation of representations, depictive *vs* propositional formats of representations (Bechtel 1998). In a Kantian understanding humans' sensation and cognition solely deal with representations of appearances/phenomena, not the-things-in-themselves (*noumena*) (Kant 1998, see 2.2). Representation and visualization of music is part of the arts as symbolic (notation-based) languages and (extended) cognitive systems (e.g. Goodman 1968, Honing 1993, Kaipainen 1994, Pearce 2005, Nussbaum 2007, Acotto & Andreatta 2012, Chouvel 2014, Kersten 2014/2015). Particular approaches including specific types of visualization and representation are observed in auditory and mental imagery (Phillips 2018) and the Conceptual Blending Theory (Antović 2018, 2022). The MusicVis Project "enriches music visualization and opens new research paths in Visual Musicology, tackling challenges in music analysis and theory. It exemplifies the synergy between musicology and visualization, fostering innovation and artistic exploration" (Miller, Fürst and El-Assady [2025]), e.g. analyzing visual mappings of traditional and alternative music notation (Miller *et al* 2018). See more on methodological aspects of visualization of data in 4.5.

In order to achieve an own theoretical, conceptual and methodological identity "to become part of cognitive science" (Núñez *et al* 2019: 784) this dissertation approaches Cognitive Musical Tension (CMT) via cognitive modeling: 1) introducing principles of modeling in science and music while developing the Twelve Strategic Steps for modeling/analysis of models (TSSM) (see publication III); 2) based on TSSM the development and application of a listening-based analysis model COSM: Cognitive Octagonal Slice Model (see publication IV and section 4). As follows I present basic methodological aspects: introspection, intuition, perceptual experience, triangulation.

1.4 Basic methodological aspects

This dissertation is restricted to basic methodological aspects of music theory/analysis, empirical musicology/music psychology and philosophy/Kantian *analytic idealism* (Kant 1998, see 2.2).[23] The philosophical-methodological background of this dissertation is introspection and intuition, perceptual experience and perceptual justification (Clifton 1975; Morales 2018, 2024; Schwitzgebel 2024; Silins 2024). Their elementary roots I see in Kant's (1998) *analytic idealism*. Furthermore, I assume that music can only be grasped adequately via listening (auditory, perceived domain, Wiggins 2009) based on experience as intuitive cognition/consciousness/(musical) thinking (Clifton 1975, Kant 1998), musical sense-making (in- and out-of-time) (Reybrouck 2019), and detection that leads to the realm of psychophysical measurement (e.g. Steingrimsson 2016, Michell 2020, Vessonen 2021, see 2.2). The "more-or-less abstract representations of musical events or phenomena" (Cross 1998b: 211) can be realized with visualization (publication II) and modeling (publication III) supporting the understanding of music via theoretical analysis and empirical experiments also with visualizing means (Levitin 2002: 128–129, see 4.5) while aiming at systematization to gain scientific knowledge (Hoyningen-Huene 2013).

In this dissertation I use a *triangulation* (Denzin 2012, Heesen, Bright & Zucker 2019, Arias Valencia 2022/2023, Kaipainen & Tikka 2025, forthcoming) of three (of each other) not fully independent empirical data retrieval methods via attentive listening. Method 1 is the detection of the increase, decrease or remaining state of perceived musical tension – the data is gained with "continuous" slider-controllers (see 4.3). Method 2 splits up into two interconnected sub-steps that take salience as their main mechanism (see publication IV). Method 2a is detection

---

[23] Although they might be fruitful in future, I have not used methods of neurophysiology, psychoacoustics or biometrics because these are not my fields of expertise or training.

of salient Impulses), method 2b is detection of the salience of musical parameters in the Impulses (data points) enlarged into 14 seconds chunks/time windows (see publications IV). Furthermore, the described methods are interconnected by a principle I call Impulse or "moment of change."[24] Salience also plays a role in the slider-controller method, although its detection is not explicitly asked from the participants. A more detailed explanation and justification of my triangulation-type research design see in 4.1.

In addition, this dissertation includes also two methods from music analysis that are subjective and speculative (however, still using perceptual experience, introspection and intuition): 1) detecting salient features (local: Impulse and Culmination; global: Contrast and Contrast-Culmination) during listening based on the musical score (publications I, Ia) that I later have abandoned (see Appendix C). 2) Results of the rhetorical/rotational form-analysis-method applied by Kotta (2011, based on Hepokoski & Darcy 2007, see publication IV) are included in order to achieve a comparison between empirical and music theoretical results in the context of the *graphemic* (notation) vs *auditory* (perceived) domains (Wiggins 2009, see 4.1, publication V).

The following section 2 presents relevant theoretical background on music-theoretical and music-psychological thinking intertwined, methodological-philosophical matters (e.g. Kantian *analytic idealism*), temporal dynamic cognitive processes and *microgenesis*, psychophysical measurement principles, cognitive dynamic listening models, and *protonarratives* as a "pre-definition" of musical tension.

---

[24] For the concept of salience of "moments of change" also the Ecological Approach (e.g. generally and for vision see Michaels & Carello 1981, Warren 2005, for music see Clarke 2005) is relevant. Salience emerges through attentive listening enabling us to approach music beyond prefixed and culturally-historically "limited" knowledge concepts of style and form in music analysis. While juxtaposing (absolute) time (as abstraction from change) with change (making time exist) Claire Michaels and Claudia Carello (1981: 13) define change as "events in space-time" and claim that "change is what is perceived." This is exactly what I apply in the methodology in my dissertation (see section 4).

2. Theoretical background

This rather methodological interdisciplinary, cross-domain and cognitive dissertation in music analysis and music psychology is written in the context of broadly accepted scientific paradigms of music as well as human perception and cognition research based on formalist-mechanistic/materialist,[25] dynamicist-environmental, connectionist, and *Gestalt* (and holistic) principles. In the following subsections I discuss music-theoretical and music-psychological thinking intertwined; methodological-philosophical matters (including Kant's *analytic idealism*; temporal dynamic cognitive processes, psychophysical measurement and meaningfulness; cognitive dynamic listening models; and *protonarratives* as a "pre-definition" of musical tension. This latter "pre-definition" aims to prepare the reader for the fully developed cross-domain definition of musical tension presented in 3.5.

In this dissertation I explain, define, or border, chosen terms and concepts at certain points when necessary (e.g. also terms included into the main objective, see 3.7). See for instance the overview of the definitions of musical tension I have developed between 2010 (publication I) and 2020 (publication IV) (Lock 2022, in Estonian). The phrasing "perceived structural musical tension" I use in publication IV (in general defined already in publication Ia, see 3.1) I have defined in this Introductory chapter from a cross-domain/cognitive viewpoint (see 3.5). However, before arriving there I present the concept of *protonarratives* (Imberty 2000) as a "pre-definition" of musical tension (see 2.6). In this dissertation there is no space to deal with the language matter crucial for science more deliberately. However, for instance, *cognition* (Cognition O-ED [2025]) as a term has been used widely but often ambiguously in everyday language (especially in English) and in scientific context in many countries through many centuries. English is the dominant language for research and publication in cognitive science (Blasi *et al* 2022: 1153) since the establishment of it as a research realm during the second half of the 20th century. Therefore it seems inevitable that research publications will be understood and reviewed internationally in English. Interestingly, Damián Blasi *et al* (2022: 1153) state

---

[25] The utterly materialist concept of science has been criticized e.g. as dogmatized by the biologist Rupert Sheldrake (2012a & b, 2020) and by the philosopher Bernardo Kastrup (2019) from the viewpoint of Immanuel Kant's (1998) *analytic idealism* (see 2.2) The absoluteness of science in terms of methodology, "truth," reliability and objectivity as well as the outdated distinction between art and science has been criticized by the philosopher of science Nancy Cartwright *et al* (2022) underlying that science is multifold and a practice similar to art.

that over-reliance on English *hinders* cognitive science (mentioned also in Lock 2023b, in Estonian). Despite this critical viewpoint, a considerable part of this dissertation is still written in English (Introductory chapter, publications Ia, III, V). Furthermore one publication is issued in Estonian (publication I), two others are printed in German (publications II, IV). A researcher has to be aware of issues and matters concerning a) the language background (in my case German as mother tongue, Estonian as adopted language in the place of residence and English as international research-dominating learned foreign languages), and b) methods and traditions (also related to different languages) of the several fields of research that may stem from linguistic as well as (often unreflected) common use aspects. These aspects in turn stem from etymological meanings as well as conventions of use through their history and development. This leads us further to the next subsection on music-theoretical and music-psychological thinking intertwined.

2.1 Music-theoretical and music-psychological thinking intertwined

As mentioned in the beginning of the Introduction section, the basic idea of this dissertation is to intertwine music theoretical and music psychological methodological thinking. Music theory is devoted to the analysis of concrete musical works and to discover general principles (of form, harmony, melody, rhythm etc) using speculative (Pearce 2005) and even "folk psychological" approaches (Cross 1998a). Music theory has developed historically grounded and also modern analysis methods. It gains knowledge about the music/works/repertory, disputes between different interpretations of analysts enrich the discourse. Music psychology applies methods of empirical science and uses music/sounds mainly as stimuli while asking questions and posing hypotheses on musical/sound phenomena in tests and experiments. It mainly observes the behavior of participants, but it draws conclusions also about music/musical structures.

In this context two important questions can be posed: (1) How able is the psychology of perception to contribute to the definition of analytical rules, especially for nontonal music with a not widely accepted grammar structure (Addessi & Caterina 2000: 32)? (2) Does traditional music theory assume a somehow fundamental holism via understanding the work as a whole before comprehending its parts (Dahlstedt 2007: 254)?

The first question was posed and answered by the musicologist and music educator Anna Rita Addessi and the psychologist and psychotherapist Roberto Caterina (2000: 33) based on previous studies arguing that common perceptual patterns are identifiable both for tonal and post-tonal music due to nontonal-harmony-related factors while *cues* (a concept introduced by the cognitive music researcher Irene Deliège 1989, 1996, 2001, 2007) are abstractable from the surface rather than based on features of musical structure (see more in 2.5). They conclude that this also encloses aspects of musical memory that are influenced by (special) *cues*, and is related to Michael Imberty's dual function of musical memory: "informative memory" organizing sound into syntactic structures (allowing segmentation and sectioning based on tonal/atonal style); "dynamic memory" dividing sound temporally – here tension and relaxation can be found. Addessi and Caterina (2000: 45–46) conclude that "segmentation and the perception of tension and relaxation are the basis of musical listening and of musical memory. These moments (segmentation, tension and relaxation) constitute a common perceptual process in which constant and variable elements are both maintained, even in the absence of an explicit musical grammar as in post tonal music" – principles of "sameness" and "difference" cause segmentation borders based on surface aspects (invariants) that need not to be grammar or syntax rules. In a later study, Addessi (2010: 228) underlines that "although belonging to two different domains (the psychology of perception and theory of music), the two analyses (score analysis and auditive analysis), have several activities in common: the segmentation into small, medium or large pieces; the grouping formation; the organizations of hierarchies, and so on." According to Addessi (2010: 228) the psychological approach can "render explicit the perceptual supposition present in a piece of musical analysis, and in so doing make them debatable and verifiable." Addessi (2010: 228) aimed to explore „rules" useful for score analysis via auditive analysis of macroform (see 4.1).

The difference between music psychology and analysis of music has been described by the cognitive musicologist Emilios Cambouropoulos and music theorist and composer Costas Tsougras (2009: 119) saying that "the way a musical work is perceived by a listener may be significantly different from the organization of notes suggested by a score, or even from analytic results given by different musical analytic methodologies." They confirm, based on their interdisciplinary research on Ligeti's cembalo piece *Continuum* (1968) and applying audiovisual Gestalt principles, that "musical psychology can offer a very fruitful way of looking directly into certain structural features of music that other analytic methodologies have difficulty dealing with" (Cambouropoulos & Tsougras 2009: 119).

The second question was posed and answered by the musicologist and researcher of the history of philosophy Sten Dahlstedt (2007: 262) proposing that the musical material can be understood on two levels: mainly as a transformational process via the mind “from something strictly material to something musical.” Only at the following level the musical material goes through a joining process that turns it into a work of music “with aesthetic qualities and value.” Similarly, Clifton (1975: 70) has music described as intuitively experienced via its inherent meaning independently of later objectification or justification “by scientific description.” This, in my view, can be traced back to Kant’s (1998) analytic idealism: the joining process of the musical material into music (Dahlstedt) as synthetic unity of apperception, and intuitive experience relying on inherent meaning (Clifton) as empirical (via sensation) and pure intuition (via thought) (see more in 2.2).

Dahlstedt (2007: 262) underlines that “what is presented to our perception when we are listening to music is something other than physical attributes.” In his opinion the re-integration of “human feeling and thinking” into a physical understanding is “a rather simple and reductive materialism that has been too common during the twentieth century” (Dahlstedt 2007: 262). He continues that while retaining a materialistic viewpoint one must admit that we actually do “not know the whole process between the changes in air pressure and the experience of music.” He suggests that “the formation of musical entities could take place already within the act of perception” (Dahlstedt 2007: 262).

The complicated relationship of music theory with science takes place between the singularity of music and its analysis on the one hand, and scientific norms on the other hand (Clifton 1975, see 2.2). The music theorists Per Broman and Nora Engebretsen (2007: 11–12) have critically underlined that “music theorists commonly measure their models against scientific standards of logical rigor and terminological precision” while struggling to negotiate “systemic coherence and economy” with the data they have gained.” Furthermore, music theorists have to reconcile “the normative tendencies of scientific theory with the individuating goals of music analytic practice” on a continuum between “objective, verifiable truth uncovered through a scientific enterprise and *ad hoc* subjectivity” (Broman & Engebretsen 2007: 11–12).

A critical position towards the scientific validity of music theory in relation to missed opportunities of a psychology of music based on an examination of the history of music theory was expressed by Dahlstedt (2007: 262) as follows:

> “The terms of music theory do not represent unambiguous physical entities. Rather they are defined in respect to perceptions and sometimes they are used in accordance with a blurred mixture of sensations and perceptions. This is the result of the marriage between old traditions of music education and an urge to base music theory on scientific empiricism and a scientific realism. This should have resulted in a psychology of music that admitted the complete translation of old music terms into the language of science, but it did not. A general problem with this scientific theory of music was the ambition to present law-like propositions of more or less universal relevance.”

Therefore, one should make a distinction between the terms and concepts applied in music perception, cognitive psychology, music theory, and cognitive musical analysis. If perception is seen as rather involuntary or “reflexive” involving non-conscious processes (Cross 1998a: 4, Purwins *et al* 2008b: 170) in contrast to music theorists’ rather conscious and voluntary processes applied while analyzing music (Purwins *et al* 2008b: 170) then an appropriate cognitive-scientific method is able to reveal “those aspects of the perceptual process that are not amenable to conscious introspection” (Cross 1998a: 17). Both referred articles underline the (well-known) mutual criticism in methodological approaches between music theory and cognitive psychology (“banal music-theoretical concepts that lead only to a ‘psychology of ear training’,” Gjerdingen 1999: 165 based on Cook 1994; “focused on too low level, perceptual processes, using non-realistic stimuli,” Honing 2006: 4; “lack of reality”, Gjerdingen 1999: 163). The researcher of music theory and music perception Robert Gjerdingen (1999: 163) describes how a psychologist and a music theorist would read an empirical research article differently: the psychologist would value the ‘results’ section with statistical proofs and hypotheses refutation/support, the music theorist would rather be interested in the narrative ‘discussion’ section that includes speculations based on data results with the aim to carefully double-check rather symbolic musical terms that are maybe used to generalize psychological phenomena.

The music and science researcher Ian Cross (1998a: 17) advocates that the cognitive science of music should not push “folk psychologies” (music theorists’ understandings of how music is perceived and cognized) aside, instead it could “bridge the gap between what music feels like – its experiential texture – and the language that is used to describe it and to teach it.” As a reminder: the basic goal of this dissertation (at the beginning of the Introduction section) is to intertwine music analytical (theoretical) and music psychological (empirical) thinking with the aim to “bridge the gap” (Cross 1998a: 17) between these domains. As a tool to achieve this I have chosen in the dissertation narrativity, that is, indeed (especially in music theory), common in language form to describe (and teach) the “experiential texture” of music (Cross 1998a: 17). However, narrativity can be used methodologically as the level connecting perception with cognition in intuitive listening strategies and listening stories [*Hörgeschichten*] (Gruhn 1992, see publication IV, see 3.5) and furthermore as abstract and structural principle to analyze and interpret empirical data (see publication IV).

The so-called cognitive musical analysis approach by the cognitive musicologist Jean-Marc Chouvel (2014: 17) advocates that “the analytical process must follow as closely as possible the behavior of the musical one. This means that we have to be aware of the cognitive processes at play *when one is listening.*” Chouvel (2014: 17–19) doubts that the inevitably improving accuracy of experimental psychology and neuroscience reveals more of the far greater complexity of musical experience in comparison to the lower complexity of the musical score. He advocates “cognitive analysis” as an immanent musical analysis of the musical phenomenon (“as indeterminate flux of “events” into a collection of structured “objects””) and its “temporal configuration” in real time based on external and internal knowledge (paradigmatic and syntagmatic recognition), similarity, memorisation, protention and retention, “completeness of one of the hypothetical objects,” and recursivity. Several cognitive listening models are introduced in 2.5.

In general, I advocate that music should be analyzed music-theoretically and observed empirically both as complete works (holistic information processing, form analysis, global viewpoint) and as separate structural aspects with its properties and interrelations (analytic information processing, analysis of musical material, local viewpoint) (holistic *vs* analytic approaches see Eitan & Granot 2009: 165, Deliège & Mélen 1997: 387).

Similarly to Cambouropoulos and Tsougras (2009: 119) I advocate that music theoretical and music psychological methodological approaches are different but can inform and support each other, the process of listening is, however, connecting them. Therefore, in this dissertation the principle of attentive (focused) intuitive listening as listening strategy[26] is regarded as the

---

[26] The term phrase 'intentional act of listening' refers to Clifton’s (1975: 70) term phrase *intentional acts* to approach music’s many experiential gestalts. The concept of 'attentive listening' has been used by the cognitive musicologist Irène Deliège and the music psychologist Marc Mélen (1997: 388). The term *listening strategy* was suggested in the doctoral seminars by Prof. Dr. Urve Lippus (Head of the musicology department at the Estonian Academy of Music and Theatre) to characterize the outcome of participants using slider-controllers.

central method. The reasons are the following: (1) In my experiments[27] I let participants listen to whole pieces, not just short chunks. Concentration on brief chunks is clearly possible, but concentrating on listening to a 30 min long full work requires a different kind of listening attitude. (2) Intuitive (in the context of Kant's definition of the cognitive process) and attentive means "active" listening based on knowledge as "the assertion of judgements, which combine these sensible representations in some definite manner or other" (Wolff 1963: 149, see 2.2). I also have applied a systematic, but speculative approach combining a score-based reading method that requires listening as confirmation (see publication I), but later abandoned it because of its speculativeness – it however brought me to the salient feature Impulse that I have applied in my new empirical model COSM: Cognitive Octagonal Slice Model (see publication IV). The next subsection provides the methodological-philosophical underpinnings for the cognitive (1.3) and especially intuitive aspects this dissertation suggests for achieving its goals, aims and objectives.

2.2 Methodological-philosophical matters

In this subsection the relation between philosophy of mind, brain-machine-metaphor and science, and the relation between music/musicology/music theory and science is discussed – building the fundament and actuality this dissertation is embedded in. A cognitive approach is suggested to negotiate between the different realms. In general, one can find two-part or three-part differentiations in philosophy of mind and science. The most famous dichotomy is the so-called mind *vs* body/world problem that was invoked by the philosopher, scientist and mathematician René Descartes (1596–1650) (Mind-body problem APA 2018, Westphal 2016, [2025]). A rather methodological explanation of this two-part view can be found in the APA Dictionary of Psychology entry on Consciousness (APA 2018) as (a) function or behavior (viewed "from the outside" – the observable organism, a reductionist or materialist perspective) vs (b) experience or subjectivity (viewed "from the inside" – the mind, an immaterialist perspective). This dichotomy can be simplified as physicalism/materialism *vs* Kant-based

---

[27] In those publications that include experiments with music (publication I, Ia, and the empirical methodological design in publication IV.

‘transcendental aesthetic’/‘analytic idealism’[28] and was proposed as part of a recent renaissance of Immanuel Kant’s ideas by the philosopher and computer scientist Bernardo Kastrup (2019) (see below).

Cross (1998b: 207 etc.) has proposed the following three-part view: “three loci between which can be positioned many of the views that have been held – and are held – about what music is or can be”: i) the physicalist position – physical facts as a base for sounds/structures as material (“given by nature”) of music utilized and experienced; ii) the immanentist position – music is not existing in physical reality, it's a concept making use of the “meaning-seeking capacity” of humans; iii) the cognitivist position – music as cultural convention, embodiment, “instantiated in the cognitions of the members of a culture” of “more-or-less abstract representations of musical events or phenomena” both stable and changeable via cultural selection. Cross has given assessments concerning possibilities of scientific inquiry of “music” for each position: i) music understood via physical science – seems possible, but remains unfeasible because a claimed “direct, one-to-one correspondence between objects and events in the physical world and our sensations and perceptions” is not existing (Seashore 1938: 382, see next paragraph), acoustical information transformed into sense-data may just reveal “constraints on that experience,” science does not exist as mere physics; ii) “impenetrable by scientific means of enquiry,” – impossible, because (based on Cook 1990) reductionism is “antithetical to the nature of music as experienced” and grounded in a “logical positivist tradition” of science; iii) “susceptible to scientific explanation” – likely possible.

According to the music psychologist Carl Seashore (1938: 382) it is fundamental that “there is not a one-to-one relationship between music as performed and music as experienced. The hearing of music is a response to a stimulus.” This has led to a shift from music as denotative (understood by universal standards – with means of music composition and theory/analysis) towards music as connotative (individual understanding by the listener – with empirical means of music psychology) incorporating the effect of context (familiarity vs unfamiliarity of the stimulus) and the role of cognitive factors (meaning, mood) on the perception of music (O’Briant & Wilbanks 1978: 441–443).

A three-part view is introduced by the cognitive musicologist Mauri Kaipainen and the neuroscience researcher Pia Tikka (2022: 129–131) as follows: 1) Physicalism based on

---

[28] Also called ‘transcendental idealism.’ In this dissertation I use further only the term-phrase *analytic idealism*, because I do not go beyond the basic analytic and synthetic aspects related to perception and cognition.

explanations via the physical world only – in the 1950ies expressed as mind-as-machine/brain-as-machine metaphor; since the 1980ies changed into the machine-like-brain metaphor, developments towards artificial neural networks and deep learning. 2) The existence of a reality outside the physical world stands in contrast to physicalism. 3) A holistic view requires a biological ground for technological applications.[29]

My standpoint to this three-part view is the following: 1) I do not deny powerful supportive possibilities of music analysis created in the physicalist machine-like-brain metaphor context, but they are still limited to certain aims and deal with data and decoding problems.[30] I do not think that they provide satisfying solutions for the objective and the Dilemmas (see publication V) I deal with in this dissertation. 2) I share the "reality outside the physical world" view in the sense that my objects of research, music and tension, are cognitive.[31] they take place in the mind, they are not sufficiently comprehensible applying only physicalist-materialist explanations. 3) The holistic view is outside the scope of this dissertation. The first view is supported by Integrative Musicology (Dahan 2012: 3) suggesting that a so-called representational barrier seems nowadays crossable via combining cognitive science, digital signal processing, hardware and software engineering – only a clear musicological agenda is needed to coordinate and realize this development. The second view is supported by the problem of the so-called *explanatory* or *semantic gap* faced within the MIR (Music Information Retrieval) community that cannot be overcome, because music is cognitive (Wiggins 2009) (see next paragraph).

The understanding of music theory as a scientific method has been criticized by the musicologist Sten Dahlstedt (2007: 254) underlining the fundamental difference of both:

> "Certainly music theory takes music's physical, physiological, and other conditions into consideration, but that still does not make music theory a science. While science relies on empirical information and on a mind-dependent theoretical apparatus with high general acceptability, music theory is the theory of an art form." (Dahlstedt 2007: 254)

---

[29] See actual overviews on topics in all three approaches in "The Science-Music Borderlands. Reckoning with the Past and Imagining the Future" (Margulis, Loui & Loughridge 2023).

[30] See a methodological classification of computer-assisted music analysis (Schüler 2000a&b), "How to Think Music with Data. Translating from Audio Content Analysis to Music Analysis" (Andersen 2017) and "Diginalysis: The Man-Machine Collaboration in Music Analysis" (Onwuegbuna 2020).

[31] See an overview on musical grammars and music cognition developments in the 1980s and 1990s (Schüler 2007).

“Music” can be understood as “a middle term between stimuli (sounds) and responses (listening)” [...] “and music is where the mind is” (McLarty, Braun & Benitez 1990: 500–501) advocating a participatory view: “the listener is a part of the system being listened to” (McLarty, Braun & Benitez 1990: 498).

An explicitly cognitive approach has been put forward by the researcher of music cognition and computational creativity Geraint Wiggins (2009: 477):

> “*Music*, as opposed to *Sound*, cannot be effectively studied from the standpoint of pure audio engineering – nor, indeed, from that of pure music theory. We argue that the cognitive mechanisms involved in human music perception and cognition must be taken into consideration, for a realistic account to be given.” [italics and capitalized M in Music and S in Sound by Wiggins]

In this dissertation I take (as mentioned in 1.3, see publication V) a cognitivist position and understand music based on listening as cognitive (Cross 1998b, Wiggins 2009). Furthermore, I hold with the analytic idealism of Immanuel Kant’s (1724–1804) *Critique of Pure Reason*.[32] Analytic idealism assumes synthetic *a priori* cognition based on mathematical means (positive viewpoint)[33] and limiting human cognition to sensible appearances admitting that humans cannot cognize things-in-themselves (*noumena*, negative viewpoint) (Kant 1998: 8). The influence of Kant’s thinking on Western philosophy and science is undoubted both affirming and contravening it.[34]

In my view Kant’s understanding is often rather implicitly applied by researchers (e.g. Clifton 1975) without referring to basic sources by Kant. Such an influence of the importance of mind and the human (observer) viewpoint can be found e.g. in publications by the computer scientist and mathematician Steven Wolfram (2002) (see also Wolfram 2020, 2021, 2023) on science and fundamental theory of physics. Another example is the Integrated Information Theory (IIT) of Consciousness (e.g. Tononi *et al* 2016, Hendren *et al* 2024, ITT Wiki [2024], critically analyzed by Pautz 2019) that incorporates concepts like immediacy, intrinsicality, introspection

---

[32] First edition 1781, second edition 1778, English translation by Paul Guyer and Allen Wood (Kant 1998). Although I could read Kant in German, I deliberately have chosen the English translation, because 1) my dissertation is mainly in English, 2) thinking in different languages is enriching and a mediation via translation can support the development of understanding of complex topics (Lock 2022).

[33] E.g. the concept of inner and outer sensation/magnitudes proposed as early psychophysical measurement principles (Kraus 2013/2016).

[34] Pointed out by the translators and editors of Kant’s *Critique of Pure Reason* Paul Guyer and Allen Wood (Kant 1998: 22–23).

and existence[35] reminding clearly of Kant's understanding of the unity of apperception[36] (Kant 1998: 9 etc) as well as the immediacy of cognition and intuition in relation to empirical sensibility (e.g. Kant 1998: 173).

Despite of controversies around Kant's understanding of *analytic* or *transcendental idealism*[37] I shortly summarize the terms and concepts[38] that in my view are relevant as a methodological-philosophical background for this dissertation. According to Kant (1998: 155, 172) intuition takes place immediately when a cognition is related to objects while "all thought as a means is directed as an end."[39] Objects are given[40] via sensibility but affect the mind via representations.[41] The capacity or receptivity for acquiring these representations is called sensibility, this enables intuitions. Intuitions are "thought through understanding, and from it arise concepts."[42] All (direct and indirect) thoughts (via certain marks[43]) have to be connected ultimately to intuitions triggered by sensibility, "since there is no other way in which objects can be given to us." The influence an object has on the capacity/receptivity for representation works via sensation. "That intuition which is related to the object through sensation is called empirical. The undetermined object of empirical intuition is called appearance." Matter are elements in appearance corresponding to sensation. The form of appearance emerges when the manifold of appearances is intuited while being "ordered in certain relations." The matter of all appearances is given *a posteriori*, form must be existing in the mind *a priori*. Pure (transcendental) representations exist without connection to sensation. The pure form of sensibility is pure intuition. While separating the thoughts (understanding) (substance, force,

---

[35] Existence (0th axiom): Experience *exists*: there is *something* – Existence (0th postulate): The substrate of consciousness can be characterized operationally by *cause–effect power*: its units must *take and make a difference*. Intrinsicality: Experience is intrinsic: it exists for itself – Intrinsicality: Its cause–effect power must be *intrinsic*: it must take and make a difference *within itself*. (See ITT Wiki 2025.)

[36] Of course, the translators and editors admit that "what the exact relation between apperception and the representation of objects is, are obscure and controversial, and continue to generate lively philosophical discussion even after two centuries of interpretation." (Kant 1998: 9).

[37] "Kant's readers have wondered, and debated, what exactly transcendental idealism is, and have developed quite different interpretations" (Stang 2024).

[38] From Kant's (1998: 173) second edition of the *Critique of Pure Reason* (1787).

[39] I interpret this as belonging to the realm of narrativity.

[40] Things-in-itself (*noumena*) cannot be perceived, only representations can be intuited and apprehended (Kant 1998).

[41] Representations in Cognitive Science (Shea 2018) are applied to comprehend the behaviour of organisms as well as artefacts like computers and control devices. „When a scientific explanation points to representational content to explain behaviour, we need to get inside that explanation to see how it works" (Shea 2018: 28).

[42] The distinction between concept and intuition is of the utmost importance for understanding Kant's critical philosophy" (Smit 2000: 235).

[43] "Intuitive marks are singular instances of properties, as they are represented in, and make up the content of, our intuitions. Discursive marks, in contrast, are general properties as they are represented in, and make up the content of, our concepts" (Smit 2000: 266).

divisibility) and sensation (impenetrability, hardness, color) from the representation of a body leads to empirical intuition. However, extension and form of such a representation of a body remains to be pure intuition *a priori*, especially when excluding the actual object from the senses/sensation, being "a mere form of sensibility in the mind." Kant (1998: 248–249) underlines that

> "the supreme principle of all intuition in relation to the understanding is that all the manifold of intuition stand under conditions of the original synthetic unity of apperception. All the manifold representations of intuition stand under the first principle insofar as they are given to us, and under the second insofar as they must be capable of being combined in one consciousness; for without that nothing could be thought or cognized through them, since the given representations would not have in common the act of apperception, I think, and thereby would not be grasped together in a self-consciousness." (Bold words given by Kant – G.L.)

In this dissertation the object of *immediate intuition* is *music*. Music provides the manifold of intuition. After being composed, it is given to sensibility through listening via performance/interpretation affording empirical intuition and the active combining and synthesizing process. The listeners apply concepts to understand the music via their thoughts (attentive listening, listening strategy). The phenomenon in focus of this dissertation is musical tension. Musical tension is approached structurally via musical form (global level) and musical parameters (local level) (see publications I, Ia, IV). Kant's "intuitive marks" as "singular instances of properties making up the content of intuition" (Smit 2000: 235) are in my view *Impulses* ("moments of change"). Kant's "discursive marks" as "general properties making up the content of concepts" (Smit 2000: 235) are in my view *salience* (see 3.2, 3.3) and *musical parameters* (see 3.4). In the following subsection on temporal dynamic cognitive processes and *microgenesis* accounts of the above presented cognitive, *analytic-idealism*- and and intuition-based understanding are given in the visual and the sound/music realms.

2.3 Temporal dynamic cognitive processes and *microgenesis*

Temporal dynamic cognitive processes are explained as a first-person account in the concept of ‘microgenesis’ by the psychologist Talis Bachmann (2000). As a term it appeared in the 1950ies to explain the German term *Aktualgenese* (having roots in the 19th century, Bachmann 2000: 23). It has become part of the ongoing consciousness research (Aru & Bachmann 2017).

Microgenesis aims to explain the *immediate experience* of short now-moments via temporal dynamic unfolding and differentiation into meaningful objects that break up “the continuous fabric of reality into individuated forms.”[44] In the arts and reading realms *microgenesis* claims that *genetic* (the developmental dynamics of a process) “segmentation of the perceptual field into individual objects is thus the result of perceptual differentiation, and not the objective state of affairs that perception would merely seek to detect and acknowledge” (Rosenthal 2004: 221, 224). Primary categorization in visual cognition needs completion via local discrimination (Rosenthal 2004: 224). For visual cognition the philosopher of science Arturo Carsetti (2004b: 314) underlines that in self-organizing emergence processes of assimilation, growth, and stabilization and reduction via fixed points take place simultaneously. He underlines that “function and meaning articulate together, but in accordance with the development of a process of *adequatio*, and not of autonomous and direct creation” claiming that one cannot think of vision during the emergence process “but will be able to use it, once realized, to construct further simulation models” (Carsetti 2004b: 314). He admits that these processes are not easy to describe mathematically and it is impossible to understand the processes of self-reflection and assimilation within the mechanistic reductionism paradigm. He suggests that internal and external selection processes are deeply creative (Carsetti 2004b: 314). The concept of ‘genetic’ reminds of Kant’s (1998) *synthetic unity* of *apperception*, the perceptual differentiation via segmentation of the perceptual field into individual meaningful objects (forms) belongs to Kantian analytical concepts. Perception is not just a process of detection and analysis, instead it means *active synthesizing* by the perceiver/cognizer.

Bachmann (2000: 1) has shown with the concept of ‘microgenesis,’ based on psychophysical experiments (especially in subjective experience, see more on psychophysical measurement

---

[44] Compare these with the notion of qualitatively distinct forms proposed in linguistics (Browman & Goldstein 1990: 413).

below), that a sometimes claimed "directness"[45] in perception and cognition is misleading: there is neither immediacy of perception as automatic nor simultaneity between the phenomenal percept and the depicted/presented: "with most of the known realities, perception does not appear, *voilâ*, within the infinitely short slice of time, but it takes time for it to manifest" (Bachmann 2000: 1). Furthermore, Bachmann (2000: 1) underlines that "the object or brief physical event, or what in the tradition of psychological science could be called *the stimulus* (stimulus event), is not directly represented in the mental experience of the perceiver. Bachmann (2000: 1–11) underlines that there is no immediate instantaneity; there is a latency time (around 50–250 milliseconds); "prefix states"/representation of perceptual "proto-objects" appear (see more below); the latency time "is itself a psychological event extended in time" (see also Jensen 1979, Libet 2004, Canales 2010); brain research shows time developments of the conscious percept emerging between 300–500 milliseconds. On latency times and the appearance of meaning in music perception and cognition see Seashore (1938), Snyder (2000), Koelsch and Siebel (2005), Koelsch (2013a & b).

In music one can speak of *emergent qualities* (*properties*) similar to the measurement of temperature or humidity "whose flux can be sensed approximately but also computed precisely" (Mailman 2010: vi). The music theorist, researcher and composer Joshua Mailman underlines that "contrary to the assumption that there is a small fixed set of quantitative dimensions in music, [his] dissertation shows there are innumerable quantifiable qualities whose flux of intensity bears form in music," the latter he calls *temporal dynamic forms* (Mailman 2010: vii–viii). He investigates the role of emergent properties in building temporal dynamic form, he presents the Discreteness-Continuity Salience Question (Mailman 2010: 119) (discussing its relation to the dynamic-*vs*-static dichotomy) asking whether discrete (abrupt, crisp, categorial) or continuous (smooth, gradual, statistical) change is more salient in our consciousness. He deals with this question in further chapters in relation to "statistical change and the advancing continuum of nowness" proposing a Continually Updated Continuum of Nowness (CUCoN)

---

[45] The discussion about perception as "direct" or indirect (directness negating the need for representation in the perceptual system, e.g. generally and for vision see Michaels & Carello 1981, Warren 2005, for music see Clarke 2005) will not be elaborated here further, because in my opinion perception is never direct in the literary sense due to biochemical processes causing delays in the perceptual system. A reason for believing in whatever directness can be seen also in concepts of "naïve realism" or "direct realism" or "common sense realism" that are in opposition to scientific knowledge acquisition (Hoyningen-Huene 2013: 194–195). Furthermore, as mentioned above, there is "no one-to-one relationship between music as performed and music as experienced" (Seashore 1938: 382). However, the theory of "direct" perception belongs to the so-called Ecological Approach of perception. In my opinion the notion of "direct" has its roots in Kant's immediate pure intuition of concepts via thoughts and the ideas that one can perceive the world only as representations (see previous subsection).

principle (Mailman 2010: 235)[46] and to “emergent properties and contour smoothing” (Mailman 2010: 251–363). The results of a perceptual (empirical) study on dynamic forms as “subjective perception of temporal events in music (*explosive*, *fading out*, *rising* etc)” indicate “that subjects are sensitive to dynamical forms, but were particularly sensitive to a specific one (*suspense*)” or a state of tension based on its higher salience (Hjortkjær & Nielbo 2010: 1) lead us to the concept of *vitality forms* (referred to by the authors, too) – on salience see more in 3.2, 3.3.

Dynamic forms of vitality as the experience of the act of movement of humans have been observed by the psychologist and medical doctor Daniel Stern (2010: 4–5),[47] they create the impression of a holistic, therefore continuously perceived world. He grounds his theory and observations since the 1980ies in the merged experience of five aspects of *vitality* (dynamic events: theoretically different), the “fundamental dynamic pentad“ of movement, force, time, space, and intention/directionality that especially works also in the perception of the arts (Stern 2010: 4–5). Stern understands the *Gestalt* or “emergent property” as “the most useful concept for dealing with holistic experience” (Stern 2010: 5). The same is underlined also by Carsetti (2004a: 12) saying that “gestalt is not only a mathematical or computational construction: it is something that lives, of which we have direct and holistic experience.” For Stern (2010: 20–21) the mind functions via embodiment and “movement is everywhere” appearing also mentally in imagination and virtual body movement. Stern refers to Ernst Kurth’s notion of “sound in motion,” we can add Eduard Hanslick’s (1854) even earlier understanding of music as “sounding moving forms” (German *tönend bewegte Formen*). In order to understand movement and motion, Stern (2010: 20) underlines that we need to ask in addition to the *what* (see 2.5) and *why* (see 3.7) also the *how*-question (see 2.5). Especially the latter is observable in temporal dynamics, e.g. in curves and trajectories (profiles) of *vitality forms* that have been also graphically visualized under simplified conditions (Stern 2010: 8), looking similar to simplified tension curves. The possibility of the recognition of vitality forms has been investigated in neurological research observing participants’ real gestures (Di Cesare *et al* 2014, including Stern as one of the authors).

---

[46] “It should be clear by now that besides being a series of “nows,” our temporal consciousness may consist of the aggregate of recent “nows” we contemplate. Change may be observed within a short “now” or an inclusive longer “now”” (Mailman 2010: 236). See also Husserl's concept of loops of dynamical *retention* (experience fading into the past) and *protention* (implied next-moment-anticipation) within which the continuously moving “now” is located (explained in music context in Mailman 2010: 233), used also to explain the experiential moment of “narrative nowness” in neurocinematic experiments (Kauttonen, Kaipainen & Tikka 2014: 79).

[47] Based on his research on newborns and infants since the 1980ies. Later, Giuseppe Di Cesare *et al* (2014: 951) have used fMRI to study the neural correlates of vitality form recognition based on what-how-tasks that revealed that the observer is able to understand the vitality forms conveyed by the observed action.

Stern (2010: 5) calls the step from vitality to its gestalt-based perception “as mysterious as the appearance of an emergent property” while asking: “do we learn that these five events belong together or is the mind designed to fuse them?” bringing out here the ongoing struggle in science (e. g. dynamic systems theory, complexity theory, and chaos theory). Again, Stern’s and Carsetti’s understandings of directness, holistic experience and the fundamental dynamic pentad clearly remind of Kant’s (1998) unity of perception into one consciousness and Clifton’s (1975) immediate perception of music without the need of scientific explanation.

Methodologically, Bachmann (2000: 11) advocates that the “perceiver as a subject in experimental-psychological studies becomes an “instrument” of observation and measurement himself (herself)” (see more in 2.4, see also 3.2). Furthermore, Bachmann (2000: 12) underlines that “regardless of the methodological complications, if the assumptions about the validity of first-person description for third person purposes have been accepted and apt methodologies worked out, several topics can be and have been investigated in microgenetic research.” While focusing on the analysis of the music, the data obtained from empirical experiments in this dissertation will neither be used to analyze the participants' (psychological, emotional) responses from a behavioristic point of view nor qualitatively/subjectively (observing each participant’s strategies separately).

In this dissertation I treat sound and music as private sensations (O’Callaghan 2021). This is supported also by the following claims by Wiggins (2009: 479) saying that the “essence of Music” appears in perception and “not sound *per se*,” and “the music experienced by a listener *can only* be a perceptual or cognitive construct of their own making.”[48] Sound and music are experienced, perceived and cognized as a continuum and as events, not as (“static”) objects (Casati, Dokic & Di Bona 2020). However, the Kantian ‘object’ introduced in the previous subsection is a wider and more general concept embracing everything given and appearing to the senses and sensibility both to the outer and inner senses and their measurability (Kraus 2013/2016, Kant 1998).

Although *microgenesis* does not advocate “direct” perception (see above), introspection is still sometimes understood as “direct,” its difference from a behavioristic understanding is rarely considered (Shepard 1978: 445) and often unreflectively mixed up (Savage 1966, even in actual visual psychophysical methodology, e.g. Signal Detection Theory (SDT), Kingdom & Prins

---

[48] The term “own” points to the first-person (so-called subjective) perspective until recently underrated in comparison to the third-person or so-called objective perspective and methods traditional science has been ruled by generally (Valenzuela-Moguillansky & Demšar 2023).

2016). This contradiction in relation to “direct” understanding has been discussed in auditory perception categorizing sounds as either private (sensations, “direct”) or public (pressure waves causing auditory experiences) (O’Callaghan 2021). It has appeared also in the context of psychophysical measurement (see below) critically addressed by Clarence Wade Savage (1966: 1): “The psychological magnitudes involved in perception are observable but private according to the introspectionist, public and nonobservable according to the behaviorist.” The private *vs* public distinction leads further to the methodological dichotomy of intuitive (“inner”) *vs* scientific (“outer”) description of music (Clifton 1975: 70) whereas the “inner” is conventionally treated as subjective (requiring first-person methodology), the “outer” as objective description (Clifton 1975: 72) approached with third-person methods. The “inner” *vs* “outer” dichotomy refers to Kant’s concept of objective cognition based on inner and outer sensation, intensive (inner sense) and extensive (outer sense) magnitudes (Kraus 2013/2016).

According to the researcher of the history of philosophy Katharina Kraus (2013/2016: 350) Kant’s idea of objective cognition is based on “formal determination of experience in accordance with the transcendental principles”[49] and “by a material condition, which requires something mind-independent that affects the senses.” Kraus (2013/2016: 350) suggests that “there is a transcendental relation between intensive and extensive magnitudes, which makes possible the numerical quantification of intensities. Yet a fully fledged objectivity of such claims can only be guaranteed by external, intersubjectively shared standards that have to be spatial.”[50] Kraus (2013/2016: 350) interprets this as an alternative conception of psychology that comes close to later psychophysics and psychophysical measurement. Kraus understands her analysis of quantifying inner experience based on Kant’s mathematical principles – intensive magnitudes as degree-based and extensive magnitudes as additive successions – in Kant’s (1998) *Critique of Pure Reason* (1781[1], 1787[2]) as “a first step towards a transcendental foundation of psychology on Kantian grounds” (Kraus 2013/2016: 351, see also publication V). The following subsection introduces basic aspects of psychophysical measurement and meaningfulness.

---

[49] “The Analogies of Experience, which correspond to the three schematized categories of substantial permanence, causal sequence, and communal interaction, are rather like a proportion in which there is no *a priori* method of determining the fourth proportional, as there is in the case of the mathematical categories. They only tell us, for example, that *something* stands to this thing as cause stands to effect, but what this something is experience must reveal.” (Findlay 1981/2011)

[50] This claim seems to be confirmed by clinical research: “The current research proposes that prothetic and metathetic continua are represented in the brain in a fashion similar to that of spatial and object representations” (Jewell 2005: 5). See more on *prothetic* (how much, intensity, e.g. dynamics) and *metathetic* (what or where, e.g. pitch) *continua* in 3.4.

2.4 *Psychophysical measurement* and meaningfulness

Scientific *psychophysical measurement* is a challenge around the following question: How to measure the not directly observable sensation that intervenes between physical stimuli and behavioral responses? (Steingrimsson 2016: 205). Human-conducted detection and measurement procedures have become a part of science since the second half of the 19th century when psychophysicists started to try to comprehend empirically (testing "just noticeable differences" (JND's) between the magnitudes of perceived/cognized phenomena) and via mathematical modeling/functions how sensations elicited by physical phenomena and their intensities (magnitudes) can be captured as phenomenal judgments in experimental phenomenology as a psychophysics application (Burro 2016: 9).

The most widespread definition of *psychophysical measurement* is taken from the psychologist Stanley Smith Stevens' (Stevens 1951, Michell 2020) meaning "the assignment of numerals to objects or events according to rules" using different scale classes (nominal, ordinal, interval, ratio scales) (Steingrimsson 2016: 205). The mathematical foundation-based Representational Theory of Measurement (RTM) defines psychophysical measurement as

> "mapping between two relational structures, an empirical one and a numerical one. For a given empirical system to be measurable, it needs to fulfill the properties of a formal system. In practical terms, the approach consists of formulating qualitative (testable) axioms involving ordering [...] that together imply a numerical, vectorial, or geometric numerical representation" (Steingrimsson 2016: 206).

The social science philosopher Elina Vessonen (2021) claims that measurement equals 'representation' and it is sufficient to base it on minimum mirroring between the empirical and the numerical (mathematical or visual/geometrical) system (Representational Minimalism, ReM, based on Representational Theory of Measurement, RTM). On the contrary, if acknowledging (in Patrick Suppes' version of representational theory) that "mathematical structure is already *embedded* in empirical systems," one should also take into account that measurement deals only with attributes (of objects and events), that means, according to the psychologist Joel Michell (2020: 3), that "empirical systems sustaining measurement already instantiate positive real numbers. Thus, in measurement, real numbers are estimated, not assigned from without. Representational theory not only misrepresents measurement; it refutes itself." Read more about this debate (e.g. Heilmann 2015, Humphry 2017, Michell 2017, van Fraasen 2018, Teller 2018).

Both psychophysical and instrument-based physical measurement are neither direct nor accurate. In both one must deal with errors (for psychophysical measurement see Shepard 1978, Levitin 2002: 126–127, Hohwy 2012,[51] Kingdom & Prins 2016; for physical measurement see Bentley 2005). Of course, physical measurement instruments have become more and more accurate through laser and other technology applicable in more and more fine-grained dimensions. But even the accuracy of measurement of physical quantities (e.g. mass and temperature, but also multivariable quantities) is still a challenge and (on a grounding level) even basic units and values of physical measurement as well as their definitions and relations among each other are in an ongoing process of concretization and unification (CGPM R1 2018, Davis 2019, 2023). The comparison between psychological and physical measurement and psychophysical scaling principles show similarities and differences. However, no "direct" measurement is possible for either one (Shepard 1978: 445), equivalence between both is a rather technical concept (Narens & Mausfeld 1992: 472), that anyway transmit some empirical significance (Shepard 1978: 485).

Another problem especially relevant in psychophysical measurement is posed by the statistician Paul Velleman and the statistician and computer scientist Leland Wilkinson (1993): one often only implicitly assumes the identification of relevant attributes, the establishment of relevant questions as well as understanding of attribution rules for measurement of empirical phenomena: "However, in real-world data analyses we often do not know the attributes we wish to measure or what questions we intend to ask. New questions may arise from the initial analyses or from external influences" (Velleman & Wilkinson 1993: 14). Relying on theories can be problematic, too: "Ostensibly scientific descriptions refer to concrete entities and quantities in nature, but what are the referents if those descriptions can only be understood within their theoretical context?" (van Fraassen 2018: 271).

In this dissertation I rely on quantitative psychophysical measurement principles, admitting that they are estimations and instantiations (Michell 2020). Measurement principles are a methodological background for the detection process in experimental psychology (Burro 2016) and philosophy-of-consciousness research methods (introspective Signal Detection, iSDT,

---

[51] Philosophically one can say that the brain is constantly testing hypotheses based on attention and conscious perception using prediction error minimizing (Hohwy 2012: 2).

Theory Morales 2018, 2024). The current state of psychophysical methods of measurement of auditory (e.g. based on sound production) and especially cross-model stimuli is described by the cognitive researcher Ragnar Steingrimsson (2016) and the experimental psychologist Jürgen Heller (2021). Indeed, psychophysical measurement is *per se* cross-modal and cognitive, because "measuring" using assignation of numbers to some continuum involves knowledge and functions as follows: matching the magnitudes of one continuum (music, unprimed) with those of another continuum (numbers, primed, learned proportion) to derive an "ordinal" scale, but one cannot tell *how much* greater/smaller the sensation is (Shepard 1978: 470, 480–481). However, "these operations do not in themselves permit us to measure inner sensations in any quantitative sense" […] "what the psychophysicist measures is an important constant governing how a subject transduces any stimulus from a particular sensory continuum, such as the continuum of lights varying in intensity, tones varying in amplitude, or lines varying in length" (Shepard 1978: 481).

In science and analysis in general, and particularly in measuring as well as in representation, visualization and modeling one looks for *meaningfulness* based on either presumed (first of all qualitative) "primitives" that are mathematically definable or seeks for invariants (to take the role as primitives) that are preserved through mathematical-statistical transformations (Narens 2001: 2). In both *physical* and *psychophysical measurement* procedures the aspect of stability (unchangeable elements) is important to gain meaningfulness: in physical measurement they are constants, in psychophysical measurement they are called invariants (Narens & Luce 1987/2016, 1990, 2008/2017). Invariants are explained as "patterns of stimulation over time and/or space that are left unchanged by certain transformations" either as "perceptual constancy," related to the laws of physics, and quantified and described mathematically (Michaels & Carello 1981: 20).

Concerning invariants in music and perceptual processes within the Ecological Approach of perception the music psychologist Eric Clarke (2005: 35) explains that music "offers a particularly clear example of invariance in the perceived identity of material under transposition and other kinds of transformations" taking as examples themes or motifs, style specific elements or harmonic invariants (e.g. tonic/dominant alternation); he also makes clear the role of implicit learning through attunement of the listener "through exposure to a particular repertoire, whether that exposure is accompanied by direct instruction or not" (Clarke 2005: 35). Clarke also mentions that "these higher order invariants are no more abstract than the most specific and local invariant that is unique to one particular context, even if some of them may be more extended in time: in every case the invariant is a set of relationships that is available in the stimulus informations" (Clarke 2005: 35–36).

To achieve meaningfulness in a real-world context is also the aim of modeling (including representation and visualization). Theories are important back-bones in modeling as a non-linguistic concept via generalizations and simplifications (publication III). The weaker versions of the linguistic and non-linguistic concepts complement each other (Hendry & Psillos 2007) and modeling should retain the connection to the "real" phenomena. This is expressed by the philosophers of science Robin Hendry and Stathis Psillos (2007: 140–141, based also on other authors) saying that models themselves do not say anything about the world, theoretical hypotheses "tying free-floating structures to the empirical world" as "bridge principles which give the theory whatever empirical content it has [...]." Hendry and Psillos (2007: 142) underline that "in so far as the theory embodies a representational relationship between model and world, it must reach out beyond the model to the world itself." Meaningfulness can be explicit (verbalized, linguistically theorizable) or implicit (Nickols 2001/2010). In the implicit case a listener perceiving/cognizing music intuitively has often no need (and see no means) to explain one's experience verbally (see Clifton 1975, immediate intuitive experience, see Kant 1998). However, a researcher must be able to verbalize, e.g. give reasons for generalizations and simplifications for the choice of relevant attributes in analysis and especially in modeling (see publication III). Furthermore, modeling can be understood as working in the "cognitive space of ideas" and can be understood as "storytelling" (Rothbart 2004, xi, see publications IV, V). Storytelling as creating a narrative is conducted by attentive listeners when discovering and applying their individual listening strategies. In order to understand musical tension based on the salience of musical parameters via the salience (perceived "relative change") of musical events as Impulses/"moments of change," Impulse as an analytical and empirical concept is the only salient feature (taken from publications I, Ia, see Appendix C) that is more or less clearly measurable, see the scientific Dilemmas in publication V.

The empirical methods applied in this dissertation (see section 4) are actually based on psychological/psychophysical measurement principles[52] conducted by the participants of the experiments while using slider-controllers and the cross-domain audiovisual analysis model (COSM) developed for this dissertation. Because the cognitive comprehension of sound and

---

[52] Compare with Introspective Signal Detection Theory (iSDT) (Morales 2018, 2024).

music works via listening, the next subsection presents some cognitive dynamic listening models.

2.5 Cognitive dynamic listening models

Music cognition and musical structures have been approached using several kinds of models. Kaipainen's (1994) dynamic theory of musical knowledge ecology describes musical experience based on "the tension between anticipation and its actualization as either confirmation or denial, following Leonard B. Meyer" (Kaipainen 1994: abstract). It works with the interaction of knowledge-acquisition and knowledge-use and takes into account expectation/anticipation, recognition as well as autocommunication in music as cognition, dynamics and holism. As a computational cognitive-dynamics-of-the-mind model of an "ideal" human listener it deals with the mutual relationship of 'knowing-what' (KWhat – momentary musical situation: listening) and 'knowing-how' (KHow – processing perceptual, cognitive or motor aspects: acting, reconstructing, generalizing) in an environment/music.[53]

Kaipainen (1994: 3, 15–16) assumes that music processed by the human mind works always via_cognitive involvement. But instead of propositions, logical deductions or syntax it is based on an individual's "holistic experience that emerges from the world through our own intentional actions." Kaipainen (1994: 3) explores the phenomenon of music "as cognitive dynamics of the individual mind" taking as the central point of discussion the usefulness of knowledge (including aspects of acquisition and use) gained through "sequential continuity of thought and action" based on "constructive perception in conditions of continual variation, distortion and noise." If KHow dominates (over KWhat: input via the environment/music), the model works on itself even without external input. If the model does not self-actualize (KHow disabled), it becomes a "passive pattern-recognizer" of the (full fidelity of the) KWhat stream of input. According to Kaipainen (1994: 3) the most ideal human music-cognizer can be modeled between these extremes: while ignoring deviations from what the model interprets as intended, it still stays open to the environmental input. Similar to Kaipainen's notion of "interpretation as

---

[53] In addition to the "what" and "why" also the "how" in relation to motion and movement is needed for the mind, see Stern (2010: 20). Also in the context of the ITPRA theory (Huron 2006: 16) the "how" plays a crucial role.

intended"[54] there is a similarity to the cognitive scientist Jakob Hohwy's (2012: 2–3) understanding of how the brain deals with *prediction error* and precision while constantly attempting to represent the world. Hohwy's internal model stays either constant (Kaipainen's KWhat input disabled) or it is generative in changing its state and parameters in response to *prediction error* (intention not fulfilled) (KWhat and KHow aspects are mutually at work). The dissolving of so-called errors, faults, accidents and related (general psychological) tension (triggered in listeners and performers based on cognition and previous experience) into creative opportunities is practiced and researched in improvisation, composition and performance art in the recent new realm of *serendipity* (Lock & Sikk 2022).

Pearce's (2005–2018) IdyOM: Information Dynamics of Music model is a statistical model of pitch expectation. The music psychologists and computer scientists Geraint Wiggins, Marcus Pearce and Daniel Müllensiefen (2012: 383 etc) claim that via unsupervised machine-learning, a computer can create music. But they take only tonal melody into consideration, because music in its totality is too complex. However, they start with a perceptual model simulating listening that deals with musical phrase segmentation. They also take into account *descriptive* ("modeling *what a phenomenon does*") and *explanatory* ("modeling *how it does it*") principles of modeling. As a third concept they use the so-called *meta-model* (meta meaning after or beyond) as "a model intended for and validated with respect to a particular (cognitive) phenomenon, which is then (directly or indirectly) able to predict the behaviour of another related but different phenomenon for which it was neither intended nor designed" (Wiggins, Pearce & Müllensiefen 2012: 385). The authors provide an overview of non-computational and computational cognitive models of music perception as well as an introduction of computational cognitive models of musical composition. What follows is the presentation of the IDyOM model based on melody perception and its application as a simple computational model of musical creativity, together with its further system development. They finally conclude that the composed melodies were stylistically just acceptable; more capable grammars are required for creating valid chorale melodies. The model performed at a high level for melody segmentation meaning that "the model seems to capture a more general perceptual mechanism and we have, therefore, designated it a *meta-model*" (Wiggins, Pearce & Müllensiefen 2012: 415).

---

[54] Kaipainen (1994: 3) understands (based on Leonard B. Meyer) musical experience as "tension between anticipation and its actualization" as either confirmed or denied. This works at the local level (micro dimension), because Kaipainen has rejected in his definition global-level (macro dimensional) propositions, logical deductions or syntax to be relevant for the model. However, this understanding leads further to definitions of general tension and suspense (Lehne & Koelsch 2015) and musical tension (e.g. ITPRA by Huron 2006).

The computational and interactive composing model creating avangard-style music called SMuSe (Situated Music Server) by the computational and cognitive music researcher Sylvain Le Groux (2011) “is grounded on principles of modern cognitive science, provides perceptual control of sound synthesis and includes emotional feedback” (Le Groux 2011: v); particularly musical tension has been used in his experiments as a reward function to real-time modulating three musical parameters (tempo, articulation, dynamics) in the Reinforcement Learning (LR) framework (Le Groux 2011: 198–204, see also Le Groux & Verschure 2010). The computational music researcher Chih-Fang Huang (2020) has proposed a musical composing algorithm that operates by changing musical tension energy based on analyzing chord progressions, pitch and rhythm. A spectral centroid analysis of energy in sound in relation to dynamic acoustic salience for motor responses see Schultz, Brown and Kotz (2021). Energy from a music theoretical point of view[55] has been analyzed also in publication Ia and in Kotta (2013) based on Tüür’s music. Wiggins (2010) has also proposed a cognitive model exploring Deliège's (1987, 1996, 2001, 2007) idea of *cue abstraction* using information-dynamic modeling of musical behavior (Wiggins 2010) in connection with paradigmatic analysis (Nattiez 1975,[56] after Ruwet 1972). The *cue-abstraction* mechanism has been hypothesized and applied by Deliège since the end of the 1980ies departing from testing GTTM's grouping structures (Lerdahl & Jackendoff 1983) empirically (Deliège, Mélen, Stammers & Cross 1996: 122) while defining a *cue* as a “very restricted entity [...] often shorter than the group itself, but always embodying striking attributes” (Deliège 1989: 214). The *cue* is implicitly detectable based on sameness vs difference via the perception of relatedness and groupings of the musical structure “as long as an invariant potential is perceived; the process stops when a contrasting structure appears” and a new *cue* appears (Deliège 2007: 15). It was empirically tested on post-tonal as well as Wagner’s music with the instruction “to listen to these pieces as if they were a story, or discourse [and] to note the punctuation that they heard: commas, periods, next paragraphs, etc by pressing a key on the computer keyboard” (Deliège 2007: 15). This is clearly a narrative approach (see 3.5). For me the earlier (and tonal music related) *cue* concept is too blurry as a phenomenon and too much related to formal principles in classical music, such as motifs. The later *cue* concept applies the principle of invariance and is therefore relatable to psychophysical measurement (see 2.4); furthermore it is tested on post-tonal music. However,

---

[55] See also the energetics and phenomenology approaches from the first half of the 20th century, e.g. Ernst Kurth, Hans Mersmann, Andrzej Dobrowolski or Erhart Karkoschka in relation to musical form, musical tension and force (Rothfarb 2002, Kirschbaum 2001, Stratilková 2015).

[56] A short review in English see Nattiez and Dunsby (1977).

such *cues* are durational phenomena related to memory. I will not expand in this dissertation on memory aspects and in my view it is too complicated to define the borders of *cues*, similarly to the complicatedness of defining tension as a durational phenomenon in its start/ending (see the five Dilemmas in 3.6 and in publication V). This is the reason I turned to other phenomena that are more clearly definable independently of preloaded music theoretical concepts: salience of Impulses and musical parameters.

Another duration-based concept including tension is that of Huron (2006) who has developed the ITPRA theory or model of expectation embracing the following steps while auditorily cognizing music: imaginative response (I), tension response (T), prediction response (P), reaction response (R), appraisal response (A). He binds together five functionally different neurophysiological systems (Huron 2006: 7, 17).

The KWhat/KHow, IDyOM, SMuSe, ITPRA and the general psychological model of tension and suspense are dynamic models because they enhance the understanding of the dynamic phenomenon of music via its dynamic listening process. Although I do not deal in my dissertation with algorithmic, computational probabilistic-prediction- or Artificial-Intelligence-based modeling that develops predictive models[57] to be tested empirically, I cannot deny the enculturation-based general mechanisms of music perception explained by Pearce (2018: 1). This is as follows: "a broad range of psychological processes involved in music perception – expectation, emotion, memory, similarity, segmentation, and meter – can be understood in terms of a single, underlying process of probabilistic prediction using learned statistical models." This means that a cognitive model which enhances the understanding of the listening process of music must be 1) dynamic, and 2) take into account probabilistic predictions the listener "automatically" and intuitively performs. While understanding music perception as involuntary/"reflexive" in contrast to conscious/voluntary processes applied in music theory/analysis (Cross 1998a, Purwins *et al* 2008b), cognition processes during the listening process (e.g. "immanent musical analysis") (Chouvel 2014), and ecologist cognition/knowledge of music are based on the individual's mind's cognitive dynamics (Kaipainen 1994) (see also publication V). Purwins *et al* (2008a: 154) propose the following four aspects of cognitive modeling: (1) predictive power, (2) generalizing power, (3) simplicity, and (4) relation to existing theories. In this dissertation (publication III) the predictive aspect applies only partly,

---

[57] My Twelve Strategic Steps for modeling/analysis of scientific models (TSSM) are not primarily concerned with the predictive power of models (publication III).

especially because I do not develop a prediction model (publication IV). The aspects of generalizing, simplicity and relation to existing theories are crucial in my COSM: Cognitive Octagonal Slice Model (see more in 4.4 in Q 9 (E). What type (function/method) of model makes sense to apply?). The next subsection introduces *protonarratives* as a "pre-definition" of musical tension.

2.6 *Protonarratives* as a "pre-definition" of musical tension

For the analysis of musical tension in the context of a temporal analysis of music (their relation to cognitive tension/relaxation and expectation, see publication IV), temporal dynamic forms play a crucial role (Mailman 2010). As mentioned in 2.3 they are based on *emergent qualities/properties* measurable as magnitudes and have been analyzed and visualized as "retrospective contour of the flux of intensity of qualities" (Mailman 2010: vi, vii). Mailman's holistic musical form models "demonstrate new connections between composition, interpretation, and listening. Some of these pertain to narrative and cross-modal expressivity" and advocate a pluralistic understanding of music as musical open-ended flux whose many elements can be modeled computationally (Mailman 2010: viii).

As important for innate competencies in musicals communication the music psychologist Michel Imberty (2000) has discussed the following concepts and objects of observation in relation to music: gestaltism and grammar systems, universality and innateness, neurophysiological cognitivism, the question of psychological reality, the question of atonal music, dynamic aspects of salience cues and the concept of macrostructure, and the psychological foundation of macrostructure (on "macroform" see Baroni 2003, Addessi & Caterina 2000, 2005; Addessi 2010). In the last section, Imberty (2000: 459) refers to Stern's concept of "vitality affects" (see its use since 1985, Stern 2010: 40–41) as *temporal dynamic experience* (later called *dynamic forms of vitality*, see 2.3) that Imberty understands as *dynamic vectors*. These are "musical elements that transport temporal significations of orientation, progression, diminution or growth, and repetition or return.[58] Perceived and felt change is thus

---

[58] Compare the notion of vectors in relation to musical elements/parameters with Tüür's "vectorial" compositional method (see 1.1). See Eitan and Granot's (2007) *intensity contours* of different musical parameters (see 3.1).

a dynamic vector that orients the listener's perception, anticipation, and internal representations" (Imberty 2000: 459). Furthermore, Imberty (2000: 460) transfers Stern's term-phrase *dynamic vitality affects* into "protonarrative envelope" and interprets it as a "quasi-plot, a line of intuitive dramatic tension" to explaining music in relation to cognition and musical tension that, in my dissertation, functions as a kind of "pre-definition" of musical tension:

> "A musical piece is first of all an ordering of acoustic events in time. The macrostructure is a simplified schema type, an ordering a priori that later will be filled by the concrete acoustic events of which the progression may be defined for the listener as a structure and hierarchical succession of tensions and relaxations. In consequence it is simultaneously defined at the level of musical grammar, cognitive operations active in composition and comprehension of the piece, but also at the level of expressivity and feelings of the listener. The temporal progression through tensions and relaxations, and formal patterns that evoke what I can now call vitality affects, takes meaning in oriented continuity from the beginning to the end of the piece. It finds its coherence in this temporal web that links melodic, rhythmic, and harmonic gestures, telling protonarratives of a thousand nuances." (Imberty 2000: 460)

The idea of a *pre-state* or *proto-state* in *protonarratives* can be found also in Bachmann's (2000) "perceptual prefix state" and perceptual "proto-objects" in *microgenesis*. Furthermore the *proto-idea* is analogous to the *pre-(or semi-)categorial object* or so called *object** proposed by Kraus (2013/2016: 335) enlarging Kant's understanding of an early psychophysical measurement perception/cognition mechanism in order to discriminate an object of inner sense or sensation (music as CPTOM, musical tension as CMT) from a "fully fledged" object of experience perceivable via outer sensation – an intersubjectively shared standard, a spatial object. This foreshadows and supports my own cross-domain definition of musical tension ('Perceived structural musical tension', PSMT) based on narrativity (see publication IV) presented in 3.5 – see the following section 3.

3. Key concepts, dilemmas and the objective

This section introduces the key concepts, definitions and methodological principles around the general object – contemporary post-tonal orchestral music (CPTOM) – and the special object of this dissertation – musical tension. It deals with musical parameters as structural (compositional) and perceptual-cognitive (listening) attributes, salience as perceptual-cognitive mechanism, and finally narrativity as a cross-domain cognitive way of thinking to enhance the understanding of the captured and compared data. The key concepts relevant in this dissertation are the following:

- general psychological tension (and suspense), musical tension, structural musical tension, perceived structural musical tension;
- salience and attention: auditory, visual and relative salience, bottom-up and top-down attention, salience in cognition and as an important component for a composer;
- musical parameters and their compositional and perceptive dimensions;
- narrativity.

The following subsections present research and concepts around music and tension, salience and attention, salience in complex sounds and contemporary music, musical parameters and narrativity. Together with narrativity and in the context of music and musical tension as cognitive (CMT) I present my fully fledged cross-domain definition of the 'Perceived structural musical tension' (PSMT). Furthermore, I introduce the five scientific Dilemmas that are accompanied with Research Questions (RQs) and Statements (STs) (details and Answers see publication V) aiming to enlighten the challenges and opportunities while observing musical tension in complex musical structures. Finally the main objective of this dissertation is presented in the context of central terms relevant for the research methodology (*purportedly*, *trigger*, the *why*-question, *segmentation*).

3.1 Music and tension

The musicologist and composer Leonard B. Meyer (1956) and the cognitive musicologist David Huron (2006) treat musical tension in the context of emotions. In contrast, Lehne's and Koelsch's (2015) general psychological model of tension and suspense understand perceived tension in general (in the arts and film domains) as a meta-phenomenon, not only as one of many emotions. Tension can be analyzed as "felt" or "perceived" (Lehne & Koelsch 2015: 3, 5, publication IV), in this dissertation "perceived" tension is in focus (see definitions in publications I, Ia and IV; the development of the definitions see in Lock 2022). Other relevant approaches to research musical tension see in publications I and IV.

The music theorists and cognitive researchers Zohar Eitan and Roni Y. Granot (2007: 39) describe "diverse musical processes in terms of changes (increase or decrease) in "intensity" while examining "that analogous intensity changes in different musical parameters can be perceived as similar." They empirically tested melodic direction, and attack rate and their combinations, pitch interval size, motivic pace, and harmonic tension in comparison to "standards" of these processes. In their results they conclude that *intensity contours*[59] of different musical parameters can be perceived as analogous and as musical "gestures"[60] no matter the specific parameter triggering them.

Addessi and Caterina (2000: 34) have observed the correlation between perceptual segmentation points (as local continuity/discontinuity points of separation between two sections) and perceived points of tension and relaxation in post-tonal music. The procedural/temporal/dynamic aspect of musical tension/relaxation in relation to stability and instability of contrasting zones in post-tonal/atonal music has been expressed by Addessi (2010) with the metaphor of breathing: "the tension zones were to be characterised by instability, as if taking a breath in; the relaxation zones were to be characterised by stability, as if exhaling. We asked [the listening test participants] to identify only wide tension/relaxation zones and not the smaller tension/relaxation changes" (Addessi 2010: 230). The criteria causing the listeners to divide (segment) parts from each other were collected with verbal feedback asking for

---

[59] Compare these *intensity contours* with Mailman's (2010) *temporal dynamic forms*, Bachmann's (2000) and Rosenthal's (2004) *microgenesis*, Stern's (2010) *dynamic forms of vitality* (see 2.3) and Imberty's (2000) interpretation of the latter as *dynamic vectors* (see 2.6).

[60] See also the theory of 'musical gesture' by Hatten (2004) and the notion of direction of perception as gesture in Lock (2017b), see 1.2.

descriptions of the "quality" of each part in adjectives and sound parameters (Addessi 2010: 228). This is a language-based understanding of musical tension opening up for a semantic dimension.

According to the music and technology researcher Dorien Herremans and the interdisciplinary researcher Ching-Hua Chuan (2017: 1), tension is "a semantic characteristic of music that directly shapes the music experience and thus forms a crucial topic for researchers in musicology and music cognition." The cognitive and computational musicologist Morwaread Farbood (2012: 387) has called musical tension to be evident to listeners and "relatively easy to define in informal, qualitative terms. [...] However, formalizing and quantifying such a description is a difficult problem." According to Farbood (2012: 387) the "flux inherent in music is what defines it as an immediate, temporal experience. It is also a key element in the perception of musical tension, a sensation that arises from the combined interaction of various musical parameters." This understanding is similar to the concept of immanent cognitive musical analysis based on the flux of musical "events" as structured "objects" (Chouvel 2014). The mentioned concepts remind of Kant's understanding of *analytic idealism* (Kant 1998, Kraus 2013/2016, see 2.2). Intriguingly, Farbood (2012: 387) claims that "tension as a concept is an emergent phenomenon that previous studies have shown to be judged with considerable consistency by listeners" no matter of their familiarity (via listening and performing) with the music before.

The most detailed insight into the variety of research approaches to musical tension (from solo works to orchestral music), especially in respect to musical parameters, is given by Farbood (2006, 2012). Farbood (2012: 388) categorizes musical tension research into "domain-general psychological" (low-level auditory perception, higher-level cognitive features including loudness and dynamics, timbre, pitch register, as well as Gestalt principles, expectation and emotion) and "domain-specific musical" (discrete-pitched and rhythmic elements of music, including melodic contour, harmony, tonal perception, harmonic (tonal) tension, rhythm, meter, and timing). Dorien Herremans and Elaine Chew (2016: 1) categorize musical tension into the realms of psychology (models of expectation and emotion, semantic meaning of lyrics) and music (low-level structural examined features like rhythm and timing, harmonic tonal perception, pitch height/melodic contour, dynamics, timbral elements (roughness, brightness, and density), pitch register). Farbood (2012: 404) admits that "it was difficult to evaluate relative contributions of each feature in interactive situations because in most cases there was no way to measure the impact of salience. Most of the stimuli were designed to have changes that were as obvious as possible for the various parameters involved." Herremans and Chew (2016: 1) refer to Farbood (2012) underlining that "listening to music is an aggregate experience that requires the integration of many different features. A listener's attention can focus on one feature at a particular time and then shift to a different feature or combination of features."

An advanced technological approach to tonal musical tension research with a practical outcome has been achieved by Herremans and Chew (2016) who developed the so-called tension ribbon method based on the harmonic structure of the musical score to quantify and visualize tonal tension. Herremans and Chuan (2017) have even developed a multi-modal platform for semantic music analysis visualizing audio- and score-based tension. Both approaches have been tested empirically. Herremans and Chuan's (2017: 1) statistical analysis of the results of tonal (from the score) and timbral tension (from the performance) discovered correlations between them – they finally used a clustering algorithm and musical score-following to visualize similar segments in their online Interactive Multi-modal Music Analysis (IMMA) system.

The computational researcher Aozhi Liu *et al* (2020) offer a short overview of the research on the perceptual dimension of musical tension (e.g. tension-relaxation relations, emotional aspects) as multidimensional and auditory and music-feature (musical parameter) related. The neuroscientist Lijun Sun *et al* (2020) observe musical tension in association with violations of hierarchical structure using "continuous" behavioral rating and electroencephalography (EEG).

In terms of neo-, atonal and nontonal music the amount of studies is smaller. In a quasi-empirical approach the composer Martin Kirschbaum (2001) has observed musical form dramaturgy via musical tension in a whole repertory of contemporary music from the second half of the 20th century. Empirical analysis of musical tension of atonal and non-triadic music (Milhaud, Maderna, Webern) through observing the segmentation process of their participants in relation to musical structure (musical parameters) and "macroform" has been conducted by Addessi and Caterina (2000). Furthermore, Addessi (2010) examined the correlation between "macroform" and perception of tension/relaxation. Yvonne Teo's (2020) study on theoretical and perceived harmonic tension in neoclassical music (Hindemith, Ravel, Stravinsky) discusses tonal and atonal tension and shows the importance of physical and acoustical factors (e.g. roughness). For an overview of psychoacoustic aspects of tension (roughness, sensory consonance, and timbre) see the auditory hearing researcher Daniel Pressnitzer *et al* (2000, perception of musical tension for nontonal orchestral timbres and psychoacoustic roughness).

On consonance and dissonance perception see Nicola Di Stefano, Peter Vuust and Elvira Brattico (2022).

The phenomenon of salience (of pitch and harmony) to analyze atonal and later nontriadic chromatic tonal music was suggested by Lerdahl (1989, 2001) in relation to the concepts of prolongation (taken from GTTM, Lerdal & Jackendoff 1983) and attraction points – this has inspired me to move on in the cognitive direction. As Farbood (2012: 404) has mentioned, the empirical measurement of the impact of salience on musical tension is something that has not been done, yet. This dissertation therefore offers a possibility to approach musical tension from the salience viewpoint.

However, initially I have been inspired by Kirschbaum (2001), who, in his study on "formal tensional developments" (*"formale Spannungsverläufe"*) as tensional (or wave-like) developments of contemporary or new music (*Höhepunktbildung und Dramaturgie in Neuer Musik*), discriminates (more simplifying then Farbood 2012) between the following groups of musical parameters relevant to an understanding of musical tension: psychological (non-exactly measurable aspects like emotions), and acoustic-physical (exactly measurable parameters including pitch, loudness, rhythm, density, texture etc). I have slightly modified Kirschbaum's (2001: 23–26) list of musical parameters but his (in my view) problematic quasi-empirical approach – conducting a self-experiment without a clear explanation of the theoretical basis of his empirical methodology – has triggered me to dive deeper into the research methodology of musical tension, ways of analyzing/observing music and musical tension in both music psychology and music theory, that has led me to a cross-domain, cognitive understanding.

In this dissertation I assume that musical tension is co-constituted by both structural features of the musical composition and the experience its performance evokes in the listener – enabling therefore the application of both music theoretical and empirical analysis methods (proposed first in publication I, see music-theoretical and music-psychological matters intertwined in 2.1). Structural musical tension can be understood as a listener's response to formal patterns whereby "the intensity" of the experience "of the perceived musical tension is proportional to the structural (or hierarchical) significance of the corresponding musical event" (see publication Ia). Perceived structural musical tension as I have arrived at it in the main empirical study of this dissertation (publication IV), is a global (temporal, musical-form-related) as well as a local phenomenon that can be observed (experienced, analyzed through listening) first of all as a compound whole (via "continuous" data capture). Secondly, auditory salience (see 3.2, 3.3) of musical events I call Impulses/“moments of change” as “relative change” (“just noticeable difference” (JND) as principles of psychophysical measurement, see 2.4, 4.3 and publication V) enables to gain insight into the compound whole of a musical event. Finally the detection of the salience (see 3.2) of “primary” and “secondary” musical parameters (see 3.4) via the mediation of auditory via visual salience is realized with the new COSM: Cognitive Octagonal Slice Model – basic modeling principles have been developed in publication III, its practical application has been conducted in publication IV. The methodological challenges to scientifically observe and detect musical tension are presented in publication V. As follows I introduce the aspect of salience in relation to attention.

3.2 Salience and attention

Salience (or *saliency*)[61] can appear in different domains, especially in sound and vision. In this dissertation I assume that the phenomenon of *auditory salience* (Jensenius 2002; Lerdahl 1989, 2001) can be mediated via the phenomenon of *visual salience* (Jahanian, Vishwanathan & Allebach 2015). Hence, *relative salience* plays a role in detection of change in *visual signals/sensory events* meaning that “relative saliency being determined by both physical salience and subjective weight of feature information” (Yang 2011: 1708). Salience as bottom-up attention has been understood as “automatic responses to salient events” in contrast to top-down attention as focusing volitionally (Noyce, Kwasa & Shinn-Cunningham 2022: 2 etc). Both forms of salience should be taken into account: the former works at the level of sensation and basic perception mechanisms, the latter works based on cognition.

The question of the role of *salience* in relation to *attention* in perception and cognition is based on “scientific evidence [suggesting] that our sampling of the world is highly selective and proactively shaped by numerous external and internal factors” in a way that we do not have direct access (see 1.2, 2.2, 2.3) and apprehending abilities to the world around us (Nobre 2018: 241). The question therefore is what are the mechanisms of this sampling selection process, especially in regard to experiencing music via listening.

---

[61] In this dissertation I mainly use the term *salience*, in quotes also the term *saliency* can appear.

According to the psychologist William James (1890) the classical definition of *attention* includes the following aspects: (1) process, not mental representation, (2) prioritization and selection, (3) only one item is chosen at a given time, (4) can select items from the external environment (objects) or from the internal mental landscape (trains of thought), (5) involves focusing, includes inhibiting (withdrawing from) distracting items, (6) to guide adaptive behavior (to deal effectively), (7) essential for normal, healthy cognition.

The cognitive neuroscientist Britt Anderson (2021/2023) makes the bold claim that *attention* as a modern psychological term is fragmented with overlapping and confusing meanings – mirroring the also bold claim that "no one knows what attention is" (Hommel *et al* 2019). The philosopher Sebastian Watzl (2023), on the other hand, advocates a *priority structure account* of attention based on an understanding of attention as a first-person conscious experience and suggests to unify attention research advancing especially the computational neuroscientist David Marr's (1982) theory's computational level. For Watzl (2023: 3–4) the role of attention depends on four aspects applicable in computation: 1. Activation of current states of the mind in the here and now, 2. Guided 'activity' by an agent or organism, sensitive to executive control, 3. Meta-representationality in organizing information, not being accurate or inaccurate, satisfied or unsatisfied, 4. Usability making information accessible or usable as 'priority' in the mind with neutrality towards its interpretation. See also the lack of truth or accuracy in the biological function of sound localization for survival (Huron 2006: 105).

Actually in cognitive science, attention stands for *selective attention* (also *prioritization*) assuming that "the selective and proactive qualities of our perception and memories are so intrinsic and entrenched that we hardly notice them" (Nobre 2018: 244). Further elaboration classifies attention as a proactive process that "prepares and guides perception in an anticipatory way based on what we have in mind and what we have experienced" (Nobre 2018: 244). "Perception of external objects results from the interaction between an internally driven ideation or memory and incoming external sensory stimulation" (Nobre 2018: 244). See in this context also the Kantian understanding of cognition (Wolff 1963) enlarged by Kraus 2013/2016 to measure inner sensations (magnitudes) via outer sensations (magnitudes) (see 2.3). In a cognitive research methodology the main focus is on the performance aspect of attention (Nobre 2018: 250), less on the learning aspect that includes uncertainties, especially learning in a new environment: "it is what you *don't* know that becomes important" (Nobre 2018: 251). According to the neuroscientists Thomas Parr and Karl Friston (2019: 2–4) *attention* and *attentional gain* mean emphasizing one type of sensory data and need contextualization whereby *salience* deals with controllable states by the perceiver and is based on action. Both can, of course, fail, the former as false precepts, the latter due to active scene construction problems.

While considering music and musical tension as cognitive that requires listening as an active process the cognitive musicologist Mauri Kaipainen's (1994: 3) computational cognitive model of 'knowing-what' (KWhat, momentary musical situation: listening) *vs* 'knowing-how' (KHow, processing perceptual, cognitive or motor aspects: acting, reconstructing, generalizing) (see 2.5) can be enlarged beside aspects of expectation/anticipation, recognition, and autocommunication (Kaipainen 1994: 22–23) with the juxtaposition of *attention* (something rather unexpected in the music catches the listener's awareness) *vs salience* (a rather controllable process: the listener turns towards something that has catched interest). Compare this with the understanding of *bottom-up attention* as "automatic responses to salient events" *vs top-down attention* as focusing volitionally (Noyce, Kwasa & Shinn-Cunningham 2022: 2 etc). The volitional component means a *first-person* (agent of action) research approach in contrast to a rather reflective, automatic *third-person* (observer, behaviorist) research approach (see Valenzuela-Moguillansky & Demšar 2023). However, observing attentional processes means to me to research the observers with their abilities. Observing salient processes means to me to research the object of research (music, musical tension) and its affordances while treating the observer as "instruments" of observation and measurement (see Bachmann 2000, 2.3) co-analysts (similar to Addessi & Caterina 2000, 2005; Addessi 2010, 4.1). The latter is the focus of this dissertation.

Attention in the form of *auditory* and *listening awareness* plays a crucial role in the explanation of the musicologist Erik Christensen's (1996: 10 etc) *basic listening dimensions* (see more in 3.5). Christensen (1996: 10) claims that "hearing is not designed for music listening," instead it functions in natural environment survival as attention trigger of events and danger as well as in spatial orientation supporting localization, distinction and movement of sound and sound sources. See also the importance of anticipating future events for survival underlined by Huron (2006: 7). By physical sound waves reaching the ears "the auditory system is activated" and "aroused to a state of attention," even during sleep. The "listening mind" becomes auditorily aware of and emotionally attached to the event (Christensen 1996: 10) (see more in 3.5). *Auditory awareness* is related to the *macrotemporal listening dimensions* 'movement' and ‘pulse’ responsible for “the experience of time in the listening process,” ‘movement’ triggers the awareness of change, ‘pulse’ the awareness of regularity[62] (Christensen 1996: 13–14). As follows I introduce research on salience in complex sounds and contemporary music.

3.3 Salience in complex sounds and contemporary music

Salience in music can be understood structurally as a *musical event* that protrudes and catches *attention* in the context of other musical events. Analytically salience can also be understood as “important perceptual points” that are preserved in segmentation- and reduction processes over several layers (see publications Ia, IV). Protruding and catching attention requires *attraction points* to be integrated into and detectable in the music (Lerdahl 1989, 2001; Kirschbaum 2001). In this dissertation I define a ‘salient musical event’ as ‘Impulse’ that can be understood as a “moment of change” that (due to its slowness in comparison to impulses observed in neurology) requires *mental pointing*, *focal attention* and *attentive listening* (see publication IV).

A study by the transdisciplinary researcher Martin Coath *et al* (2009) has aimed to measure *auditory salience* (based on a general model of auditory processing), exploring that a “biophysically motivated approach for the segmentation and identification of salient events in a stimulus is suitable for the field of automatic music processing” (Coath *et al* 2009: 193–194). To “model cortical responses for the detection of perceptual onsets and beat tracking in singing” they compared results from two separate approaches to enhance the understanding of complex sounds: a model of auditory salience (including a cochlear, a transient sensitivity, and a cortical model) resulted in a summed-up salience graph via transient enhancement; the multi-resolution representation of rhythmic structure was based on a wavelet transform (Coath *et al* 2009: 195–196). In another research they has tested “a robust sound perception model suitable for neuromorphic implementation” that deals with “the emergence of dynamic feature sensitivity through exposure to formative stimuli in a real-time neuromorphic system implementing a hybrid analog/digital network of spiking neurons” (Coath *et al* 2013: 1). The robustness and

---

[62] From a composers’ point of view see more on “the rhythmic idea and the musical representation of time” as dynamism and multiplicity in a post-Kantian philosophical understanding elaborating further Schoenberg’s and other’s “idea” concepts (Rodrigues 2019).

"variability of the responses" of their system or network to "noisy" stimuli (e.g. of naturalistic and speech samples, ecological settings) enabled them to claim that their approach "could lead to a neuromorphic sub-system engineered for dynamic pattern recognition in real world applications" (Coath *et al* 2013: 9). Research on auditory salience has been conducted to analyze *natural soundscapes*, *natural scenes* and *real-life scenarios* (Kothinti, Huang & Elhilali 2021) as well as *natural scenarios* (Tordini *et al* 2023) (music as environment/ecological system see 1.2, 1.3, 2.4, 2.5).

Furthermore, the computer scientist and composer Shlomo Dubnov and the music cognition and perception researcher Stephen McAdams (2006) analyzed *structural* and *affective* aspects of music from statistical audio signal analysis in contemporary (and electronic) music observing in addition to the signal analysis familiarity ratings (FR) and emotional forces (EF) reported empirically with slider-controllers during the live-performance of two versions of a contemporary orchestral work by the composer Roger Reynolds. Their findings revealed a strong relation between signal properties (statistical similarity and repetition aspects) and human responses: "a significant correspondence for FR and a high correspondence for EF between our simple statistical audio signal analysis and experimental human responses" (Dubnov & McAdams 2006: 2, 26–27). They underline that the listening experience based on perception and cognition plays with building and violating of expectations "create tension and resolution over time" (Dubnov & McAdams 2006: 27). They underline that musical parameters like melody, harmony and rhythm may improve the results, however, they admit that "these musical features are complicated for estimation from a raw audio signal. A further subject for future research is to find a method for machine learning of regression coefficients" (Dubnov & McAdams 2006: 27).

Salience has also been applied to interpret the results of a perceptual (empirical) study on *dynamic forms* (see 2.3) in excerpts of 19th, 20th and 21th century tonal and atonal solo- and orchestral music from Beethoven (1803) to Messiaen (1944) to Alva Noto (2009). *Suspense* or a *state of tension* (see also the general model by Lehne & Koelsch 2015, see 3.1) on a more global level has appeared to be more salient than *dynamical forms* (triggered by sudden changes in instrumentation and loudness) (Hjortkjær & Nielbo 2010: 4).

The salience of *musical parameters* is related to the aspect of *perceptual accent* as "an event that occurs when a note or sonority appears to be perceptually more important than the other notes" emerging with variations in musical parameters (e.g. note duration, pitch, dynamics) categorized either as “immanent” or “performed,” or as phenomenal accents (Friberg *et al* 2019: 2). Friberg *et al* (2019: 3) have further developed and applied a quantitative model of *pulse salience* and *metrical accent* in musical rhythms (based on Parncutt 1994) to observe tempo changes in performances. Computational research based on musicological concepts like metrical (e.g. hypermeter) and melodic contour accents (e.g. melodic climax) in relation to salience has been conducted by several researchers (see Friberg *et al* 2019: 3).

Salience and *relative salience* (see 3.2) have been addressed also by the composer Alan Belkin expressing that music as a meaningful way of communication must be interesting for the listener via “the relative salience of musical phenomena” (Belkin [2025]). He propagates to put greater focus on salience because it connects music analysis and cognition – practical examples how to apply salience he gives also in Belkin (2018). Furthermore, “when attempting to sensitively describe and explain musical phenomena, prioritization is necessary, because that is how the human brain works” (Belkin [2025], on prioritization/selective attention and salience see 3.2). In this dissertation I approach the notions of “triggering attention” and “prioritization” through the concept of salience also, because Belkin [2025] suggests that “a theory of salience could thus be a bridge between the fields of analysis and music cognition.” Salience is “too often dismissed as unimportant” in music theory and composition teaching (Belkin [2025]).

As follows I introduce the (changing) role of musical parameters in (contemporary) music in the context of perception/cognition, “primary” and “secondary” parameters’ categorization, extensive and intensive magnitudes of psychophysical measurement based on a Kantian understanding. This subsection ends with an introduction of Christensen’s (1996) listening dimensions for an alternative perception-based musical parameters’ categorization.

3.4 Musical parameters

Musical parameters are traditionally categorized in music theory/analysis/composition and in music psychology as “primary” and “secondary” (Meyer 1956, Snyder 2000: 195 etc.). According to the composer, sound and video artist Bob Snyder (2000: 195) “primary parameters” possess “relatively fixed proportional relationships between” values of “musical variables” (e.g. in tuning systems of pitches and systems of temporal intervals and durations of notes) which's patterns and variations can be identified as learned "*conceptual* categories." In contrast, "secondary parameters" do not have "standardized scales of proportional values," therefore they "cannot easily be divided up into very many clearly recognizable categories"; they are perceivable rather in "*relative amounts*" like "much"/"not much" or "more"/"less"with the ability to "reinforcing patterns in primary parameters" (Snyder 2000: 196–197).[63] The "primary" parameters (PP) I have chosen for this dissertation are the commonly regarded pitch, rhythm and harmony; the "secondary" parameters (SP) are dynamics, tempo, instrumentation/timbre, texture, effects based on the modified list taken from Kirschbaum (2001).

However, *dynamics* as loudness changes, physical sound pressure level (and amplitude) is considered as essential in early processing together with pitches and simultaneous intervals (event fusion) (Snyder 2000: 11 etc., 25 etc., 124 etc.). According to Alison Abbott (2002: 13) *pitch* and *loudness* are fundamental elements of music that in the brain are identified by the primary cortex (thalamus): "The secondary cortex is believed mainly to focus on harmonic, melodic and rhythmic patterns, and the tertiary auditory cortex is thought to integrate these patterns into an overall perception of the music."

Furthermore, the not absolute categorization into "primary" (PP) and "secondary" (SP) parameters (Snyder 2000: 195) seems having its roots (not to mention underlying older philosophical backgrounds) in an Kantian (pre-psychophysical-measurement-type) understanding of *extensive* (outer sense) and *intensive* (inner sense) *magnitudes* and their possible measuring (Kraus 2013/2016) based on Kant's mathematical principles: *extensive magnitudes* as additive successions – e.g. pitch, rhythm, and harmony as pitch combinations as PP's, and *intensive magnitudes* as degree-based – e.g. dynamics/loudness, tempo etc. as SP's.

Hence, in the larger methodological context of *introspection* and *intuition*, *perceptual experience* and *perceptual justification* (see 1.4, 2.2, 2.3) and in the narrower context of *psychophysical measurement* we experience/perceive/cognize sound and music as a *continuum* – more specifically as *prothetic* (including intensive parameters, e.g. dynamics) and as *metathetic* (including extensive parameters, e.g. pitch; Shepard 1978: 451–452) (see principles of psychophysical measurement in 2.4). This is explained more simply by the clinical neuropsychologist George Jewell (2005: ii): "Prothetic continua involve judgements of how

---

[63] "Note that it is also possible to use primary parameters as secondary parameters, although the reverse is usually not true" (Snyder 2000: 197).

much (i.e., stimulus intensity) and correspond to power (log-log) functions. Metathetic continua involve judgments of what or where and correspond to linear functions."[64]

The understanding of musical parameters has changed with the emergence of so-called *atonality* at the beginning of the 20th century. Especially the rules of pitch organization and structure, fundamental to Western music composition, were questioned and re-organized. Most significant in the first half of the 20th century have been Arnold Schoenberg's dodecaphony and Josef Matthias Hauer's twelve-tone music based on hexachordal tropes. After WWII developments in *serial music* (Olivier Messiaen inspired Karlheinz Stockhausen, Pierre Boulez and others) not only the serialization of pitch but also of other musical parameters was applied.[65] Less known is a variety of so-called microtonal approaches.[66] These include in Europe quarter-tone music by Alois Hába. In recent decades microtonal approaches have been used more extensively, e.g. by Georg Friedrich Haas, Enno Poppe, Georg Hajdu, Manfred Stahnke. In America (microtonal and rhythmical) alternative approaches by Charles Ives were more known than Harry Partch's or in particular Ivor Darreg's achievements in theorizing and compositionally applying microtonality. To mention just some studies dealing with microtonal systems: Paul Erlich (2006) suggests a middle path between Just Intonation and the Equal Temperaments, Thomas Nicholson and Marc Sabat (2018) explore fundamental principles of Just Intonation and Microtonal Composition, and Peter Thoegersen (2022) discusses polytempic polymicrotonal music. New approaches to musical parameters were developed in (serial) *electronic music* (since the 1950ies Stockhausen and others), but also in (orchestral) *sonorism* (since the 1960ies György Ligeti, Witold Lutosławski, Krzysztof Penderecki, Arvo Pärt and others), Iannis Xenakis' (architectonic and stochastic) orchestral mass structures, Helmut Lachenmann's (noise cultivating) string playing techniques *(musique concrète instrumentale*), in *spectral music* (since the 1970ies Tristan Murail, Gérard Grisey) or the *new complexity* (since the 1980s Brian Ferneyhough, Michael Finnissy and others).

Snyder (2000: 195) has claimed that "primary parameters are not more important than secondary parameters; rather, they operate in a different way" but this obviously concerns mainly classical music before the 20th century second half academic avant-garde music. The twentieth-century music researcher Amy Bauer (2001), based on Ligeti's and others'

---

[64] See also the chapter "Distinctions between metathetic and prothetic continua" (Jewell 2005: 1, 20–21).

[65] The "vectorial" method of aggregating intervals and the "tone" (timbre as depth dimension) applied by Erkki-Sven Tüür continues this tradition in a personal way (see 1.1).

[66] See more in the Xenharmonic Wiki [2025] and the Tonalsoft Encyclopedia (Monzo [2025]).

explanations and analyzing a number of Ligeti's micro-polyphonic orchestral works from the 1960ies, underlines that "the "new language" has much to do with an ironic reversal of the traditional relationship between "primary" (PP) and "secondary" (SP) parameters: the elevation of timbre, articulation and dynamics over pitch and rhythm as determinants of musical structure" (Bauer 2001: 38). Sound mass music is indeed complex and rather complicated to analyze, and requires computational and even neuroscientific methods (e.g. see Noble *et al* 2020, see 4.1). I have tested Bauer's hypothesis of SP's ruling over PP's for Tüür's Fourth Symphony in the context of the salience of musical parameters and in comparison with musical tension and found that this change of relationship didn't appear in general, only partly in the middle of the music before the cadence (see publication IV).

While analyzing contemporary post-tonal music the musicologist Erik Christensen (1996: 44–45, 145, 2012: 321–322, 362) has proposed a different categorization of musical parameters in the form of *eight/nine listening dimensions*. He discriminates *five basic dimensions* (intensity, movement, timbre, pitch height, pulse) and *three/four secondary dimensions* (rhythm, melody, harmony, micromodulation as a ninth dimension). Christensen categorizes these into *micro-* and *macrotemporal listening dimensions*: 1) microtemporal – basic listening dimensions timbre (prominent quality) and pitch height (pitch continuum), between them the secondary listening dimension harmony); 2) macrotemporal – basic listening dimensions movement (shape) and pulse (tempo), between them the secondary listening dimensions rhythm (temporal dimension) and melody (spatial dimension). According to Christensen (1996: 11–12, 144 etc, 2012: 361) the three microtemporal dimensions of listening are *intensity*, *timbre* and *space*. Christensen (1996: 11) claims that only above a certain threshold of physical intensity of sound the mind gets involved. He states that "as a listening dimension, intensity is a subjective quality, largely dependent on the loudness of sound", because also other factors shape the sound events: *distinctness*, *sharpness*, *duration*, *temporal density*. *Sharpness* and *softness* (see their relations to the nature of the sound object, Vicario 2003) are characteristics of timbre and help the listening mind to judge the relevance (selective attention, prioritization) and identity of sounds, whether to focus on parallel or single events (Christensen 1996: 11). Furthermore, Christensen (1996: 14–16) understands intensity both as a microtemporal and a macrotemporal listening dimension, because, based on auditory awareness (see in 3.2), intensity informs us about successive changes of states and events of sound sources. Pitch as a predominant parameter in (Western) music is much more related to timbre than recognized in traditional music analysis (Christensen 1996: 16).

The role of musical parameters in contemporary post-tonal music is observed in this dissertation both from the traditional "primary"/"secondary" musical parameters' (Meyer 1956, Snyder 2000, see publication IV) and an alternative listening dimensions' (Christensen 1996, 2012, see above) viewpoint. Both are supported by basic perception aspects, e.g. in the relations between discernable and changeable "perceptual qualities" referred to as "parameters" (pitch and timbre)[67] for establishing musical form (McAdams 1989: 182, Snyder 2000: 194), more precisely temporal dynamic form (Mailman 2010) (see 2.3). According to McAdams (1989: 183) one has to understand the "perceived qualities of musical events" as related to a "learned system of relations (scale, meter, harmonic field, etc.) that is more or less strongly evoked by the relations among events in the musical context" as an abstract structural knowledge gained via listening experience of a wider body of musical works "to establish the relative stability or salience relations among the values along a given dimension."

In the next subsection I bind together *protonarratives* (Imberty 2000, see 2.6) and relative stability/salience relations of perceived qualities of musical events (McAdams 1989) with several concepts of narrativity and music finally arriving at my cross-domain "fully fledged" (Kraus 2013/2016) definition of 'Perceived structural musical tension' (PSMT).

3.5 Narrativity and a "fully fledged" definition of musical tension

This subsection discusses several concepts dealing with *narrativity* and music in order to arrive at a structural understanding of narrativity in music in relation to perception and cognition. More generally, Mailman (2012/2013: 127) proposes that despite paradoxes (e.g. related to aspects of time, duration, repetition, determinism, teleology, normativity and agency) in understanding narrativity in music "narrative experience survives among the repeatable experiences of musical works afforded by the unprecedented modes of music's mechanical reproduction in the twenty-first century." However, Mailman (2012/2013: 127) underlines that "approaches to narrative in twentieth-century music may differ significantly from those for common-practice repertoire," a narrative appears via the interpretative interaction of the listener

---

[67] "Some perceptual dimensions are strongly correlated with the sensory dimensions along which grouping decisions are made, such as pitch and timbre being strongly correlated with the spectral changes that affect sequential and segmentational grouping" (McAdams 1989: 182).

with the music and a narrative analysis is related to a "critical metalanguage" influenced by intertextuality and non-musical/-sound context. "For less normative music, such as that of the twentieth century, a narrative cannot rely on an intertextually fueled conventional metalanguage and therefore must develop in a suitably particularized ad hoc fashion [...]" (Mailman 2012/2013: 127).

According to the cognitive musicologist Candace Brower (2000: 325) a cognitive theory of musical meaning can be built based on "intra-opus patterns" of music. These repeating patterns (or paradigms) in a work map onto each other during the listening process creating a *paradigmatic axis* (based on Ruwet and Nattiez, Brower 2000: 325). See also the "zygonic" theory (repetition, parallelism, imitation: a union of two identical things) by the composer, researcher and music educator Adam Ockelford (2009: 97) and its empirical testing as a model of expectation (Thorpe, Ockelford & Aksentijevic 2012). Brower (2000: 325–326) understands musical events as actions and a musical work as a journey that can be interpreted as "successive statements of a pattern as a sequence of related actions making up a musical narrative," this match of intra-opus patterns "allows us to assign relatively precise meanings to surface-level musical events" – also tiny alterations in these patterns contribute to their narrative interpretation. Brower (2000: 326 etc) includes into this theory embodied image schemas (container, cycle, verticality, balance, center-periphery, source-path-goal) that "reflect basic features of our bodily experience of *space*, *time*, *force*, and *motion.*" These are also part of the Cognitive Blending Theory (CBT, Antović 2018, 2022)[68] and belong to the "fundamental dynamic pentad" of Stern's (2010: 4–5) dynamic vitality forms: movement, time, force, space, intention/directionality. Following Brower (2000: 327) we map "these features of our bodily experience of the physical world (the *source* domain) onto music (the *target* domain)." Cross-domain mapping and analogies are basic principles of modeling (see Zbikowski 1997, 2002; see publication III).

Although the *narrative constructivism* (composing of events) is debatable (Saupe & Wiedemann 2015) and it has been doubted that a *musical narrative* can be found in music directly (e.g. Almén 2003: 3, Kramer 1991: 143),[69] there are definitions that can be interpreted from a structural viewpoint or are itself structural approaches to narrativity. The following

---

[68] Force (intensity), discrete distance, path, link, oscillation; verticality, stability; repetition, musical movement; generic, input 1 and 2 and blend spaces.

[69] Compare this doubt with the music historian and critic Eduard Hanslick's (1854) claim about the absence of emotions in the music that has been called by Leonard Meyer (1956: x, 1–3) and Amanda Staufer (2018: 38–39) a formalist and absolute understanding of music.

structural concept arose from the musicologist and composer Lawrence Kramer's (1991: 143) general definition of narrativity types gained from abstract sequences of repeated events that are variously concretized; these concretizations are the individual stories.[70]

Kramer (1991) interprets narrativity as a *dynamic principle* and *teleological* (direction towards a purpose) *impulse* (driving force behind or reason for something else) embracing all narratives.[71] The teleological meaning can be found in tonal music clearly in harmony, melody (pitch and rhythm) (primary musical parameters, PP's) and melodic arches (up/down contours) as well as in changing dynamics (de/crescendo) and tempo (accelerando/decelerando) (secondary musical parameters, SP's), and in developmental and culmination-based form principles. In contemporary post-tonal music the clear teleological structure of classical functional harmony is absent or the relations between chords are organized differently in relation to *temporal dynamic forms* in music (Mailman 2010, see 2.3). A kind of *impulse* (here in Kramer's philosophical meaning) I propose to find in the musical parameters that can be seen as the driving force/reason behind the unfolding of both the narrative and perceived structural musical tension.

Furthermore, composers have developed and tested algorithms to shape their music and its narrative via manipulating musical tension. Huang (2020) has connected narrativity, musical structure (parameters) and musical tension energy in algorithmic composition while testing it empirically with two questionnaires. Musical tension, more concretely musical tension energy based on the analysis of the primary musical parameters pitch, rhythm and harmony (chord progressions) was manipulated with an automated composing algorithm. This research is based on narratology/narrative theory assuming that music conveys a story consisting of diverse tensions. However, this is a rather practical research of tonal and film music composition using perception/cognition and tension energy as arousal and emotion similar to Le Groux' (2011) SMuSe (Situated Music Server) system composing avant-garde music (see 2.5).

I have applied Byron Almén's (2003: 12–13) five aspects of a musical narrative together with Kramer's (1991) definition of narrativity (dynamic principle and teleological, impulse) to explain the principles of my COSM: Cognitive Octagonal Slice Model (see more in section 4),

---

[70] The *narrative* as an „acknowledged story, whether typical (an abstract sequence of events repeatedly and variously concretized within a given historical frame) or individual (one of the concretizations)" (Kramer 1991: 143).

[71] "Narrativity is the dynamic principle, the teleological impulse, that governs a large ensemble of all narratives, up to and including the (imaginary) ensemble of all narratives" (Kramer 1991: 143–144).

in the context of the music under observation, symphonic pieces by Tüür (see publication IV). Almén's (2003: 12–13) aspects can be interpreted in relation to perception and cognition aspects as follows: 1) musical elements interact during the perception and cognition process; 2) early processing of pitches, simultaneous intervals and loudness changes (event fusion) as well as melodic and rhythmic grouping (Snyder 2000) built limited patterns that are archetypal; 3) these patterns can be perceptually in conflict on various hierarchical levels (e.g. appearing simultaneously in different voices, with contrasting dynamics and timbres) causing a constant revaluation of the elements building these patterns; 4) the unique surface features and effects of these patterns are clearly falling into the perception/cognition domain of music; 5) these patterns must be psychologically meaningful via frequent appearances (including repetition) as formative principles. The only concession must be made in the last argument: contemporary post-tonal music in particular is often composed based on logical necessities of the compositional system (not always taking the listener into account, as it is said about dodecaphony or serialism) or using chance or aleatorics as a compositional principle (e.g. Cage, Stockhausen and others) either not caring about the "significance of these patterns" for the listener or playing with patterns appearing at the moment of performance that are only partly pre-composed. However, the listener will go through a sense-making cognitive process searching for Almén's aspects 1–4 and also for impulses, salient elements (in atonal music, Lerdahl 1989, 2001), or other extractable cues (Deliège & Mélen 1997, Reybrouck 2010, Wiggins 2010). This leads to my following definition of 'Perceived structural musical tension' (PSMT) (first presented in Estonian in Lock 2022):[72]

> PSMT can be comprehended based on general cognitive phenomena (musical events as impulses) and processes (focal attention and listening, mental pointing) via a cognitive mechanism (salience) as a cross-domain concept (narrativity) connecting perception and cognition as a listening story based on structural elements/musical parameters.
>
> PSMT (like narrativity) is a dynamic principle with teleological impulses, musical events/structural patterns/a piece are actions in flux and a journey, a "real-time" discovery process/algorithm (requiring attentive listening, mental pointing) of musical events/parameters that built limited archetypical patterns that are in conflict on various hierarchical levels causing constant revaluation of these elements, assuming that the

---

[72] In Estonian, abbreviated as TSMP: 'Tajutud strukturaalne muusikaline pinge', see also in Lock 2023a; in English see publication V.

> unique surface features (cues, salient features) and effects of these patterns (salience principle) should be psychologically meaningful.

‘Perceived structural musical tension’ (PSMT) merges together dynamic and temporal, cognitive and narrative concepts based on focal attention and immediate experience of the perceiver, the attentive listener. The aim is to observe meaningful (limited archetypical) patterns based on emergent properties of Impulses and salient musical parameters (theoretical principles and empirical application see publication IV). Unlike Imberty’s (2000) “pre-definition” of *protonarratives* (musical events, vitality forms, tension-relaxation, see 2.6). PSMT defines the special object under observation – musical tension –, belonging to the general object of this dissertation – post-tonal orchestral music –, as “fully fledged” (Kraus 2013/2016) cross-domain definition and (abstract) narrative negotiating between perception (as sensing) and cognition (as understanding) (Gruhn 1992, see publication IV).

In the light of the complexity to pin down music and musical tension systematically and therefore scientifically (see sections 2 and 3 so far) I proceed in the following subsection with introducing the crucial methodological challenges for the research of musical tension I call the five scientific Dilemmas (see publication V).

3.6 Five scientific Dilemmas

In this subsection the five scientific Dilemmas are presented shortly but explained in detail in publication V. The Dilemmas pave the way for the five Research Questions (RQs) and the five Statements (STs) in publication V and constitute the methodological frame for this dissertation. Furthermore, the Answers to these Dilemmas/RGs/STs constituting the methodological solutions for this dissertation can be found in publication V.

Dilemma 1: Based on the understanding of music as a cognitive phenomenon (e.g. Wiggins 2009, Addessi 2010) I have also identified musical tension as cognitive based on listening[73]

---

[73] See also tension understood as an experience-shaping semantic characteristic interesting for cognitive science (Herremans & Chuan 2017).

using further the term phrase Cognitive Musical Tension (CMT). CMT is well commonly experienced (Farbood 2012, Lehne & Koelsch 2015), but complicated to describe and formalize (Addessi & Caterina 2000, Addessi 2010, Farbood 2006, 2012), it cannot be examined directly (see 1.2, 2.2, 3.1).

Dilemma 2: Intensity or energy measured at single time point values in the acoustic/physical domain (in relation to a technically absolute zero) is not sufficient and does not equal cognitive musical tension (CMT).

Dilemma 3: Instead of start/end points of a tensional development (CMT) in contemporary post-tonal orchestral music (CPTOM) only maxima and plateaux of a culmination "arch" of a temporal dynamic wave-like phenomenon (TDWP) can be detected via musical events as Impulses (Im-s)/"moments of change" as "relative change" as "Just noticeable difference" (JND) triggered by salience.

Dilemma 4: CMT as a TDWP can be detected only as relative change of to-be-defined musical parameters via attentive listening as intuitive experience, audiovisual salience, microgenesis, temporal dynamic forms and dynamic forms of vitality (see 2.2) applying a Kant-based early psychophysical perception/cognition mechanism detecting inner- via outer magnitudes (see publication V and 2.2). To achieve this detection mechanism, this dissertation develops an external, spatial and intersubjectively agreeable standard – a spatiotemporal object (based on a Kant-based understanding suggested by Kraus 2013/2016) called COSM: Cognitive Octagonal Slice Model and applied in publication IV.

Dilemma 5: The 'content' of musical events/Impulses (Im-s) are musical parameters (MP-s) that "create" "perceived intensity" – the role of "primary"/"secondary" MP-s parameters is unclear. In CPTOM mainly secondary musical parameters are hypothesized to cause (trigger) Impulses (Im-s) while enabling to comprehend CMT – this hypothesis has been confirmed only partly (see publication IV).

The Dilemmas border and narrow down the special object of investigation – musical tension – as 1) compound and cognitive via listening, 2) not measurable in relation to an absolute acoustic/physical zero, 3) not measurable based on clear start/end-points but as a temporal dynamic musical event detectable as Impulses or "moments of change" via salience, 4) detection of 'relative change' ("just noticeable difference" JND) of musical parameters using a external and intersubjectively agreeable spatiotemporal object COSM mediating audiovisual salience of inner and outer magnitudes, 5) being unclear what type of musical parameters (primary or secondary) is salient leading to the main study in publication IV.

The Dilemmas lay out the methodological problems and questions, and offer solutions to observe/analyze the compound phenomenon musical tension as cognitive proposing the psychophysical measurement principle of ‘relative change’ detectable via listening in Impulses and musical parameters via audiovisual salience mediated by a spatiotemporal standard object called COSM: Cognitive Octagonal Slice Model (publication IV) specifically investigating what musical parameters are more/less salient. This is the base for the verbalization of the main objective presented in the next subsection in the context of methodologically relevant terms: *purportedly*, *trigger*, the *why*-question, *segmentation*.

3.7 Main objective of this dissertation

The main objective of this dissertation is to detect, describe, make sense of and comprehend something in music – tension – that is compound and a whole in itself. Analogously, tension can be experienced also in everyday life (real-world) context. As mentioned in 1.2 and more in detail in 3.1 tension can be understood in several ways: as intensity and/or energy and/or as a temporal dynamic wave-like phenomenon (TDWP).[74] These aspects are known and researched in the physical dimension, e.g. mechanical, water and electro-magnetic waves. More concretely musical tension can be analyzed as an emotional aspect or as something beyond emotions, as a semantic characteristic of music or on the structural level as intensities and a process triggered by the interplay of musical parameters. However, I approach musical tension in the psychological realm in the context of the general psychological model of tension and suspense (Lehne & Koelsch 2015) as a phenomenon *beyond emotions*.

In this dissertation I examine *structural aspects* (musical parameters) observable in musical events I call *Impulses*/“moments of change” that purportedly *trigger* (via salience) the

---

[74] See publication V: Statement 1 of Dilemma 1, as the concept of a culmination “arch” of Dilemma 3 and Cognitive Musical Tension (CMT) in the Research Question 3, as well as CMT and TDWP in Dilemma 4; see also subsection 5.6.

experience of musical tension during *attentive listening*[75] (Deliège & Mélen 1997: 388) to contemporary post-tonal orchestral music (CPTOM). The means to achieve this objective involve defining as well as empirically detecting or measuring musical tension as 'Perceived structural musical tension' of CPTOM (see 3.1 and publications IV, V) by Erkki-Sven Tüür (b 1959) while observing and analyzing *Oxymoron* (2003, for large ensemble) and the 4th Symphony/Percussion Concerto *Magma* (2002) (characterization of his music and his attitude towards the importance of cognitive listening see in 1.1).

As follows I will briefly discuss two crucial terms being part of the formulation of the main objective – *purportedly* and *trigger* – in relation to a third aspect important in analysis and science: the *why*-question. This main objective deliberately applies the term *purportedly*, because science mainly deals with hypotheses and correlations that often do not explain direct causations. Both correlation and causation trigger *why*-questions but these are often rather complicated to answer. The cognitive musicologist David Huron (2006: 6)[76] underlines that "along with *what* and *when*, brains also predict *where* and *why* – but these are more specialized operations." Furthermore, Huron (2006: 7) explains that "the *why* expectations are associated with physiologically recent structures associated with conscious thought. In contrast to the *what* and *when* of prediction, the *where* and *why* components of auditory expectation have played little role in musical organization and experience. But they represent opportunities for future enterprising composers."

Alongside with the *what*-question – what musical parameters trigger via salience the experience of musical tension during attentive listening – this dissertation asks, *why* the intensity level of PSMT as a compound phenomenon measured and visualized as perception experiment curve data has changed in the "moments of change" called Impulses. Do musical parameters can answer this *why*-question: *why* the intensity of TSMP has increased, decreased or remained the same in Impulses/"moments of change"? The usual research approaches (theories and models) in music psychology – also in tension research (e.g. Lerdahl 2001, Krumhansl & Lerdahl 2007, Krumhansl & Lerdahl 2011, Farbood 2012) – often derive at prediction models in order to compare their outcome with gained empirical data. Instead, this dissertation deliberately *does*

---

[75] See also focal attention, mental pointing (Reybrouck 2010, 2015). The term phrase 'intentional act of listening' refers to Clifton's (1975: 70) term phrase 'intentional acts' to approach music's many experiential gestalts.

[76] Huron (2006) has developed the ITPRA: Imagination–Tension–Prediction–Reaction–Appraisal psychological theory of expectation that also includes tension.

not develop a prediction model.[77] It observes and analyzes solely the dynamic listening strategy as attentive (focused) intuitive listening[78] process of the experiment participants. While acknowledging that musical tension is a temporal dynamic wave-like phenomenon (TDWP) that includes in itself the process of expectation, I do not attempt to model the expectation process (via prediction) that goes with the increase and decrease of tension. I only look at the (in my view) rather detectable/measurable "moments of change" called Impulses (see publications IV, V) and try to grasp the *why*-question "in reverse."

The *why*-question cannot be answered sufficiently at the level of causation (only via interpreting correlation). The music phenomenologist Thomas Clifton (1975: 72) underlines that science determines the "mode of existence" of its object via the dialectic of analysis and synthesis. However, Clifton proposes that intuitive consciousness does not need this kind of scientific determination. For Clifton the object of consciousness is simply accepted as real and this acceptance is independent of the actual mode of existence of the object, the psycho-physiological conditions during the experience. In my opinion Clifton's way of thinking is clearly a form of Kantian *analytic idealism* (see Kant 1998, see 2.2), but for some reason he does not mention Kant in his article.

This main objective deliberately embraces also the term *trigger*. This is due to the following understanding of musical tension: although it is a temporal dynamic and therefore a durational phenomenon, it cannot easily be observed/measured as such (over time), because in Tüür's music the borders of formal units (not created or appearing in the traditional meaning of a theme or motif) are ambiguous or veiled (see publication I). The beginning of a durational development cannot be clearly detected (let alone be unambiguously defined), only some kinds of maximums (highpoints) and minimums (lowpoints) of the tensional development in Tüür's music can be more or less clearly detected (see the scientific Dilemmas in 1.5 and in publication V) as Impulses/"moments of change."

A detection of Impulses/"moments of change" is a kind of segmentation process. However, segmentation in music is complicated concerning score- and audio-file-based approaches segmenting complex avant-garde orchestral music (Guigue & de Paiva Santana 2018, de Paiva

---

[77] See several types of models in publication II, see also Lehne and Koelsch's (2015) general psychological model of tension and suspense.

[78] The term *listening strategy* was suggested in the doctoral seminars by musicology Prof. Dr. Urve Lippus (Head of the Musicology Department at the Estonian Academy of Music and Theatre) to characterize the outcome of participants using slider-controllers.

Santana & Guigue 2020, 2022), and combining music theoretical and perceptual empirical methods based on a cognitive understanding while analyzing musical tension and “macroform” in tonal and post-tonal/atonal (chamber) music (Addessi & Caterina 2000, 2005; Addessi 2010). In this dissertation I have abandoned to apply segmentation as an explicit concept, because 1) Tüür’s music shows no classical themes or motives, although, rhetoric rotational form analysis can be applied to detect its rather classical global (macro)form (Kotta 2011); 2) definable segmentation of Tüür’s music is rather impossible both in the score and via listening. In my view the concept of segmentation assumes clear pre-given formal units into which one can segment the music. As mentioned above, in Tüür’s music clearly bordered formal units are rare or even absent, so one has to depart from something that might or seems to have potential characteristics to qualify as a possible border: musical events detectable via salience as Impulses/”moments of change”, e.g. invariants (Clarke 2005, Imberty 2000) in the music understood as environment (publications III, IV, see 1.3) – looking for perceptual-cognitive phenomena/mechanisms (Impulse/“moments of change” via salience).

This dissertation contributes to the research of complex musical structures and musical tension from angles that so far approaches have considered partly or have not been used so far, see the need for investigating salience in tension research expressed by Farbood (2012) (see 3.1) and salience as a psychoacoustic descriptor to comprehend the perception of musical dimensions of complex contemporary orchestral music (see 4.1), descriptions of salience and attention (see 3.2), salience research in complex sounds and contemporary music (see 3.3), narrativity (see 3.5), the scientific Dilemmas (3.6), and finally presented as the novelties of this dissertation in the conclusion section 6. Furthermore, the question of the role (the importance via salience) of primary vs secondary musical parameters in such music has been observed in the main study of this dissertation (publication IV, musical parameters see in 3.4). These angles encompass methodological-philosophical thinking (intuitive thinking, Kantian analytic idealism, psychophysical measurement principles) that have been part of research (e.g. Clifton 1975) rather implicitly, unverbalized and in a marginal position, although these principles are clearly applied and recently also brought up explicitly (Kraus 2013/2016; Morales 2018, 2024). As mentioned in 1.2 and 3.1 musical tension is universally understandable but less investigated systematically and structurally. Salience (as suggested by Lerdahl 1989, 2001 for analyzing atonal music and Farbood 2006, 2012 for observing musical tension) seems to be the most promising perceptual-cognitive mechanism capable of offering solutions to the above introduced Dilemmas. Similarly to earlier research on musical tension of post-tonal music

(Addessi & Caterina 2000, Addessi 2010) I also do not use (computational) predictive modeling, because also my basic interest is to understand the mechanisms of listening in relation to music theoretical aspects and their mechanisms. The score analysis dimension (explicitly included by Addessi & Caterina 2000, 2005; Addessi 2010) I generally have abandoned after the first two publications (I, Ia, see App. C) and apply in the main study only to a small amount for back-mapping and comparison reasons (publication IV). Instead, in contrast to Addessi and Caterina, my cognitive modeling approach since publications III and IV deals solely with observing the actual process of the listening experience and contemplates rather on the empirical side – including how to gain and analyze such data. Although the empirical gained data could be analyzed from a qualitative and individual standpoint (to explicitly understand the listening strategies) this dissertation is restricted to analyze rather quantitative data (e.g. means of the samples and subsamples per background). In addition, with narrativity a cross-domain and therefore cognitive-science-realm way of thinking seems to be useful to extend existing methods of understanding such an universal and at the same time individual and intuitive phenomenon like musical tension. The following section 4 introduces the research methodology, design and modeling steps of this dissertation.

4. Research design, modeling and data analysis

This section presents research methodology matters and the design of the main study (publication IV), the applied principles of modeling (publication III) as well as some data analysis and visualization aspects (publication II) developed for and implemented in this dissertation. The main objective of this dissertation – analyzing musical tension (PSMT) of contemporary post-tonal orchestral music (CPTOM) via detecting, measuring and understanding structural aspects that purportedly trigger its experience during attentive listening (see 3.7) – is achieved with an empirical triangulation-type attentive-listening-experience research design (see 1.4, 4.1). Central to this dissertation is the development and application of my COSM: Cognitive Octagonal Slice Model.[79]

In subsection 4.1 some methodology related matters and examples concerning musical properties/parameters in tonal and post-tonal/atonal music as well as sound mass and complex orchestral music is presented. These approaches apply semantics and neuronal modeling as well as musical score and audio file analysis in relation to musical properties/parameters based on the graphemic, acoustic, and auditory domains of music (Babbitt 1965, 2003 in Wiggins 2009: 477, see also publication V). In 4.2 my research design is located within current research/analysis of (contemporary) orchestral music showing in what direction my dissertation extends and complements existing methodology. In 4.3 insight is given into the application and methodological principles of my two main data retrieval methods – **TEDEA: Tension Design Experimental Apparatus** (Lock, G. & Lock, H.-G. 2011) using slider-controllers (used to analyze Tüür's *Oxymoron* and the 4th Symphony/Percussion Concerto *Magma*), and **COSM: Cognitive Octagonal Slice Model** (used to analyze Tüür's 4th Symphony/Percussion Concerto). There one also finds explanations concerning the *psychophysical measurement* aspects (see 2.4) that are either directly or indirectly part of the data retrieval methods. In 4.4 the modeling steps for developing COSM are shortly presented and answers provided to the questions that are part of my Twelve Strategic Steps for modeling/analyzing scientific models (TSSM) (see publication III) that have supported me to develop COSM. In 4.5 some principles

---

[79] First presented in Lock and Kotta (2015), Lock (2017b) and Lock (2019), development principles see publication III, empirical application and detailed results see publication IV (see also Lock 2021, 2023a). See some overall results in Appendix D.

of visualization and automatization of “multi-dimensional time series” and pattern mining are presented that are relevant for data gained via perception-cognition experiments. In 4.6 some data analysis principles (cluster and pattern finding) are presented concerning the DBSCAN: Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise clusters (Ester *et al* 1996) applied in this dissertation. The next subsection continues with chosen research methodology matters and examples.

4.1 Research methodology matters and examples

Recalling the first subsections of the Introduction (1.1, 1.2) and the five scientific Dilemmas (3.6) contemporary post-tonal orchestral music (CPTOM) and ‘Perceived structural musical tension’ (PSMT) are complex and rather difficult to study systematically. The general aim of this dissertation (1.2) – to advance analytical methods to enhance the understanding of CPTOM that poses challenges to the analyst – calls for technological-computational assistance in data gaining and analysis up to different modeling approaches (e.g. Addessi & Caterina 2000, Schüler 2000a&b, Dubnov & McAdams 2006, Schüler 2007, Dahan 2012, Friberg *et al* 2019, de Paiva Santana & Guigue 2020, Noble *et al* 2020, Onwuegbuna 2020, Antunes *et al* 2021/2023). It seems unavoidable to apply computerized aids to handle a high amount of data retrieved in either of Babbitt’s (1965, 2003) three domains of music (Wiggins 2009, publication V): the *graphemic* (notated), the *acoustic* (physical) and the *auditory* (perceived) *domains*. In the next paragraph the graphemic domain (notation) aspects in some studies on tension and “macroform” are presented. Examples of research designs in the *acoustic* and the *auditory domains* see several model-based approaches to analyze and predict musical tension see in 3.1, computational and neuroscientific model-based approaches to analyze auditory salience see in 3.3, studies of sound mass/contemporary orchestral music using computational means see here in 4.1 below.

For tension research of tonal music the graphemic domain (notation) has been taken as a basis by Herremans and Chuan (2017: 1) claiming that in order “to truly understand semantically meaningful concepts such as tension, relaxation, and closure, it is necessary to consider the manner in which music theorists and musicologists study music, i.e., to study the score.” For harmony, roughness and tension research of neoclassical music including post-tonal harmonies (Hindemith, Ravel, Stravinsky) the *graphemic domain* (notation) has been used by Teo (2020) to compare the results of the empirical online-survey study with the music theoretical analysis of the works. For "macroform" and tension research of post-tonal and atonal music (Kurtág, Maderna) the *graphemic domain* (notation) has been taken as the object of one of two empirically juxtaposed methods by Addessi and Caterina (2005) and Addessi (2010): in both studies a group of expert music theorists have analyzed the music based on the score and via listening,[80] and the most fitting versions have been taken as comparison against the results of empirical listening tests with listening-only conditioned musicians' and non-musicians' participants. The results of Addessi's (2010: 229) study while applying supportively perception (auditive analysis) and theory of music (score analysis) revealed that there could not be found a particularly significant difference between the analytical and listening procedures for score analysis conducted on explicit criteria based on theories using perceptual aspects of sound (especially the *cue* abstraction theory by Deliège 1989, 1996, 2001, 2007). Addessi (2010: 232) took "into account [the professional analysts'] personal listening experience" underlining that "the analyst should nevertheless be aware of the perceptual difficulties and problems related to memory encountered by a listener and his/her task is in fact precisely to render these problems explicit (see also Cross 1998[a])." In contrast to the auditive analysis procedure (with a limited amount of "in-time" [see Reybrouck 2019] listening tasks) the professional analysts were enabled to listen to the music repeatedly as much as they liked, the score was taken as a helping medium (Addessi 2020: 233) in order to gain also an "out-of-time" [see Reybrouck 2019] understanding of the music). However, Addessi's (2010: 233) "objective was not to use theoretical methods of analysis of the macroform, but to hypothesise analysis criteria of the macroform as close as possible to those of a listener."

In the context of this dissertation I understand music as environment (see 1.2, 1.3 and publications III, IV) and a cognitive phenomenon (see 1.3, 2.5 and publication V) and first of all comprehensible via human listening based on perception-cognition acknowledging the power of the human ear and conceptualizing abilities (Wang 2007: 164, see 4.4 Q 8 (E)). Secondly, music cannot be solely understood in neither the graphemic, physical nor the

---

[80] A similar mixed score-based and listening approach conducted by myself (see publication I) I decided to abandon (see App. C) because of its ambiguous 'speculative' nature (see 1.2). However, similarly to Addessi's above shown approaches I still have included the results of one expert music theorist's system-based analysis (see publication II) as comparison with results of my (own developed) empirical methods (see next subsections) – the rhetorical/rotational form analysis by Kotta (2011, based on Hepokoski & Darcy 2007, see publication IV).

perceived domains separately. Furthermore, music analysis – if it is not just statistically analyzing MIDI or via other automatized notation/representation data (possibilities see in publication II) –, is first of all conducted by humans based on music theoretical methods located between art and science (Broman & Engebretsen 2007). In contrast to purely "behavioristic" studies in music psychology the research method examples shown above give the human analyst a strong and decisive role both while analyzing the musical score and relying on the process of listening (see music as cognitive, see 1.3, 2.2, Kant 1998, Clifton 1975 and publication V). Technological-computational means have been applied to gain data in various consecutive tasks during the experiments while allowing repeated listening partly as much as needed (score-based and listening tasks by experts) and partly with a limited amount of listening only tasks.

As follows I introduce two studies on post-tonal orchestral music where human activity is still a deliberate part of the research design. The reason I underline this is because computational tools should be seen as a supporting component of music analysis (Dahan 2012), a computer program (and automatization, see publication II) does not do everything on its own, its outcome always depends on the context and human interpretation in accordance with the aims of the analysis. Furthermore, computer algorithms are programmed by humans and do what has been programmed – this even applies so far for self-learning AI (Kaipainen & Tikka 2022). The human component (participants in perception tests, analysis via cross-domain concepts) also plays a crucial role in the research design of my dissertation (see 4.2, 4.3, 4.4).

The first example is a study by Noble *et al* (2020: 214) on the semantic dimension of sound mass music[81] where a mapping between perceptual and acoustic domains has combined human ratings based on batteries of semantic scales, its statistical analysis revealed a coherence in "cross-domain mappings between semantic categories and musical properties" among participants "implying non-arbitrary relations." While a correlation between semantic ratings provided by human participants and classical audio descriptors (see several chapters on the topic of MIR in e.g. Polotti & Rocchesso 2008) was not detectable, results via "a neuromimetic representation called spectrotemporal modulations (STMs)" ("mimic sound processing in the cochlea and the primary auditory cortices" (Noble *et al* 2020: 224), allowed the authors to suggest that "semantic interpretations of music may involve qualities or attributes that are objectively present in the music, since computer simulation can use sound signals to partially

---

[81] Orchestra, chamber or string orchestra, choir, organ, string quartet or electronic/amplified works (excerpts) by Xenakis, Lutosławski, Penderecki, Ligeti, Tormis, Stockhausen, Grisey, Górecki, Takemitsu, Harvey, Dhomont, Risset, Truax, Wishart, Saariaho, Hurel, Roy, Haas, Bedrossian and Normandeau composed between 1953–2009.

reconstruct human semantic ratings" (Noble *et al* 2020: 214). This approach compares pre-formatted human ratings with a computer model based on predicting (simulation and reconstructing) of the same aspect: semantic categories in relation to musical properties.

As a second example Didier Guigue and Charles de Paiva Santana (2018) and Charles de Paiva Santana and Didier Guigue (2020, 2022) analyzed complex contemporary orchestral music[82] based on the musical score (graphemic domain) and audio files (acoustic domain) applying manual segmentation of the texture in a complex musical score (cataloging and representing the musical parameters' data in a matrix) and analyzing the retrieved data computationally (with SOAL, Sonic Object Analysis Library based on CSU, Compound Sonic Unit).[83]

Guigue and de Paiva Santana (2018: 102) underline that their "experimental model for computer-assisted analysis proposes a formal strategy for evaluating the role of orchestration in musical structure and perception. It works by looking at the symbolic level of the score's prescriptions and the "perceptual level" of the performed music" (for details see Figure 1 in 4.2). However, they admit that "the audio branch of our model [...] is at a very incipient stage: choosing, importing, developing and implementing appropriate descriptors is not a trivial task and demand further extensive research and experiments. Furthermore, the impact of an LSS' duration[84] upon the overall musical perception should be properly addressed in the next stages of our work" (Guigue & de Paiva Santana 2018: 103). They later confirm that "information which is easily accessible in a musical score or musicological transcription may be very difficult to get by listening to a performance alone or through computational means with sound files" (de Paiva Santana and Guigue 2020: 121). However, they propose to combine LSS (Local Sonic Setup) data from the musical score with data from the corresponding LAU (Local Audio Unit) while extracting audio features like spectral centroid, zero-crossing rate, spectral roll-off, MFCC (Mel-frequency cepstrum coefficients, a mathematical coefficients for sound modeling), etc whereby "a number of numerical descriptors for LSS and LAU are combined to achieve a measure of relative complexity of orchestration and musical texture" (de Paiva Santana & Guigue 2020: 122). De Paiva Santana and Guigue (2022: 146–147) further have applied

---

[82] Orchestra and opera as well as chamber ensemble works (excerpts) by Rameau, Beethoven, Schoenberg, Webern, Xenakis, Penderecki, Pascoal from the 18th to the 20th centuries.

[83] CSU is "the combination and interaction of musical primary and secondary components: the former refers to collections of pitches and the latter to aspects as intensities, ranges, registers, densities and so forth. We also recognize as secondary components statistical measurements such as distributions, deviations, and entropy, among others" (Guigue & de Paiva Santana 2018: 101).

[84] Segmentation of the musical score via LSS (Local Sonic Setup), its audio file equivalent segmentation is called LAU (Local Audio Unit) (Guigue & de Paiva Santana 2018: 98).

IRCAM’s Audiosculpt Chord Sequence Analysis function to calculating the average spectrum for LAUs and further analyze them in OpenMusic in order to compare the composer’s prescriptions (in the score) with performance audio data. Psychoacoustic descriptors to approach the perception of musical dimensions, e.g. timbre or salience, are the aim of their further research, but “the problem lies in the quantity of descriptors available, and the heterogeneity of the information produced,” taking into account that “the basic concept of a descriptor is to “flatten” a multi-dimensional component, the stamp, on a single dimension” – they offer “to constitute a modular network of descriptors that would be used according to contexts, and to standardize their results on the same vector of relative complexity to which the other components of the model have been subjected” (de Paiva Santana & Guigue 2022: 148).

Guigue’s and de Paiva Santana’s approaches – extracting data from the musical score and applying automated spectral and other features’ analysis on audio files – are a realization of what I also had tried out early during the development of methods for my dissertation. However, in the present stage of my dissertation I do not apply either of them in a systematic manner, because I am not a computer programmer and, despite knowing a number of possibilities, feel unsure about appropriate tools and methods for further data analysis. This skepticism is confirmed also by de Paiva Santana and Guigue’s (2022: 148) description of the difficulties they still encounter in their approach. In a similar way Farbood (2012) has modeled musical tension based on score-based calculated musical parameters (e.g. pitch, rhythm and harmony, the latter based on calculations by Lerdahl 2001) as a prediction model to compare it with empirical data. However, in the main publication IV of this dissertation I generally have abandoned a systematic score-based approach, because the localization (both exact pinpointing and approximate bordering, compare with the LSS and LAU concepts of de Paiva Santana and Guigue) of the salient features (developed and presented in publications I, Ia) in respect to the bar structure of the music (in relation to the audio file) remained in my view speculative (see 1.2) and were neither exactly measurable nor reasonable (see the scientific Dilemmas in 3.6 and in publication V). I also did not want to create another prediction model of musical tension in relation to musical parameters etc based on in my understanding limited input data derived solely from the musical score (see publication IV).

In general my research design (see more in 4.2) is similar to Farbood’s (2012: 389) two different data retrieval approaches to explore musical tension in relation to musical parameters: 1) continuous, online or real-time data-capture, and 2) discrete, retrospective judgments. In general also Addessi and Caterina (2000, 2005) and Addessi (2010) have used a similar design: 1) computer-based data capture, including "continuous" tension/relaxation ratings (via computer mouse[85]), and 2) retrospective verbal judgments (and in the case of musical analysts professional score-based analysis). For this dissertation I have ordered (by Hans-Gunter Lock) and applied two experimental hard- and software solutions. This equipment has enabled three empirical data capture methods using slider-controllers (sound mixer slider-controller, hard- and software called TEDEA since 2011) as well as a cognitive application of Impulse and salience detection (software called COSM since 2017) (see 4.2, 4.3 and 4.4). The graphemic (or notated) domain in form of analysis of the musical score I have integrated while comparing my results with that of the rhetorical/rotational form analysis by Kotta (2011) (see publication IV). Furthermore, I have mapped chosen results of clusters and tension points to a *particell*[86] of Tüür' works (see publications Ia, II, IV).

As follows I locate and identify the methodology of TEDEA and COSM within previous research, the triangulation principle, and modeling and the Conceptual Blending Theory (CBT, Antović 2018, 2022).

4.2 Design, context and modeling aspects

Generally my research design is based on the *triangulation principle* (Denzin 2012, Heesen, Bright & Zucker 2019, Arias Valencia 2022/2023, Kaipainen & Tikka 2025, forthcoming), but I do not use three distinct methods from three different realms. I basically apply two methods of empirical data-retrieval: TEDEA slider-controllers and COSM salience detection. Slider-controllers (or different types of dials or computer-screen-based applications) for data capture have been used for quite a time already for tension (and more widespread for emotion) research (see Farbood 2012, Schubert 2010). TEDEA methodology also fits behavioristic standards, because the response expected does not require the participants to apply explicit pre-knowledge to analyze their response to the musical tension as a compound and multilayered phenomenon. They are asked to react like a "black box" getting triggered by the input from the music as an

---

[85] Technical solutions to capture "continuous" ratings of musical tension include software-based computer-mouse, special hardware (several types of dials) and sound mixer slider-controller (an overview see Farbood 2012).

[86] A condensed score and fragments of the score provided by Kerri Kotta.

environment (no need to use knowledge/cognition of music/theory-specific concepts) and respond immediately via hand movement to a continuous stimulus without interruption possibility. COSM I have developed myself (Lock & Kotta 2015, Lock 2017b, Lock 2019, publication IV, Lock 2021, 2023a), it contains two different, but interconnected methods based on the detection of Impulses/"moments of change" via salience and salient musical parameters. Results of a third, distinctive, method, a rhetorical/rotational analysis (Kotta 2011, based on Hepokoski & Darcy 2007) of especially the symphonic music by Tüür I have added in publication IV for juxtaposition aims. However, I neither have developed nor conducted this method myself – therefore I cannot claim it to be part of my own original research design. The methodological principles of modeling COSM based on my Twelve Strategic Steps for modeling/analysis of models, (TSSM) are explained in publications III and first results are presented in publication IV. Modeling aspects in a wider methodological context see below in this subsection.

In short, **Experiment I TEDEA: Tension Design Experimental Apparatus** uses slider-controller "continuous"[87] data capture hard- and software[88] to analyze "intuitively" perceived musical tension (holistic approach, attentive listening). **Experiment II COSM: Cognitive Octagonal Slice Model** aims at detecting in **part 1** musical events as Impulses ("moments of change") via their salience (holistic approach, attentive listening); in **part 2** detecting the salience of musical parameters (attentive listening, analytical listening) within a 14 sec chunk/window (+/–7 sec) around each Impulse (Im) detected in part 1 by each participant

---

[87] The data capture method is "continuous" only metaphorically (computers do not calculate continuously, they work on a discrete basis). The Max/MSP software for TEDEA records data points every 500 milliseconds and repeating data points are not recorded. This was necessary to lower the amount of data that has an effect on the speed of the recording and on the analysis of the data. See more on "continuous" data capture methodology in Schubert (2010).

[88] The TEDEA: Tension Design Experimental Apparatus (Lock, G. & Lock, H.-G. 2011) was built by Hans-Gunter Lock for my dissertation in order to enlarge the sample size of participants in listening tests. It was planned for up to 16 slider-controllers, but the system has worked stable so far with up to 10 slider-controllers. It consists of Arduino technology and works with Max/MSP programming based on the Cycling 74 object-based programming environment. The Max/MSP-based software component was built and applied already for earlier studies related to and part of my dissertation. The development of Tension Design (TD) was abandoned, because the proposed ideas related to a possible TD method introduced in a conference publication (Lock & Valk-Falk 2008) could not be developed further (they were not considered in the dissertation seminars) and the development of the dissertation took a different direction with my publication I towards salient features analysis instead. The idea to also explore the role of musical parameters in musical tension was empirically tested already during the first experiments on Tüür's *Oxymoron* in Lock and Valk-Falk (2008), but these results were too imprecise and problematic for further treatment and analysis. Its methodology remained highly questionable, but triggered a long thought process that took concrete shape since Lock and Kotta (2015) in the development of COSM: Cognitive Octagonal Slice Model. It initiated the theoretical modeling article (publication III) and the empirical experiments in the same year. Concrete theoretical background and explanation of the computerized COSM together with cluster finding using the DBSCAN (data mining) algorithm (Ester *et al* 1996) as well as comparative analysis of the empirical data and their narrative interpretation were finally issued in publication IV.

individually. These chunks can be listened to repetitively by the participants as much as they need (compare this with Albrecht's 2012 chunking methodology, see more in publication IV; and with the methodology by Addessi & Caterina 2000, 2005; Addessi 2010). Details and the technical steps to apply the TEDEA/COSM experiments (including activities by the participants and by the researcher) see in Appendix B.

The participants of the experiments have been previously informed[89] about the phenomena they are supposed to detect. In **Experiment I** (N=26) they have been asked to respond by moving the slider-controller up/down when the perceived musical tension is in-/decreasing or keep it in the momentary position if nothing changes. They have been informed to not immediately (from the beginning on) use the whole slider-controller range, because they may encounter later even higher culmination points in the music. In **Experiment II** (N=14) the participants have been informed about the terms *Impulse* (Im) and *salience* in music context: I call an Impulse (Im) a clearly perceivable and relatively salient short and local musical event, that stands out (protrudes) from other events in a certain manner (e.g. in timbre, pitch height, dynamics). This can be a single musical element, a sound, voice or parameter (e.g. in the beginning of the piece or a formal part of the piece, if relative calmness is dominating), that changes, accrues or disappears. These can be clearly outstanding sounds by percussions, a characteristic motif that is heightened by dynamics and register, a tutti chord etc.

In order to locate my own cognitive approach within previous methodology and concepts I have compiled two comparison schemas enlarging Guigue and de Paiva Santana's (2018: 99) mainly score-based analysis model (but including also audio file analysis) with concepts of Cognitive Complexity of Analysis and Listening (Figure 1) and the activities of my research design (Figure 2).

---

[89] Similarly, Addessi (2010: 230) informed their participants of the nature of the tension (inhaling) vs relaxation (exhaling) zones she expected them to detect. This may be understood as influencing (biasing) the participants with pre-information. However, both Addessi (2010) and this dissertation are not classical behavioristic studies solely observing the participants themselves like "black boxes" – instead, both methods rather treat the participants as co-analyst.

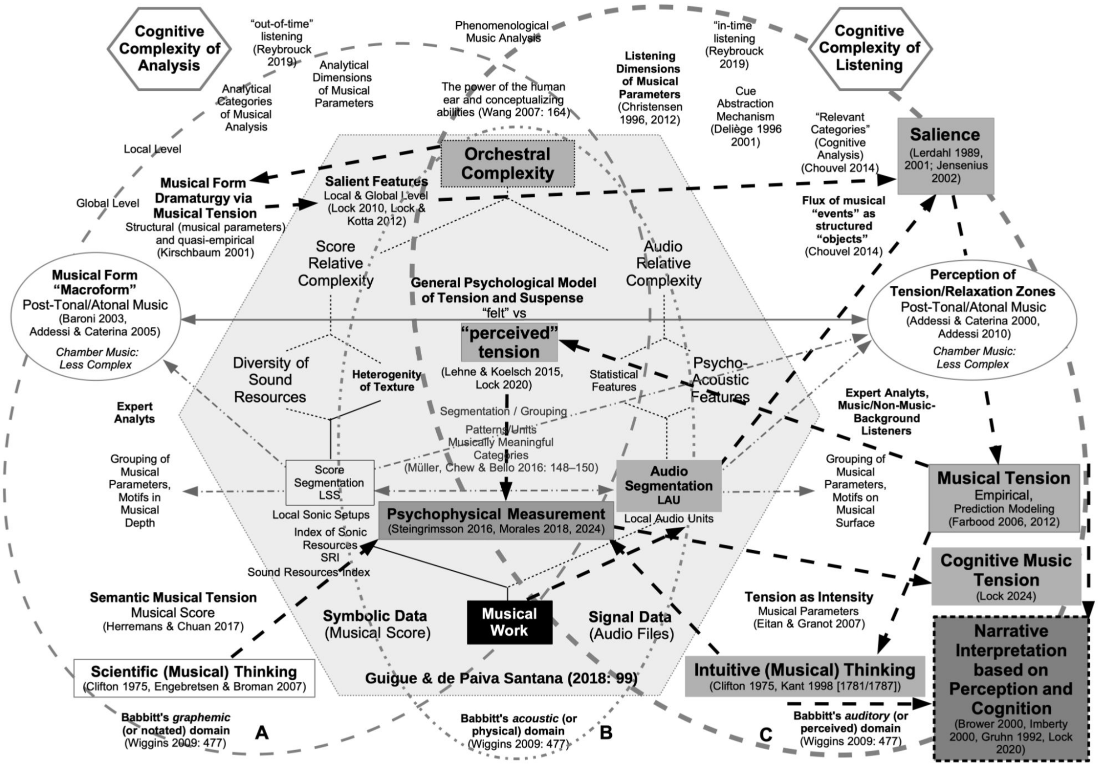


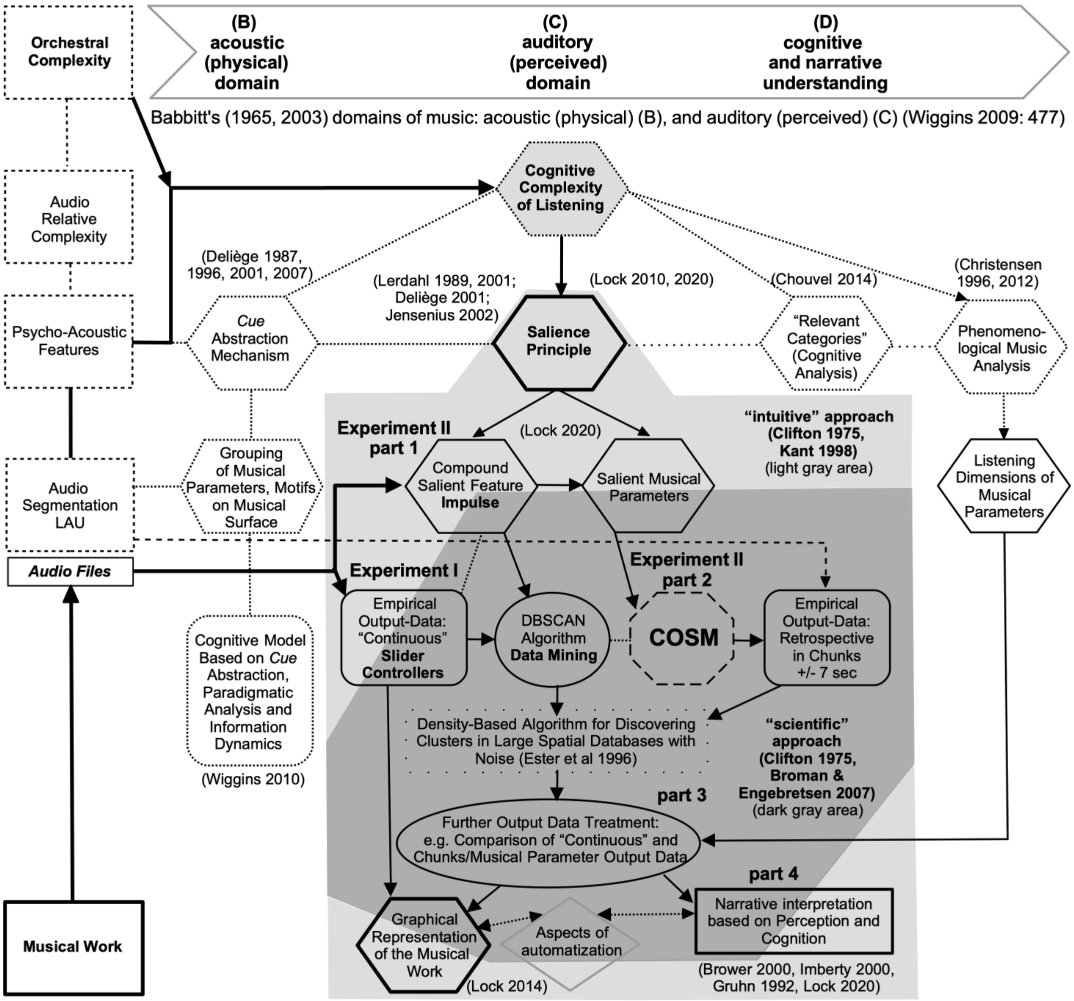
Figure 1. The Musical Work, Orchestral Complexity and Cognitive Complexity of Analysis and Listening. This scheme enlarges Guigue and de Paiva Santana's (2018: 99) model for computer-assisted analysis of orchestration (light gray hexagonal area in the middle of the scheme). The dashed arrows show the concepts that are dealt with in this dissertation.

In Figure 1 methodological approaches within Babbitt's (1965, 2003) domains of music are presented as three large and merging ellipses: graphemic (notated) (A), acoustic (physical) (B), and auditory (perceived) (C) (Wiggins 2009: 477). Guigue and de Paiva Santana's (2018: 99) model located in the light grey hexagonal area takes into account symbolic data (musical score) and signal data (audio files) and is located in the middle of Figure 1. My enlargements are placed within and around this hexagonal middle shape. The hexagonal boxes on the top of the scheme show the Complexity of both Listening (right side) and Analysis (left side) in relation to Guigue and de Paiva Santana's scheme on Orchestral Complexity. On both sides of the scheme there are studies in ellipses that partly jointly analyze "Macroform" and Tension Relaxation Zones of Post-Tonal/Atonal Music (Baroni 2003, Addessi & Caterina 2000, 2005; Addessi 2010) based on listening experiments and music analytical/theoretical understanding of experts. The concepts in darker grey rectangles are those that I have concentrated on in this dissertation. The wider black dashed arrows between them show the trajectories I, roughly speaking, have taken from Orchestral Complexity via Musical Form Dramaturgy via Musical Tension (structural (musical parameters) and quasi-empirical, Kirschbaum 2001) at the left side of the scheme via Salience (Lerdahl 1989, 2001; Jensenius 2002), "perceived" tension" (Lehne & Koelsch 2015, publication IV), psychophysical measurement (Signal Detection Theory, SDT, Steingrimsson 2016; introspective SDT aka iSDT Morales 2018, 2024), Intuitive (Musical) thinking Clifton 1975, Kant 1998) and Cognitive Musical Tension (Lock 2024) to a Narrative Interpretation based on Perception and Cognition (Brower 2000, Imberty 2000, Gruhn 1992, Lock 2020). This scheme includes concepts and their relations that I touch only partly and do not expand on.

The following Figure 2 accompanies Figure 1 while also taking as the basis Guigue and de Paiva Santana's (2018: 99) model. Figure 2 is an expansion of their audio signal data side into the cognitive complexity of listening. It is a schematic overview of the research design of the main study (publication IV) of this dissertation locating its activities applied in the experiments:

Experiment I – TEDEA (slider-controllers)**:** "Intuitive" perceived musical tension detection process with slider-controllers ("behavioristic" and "intuitive" consciousness approach – term inspired by Clifton 1975) [Participant activity], inspirations see 4.4, activities see Appendix B.

Experiment II – part 1: Impulse ("moment of change") detection – "primary categorization" (term inspired by Rosenthal 2004: 224) [Participant activity]; **COSM part 2:** Impulse 'content' (musical parameters' salience) detection – "local discrimination" (term inspired by Rosenthal 2004: 224) [Participant activity], inspirations see 4.4, activities see Appendix B.



Fragment of the audio signal data side by **Guigue & de Paiva Santana (2018: 99)**

Expansion of the audio signal data side into the cognitive complexity of listening. Schematic overview of the research design of the main study (Lock 2020) of this dissertation.

Figure 2. Design of the experiments and data analysis of this dissertation. This scheme enlarges the audio signal data side Guigue and de Paiva Santana's (2018: 99) model for computer-assisted analysis of orchestration towards the Cognitive Complexity of Listening Salience, cognitive and narrative approaches.

Similarly to Figure 1 also Figure 2 departs from Babbitt's (1965, 2003) domains of music (here located in the wide arrow on the top of the scheme: *acoustic* (physical) (B), and *auditory* (perceived) (C) (Wiggins 2009: 477) enlarged with a fourth domain dealing with a cognitive and narrative understanding (D). The hexagram shaped and rounded (dotted or solid line) boxes show principles and examples within which I locate my own approach. Dotted lines show connections that exist between concepts but that I do not apply in my design. Solid lines with arrows indicate the concepts and trajectories of my experimental design. I differentiate "intuitive" (light gray) (Clifton 1975, Kant 1998) from "scientific" (dark gray) (Clifton 1975, Broman & Engebretsen 2007) approaches. Experiment I is both an "intuitive" and a "scientific" approach (fitting also "classical" behavioristic" approaches), because the listeners respond to the music intuitively (not exactly knowing what one is doing, similar to a black-box research design) but the slider controller enables data output for further analysis with scientific methods (e.g. cluster finding with data mining algorithms). In Experiment II part 1 the listeners also respond intuitively, the data output is the basis for part 2 and data analysis preparation via clustering. Part 3 (output data treatment) and part 4 (cross-domain narrative interpretation based on perception and cognition and the PSMT definition see 3.5) are partly explained in 4.6 and extended with some general data in Appendix D.

COSM as a cognitive model (see publication III) is in the following Figure 3 put into a wider methodological context. The three-part relationship of modeling (Brodie *et al* 1994 and Duit 1991, juxtaposed in Coll & Lajium 2011: 3–4) combined with the four-part relationship of the Conceptual Blending Theory (CBT, Antović 2018, 2022) includes now in Figure 3 aspects of the general object of this dissertation (CPTOM): basic terms and principles applied in COSM: musical tension; salient musical features of musical events: Impulse (Im), musical parameters (MP's) and their salience, the wave-like character of Erkki-Sven Tüür's music (as a TDWP). Salient musical features appear during the listeners' intuitive attentive/conscious perception of musical structures. Impulse/"moment of change" is the chosen salient musical feature (publications I, Ia) under observation (publication IV) using principles of "psychophysical measurement" (see 2.4, publication V). Musical parameters and their salience are the 'content' of each Impulse (Im).

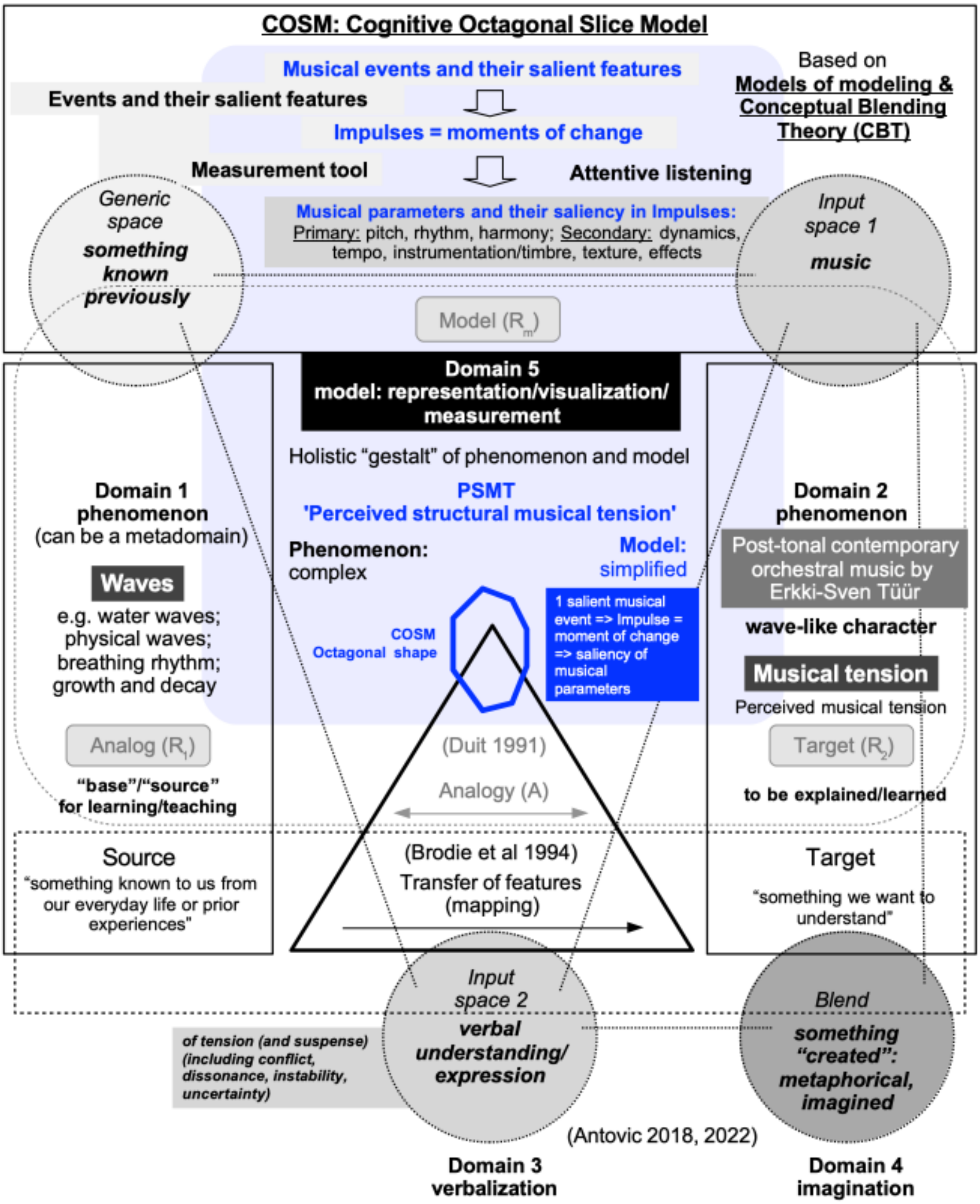


Figure 3. Modeling aspects of COSM in the context of models of modeling and Conceptual Blending Theory (CBT) (Antovic 2018, 2022). Figure 2 includes information on musical tension. The four-part relationship of CBT is merged with the three-part models-of-modeling schemes, as four-part relationship joined with the Models of modeling combining Brodie's *et al* (1994) (black information embraced by the dotted black oval) and Reinders Duit's (1991) (gray information embraced by dotted gray the smooth corners rectangle) three-part relationship modeling schemes (juxtaposed in Coll & Lajium 2011: 3–4, Duit 1991: 649–650, Brodie *et al* 1994).

The following subsection gives insight into the methodological principles of TEDEA and COSM, it recalls relevant theoretical background aspects and gives basic information on the main Experiment session conducted in 2017 (participants, technology) as well as the principles of the visual interface of the COSM slice based on Gestalt principles. Furthermore, “idiographic” vs “nomothetic” research approaches as well as lab-based controlled stimuli vs “ecological” music listening principles are discussed to explain the reason for using “attentive listening” (Deliège & Mélen 1997: 388) to observe complete musical works as stimulus.

4.3 Procedure of the data retrieval

In this subsection the methodological principles of the research design of this dissertation is presented. The two main data retrieval methods are **TEDEA: Tension Design Experimental Apparatus** (Lock, G. & Lock, H.-G. 2011) to detect perceived musical tension using “continuous” data capture technology (slider-controllers)[90] on Tüür’s “Oxymoron” (N=7)[91] and the 4th Symphony/Percussion Concerto *Magma* (N=6,[92] N=26[93]), and **COSM: Cognitive Octagonal Slice Model** (Tüür’s 4th Symphony/Percussion Concerto N=14). COSM principles were first introduced in the paper presentations Lock and Kotta (2015, manual version) and Lock (2017b, computerized empirical version). The final methodology together with first data results and their analysis with the DBSCAN algorithm (Ester *et al* 1996) has been issued in detail in publication IV (see also paper presentations Lock 2019, 2021, 2023a). With COSM one can detect the salience of the ‘content’ (musical parameters, MP’s) of Impulses (Im-s) in experiment **part 2** chosen by the participants themselves in **part 1** (Impulse detection via

---

[90] Data capture is still discrete, every 500 milliseconds, repeating data points are not recorded.

[91] The visualized data of this first experiment (N=7, 3 experts) of this dissertation (using a Behringer slider-controller) see in publications I, Ia, II. My earlier attempts to analyze this data during submitting to doctoral studies see in Lock and Valk-Falk (2008).

[92] Data fragments of a follow-up experiment (conducted 2010, still using a Behringer slider-controller) on Tüür’s 4th Symphony/Percussion Concerto *Magma* (N=6, 3 experts) are shown in publication Ia (see 5.2) to demonstrate a potential combination of slider-controller perception curves reductions with a 15 seconds “moving window” principle. This was later abandoned in favor of DBSCAN, see 4.6. Details of the 2010 data and participants together with another data-analysis try-out using trend/regression analysis in comparison with Kotta’s (2011) rhetoric/rotational form analysis, see Lock (2019/2011). Preliminary results combining 15-sec.-windows curves’ reductions, trend/regression, rhetoric/rotational form analysis and salient features’ analysis were presented also in an unpublished presentation (Lock 2012). See also 4.4 FN 105.

[93] The data of this main experiment using 10 (out of 16) smoothly working slider-controllers of TEDEA was gained in 2017 and is analyzed/visualized in publication IV.

salience). Except for the 4 experts, Tüür's 4th Symphony/Percussion Concerto was unfamiliar to the rest of the sample: n=22 for the TEDEA (BA level music and art students) and n=10 for the COSM (BA level music students only) experiment. The number of participants has been rather small. For the TEDEA experiment data (gained in 2017) I tried to reach a higher number to achieve a kind of "nomothetic approach" (Harris 2003/2008). For the COSM experiment data (gained 2017, see publication IV) this number is again smaller due to the restriction of the design to experts and musicians and the time-consuming structure (two-part experiment, repeated listening) that qualifies this experiment rather as a kind of "idiographic approach" (small-N research design, Saville & Buskist 2003/2008).[94] According to Richard Harris (2003/2008: 41), a pure distinction between "idiographic" and "nomothetic" research approaches is not realistic, because 1) "idiographic" results of single (or small collections of) participant(s) without an attempt to generalize is not scientific; 2) "nomothetic" results from large groups of participants and the analysis of relations between groups leave the individual participant out, this can be done in sociology or economics, but not in psychology.

> "In short, it is probably more necessary to remind "traditional nomothetic" researchers of the need to include individual-level analyses and explanations than it is to remind idiographically inclined researchers of the need to consider the generalizability of their findings across individuals. After all, the latter's audiences will test that generalization, whether the researcher does or not." (Harris 2003/2008: 42)

Generally, the data in this dissertation are analyzed based on numerical results ("quantitatively"), not focusing on single participants' data interpretation ("qualitatively"), although TEDEA and COSM data would enable a "qualitative" treatment, especially while having included narrativity into data analysis. However, such a "qualitative" approach would require different kinds of analysis methods and go beyond the scope and volume of the present dissertation.

The methodological data retrieval tools (TEDEA slider-controllers and my developed new

---

[94] Small-N research design or behavior analysis deals with single or few participants "under tightly controlled experimental conditions in which the independent variable(s) is repeatedly manipulated over successive trials or conditions and in which the dependent variable(s) is repeatedly measured" (Saville & Buskist 2003/2008: 66). However, the research design of my dissertation (see in this section) differs from these traditional laboratory settings and aims due to 1) the use of whole musical works as stimuli, 2) a rather ecological research attitude during the listening tests (Lock 2011a), 3) repetitions of self-chosen chunks by each individual participant (window around Impulse) that have neither been forced-choice nor pre-given by the researcher, 4) no manipulation of the stimulus has been done by the researcher, and 5) the aim of this research to mainly analyze the stimulus (music), although the data also enable to say something about the participants.

cognitive attentive-listening model COSM) presented and applied in this dissertation enable the listeners (participants in empirical experiments) to respond to the music intuitively (Clifton 1975), bodily ("bodiliness" see O'Regan, Myin & Noë 2004: 106–107, "muscular response" see Addessi & Caterina 2000: 34, "motor responses" in relation to dynamic acoustic salience see Schultz, Brown & Kotz 2021) and without words while using slider-controllers and the audiovisual COSM shape. This takes away the linguistic and terminological constraints and barriers that often appear with verbal and linguistic descriptions (Hendry & Psillos 2007). Hence, these tools turn the intuitive listening experience into data that is then visualized, analyzed and interpreted by the researcher. Based on Clifton (1975: 74) this interpretation should on the one hand take into account the singularity of the music and musical tension, and look for its substantiation. On the other hand it should balance between the singular thing (musical work) and the (musical) event (within the work), and its general causal conditions, as well as transferring its thinghood and its cognitive experience to pure functional relations (elements/parameters causing musical tension).

The narrative level, created either spontaneously (*ad hoc*) – verbalizing the intuitive listening experience – or systematically or using system-based or even automatizing tools (see publication II) – turning the intuitive listening experience into data – enables us to communicate both the experience and the scientific data. I claim that the data retrieved with the methodological tools (slider-controllers and COSM software) are themselves narratives (intuitive listening strategies and listening stories [*Hörgeschichten*]) (Gruhn 1992, see publication IV). Furthermore, creating and applying models is also a narrative activity (Rothbart 2004), and cognitive science as a field explains (and narrates) human cognition based on models, especially when scientifically explaining behaviour with representational content calling for the need "to get inside that explanation to see how it works" (looking at causal underpinnings, Shea 2018: 29). Therefore, it seems reasonable to interpret the empirical data derived in this dissertation based on narrativity principles (see 2.6, 3.5, publication IV). However, I am aware that the amount of data used in the dissertation rather shows the particular and substantiates musical tension of first of all one chosen composer's (Tüür's) music; it hardly enables yet to generalize causal conditions or pure functional relations. This dissertation rather shows the development of a methodology of empirical music analysis in the form of a cognitive attentive listening model that can be later applied to a corpus of works of a composer or a style in order to test its wider functionality.

The momentary versions of both methods and tools are applied for the first time in combination in publication IV. In earlier studies (publications I, Ia) and publications referring to earlier data (publications Ia and II, Lock 2011a) I have used as slider-controller hardware a Behringer controller but basically the same Max/MSP software (written by Hans-Gunter Lock) that was 2011 further developed for TEDEA. The COSM development started in 2015 (first presented in Lock & Kotta 2015), the Max/MSP software was written by Hans-Gunter Lock to turn COSM into an digital empirical tool. Its principles I have introduced[95] in three paper presentations (Lock 2017b, 2019, 2023a) including first overall results, in detail results are issued in publication IV. TEDEA and COSM were applied in the main empirical experiment-session for this dissertation between September and December 2017. The main object of analysis is Tüür's 4th Symphony/Percussion Concerto *Magma* (2004) (30'58"). The sample for **Experiment I TEDEA** was N=26 (experts n=4, music students n=16, art students n=6) with average age 24,3), the sample for **Experiment II COSM** was N=14 (same experts n=4, same music students n=10) with average age 23,3) (see more details in publication IV). **Experiment I** took place in small groups (up to 10 participants), **Experiment II** took place individually, each participant was placed in front of the MacBook Pro and guided to perform the task. For music listening, comfortable fully-enclosed dynamic Pioneer Powerful Bass DUCT Stereo Headphones SE-M290 were used.[96]

The **Experiment II COSM part 1** (Impulse detection) experiment took place with the DAW Reaper, the **Experiment II COSM part 2** ('content' of Impulse musical parameters' salience detection with the octagonal slice shape) took place with Max/MSP software, see Figure 4. Impulse 'content' refers to the eight musical parameters (and contrast as separate "parameter") iconized as colored dots on the left side of the interface that are required to be placed with mouse clicks at the corners of the slice in accordance to the approximated grade of salience either on the left side (more close and more clearly perceivable as salient to the listener, in visual analogy nearer to the viewer) or at less visible corners on the right side (less salient to the listener, in visual analogy located more in a background position for the viewer). Details concerning the technical steps to conduct TEDEA and COSM see Appendix B.

---

[95] Based on modeling principles partly issued in publication III and enlarged with Conceptual Blending Theory (CBT, Antović 2018, 2022) (see Figure 3 in 4.2).

[96] Ported bass duct engineered for deep dynamic performance. impedance 32 Ω, sensitivity 102 dB, frequency response 5 Hz to 25,000 Hz, maximum input power 1,200 mW. Weight 225g without cord, comfortable softly cushioning around the ears (SE-M290 [2025]).

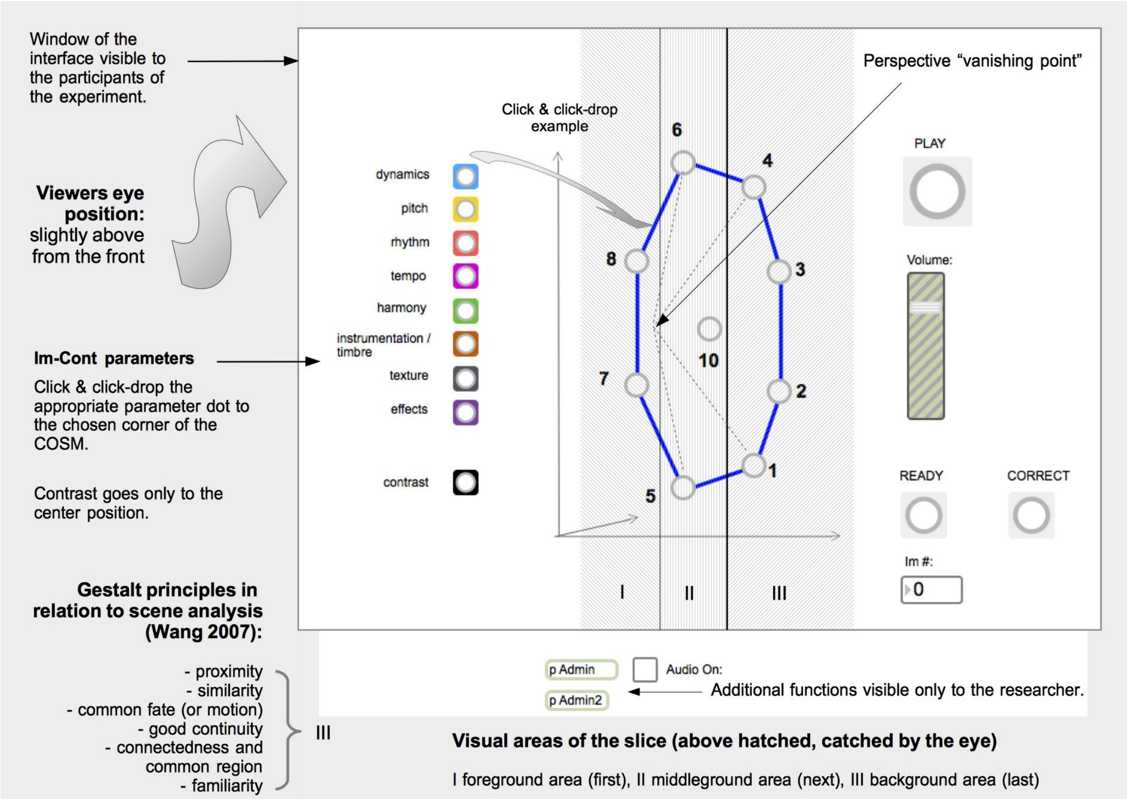


Figure 4. Principles of the visual interface of the COSM measurement instrument based on *Gestalt* principles. This Figure combines the earlier issued separate Figures: principle of the octagon with gestalt principles (Figure 2 in publication IV, Figure 6 in publication V), COSM slice Max/MSP interface used by the experiment participants (Figure 3 in publication IV). Im = Impulse, Cont = ‘content’ = musical parameters.

As follows I discuss laboratory *vs* ecological research settings. Because laboratory controlled stimuli lack the context of “ecological” music listening (Eitan & Granot 2009: 170–171), researchers suggest using rather complete musical pieces and applying “attentive listening” (Deliège & Melen 1997: 388). The “ecological” musical listening setting is related to the ecological nature of creativity that is part of the five principles of cultural psychology of creativity (Glăveanu 2010). I have hierarchically visualized the creativity researcher Vlad-Petre Glăveanu’s principles treating the ecological nature of creativity as basically embracing the other four principles: the contextual principle that is in mutual relationship with the generative, the meaning-oriented and the developmental principles (Lock 2011a: 127). In this creativity-principles commentary (Lock 2011a: 130–133) I have demonstrated my research design understanding based on examples from my earlier slider-controller studies on Erkki-Sven Tüür’s *Oxymoron* (2003) (Lock & Valk-Falk 2008, publications I, Ia) and the Estonian contemporary composer Helena Tulve's orchestra piece Sula (Melting, 1999) (Lock & Valk-Falk 2009).

For this dissertation I have conducted the experiments (2008, 2010, 2017) mainly in small university lecture rooms, partly in highschool and partly in non-education-related rooms. The education-institution rooms are functional but they are still non-clinical (not like experimental labs). In the context of education they are a quasi-natural environment where people meet to learn, in this respect they could be somehow called socio-cultural 'real-life' (Lock 2011a: 130–133) or real world settings (Coath *et al* 2013: 9). In general, I have provided my participants quiet rooms and a permissive atmosphere to concentrate. This, of course, differs from so-to-speak real socio-cultural 'real-life' situations at a party or a concert, or real-world-settings in noisy environments (see related terms/concepts for the research of complex sounds and salience in 3.3). Concerning 'real-life' settings I rather refer to a situation where the participants are put in a convenient space (for TEDEA in small groups, for COSM separately together with me) to enjoy music using an "attentive listening" strategy – for the student participants the rooms were familiar to them. I claim that my settings have enabled them to concentrate on the music to such an extent they may have not done before, especially while involving tools that enabled them to express their immediate listening experience: using slider-controllers (TEDEA) or focus on details (musical parameters) of the music (COSM). It partly enabled them to become familiar with so far unknown music in a permissive situation. The criticism towards artificiality of traditional music psychological experiments concerns rather the object of research and the need to observe complete musical pieces instead of singular stimuli. The environmental, 'real-life' and real-world aspect also concerns rather the object of research calling for the need to analyze the interplay of elements (and participants' responses) in music as environment (1.2, 1.3, 2.5, 3.2, publications III, IV) that is provided only by whole works (examples on salience research see 3.3, examples on musical "macroform" and tension/relaxation research see 4.1).

In this paragraph I discuss the principles of "attentive listening" and chunking**.** To analyze complete musical works based on "attentive listening" a two-stage design of stimulus comparison (Eitan & Granot 2009: 165) is suggested: a fast, pre-attentive "comparator" which operates on wholes, and a second, later comparator, which checks feature by feature" (after Krueger 1973), used also by Deliège, Melen, Stammers and Cross (1996: 127). Farbood (2012: 389) explains two different data retrieval approaches to explore musical tension in relation musical parameters: 1) asking listeners to make continuous, online or real-time data-capture-based, or 2) obtain discrete, retrospective judgments. The attentive listening (attention and salience, see 3.2) stage can also be called an "intuitive" consciousness approach (Clifton 1975) (see 2.1, 2.2). Compare Lester Krueger's stages with the visual-domain cognitive microgenesis approach (Rosenthal 2004: 224) that I have applied for COSM: "Primary categorization" (salience detection of Impulses) and "Local discrimination" (musical parameters' salience detection per Impulse). The latter is a methodological compromise between listening to a complete work and excerpts that has been applied by Albrecht (2012: 17–20) in his Progressive Exposure Method (PEM). Albrecht uses rather arbitrary same-length fragments/chunks (segmentation) of the work instead of music-theory-informed chunks as well as a rather random or mosaic-ordering of the chunks to avoid biasing effects.

However, in **Experiment II COSM part 2** I still use the chronological order of the chunks, because I am rather interested in the development and listening strategy the listener exhibits (via learning) in order to get to a narrative analysis. Of course, a comparison experiment could be conducted later to test if the order of the chunks affected the results. Observing repetition and priming aspects (Russell 2019; Bigand *et al* 2005) in this context is a future task. In the research by Addessi and Caterina (2000, 2005) and Addessi (2010) the expert listeners could listen to the music as often they wished, the listening test participants, however, could use only a limited amount of re-listenings.

As follows I explain the *psychophysical measurement*[97] principles relevant in my data-retrieval methods. The first is the "just noticeable differences" (JND) principle between minimal-different magnitudes on a "yes-no" level; this applies intuitively during detecting "moments of change" – the participants (listeners) should become aware of a difference that is significant for them personally ("yes" or "no"). The second is the assignment principle of the "not directly observable sensation" to another medium (e.g. numbers); this applies implicitly – the participants must not be aware of the numbers (it becomes explicit only for the researcher). The third is the representation principle, this applies visually and bodily – again, the participants must not be aware of its mechanisms (it will be made explicit by the researcher via data-analysis and visualization).

In **Experiment I TEDEA** the detection of musical tension with slider-controllers (increasing, decreasing, remaining) makes use of the first principle, "just noticeable differences" (JND),

---

[97] How to measure the not directly observable sensation that intervenes between physical stimuli and behavioral responses? (Steingrimsson 2016: 205) (see in 2.4).

during detecting “moments of change.” The second principle, assignment to another medium, remains implicit and unaware for the participant. The third principle, ways of representation, is at work while the participant moves the slider-controller visually and bodily up and down, but the large time-scale of the 30 min music normally will the participant not enable to memorize and rationalize this principle (dancers, indeed, can be more aware of the movement-aspect).

In **Experiment II COSM part 1** the Impulse detection of musical events makes use of the first principle during detecting “just noticeable differences” (JND) as “moments of change.” During this process the participant may develop thoughts concerning the reasons (parameters) that may cause this Impulse, but the methodology does not enable or require awareness/reporting these thoughts to the researcher. The second principle is irrelevant for the participant because the experimental setting does not allow assignment to a reference-medium: the detection process takes place in real-time in a DAW, but the visual wave-form is disabled to not influence the response. The third principle is also irrelevant, because visual and bodily references are disabled or minimized by the experiment settings.

In **Experiment II COSM part 2** the ‘content’ of Impulses, musical parameters’ salience detection with the octagonal slice shape makes use of the first principle via the visual salience requiring the listener to decide their salience in analogy to the auditive salience of the musical parameters. It remains, however, intuitive, how clear and what “size” the “just noticeable differences” (JND) (magnitudes) are for the participants personally. The second principle applies with the full awareness of the participant in using the octagonal shape of COSM as the other (visual) medium. The difference and relative “scale” of the magnitudes are given with the octagonal shape and the participant is asked to imagine an analog in difference in the weight of the musical parameters in respect to their auditive salience. The third principle applies visually in the digital (computer-screen-based) 2D interface as an imaginary 3D octagonal shape, the bodily aspect is minimized.[98]

Here I will briefly mention some general data analysis issues. Although the TEDEA slider-controller apparatus was especially developed and built (by Hans-Gunter Lock) for my dissertation one can say that slider-controllers have already been used widely to analyze musical

---

[98] In a so far not yet realized physical COSM model the bodily movement of pressing buttons at different corners of the octagonal shape would become more haptic and sensible. It would also enable to minimize the visual aspect of looking at the screen in order to find and assign the right corner. A physical shape prospectively would enable to use the model also for (audio)visual stimuli, not only for auditory stimuli. On the other side, also eye tracking technology could be applied to observe at what moment the listener looks at what corner of COSM. This would take the body movement out of the process and ideally could provide more immediate/faster responses.

tension (see Farbood 2012) and in emotion research (see in Schubert 2010). This is the reason I will not provide more details on this data retrieval technology. In order to arrive at something original that aims at developing the fields of music psychology and music theory further I have put my focus on developing my own concept and methodology especially focusing on developing COSM. This requires also to deal with data analysis issues. Data analysis matters are partly presented in publications II, IV and more actually in 4.5. An overview of the data-analysis principles and the clusters' statistics derived from the DBSCAN algorithm (Ester *et al* 1996) is given in 4.6. The following subsection explains in detail modeling steps, reasons and theoretical background for developing COSM based on eight stages and 12 questions.

4.4 Modeling steps for developing COSM

In this subsection I explain the steps, reasons and theoretical background for developing **COSM: Cognitive Octagonal Slice Model** based on my **Twelve Strategic Steps for modeling/analysis of scientific models (TSSM)** (see publication III). The model has been developed in parallel with the theoretical base, this means that the answers to the following questions have been partly written retrospectively, however they reflect my thoughts throughout the whole process.

The TSSM depart from the following four requirements:

a) A scientific model needs to have one or more accepted/proven theories as backbone(s);

b) The phenomenon observed/modeled must be able to be interpreted, translated and expressed into explicit or implicit 'input' data, the latter should be able to be made explicit, turned into numbers; the explicit 'output' data must be analyzed and interpreted reasonably and simply without losing touch with the phenomenon to be modeled;

c) The model must be consistent, a whole in itself. Gaps or inconsistencies must be minimized or turned off (eliminated);

d) As an exploratory or new model, it should explain something generally known, but not yet explained in a similar systematic way from the same angle, from a new viewpoint based on accepted/proven theories, but offering also new or even unexpected aspects using e.g. cross-domain mapping approaches between different fields.

The eight steps (ordered with the letters A→H) appear as one or more questions and sub-questions in eight developing stages/phases (publication III):

A) Initial situation (problem) stage,

B) Prospective outcome (hypothesis) stage,

C) Data developing stage 1 – "input" data for the model,

D) Object of the model/research in relation to the data,

E) Model(ing) application phase,

F) Data developing stage 2 – "output" data of the model,

G) Data analysis phase,

H) Data comparison phase (cross-domain mappings).

Twelve Strategic Steps for modeling/analyzing of scientific models (TSSM) in eight stages/phases

A) Initial situation (problem) stage

Q 1 (A). What kind of phenomenon one observes/investigates?

The phenomenon taken as source (Brodie *et al* 1994) or as target (R_2) (Duit 1991) for modeling is "something known to us from our everyday life or prior experiences" (Brodie *et al* 1994) (see Figure 3 above): the intuitively perceived/understood wave-like character of contemporary post-tonal orchestral music by Erkki-Sven Tüür. The phenomenon I investigate is "something we want to understand" (Target, see Figure 3): musical tension. Musical tension can be defined in various ways. In publication Ia I have defined musical tension in relation to form as a hierarchical phenomenon: "perceived musical tension can be seen as being dependent on the hierarchical aspects of form. We hypothesize that the intensity of the perceived musical tension is proportional to the structural (or hierarchical) significance of the corresponding musical event" (publication II). In this Introductory chapter of this dissertation I have arrived at the definition of ‘Perceived structural musical tension’ (PSMT) based on cognition and narrativity principles (see 3.5).

Q 2 (A). What is the problem/aim to be researched/explored?

The problem/aim to be researched is how to detect, measure and conceptualize ‘Perceived structural musical tension’ (PSMT) of the wave-like contemporary post-tonal orchestral music of Erkki-Sven Tüür combining music theory and music psychology approaches together while arriving at a cross-domain/interdisciplinary interpretation of the outcome of the research – applying cognitive science methods. To specify the definition of musical tension given in Q 1 above I assume that the ‘corresponding musical event’ should be understood as ‘salient event’ (see in 3.2). As the most sense-making salient feature and musical event the Impulse (Im) has been chosen (first proposed in publication I and in Lock 2010b), because it is a perceptual phenomenon representing a “moment of change” which can be detected relatively clearly in empirical perception tests. This “moment of change” can be caused by a variety of aspects in the music, among them physiological (air pressure) and psychological (“felt tension”) aspects as well as intramusical structural parameters (PP: pitch, rhythm, harmony; SP: dynamics, tempo, instrumentation/timbre, texture, effects) allowing to approach musical tension as a structural phenomenon. An important concession hereby is the relativity of these “moments of change” which cannot be measured neither empirically nor in the score purely visually/theoretically as absolute values. An Impulse just appears to be perceptually detectable as a “moment of change,” due to the illusion of “directness” in perception (no one-to-one relationship between the performed and the perceived, Seashore 1938: 382, see 2.3), the moment cannot be exactly detected (see publication V), and one cannot measure at the same time the concrete aspects causing it, because we experience musical events as holistic (see publication IV).[99] With the help of slider controller experiments (Experiment I – TEDEA) the listener creates data on the relative intensity of the perceived musical tension that can be visualized in a curve – these curves, of course, clearly represent visually the wave-like structure of musical tension. Analytical examples of Tüür’s music see in publications I and Ia, II and IV.

B) Prospective outcome (hypothesis) stage

Q 3 (B). What may be the postulated outcome to be reached?

---

[99] Concerning visual cognition in Carsetti 2004b, dynamic vitality forms in Stern 2010, temporal dynamic forms in music in Mailman 2010); visual cognition in Rosenthal (2004: 224): primary categorization needs completion by local discrimination (see in 2.3).

3a. What kind of theory/theories relate/s to/explain/s the phenomenon/problem/aim?

3b. Is a hypothesis necessary?

The prospective outcome to be reached in this dissertation is a conceptualization of how to perceive, cognize, and make available in data form salient features as Impulses/"moments of change" in contemporary post-tonal orchestral music (CPTOM) using COSM. COSM models the perception and cognition process in connection to the musical structure – musical parameters as 'content' of Impulses.

3a) A theoretical base is the analogy of strategies in auditory and visual perception/cognition processes (Vicario 2003, Athanasopoulos & Antović 2018). Examples of relevant theoretical background are (1) music listening as ongoing (real-time) process of sense-making via focal attention (Reybrouck 2010: 191–192), cognitive analysis as elaboration ("structuring process") of "relevant categories" (Chouvel 2014); (2) seeing, thinking, and knowing as meaning and self-organization in visual cognition and thought (Carsetti 2004a), *microgenesis*, immediate experience and visual processes (Rosenthal 2004, Bachmann 2000) (see 2.3) – here transferred as analogy to listening; (3) Marr's (1982) computational approach to understanding vision, Visual Scene Analysis and Gestalt Analysis principles (Wang 2007) as well as Auditory Scene Analysis (Bregman 1990).

The definition of musical tension in Q 1 includes the hypothesis that "the intensity of the perceived musical tension is proportional to the structural (or hierarchical) significance of the corresponding musical event" (see publication Ia). The significance of the corresponding musical event (Impulse) is expressed in its salience during the perception process. The "proportionality" of the structural aspects of the musical event is expressed in the relative hierarchical salience of the musical parameters which's correlations and other mathematical relations can be detected in the data provided by the COSM empirical experiments using cluster finding (with a data mining algorithm). An Impulse is significant if a sufficient (significant) number of participants detects it. The musical parameters detected to be hierarchically salient in each Impulse, express to which extent what parameter was detectable and on what hierarchical level (proportionally) in accordance with psychophysical measurement principles (see in 2.4). Furthermore the 'Perceived structural musical tension' (PSMT) definition (see in 3.5) assumes "that the unique surface features (cues, salient features) and effects of these patterns (salience principle) should be psychologically meaningful," therefore triggering attention and prioritization based on attraction points (see in 3.1, 3.2, 3.3).

3b) Several hypotheses can be formulated depending on the object of research. In publication IV I have posed for this dissertation the following questions and hypotheses (QH's): (QH1) Whether secondary musical parameters (SP) are more salient than primary musical parameters (PP)? Hypothesis 1: The change of musical tension shows a parallel change (in start→end values) in SP and vice versa. The answer is that SPs are not generally more salient than PPs in Tüür's analyzed pieces and there is no clear correlation with the slider controller data (by experts [n=4], music [n=16] and art [n=6] students), several clusters, cluster pairs or sections show partly correlating patterns (see publication IV). (QH2) How is it possible to interpret the role of musical parameters for the development of structural musical tension narratively? Hypothesis 2: Salient changes in the flux of the music (influencing musical tension) detectable with slider controllers (global and local) and salient musical parameters with COSM (local) show clear regularities in their start→end values in parallel and cross-trends. The answer is that the slider controller average data (musical tension) are partly narratively interpretable via in-/decreasing trends: in cluster 21 the experts' sample average shows a maximum intensity level that indicates the maximum tension at the end of the first half of the piece (in part B between Cluster Areas Correlated (CAC) e and f) that head towards the percussion soloist cadence as a kind of caesura in the otherwise attacca piece. The other samples' average results (music and art students) show a lower maximum in cluster 21, their local maximum in the first half of the piece is located at the end of part A between CAC d and 3 in the cluster pair 17–18 (art students show a similar maximum in the middle of part A in cluster pair 9–10). See detailed results in Figures 10c/d in publication IV.

C) Data developing stage 1 – "input" data for the model

Q 4 (C). What kind of data does the phenomenon provide?

The data (arbitrarily chosen natural numbers ranked as decreasing from (10) 8 to 1) the Impulse (Im) phenomenon provides via COSM is based on both auditory and visual perception and cognition processes and is expressed accordingly to the salience of the musical event (Impulse/"moment of change") and its 'content' (salient aspects/parameters) both visually and in relative ranking numbers. The central graphical element of COSM, the virtual 3D-slice octagonal shape corners' relational weight is graphically based on gestalt analysis principles (see more in Figure 4 above and in publication IV) to express visually/in relative space/distances their position, this means their salience in relation to the auditorily/visually perceived salience. The corners in the model are equipped with natural numbers counting backwards from 8–1 (number 10 in the middle of the slice expresses the compound salient feature called ‘contrast’) to express the hypothetical proportional relation of the salience of these parameters later also separately and in comparison in 2D diagram form. This data can be analyzed using cluster finding (data mining) principles, see in publication IV.

Q 5 (C). How to retrieve/record appropriate data, portions of data (e.g. feature analysis)?

5a. Is computation of data retrieval possible?

The first portion of COSM data is retrieved in the first part “one”-dimensionally (binarily, digital standard on/off): Impulse/“moment of change”) salience – yes (1) or no (0). This data is recorded with key M by the participants “visually” in the label track of the sound file (its waveform is turned invisible) in the DAW Reaper (placed at the corresponding time-point position). These time-points are exported from Reaper to be available in data analysis programs as well as in the second part of COSM, where the second portion of the data is retrieved “multi”-dimensionally: auditory detection of the salience of the musical parameters around each Impulse, putting them according to their visual analogical salience to the appropriate corner of the 3D octagonal slice, where they automatically get attached to a natural number from 8–1 which allows to express the salience of every parameter (for each participant individually) both in the octagonal slices as well as in 2D diagrams (especially if analyzing results of participants together).

5a. Computation of data retrieval is possible through the application of Max/MSP software which allows the participant to choose each Impulse sound chunk separately, re-listen to it as much as needed and put via mouse click the appropriate musical parameters to the appropriate salience position in the virtual 3D octagonal slice. The data is then extracted automatically into a text-file containing each Impulse' timestamp together with the corresponding salience number for each parameter, which can be imported into data analysis programs.

Q 6 (C). To what requirements need the data to be restricted, portioned, segmented?

The data expresses proportional (relative) weight of salience (in COSM part II), it can not express absolute values. Therefore the data is sufficient in natural numbers. But, these enable one to focus on simple proportions of parameters of the music while ‘stripping away’ not necessary aspects – see ‘Aristotelian idealization’ of models (Frigg & Hartmann 2024, publication III). Hence, the musical parameters are simplified also according to their “meaning” level: each parameter is treated as a "simple" phenomenon in the sense of not taking into account its real-world complex structure. They are reduced to the common general concept musicians have when speaking of pitch, rhythm, dynamics etc, hence primary and secondary parameters are treated on the same level of simplification. This, of course is a kind of 'Galilean idealization' (Frigg & Hartmann 2024, publication III), but in modeling, and the case of COSM it is a 'caricature model' (including both 'Aristotelian' and 'Galilean idealization'), constraints and idealizations are inevitable as well as needed/useful from the point of data retrieval and reduction.

D) Object of the model/research in relation to the data

Q 7 (D). Does the data enable to research the phenomenon or its observer, or both?

The data enable to research the phenomena or general and special object of this dissertation (music, musical tension) and its observers/listeners (empirically quantitatively or for individual participants qualitatively), but in this dissertation the focus is put on the development of the model and on the phenomenon (Impulse, Impulse content). Outcomes/results of the COSM part II experiment results (N=14, experts [n=4] and music students [n=10]) see publication IV. They enable to compare the participants' cognitive listening strategies analyzed with statistical methods: averages, means, deviations, correlations, DBSCAN algorithm (Ester *et al* 1996) for cluster finding.

E) Model(ing) application phase

Q 8 (E). In which stage of the research process (and why) it is possible/necessary to apply modeling?

In the **first part** of COSM the phenomena under observation take effect on different levels: Impulse/"moment of change" can be detected applying the concept of salience. This is achieved through reduction (simplification) of the information the music as an auditory stream, an environment or dynamical system provides. The Impulse is also part of model-like thinking including a discretization process: Impulse metaphorically seen as "geological drilling core" in "music as a landscape" based on the broader analogy of "music as environment" (publications III, IV) (which is not a model, but suggests the infinite/non-discret/constant richness of auditory experience), Impulse as a "still-frame" of a film sequence (which is a kind of "conceptual analogy model") (modeling principles see publication III, metaphors presented in Lock 2017b).

The further deeper concern is, that the Impulse (Im) is still a compound and (on a more detailed level) still not a just simple/“one”-dimensional/pointillistic/clearly bordered phenomenon: especially if one takes the time dimension into account, which turns every sounding phenomenon into a temporal one, which can be perceived and conceptualized only through its duration, if experienced in a minimum amount of time (short now-moment, “a brief present-time scale” in microgenesis see in 2.3; “now moment,” “narrative nowness,” (Kauttonen, Kaipainen & Tikka 2014) Continually Updated Continuum of Nowness (CUCoN) (Mailman 2010: 235). In this context the decision to take the Impulse/“moment of change” as the central salient event as a basic aspect is an act of modeling and it was the first crucial step towards further analysis of musical tension.

In the **second part** the next step means gaining insight into the ‘content’ of each Impulse, to understand more in detail why the participant has chosen this Impulse (musical event, “moment of change”) to be salient. This needs higher level conceptualizing: the “moment of change” requires one to become aware, what are the components which cause this “moment of change” to happen/appear. Compare this with principles in visual perception underlining that primary categorization needs completion by a process of local discrimination (Rosenthal 2004: 224, see 2.3). These components in music can be detected (based on sound recording or simplified MIDI data) either through a complex software (but still under constraints depending on the program’s function and aim)[100] or cognitively, admitting the high capability of the human ear and conceptualizing power (Wang 2007: 164). However, the human listening, perceiving and cognizing process (its comprehension) is explainable only through models, including reductions, filtering, percept/concept contrasts, expectations etc,[101] which led me directly to (implicit and explicit) modeling. My decision for this dissertation was to trust the participants’ ears and perceiving/cognising capability and observe these based on already existing (and practically applied) music theoretical concepts (musical parameters) (see also Addessi & Caterina 2005, Addessi 2010) as well as using analogies of auditory and visual perception principles (Impulse, salience).

Again salience is here the key concept which allows to make a distinction – what to choose in the analytical process, to differentiate between the more extruding or protruding and the less

---

[100] Which, in turn, works only according to predefined algorithms, also based on complicated models like Markov-chains etc, that have not yet reached the capability of human perception, cognition and understanding, see Wang (2007: 164).

[101] Physiological – ear, brain (e.g. Bregman 1990, 1993; Rocchesso & Fontana 2003; Wang & Brown 2006; Wang 2007); biological-genetically – evolution, living/surviving conditions, soundscape, music as Humboldt system, music and sound/the sonic as environment or dynamical system (e.g. Schafer 1994/1977, Huron 2006, Merker 2002, Reybrouck 2015, Burrows 1997, Kaipainen 1994).

extruding or protruding parameters to focus on. Even more important here is to admit the idea (and assumption) that in every moment alla of the musical parameters (we are commonly used to apply and analyze in music) are at work, exist, are part of the whole musical event, no matter if we may not be able to perceive/(re)cognize all of them with the same intensity or at all. Reasons may be masking effects in the music, technical aspects caused in the recording technique, or our own rather unaware physiological constraints (hearing mechanisms, ASA: Auditory Stream Analysis, Bregman 1990; Wang 2007). Furthermore the more or less aware 'act of focal attention' or 'act of mental pointing' plays a major role: the listener's own "filter," that depends on experiences, training, learning, and both the ability and the will to concentrate on chosen parameters.

In general, the idea that all (at least the eight most important) musical parameters are virtually existing and working independently of being perceived or not, was the decisive aspect leading to the development of COSM, including the central octagonal shape based on analogies with visual perception principles. This is connected to the way of representation of the data (see next paragraph).

In the **third part**, the data visualization and analysis, modeling takes place while applying cluster finding (using the data mining algorithm DBSCAN) on the COSM experiment participants' data (N=14) (publication IV). This includes detecting patterns (un/regularities), analyzing averages, correlations, trends and global principles in the data. Such patterns emerge and can be analyzed first as parameters' development trends separately, secondly parameters can be also analyzed together. Looking at those patterns in each particular situation of the music under observation (in comparison with the form analysis of Tüür's orchestral music, Kotta 2011, see publication IV) one can make generalizations towards strategies (a) the listeners (experiment participants) have applied to conceptualize this particular music, and, with constraints, (b) the composer Erkki-Sven Tüür has used in a particular piece working with special salience patterns of musical parameters. Tools to analyze those patterns are applicable based on cluster finding (the data mining algorithm DBSCAN): see visualization and analyzing empirical data of slider-controllers as well as COSM experiments together in publication IV.

Q 9 (E). What type (function/method) of model makes sense to apply?

The model type appropriate to apply depends on the aim and the nature of the phenomenon. In general COSM is a *cognitive model*, because it deals with perceptual and cognitive phenomena (sound, music), more concretely with a phenomenon – Impulse (Im) – which we try to understand using *salience* as well as searching for its 'content' (musical parameters). In the context of the modeling and sound researcher Hendrik Purwins' *et al* (2008a: 154) four aspects of a cognitive model, we can say that: (1) Its predictive power applies partly (especially while analyzing perceived musical tension in combination with slider controller TEDEA experiment data): one could predict that mostly dynamics (loudness) would be the most salient parameter and therefore overrule other parameters, but this is too simple. The model does not predict something like the harmonic tension prediction model by Lerdahl (2001) or the musical tension prediction model of Farbood (2006, 2012). (2) Its generalizing power applies well to COSM in several stages: Impulse (Im) as complex, but general phenomenon in connection to microgenesis, cue abstraction, percept/concept principles; salience as audiovisual perception/cognition analogy. (3) Its simplicity applies in for detecting the holistic phenomenon Impulse (**part 1**) and eight known musical parameters (**part 2**), the ranking numbers 8–1 for salience of the musical parameters via audiovisual perception analogy is also an arbitrary but simple solution. (4) Its relation to existing theories applies well while taking into account theories listed under aspect 2.

The outcome of COSM is based on Frigg's and Hartmann's (2024) categorisation on the semantic level rather a representational model (representational model of phenomena), because the object under observation is a phenomenon and its observation provides data that is later visualized and can be compared with the recorded music. On the ontology level it is more complicated to categorize – at a first glance the gerrymandered ontologies category seems to apply: a mixture of elements belonging to different ontological categories; models involve structural as well as narrative elements. During **part 1** ("primary categorization," term inspired by Rosenthal 2004) the Impulse should usually be perceived as a compound and holistic phenomenon but its structure ('content') is complex. Maybe an expert listener already (re)cognizes those structural elements. During **part 2** ("local discrimination", term inspired by Rosenthal 2004) the Impulse 'content' is required to be "broken up" into structural elements, the eight given musical parameters via salience ranking. This ranking is conceptualized not just as linear hierarchy, but, using audiovisual perception and cognition analogies as well as gestalt principles, into the eight corner (octagonal slice) shape (as included into the abbreviation COSM), which allows all parameters to be existent in the same instant. No parameter is excluded *per se* or claimed to be not existent, but it depends on the individual listener, which parameter is more salient.[102] Hereby the octagonal slice of the model is conceptualized as a virtual 3D structure projected as a central perspective shape on a 2D computer screen.[103] In part 3 (data visualization and analysis) the data is visualized in 2D diagrams. While looking at the Impulse 'content' data of the participants for each parameter separately a narrative aspect appears: the data reveals when which parameters was absent or detected as salient and to what amount.

As a specification of the *representational models* of phenomena COSM can be categorized as an *idealized model* (a deliberate simplification of something complicated with the objective of making it more tractable; including both, 'Aristotelian' and 'Galilean idealization': caricature models, Frigg & Hartmann 2024): the complex phenomenon of the Impulse (Im) and its 'content' is simplified to make it more tractable. On the other hand it can be seen as a *phenomenological model* in both versions: (a) traditionally: representing only observable properties of their targets and refrain from postulating hidden mechanisms etc – in COSM **part 2** I postulate that only these eight parameters (plus contrast as compound/complex phenomenon) play a role without specifically addressing their 'content' or hidden mechanisms, also the perception of gestures and dynamic forms of vitality (Stern 2010) work like that; (b) models that are independent of theories, but "many phenomenological models, while failing to be derivable from a theory, incorporate principles and laws associated with theories" – this applies exactly in the sense the quotation underlines, taking into account a number of theoretical principles (perception, cognition, gestalt laws, audiovisual analogies, microgenesis (Rosenthal 2004) etc). COSM also includes features of an analogical model (Frigg & Hartmann 2024) using the analogy of mechanisms of auditory and visual perception and cognition (event, impulse, salience).

Based on Lawrence Zbikowski (1997, 2002) COSM can be seen also as a *conceptual model* based on image schemata that are not necessarily visual but mental images, in a specific domain

---

[102] This is intuitive and subjective, but seems to be the only possibility to access these phenomena via perception and cognition based on empirical experiment subjects, the same is pointed out by Vicario (2003: 18) referring to Benjamin Libet (1990): "in order to describe perceptual facts at the perceptual level, there is only a way: to ask experimental subjects for the description". Actually self-learning AI programs are quite far developed, see analyzing audio-visual associations through self-supervised training (Owens & Efros 2018). However, a computer program, also AI, is still limited to the conditions it has been written for. For instance Andrew Owens and Alexei Efros' (2018: 14) method detects temporal misalignment "but one could also incorporate other learning signals, such as the information provided by ambient sound".

[103] Later it could be realized even as real 3D shape to be touchable by the participant to place the saliency-of-aspects-decision not only virtually via mouse-click onto a screen, but as tactile button or sensitive push area on a physical model using (see aspects of "bodiliness" O'Regan, Myin & Noë (2004: 106–107) and embodiment).

(here music), but not in the cognitive processes that it involves. In COSM the analogical visual component is included, too, to support the auditory cognitive process of detecting salience. This applies more in general to observable phenomena of different domains – it works with cross-domain mappings. The octagonal slice (as a visual image schema) functions as a cognitive summary of repeated patterns.

Based on my two-dimensional classification of modeling and models (publication III) the outcome of COSM can be understood in the method category as b) representational model (including cross-domain modeling graphically expressed) and the working principle of COSM as c) cognitive model (including analogy, cross-domain, and using graphical tools). In the function category the outcome of COSM can be categorized as exploratory (III) and explanation (II) model. It is exploratory, because it tries to understand not yet formalized phenomena (Impulse and the 'content' of Impulses: eight musical parameters) and, together with slider controller (TEDEA) data, the phenomenon of perceived musical tension from a yet unexplored angle – using cross-domain mappings (audiovisual analogies: event impulse, salience), the understanding of music as environment, the metaphors of a geological drilling core and "still-frame" of a film sequence (the event/Impuls and its 'content'), principles from microgenesis. On the other hand the outcome of COSM is that of an explanation model, because it aims to explain the Impulses with a defined 'content' (musical parameters) as well as perceived musical tension in a further step. In this sense it aims enhancing the understanding of (i) musical tension as a well known (but complex) phenomenon, (i) the phenomenon of salience in music cognition focusing on the general but complex aspect of Impulse using cognitive (empirical), mathematical and computational (e.g. data mining) tools. Its goal is to enable more than one analyst to express immediately the salience of musical parameters in each self-detected Impulse. A "physicalized," tactile version of the octagonal slice would improve the audiovisual analogy in real touch-space, and would include the aspect of embodiment ("bodiliness," O'Regan, Myin & Noë 2004: 106–107) into the audiovisual perception process of the analysis. However, COSM can even be considered as a (IV) model in composition and analysis of music, because its principles could be incorporated into music composing systems. One realized example is the Reinforcement Learning (LR) framework using musical tension as reward function to modulate in real time three musical parameters (tempo, articulation, dynamics) (SMuSe, Situated Music Server), Le Groux 2011). Another not yet realized concept that could use especially COSM input is a symbiotic composing idea applying man-machine symbiosis for learning as a compositional model using biofeedback and machine learning (Kosunen 2018, Kosunen &

Väljamäe 2020). Although several previous researchers (Lerdahl 2001, Farbood 2012, Herremans & Chuan 2017) have analyzed and conceptualized musical tension based on musical parameters and created prediction models to compare with empirical data, COSM is neither a (I) prediction nor expectation model. Of course one may expect or predict several musical parameters to be more or less influential for creating an Impulse or even musical tension, but expectation and prediction are not a part of the model's functionality.

F) Data developing stage 2 – "output" data of the model

Q 10 (F). What kind of data are the output of this model?

10a. Does it represent the phenomenon well?

10b. How generalizable are the results?

COSM provides in **part 1** (Impulse salience detection) binary data (1=yes and 0=no). In **part 2** the salience ranking of the eight chosen parameters (contrast) is achieved through the octagonal slice numerical data for each corner of the shape from 10 (contrast) and 8 (most salient) to 1 (least salient). The decision whether a parameter is less or more salient in comparison with other musical parameters within the musical context of the Impulse (Im) chunk works similarly to the principle of "just noticeable difference" (JDN) known from psychophysical measurement (see 2.4, explanations of the psychophysical measurement principles relevant in my research design see also 4.3). This means that the listener has to still make binary decisions: either "yes" (this parameter is salient) or "no" (this parameter isn't salient). However, this "yes-no" decision must take place in the real music context where several musical parameters are in a contest with each other – the listener also decides this based on one's expectations, priorities and preferences: the listening strategy or narrative. JDN "yes-no" decisions can create conflicts when linearly compared: the question appears, what parameters' saliencies (JDN decisions) have more/less priority. The visual salience-based salience/gestalt principle per Impulse/slice enables the listener to make a JND decision mediated via the visual salience analogy.

The octagonal-slice-based data can be represented visually either using the raw (non-transformed) data of the octagonal slices, which creates a kind of virtual 3D tube model for each participant, where we see how the different parameters of the Impulses change their salience over time. The further treated data visualized in 2D diagrams show tendencies of the parameters' salience over time, based on cluster finding (with the data mining algorithm

DBSCAN, Ester *et al* 1996) several types of correlations between the parameters and/or participants can be revealed.

The binary data of the Impulse 'content' (musical parameters) is visualizable sufficiently in a 2D graph either as detailed raw data per cluster or more generalized as summed or averaged data. A possible virtual 3D-tube visualization of the raw data represents the salience of the parameters in direct relation to the audiovisual analogy, but it is relatively detailed, multilayered and requires at best a virtual 3D representation (e.g. with Blender software). However, a long object of research (a 30 min symphony) provides a huge amount of data that also needs dynamic visualization, e.g. interactive turning options of the slices and temporal development opportunities.

G) Data analysis phase

Q 11 (G). What is the concrete aim of the "output" data?

11a. How to treat (deduce, reduce, transpose, analyze) this "output" data further?

11b. How to use computational tools for data analysis?

The "output data" can reveal patterns on how the musical parameters of the recording of Erkki-Sven Tüür's 4th Symphony/Percussion Concerto *Magma* (2002) are perceived/cognized and unfold during the listening process – showing the listening strategy of N=14 that can be interpreted as narrative. The main data treatment methods (cluster finding with the data-mining algorithm DBSCAN) are averaging and correlation analysis of the binary data (Impulses) as well as summing up, averaging and correlating the octagonal slice salience results within each cluster. A reduction method I have applied in publication IV is filtering the start→end intensity values of each parameter per cluster in order to compare their increasing/decreasing trends within and among clusters. The same is done with the slider controller data per cluster (with Max/MSP software). The musical parameters data can be analyzed also in accordance to their type (e.g. primary, secondary) and I have applied a number of further reductions to detect tendencies in the data (see publication IV). Toby Gifford programmed me (since 2012) a Java script with Max/MSP interface to realize the maximum/minimum (Highpoint, Lowpoint, High Reverse Point, and Low Reverse Point) reduction and averaging of tension curves algorithm introduced as manual version in publication Ia.[104] However, there remained bugs that caused

---

[104] Later Joshua Mailman suggested comparing it to a similar algorithm his supervisor Robert Morris' (1993) developed for melody contour reduction, see more on contour reduction algorithms in Bor (2009).

missing reverse points as well as the technical question appeared, why the reduction algorithm would in one case keep preserving high reverse points while ignoring low reverse points and vice versa. However, the work with this algorithm and the idea of 15 seconds moving averaging windows to achieve a comparison of the different participants' tension curves, already was a kind of clustering principle. To realize this (since 2017, issued in publication IV) I am very thankful to Ilkka Kosunen who suggested (and applied) for the COSM **part 1** Impulse data the **DBSCAN: Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise** (Ester *et al* 1996). I also presented comparisons of so far explored analytical approaches to the slider-controller data (reduction and averaging tension curves algorithm results from publication Ia with wave-forms (Audacity free software) and rotational rhetorical form analysis (Kotta 2011) in a conference in Lüneburg (Lock 2019/2011, paper presentation took place in 2011) and I applied a trend/regression algorithm (with MS Excel software) to the slider-controller data (comparing the results to the other methods) in a conference in Belgrade (Lock 2012) (see 4.3 FN 92, 4.4, 5.2, 5.7, see below Q12a).[105] Thanks to Hans-Gunter Lock I could use a Max/MSP program to normalize TEDEA slider-controller data and filter from the slider-controller data the start→end (and min→max) intensity values per cluster. Thanks to Mustafa Can Özdemir I could use the Tableau visualization software for basic cluster analysis (see 4.6).

H) Data comparison phase (cross-domain mappings)

Q 12 (H). How able are the data to be compared with output from other research methods applied on the same phenomenon?

12a. How to use computational tools for data comparison?

My earlier comparison-type data-analysis try-outs include a 10-sec.-static-window principle with salient features analysis (publication I), a 15-sec.-moving-window principle (publication Ia, 5.2, 5.7) and trend/regression tools (see Lock 2019/2011, Lock 2012, see above Q11b and 4.3 FN 92). Comparisons I have done with annotated (reduced) score examples as well as the waveform visualization (Audacity) of the recording for Tüür's 4th Symphony/Percussion concerto *Magma* Cluster 21 and Cluster pair 22–23 (see Figure 4 in publication IV, see 5.5, see

---

[105] The Lüneburg conference took place in 2011 and the proceedings were published much later in 2019, because the editor-in-chief passed away. The Belgrade publication (2012) I did not finalize because I was unsure about the reasonability of these analyses, in particular because of the problem of how to compare them meaningfully. Also the sample size (N=6, N'=7 – one data set had errors) of the slider-controller experiments conducted 2010 with Tüür's 4th and 6th Symphony did not satisfy me, because in the same year TEDEA was finished, but not yet applied on Tüür's music.

also 4.3). Further comparisons could be done with recording-based automatic spectral (FFT) analysis (on complex sounds and contemporary music, see also 3.3). In future also bio-feedback data could be gathered with HCI (Human Computer Interface) technology.

In the next subsection principles of data analysis and visualization are presented in the context of "multi-dimensional time series" and pattern mining that offer more insight into possibilities to apply the above last two phases: G) Data analysis phase, H) Data comparison phase (cross-domain mappings).

4.5 Data analysis and visualization principles

In this dissertation the results of the analysis of music and in particular the data gained via perception-cognition experiments are visualized as "multi-dimensional time series" (Wiersma 2024). However, even in computer science pattern mining in multi-dimensional time series and event logs has been a challenge (Fu 2011, Feremans *et al* 2019). Time series analysis has been a challenge also in psychological research while more and more recognizing "the importance of integrating temporal dynamics into its theories" (Jebb *et al* 2015: 1). The search for *meaningful patterns* with *semantic relevance* pave the way to including *narrativity* as a means of interpreting music (as structure) and its perception-cognition-based experience, see the definition of 'Perceived structural musical tension' (PSMT) proposed in this dissertation (see 3.5). Applying a model to gain empirical data (publication III, IV) requires making sense of this data. Visualization (to graph, to display visually) of such data is a widespread method (especially for temporal dynamic data, Mailman 2010) that may be even more effective than pure mathematical-statistical means (Levitin 2002: 128–129). Therefore auditory and visual domains are intertwined in several stages of research and modeling in data retrieval (e.g. via the musical score or other mediating tools) as well as in data analysis.

I have categorized music analysis into 'intuitive,' 'systematic,' 'system-based' and 'automatized' approaches (see publication II) that are applicable also for data retrieval as well as for data analysis and visualization in music psychology. In publication II basic principles of the visualization of music (as pre- and descriptive approaches) and a systematization of its automatization (with examples of systematic, system-based and automatized visualization) are described. Automatization in data retrieval and analysis in music has been researched also as follows: the computational musicologist Wei Chai (2005: 13) has applied "automatic music segmentation, summarization and classification [...] combining music cognition, machine learning and signal processing" including music perception and pattern discovery. While looking at the "repetitive properties of music at different levels [...] organized in a hierarchical way" Chai (2005: 13) poses two questions that are relevant also for my dissertation: "Which parts are most "informative" for the listeners to make judgments? What are the most "representative" parts that make the piece unique or memorable?" The most "informative" part of music is related to musical salience detection for classification purpose while listeners directing their attention willingly to information they desire during listening – thereby the distinct weightings for divergent classification tasks and its consistency with human intuition is under observation (Chai 2005: 13–15, 75 etc). The most "representative" relates to music summarization (or music thumbnailing in Music Information Retrieval MIR) "involving human musical memory and the attentive listening process" (Chai 2005: 63 etc). Jamie Forth and Geraint Wiggins (2009: 1) attempt to identify "the repetition of perceptually salient patterns in symbolically represented music" geometrically and with specific algorithms in multidimensional representations (datasets) of polyphonic music. The audio research engineer Oriol Nieto (2015: 1) explores musical structure based on algorithmic and perceptual approaches including repetition and pattern discovery. Nieto's (2015: 2) approach is located "somewhere in between the fields of music information retrieval (MIR) and music perception and cognition (MPC)." Nieto (2015: 4) suggests that "the analysis of the structure of the organized sounds from which music is constructed can help us better understand how humans generate and process musical information." Furthermore, Forth and Wiggins (2009: 2) underline that the usefulness of automated music analysis software has to fit work-specific needs including "varying degrees of ambiguity, and necessarily must scale to real-world musical problems."

The understanding and visualization of music is located between the compositional "forward" (by the composer) and the analytical "in reverse" (by the music theorist) structuring of music, both should be regarded as creative, but not as reversible (see publication II, Lock 2011a: 124–125, Lock 2009: 155). This analytical "in reverse" structuring can take place also in a purely creative manner, a listener may want to visualize the music not in accordance with scientific aims but arriving at more artistic outcomes applying a mapping between music and visualization that may be more loose, intuitive, subjective, emotion-based while working more on the artistic and even independent artistic value of the visual part. The way of a scientific visualization depends on the affordance of the style of the music, that of an artistic visualization rather depends on imagination and the desired style and quality of the visuals. Automatizations can be applied to a "classical" music theoretical and a software-based analysis (e.g. MIDI, sonogram etc) of music as well as for empirical data retrieval and for the related visualization of that data (see publication II, Figure 2, see more in 5.3).

Music-theoretical thinking approaches (see 2.1) mainly visualize structural musical parameters separately or in combination. Music-psychological thinking approaches (see 2.1) often visualize imaginations or emotions that are evoked by the structural musical parameters but the relations to these parameters may be weak; and it may be a deliberate aim of the research to explore them. Artistic thinking approaches often visualize imaginations and emotions whereby the relations to the musical parameters that have triggered them may stay on a simplifying surface level. See an overview of graphical approaches to music analysis in the context of energetics, phenomenology and *Gestalt*-analysis in music in Rothfarb (2002) (see 2.5). The relation between parameters and their visualization may remain either unexplored or explored in its function as inspiration (e.g. with improvisation, Mailman 2013). Furthermore, artistic "sublime" visualization (in music) may become unrecognizable and not readable in contrast to pragmatic "utilitarian" visualization (Wikgren 2019: 29). In the latter case the music and the visuals unfold quasi-independently from each other, however they are mediated by the mind of the visualizer, so, there is still a connection between them. However, actual methods developed and explored in data visualization and analytics (e.g. Viégas & Wattenberg [2025]) and within the uprising artistic and creative research field enable to observe, document, describe and analyze diverse types of data-based relations between the arts and the process of their creation (e.g. Li 2018). In the next subsection the data analysis of this dissertation conducted with the DBSCAN algorithm clusters is presented.

4.6 Data analysis with the DBSCAN algorithm clusters

This subsection deals with the data analysis principles applied in this dissertation. In general, not specifically in music, pattern mining and anomaly detection in heterogeneous, multi-dimensional time series and event logs have been challenging in computer science until recently (Fu 2011, Feremans *et al* 2019). Furthermore psychological research has integrated aspects of temporal dynamics to overcome limiting constraints in data gathering especially, “however, these practical limitations do not eliminate the *theoretical* need for understanding patterns of change over long periods of time [...]” (Jebb *et al* 2015: 1). Music in fact is heterogeneous, multi-dimensional and (as whole works) longer lasting based on temporal dynamics (Mailman 2010). One can understand the salience of musical events treated as Impulses (“moments of change”) and musical parameters as kind of anomalies (Feremans *et al* 2019).

This subsection presents some **DBSCAN: Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise (a data mining) analysis principles** (Ester *et al* 1996) (as a short version in English, with examples visualized and explained in publication IV) and the general characterization (statistics) of the derived clusters. This enhances the understanding of how the data gained from COSM can be treated further based on such clusters. While acknowledging that computer software has taken a significant role in computational musicology (Dahan 2012) my proposed system of music analysis methods opens up the whole range of possibilities: an analyst can approach music either with ‘intuitive,’ ‘systematic,’ ‘system-based’ or ‘automatized’ methods (publication II, on automatization see 4.5).

Automatization in data retrieval and analysis of music (see publication II, Chai 2005, Nieto 2015) takes place at different levels of the process: data retrieval, input-data representation, further treatment, output-data representation and further analytical treatment (in the context of modeling see publication III, see Appendix B). At the representation level there are several stages where different types of automatization aids can be applied (publication II), too. Among the cross-doman aspects appearing important in music and data is *density* (for music see e.g. Kirschbaum 2001, Addessi & Caterina 2005, Addessi 2010). Especially in contemporary post-tonal orchestral music (CPTOM) one finds texturally both loose and dense parts, the same can appear in the retrieved data derived from observing this music. Density therefore is an important aspect to analyze. The DBSCAN has been chosen for data analysis in this dissertation aiming at formalizing the “intuitive notion of “clusters” and “noise” in a database” (Ester *et al* 1996: 227). The authors point out that “for each point of a cluster the neighborhood of a given radius has to contain at least a minimum number of points, i.e. the density in the neighborhood has to exceed some threshold” (Ester *et al* 1996: 227). Expressed simplified, the DBSCAN algorithm detects points that belong (are density reachable) to the cluster, called core points (p), and points that are located at the border of the cluster (indirect reachable), called border points (q). Points that are not reachable from the cluster, are outliers, don't belong to the cluster. The threshold MinPts (minimum number of points in the cluster) and the Eps (ε is the maximum radius of the neighborhood from p), in this case the time span that has been chosen in this study, is derived from the number of participants in the empirical study (N=14) (visualized examples see in publication IV).

In the main study of this dissertation (publication IV) I have chosen a highly significant threshold of 10 responses and time points created by the participants (N=14). This ensures that the event that has been indicated is of high significance among the sample. The *Eps* (maximum radius between core points) is chosen 2 seconds, because this lies within the time-span of the perceivable Present (Snyder 2000: 26) and even below the STM (Short Time Memory) average time span (3–5 seconds) (see Figure 1 in publication IV). This corresponds with the chunk/window length 14 seconds (+/–7 sec) provided in COSM around each Impulse. A wave-form-based (Audacity) real time-span-based annotated example from Experiment II COSM **part 1** (see Figure 4 in publication IV) visualizes with circles and braces how many points with their radius establish the Cluster 21.

The DBSCAN algorithm was applied[106] with the following features: 2000 ms = 2 seconds and 10 = minimum of 10 events required to be a core point (p) within 2 seconds. This has generated 48 clusters among the sample of N=14 (see Figure 5). The cluster numbering follows their chronological order in the course of the music. As follows I use for the core and border points of the DBSCAN algorithm the term *event*. These are practically the salient Impulses chosen by the participants in the COSM experiment **part I**.

---

[106] Thanks to Ilkka Kosunen, doctoral student at Helsinki University and working at this time at Tallinn University (TLU) Digital Technology Institute (DTI) and at the Baltic Film, Media and Arts School (BFM) e.g. with Dr. Aleksander Väljamäe and Prof. Dr. Pia Tikka.

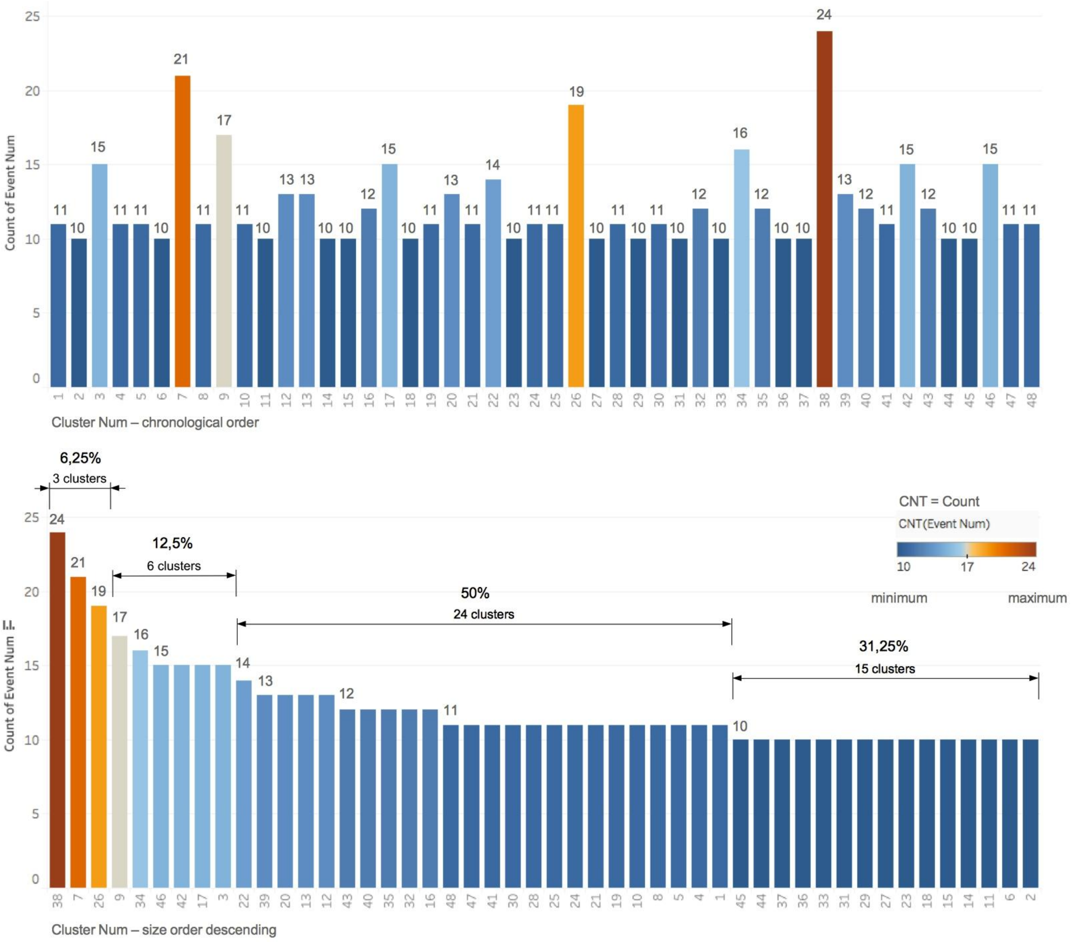


Figure 5. COSM DBSCAN (10 events, 2000 ms, Ester *et al* 1996) general statistics (experiment part I detecting Impulses): 48 clusters – count of events. The upper graph shows the cluster numbers in chronological order, the lower graph orders them in descending size. The graphs are generated with the Tableau visualization software,[107] but percentage and cluster summation details I have added myself manually. This figure was presented first in Lock (2019, 2023a).

In Figure 5 three clusters (number 7, 26, 38) (6,25%) show a significantly[108] higher cluster size (count of events/Impulses) than the required minimum threshold (10). They show almost 20 responses (19) and more than twice (20 responses → 21, 24) responses (this is more than the number of participants): Cluster 7 (21 events/Impulses/responses), Cluster 26 (19

---

[107] Thanks to Tallinn University Digital Technology Institute (DTI) doctoral student Mustafa Can Özdemir.

[108] This doesn't mean that the events/Impulses are marked only by different participants. There can also be events/Impulses indicated from the same participants within the same cluster. Significance means that the number of events/Impulses within a cluster is: 1) nearly or more than double of the DBSCAN required number of events (10), 2) is larger than the number of participants (N=14) of this study. Hereby shows the count of 17 the middle between the minimum 10 and the detected maximum 24.

events/Impulses/responses), Cluster 38 (24 events/Impulses/responses). Further there are six clusters (12,5%) that show slightly more events than the number of participants (14): Cluster 3 (15 events/Impulses/responses), Cluster 9 (17 events/Impulses/responses), Cluster 17 (15 events/Impulses/responses), Cluster 34 (16 events/Impulses/responses), Cluster 42 (15 events/Impulses/responses), Cluster 46 (15 events/Impulses/responses). One can also find 24 clusters (50%) that show 11 to 14 events/Impulses/responses (identical with N=14 but not necessarily detected by 14 participants). 15 clusters (31,25%) show exactly 10 events/Impulses/responses (the minimum required). This means that a third (1/3) of the clusters (15) fulfill minimum requirements of the algorithm settings. Half (1/2) of the clusters (24 clusters) show an event/Impulse/response number larger than the minimum, but formally not larger than the number of participants. An eighth (1/8) of the clusters (6 clusters) formally shows more events/Impulses/responses than the number of participants – up to half of the events/Impulse/responses numbers in this sample. A sixteenths (1/16) of the clusters (3 clusters) shows a significantly higher number of events/Impulses/responses. As an even more generalizing results one can say that roughly a fifth (1/5) (18,75%) of the clusters (9 clusters) shows formally more events/Impulses then the number of participants (N=14) and invite the researcher to look more in detail into their content and categories of musical parameters attached to them. Roughly a fraction of 1.23 (81,25%) of the clusters (39 clusters) show events/Impulses/responses in the range between the minimum of 10 events/Impulses/responses (lowest threshold number to ensure significance) and the maximum of participants (N=14). One can even further simplify this result saying that only a sixteenth (1/16) (6,25%) of the clusters (3 clusters) shows a significantly higher number of events/Impulse/responses worth analyzing its content (musical parameters) in detail. The majority of the remaining clusters, a fraction of 1.06 (93,75%) of the clusters (45 clusters) have been perceived as less or little significant between the minimum of the required 10 events/Impulses/responses and the maximum of 24 detected with this sample in this music. In addition see another visualization of the clusters in Figure 6 that shows Impulse summas per cluster and the length/shortness of the clusters and events/Impulses within “common seconds.”

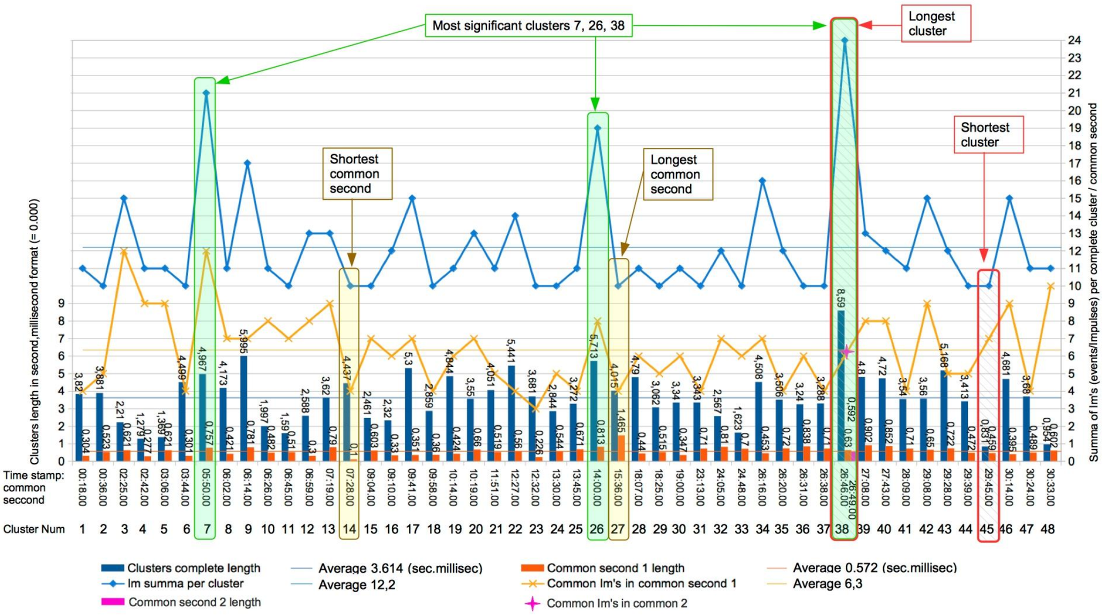


Figure 6. COSM DBSCAN (10 events, 2000 ms, Ester *et al* 1996) detailed statistics of the 48 clusters (experiment part I detecting Impulses) in chronological order (visualized with OpenOffice Spreadsheet). The figure shows cluster length and common second length within clusters (bar lines) together with total Impulse summa and Impulse summa within common second per cluster (line graphs). The graph is annotated with info concerning significant clusters concerning their length/shortness. This figure was first presented in Lock (2019).

The following section 5 presents in extended summaries the goals, methodology and relations of the included publications aiming to enable the reader a concise overview of the development of and between the studies that I have bound together in this dissertation.

5. Included publications: goals, methodology and relations

This section contains extended summaries of the included publications in English. One reason to provide these English abstracts is that part of the publications are issued in other languages: Estonian and German. Another reason is that publication I (Lock 2010a, in Estonian, 5.1, A.1), publication III (Lock 2017, in English, 5.4, A.5), publication IV (Lock 2020, in German, 5.5, A.5) and publication V (Lock 2024, in English, 5.6, A.6) as well as publication Ia (Lock & Kotta 2012, in English, 5.2, A.2) have only very short English abstracts included in their first printed versions. These short abstracts do not open the content sufficiently if the reader of this Introductory chapter seeks for a compound overview of the publications of this dissertation. Publication II (Lock 2014, in German, 5.3, A.3) has no abstract.

This subsection should enable the English reader to gain compound and extended insight into the topics, components, methods and their development and relations relevant for this dissertation without the need to contemplate previously on the separate publications. Publication V has an embracing function: it formalizes the five scientific Dilemmas, Research Questions (RQs) and Statements (STs) and Answers that collocate the methodological issues that have motivated this dissertation to finally arrive at such an empirical analysis methods of musical tension as a temporal dynamic wave-like phenomenon (TDWP).[109] Furthermore, the main study, publication IV, presents for the first time the theoretical-methodological principles and first detailed results including combining TEDEA and COSM data: analyzing musical tension empirically via detection/measurement of salience of Impulses (Im) and musical parameters interpreted through the lens of narrativity. However, to ensure a better insight into the capabilities of COSM I have added some extended results (see Appendix D). Finally, subsection 5.7 describes the methodological relations between the publications.

---

[109] The term-phrase *temporal dynamic wave-like phenomenon* (TDWP) was introduced in publication V (Lock 2024).

5.1 Publication I (Lock 2010a, A.1)

This journal article (1.2, in Estonian, short abstract extended) on musical tension, musical form and salient features (*Oxymoron*) is my first analytical/theoretical and empirical study on *musical tension* applying a "salient feature" analysis approach.[110] The summary presented here is extended, because the whole article is in Estonian and the English abstract of the original article is too short to give sufficient insight. This study treats musical tension[111] as a complex/compound phenomenon in relation to musical form. Its general aim is to develop a method to describe and integrate the structural aspects that cause the perception of musical tension especially in contemporary post-tonal music by Erkki-Sven Tüür. In this publication so-called salient features are introduced as a speculative subjective theoretical approach (see 1.2, 2.1) – local level: Impulse [Im], Culmination [Cu]; global level: Contrast [Co], Contrast-Culmination [CC]). The research questions were the following: i) what role play culminations in the form of the work (more concretely in its final section); ii) how the perception curves correlate among the participants; iii) how the score- and listening-based form analysis (applying the aforementioned salient features) relates to the average curve of the perception test experiment? As a hypothesis I proposed that the listeners perceive the development of tension in music in general and at important areas and points of the musical form similarly. This furthermore suggests that the increase and decrease of tension are generally relatively objective phenomena and it would be possible to develop a system-like analytical method I called "tension design" (see also Lock & Valk-Falk 2008). This approach assumes that tension is both a structural (music theoretical) and perceptual (music psychological) category therefore requiring a combined analytical-empirical approach. This article compares empirical results of a slider-controller listening experiment on musical tension (N=7, 4 music experts, 3 non-

---

[110] My initial idea to characterize a musical event as "impulse" or "culmination" while using these terms for the analysis of musical tension was further systematized by my supervisor Dr. Kerri Kotta and refined by my former Master's Thesis supervisor Prof. Mart Humal, suggesting the term "contrast-culminations." While the idea of salience was obtained from an article by Lerdahl (1989) on the salience of musical parameters in atonal music, Kotta suggested to call "impulses" and "culminations" on the local level "salient features." Global level "salient features" ("contrast" and "contrast culmination") complemented this approach towards an analytical system that also includes a macro-form level. See this "salient features" study also in English in an unpublished extended abstract of the conference paper Lock (2010b).

[111] In this publication the term phrase *musical tension* appears together with structural aspects causing it and musical tension as structurally important phenomenon; several musical parameters are mentioned in connection with musical tension, especially based on Kirschbaum (2001) and Jere Hutcheson (1972). The aspect of perception as perceived is described in the procedure of the experiment with the aim to investigate *perceived tension*, not *felt tension* (see publication IV, Lehne & Koelsch 2015).

experts, average age 31,7) with the results of the intuitive and score- and listening-based approach analyzing significant local (Im and Cu) and global salient features (Co and CC) based on the visual and auditory imagery[112] information provided by the musical score, and seeking confirmation through repeated listening to the recording while relating it explicitly to the musical form and implicitly also to the role of musical parameters. The object of analysis and a perception experiment with slider-controllers (2008, N=7) is the final section (08’34’’) of Erkki-Sven Tüür’s work *Oxymoron* (2003, 19’40’’) for large ensemble, the first work composed completely using his new “vectorial” method (see 1.1). Musical parameters obviously are important in the salient feature analysis, but in this article their role could neither be separately described nor explained, the formalization focused on the generalization of the salient musical events as such. Also in the slider-controller results the musical parameters as potential triggers of the change in the tensional development detected by the participant remained unreflected and could not be analyzed. *Oxymoron* has no clear (classical form-like) boundaries, but, as a result of this study, I suggest that formal parsing of this work can be achieved based on maxima and minima of the musical tension development. The simplified results (x – static standardized 10 seconds time window, y – 0–100 intensity level/arbitrarily defined salience values 10, 20, 30, 35, 40 for the weight of the salient features) show that the single tension curves of the participants develop rather individually, but show lower and higher correlations in several areas of the tensional development. This correlation can be illustrated by the standard-deviation curve at several points during the experiment: its lowest level (below 10) appeared between 5’35’’–6’35’’ (approx. 60 seconds) before the main culmination area between 6’35’’–7’15’’ (approx. 45 seconds) (see Figure 2 in publication I). The average-curve maxima of the participants coincided with the simplified salient-feature analysis graph derived from my own score- and listening-confirmed speculative analysis (see Figure 1 in publication I). The hypothesis of this article was confirmed partly based on the correlating areas and salient-feature points just mentioned. But the low number of participants (N=7) as well as the short (and incomplete) object of analysis (final section only) did not allow a broader generalization towards whole works, yet. However, referring to the wave-like character of Tüür’s post-tonal *Oxymoron*, the results allow me to draw the conclusion that the formal boundaries of this music are constituted rather based on maximum and minimum points of the tensional development. Furthermore, this

---

[112] The term ‘auditory imagery’ (Phillips 2018) is applied in the late phase of this dissertation. It offers theoretical background for the way how silent score analysis might be understood that I did not have at the time of this publication.

study posed a number of methodological questions (separately formulated in Appendix C.1) and the development of a envisioned “tension design” method was abandoned at this point. Upcoming studies should be conducted with complete musical works as well as an increased number of participants. The following methodological aspects remained in use for further investigations: the salient feature “impulse” and the slider-controller listening experiment method.

Keywords: musical form, salient features, perceived structural musical tension, slider-controller experiment results, final section of *Oxymoron*.

5.2 Publication Ia (Lock & Kotta 2012, A.2)

This rather theoretical conference paper (3.4 in English, short abstract extended) on musical tension, musical form, salient features and data reduction (*Oxymoron* and 4th Symphony/Percussion Concerto *Magma*) firstly explains the visualization principles of the salient features analysis applied in publication I (in English see Lock 2010b) juxtaposed with commented musical-score examples from Tüür's *Oxymoron* (2003). Secondly, it introduces the basic principles of a hypothetical (visual, at this moment yet manual) algorithm-based method of reducing and averaging tension curves[113] obtained from listening tests using slider-controllers for “continuous” data capture of (complete) musical work to enhance mass-data analysis. In addition, new empirical data (collected 2010, N=6, musicians, artists, dancers)[114] was used to create a real-data example based on Tüür’s 4th Symphony/Percussion Concerto *Magma* (2002). It shows how the algorithm-based reduced data of several participants (appearing on different levels of detail and meaning like in Schenkerian analysis) could be meaningfully joined using a 15-seconds moving-window principle. However, this continuing study posed also a number of further methodological questions (separately formulated in

---

[113] During the conference Dr. Toby Gifford (Queensland Conservatorium, Griffith University, Australia) agreed to program the algorithm with Java/Max MSP, but this took several years (until 2015) and bugs remained unsolved. Furthermore, Dr. Joshua Mailman mentioned at this conference that this algorithm is similar to the melody-reduction method proposed by his doctoral supervisor Prof. Dr. Robert Morris (1993), see more on melody reduction in Bor (2009). Also, Mailman sent me later his extensive dissertation (2010).

[114] See 4.3 FN 92, 4.4 FN 105, 5.7. Behringer controller (Max/MSP software by Hans-Gunter Lock) for single participants. The TEDEA apparatus (Lock, G. & Lock, H.-G. 2011) was later applied (experiments 2017) for publication IV (Lock 2020).

Appendix C.2) and required me to look for more appropriate solutions. The following methodological aspects remained to be solved in upcoming studies: what type of algorithm is useful to reduce mass-data, what type of window-principle is useful to meaningfully join several participants' tension curves.

Keywords: score-based visualization of salient features, algorithm-based analysis of complete musical works as mass-data retrieval (manual principles), analytical and empirical (slider-controller) analysis examples of *Oxymoron* and 4th Symphony/Percussion Concerto *Magma*.

5.3 Publication II (Lock 2014, A.3)

This book chapter (3.1 in German, separately generated abstract) on visualizing as intuitive, systematic, system-based and automatized music analysis (*Oxymoron*) is a theoretical overview of basic principles of 2D visualization of music in my proposed framework of intuitive, systematic, system-based and automatized analysis of music. It embraces prescriptive (musical score) and descriptive forms (e.g. diagrams, sonograms) of representational/visualization of music in a framework from undefined, symbolic, referential and proportional to precise, discrete and "continuous" definitions of x-/y-/z.axes (see Figure 1a in publication II). I take into account that (in analogy to modeling) only a chosen part of the available information (music theoretical, perceptive and cognitive aspects) detectable in the music can be visualized. I discusses didactical aspects and offer a novel system of levels of automatization tools (both data retrieval and visualization of musical tension via slider-controllers) for visualizations between the lack of and maximum invasive (control) opportunities of the analyst concerning the automatization result either in real-time or retrospectively. In Figure 2 in publication II I describe and visualize possibilities of scientific automatization during the analysis- and visualization process of music as depending on either process-related *mediate* (static, non-automated *vs* dynamic automatizable/automatic) or *immediate* (in real-time, automatic) visualization methods. *Static* means that the processes leading from the Analysis Object (AO) to the Visualization Result (VR) are not related timewise, no feedback is enabled between the analyst (A) and the Automatization Helper method (AHm). *Dynamic* means that these processes are not related timewise, but include a time component taking minimal time or are delayed, feedback is enabled between A and AHm. *In real-time* means direct active/passive-invasive feedback between A and AHm during processes leading from AO to VR.

This chapter presents visualizations of the main culmination area (bars 387–391) of Erkki-Sven Tüür's large ensemble piece *Oxymoron* (2003) spanning from summarized traditional notation (*particell*, Figures 3a–c in publication II) to MIDI (see Figure 4 in publication II), timbre-groups (see Figure 5 in publication II) to spectral (sonogram) analysis (musical and physical sound parameters) (Figure 6 in publication II) as well as a comparing salient features and slider-controller-based tension analysis (N=7) with average and standard-deviation data from publication I (Lock 2010a&b) (see Figure 7 in publication II).

Keywords: basic principles of representation/visualization of music; intuitive, systematic, system-based and automatized methods (data retrieval and analysis); a novel system of automatization in visualisation processes; analytical, automatized and empirical (slider-controller) analysis examples of *Oxymoron*.

5.4 Publication III (Lock 2017a, A.4)

This journal article (1.2 in English, long abstract) is a theoretical overview of basic principles of models and scientific modeling in music, the sciences and education that, as the final outcome, introduces my Twelve Strategic Steps for modeling/analysis of models (TSSM) in eight stages/phases (posed as questions) based on which the COSM: Cognitive Octagonal Slice Model has been developed.[115] The object of modeling in this article, music, can be understood in analogy to the natural environment (Reybrouck, 2015) and as a dynamical system (Burrows, 1997). Scientific models have diverse functions (aims, purposes), but also constraints, they work with different methods. They derive from implicit and explicit knowledge, appear between cognitive experience and implicit and explicit data within the paradox of knowledge as 'world first' *vs* 'mind first,' manifested in two extremes: existence and technique (Varto, 2009). In the theoretical part, the author introduced his multi-dimensional scheme of modeling and models incorporating models from music theory (analysis) and music psychology (perception and

---

[115] COSM's initial manual version was first presented in an unissued conference paper (Lock & Kotta 2015), its first full application as a computerized version applied in empirical experiments see publication IV.

cognition) including the creative approach (musical composition/improvisation). Its aim is to position modeling used in (and for understanding of) music within the context of modeling in science in general. Also, the relation of models to theory, as well as computational and cognitive models in music are discussed briefly. In the main contribution, I introduce my Twelve Strategic Steps for modeling/analysis of models (TSSM) (partly inspired by Purwins *et al* 2008a), present my enlargement of Rosária Justi and John Gilbert's (2002) "Model of Modelling" and compare it with the TSSM. In the implications and conclusion, I propose that modeling in science is resultful if an idea can become explained with a concept, which can then be approached systematically, further developed as a method, includes connections/backups to one or more theories, and is consistent (a whole), either representing or exploring a known (but less systematically explained) phenomenon including cross-domain mappings. I hope that this article enhances the understanding of modeling and assists in creating models in sciences in the field of music and music education.

Keywords: scientific modeling, strategic steps (stages/phases) to develop/analyze a model.

5.5 Publication IV (Lock 2020, A.5)

This journal article (1.1 in German, short abstract extended) on musical tension, salience, narrativity and musical parameters (4th Symphony/Percussion Concerto *Magma*) explores cognitive listening strategies (attentive listening, salience, musical event as Impulse), perceived structural musical tension,[116] musical parameters (their salience in post-tonal orchestral music (PSMT) in comparison to tonal music), narrativity (as cross-disciplinary and meta-approach) and presents the theoretical background and functioning principles of my new cognitive empirical analysis model COSM: Cognitive Octagonal Slice Model. This study poses two research questions/hypotheses: 1) What is the role of musical parameters: are secondary more salient than primary parameters? 2) What is the possibility of narrative interpretation (based on cluster finding) for the role of musical parameters in contemporary post-tonal music?

---

[116] This term phrase was developed in the late phase of the dissertation in German as "wahrgenommene strukturelle musikalische Spannung" (publication IV). Since Lock (2022, development of the definitions, terms and term phrases around the basic term "musical tension") it is defined in Estonian as 'Tajutud strukturaalne muusikaline pinge' (TSMP), in English as 'Perceived structural musical tension' (PSMT), see the definition in English in 3.5 and in publication V.

Characteristic for Tüür's music since 2002/2003 is that transitions between different states/textures/temporal organization of the musical material are dominating instead of contrasting them like in earlier music by Tüür (see also 1.1). Despite of this surface treatment different from classical music, the 4th Symphony/Percussion Concerto (2002) can be analyzed based on the three types of musical time (continuous, slow, quick) characteristic für Tüür's music either as sonata cycle movements or detecting sonata allegro formal section features based on their rhetorical and rotational structure (Kotta 2011). In this article the data of the slider-controller listening experiment (TEDEA)[117] is compared with the data of the COSM experiment using the DBSCAN (data-mining) algorithm (Ester et al 1996)[118] and several other ways of data visualization. In the results the data of empirical slider-controller listening experiments – TEDEA, perceived structural musical tension, N=26, music experts (n=4), music (n=16) and arts students (n=6) – is combined with the data of COSM – salience of part 1 Impulses (Im's)/"moments of change", part 2 salience of their 'content': musical parameters; N=14, experts (n=4) and music students (n=10) – conducted on Erkki-Sven Tüür's 4th Symphony/Percussion Concerto *Magma* (2002, 30'58''). The results show that the average curves of the slider-controller experiment by experts, music and art students are similar in global tendency, but differ significantly at the end of the work: musicians and especially experts are obviously more used to keeping the tension until the real end. Concerning the role of the musical parameters in relation to the musical tension the results show not many clear data patterns[119] and the "secondary" musical parameters (SP) are not more salient than the "primary" musical parameters (PP).[120] Only in clusters 21–23 (from 48 clusters) located before the percussion solo cadenca in the middle of the work the hypothesized higher salience of "secondary" over "primary" parameters could be detected clearly based on their correlative in- and decreasing tendency within the clusters that partly coincide with the same tendency in the averaged musical tension results for the three participant groups. The meaningfulness in the data (tendency of changes in tension in correlation with changes in the salience of musical parameters) is interpreted via the following data-mining principles: more clear regularity (invariance) means higher level of narrativity. Meaningfulness (see 2.4) is part of the narrative interpretation of the

---

[117] Tension Design Experimental Apparatus: soft- and hardware built by Hans-Gunter Lock (Max/MSP, Arduino) (Lock, G. & Lock, H.-G. 2011).

[118] On data analysis with the DBSCAN algorithm clusters see 4.6,

[119] On meaningfulness in psychophysical measurement see 2.4, on data analysis and visualisation principles see 4.5.

[120] "Primary" musical parameters (PP): pitch, rhythm, harmony; "Secondary "musical parameters (SP): dynamics, tempo, instrumentation/timbre, texture, effects. See more on musical parameters in 3.4.

data, realizing Gruhn's (1992) suggestions of a listening story (*Hörgeschichte*) as an interpretive and sense-giving narrative level connecting perception- and cognition-processes to enable a cross-domain/meta-domain understanding of perceived structural musical tension in relation to the role of musical parameters and musical form. Although this study did not reveal a clear global teleological pattern/narrative in relation to the salience of musical parameters, interesting patterns/narratives at local level[121] have appeared that need further analysis. Some overall results of the COSM experiment have been previously presented in two unpublished papers (Lock 2019, 2023a) but have not been issued in a printed publication,[122] see Appendix D.

Keywords: perceived structural musical tension, cognitive listening strategies, musical form, salience of impulses and musical parameters, cross-domain understanding through narrativity, results of empirical experiments (TEDEA and COSM) on 4th Symphony/Percussion Concerto *Magma*.

5.6 Publication V (Lock 2024, A.6)

This conference paper (3.4 in English, short abstract extended) on five scientific Dilemmas, Research Questions (RQs), Statements (STs) and Answers intertwines music analysis and music psychology with a cognitive approach to better comprehend contemporary post-tonal orchestral music. It presents five scientific Dilemmas crucial to achieve an empirical analysis of musical tension as a temporal dynamic wave-like phenomenon (TDWP). Considerable theoretical and philosophical-methodological issues are 'experience' and 'intuition,' an enlarged Kantian understanding of perception and cognition while measuring inner sensations via outer sensations (magnitude detection) (Kraus 2013/2016, Wolff 1963), "relative change," audiovisual salience and psychophysical measurement principles (Shepard 1978; Steingrimsson 2016; Michell 2020; Vessonen 2021). Based on the Dilemmas five Research Questions (RQs) and Statements (STs) are presented and their Answers are provided. As a result I have developed

---

[121] On *protonarratives* see 2.6, on narrativity in music in 3.5.

[122] In 2019 I was invited by the RATM Journal to issue such results, but at the same time I continued compiling the also invited publication IV (Lock 2020) that aimed to provide (in a longer article) basic methodological insight and the focus on narrativity as cross-domain data interpretation. Further developing analytical results based on these general results (see Appendix D) still remain in a preliminary state – this is work in progress and should be issued later.

and applied a new model – COSM: Cognitive Octagonal Slice Model – which's outcome data (in combination with other methods) is suggested to be analyzable via narrativity in order to comprehend Cognitive Musical Tension (CMT). I assume that *Sound* and *Music* may appear in several domains,[123] but *Music*, based on human perception and cognition via the *auditory* (perceived) domain, therefore via listening is best comprehended as cognitive (Wiggins 2009). Music has been analyzed cognitively, e.g. with principles of "immanent musical analysis" during listening as elaborating/structuring process of "relevant categories" (Chouvel 2014) and modeled computationally as an ecologist cognition/knowledge of music based on of the individual's mind's cognitive dynamics (Kaipainen 1994). *Cue* abstraction and information dynamics have been dealt with while modeling cognitive music analysis (Wiggins 2010). An overview of computational modeling of music cognition and the process of musical composition see Wiggins, Pearce and Müllensiefen (2012). The first Dilemma/RG/ST/A encompasses musical tension as cognitive (CMT). The second Dilemma/RG/ST/A outlines that intensity or energy measured in music do not equal cognitive musical tension (CMT). The third Dilemma/RG/ST/A states that a culmination "arch" based on maxima and plateaux of CMT can be detected well via salience, but not its beginning or ending points. The fourth Dilemma/RG/ST/A claims that CMT can only be experienced based on at least two time points (musical events/Impulses) as perceived "relative change" of musical parameters (MPs) of TDWP. The fifth Dilemma/RG/ST/A deals with the 'content' of the Impulses/"moments of change" – musical parameters (MP's) – and possibilities of their salience weights detection applying intuitively the psychophysical measurement principle "just noticeable differences" (JND) in the form of "relative change" with COSM (see publication IV).

Keywords: Cognitive Musical Tension (CMT), musical tension as a temporal dynamic wave-like phenomenon (TDWP), enlarged Kantian understanding of perception and cognition as forerunner of psychophysical measurement (inner *vs* outer sensation), psychophysical measurement principles, Impulses (Im's), musical parameters (MP's), *precategorial objects**, *protonarratives*.

---

[123] The *graphemic* (or notated/also digitally encoded), the *acoustic* (or physical), and the *auditory* (or perceived) domain (Wiggins 2009: 477 after Babbitt 1965, 2003: 204, see 1.3, 2.2, 4).

5.7 Relations between the included publications

In this subsection the relations between the publications of this dissertation are explained based on the following categorization into realms, approaches and phenomena. Music theory, music psychology, perception/cognition and cognitive science; modeling and cross-domain cognitive phenomena: musical tension; salience; musical event as Impulse; narrativity. As stated in the introduction of this Introductory chapter (see 1.2, 1.3), representation and visualization in music (publication II, A.3) play an important role in modeling and cognitive science; modeling itself is part of cognitive science. This subsection shows why these publications have been chosen for this dissertation and what questions and reasons led from one publication to the next. These questions and reasons will be opened especially concerning the development of COSM: Cognitive Octagonal Slice Model. More details see in section 4 based on questions proposed in the Twelve Strategic Steps for modeling/analysis of models (TSSM) in eight stages/phases (publication III, A.4).

The methodological trajectory of this dissertation to explore musical tension can be subsumed as follows: Publications I and Ia (a) explore the cognitive aspect of salience relevant for comprehending perceived structural musical tension using a speculative and intuitive score- and listening-based music theoretical approach (see 1.2) to analyse “salient features” (see I, 5.1, A.1), (b) introduce a slider-controller empirical experiment of musical tension with data analysis of not yet a complete work (final section of Tüür’s *Oxymoron*) while presenting short score-based examples for illustrating salient features and an abstract “continuous”- data-capture-curves reducing (visual and manual) algorithm (see Ia, 5.2, A.2). This has revealed challenges concerning data analysis in general and questions of comparison of data of different methods especially. The outcome data analysis of both methods has been resolved with simple mathematical means (simple smoothing, averaging, standard deviation) and needed further improvement (see Appendix C). Publications II and III give methodological insights into the functioning of two important principles in music analysis and science: visualization/representation of music – presenting automatization approaches and short musical examples from Tüür’s *Oxymoron* (based on data from publication I, see II, 5.3, A.3) – and modeling/analysis of models in science and music – including a commented model of modeling, a systematization of models in music and science and my Twelve strategic steps for modeling (TSSM) (see III, 5.4, A.4). Publication IV (see 5.5, A.5) introduces theoretical aspects for the comprehension of musical tension – via salience and narrativity – as well as empirical results of my new model COSM that are compared with TEDEA slider-controller results using DBSCAN algorithm clusters. Publication V presents the basic understanding of musical tension as cognitive (CMT) and as a temporal dynamic wave-like phenomenon (TDWP), it also provides theoretical aspects of an enlarged Kantian understanding of perception and cognition (based on experience and intuition) as well as psychophysical measurement principles. With its Dilemmas, Research Questions (RQs), Statements (STs) and Answers it collocates matters presented in the previous publications and can be considered as a condensed presentation of the methodology of this dissertation.

Publication I introduces musical tension in relation to musical form and applies a preliminary speculative and intuitive score-based and listening analysis method called salient features analysis of musical events (Impulse, Culmination, Contrast, Contrast-Culmination) as well an empirical method – "continuous" data-capture with slider-controllers (intuitive listening approach) – to analyze perceived musical tension. The aim of this study was to analyze how musical tension is perceived in real-time experiments with a slider-controller claiming that musical tension is a structural and a perception category. Analysis and experiments (N=7, data retrieved 2008) have been conducted on Tüür's *Oxymoron* (2003) (final section 8'34''). The methods applied and the theoretical background are taken from music theory and music psychology. Musical parameters are important in the salient features analysis, but in this publication their role is neither separately described nor explained, the formalization focuses on the generalization of the musical events as such. Also in the slider-controller results the musical parameters as potential triggers of the change in the tensional development detected by the participant remain unreflected and are not analyzed. The results showing partly correlation between slider-controller and salience values are presented in a joint diagram with simplified data visualization: x – static standardized 10 seconds time window, y – 0–100 intensity level/arbitrary salience values 10, 20, 30, 35, 40. Here the idea of an arbitrary static time window enabling comparison of data from different methods is first introduced. Methodological questions see Appendix C.1.

Visualization uses a simplified 2D diagram (time window *vs* tension intensity/salience value). Concepts used and developed further: salience, musical event: impulse. Methods used and developed further: Impulse detection (later only as listening analysis), slider-controllers; static time window idea.

Publication Ia extends publication I offering methodological insight into the salient features analysis method: it firstly shows examples of how the cognitively (combined score-based and listening analysis) detectable musical events (Impulse, Culmination [local level]; Contrast, Contrast-Culmination [global level]) are detectable in the musical score and via listening. In this article the role of musical parameters is shown (but not discussed in detail further) in the score examples that visualize the chosen musical events exemplifying the proposed salient features types. Secondly, this publication introduces a theoretical method of analyzing the data of slider-controller musical tension curves suggesting, how multiple participants' curves could be joined together via a high-/low-point (maximum/minimum) reduction and averaging algorithm to achieve meaningful comparison of results derived from observing longer whole works and a larger number of participants (here still manual). Beside abstract reduction/averaging examples also real-data results of a next empirical study (N=6, data retrieved 2010, analyzed with trend/regression tools and rhetorical/rotational from analysis in Lock 2019/2011, see) on Tüür's 4th Symphony/Percussion Concerto *Magma* (2002) (the whole work, 30'58'') are presented, using a 15 seconds static or moving window to join the participants' curves together (see 4.3 FN 92, 4.4 FN 105, 5.2). However, the unsolved bugs of the software made it ineffective for this dissertation. Methodological questions see Appendix C.2.

Visualization uses a simplified 2D diagram (time window *vs* tension intensity/salience value), conceptual visualization of the reduction principle of points of a line-curve, annotated musical score examples. Concepts used and developed further: salience, musical event: Impulse. Methods used and developed further: moving time window idea.

In the above described two publications I have attempted to compare results from an intuitive music theoretical (speculative) score- and listening-based salient features approach (conducted by a single analyst, aspects "mediately" relatable to musical tension) with intuitive slider-controller empirical experiments (multiple co-analysts, immediate musical tension detection). However, these approaches seemed to be ineffective methodologically because the methods applied have been too imprecise and different from each other, results have been too complicated to be meaningfully compared. They did not enable me to get sufficient data neither on concrete salience nor on concrete musical parameters (see questions after publications I and Ia in Appendix C). In a next step I contemplated on methodological principles of music analysis on the one hand (principles of visualization/representation of music) and of scientific inquiry on the other hand (basic principles of scientific modeling in music).

Publication II offers methodological insight into the principles of visualization/representation of music. Its aim is to systematize visualization/representation issues from the musical score (prescriptive) to several analytical approaches (descriptive) and to understand in what way such methods (intuitive, systematic, system-based and automatized analysis) may contribute to this dissertation. These methods concern music analysis/theory and partly also music psychology (visualization of empirical data). Aspects of cognition (how the human mind works via representation and modeling) are touched mainly through visualization/representation, partly through modeling, because (similar to modeling) a score and visualized analysis show only chosen elements of complex contemporary orchestral music. In the case of this dissertation so far salient features (publication I) and partly musical parameters – musical pitch, harmony and rhythm ("primary" parameters) as well as dynamics (sound pressure level, "secondary" parameter) (publication Ia) have been visualized/represented based on the musical score. Furthermore, this publication proposes a system of levels of automatization of both data retrieval and visualization for musical tension via slider-controllers and aspects of the analyst's invasive (control) opportunity over the automatization result either in real-time or retrospectively. This book chapter presents different visualizations of the main culmination area from Tüür's *Oxymoron* (2003): musical and physical sound parameters (from the score to MIDI to timbre groups to spectral analysis) and a comparison of the salient features and musical tension slider-controller empirical experiment (N=7) with average and standard deviation (data from publication I).

Concepts used and developed further: visualization/representation of musical parameters and data. Methods used and developed further: modeling.

Publication III offers insight into basic principles of scientific modeling in music, the sciences and education. Its aim is to explore what is modeling and how it can be applicable for this dissertation. It contains conceptual vertical-horizontal-type schemes: the multi-dimensional scheme of modeling and models incorporating models from music theory (analysis) and music psychology (perception and cognition) including the creative approach (musical composition/improvisation), the enlargement of Justi's and Gilbert's (2002) and Gilbert's (2010) "Model of Modeling." Aspects of cognition (how the human mind works via representation and modeling) are directly touched with modeling and representational issues. It introduces also my Twelve Strategic Steps for modeling/analysis of models (TSSM) in eight stages/phases (posed as questions) that were the basis for the COSM: Cognitive Octagonal Slice Model development (manual version presented in Lock & Kotta 2015, computerized empirical version see Lock 2019). For the first full application of the computerized COSM with empirical experiments see publication IV.

Concepts used and developed further: visualization/representation of data. Methods used and developed further: modeling.

In the above described two theoretical publications I have attempted to better comprehend principles (visualizing/representation) of music and data retrieved from empirical experiments as well as scientific methods (modeling) to gain more precise data from empirical experiments to get closer to an enhanced understanding of musical tension and the role of musical parameters in contemporary post-tonal music. The next step was to apply the principle of cognitive modeling while designing and conducting a scientifically more precise empirical experiment. The following publication IV introduces theoretical background on cognitive listening strategies and salience, perceived structural musical tension and musical parameters. It presents the empirical version of my new cognitive analysis model COSM. The applied cluster-finding DBSCAN (data-mining) algorithm was useful to achieve a more effective mass-data analysis and visualization/representation as well as to enhance the comparison of data from different types of analysis/data retrieval methods. Furthermore, narrativity as a cross-domain concept is introduced to more meaningfully interpret joined data from different methods.

Publication IV is the first full application of the computerized and empirical COSM. It explores cognitive listening strategies (attentive listening, salience, musical event as Impulse/"moment of change"), perceived structural musical tension (since Lock 2022 PSMT, see also 3.5 and publication V**)**, the role of musical parameters (their salience in contemporary post-tonal orchestral music (CPTOM) in comparison to tonal music) as well as narrativity. It introduces the theoretical background and functioning principles of my TSSM-based (publication III) COSM: Cognitive Octagonal Slice Model. It merges together music theoretical and music psychological approaches and the results are interpreted cross-disciplinarily with narrativity. Its results conducted with Tüür's 4th Symphony/Percussion Concerto *Magma* (2002) visualize combined data of empirical slider-controller listening experiments (TEDEA, musical tension N=26) with results of COSM (salience of part 1 Impulses and part 2 musical parameters N=14). Tüür's 4th Symphony is predestined as object of this main study of this dissertation and as model-development basis, because the music develops clearly wave-like (musical tension) and it mixes different styles and genres (post-tonal contemporary orchestral music and rock-music rhythms). COSM is holistic in its structure, definition and theoretical background (qualifying it as model, not as theory). The data analysis and visualization uses data-mining (DBSCAN algorithm, Ester *et al* 1996) as well as 2D time-intensity/value diagrams. The DBSCAN algorithm enables analytically joining together mass-data both from longer works and higher numbers of participants using a kind of flexible time window principle – earlier attempts used a static time window (I) and a moving time window (Ia). This publication merges different fields and methods interdisciplinarily with an own identity within "cognitive science" (Núñez *et al* 2019: 784) via developing a new model (COSM), its empirical testing uses general and cross-domain cognitive phenomena: audiovisual salience, musical event as Impulse, narrativity.

Visualization uses several simplified 2D diagrams (time window vs tension intensity/salience value), line-curve and matrix visualizations of the COSM musical parameter data, waveform-based visualization of the principles of the DBSCAN algorithm, the methodological principles and the Max/MSP interface of the COSM octagonal shape; annotated musical score examples, and musical rhetorical form-analysis schemes (Kotta 2011, based on Hepokoski & Darcy 2007). Concepts used and developed further: salience, musical events as Impulses; narrativity; visualization/representation of musical parameters and data. Methods used and developed further: Impulse detection (listening analysis only), slider-controllers; DBSCAN cluster finding algorithm (flexible time window); modeling.

Before publication V I presented in a Tallinn University Academic Language Conference paper (Lock 2022, in Estonian) the evolution of my definitions of musical tension 2010–2020 that enabled me to finally arrive at a cross-domain definition of 'Perceived structural musical tension' (PSMT)[124] revealing that the processes of perception, cognition and narrativity exhibit analogous principles with the process of experiencing musical tension via listening.

Publication V intertwines music analysis and music psychology with a cognitive approach to better comprehend contemporary post-tonal orchestral music (CPTOM). It presents five scientific Dilemmas, Research Questions (RQs), Statements (STs) and Answers as challenges for systematically analysing Cognitive Musical Tension (CMT). They collocate and conclude the theoretical problems that have been the intuitive trigger of this dissertation until its formulation in this paper. The aim was to formalize five research-methodological Dilemmas

---

[124] In Estonian: 'Tajutud strukturaalne muusikalise pinge' (TSMP). See the Estonian version of the definition also presented in a paper at the Estonian Academy of Music and Theater (EAMT) conference (Lock 2023a, in Estonian).

crucial for an empirical analysis of musical tension as a temporal dynamic wave-like phenomenon (TDWP). It enabled me to methodologically interconnect all methodological principles developed so far with basic principles of ‘experience’ and ‘intuition,’ “psychophysical measurement” and a Kantian inner/outer sensation (magnitude detection) (Kraus 2013/2016) in order to deepen the methodological-philosophical ground and to round up the analytical/empirical experimental methodology. On the one hand this paper identifies musical tension as cognitive, therefore it remains complicated to research. On the other hand it suggests theoretically/methodologically grounded “solutions” to the Dilemmas while encouraging a modeling approach as well as applying “psychophysical measurement” principles that are both closely related to comprehend an everyday real-life phenomena, the temporal dynamic wave-like phenomenon (TDWP) musical tension.

Visualization uses several simplified 2D diagrams: domains of music as Venn-diagram, generalized tension-intensity/salience-values time-windows; the methodological principles and the Max/MSP interface of the COSM octagonal shape in a combined figure. Concepts used and developed further: psychophysical measurement, Kantian external object for inner-outer magnitude comprehension. Methods used and developed further: audiovisual salience, psychophysical measurement principles

In CPTOM represented by the music of Erkki-Sven Tüür (b 1959) merging several contrasting styles and composition techniques and differing from clear classical/tonal structures (bordered motifs, themes, harmonic schemes) and/or form it is the analogy (i) of such music to a wave-like phenomenon (via modeling) that enables to focus on the detectable Impulses (Im) (either high points/maxima/culmination areas, or other salience triggering musical events; (ii) of wave-like phenomena and Cognitive Musical Tension (CMT) to narrativity that enables to comprehend/interpret CMT with cross-domain/modeling/visualization tools. Therefore it was necessary to abandon the detection of segmentation points in classical/tonal musical form (their role in Tüür’s music is taken over by the maxima and minima of the tensional development of the music as well as the durational salient features (culminations, see publication I). Salience is the key concept for solving the scientific Dilemmas and answering the research questions posed in publication IV and V. The following section 6 presents the conclusions and implications concerning general and specific results. It concludes on the actuality (see 1.3, 2.2) and the methodological novelties as the original contribution (see 3.7) this dissertation offers to the research of complex musical structures and musical tension.

6. Conclusions and implications

The basic goal of this dissertation – to intertwine music analytical (theoretical) and music psychological (empirical) thinking – has been fulfilled via developing and applying a cross-domain approach to understand musical tension as ‘Perceived structural musical tension’ (PSMT, the special object) of contemporary post-tonal orchestral music (CPTOM, the general object) of the Estonian composer Erkki-Sven Tüür (b. 1959) via the salience of Impulses (Im’s) and musical parameters (MP’s) (see main study in publication IV).

The general aim to advance analytical methods to enhance the understanding of complex, multilayered and dense CPTOM (the general object) while observing/analyzing musical tension (the special object) has been achieved via identifying musical tension as Cognitive Musical Tension (CMT) and formulating the five scientific Dilemmas with five Research Questions (RQs), five Statements (STs) and Answers to each (see publication V). The acknowledging of intuitive *attentive listening* based on the methodological-philosophical background of Kant’s (1998) *analytic idealism* (developed in the *Critique of Pure Reason*, 1781[1], 1787[2]) and aspects of psychophysical measurement (“just noticeable difference” (JND) as perceived “relative change” as “moments of change”/Impulses, and meaningfulness, see 2.4) specify, underpin and even extend speculative music theory/analysis approaches (see 2.1).

The main features of the general object, Erkki-Sven Tüür’s CPTOM (2002/2003) analyzed/observed in this dissertation are the increasing role of its converging textures (from “slow” to “quick” to “continuous” musical time organization, Kotta 2008, 2011) and Tüür’s specific form-building aspect called “tone” (“depth-timbre,” pre-planned by the composer via visualization) establishing the dramaturgy and narrative of his symphonic and orchestral music. Tüür’s music merges classical hierarchical (structural, teleological) archetypes with sound-centered multidimensional (spatial, non-hierarchical) thinking (Kotta 2009, 2009/2010) (see 1.1). Furthermore, Tüür’s music, like musical tension, too, is a temporal dynamic wave-like (real-world) phenomenon (TDWP) (see 2.3, publication V).

In order to enhance the understanding of the general object of this dissertation, the contemporary post-tonal orchestral music (CPTOM) by Tüür – *Oxymoron* (2003) for large ensemble and the 4th Symphony/Percussion Concerto *Magma* (2002) – I have proposed and conducted this dissertation in accordance with the following requirements to: 1) understand music as environment (publications III, IV), 2) apply modeling (publication III) including representation and visualization (publication II) including computational means; 3) consider intuitive musical thinking (a so-called first-person methodological approach) (see 2.1) based on Kant's *analytic idealism* (see 2.2). My framework of intuitive, systematic, system-based and automatized analysis methods (publication II) covers so-called speculative (music theoretical/analytical) methods and my preferred intuitive (*analytic idealist*, introspection and experience based) approach together with empirical, modeling and computational means as the core of this mainly methodological dissertation. The understanding of music as environment has been justified in addition to research-based methodological reasons also by the composer's own goals to embrace sensation and integrated cognition in a modern acoustic world (see 1.1).

This dissertation has brought out that its general and the special object of observation – music (CPTOM) and musical tension (CMT) – are both complicated to pin down analytically and scientifically. Empirical and quasi-empirical studies on musical tension of atonal, non-triadic, neoclassical, and nontonal (orchestral) music have been conducted analyzing/observing e.g. semantics, intensities and musical structure (musical parameters, dramaturgy, form) as well as timbre/roughness and harmonic tension (see 3.1). This dissertation underlines that musical tension research faces analytical and scientific difficulties and dilemmas (publication V).

The main objective of this dissertation was to *detect*, *describe*, *make sense of* and *comprehend* musical tension as a compound and whole phenomenon. This was done via examination of structural aspects observable in musical events that *purportedly trigger* the experience of musical tension during *attentive listening* to Tüür's CPTOM (works mentioned above). The core of the main objective of this dissertation – analyzing/observing 'Perceived structural musical tension' (PSMT, a cross-domain definition underpinned in 3.5) are structural aspects (musical parameters) detected/measured in Impulses/"moments of change" via salience during *attentive listening*. The *why*-question (why musical tension changes at Impulse moments, see 3.7) is complicated and can be answered only by *interpreting* correlations found from the empirical data (publication IV). Musical tension (PSMT) as a temporal dynamic wave-like (real-world) phenomenon (TDWP) is highly related to expectation, it calls for prediction models and segmentation in the musical score and perceptually (e.g. partly tension-including cognitive models/approaches see 2.5; directly musical-tension-related research see 3.1, 4.1). However, prediction and segmentation are challenging, therefore the research methodology of this dissertation has dealt mainly with an indirect kind of segmentation process: i) observing high- and low-points, increasing, decreasing and remaining of tension as a compound phenomenon related to musical form ("macroform") traceable with slider-controllers; ii) detection of Impulses/"moments of change" and meaningfulness as psychophysical measurement principles via salience of Impulses (Im-s) and musical parameters (MP-s) using my COSM: Cognitive Octagonal Slice Model. The *why*-question (why is tension in-/decreasing/remaining) is approached in the main study of this dissertation through measuring the role aka salience of musical parameters in Impulses (publication IV).

The detection of Impulses/"moments of change" as "relative change" as "just noticeable difference" as well as measuring inner quantities based on outer quantities qualifies both the slide-controller-based TEDEA: Tension Design Experimental Apparatus and especially COSM as a psychophysical measurement tool (see 2.4). COSM qualifies as an external, intersubjectively shared spatial standard-object (Kraus 2013/2016) to measure inner sensation (*precategorial objects**) based on outer sensation of Impulses (Im-s) and musical parameters (MP's). So-called musical *protonarratives* created by patterns of tension and relaxation (Imberty 2000) trigger "perceived intensity" that enables the measurement of 'Perceived structural musical tension' (PSMT), my fully fledged definition joining basic features of tension, cognition and narrativity (see 3.5).

The actuality of this dissertation lies in the methodological-philosophical background of introspection, intuition and immediate experience being supported by a recent renaissance of Kant's *analytic idealism* especially on the level of elementary sensation and apperception (see 2.2). Hence, a cognitive approach also includes representation and visualization (publication II) that are part of modeling as a core and actual tool of science, analysis, composition and education (publication III) – especially in the light of the recent growing and ubiquitous AI modeling approaches and tools. My own theoretical, conceptual and methodological identity "to become part of cognitive science" is achieved in this dissertation via developing the Twelve Strategic Steps for modeling/analysis of models (TSSM) (publication III) and the listening-based COSM: Cognitive Octagonal Slice Model (publication IV).

The general results of the COSM musical parameters' salience for the 4th Symphony/Percussion Concerto *Magma* (see Appendix D) show that the musical parameters dynamics (SP), rhythm (PP) and instrumentation/timbre (SP) as so-called All salience level parameters (AL) are more salient than other, so-called Basic salience level parameters (BL).

This result is in line with general observations/expectations about musical tension, musical parameters as well as Tüür's compositional tools and goals on the cognitive level (see 1.1). These fit to three basic microtemporal listening dimensions (Christensen 1996, 2012) intensity, timbre (and partly space). Dynamics as loudness (SP) and intensity is a basic perceptual parameter and a leading secondary parameter (SP), intensity as micro- and macrotemporal listening dimension via auditory awareness signifies changes especially in Tüür's CPTOM as TDWP. Rhythm (SP) is a particularly prominent/salient musical parameter in the 4th Symphony/Percussion Concerto *Magma* especially via the percussion solo part. Furthermore, through converging of "quick," "slow" and "continuous" time textures and the wave-like character as in- and decreasing phenomena are composed out in Tüür's music. Also, instrumentation/timbre (SP) is prominent/salient in Tüür's music particularly via the percussion instruments, especially in the solo part; timbre and roughness have prominent/salient "qualities" in relation to musical tension. Tüür's "tone" as "depth-timbre" parameter (visualizable as z-axis) is more than surface-timbre and leads to Christensen's (1996) listening dimension called "space." "Space" is salient also metaphorically/via conceptual blending (CBT) as up-down/room-related movement of musical tension (PSMT) as temporal dynamic wave-like (real-world) phenomenon (TDWP) in Tüür's music using in the title *Magma* of the 4th Symphony/Percussion Concerto a natural environment-analogy appropriately fitting also to tension as general phenomenon.

This dissertation has aimed to advance analytical methods to comprehend complex musical structures via musical tension (general aim) from angles so far considered less or not at all: music understood as environment (publication III, IV, see 1.2, 1.3), the role of salience in tension research (see 3.1 and publication I, more specifically 3.2 and 3.3), narrativity (see 3.5 and publication IV) and the five scientific Dilemmas (see 3.6 and publication V). The methodological-philosophical basis has been built on intuitive thinking (Clifton 1975), Kantian *analytic idealism* (Kant 1998) and psychophysical measurement and signal detection principles (Steingrimsson 2016; Kraus 2013/2016; Morales 2018, 2024) that, if at all, are applied in music theory/analysis often indirectly or implicitly – intuitively as speculative approach.

The methodological novelties this dissertation has suggested and applied, are the following: 1) Musical Tension is identified as Cognitive (CMT, publication V) via intuitive attentive listening, this requires an enlarged definition joining aspects of tension, cognition and narrativity (PSMT) (see 3.5, publication V); 2) musical tension as a compound temporal dynamic wave-like (real-world) phenomenon (TDWP) is complex to research (Dilemmas publication V), instead the detection of the salience of Impulses/“moments of change” and musical parameters as a *psychophysical measurement* procedure is applied (publication IV); 3) such a psychophysical measurement procedure has included the search for an external, “fully fledged,” intersubjectively shared spatial standard-object (Kraus 2013/2016) I have called COSM: Cognitive Octagonal Slice Model, against which the detection of inner (intuitive) and outer (“objective”) magnitudes can take place systematically.

As a final condensed conclusion I repeat that music and musical tension are cognitive, perception is not direct, it is mediated by the cognition, experience and implicit knowledge of the enculturated listener. Furthermore, listening to music can be seen as a detection, a kind of psychophysical measurement procedure that, of course, includes expectations and assessment decisions (e.g. liking or not liking, being captured by its tensional development or not), comparisons and conceptual blending – therefore it depends on (pre-)knowledge and cognition. However, in my opinion it does not matter whether meaningful invariants can be found either mentally “in the mind” of the listener (primed) or in the music understood as an environment (unprimed or primed). Hence, one should admit the following limiting aspects in scientific methodology: 1) physical and psychophysical measurement itself is never neither direct nor absolutely accurate or without errors, 2) modeling and representation deal with analogies and simplification, and 3) analysis relies on the search for generalized aspects (e.g. characteristic features, salient attributes, Impulses/“moments of change”) in objects of research based on relevant realm-related theories. While taking into account all three limiting aspects (including the five scientific Dilemmas proposed in this dissertation) I have combined analytical with empirical research methods. I have coped with these limitations through observing, detecting, analyzing and comprehending characteristic features, salient attributes and Impulses/“moments of change” of CPTOM arriving at a “fully fledged” definition of ‘Perceived structural musical tension’ (PSMT) via a cross-domain cognitive and Kantian *analytic-idealism* based approach, and developing/applying my COSM: Cognitive Octagonal Slice Model.

References

Abbott, Alison 2002. Music, Maestro, Please! – *Nature* 416: 12–14. https://doi.org/10.1038/416012a

Acotto, Edoardo & Andreatta, Moreno 2012. Between mind and mathematics. Different kinds of computational representations of music. – *Mathématiques et Sciences Humaines Mathematics and Social Sciences* 199: 7–25. https://doi.org/10.4000/msh.12254

Addessi, Anna Rita 2010. Auditive analysis of the *Quartetto per Archi in due tempi* (1955) by Bruno Maderna. – *Musicae Scientiae* 14 (2): 225–250. https://doi.org/10.1177/10298649100140S213

Addessi, Anna Rita & Caterina, Roberto 2000. Perceptual Musical Analysis: Segmentation and Perception of Tension. – *Musicae Scientiae* 4 (1): 31–54. https://doi.org/10.1177/102986490000400102

Addessi, Anna Rita & Caterina, Roberto 2005. Analysis and perception in post-tonal music: an example from Kurtág's String Quartet Op.1. – *Psychology of Music* 33 (1): 94–116. https://doi.org/10.1177/0305735605048015

Albrecht, Joshua 2012. *Affective Analysis of Music Using the Progressive Exposure Method: The Influence of Bottom-Up Features on Perceived Musical Affect* (Doctoral Dissertation). Columbus, OH: The Ohio State University. https://library.ohio-state.edu/record=b7147732~S7

Almén, Byron 2003. Narrative Archetypes: A Critique, Theory, and Method of Narrative Analysis. – *Journal of Music Theory* 47 (1): 1–39. https://doi.org/10.1215/00222909-47-1-1

Almén, Byron 2006. Modes of Analysis. – *Theory and Practice (Journal of the Music Theory Society of New York State)* 31: 1–38. https://www.jstor.org/stable/41054372

Anderson, Britt 2021/2023. Stop paying attention to "attention". – *WIREs Cognitive Science* 14 (1) (Special Issue: What is Attention?, e1574): 1–6. https://doi.org/10.1002/wcs.1574

Antović, Mihailo 2018. Schemas, grounds, meaning: On the emergence of musical concepts through conceptual blending. – *Musicae Scientiae* 22 (1): 57–71. http://dx.doi.org/10.1177/1029864917711218

Antović, Mihailo 2022. *Multilevel Grounding: A Theory of Musical Meaning* (SEMPRE Studies in The Psychology of Music). Abingdon, New York: Routledge, Taylor & Francis Group. DOI: 10.4324/9781003030768 https://www.taylorfrancis.com/books/mono/10.4324/9781003030768/multilevel-grounding-mihailo-antovi%C4%87, https://www.researchgate.net/publication/361265523

Antunes, Micael; Espirito Santo, Guilherme Feulo do; Manzolli, Jônatas & Queiroz, Marcelo 2021/2023. A psychoacoustic-based methodology for sound mass music analysis. – *Music in* *the AI Era. 15th International Symposium, CMMR 2021.* Virtual Event, November 15–19, 2021, Revised Selected Papers. Dordrecht: Springer, 267–281. https://link.springer.com/chapter/10.1007/978-3-031-35382-6_21

Arias Valencia María Mercedes 2022/2023. Principles, Scope, and Limitations of the Methodological Triangulation. – *Investigación y Educación en Enfermería – Research and Education in Nursing* 40 (2): 1–14, e03. https://doi.org/10.17533/udea.iee.v40n2e03, https://revistas.udea.edu.co/index.php/iee/article/view/350172/20812334

Aru, Jaan & Bachmann, Talis 2017. In and Out of Consciousness: How Does Conscious Processing (D)evolve Over Time? – *Frontiers in Psychology* 8: 1–4 (128). https://doi.org/10.3389/fpsyg.2017.00128

Athanasopoulos, George & Antović, Mihailo 2018. Conceptual integration of sound and image: A model of perceptual modalities. – *Musicae Scientiae* 22 (1): 72–87. https://doi.org/10.1177/1029864917713244

Babbitt, Milton 1965. The use of computers in musicological research. – *Perspectives of New Music* 3 (2): 74–83. https://doi.org/10.1515/9781400841226.202

Babbitt, Milton 2003. The Use of Computers in Musicological Research 1965. – *The Collected Essays of Milton Babbitt*. Eds. S. Peles, S. Dembski, A. Mead, J. N. Straus. Princeton: Princeton University Press, 202–212. https://doi.org/10.1515/9781400841226.202

Bachmann, Talis 2000. *Microgenetic Approach to the Conscious Mind.* Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/aicr.25

Baroni, Mario 2003. The Macroform in Post-tonal Music: Listening and Analysis. – *Musicæ Scientiæ* 7 (2): 219–40. https://doi.org/10.1177/102986490300700202

Bauer, Amy 2001. "Composing the Sound Itself": Secondary Parameters and Structure in the Music of Ligeti. – *Indiana Theory Review* 22 (1): 37–64. https://journals.scholarsportal.info/details/02718022/v22i0001/37_tsispasitmol.xml, https://www.jstor.org/stable/24054247

Bechtel, William 1998. Representations and cognitive explanations: Assessing the dynamicist challenge in cognitive science. – *Cognitive Science* 22, 295–318. http://mechanism.ucsd.edu/research/REPRESENT.html [28.09.2024].

Belkin, Alan 2018. *Musical Composition. Craft and Art.* New Haven, London: Yale University Press. https://yalebooks.yale.edu/book/9780300218992/musical-composition/

Belkin, Alan [2025]. On salience – what do we hear? Pedagogical Writing. Teacher. Alan Belkin Music, https://alanbelkinmusic.com/on-salience/ [17.03.2025].

Bentley, John P. 2005. *Principles of Measurement Systems* (4th Edition). Harlow, England: Pearson Education, Prentice Hall. https://uvceee.wordpress.com/wp-content/uploads/2016/09/principles-of-measurement-systems-20051.pdf

Bigand, Emmanuel; Tillmann, Barbara; Poulin-Charronnat, Bénédicte & Manderlier, Deborah (2005). Repetition priming: Is music special? – *The Quarterly Journal of Experimental Psychology Section A* 58 (8): 1347–1375. https://doi.org/10.1080/02724980443000601

Blasi, Damián E.; Henrich, Joseph; Adamou, Evangelia; Kemmerer, David & Majid, Asifa 2022. Over-reliance on English hinders cognitive science. – *Trends in Cognitive Sciences* 26 (12): 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015

Bor, Mustafa 2009. *Contour Reduction Algorithms: A Theory of Pitch and Duration Hierarchies for Post-Tonal Music* (Doctoral Dissertation). Vancouver: University of British Columbia. https://dx.doi.org/10.14288/1.0067104

Bregman, Albert S. 1990. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/1486.001.0001

Brodie, T.; Gilbert, J.; Hollins, M.; Roper, G.; Robson, K.; Webb, M., *et al* 1994. *Models and Modelling in Science Education*. Hatfield Herts: Association for Science Education.

Broman, Per F. & Engebretsen, Nora A. 2007. *What Kind of Theory is Music Theory? Epistemological Exercises in Music Theory and Analysis* (Stockholm Studies in Musicology 1). Stockholm: Stockholm University. https://www.diva-portal.org/smash/get/diva2:328167/FULLTEXT03.pdf

Brower, Candace 2000. A Cognitive Theory of Musical Meaning. – *Journal of Music Theory,* 44 (2): 323–379. https://doi.org/10.2307/3090681

Browman, Catherine P. & Goldstein, Louis 1990. Representation and reality: physical systems and phonological structure. – *Journal of Phonetics* 18: 411–424. https://doi.org/10.1016/S0095-4470(19)30382-1

Burro, Roberto 2016. To be objective in Experimental Phenomenology: a Psychophysics application. – *SpringerPlus* 5: 1–15 (Article 1720). https://doi.org/10.1186/s40064-016-3418-4

Burrows, David 1997. A dynamical systems perspective on music. – *The Journal of Musicology* 15 (4): 529–545. https://doi.org/10.2307/764006

Cambouropoulos, Emilios & Tsougras, Costas 2009. Auditory Streams in Ligeti's *Continuum*: A Theoretical and Perceptual Approach. – *Journal of Interdisciplinary Music Studies* 3 (1&2): 119–137 (art. #0931207). https://musicstudies.org/wp-content/uploads/2017/01/Cambouropoulos_JIMS_0932107.pdf

Canales, Jimena 2010. *A Tenth of a Second: A History*. Chicago: University of Chicago Press. https://press.uchicago.edu/Misc/Chicago/093185.html

Carsetti, Arturo 2004a. *Seeing, Thinking and Knowing. Meaning and Self-Organisation in Visual Cognition and Thought*. Dordrecht, New York: Kluwer Academic Publishers, Springer. https://doi.org/10.1007/1-4020-2081-3

Carsetti, Arturo 2004b. The embodied meaning: self-organisation and symbolic dynamics in visual cognition. – *Seeing, Thinking and Knowing. Meaning and Self-Organisation in Visual*

Cognition and Thought. Ed. A. Carsetti. Dordrecht, New York: Kluwer Academic Publishers, Springer, 307–330. https://doi.org/10.1007/1-4020-2081-3_15

Cartwright, Nancy; Hardie, Jeremy; Montuschi, Eleonora; Soleiman, Matthew & Thresher, Ann C. 2022. *The Tangle of Science. Reliability Beyond Method, Rigour, and Objectivity*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198866343.001.0001

Casati, Roberto; Dokic, Jerome & Di Bona, Elvira 2020. Sounds. – *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition). Eds. E. N. Zalta, https://plato.stanford.edu/archives/win2020/entries/sounds/ [28.10.2024].

Chai, Wei 2005. *Automated Analysis of Musical Structure* (doctoral dissertation). Cambridge, MA: MIT Press. https://dspace.mit.edu/handle/1721.1/33878

CGPM R1 2018. Resolution 1 of the 26th CGPM. On the revision of the International System of Units (SI). The General Conference on Weights and Measures (CGPM), 26th meeting, Bureau International des Poids et Mesures (BIPM), https://www.bipm.org/en/committees/cg/cgpm/26-2018/resolution-1 [28.10.2024]. DOI: 10.59161/CGPM2018RES1E

Chai, Wei 2005. *Automated Analysis of Musical Structure* (Doctoral Dissertation). Cambridge, MA: MIT Press. https://dspace.mit.edu/handle/1721.1/33878

Chen, Chih-Ming & Sun, Ying-Chun 2012. Assessing the effects of different multimedia materials on emotions and learning performance for visual and verbal style learners. – *Computers & Education* 59: 1273–1285. https://doi.org/10.1016/j.compedu.2012.05.006

Chouvel, Jean-Marc 2014. Categories and Representation in Musical Analysis. – *Sonus* 35 (1): 17–35. http://jeanmarc.chouvel.3.free.fr/textes/English/CategoriesAndRepresentation.pdf

Christensen, Erik 1996. *The Musical Timespace. A Theory of Music Listening*. Aalborg: Aalborg University Press. https://www.researchgate.net/publication/209436138_The_Musical_Timespace_A_Theory_of_Music_Listening

Christensen, Erik 2012. *Music Listening, Music Therapy, Phenomenology and Neuroscience* (Doctoral Dissertation). Aalborg: Aalborg University Press. https://vbn.aau.dk/en/publications/music-listening-music-therapy-phenomenology-and-neuroscience

Clarke Eric F. 1989. Mind the gap: formal structures and psychological processes in music. – *Contemporary Music Review* 3 (1) (Music, Mind and Structure): 1–13. https://doi.org/10.1080/07494468900640021

Clarke, Eric F. 2005. *Ways of Listening. An Ecological Approach to the Perception of Musical Meaning*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195151947.001.0001

Clifton, Thomas 1975. Some Comparisons between Intuitive and Scientific Descriptions of Music. – *Journal of Music Theory* 19 (1): 66–110. https://doi.org/10.2307/843751, https://www.jstor.org/stable/843751

Coath, Martin; Denham, Susan L.; Smith, Leigh M.; Honing, Henkjan; Hazan, Amaury; Holonowicz, Piotr & Purwins, Hendrik 2009. Model cortical responses for the detection of perceptual onsets and beat tracking in singing. – *Connection Science* 21 (2-3): 193–205. https://doi.org/10.1080/09540090902733905

Coath, Martin; Sadique, Sheik; Elisabetta, Chicca; Giacomo, Indiveri; Susan, Denham & Wennekers, Thomas 2013. A robust sound perception model suitable for neuromorphic implementation. – *Frontiers in Neuroscience* 7: 1–10 (Article 278). https://doi.org/10.3389/fnins.2013.00278

Cognition O-ED [2025]. Cognition. Etymonline. Online Etymology Dictionary. Douglas Harper, https://www.etymonline.com/word/cognition [15.04.2025].

Cook, Nicholas 1994. Perception: A perspective from music theory. – *Musical Perceptions*. Eds. R. Aiello, & J. Sloboda. Oxford: Oxford University Press, 64–95. https://philpapers.org/rec/COOPAP-7

Cook, Nicolas & Clarke, Eric F. 2004. *Empirical Musicology: Methods, Aims, Prospects*. Oxford: Oxford University Press. https://global.oup.com/academic/product/music-imagination-and-culture-9780198163039

Cross, Ian 1998a. Music analysis and music perception. – *Music Analysis* 17 (1): 3–20. https://doi.org/10.2307/854368

Cross, Ian 1998b. Music and Science: Three Views. – *Revue Belge De Musicologie / Belgisch Tijdschrift Voor Muziekwetenschap* 52: 207–214. https://doi.org/10.2307/3686926

Dahan, Kevin 2012. Of music, mind, and machines: Towards an integrative musicology. – *Music, Mind, and Invention Workshop* (March 31, 2012). Ewing Township: The College of New Jersey, 1–4. https://owd.tcnj.edu/~mmi/papers/Paper64.pdf

Dahlstedt, Sten 2007. The Contribution of the Mind. – *What Kind of Theory is Music Theory? Epistemological Exercises in Music Theory and Analysis* (Stockholm Studies in Musicology 1). Eds. P. F. Broman & N. A. Engebretsen. Stockholm: Stockholm University, 253–264. https://www.diva-portal.org/smash/get/diva2:328167/FULLTEXT03.pdf

Davis, Richard 2019. An Introduction to the Revised International System of Units (SI). – *IEEE Instrumentation & Measurement Magazine* 22 (3): 4–8. https://doi.org/10.1109/MIM.2019.8716268

Davis, Richard 2023. The Revised SI Compared to Natural Systems of Units. – *IEEE Instrumentation & Measurement Magazine* 26 (1): 5–11. https://doi.org/10.1109/MIM.2023.10034958

DeBellis, Mark 2002. Musical Analysis as Articulation. – *The Journal of Aesthetics and Art Criticism* 60 (2): 119–135. https://doi.org/10.1111/1540-6245.00060

Deliège, Irène 1987. Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules. – *Music Perception* 4 (4): 325–359. https://doi.org/10.2307/40285378

Deliège, Irène 1989. A perceptual approach to contemporary musical forms. – *Contemporary Music Review: Music and the Cognitive Science* 4 (1): 213–230. https://doi.org/10.1080/07494468900640301

Deliège, Irène 1996. Cue Abstraction as a component of categorisation processes in music listening. – *Psychology of Music* 24 (2): 131–156. https://doi.org/10.1177/0305735696242007

Deliège, Irène 2001. Similarity Perception ↔ Categorization ↔ Cue Abstraction. – *Music Perception* 18 (3): 233–243. https://doi.org/10.1525/mp.2001.18.3.233

Deliège, Irène 2007. Similarity relations in listening to music: How do they come into play? *Musicae Scientiae*, Discussion Forum 4A: 9–37. https://doi.org/10.1177/102986490901300104

Deliège, Irène; Mélen, Marc 1997. Cue abstraction in the representation of musical form. – *Perception and Cognition of Music*. Eds. I. Deliège & J. Sloboda. New York: Psychology Press, 387–412. https://doi.org/10.4324/9780203344262

Deliège, Irène; Mélen, Marc; Stammers, Diana & Cross, Ian 1996. Musical schemata in real-time listening to a piece of music. – *Music Perception* 14 (2): 117–159. https://doi.org/10.2307/40285715

Denzin, Norman K. 2012. Triangulation 2.0. – *Journal of Mixed Methods Research* 6(2): 80–88. https://doi.org/10.1177/1558689812437186

de Paiva Santana, Charles & Guigue, Didier 2020. A Computer Model for the Analysis of Diversity and Coordination in Orchestration. – *Proceedings of EVA 2020 Conference (AI and the Arts: Artificial Imagination)* (July 6th – 9th, 2020). London: EVA (Electronic Visualisation and the Arts), BCS Learning & Development Ltd, 120–127. http://dx.doi.org/10.14236/ewic/EVA2020.21

de Paiva Santana, Charles & Guigue, Didier 2022. The role of orchestration in shaping musical form: Theory and practice of a methodological proposal and its computational implementation. – *Musicological Annual* 58 (2): 121–153. https://doi.org/10.4312/mz.58.2.121-153

Deutsch, Diana 1984. Psychology and Music. – *Psychology and its Allied Disciplines*. Ed. M. H. Bornstein. Hillsdale: Erlbaum, 155–194. https://deutsch.ucsd.edu/pdf/Ch-PsychMus.pdf

Di Cesare, Giuseppe; Di Dio, Cinzia; Rochat, Magali J.; Sinigaglia, Corrado; Bruschweiler-Stern, Nadia; Stern, Daniel N. & Rizzolatti, Giacomo 2014. The neural correlates of 'vitality form' recognition: an fMRI study. – *Social Cognitive and Affective Neuroscience* 9 (7): 951–960. https://doi.org/10.1093/scan/nst068

Di Stefano, Nicola; Vuust, Peter & Brattico, Elvira 2022. Consonance and dissonance perception. A critical review of the historical sources, multidisciplinary findings, and main hypotheses. – *Physics of Life Reviews* 43: 273–304. https://doi.org/10.1016/j.plrev.2022.10.004

Dubnov, Shlomo & McAdams, Stephen 2006. Structural and affective aspects of music from statistical audio signal analysis. – *Journal of the American Society for Information Science* *and Technology* 57 (11) (Special Topic Section on Computational Analysis of Style): 1526–1536. https://doi.org/10.1002/asi.20429

Duit, Reinders 1991. The role of analogies and metaphors in learning science. – *Science Education* 75 (6): 649–672. https://doi.org/10.1002/sce.3730750606

Eitan, Zohar & Granot, Roni Y. 2007. Intensity changes and perceived similarity: inter-parametric analogies. – *Musicae Scientiae* 11(1_suppl, Discussion Forum 4A): 39–75. https://doi.org/10.1177/102986490701100103l

Eitan, Zohar & Granot, Roni Y. 2009. Primary versus secondary musical parameters and the classification of melodic motives. – *Musicae Scientiae* 13 (1_suppl): 139–179. https://doi.org/10.1177/102986490901300107 Farbood 2006, 2008, 2012

Erlich, Paul 2006. A middle path. Between Just Intonation and the Equal Temperaments, Part 1. – *Xenharmonikôn* 18: 159–199. https://sethares.engr.wisc.edu/paperspdf/Erlich-MiddlePath.pdf

Ester, Martin; Kriegel, Hans-Peter; Sander, Jörg & Xu, Xiaowei 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. – *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining* (KDD-96). Palo Alto: AAAI Press, 226–231. https://aaai.org/papers/kdd96-037-a-density-based-algorithm-for-discovering-clusters-in-large-spatial-databases-with-noise/

Farbood, Morwaread M. 2006. *A Quantitative, Parametric Model of Musical Tension* (Doctoral Dissertation). Cambridge, MA: MIT Press. https://dspace.mit.edu/handle/1721.1/34182

Farbood, Morwaread M. 2008. A global model of musical tension. – *Proceedings of the 10th International Conference of Music Perception and Cognition (ICMPC10)*. Sapporo, Japan: International Conference on Music Perception and Cognition, 690–695. https://bpb-us-e1.wpmucdn.com/wp.nyu.edu/dist/f/11865/files/2020/08/Farbood-ICMPC2008.pdf

Farbood, Morwaread M. 2012. A parametric, temporal model of musical tension. – *Music Perception* 29 (4): 387–428. https://doi.org/10.1525/mp.2012.29.4.387

Feremans, Len; Vercruyssen, Vincent; Meert, Wannes; Cule, Boris & Goethals, Bart 2019. A framework for pattern mining and anomaly detection in multi-dimensional time series and event logs. – *New Frontiers in Mining Complex Patterns*. 8th International Workshop, NFMCP 2019, held in conjunction with ECML-PKDD 2019 Würzburg, Germany, September 16, 2019 Revised Selected Papers. Cham: Springer Nature Switzerland, 3–20. https://doi.org/10.1007/978-3-030-48861-1_1, https://lirias.kuleuven.be/retrieve/550412 (Open Access, accepted version)

Findlay, John Niemeyer 1981/2011. The Analytic of Principles. – *Kant and the Transcendental Object: A Hermeneutic Study*. Online Edition. Eds. J. N. Findlay. Oxford: Oxford Academic, 158–194. https://doi.org/10.1093/acprof:oso/9780198246381.003.0005

Forth, Jamie C. & Wiggins, Geraint A. 2009. An approach for identifying salient repetition in multidimensional representations of polyphonic music. – *London Algorithmics 2008: Theory* *and Practice*. Eds. J. Chan, J. W. Daykin & M. S. Rahman. London: College Publications, 44–58. https://research.gold.ac.uk/id/eprint/8578/1/forth-wiggins-2009.pdf

Freytag, Gustav 1900/1863. *Freytag's Technique of the Drama: An Exposition of Dramatic Composition and Art* (translated and edited by Elias J. MacEwan). Chicago: Scott, Foresman and Company. https://archive.org/details/freytagstechniqu00freyuoft

Friberg, Anders; Bisesi Erica; Addessi Anna Rita & Baroni Mario 2019. Probing the underlying principles of perceived immanent accents using a modeling approach. – *Frontiers in Psychology* 10: 1–21 (Article 1024). https://doi.org/10.3389/fpsyg.2019.01024

Frigg, Roman & Hartmann, Stephan 2024. Models in Science. – *Stanford Encyclopedia of Philosophy* (Fall 2024 Edition). Eds. E. N. Zalta & U. Nodelman, https://plato.stanford.edu/archives/fall2024/entries/models-science [28.10.2024].

Fu, Tak-chung 2011. A Review on Time Series Data Mining. – *Engineering Applications of Artificial Intelligence* 24: 164–181. https://doi.org/10.1016/j.engappai.2010.09.007

Gallagher, Shaun & Zahavi, Dan 2008/2012/2020. *The Phenomenological Mind*. 3rd Edition. London, New York: Routledge, Taylor and Francis. https://doi.org/10.4324/9780429319792

Gjerdingen, Robert O. 1999. An Experimental Music Theory? – *Rethinking Music*. Eds. N. Cook & M. Everist. Oxford: Oxford University Press, 161–170. https://global.oup.com/academic/product/rethinking-music-9780198790044

Glăveanu, Vlad-Petre 2010. Principles for a Cultural Psychology of Creativity. – *Culture & Psychology* 16 (2): 147–163. https://doi.org/10.1177/1354067X10361394

Goodman, Nelson 1968. *Languages of Art. An Approach to a Theory of Symbols*. Indianapolis: Bobbs-Merrill. https://philpapers.org/rec/GOOLOA-2

Granot, Roni Y. & Eitan, Zohar 2011. Musical tension and the interaction of dynamic auditory parameters. – *Music Perception* 28 (3): 219–245. https://doi.org/10.1525/mp.2011.28.3.219

Gruhn, Wilfried 1992. Wahrnehmen und Verstehen. Kognitive Grundlagen der Repräsentation musikalischer Elemente und Strukturen. Entwurf zu einem Forschungsprogramm. – *Musikalische Erfahrung. Wahrnehmen, Erkennen, Aneignen* (Musikpädagogische Forschung, Band 13). Hrsg. H. J. Kaiser. Essen: Die Blaue Eule, 44–51. https://doi.org/10.25656/01:25029, https://www.pedocs.de/volltexte/2022/25029/pdf/AMPF_1992_Band_13_Gruhn_Wahrnehmen_und_Verstehen.pdf

Guigue, Didier & de Paiva Santana, Charles 2018. The structural function of musical texture: Towards a computer-assisted analysis of orchestration. – *Journées d'Informatique Musicale (JIM 2018).* May 2018. Amiens, France: L'Association Francophone d'Informatique Musicale (AFIM), 97–103. https://hal.archives-ouvertes.fr/hal-01791398

Hanslick, Eduard 1854. *Vom Musikalisch-Schönen. Ein Beitrag zur Revision der Ästhetic der Tonkunst*. Leipzig: R. Weigel. https://www.koelnklavier.de/quellen/hanslick/_index.html

Harris, Richard J. 2003/2008. Traditional nomothetic approaches. – *Handbook of Research Methods in Experimental Psychology*. Ed. S. F. Davis. Malden, MA; Oxford: Blackwell Publishing, 41–65. https://www.wiley.com/en-us/Handbook+of+Research+Methods+in+Experimental+Psychology-p-9780470756720

Hatten, Robert S. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington: Indiana University Press. https://archive.org/details/interpretingmusi00robe

Heesen, Remco; Bright, Liam Kofi & Zucker, Andrew 2019. Vindicating methodological triangulation. – *Synthese – An International Journal for Epistemology, Methodology and Philosophy of Science* 196, 3067–3081. https://doi.org/10.1007/s11229-016-1294-7

Hendren, Jeremiah; Grasso, Matteo; Juel, Bjørn Erik & Tononi, Giulio 2024. Integrated Information Theory Wiki (https://www.iit.wiki). Center for Sleep & Consciousness, University of Wisconsin-Madison. https://doi.org/10.5281/zenodo.14160284

Heilmann, Conrad 2015. A new interpretation of the Representational Theory of Measurement. – *Philosophy of Science* 82 (5): 787–797. https://doi.org/10.1086/683280

Heller, Jürgen 2021. Internal references in cross-modal judgments: a global psychophysical perspective. – *Psychological Review* 128 (3): 509–524. https://doi.org/10.1037/rev0000280

Hendry, Robin F. & Psillos, Stathis 2007. How to do things with theories: an interactive view of language and models in science. – *The Courage of Doing Philosophy: Essays Dedicated to Leszek Nowak*. Eds. J. Brzeziński, A. Klawiter, T.A.F. Kuipers, K. Łastowski, K. Paprzycka & P. Przybysz. Amsterdam, New York, NY: Rodopi, 129–164. http://users.uoa.gr/~psillos/PapersI/40-Models%20Psillos-Hendry.pdf

Hepokoski, James & Darcy, Warren 2006. *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth-Century Sonata*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195146400.001.0001

Herremans, Dorien & Chew, Elaine 2016. Tension ribbons: Quantifying and visualising tonal tension. – *Proceedings of the Second International Conference on Technologies for Music Notation and Representation (TENOR)*. Cambridge, UK, 27–29 May 2016. Ed. R. Hoadley. Cambridge: Anglia Ruskin University, 1–10. https://www.tenor-conference.org/proceedings/2016/02_Herremans_tenor2016.pdf, https://hdl.handle.net/10067/1363270151162165141

Herremans, Doriens & Chuan, Ching-Hua 2017. A multi-modal platform for semantic music analysis: visualizing audio- and score-based tension. – *Proceedings of 11th International Conference on Semantic Computing (IEEE ICSC).* San Diego, USA, 30.01–01.02.2017. Piscataway, NJ: IEEE Press, 419–426. https://www.proceedings.com/content/033/033863webtoc.pdf

Hipólito, Inês 2022. Cognition without neural representation: Dynamics of a complex system. – *Frontiers in Psychology* 12: 1–15 (Article 643276). https://doi.org/10.3389/fpsyg.2021.643276

Hjortkjær Jens & Nielbo L. Frederik 2010. A perceptual study on dynamical form in music. – *Proceedings of the 7th Sound and Music Computing Conference* (SMC, 18.–23.07.2010, Barcelona, Spain). https://www.academia.edu/320965

Hohwy, Jakob 2012. Attention and conscious perception in the hypothesis testing brain. – *Frontiers in Psychology* 3: 1–14 (Article 96). https://doi.org/10.3389/fpsyg.2012.00096

Hommel, Bernhard; Chapman, Craig S.; Cisek, Paul; Neyedli, Heather F.; Song, Joo-Hyun & Welsh, Timothy N. 2019. No one knows what attention is. – *Attention, Perception, & Psychophysics* 81: 2288–2303. https://doi.org/10.3758/s13414-019-01846-w

Honing, Henkjan 1993. Issues on the representation of time and structure in music. – *Contemporary Music Review* 9 (1-2): 221–238. https://doi.org/10.1080/07494469300640461

Honing, Henkjan 2006. On the growing role of observation, formalization and experimental method in musicology. – *Empirical Musicology Review* 1 (1): 2–6. https://doi.org/10.18061/1811/21901

Hoyningen-Huene, Paul 2013. *Systematicity: The Nature of Science*. Oxford: Oxford University Press. https://academic.oup.com/book/7464

Huang, Chih-Fang 2020. An innovative method of algorithmic composition using musical tension. – *Multimedia Tools and Applications* 79: 32119–32136. https://doi.org/10.1007/s11042-020-09506-0

Humphry, Stephen M. 2017. Psychological measurement: Theory, paradoxes, and prototypes. – *Theory & Psychology* 27 (3): 407–418. https://doi.org/10.1177/0959354317699099

Huron, David 2006. *Sweet Anticipation. Music and the Psychology of Expectation*. Cambridge, MA: A Bradford Book, The MIT Press. https://doi.org/10.7551/mitpress/6575.001.0001

Hutcheson, Jere T. 1972. *Musical Form and Analysis. A Programmed Course* (Vol. 1). Boston: Allyn and Bacon. https://archive.org/details/musicalformanaly0000hutc

IIT Wiki [2024]. Integrated Information Theory: Consciousness and Intrinsic Ontology https://www.iit.wiki/ [17.03.2025].

Imberty, Michael 2000. The question of innate competencies in musical communication. – *The Origins of Music.* Eds. N. L. Wallin, B. Merker & S. Brown. Cambridge, MA: MIT Press, 449–462. https://awspntest.apa.org/record/2000-07112-017, https://mitpress.mit.edu/9780262232067/the-origins-of-music/

Intuition O-ED [2025]. Intuition. Online Etymology Dictionary (O-ED), https://www.etymonline.com/word/intuition [17.03.2025].

Jahanian, Ali; Vishwanathan, S.V.N. & Allebach, Jan P. 2015. Learning visual balance from large-scale datasets of aesthetically highly rated images. – *Proceedings Volume 9394, SPIE/IS&T, Human Vision and Electronic Imaging XX, 93940Y*. 17.03.2015. San Francisco, CA: SPIE/IS&T Electronic Imaging. https://doi.org/10.1117/12.2084548

James, William 1890. *The Principles of Psychology, Vol. 1*. New York: Henry Holt & Company. https://doi.org/10.1037/10538-000

Jebb, Andrew T.; Tay, Louis; Wang, Wei & Huang, Qiming 2015. Time series analysis for psychological research: examining and forecasting change. – *Frontiers in Psychology* 6: 1–24 (Article 727). https://doi.org/10.3389/fpsyg.2015.00727

Jensen, Arthur R. 1979. "g": Outmoded theory of unconquered frontier. – *Creative Science & Technology* 2 (3): 16–29. https://arthurjensen.net/

Jensenius, Alexander Refsum 2002. *How Do We Recognize a Song in One Second. The Importance of Salience and Sound in Music Perception* (Master of Arts). Oslo: Music Technology Department of Music and Theatre University of Oslo. http://urn.nb.no/URN:NBN:no-7119

Jewell, George R. 2005. *Psychophysical Scaling of Prothetic and Metathetic Continua in Lateralized Brain Injury and Neglect* (Doctoral Dissertation). Birmingham, AL: The University of Alabama at Birmingham. https://www.proquest.com/openview/bb4f56d69477d0a7c523d87fe214b54b/1

Justi, Rosária S. & Gilbert, John K. 2002. Modelling: Teachers' view on the nature of modelling, and implications for the education of modelers. – *International Journal of Science Education* 24 (4): 369–387. https://doi.org/10.1080/09500690110110142

Kaipainen, Mauri 1994. *Dynamics of Musical Knowledge Ecology. Knowing-What and Knowing-How in the World of Sounds* (Doctoral Dissertation). Helsinki: University of Helsinki. https://mkaipain.wixsite.com/mkaipain/publications

Kaipainen, Mauri & Tikka, Pia 2022. Artificial Intelligence and Art. What's Left After the Hype? – *The Meaning of Creativity in the Age of AI.* Eds. R. Kelomees, V. Guljajeva, O. Laas. Tallinn: Estonian Academy of Arts, 123–149. https://www.researchgate.net/publication/375696443

Kaipainen, Mauri & Tikka, Pia 2025 (forthcoming). Towards triangulating epistemology of narrative experience. – *Gramma: Journal of Theory and Criticism*, 30 ((Re-)mediating the Unconscious) [forthcoming]. https://ejournals.lib.auth.gr/gramma/announcement/view/98

Kant, Immanuel 1998. *Critique of Pure Reason* (First edition 1781, second edition 1778, English translation by Paul Guyer and Allen Wood). Cambridge: Cambridge University Press. https://archive.org/details/critiqueofpurereason_201907

Kastrup, Bernardo 2019. *Analytic Idealism: A Consciousness-Only Ontology* (Doctoral Dissertation). Nijmegen: Radboud Universiteit Nijmegen. https://hdl.handle.net/2066/203090

Kauttonen, Janne; Kaipainen, Mauri & Tikka, Pia 2014. Model of Narrative Nowness for Neurocinematic Experiments. – *Workshop on Computational Models of Narrative. Open Access Series in Informatics (OASIcs) Volume 41*. Schloss Dagstuhl: Leibniz-Zentrum für Informatik, 77–87. https://doi.org/10.4230/OASIcs.CMN.2014.77

Kersten, Luke 2014/2015. Music and Cognitive Extension. – *Empirical Musicology Review* 9 (3-4): 193–202. https://doi.org/10.18061/emr.v9i3-4.4315

Kingdom, Frederick A. A. & Prins, Nicolaas 2016. *Psychophysics: A Practical Introduction*. 2nd Edition. Amsterdam: Academic Press, Elsevier Science & Technology. https://shop.elsevier.com/books/psychophysics/kingdom/978-0-12-407156-8

Kirschbaum, Martin 2001. *Höhepunktbildung und Dramaturgie in Neuer Musik. Die Erstellung "formaler Spannungsverläufe" als Ergänzung der musikalischen Analyse*. Köln: Dohr. https://isbn.de/person/Martin+Kirschbaum?verlag=dohr+k%C3%B6ln

Klein, Michael L. & Reyland, Nicholas 2012/2013. *Music and Narrative since 1900*. Bloomington, IN: Indiana University Press. https://iupress.org/9780253006448/music-and-narrative-since-1900/

Koelsch, Stefan 2013a. *Brain and Music*. Oxford: Willey-Blackwill. https://www.wiley.com/en-us/Brain+and+Music-p-9781119943112

Koelsch, Stefan 2013b. Neural correlates of music perception. – *Language, Music, and the Brain. A Mysterious Relationship*. Ed. M. A. Arbib. Cambridge, MA: The MIT Press, 141–172. https://doi.org/10.7551/mitpress/9780262018104.003.0006

Koelsch, Stefan & Siebel, Walter A. 2005. Towards a neural basis of music perception. – *Trends in Cognitive Sciences* 9: 578–584. https://doi.org/10.1016/j.tics.2005.10.001

Kosunen, Ilkka 2018. Sümbiootiline helilooming [Symbiotic composition, in Estonian]. – *Teater. Muusika. Kino* 4: 60–65. https://www.temuki.ee/archives/1793, https://dea.digar.ee/article/AKtmk/2018/04/0/15

Kosunen, Ilkka & Väljamäe, Aleksander 2020. Designing symbiotic composing. – *Acoustical Science & Technology (Acoustical Society of Japan)* 41 (1): 322–325. https://www.jstage.jst.go.jp/article/ast/41/1/41_E19275/_pdf

Kothinti, Sandeep Reddy; Huang, Nicholas & Elhilali, Mounya 2021. Auditory salience using natural scenes: An online study. – *The Journal of the Acoustical Society of America* 150 (4): 2952–2966. https://doi.org/10.1121/10.0006750

Kotta, Kerri 2008. Zur musikalischen Zeit: Über die formalen Grundlagen der Werke von Erkki-Sven Tüür – ein Beschreibungsversuch. – *Schwarze Milch und bunte Steine. Der Komponist Erkki-Sven Tüür*. Hrsg. H.-K. Jungheinrich. Mainz: Schott, 23–38. https://content.e-bookshelf.de/media/reading/L-3454201-1969de63fe.pdf

Kotta, Kerri 2009. Erkki-Sven Tüür – monumentaalsete kõlade looja. – *Sirp*, 15.10.2009, https://www.sirp.ee/erkki-sven-tueuer-monumentaalsete-kolade-looja/

Kotta, Kerri 2009/2010. Erkki-Sven Tüür – 50. – *Music in Estonia* 12, 2–6. https://www.emic.ee/failid/File/Music%20In%20Estonia%20No%2012_2009%20(1).pdf

Kotta, Kerri 2011. Classical formal archetypes in selected works by Erkki-Sven Tüür. – *Lietuvos muzikologija [Lithuanian Musicology]* 12: 102–110. https://zurnalai.lmta.lt/wp-content/uploads/2011/Lietuvos-muzikologija-XII-Kotta.pdf

Kotta, Kerri 2013. On “energy conversion” in the music of Erkki-Sven Tüür. – *De Musica Commentarii. Ars Nova* (4). 103−112. https://www.etis.ee/portal/publications/display/e055335f-36b2-448c-a919-ee7346172083

Kotta, Kerri 2017. *Erkki-Sven Tüür – Peregrinus Ecstaticus. Le poids des vies non vécues. Noesis* (Kerri Kotta, booklet text author pages 2–6). CD Ondine ODE 1287-2. https://www.emic.ee/cd-erkki-sven-tuur-peregrinus-ecstaticus-le-poids-des-vies-non-vecues-noesis-2

Kramer, Lawrence 1991. Musical narratology: a theoretical outline. – *Indiana Theory Review* 12: 141–162. https://scholarworks.iu.edu/iuswrrest/api/core/bitstreams/a2a76f71-c4dc-4feb-a18c-83244ed36b1f/content

Kraus, Katharina T. 2013/2016. Quantifying Inner Experience? Kant’s Mathematical Principles in the Context of Empirical Psychology. – *European Journal of Philosophy* 24 (2): 331–357. https://doi.org/10.1111/ejop.12068

Krueger, Lester E. 1973. Effect of irrelevant surrounding material on speed of same-different judgment of two adjacent letters. – *Journal of Experimental Psychology* 98 (2): 252–259. https://doi.org/10.1037/h0034358

Krumhansl, Carol L. & Lerdahl, Fred 2011. Musical Tension. – *Art and the Senses*. Eds. F. Bacci & D. Melcher. Oxford: Oxford University Press, 311–328.

Le Groux, Sylvain 2011. *Situated, Perceptual, Emotive and Cognitive Music Systems. A Psychologically Grounded Approach to Interactive Music Composition* (Doctoral Dissertation). Barcelona: Universitat Pompeu Fabra. http://hdl.handle.net/10803/32043

Le Groux, Sylvain & Verschure, Paul F. M. J. 2010. Towards adaptive music generation by Reinforcement Learning of musical tension. – *Proceedings of the 7th Sound and Music Computing Conference SMC 2010*. Barcelona, July 21–24, 2010. Eds. E. Gómez; P. Herrera & R. Ramírez. Barcelona: Universitat Pompeu Fabra, 160–165. http://smc.afim-asso.org/smc10/programme.htm, http://smc.afim-asso.org/smc10/proceedings/24.pdf

Lehne, Moritz & Koelsch, Stefan 2015. Towards a General Psychological Model of Tension and Suspense. – *Frontiers in Psychology* 6: 1–11 (Article 79). https://doi.org/10.3389/fpsyg.2015.00079

Lerdahl, Fred 1989. Atonal Prologational Structure. – *Contemporary Music Review* 4: 65–87. https://www.academia.edu/3408915

Lerdahl, Fred 2001. *Tonal Pitch Space*. Oxford: Oxford University Press. https://global.oup.com/academic/product/tonal-pitch-space-9780195178296

Lerdahl, Fred & Jackendoff, Ray S. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/12513.001.0001

Lerdahl, Fred; Krumhansl, Carol L. 2007. Modeling Tonal Tension. – *Music Perception* 24 (4): 329–366. https://doi.org/10.1525/mp.2007.24.4.329

Levitin, Daniel J. 2002. Experimental Design in Psychological Research. – *Foundations of Cognitive Psychology: Core Readings*. Ed. D. J. Levitin. Cambridge, MA: The MIT Press, 115–130. https://mitpress.mit.edu/9780262621595/foundations-of-cognitive-psychology/

Li, Qi 2018. Data visualization as creative art practice. – *Visual Communication* 17 (3): 299–312. https://doi.org/10.1177/1470357218768202

Libet, Benjamin 1990. Cerebral processes that distinguish conscious experience from unconscious mental functions. – *The Principles of Design and Operation of he Brain*. Eds. J. C. Eccles & O. Creutsfeldt. Rom: Pontificia Academia Scientiarum, 185–202.

Libet, Benjamin 2004. *Mind Time. The Temporal Factor in Consciousness*. Cambridge, MA: Harvard University Press. https://www.hup.harvard.edu/books/9780674018464

Liu, Aozhi; Zhu, Zhaohua; Cai, Zifeng; Xie, Zongyang; Mei, Yaqi & Xiaol, Jing 2020. A multidimensional model of music tension. – *Proceedings of the International Conference on New Music Concepts Inspired Education and New Computer Science Generation. Vol 7.* Ed. M. D. Ventura. Treviso: Accademia Musicale Studio Musica, 47–60. http://www.studiomusicatreviso.it/icnmc/icnmc_paper.php?id=73

Lock, Gerhard 2006. Aspekte der Visualisierung und graphischen Analyse von Musik. Die Fünfte Sinfonie von Lepo Sumera. *Musical Work: Boundaries and Interpretations* (Selected Papers of the 38th Baltic Musicological Conference, 21.–23.10.2004, Vilnius). Ed. A. Žiuraityte. Vilnius: Lithuanian Composers Union, 243−269. https://www.researchgate.net/publication/386487944

Lock, Gerhard 2009. Some preliminary concepts of graphical analysis/composition for musicians and nonmusicians. – *CFMAE: The Changing Face of Music and Art Education*, 1 (Proceedings of CFME: The Changing Face of Music "Music and Environment"): 152−173. https://www.tlu.ee/CFMAE/CFMAE_Vol_1_2009.pdf Lock 2011

Lock, Gerhard 2010a. Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses „Oxymoron" [On factors influencing musical tension and its perception in Erkki-Sven Tüür's piece „Oxymoron"]. – *Res musica,* 2: 63−74. https://resmusica.ee/wp-content/uploads/2017/04/rm2_2010_63-74_Lock.pdf

Lock, Gerhard 2010b. Tension design analysis as listening analysis strategy for contemporary music: Perception of tension in contemporary post-tonal orchestral music: a case study (unpublished extended abstract of the poster presentation). – *11th International Conference on Music Perception and Cognition ICMPC 11*. 23.–27.08.2010. Seattle, WA: University of Washington, 1–6 [issued originally in Estonian in Lock 2010a]. http://depts.washington.edu/icmpc11/Program.html, https://www.researchgate.net/publication/348008177

Lock, Gerhard 2011a. Musical creativity in the mirror of Glaveanu's five principles of cultural psychology. – *Culture & Psychology* 17 (1): 121−136. https://doi.org/10.1177/1354067X10388853

Lock, Gerhard 2011b. *Ärkamine/Awakening. Erkki-Sven Tüür* (commissioned by Tallinn – European Capital of Culture 2011; Gerhard Lock, booklet text author pages 5–8). CD Ondine

ODE 1183-2. https://www.researchgate.net/publication/378475815, https://www.epcc.ee/plaadid/awakeningarkamine/

Lock, Gerhard 2012. The Tension Design of Erkki-Sven Tüür's 4th Symphony (unpublished paper presentation). – *Abstracts of the 15th Pedagogical Forum of Performing Arts*. 14.–16.12.2012. Belgrade: University of Arts, Faculty of Music, 47−48. https://www.researchgate.net/publication/385680744

Lock, Gerhard 2014. Visualisierende Analyse [Visualizing analysis]. Erkki-Sven Tüür: *Oxymoron*. – *Musikalische Analyse. Begriffe, Geschichten, Methoden*. Hrsg. F. Diergarten. Laaber: Laaber, 287−319. (Grundlagen der Musik 8, Hrsg. Felix Diergarten & Manuel Gervink). https://www.researchgate.net/publication/284900168

Lock, Gerhard 2017a. Models and scientific modeling in music context. – *Problems in Music Pedagogy* 16 (2): 25−45. http://pmp.du.lv/wp-content/uploads/2016/08/PMP_2017_Vol-16-2.pdf

Lock, Gerhard 2017b. The direction of perception as gesture: the COSM model as example for the analysis of contemporary music (unpublished paper presentation). – *Symposium "Gesture and Music. Gesture and Musical Form". Pärnu Days of Contemporary Music "Performativity 4"*. 13.–15.01.2017. Pärnu: Estonian Arnold Schoenberg Society. https://www.schoenberg.ee/?sitesig=SB&page=SB_060_Arhiiv&subpage=ARH_PNP_2017, https://www.schoenberg.ee/SB_060_Arhiiv/ARH_PNP_2017/PNP_170114_GERHARD_LOCK_SLAIDID.pdf, https://www.researchgate.net/publication/385694030

Lock, Gerhard 2019/2011. Spannungsdesign und Hörstrategien zeitgenössischer Musik am Beispiel von Orchestermusik von Erkki-Sven Tüür. – *Musik im Baltikum: 10 Beiträge zum 23. Baltischen Seminar 2011*. Hrsg. H. Scheunchen. Lübeck: Verlag Carl-Schirren-Gesellschaft (Baltische Seminare), 155−182. https://www.db-kulturwerk.de/baltische-seminare, https://www.researchgate.net/publication/378491089

Lock, Gerhard 2019. Observing relations between perceived musical tension and musical parameters in post-tonal orchestral music (unpublished paper presentation). – *16th International Conference on Music Analysis and Theory "Dichotomous Paths and Clashes: Complexities and Multiplicities in Music Analysis and Theory"*. GATM Gruppo Analisi e Teoria Musicale. 10.–13.10.2019. Rimini: Istituto Superiore di Studi Musicali "G. Lettimi". https://www.researchgate.net/publication/385681160

Lock, Gerhard 2020. Salienz, Narrativität und die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung von post-tonaler Orchestermusik [Salience, Narrativity and the Role of Musical Parameters in the Analysis of Musical Tension in Post-Tonal Orchestral Music]. – *ZGMTH: Zeitschrift der Gesellschaft für Musiktheorie* 17 (2): 311–349. https://doi.org/10.31751/1074

Lock, Gerhard 2021. The role of musical parameters analyzing musical tension in Erkki-Sven Tüür's post-tonal orchestral music (unpublished paper presentation). – *10th European Music Analysis Conference.* 20.–24.09.2021 (Online) (Abstract-Book). Moscow: Society for Theory of Music, Moscow State Tchaikovsky Conservatory, 292–294. https://euromac.mosconsv.ru/en/home-0, https://www.researchgate.net/publication/378593136

Lock, Gerhard 2022. Mitmekeelse teekonna kogemus muusikalise pinge uurimisel ja määratlemisel [The experience of a multilingual journey in exploring and defining musical tension (unpublished paper presentation)]. – *Lühitutvustused. VIII mitmeteaduslik eesti teaduskeele konverents* [Abstracts. VIII Multidisciplinary Conference on Estonian Scientific Language]. 02.12.2022. Tallinn: Tallinna Ülikool, 12−13. https://teaduskeelekeskus.weebly.com/viii-konverents-2022.html, https://www.researchgate.net/publication/385680809

Lock, Gerhard 2023a. Metodoloogilisi arusaamu tajutud kaasaegse posttonaalse orkestrimuusika strukturaalse muusikalise pinge kognitiivseks analüüsiks [Methodological understanding of the cognitive analysis of perceived structural musical tension in contemporary post-tonal orchestral music (unpublished paper presentation)]. – *Abstraktid. 3. Muusikauuringute kevadkonverents: diskussioonid ja dialoogid* [Abstracts. 3rd Spring Conference on Music Studies: Discussions and Dialogues]. 31.05.2023. Tallinn: Eesti Muusika- ja Teatriakadeemia, 12−14. https://eamt.ee/muusikauuringute-kevadkonverents-diskussioonid-ja-dialoogid/, https://www.researchgate.net/publication/385682067

Lock, Gerhard 2023b. Psühhofüüsiline mõõtmine tajutud strukturaalse muusikalise pinge uurimise alusena [Psychophysical measurement as a basis for studying perceived structural musical tension (unpublished paper presentation)]. – *Lühitutvustused. IX mitmeteaduslik eesti teaduskeele konverents* [Abstracts. IX Multidisciplinary Conference on Estonian Scientific Language]. 24.11.2023. Tallinn: Tallinna Ülikool, 14−15. https://teaduskeelekeskus.weebly.com/ix-konverents-2023.html, https://www.researchgate.net/publication/385693729

Lock, Gerhard 2024. Musical tension: Methodological dilemmas. – *Actes des Journées d'Informatique Musicale [Proceedings of the Musical Computer Days] JIM 2024.* 06.–08.05.2024. Marseille: PRISM (Perception, Representations, Image, Sound, Music) Laboratory, CNRS (The National Centre for Scientific Research), Aix-Marseille University, 303–312. https://cloud.prism.cnrs.fr/index.php/s/8aX4YaDXArst2xQ, https://www.researchgate.net/publication/380726031

Lock, Gerhard; de Paiva-Santana, Charles & Kotta, Kerri 2021. Musical Form, Tension and Texture in Tüür's "Flamma" for String Orchestra: A Multiperspective Approach (unpublished paper presentation). – *Tenth European Music Analysis Conference* (20.–24.09.2021) (ONLINE, Abstract-Book). Moscow: Moscow Conservatory, 291−292. https://www.researchgate.net/publication/378592969

Lock, Gerhard; Kotta, Kerri 2012. Musical tension as a response to musical form. – *Music Cognition: 12th International Conference on Music Perception and Cognition (ICMPC) / 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM).* 23.–28.07.2012. Thessaloniki: Aristotle University of Thessaloniki, 612−617. http://icmpc-escom2012.web.auth.gr/files/papers/612_Proc.pdf

Lock, Gerhard & Kotta, Kerri 2015. A preliminary cognitive model of musical tension in contemporary orchestral music (unpublished paper presentation). – *12th International Music Analysis and Theory Conference*. GATM Gruppo Analisi e Teoria Musicale. 24.–27.09.2025. Rimini: Istituto Superiore di Studi Musicali "G. Lettimi". https://www.researchgate.net/publication/385680511

Lock, Gerhard & Lock, Hans-Gunter 2011. Tension design. TEDEA Tension Design Experimental Apparatus, https://www.schoenberg.ee/tension-design/tension-design.html [15.04.2025].

Lock, Gerhard & Sikk, Jaak 2022. Accident and serendipity in music composition, improvisation and performance art. *The Art of Serendipity*. Eds. W. Ross, S. Copeland. Cham: Palgrave Macmillan, Springer Nature Switzerland, 191−219. https://doi.org/10.1007/978-3-030-84478-3_8

Lock, Gerhard & Valk-Falk, Maris 2008. Investigating structure of contemporary music applying tension design and empirical perception analysis. – *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08).* 02.–06.07.2008. Thessaloniki: Aristotle University of Thessaloniki, 1−15. https://www.researchgate.net/publication/378746650

Lock, Gerhard & Valk-Falk, Maris 2009. Tension design of contemporary orchestral music in different performances of "Sula" by Helena Tulve. – *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009).* 12.–16.08.2009. Jyväskylä: University of Jyväskylä, 596–603. http://urn.fi/URN:NBN:fi:jyu-2009411337

Mailman, Joshua B. 2010. *Temporal Dynamic Form in Music: Atonal, Tonal, and Other* (Doctoral Dissertation). Rochester, NY: University of Rochester. http://www.joshuabanksmailman.com/

Mailman, Joshua B. 2012/2013. Agency, Determinism, Focal Time Frames, and Processive Minimalist Music. – *Music and Narrative Since 1900*. Eds. M. Klein & N. Reyland. Bloomington, IN: Indiana University Press, 125–144. https://iupress.org/9780253006448/music-and-narrative-since-1900/, https://www.academia.edu/2315576

Mailman, Joshua B. 2013. Improvising Synesthesia: Comprovisation of Generative Graphics and Music. – *Leonardo Electronic Almanac* 19 (3) (special issue on Live Visuals): 352–384. https://www.leoalmanac.org/wp-content/uploads/2013/07/LEAVol19No3-Mailman.pdf

Mailman, Joshua B. 2016. On a cybernetic phenomenology of music, embodied speculative realism, and aesthetics-driven techné for spontaneous audio-visual expression. – *Perspectives of New Music* 54 (1): 5–95. https://doi.org/10.7757/persnewmusi.54.1.0005

Margulis, Elizabeth H.; Loui, Psyche & Loughridge, Deirdre 2023. *The Science-Music Borderlands. Reckoning with the Past and Imagining the Future*. Cambridge, MA: The MIT Press. http://direct.mit.edu/books/book-pdf/2081702/book_9780262373043.pdf

Marr, David 1982. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, CA: W.H. Freeman & Company, The MIT Press. https://mitpress.mit.edu/9780262514620/vision/

Mayer, Richard E. & Massa, Laura J. 2003. Three facets of Visual and Verbal Learners: Cognitive ability, cognitive style, and learning preference. – *Journal of Educational Psychology* 95 (4): 833–846. https://doi.org/10.1037/0022-0663.95.4.833

McAdams, Stephen 1989. Psychological constraints on form-bearing dimensions in music. – *Contemporary Music Review* 4 (1): 181–198. https://doi.org/10.1080/07494468900640281

McLarty, Michael H.; Braun, William H. & Benitez, Vincent B. 1990. The foundation of a scientific theory of music. – *The Psychological Record* 40: 491–504. https://www.academia.edu/82595271

Merker, Björn 2002. Music: the missing Humboldt system. – *Musicae Scientiae* 6 (1): 3–21. https://doi.org/10.1177/102986490200600101

Merleau-Ponty, Maurice 1945/2005. *Phenomenology of Perception* [translated by Colin Smith]. London, New York: Routledge. https://www.academia.edu/12320612/Merleau_Ponty

Meyer, Leonard B. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/E/bo28551887.html

Michaels, Claire F. & Carello, Claudia 1981. *Direct Perception*. Englewood Cliffs, NJ: Prentice-Hall. https://cespa.uconn.edu/wp-content/uploads/sites/2363/2018/04/Michaels-Carello-1981-Direct-Perception.pdf

Michell, Joel 2017. On substandard substantive theory and axing axioms of measurement: A response to Humphry. – *Theory & Psychology* 27 (3): 419–425. https://doi.org/10.1177/0959354317706746

Michell, Joel 2020. Representational measurement theory: Is its number up?. – *Theory & Psychology* 31 (1), 3–23. https://doi.org/10.1177/09593543209308

Mind-body problem APA 2018. Mind-body problem. American Psychological Association Dictionary of Psychology (APA), https://dictionary.apa.org/mind-body-problem [17.03.2025].

Miller, Matthias; Häußler, Johannes; Kraus, Matthias; Keim, Daniel & El-Assady, Mennatallah 2018. Analyzing Visual Mappings of Traditional and Alternative Music Notation. – *arXiv:1810.10814 [cs.IR]*, 1–5 (IEEE VIS 3rd Workshop on Visualization for the Digital Humanities (21.10.2018, Berlin, Germany, https://vis4dh.dbvis.de/2018/). https://doi.org/10.48550/arXiv.1810.10814

Miller, Matthias; First, Daniel, & El-Assady, Mennatallah [2025]. Music Visualization (MusicVis Project)/Visual Musicology, Interactive Visualization & Intelligence Augmentation (IVIA) Lab (Bridging Human and Artificial Intelligence), ETH Zürich, https://ivia.ch/music, https://visual-musicology.com/ [17.03.2025].

Monzo, Joe [2025]. Tonalsoft. Encyclopedia of Microtonal Music Theory, http://tonalsoft.com/enc/encyclopedia.aspx [17.03.2025].

Morales Ladrón de Guevara, Jorge Francisco 2018. *The Strength of the Mind: Essays on Consciousness and Introspection* (Doctoral Dissertation). New York: Columbia University. https://academiccommons.columbia.edu/doi/10.7916/D8X93T9M/download

Morales Ladrón de Guevara, Jorge Francisco 2024. Introspection is signal detection. – *British Journal for the Philosophy of Science* 75 (1): 99–126. https://doi.org/10.1086/715184

Morris, Robert D. 1993. New directions in the theory and analysis of musical contour. – *Music Theory Spectrum* 15 (2): 205–228. https://doi.org/10.2307/745814

Müller, Meinard; Chew, Elaine & Bello, Juan Pablo 2016. Computational Music Structure Analysis (Dagstuhl Seminar 16092, Schloss Dagstuhl: Leibniz-Zentrum für Informatik). – *Dagstuhl Reports* 6 (2): 147–190. http://www.dagstuhl.de/16092

Nanay, Bence 2007. Four theories of amodal perception. – *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. Nashville, TN: Cognitive Science Society. https://cognitivesciencesociety.org/past-conferences/, https://www.proceedings.com/03510.html

Nanay, Bence 2009/2010. Perception and imagination: amodal perception as mental imagery. – *Philosophical Studies* 150: 239–254. https://doi.org/10.1007/s11098-009-9407-5

Narens, Louis 2001. *Theories of Meaningfulness*. 1st Edition. New York: Psychology Press. https://doi.org/10.4324/9781410604156

Narens, Louis & Luce, R. Duncan 1987/2016. Meaningfulness and Invariance. – *The New Palgrave Dictionary of Economics.* 1st Edition. Eds. J. Eatwell, M. Milgate & P. Newman. London: Palgrave Macmillan, 417–421/1–7. https://doi.org/10.1057/978-1-349-95121-5_826-1

Narens, Louis & Luce, R. Duncan 1990. Meaningfulness and Invariance. – *Time Series and Statistics.* Eds. J. Eatwell, M. Milgate & P. Newman. London: The New Palgrave. Palgrave Macmillan, 140–148. https://doi.org/10.1007/978-1-349-20865-4_19

Narens, Louis & Luce, R. Duncan 2008/2017. Meaningfulness and Invariance. – *The New Palgrave Dictionary of Economics*. 2nd Edition. Eds. S. N. Durlauf & L. E. Blume, 1–9. https://link.springer.com/referenceworkentry/10.1057/978-1-349-95121-5_826-2

Narens, Louis & Mausfeld, Rainer 1992. On the relationship of the psychological and the physical in psychophysics. – *Psychological Review* 99 (3): 467–479. https://doi.org/10.1037/0033-295X.99.3.467

Nattiez, Jean-Jacques 1975. *Fondements d'une Sémiologie de la Musique*. Paris: Union Générale d'Editions. https://ulysse.univ-lorraine.fr/discovery/fulldisplay?docid=alma991001560369705596&context=L&vid=33UDL_INST:UDL&lang=fr&search_scope=MyInst_and_CI&adaptor=Local%20Search%20Engine&tab=Everything&query=undefined&mode=advanced

Nattiez, Jean-Jacques & Dunsby, Jonathan M. 1977. Fondements d'une sémiologie de la musique. – *Perspectives of New Music* 15 (2): 226–233. https://doi.org/10.2307/832821

Nicholson, Thomas & Sabat, Marc 2018. *Fundamental Principles of Just Intonation and Microtonal Composition*. Berlin: Universität der Künste Berlin (Studio für Intonationsforschung und mikrotonale Komposition). http://masa.plainsound.org/pdfs/JI.pdf

Nickols, Fred 2001/2010. The Knowledge inKnowledge Management. – *Knowledge Management Yearbook 2000–2001*. Distance Consulting.

https://www.researchgate.net/publication/229101067, https://www.nickols.us/knowledge_in_KM.pdf

Nieto, Oriol 2015. *Discovering Structure in Music: Automatic Approaches and Perceptual Evaluations* (Doctoral Dissertation). New York: New York University. https://www.proquest.com/openview/09f67403121bcbc7d2ee431985bf0568

Noble, Jason; Thoret, Etienne; Henry, Max & McAdams, Stephen 2020. Semantic dimensions of sound mass music: Mappings between perceptual and acoustic Domains. – *Music Perception* 38 (2): 214–242. https://doi.org/10.1525/mp.2020.38.2.214

Nobre, Anna C. 2018. Attention. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience, Sensation, Perception, and Attention*. *Volume 2*. 4th Edition. Eds. J. T. Wixted & J. Serences. New York: John Wiley & Sons, 241–315. https://www.wiley.com/en-us/Stevens'+Handbook+of+Experimental+Psychology+and+Cognitive+Neuroscience%2C+Volume+2%2C+Sensation%2C+Perception%2C+and+Attention%2C+4th+Edition-p-9781119170044

Noyce, Abigail L., Kwasa, Jasmine A. & Shinn-Cunningham, Barbara G. 2022. Defining attention from an auditory perspective. – *WIREs Cognitive Science* 14 (1): 1–11. https://doi.org/10.1002/wcs.1610

Núñez, Rafael; Allen, Michael; Gao, Richard; Miller Rigoli, Carson; Relaford-Doyle; Josephine & Semenuks, Arturs (2019). What happened to cognitive science? – *Nature Human Behaviour* 3: 782–791. https://doi.org/10.1038/s41562-019-0626-2

Nussbaum, Charles O. 2007. *The Musical Representation: Meaning, Ontology, and Emotion*. Cambridge, MA: The MIT Press. https://mitpress.mit.edu/9780262517454/the-musical-representation/

O'Briant, Mary P. & Wilbanks, W. A. 1978. The effect of context on the perception of music. – *Bulletin of the Psychonomic Society* 12 (6): 441–443. https://doi.org/10.3758/BF03329731

O'Callaghan, Casey 2021. Auditory perception. – *The Stanford Encyclopedia of Philosophy (Fall 2021 Edition)*. Ed. E. N. Zalta, https://plato.stanford.edu/archives/fall2021/entries/perception-auditory [19.04.2025].

O'Regan, J. Kevin; Myin, Erik & Noë, Alva 2004. Towards an analytic phenomenology: The concepts of "bodiliness" and "grabbiness". – *Seeing, Thinking and Knowing. Meaning and Self-Organisation in Visual Cognition and Thought.* Ed. A. Carsetti. Dordrecht, New York: Kluwer Academic Publishers, 103–114. https://doi.org/10.1007/1-4020-2081-3_5

Ockelford, Adam 2006. Implication and expectation in music: a zygonic model. – *Psychology of Music* 34 (1): 81–142. https://doi.org/10.1177/0305735606059106

Ockelford, Adam 2009. Zygonic Theory: Introduction, Scope, and Prospects. – *Zeitschrift der Gesellschaft für Musiktheorie* 6 (1): 91–172. https://doi.org/10.31751/400

Onwuegbuna, Ikenna Emmanuel 2020. Diginalysis: The Man-Machine Collaboration in Music Analysis. – *Toward Super-Creativity – Improving Creativity in Humans, Machines,* *and Human-Machine Collaborations* (online version). Ed. S. M. Brito. London: IntechOpen, 1–7. https://doi.org/10.5772/intechopen.84355

Owens, Andrew & Efros, Alexei A. 2018. Audio-Visual Scene Analysis with Self-Supervised Multisensory Features (arXiv:1804.03641v2 [cs.CV]). – *arXiv: Computer Science, Computer Vision and Pattern Recognition*. https://doi.org/10.48550/arXiv.1804.03641

Oxymoron DC [2025]. Oxymoron. Dictionary.com (DC), https://www.dictionary.com/browse/Oxymoron [17.03.2025].

Parr, Thomas & Friston, Karl J. 2019. Attention or salience? – *Current Opinion in Psychology* 29: 1–5. https://doi.org/10.1016/j.copsyc.2018.10.006

Pautz, Adam 2019. What is the Integrated Information Theory of Consciousness? A Catalogue of Questions. – *Journal of Consciousness Studies* 26 (1–2): 188–215.

Pearce, Marcus T. 2005. *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition* (Doctoral Dissertation). London: City University London. https://openaccess.city.ac.uk/id/eprint/8459/

Pearce, Marcus T. 2018. Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation. – *Annals of the New York Academy of Sciences* 1423 (1) (Special Issue: The Neurosciences and Music VI): 378–395. https://doi.org/10.1111/nyas.13654

Pearce, Marcus 2005–2018. IDyOM Information Dynamics of Music, IDyOM Project. Marcus Pearce, http://www.marcus-pearce.com/idyom/ [28.10.2024].

Phillips, Keith 2018. *A Neuropsychophilosophical Investigation of Musical Improvisation* (Doctoral Dissertation). Manchester: Manchester Metropolitan University in collaboration with The Royal Northern College of Music. https://e-space.mmu.ac.uk/622845/

Polotti, Pietro & Rocchesso, Davide 2008. *Sound to Sense, Sense to Sound. A State of the Art in Sound and Music Computing*. Berlin: Logos Verlag. https://www.logos-verlag.de/cgi-bin/engbuchmid?isbn=1600&lng=eng&id=, https://www.academia.edu/37605984/

Pressnitzer, Daniel; McAdams, Stephen; Winsberg, Suzann & Fineberg, Joshua 2000. Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. – *Perception & Psychophysics* 22: 66–80. https://doi.org/10.3758/bf03212061

Purwins, Hendrik; Herrera, Perfecto; Grachten, Maarten; Hazan, Amaury; Marxer, Ricard & Serra, Xavier 2008a. Computational models of music perception and cognition I: The perceptual and cognitive processing chain. – *Physics of Life Reviews* 5: 151–168. https://doi.org/10.1016/j.plrev.2008.03.004

Purwins, Hendrik; Herrera, Perfecto; Grachten, Maarten; Hazan, Amaury; Marxer, Ricard & Serra, Xavier 2008b. Computational models of music perception and cognition II: Domain-specific music processing. – *Physics of Life Reviews* 5: 169–182. https://doi.org/10.1016/j.plrev.2008.03.005

Reybrouck, Mark 2010. Music cognition and real-time listening: denotation, cue abstraction, route description and cognitive maps. – *Musicae Scientiae* 14 (2): 187–202. https://doi.org/10.1177/10298649100140S211

Reybrouck, Mark 2015. Music as environment: an ecological and biosemiotic approach. – *Behavioral Sciences* 5 (1): 1–26. https://doi.org/10.3390/bs5010001

Reybrouck, Mark 2019. Experience as cognition: musical sense-making and the 'in-time/ outside-of-time' dichotomy. – *Interdisciplinary Studies in Musicology* 19: 53–80. https://doi.org/10.14746/ism.2019.19.4, https://bibliotekanauki.pl/articles/780097.pdf

Rocchesso, Davide & Fontana, Federico 2003. *The Sounding Object*. Firenze: Edizioni di Mondo Estremo. https://www.academia.edu/1926068/The_Sounding_Object

Rodrigues, Indioney Carneiro 2019. *The Rhythmic Idea and the Musical Representation of Time* (Doctoral Dissertation). London: Goldsmiths College, University of London. https://doi.org/10.25602/GOLD.00026877

Roig-Francolí, Migul A. 2008. *Understanding Post-Tonal Music*. McGraw-Hill, Boston. https://archive.org/details/understandingpos0000roig

Rosenthal, Victor 2004. Microgenesis, immediate experience and visual processes in reading. – *Seeing, Thinking and Knowing. Meaning and Self-Organisation in Visual Cognition and Thought*. Dordrecht, New York: Kluwer Academic Publishers, 221–244. https://doi.org/10.1007/1-4020-2081-3_11

Rosin, Heigo 2024. *Kaasaegse instrumentaalkontserdi mõtestamisest Erkki-Sven Tüüri "Magma" näitel – interpreedi vaade* [Considerations on modern instrumental concerto using Erkki-Sven Tüür's "Magma" as an example – an interpreter's view] (Doktoritöö) [Doctoral Dissertation]. Tallinn: Eesti Muusika- ja Teatriakadeemia. https://repo.eamt.ee/r1/2024/04/29/_29.04.2024_12.52.57_Rosin_H_Kaasaegse_instrumentaalkontserdi.pdf

Rothbart, Daniel 2004. General introduction. – *Modeling: Gateway to the Unknown. A Work by Rom Harré*. Ed. D. Rothbart. Amsterdam: Elsevier, vii–xiii. https://shop.elsevier.com/books/modeling-gateway-to-the-unknown/rothbart/978-0-444-51464-6

Rothfarb, Lee 2002. Energetics. – *The Cambridge History of Western Music Theory.* Ed. T. Christensen. Cambridge, UK; New York: Cambridge University Press, 927–955. https://archive.org/details/christensen-2006-the-cambridge-history-of-western-music-theory

Roy, Asim; Perlovsky, Leonid; Besold, Tarek R.; Weng, Juyang & Edwards, Jonathan C. W. 2018. Editorial: Representation in the Brain. – *Frontiers in Psychology (Section Cognition)* 9: 1–3 (Article 1410). https://doi.org/10.3389/fpsyg.2018.01410

Russell, Christine R. 2019. Effects of pitch and rhythm priming tasks on accuracy and fluency during sight-reading. – *Journal of Research in Music Education* 67 (3): 252–269. https://doi.org/10.1177/0022429419851112

Ruwet, Nicolas 1972. *Langage, Musique, Poésie.* Paris: Editions du Seuil. https://archive.org/details/langagemusiquepo0000unse

Saldaña, Johnny 2024. *Developing Theory Through Qualitative Inquiry*. Thousand Oaks, CA: Sage Publications. https://collegepublishing.sagepub.com/products/developing-theory-through-qualitative-inquiry-1-286499

Saupe, Achim & Wiedemann Felix 2015. Narration und Narratologie. Erzähltheorien in der Geschichtswissenschaft. – *Docupedia-Zeitgeschichte*. Hg. vom Zentrum für Zeithistorische Forschung Potsdam, Institut der Leibniz-Gemeinschaft (Version 1.0). http://dx.doi.org/10.14765/zzf.dok.2.580.v1 [17.03.2025].

Savage, Clarence Wade 1966. Introspectionist and behaviorist interpretations of ratio scales of perceptual magnitudes. – *Psychological Monographs: General and Applied* 80 (19): 1–32. https://doi.org/10.1037/h0092980

Saville, Bryan K. & Buskist, William 2003/2008. Traditional Idiographic Approaches: Small-N Research Designs. – *Handbook of Research Methods in Experimental Psychology.* Ed. S. F. Davis. Malden, MA; Oxford: Blackwell Publishing, 66–82. https://www.wiley.com/en-us/Handbook+of+Research+Methods+in+Experimental+Psychology-p-9780470756720

Schafer, R. Murray 1994/1977. *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, VT: Destiny Books. https://archive.org/details/soundscapeourson0000scha

Schubert, Emery 2010. Continuous self-report methods. – *Handbook of Music and Emotion: Theory, Research, Applications*. Eds. P. N. Juslin & J. A. Sloboda. Oxford: Oxford University Press, 223–253.

Schüler, Nico 2000a. Towards a general methodological classification of computer-assisted music analysis. – *Proceedings of the International Computer Music Conference ICMC 2000*. 27.08.–01.09.2000, Berlin. Ann Arbor: Michigan Publishing. https://researchr.org/publication/icmc-2000

Schüler, Nico 2000b. Towards a methodological view on (computer-assisted) music analysis. – *Bridges. Mathematical Connections in Art, Music, and Science.* Proceedings, Southwestern College Winfield, Kansas, USA, 28–31 July 2000. https://archive.bridgesmathart.org/2000/bridges2000-99.pdf

Schüler, Nico 2007. From musical grammars to music cognition in the 1980s and 1990s: highlights of the history of computer-assisted music analysis. – *Musicological Annual* 43 (2): 371–396. https://doi.org/10.4312/mz.43.2.371-396

Schultz, Benjamin G.; Brown, Rachel M. & Kotz, Sonja A. 2021. Dynamic acoustic salience evokes motor responses. – *Cortex* 134: 320–332. https://doi.org/10.1016/j.cortex.2020.10.019

Schwitzgebel, Eric 2024. Introspection. – *The Stanford Encyclopedia of Philosophy* (Fall 2024 Edition). Eds. E. N. Zalta & U. Nodelman, https://plato.stanford.edu/archives/fall2024/entries/introspection/ [04.11.2024].

Seashore, Carl E. 1938. *Psychology of Music*. New York: McGraw-Hill. https://archive.org/details/psychologyofmusi030417mbp

SE-M290 [2025]. SE-M290, Fully-Enclosed Dynamic Headphones. Pioneer, https://www.pioneerelectronics.ca/POCEN/Portable/Headphones/Headband/SE-M290 [15.04.2025].

Shea, Nicholas 2018. *Representation in Cognitive Science*. Oxford: Oxford University Press. doi.org/10.1093/oso/9780198812883.001.0001

Sheldrake, Rupert 2012a. *The Science Delusion: Freeing the Spirit of Enquiry*. London: Hodder & Stoughton. http://argos.vu/wp-content/uploads/2016/05/the-science-delusion-rupert-sheldrake.pdf

Sheldrake, Rupert 2012b. *Science set Free: 10 Paths to New Discovery* [US version of The Science Delusion]. New York: Deepak Chopra Books. https://archive.org/details/sciencesetfree100000shel/page/n5/mode/2up

Sheldrake, Rupert 2020. *The Science Delusion: Freeing the Spirit of Enquiry.* Updated Edition. London: Hodder & Stoughton. https://www.sheldrake.org/books-by-rupert-sheldrake

Shepard, Roger N. 1978. On the status of "direct" psychophysical measurement. – *Perception and Cognition. Issues in the Foundation of Psychology* (Vol IX Minnesota Studies in the Philosophy of Science). Ed. C. W. Savage. Minneapolis, MN: University of Minnesota Press, 441–490. https://philpapers.org/archive/LEWMBA.pdf

Silins, Nicholas 2024. Perceptual Experience and Perceptual Justification. – *The Stanford Encyclopedia of Philosophy (Fall 2024 Edition)*. Eds. E. N. Zalta & U. Nodelman, https://plato.stanford.edu/archives/fall2024/entries/perception-justification/ [17.03.2025].

Smit, Houston 2000. Kant on Marks and the Immediacy of Intuition. – *The Philosophical Review* 109 (2): 235–266. https://doi.org/10.2307/2693581

Snyder, Robert 2000. *Music and Memory. An Introduction*. Cambridge, MA: The MIT Press. https://mitpress.mit.edu/9780262692373/music-and-memory/

Stang, Nicholas F. 2024. Kant's Transcendental Idealism. – *The Stanford Encyclopedia of Philosophy* (Spring 2024 Edition). Eds. E. N. Zalta & Uri Nodelman, https://plato.stanford.edu/archives/spr2024/entries/kant-transcendental-idealism/ [17.03.2025].

Staufer, Amanda N. 2018. The unifying strands: formalism and gestalt theory in the musical philosophies of Aristoxenus, Descartes, and Meyer. – *Musical Offerings* 9 (1): 31–41 (Article 3). http://dx.doi.org/10.15385/jmo.2018.9.1.3

Steingrimsson, Ragnar 2016. Subjective intensity: Behavioral laws, numerical representations, and behavioral predictions in Luce's model of global psychophysics. – *Journal of Mathematical Psychology* 75: 205–217. https://doi.org/10.1016/j.jmp.2016.06.004

Stern, Daniel N. 2010. *Forms of Vitality. Exploring Dynamic Experience in Psychology, Arts, Psychotherapy, and Development*. Oxford: Oxford University Press. https://doi.org/10.1093/med:psych/9780199586066.001.0001

Stevens, Stanley S. 1951. Mathematics, measurement, and psychophysics. – *Handbook of Experimental Psychology*. Ed. S. S. Stevens. Hoboken, NJ: Wiley, 1–49.

Stratilková, Martina 2015. Hans Mersmann and the Analysis of the New Music. – *Musicologica Olomucensia* 22: 109–119. https://doi.org/10.5507/mo.2015.016, https://www.researchgate.net/publication/304462504

Sun, Lijun; Hu, Li; Ren, Guiqin & Yang, Yufang 2020. Musical tension associated with violations of hierarchical structure. – *Frontiers in Human Neuroscience* 14: 1–11 (Article 578112). https://doi.org/10.3389/fnhum.2020.578112

Teller, Paul 2018. Measurement accuracy realism. – *The Experimental Side of Modeling.* Eds. B. C. van Fraasen & I. F. Peschard. Minneapolis, MN: University of Minnesota Press, 273–298. https://doi.org/10.5749/j.ctv5cg8vk.15

Teo, Yvonne 2020. Measuring harmonic tension in post-tonal repertoire. – *Empirical Musicology Review* 15: 61–104. https://doi.org/10.18061/emr.v15i1-2.6994

Thoegersen, Peter A. 2022. *Polytempic Polymicrotonal Music*. Singapore, New York: Jenny Stanford Publishing, Taylor & Francis. https://doi.org/10.1201/9781003311683

Thorpe, M., Ockelford, A., & Aksentijevic, A. 2012. An empirical exploration of the zygonic model of expectation in music. – *Psychology of Music* 40 (4): 429–470. https://doi.org/10.1177/0305735610392103

Tononi, Giulio; Boly, Melanie; Massimini, Marcello & Koch, Christof 2016. Integrated information theory: from consciousness to its physical substrate. – *Nature Reviews Neuroscience* 17: 450–461 (2016). https://doi.org/10.1038/nrn.2016.44

Tordini, Francesco; Bregman, Albert S.; Ankolekar, Anupriya; Sandholm, Thomas E. & Cooperstock, Jeremy R. 2013. Your attention, please! Determining saliency of competing audio stimuli in natural scenarios. – *Proceedings of the Meetings on Acoustics* 19 (1): 1–9 (Article 010031). https://doi.org/10.1121/1.4800201

Tüür, Erkki-Sven 2001. Erkki-Sven Tüür. Desert Island. CD Finlandia Records 8573-82187-2. https://www.discogs.com/release/12849851-Erkki-Sven-T%C3%BC%C3%BCr-Desert-Island

Tüür, Erkki-Sven 2012/2023. Erkki-Sven Tüür – Igapäevastest asjadest. Saade Eesti Raadio saatesarjast "Raadio Ööülikool" [On everyday things. Feature of the Estonian National Broadcasting/Estonian Radio "Radio Night University"]. Eds. Joonas Hellerma, Jaan Tootsen, sound regisseur Külli Tüli, broadcasted in Vikerraadio 09.11.2012, re-broadcasted 26.08.2023, duration 52'47'', https://vikerraadio.err.ee/1609061159/ooulikool-erkki-sven-tuur-igapaevastest-asjadest, listened 28.04.2025 [in Estonian].

Valenzuela-Moguillansky C, Demšar E. 2023. Toward a science of experience: Outlining some challenges and future directions. – *Adaptive Behavior* 31 (5): 449–466. https://doi.org/10.1177/10597123221084739 van Fraasen, Bas C. 2018. Introduction: The changing debates about measurement. *The Experimental Side of Modeling*. Eds. B. C. van Fraasen & I. F. Peschard. Minneapolis, MN: University of Minnesota Press, 267–272. https://doi.org/10.5749/j.ctv5cg8vk.14

Varto, Juha 2009. *Basics of Artistic Research: Ontological, Epistemological and Historical Justifications*. Helsinki: University of Art and Design Helsinki. https://finna.fi/Record/lut.99876885406254/Versions

Velleman, Paul F., & Wilkinson, Leland 1993. Nominal, ordinal, interval, and ratio typologies are misleading. – *The American Statistician* 47 (1): 65–72. https://doi.org/10.1080/00031305.1993.10475938

Vessonen, Elina 2021. Representation in measurement. – *European Journal for Philosophy of Science* 11: 1–23 (Article 76). https://doi.org/10.1007/s13194-021-00365-6

Vicario, Giovanni B. 2003. Prolegomena to the perceptual study of sounds. *The Sounding Object.* Eds. D. Rocchesso & F. Fontana. Firenze: Edizioni di Mondo Estremo, 17–32. https://www.academia.edu/1926068/The_Sounding_Object

Viégas, Fernanda & Wattenberg, Martin [2025]. Hint.Fm (data visualization and analytics). http://hint.fm/ [17.03.2025].

Wang, DeLiang 2007. Computational Scene Analysis. – *Challenges for Computational Intelligence.* Eds. W. Duch & J. Mandziuk. Berlin: Springer, 163–191. https://doi.org/10.1007/978-3-540-71984-7_8

Wang, DeLiang & Brown, Guy J. 2006. *Computational Auditory Scene Analysis – Principles, Algorithms, and Applications*. Piscataway, Hoboken, NJ: IEEE Press, Wiley. http://dx.doi.org/10.1109/9780470043387

Warren, William H. 2005. Direct perception: the view from here. – *Philosophical Topics* 33 (1): 335–361. http://www.jstor.org/stable/43154721

Watzl, Sebastian 2023. What attention is. The priority structure account. – *WIREs Cognitive Science* 14 (1): 1–10 (Article e1632). https://doi.org/10.1002/wcs.1632

Westphal, Jonathan 2016. *The Mind–Body Problem*. Cambridge, MA: The MIT Press. https://resource.odmu.edu.ua/chair/download/155381/CkOSF0pFhP-vWQ3G5GcEmw/The_Mind-Body_Problem.pdf

Westphal, Jonathan [2025]. Descartes and the discovery of the mind-body problem. The MIT Press Reader, https://thereader.mitpress.mit.edu/discovery-mind-body-problem/ [17.03.2025].

Wiersma, Quint 2024. *Dynamic Models for Multi-Dimensional Time Series* (Doctoral Dissertation). Amsterdam Vrije Universiteit Amsterdam. https://doi.org/10.5463/thesis.500

Wiggins, Geraint A. 2009. Semantic Gap?? Schemantic Schmap!! Methodological considerations in the scientific study of music. – *Proceedings of the 11th IEEE International Symposium on Multimedia*. Piscataway, NJ: IEEE Press, 477–482. https://doi.org/10.1109/ISM.2009.36

Wiggins, Geraint A. 2010. Cue abstraction, paradigmatic analysis and information dynamics: towards music analysis by cognitive model. – *Musicæ Scientiæ* 14 (2) (Special issue): 307–331. https://doi.org/10.1177/10298649100140S217

Wiggins, Geraint A.; Pearce, Marcus T. & Müllensiefen, Daniel 2012. Computational modeling of music cognition and musical creativity. – *The Oxford Handbook of Computer Music.* Ed. R. T. Dean. Oxford: Oxford University Press, 383–420. https://doi.org/10.1093/oxfordhb/9780199792030.013.0019

Wikgren, Wolf 2019. *Beyond Shape. An Exploration in Alternative Forms for Data Visualization* (master's thesis). Aalto: Aalto University. https://aaltodoc.aalto.fi/server/api/core/bitstreams/ddbb7f5a-a494-4c97-9c82-18c5a5d66345/content

Wolff, Robert P. 1963. *Kant's Theory of Mental Activity: A Commentary on the Transcendental Analytic of the Critique of Pure Reason*. Cambridge, MA: Harvard University Press. https://archive.org/details/kantstheoryofmen0000wolf

Wolfram, Stephen 2002. *A New Kind of Science*. Champaign, IL: Wolfram Media. https://www.wolframscience.com/nks/ (Online version) [17.03.2025].

Wolfram, Stephen 2020. *A Project to Find the Fundamental Theory of Physics*. Champaign, IL: Wolfram Media. https://www.wolfram-media.com/products/a-project-to-find-the-fundamental-theory-of-physics/, https://www.wolframphysics.org/technical-introduction/ (Online version) [17.03.2025].

Wolfram, Stephen 2021. The Concept of the Ruliad. Stephen Wolfram Writings, 10.11.2021, https://writings.stephenwolfram.com/2021/11/the-concept-of-the-ruliad/ [17.03.2025].

Wolfram, Stephen 2023. Observer Theory. Stephen Wolfram Writings, 11.12.2023, https://writings.stephenwolfram.com/2023/12/observer-theory/ [17.03.2025].

Xenharmonic Wiki [2025]. Main Page. Xenharmonic Wiki, https://en.xen.wiki/w/Main_Page [17.03.2025].

Yang, Cheng-Ta 2011. Relative saliency in change signals affects perceptual comparison and decision processes in change detection. – *Journal of Experimental Psychology: Human Perception and Performance* 37 (6): 1708–1728. https://doi.org/10.1037/a0024257

Zbikowski, Lawrence M. 1997. Conceptual models and cross-domain mapping: new perspectives on theories of music and hierarchy. – *Journal of Music Theory* 41 (2): 193–225. https://doi.org/10.2307/843958

Zbikowski, Lawrence M. 2002. *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. Oxford: Oxford University Press. https://zbikowski.uchicago.edu/cm1.html

Scores and recordings

Scores

Tüür, Erkki-Sven 2002. Symphony No. 4 ‘Magma’, Solo Percussion, Orchester (full score, EP10897). Frankfurt, Leipzig: Edition Peters. https://www.erkkisven.com/worklist-and-reviews/works-for-ensemble

Tüür, Erkki-Sven 2003. Oxymoron (Music for Tirol) for large ensemble (full score, EP 10907). Frankfurt, Leipzig: Edition Peters. https://www.erkkisven.com/worklist-and-reviews/works-for-orchestra

Recordings

Tüür, Erkki-Sven 2002/2007. Tüür Magma. Symphony No. 4 ‘Magma’ for Solo percussion and Symphony Orchestra (2002) [...]. Dame Evelyn Glennie (percussion), Estonian National Symphony Orchestra (ERSO, cond. Paavo Järvi) (Martin Anderson, booklet text author pages 4–13). CD Virgin Classics (EMI Records), 0946 3 85785 2 9. https://www.erkkisven.com/recordings/24/magma, https://erso.ee/en/records/erkki-sven-tuur-magma-evelyn-glennie-paavo-jarvi-virgin-classics-2007/

Tüür, Erkki-Sven 2003/2007. Oxymoron (2003). Nyyd-Ensemble (cond. Olari Elts) (Erkki-Sven Tüür, booklet text author pages 4–8; Hans-Klaus Jungheinrich, booklet text author pages 11–22). CD ECM Records (ECM New Series 1919), 476 5778). https://www.erkkisven.com/recordings/23/oxymoron, https://ecmrecords.com/product/erkki-sven-tuur-oxymoron-pedro-carneiro-leho-karin-marrit-gerretz-traksmann-vox-clamantis-nyyd-ensemble-estonian-national-symphony-orchestra-olari-elts/

Töö lühikokkuvõte

Väitekirja "Metoodikat kognitiivselt analüüsimaks nüüdses posttonaalses orkestrimuusikas tajutud strukturaalset muusikalist pinget" taustsiht on põimida muusikaanalüüsi ja muusikapsühholoogiat kognitiivselt lähenedes. Laiem siht on arendada analüüsimeetodeid, et edendada arusaamist keerukatest muusikalistest struktuuridest ja kaasaegsest ehk nüüdsest posttonaalsest orkestrimuusikast (inglise *contemporary post-tonal orchestral music*, CPTOM) üldiselt (üldine objekt), vaadeldes ja analüüsides väitekirjas eriti muusikalist pinget kui n-ö terviknähtust (inglise *compound phenomenon*) (eriobjekt), mis on universaalselt mõistetav (Farbood 2012, Lehne & Koelsch 2015), kuid keeruline süstemaatiliselt uurida (publikatsioon V). Mõistan muusikat kui keskkonda (Kaipainen 1994, Reybrouck 2015, publikatsioonid III, IV) ning kaasan muusika visualiseerimist ja kujutamist (inglise *representation*) (publikatsioon II) mudeldamise ja analoogiate osana (publikatsioon III). Peamine siht on tuvastada, kirjeldada, mõtestada ja mõista strukturaalseid aspekte, mis väidetavalt *vallandavad* (inglise *trigger*) tähelepanelikul kuulajal (Deliège & Mélen 1997) muusikalise pinge kogemise ajalise dünaamilise lainetaolise nähtusena (inglise *temporal dynamic wave-like phenomenon*, TDWP). Selle sihini jõudmaks olen määratlenud ja empiiriliselt eritlenud muusikalist pinget eesti nüüdisheliloója Erkki-Sven Tüüri (s 1959) keerukas ja rist-stiililises posttonaalses sümfoonilises orkestrimuusikas: 4. sümfoonia-löökpillikontsert *Magma* (2002) ja *Oxymoron* (2003) suurele ansamblile. Tüüri muusikas on ühendatud klassikaline hierarhiline (strukturaalne, lineaarne, teleoloogiline) ja kõlal põhinev (ruumiline, mittelineaarne, hierarhiatu) komponeerimise viis eesmärgiga võimaldada kuulajatel tajuda tema muusika tekkelugu detailidest tervikteoseni n-ö integreeritud kognitsioonina (Tüüri sõnastuses „kuulaja jälgivate meelte kompleks", inglise *integrated cognition*) tänapäeva maailmas (Kotta 2009, 2009/2010). Muusikaline faktuur rullub (alates 2002/2003) lahti pigem „aeglase", „kiire" ja „pideva" ajakorralduse tunnuste sulandumise kaudu (Kotta 2008, 2011; publikatsioon I). Selgete klassikaliste piiride asemel moodustavad muusikalise vormi pingelised maksimum- ja miinimumpunktid (publikatsioon I). Varem on uuritud nt muusikalise pinge tajumist atonaalses ja mitte-kolmkõlalises muusikas seoses struktuuri ja vormiga (Addessi & Caterina 2000, 2005; Kirschbaum 2001; Addessi 2010), mittetonaalse orkestrimuusika kõlavärvi (Pressnitzer *et al* 2000) neoklassitsistliku muusika harmoonilise pinge (Teo 2020) empiiriliseks analüüsimiseks.

Sissejuhatava peatüki 1. osa esitleb väitekirja põhieesmärki ja üldist objekti ning laiemat sihti ja eriobjekti ja selle täpsemat piiritlust, põhjusi rakendamaks kognitiivset lähenemist ning põhilisi metoodikaküsimusi. 2. osas arutatakse muusikateoreetilise ja muusikapsühholoogilise mõtlemise põimumist (nt Clifton 1975, Broman & Engebretsen 2007), esitatakse metodoloogilis-filosoofilisi vaateid (nt Kanti (1998) *analüütilist idealismi*), ajalis-dünaamilisi kognitiivseid protsesse (nt Reybrouck 2010, Mailman 2010, Stern 2010) ja *mikrogeneesi* (Bachmann 2000, Rosenthal 2004), psühhofüüsilisi mõõtmispõhimõtteid (e.g. Shepard 1978, Steingrimsson 2016), kognitiivseid dünaamilisi kuulamismudeleid (nt Kaipainen 1994, Pearce 2005, 2018; Le Groux 2011; Wiggins 2010; Huron 2006) ning *protonarratiive* kui muusikalise pinge n-ö eeldefinitsiooni (Imberty 2000). 3. osas tutvustatakse põhimõisteid: muusika ja pinge (nt Addessi & Caterina 2000; Kirschbaum 2001; Eitan & Granot 2007; Addessi 2010; Farbood 2006, 2012; Herremans & Chuan 2017), salients ehk väljapaistvus ja tähelepanu (nt Lerdahl 1989, 2001; Jensenius 2002; Nobre 2018; Parr & Friston 2019; Noyce, Kwasa & Shinn-Cunningham 2022), salients kompleksetes helides (inglise *complex sounds*) ja nüüdses muusikas (nt Dubnov & McAdams 2006; Coath et al 2009, 2013; Belkin [2025]), muusikalised parameetrid (nt Christensen 1996, 2012; Snyder 2000) ja narratiivsus (nt Kramer 1991, Gruhn 1992, Brower 2000, Almén 2003, Mailman 2012/2013). Muusikalist pinget on muusikateoreetiliselt määratletud kui (tajutud) strukturaalset muusikalist pinget (publikatsioonid I, Ia, IV). Hiljem olen seda käsitletud kui kognitiivset muusikalist pinget (inglise *Cognitive Musical Tension*, CMT) (publikatsioon V). Tutvustan oma pinget, kognitsiooni ja narratiivsust liitvat valdkondadeülest määratlust 'tajutud strukturaalne muusikaline pinge' (TSMP, inglise '*Perceived structural musical tension*', PSMT, publikatsioon V, vt 3.5), esitades ühtlasi viit teaduslikku dilemmat, uurimisküsimust, seisukohavõttu ja vastuseid (publikatsioon V). Väitekirja jaoks olen teinud empiirilisi katseid nii üksikute (publikatsioonid I, Ia, II, N=7, N=6) kui ka mitme liugur-kontrolleriga (inglise *Tension Design Experimental Apparatus*, TEDEA, Lock, G. & Lock, H.-G. 2011, publikatsioon IV, N=26), samuti minu spetsiaalselt välja töötatud kognitiivse kaheksanurkse viilmudeliga (inglise *Cognitive Octagonal Slice Model*, COSM, N=14) muusikaliste sündmuste kui nn impulsside ehk muutmishetkede ja muusikaliste parameetrite salientsi tuvastamiseks. 4. osas on esitatud uurimismetoodika ja -disain, mis laiendab Guigue ja de Paiva Santana (2018) metoodikaskeemi, kasutab triangulatsiooni ning sisaldab ka retoorilist/rotatsioonilist vormianalüüsi (Kotta 2011, publikatsioon IV). Lisaks on esitletud COSM-i arendamist toetavad vastused küsimustele, mis kuuluvad minu teaduslike mudelite mudeldamise/analüüsi kaheteistkümne strateegilise sammu juurde (inglise *Twelve Strategic Steps for modeling/analyzing of scientific models*, TSSM, publikatsioon III). 5. osast leiab kõikide kaasatud publikatsioonide laiendatud kokkuvõtted ja selgitatakse nende omavahelisi seoseid. 6. osa esitleb väitekirja kokkuvõtteid ja järelmeid. Lisades asuvad kõik kaasatud publikatsioonid (A.1–6); ülevaade tehnilistest sammudest väitekirja uurimisdisaini läbiviimiseks (B); metoodikaküsimused, mis kerkisid publikatsioonide I ja Ia järel (C.1–2); ning mõned olulised COSMi üldtulemused (D). Andmeid olen analüüsinud n-ö tihedusel põhineva klastrite avastamise algoritmiga suurtes ruumilistes andmebaasides koos müraga (inglise *Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*, DBSCAN, Ester et al 1996) klastrite leidmise ja nn aegridade mustrite (inglise *time series patterns*) kaudu (Fu 2011, Jebb jt 2015, Nieto 2015, Feremans jt 2015). Põhiuuringu (publikatsioon IV) konkreetsed tulemused näitavad, et Tüüri 4. sümfoonias-löökpillikontserdis sai nn teiseste parameetrite (inglise *„secondary" parameters*, SP) oletatav kõrgem salientsi tase nn esmaste (inglise *„primary" parameters*, PP) parameetrite suhtes vaid osaliselt kinnitust. Üldtulemused (vt lisa D) näitavad, et muusikalised parameetrid dünaamika (SP), rütm (PP) ja instrumentatsioon/tämber (SP) kui nn kõikide salientsitasandite parameetrid (inglise *All salience level parameters*, AL) on enam salient kui ülejäänud, nn põhisalientsitasandi parameetrid (inglise *Basic salience level parameters*, BL). See tulemus on kooskõlas üldiste tähelepanekute/ootustega nii muusikalise pinge, muusikaliste parameetrite kui ka Tüüri kompositsioonivahendite ja tema n-ö integreeritud kognitsiooni taotlustega. Väitekirja uudsus ja saavutused on peamiselt metoodikas: rakendatud eksperimente (s.h COSMi) võib mõista kui n-ö psühhofüüsilist mõõtmisprotseduuri, mille abil tuvastada n-ö impulsside ehk muutmishetkede ja muusikaliste parameetrite salientsi n-ö vaevu märgatava erinevusena (inglise *„just noticeable difference"*, JND). COSM esindab välist, intersubjektiivselt jagatud ruumilist standardobjekti (Kraus 2013/2016), et saavutada audiovisuaalsel salientsil põhinev tähenduslik Kantilik objektiivne arusaam muusikalisest pingest CPTOM-is.

Märksõnad: taju, kognitsioon, kogemus, tähelepanelik kuulamine, psühhofüüsiline mõõtmine, salients, tajutud strukturaalne muusikaline pinge, posttonaalne orkestrimuusika.

Appendices

A. Reprints of the included publications

A.1 Publication I (Lock 2010a, 5.1)

“Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses “Oxymoron” [On factors influencing musical tension and its perception in Erkki-Sven Tüür’s piece “Oxymoron”]” (in Estonian)

Musical tension, musical form and salient features – *Oxymoron*

Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses „Oxymoron”

Gerhard Lock

Sissejuhatus

Käesolev artikkel keskendub muusikalisele pingele (*tension*) ja selle kujunemisele muusikateoses, kasutades materjalina Erkki-Sven Tüüri teost „Oxymoron"[1] suurele ansamblile. Pinge on muusikas niihästi strukturaalne kui ka tajukategooria ning seetõttu on selle analüüsimisel otstarbekas rakendada kombineeritult kahte eri lähenemisviisi: muusikateoreetilist[2] ja empiirilist (tajukatsed). Post- ja vabatonaalse muusika analüüsis puudub seni üldaksepteeritav meetod, mis hõlmaks analüüsitava teose kõiki olulisemaid aspekte. Ülalmainitud meetodite kombineerimise tulemusena on autori üldisemaks eesmärgiks välja pakkuda selline meetod muusikalist pinget põhjustavate struktuuriaspektide kirjeldamise ja integreerimise põhjal. Seetõttu püütakse nimetatud meetodi väljatöötamisel käsitleda peamiselt post- ja vabatonaalset nüüdismuusikat.

Nagu öeldud, keskendub antud artikkel Erkki-Sven Tüüri teosele „Oxymoron", millest leiab käsitlemist peamiselt teine pool. Uurimus kõrvutab seitsme osalejaga tehtud tajukatsete tulemusi teose struktuuriga ning selgitab, kuidas peamised kulminatsioonialad on seotud teose muusikalise materjali ja selle vormiga. Sellest ülesandepüstitusest lähtudes esitatakse siin töös järgmised küsimused:

- missugust rolli mängivad kulminatsioonid analüüsitava teose vormi kujunemisel,
- kuidas korreleerivad tajukatse osalejate kõverad omavahel ja
- kuidas suhestub partituuri ja salvestuse põhjal tehtud vormianalüüs valitud tajukatse graafide ning kõikide kõverate keskmiste näitajate kõveraga.

Võiks arvata, et kuulajad tajuvad muusikateoses pinge kulgemist nii üldjoontes kui ka vormi olulistes kohtades samamoodi, s.t. et teose pinge tõusud ja langused on üldises plaanis olemuselt võrdlemisi objektiivsed nähtused ning seega on nende põhjal võimalik välja arendada pingedisainil[3] (*tension design*) põhinev analüüsimeetod.

Pinge käsitlustest muusikas

Encyclopaedia Britannica defineerib pinget kui „kunstiteostes (nt. luuletuses, maalis või heliteoses) säilitatud tasakaalu vastandlike jõudude või elementide vahel; juhitud dramaatilist või dünaamilist kvaliteeti".[4] Üldisemas plaanis on muusikalist pinget kogu 20. sajandi vältel püütud mõista nii strukturaalselt, füüsikaliselt kui ka psühholoogiliselt. Nimetagem siin kas või eelmise sajandi esimese poole orgaanilis-energeetiliste, fenomenoloogiliste ja geštalt-psühholoogiliste suundade ehk „energeetikute" esindajaid, tuntuimad neist Ernst Kurth ja Heinrich Schenker (ülevaadet neist vt. nt. Rotfarb 2002, Thaler 1984). Mainima peab ka Paul Hindemithi (1939 ja 1940), kes reastab intervallid nende pingestatuseastme järgi, eristades seejuures intervalli esinemist kas horisontaalis või vertikaalis. Lee Rotfarbi sõnul võib „energeetikute" teoorial olla ajaloolise referentspunktina oluline roll ka nüüdismuusika analüüsimisel (Rotfarb 2002: 951). Sajandi teisel

---

[1] Helilooja ei anna CD-plaadi vihikus otseselt selgitust teose „Oxymoron" pealkirja kohta, kuid mainib vastandlikkuse printsiipi vertikaali ja horisontaali vaheldumises, mida ühendab n.-ö. peidetud muusikaline põhimaterjal (Tüür 2007: 4, 6–7). Hans-Klaus Jungheinrich selgitab samas kohas lähemalt muusikalise materjali heterogeensete või vastandlike iseloomujoonte ja muusikalist kulgemist, mis võivad igal hetkel muutuda oma vastandiks (vt. Tüür 2007: 11–12, 18). Teisalt viitab sama autor pealkirja mainides vastanduste „võimatu" ühendamise (s.t. „vägivaldne mahedus" või „lõtv tahkus") nähtusele, mis jõuab üsna lähedale mõiste lingvistilisele definitsioonile (semantiliselt vastandlike mõistete ühendus).

[2] Muusikateoreetilist analüüsi on siin mõistetud võrdlemisi laialt: see võib hõlmata nii automaatsete (nt. statistiliste) kui ka rohkem interpreteerivate analüüsitehnikate (nt. Schenkeri analüüs) kasutamise.

[3] Pingedisaini kohta vt. lähemalt veel Lock, Valk-Falk 2008. Tegemist on alles väljatöötatava meetodiga, millel käesoleva artikli autori arvates muusika analüüsis otsest analoogi ei ole.

[4] „Tension: a balance maintained in an artistic work (such as a poem, painting, or musical composition) between opposing forces or elements; a controlled dramatic or dynamic quality." <www.britannica.com> (vaadatud 22.06.2008).

poolel on aga järjest enam tähtsust omandanud muusika uurimine seoses psühholoogias rakendatavate empiiriliste meetoditega (Krumhansl 1985, Cook 1987 ja 1994, Clarke 1989, ülevaade vt. Deutsch, Sloboda jt. 2007; vt. ka Lerdahl 2001: 188–192), nt. käsitletakse segmentatsiooni ja muusikalist pinget klassikalises muusikas (Krumhansl 1996), segmentatsiooni ja vormianalüüsi (Deliège 1989, Deliège, Melen jt. 1996, Baroni 2003). Kognitiivse muusikateooria tähtsust rõhutavad tonaalse muusika generatiivse teooria rajajad Fred Lerdahl ja Ray Jackendoff (1983, ka Lerdahl 2001), kes on mõjutatud lingvistilisest strukturalismist. Seega on muusikaline pinge olemuselt väga mitmetasandiline nähtus ning selle võimalikult terviklikuks hõlmamiseks on vaja üheaegselt rakendada eri valdkondade meetodeid.

Pingest rääkides kasutatakse enamasti mõisteid pingetõus–pingelangus (*tension–relaxation*) ning kõrgpunkt, kulminatsioon jne. Autoritest, kes oma töödes käsitlevad muusikalist pinget strukturaalselt olulise nähtusena, tuleb nimetada eelkõige juba mainitud Fred Lerdahli ja Ray Jackendoffi (1983). Selles raamatus uurivad nad pinget peamiselt prolongatsiooni (harmoonilise hierarhia) kontekstis. Hiljem on Fred Lerdahl (2001) oma ideid ka edasi arendanud. Viimases on pingeaspekt võrreldes eelnevaga oluliselt tähtsamal kohal, kusjuures pinget käsitletakse siin ka 20. sajandi muusika analüüsimisel.

Lerdahl mõistab muusikalise pinge all peamiselt tonaalset ehk harmoonilist pinget, mille ta omakorda liigitab järjepidevaks (*sequential*) ja hierarhiliseks pingeks. Lerdahl seostab pinge kaht eri tüüpi, vastavalt n.-ö. naiivse ja kogenud kuulajatüübiga. Kui nn. naiivsed kuulajad jäävad kuulamisel pigem pinnatasanditele (jälgides muusikalisi sündmusi üksnes nende ajalises kulgemises), siis kogenud kuulajad kalduvad kuulatavat ka hierarhiseerima. Mõlemad mudelid on tema sõnul vaid idealiseeritud: kui järjepideva mudeli puhul võrreldakse pelgalt üksteisele järgnevaid muusikalisi sündmusi, siis hierarhilise mudeli puhul proovitakse kuulatavat sündmust sobitada kujutlusega teose tervikstruktuurist (Lerdahl 2001: 143). Lisaks eelmainitutele nimetab Lerdahl oma raamatus ka muusikalist pinget, mis tekib kasvava dünaamika (*crescendo*), tõusva meloodiakontuuri ja tiheneva faktuuri tagajärjel. Samuti mainib ta psühholoogilist pinget, mis on mõõdetav lähtudes kuulaja täituvatest või täitmatutest ootustest (*preference rules*) (Lerdahl 2001: 142–143).

Käesoleva artikli seisukohalt on oluline ka Martin Kirschbaumi (2001) käsitlus. Mainitud autori tähtsus seisneb selles, et ta muusikalise kõrgpunkti (kulminatsiooni) mõistet defineerides käsitleb viimast süstemaatiliselt.[5] Kuigi seda otsesõnu nimetamata, keskendub Kirschbaum sisuliselt muusikalise pinge loomise ja analüüsimise käsitlemisele ning näitab, missuguste struktuuriaspektidega selle loomisel või analüüsimisel tuleb arvestada. Tema arvates on olulised järgmised aspektid: dünaamika, helikõrgus, rütm, tempo, harmoonia, faktuur, instrumentatsiooni ja mänguvõtetega seotud efektid ning muusikaliste sündmuste kontrastsus. Siinkohal on oluline ka see, et Kirschbaum keskendub ainult 20. sajandi muusikale.[6] Tema kognitiivsel lähenemisel ei ole päris selge, kuidas ta saavutab kvantitatiivseid näitajaid. Seetõttu huvitab mind ka see, kas ja kuidas on võimalik autori *quasi*-empiirilisele meetodile leida n.-ö. päris-empiirilist rakendust.

Ka Jere T. Hutcheson (1972) kasutab vastandpaari pingetõus–pingelangus, käsitlemaks muusikalist pinget. Oma vormianalüüsi õpikus kirjutab ta, et muusikalise pinge suurenemine jätab tavaliselt

---

[5] Muusikalise pingega otseselt seotud on kõrgpunkti mõiste, mille puhul on kasutusel mitmed kas sünonüümsed või pisut teise rõhuasetusega mõisted: *Höhepunkt*, *climax*, *culmination point*, *highpoint*. Mõistet on seostatud muusikalise vormi ja vahel ka struktuuri erinevate aspektidega, kuid sageli pole seda tehtud kuigi süstemaatiliselt. Näiteks tonaalses muusikas võib harmoonilise pinge kasvu seostada liikumisega toonikalt dominandile (toonikalt ära) ja pinge langust liikumisega dominandilt toonikasse. Dominanti võib seega pidada tonaalse muusika omamoodi harmooniliseks kõrgpunktiks. Seega võib harmoonilise pinge kasvust ja langusest rääkida juba tonaalse teose süvatasandil, sest Schenkeri *Ursatz*'ist lähtudes võib öelda, et tonaalse teose harmoonilise tagaplaani moodustabki harmooniajärgnevus I-V-I. 19. sajandi muusikat on kõrgpunkti seisukohalt uurinud Kofi Agawu (1982). Meloodia kõrgpunkte ja nende seostamist esitusega on uurinud Zohar Eitan (1997).

[6] Pärast Teist maailmasõda on mõningad uurijad lähenenud pingele nüüdismuusikas fenomenoloogiliselt, nende hulgas Erhard Karkoschka (1974, *Klangspannungsanalyse*), kes käsitleb selliseid aspekte nagu intervallipinge (*Intervallspannung*), akordipinge (*Klangspannung*), impulsitihedus (*Impulsdichte*) ja kõlaruum (*durchschrittener Raum*) ja kujutab neid ka graafiliselt. Schenkeri analüüsist pärit prolongatsiooni mõistet on varase post-tonaalse muusika analüüsimisel hiljuti vaadelnud koos psühhoakustiliste aspektidega ka Olli Väisälä (2004).

mulje aktiivsuse kasvust, liikumisest n.-ö. (lokaalse) kulminatsioonipunkti poole, samas kui pingelangust tajutakse sellisest punktist eemaldumisena. Hutchesoni sõnul võib pinge väljenduda meloodia liikumisena (lokaalse) kõrgpunkti poole, meloodia helikõrgusliku aktiivsuse kasvus (intervalliline pinge), kõrgete registrite kasutuses (instrumentatsioon), rütmi aktiivsuse, harmoonia aktiivsuse, dünaamika (intensiivsuse) ja faktuuri tiheduse tõusus, kusjuures eelnimetatud aspektid toimivad sageli üksteisega interaktsioonis ning nende tähtsus muusikalise pinge kujundamisel muutub pidevalt. Liikumist kulminatsioonipunktide poole ja nendest eemale (mis omakorda väljendub muusikalise pinge kasvu või kahanemisena) peab Hutcheson üheks olulisemaks aspektiks vormikujundusel (vt. Hutcheson 1972: 11).

„Oxymoroni" vormist

Erkki-Sven Tüüri teos „Oxymoron" (2003) suurele ansamblile tähistab murdepunkti tema loomingulises arengus. Kui seni põhines helilooja muusika erinevatel, kuid võrdlemisi stabiilsetel seisunditel (nt tonaalsus ja atonaalsus, regulaarne ja ebaregulaarne rütm, kontemplatsioon ja plahvatuslikud pursked), mida helilooja kas vastandas või ühelt seisundilt teisele liikudes järk-järgult teisendas, siis alates sajandivahetusest on tema peamine huvi suunatud eelkõige pidevate muusikaliste üleminekunähtuste loomisele (Tüür 2007: 4, 6).

Erkki-Sven Tüüri loomingut analüüsides tuleb lähtuda muusikalise materjali ajalisest, süntaktilisest, laadilisest ja tämbrilis-faktuurilisest organiseerimisest.[7] Käesoleva artikli seisukohast oli helilooja teoste vormianalüüsi lähtepunktina oluline määratleda muusikalise materjali ajalisi organiseerimisprintsiipe (nn. perioodiline ja mitteperioodiline aeg, vt. Kotta 2008: 23–38). Siiski ei valitse ajalises plaanis Tüüri hilisemas loomingus mitte üks organiseerimisprintsiip, vaid toimub pidev üleminek ühelt printsiibilt teisele (Kotta 2008: 37), seega esinevad vormilõigu aluseks olevad ajalised organiseerimisprintsiibid siin enamasti sünteesitult. See põhjustab omakorda selgelt eristatavate vormiosade hägustumise, mistõttu eelnimetatud käsitlust ei saanud siin otse üle võtta.

Ka „Oxymoroni" muusikat iseloomustab pidev üleminek ühelt ajaliselt organiseerimisprintsiibilt teisele, mille tulemuseks on erinevate faktuuritüüpide segunemine. Ühemõtteliste liigenduskohtade puudumise tõttu on teost väga raske vormiliselt lõikudeks jagada. Traditsioonilises vormikäsitluses lähtutakse üldjuhul sellest, et mingi vaadeldav vormiüksus on selgelt piiritletud. See eeldab üheselt määratletavate algus- ja lõpp-punktide olemasolu, nii et igal hetkel on selge, missugune vormiosa kõlab (vormiosade võimalik haakumine tonaalses muusikas seda põhimõtet veel ei tühista). „Oxymoroni" puhul on aga selliseid selgeid piirjooni erinevate võimalike vormiüksuste vahel väga vähe. Arvatavasti saabki selle teose puhul rääkida vaid ühest selgemast vormilisest liigenduskohast umbkaudu teose keskpaigas (takt 218). Seetõttu saab teose jagada üldiselt kaheks osaks, kuid vormi edasiseks liigendamiseks oleks vaja leida mingi teine alus.

„Oxymoroni" kõige väljapaistvamaks omaduseks näib olevat moodustumine lühikestest, võrdlemisi kontrastsetest muusikalistest sündmustest ja kulminatsioonidest. Need sündmused ja kulminatsioonid moodustuvad motiivilises ja faktuurilises plaanis mitmel erineval viisil (millel lähemalt peatutakse allpool). Samas leidub teoses vähe selgelt artikuleeritud tsesuure. Seetõttu lähtutakse teose vormi edasisel analüüsimisel just nimetatud muusikalistest sündmustest ja kulminatsioonidest.

Kõnealuses teoses võib mingi muusikaline sündmus olla esile tõstetud, kusjuures see esiletõst võib olla eelneva muusikalise arengu poolt ette valmistatud, aga ei pruugi. Esimest liiki sündmusi tähistan ma mõistega „kulminatsioon" ja teist mõistega „impulss". Kui impulsi või kulminatsiooniga kaasneb muusikalise arengu loogikas mingi suuremat laadi ja pikemaajalisem muudatus, siis olen seda tähistanud mõistega „kontrast". Kontrast ei pruugi seega ilmneda vahetult, vaid pigem analüüsija retrospektiivse järeldusena. Kontrasti, mis langeb omakorda kokku kulminatsiooniga, olen tähistanud mõistega „kontrastkulminatsioon".[8] Järgnevalt esitan nende nelja mõiste definitsioonid. Iga mõiste järel on nurksulgudes toodud defineeritava mõiste

---

[7] Sellistest aspektidest lähtuvalt olen eesti uut muusikat, sh. Tüüri loomingut, käsitlenud näiteks oma artiklites (vt. Lock 2005: 109–112; 2008: 83–93).

[8] Tänan Mart Humalat oluliste ettepanekute ja märkuste, eelkõige aga soovituse eest võtta kasutusele kontrastkulminatsiooni mõiste.

lühend. Neid kasutatakse lisas asuvates joonistes ja edaspidi tekstis.

> **Impulss [im]** – Impulsiks nimetan ma mingis vaadeldavas kontekstis suhteliselt väljapaistvat lühikest ja lokaalset muusikalist sündmust, mis eristub teistest sündmustest mingil viisil (tämbriliselt, helikõrguslikult, dünaamiliselt). Selleks võib olla üksikheli (eriti teose või vormiosa alguses, kui valitseb suhteline vaikus), löökpillide esitatav(ad) üldisest taustast selgelt eristuv(ad) heli(d), iseloomulik motiiv, mis on dünaamiliste ja registriliste vahenditega esile tõstetud, (*tutti*-)akord jne.
>
> **Kulminatsioon [ku]** – Kulminatsiooniks nimetan ma mingi fraasi dünaamiliselt, registriliselt ja/või rütmiliselt[9] ettevalmistatud kõrgpunkti ehk nn. kõrgala. Seega pole kulminatsiooni siin mõistetud traditsiooniliselt teose või teoseosa sisulise kõrgpunktina (seda tähistatakse siin mõistega „kontrastkulminatsioon"; vt. allpool), mis sellisena leiab aset vaid üks kord, vaid lokaalse ja sagedasti korduva nähtusena.
>
> **Kontrast [ko]** – Kontrastiks nimetan ma muusika arengukäigus asetleidvat suuremat ja põhimõttelisemat laadi muudatust, eelkõige püsivat faktuurimuutust. Kontrastikohana (s.t. kontrasti alguspunktina) käsitan ma muusikateoses sellist impulssi [im], millele järgneb eelmainitud suurem muutus. Seega eeldab kontrast laiemat kui ainult lokaalset konteksti.
>
> **Kontrastkulminatsioon [kk]** – kontrastkulminatsiooniks nimetan ma kontrasti [ko] ja kulminatsiooni [ku] kokkulangemist, millega kaasneb püsiv faktuurimuutus.

Mõisted on järjestatud nende suhtelise kaalukuse järgi: impulss ja kulminatsioon esindavad muusika vormi seisukohalt võrdlemisi lokaalseid sündmusi, kontrast ja kontrastkulminatsioon aga laiemas kontekstis olulisi sündmusi.[10] Neid mõisteid rakendades peaks olema võimalik teose vormi edasine liigendamine. Joonisel 1 on teose teises pooles (taktid 218–421) sisalduvad impulsid, kulminatsioonid, kontrastid ja kontrastkulminatsioonid tähistatud tumedamate ja erineva jämedusega vertikaaljoontega. Impulsse tähistavad kõige madalamad vertikaaljooned (im, kõrgus 10 ühikut), kulminatsioone sellest mõnevõrra pikemad (ku, kõrgus 20 ühikut), kontraste veel pikemad (ko, kõrgus 30 ühikut) ja kontrastkulminatsioone kõige pikemad vertikaaljooned (kk, kõrgused 35 ja 40 ühikut).[11] Vertikaaljoonte pikkus viitab mainitud vormikujundavate muusikaliste sündmuste suhtelisele tähtsusele: impulss kui väljapaistev, aga suhteliselt lühiajaline sündmus kuulub antud gradatsioonis võrreldes kulminatsiooni kui suhteliselt pikemaajalise sündmusega madalamale tasandile; kontrast kui impulss ja/või kulminatsioon, millega kaasneb muusikalise arengukäigu muutus, kuulub kõrgemale ja kontrastkulminatsioon kui kontrasti ja kulminatsiooni ühendav nähtus kõige kõrgemale tasandile. Joonisel 1 on vormianalüüsi tulpadele lisatud teose teise osa kestusest lähtuv silutud ja standardiseeritud aeg,[12] mis tähistab mingile vormiliigenduskohale (ko ja kk) vastavat „ajaakent" (*time window*). Lisaks sellele on antud ka vastava vormiliigenduskoha taktinumber partituuris (mitme takti ulatuse puhul on orientiiriks antud vaid esimese takti number). Joonis 1 sisaldab samuti tajukatse tulemusi, seitsme tajukatse osaleja keskmise kõvera pikkade hallide vertikaaljoonte näol, mida käsitletakse hiljem lähemalt.[13]

Vastavalt joonisele 1 leidub teose teises pooles seitse kohta, mida võib käsitada kontrastidena (taktid 218 [00:05],[14] 257 [02:05], 288 [03:40], 333 [05:10], 391 [06:45], 401 [07:15], 415 [08:05]) ja mis edaspidi markeerivad teose seitsme vormilõigu „piire". Esimene kontrast (takt 218) on seotud teose teise poole algusega. Takt 218 on kontrastne sellele eelnevate taktidega, kusjuures

---

[9] Rütm võib kas kiireneda või aeglustuda (erinevates faktuurikihtides).

[10] Siin ei ole püütud hierarhiseerida muusikalisi sündmusi nende suhtelise pingelisuse põhjal.

[11] Kontrastkulminatsioonidele on antud kaks erinevat arvu, sest vormikujundavad muusikalised sündmused võivad langeda ajaliselt kas ühte või järgneda teineteisele suhteliselt lühikese aja (nt. paari takti) jooksul, moodustades mõtteliselt siiski koos vastava kontrastkulminatsiooni.

[12] Kuulamisanalüüsi aluseks on „Oxymoroni" NYYD-ansambli salvestus Olari Eltsi juhatusel (Tüür 2007).

[13] Vormianalüüsi tulpade ja tajukatseanalüüsi keskmise kõvera graafiline kujutamine ühes ja samas joonises nõudis selle realiseerimist samadel tehnilistel alustel. Seetõttu koosneb nimetatud keskmise kõver antud joonises samuti tulpadest.

[14] Edaspidi on joonistel näidatud tajukatse aluseks olnud teose teise poole kestus (Tüür 2007), mis algab sisuliselt nullist, kuid ajatelje ühtlustamise tulemusena algavad tajukatse osalejate reaktsioonid ajaaknas 00:05.

kontrast avaldub siin rütmilise ja faktuurilisena. Teine kontrast (taktid 257–264) on esmapilgul kõige märkamatum, sest kontrasti loovad muusikalise struktuuri aspektid on hajutatud kaheksa takti peale. Taktides 257–261 kõlavad lokaalsed muusikalised sündmused (impulsid ja kulminatsioonid), mis viivad muusikalise materjali põhimõttelisemat laadi muutusteni alles taktis 264. Põhjus, miks kontrasti alguspunktiks on valitud takt 257 ja mitte takt 261 (milles kõlab tegelikult viimane impulss enne põhimõttelisi muudatusi muusikalises faktuuris), on eelkõige tämbriline: teose teise poole alguses kõlavad kõigepealt keelpillid (taktid 218–227), millele nendega osalt kattudes sekundeerivad puupillid (taktid 226–229), ning nii toimub see mitu korda: vastavalt taktides 229–238 (keelpillid) ja 236–239 (puupillid), 239–247 (keelpillid) ja 246–251 (puupillid). Takt 257 artikuleerib keelpillitämbri taaskordset kõlamist ja seetõttu on järgneva vormiosa algust põhjust lugeda just sellest taktist.

Kolmas kontrast (takt 288) on teose teise poole esimene kontrastkulminatsioon. Selle olemasolu saab analüütiliselt järeldada seetõttu, et kolme takti vältel (seega suhteliselt lühikese aja jooksul) järgnevad teineteisele kulminatsioon (puupillid taktis 286 ja keelpillid taktis 287) ja kontrast (taktis 288 puupillid, löökpillid ja keelpillid).[15] Neljas kontrast on taas ambivalentne: basskitarri sisseastumisega asetleidva tämbrikontrastiga kaasnevad ühtlasi järgnevates taktides ka faktuurimuutused (taktid 333–343), kuid hiljemalt taktis 345 taastub üldjoontes mainitud kontrastile eelnevaid takte iseloomustav karakter. Viies kontrast on seotud teose ühe olulisema peakulminatsiooniga (kõige olulisem faktuurimuutus taktis 391), kus ainsana ühtivad kõik vormikujundavad muusikalised sündmused (impulss, kulminatsioon, kontrast) taktis 391 rõhulisel (esimesel) taktiosal, moodustades seega kõige selgemalt artikuleeritud kontrastkulminatsiooni (kk).[16] See kõrgpunkt on ka pöördelise tähtsusega, sest ühe hetkega leiab aset kõige olulisem faktuurimuutus, mille tagajärjel hakkavad valitsema kõlavälja loovad kolmekümnekahendikes arpedžofiguurid.

Kuues kontrast (takt 401) on taas mõnevõrra ambivalentne ning selgub alles taktides 408–410, kus põhimõtteline faktuurimuutus on saanud ilmseks. Nimetatud kontrast on samaaegselt kogu teose seisukohast kolmas (ja viimane), kuid võrreldes eelmisega taas nõrgema mõjuga kontrastkulminatsioon, sest impulss, kulminatsioon ja kontrast leiavad aset küll puu- ja vaskpillidel (takt 401), kuid keelpillidel jääb eelmisel kõrgpunktil alguse saanud arpedžotaoline kõlaväli veel mitu takti kõlama, enne kui see muutub alates taktist 408 järk-järgult reljeefsemaks ja toimub väljakomponeeritud tempoaeglustus. Seitsmenda kontrasti (takt 415) tulemusena luuakse keelpillide ja poognaga mängitava väga väikese taldriku (*crotales*, *cymbales antiques*) abil enne teose lõppu täiesti uus kõlamaailm. Lähtuvalt mainitud punktidest saab teose teise poole vormi liigendada seega seitsmeks (suhteliselt) suuremaks lõiguks.

Tajukatsed

Tajukatsed viidi läbi 17.–18. ning 25.–26. aprillil 2008 Eesti Muusika- ja Teatriakadeemia muusikateaduse osakonna ruumides. Tarvilik tehniline varustus pärines Eesti Muusika- ja Teatriakadeemia elektronmuusika stuudiost.[17] Katses osales seitse isikut (N=7), kellest neli olid naised (n) ja kolm mehed (m). Katses osalenute keskmine vanus oli 31,7 aastat. Kõigil katses osalenutel oli erinev muusikaline taust: neist neli olid nn. ekspert- ja kolm nn. mitte-ekspert-kuulajad. Iga katses osaleja on edaspidi tähistatud suure tähega: A – ekspert (m – muusikateoreetik ja helilooja), B – mitte-ekspert (n – kunstnik), C – mitte-ekspert (n – üliõpilane, amatöör), D – ekspert (n – helilooja ja interpreet), E – mitte-ekspert (m – amatöör), F – ekspert (m – dirigent, laulja), G – ekspert (n – interpreet).[18]

---

[15] Sellele on näites 1 antud number vertikaalteljel 35 (ja mitte maksimumi 40), sest vormi kujundavad muusikalised sündmused ei lange ühte (vrd. märkus 11).

[16] Sellele on näites 1 seetõttu ainsana antud vertikaalteljel maksimumväärtus 40.

[17] Võlgnen tänu Maris Valk-Falkile, kes osutas suurt abi tajukatsete läbiviimisel ning katsetulemuste vormistamisel Conference on Interdisciplinary Musicology 2008 (CIM08 Thessaloniki 1.–7.7.2008) peetud ettekandeks. Samuti tänan Hans-Gunter Locki tehnilise varustuse kasutamise võimaldamise ning vajaliku tarkvara loomise eest. Seda tarkvara on rakendatud ka kahe eelneva uurimuse vältel (Lock, Valk Falk 2008 ja 2009).

[18] Katseisikud on sulgudes olevate mõistetega kirjeldatud nende enda määratluse järgi. Nii jääb C siin üliõpilaseks, ehkki tema kohta võiks täpsustuseks öelda, et tajukatsete ajal oli ta lõpetamas eesti filoloogiat ja alustamas kompositsiooniõpinguid. E elukutse on kaugel muusikast ja ta on tegelnud muusikaga vaid kooris lauldes.

Tajukatse läbiviimise protseduur

Katse eesmärgiks oli testida ja registreerida seda, kuidas tajutakse muusikalist pinget reaalajas kulgeva muusika puhul. Katse käigus paluti katseisikutel pinget n.-ö. joonistada, s.t. kasutada selleks slaiderit (liugurit): tajudes muusikalise pinge tõusu, pidi katseisik slaiderit ülespoole ning pinge langemisel allapoole liigutama. Eksperiment algas tervikteose kuulamisega (19:40 minutit), mille raames paluti osalejaid tähele panna erinevaid pinget kujundavaid aspekte ja vahetult enne katsealgust kolmeminutilise antud lõigu kõige valjema koha eelkuulamisega. Arvesse võetud katses eneses kuulati ainult teose teist poolt (8:34 minutit).

Katse läbiviimiseks, s.t. andmete numbriliseks salvestamiseks reaalajas, loodi tarkvara graafilises programmeerimiskeskkonnas Max/MSP (Hans-Gunter Lock), mis salvestas iga poole sekundi tagant tajukatse isiku slaideri liigutused, s.t. vastavat punkti, mida graafiliselt tõlgendati hiljem intensiivsustasemena ulatusega 1–127. Riistvarana kasutati *Midi-keyboard-controller*it (Roland) ja *USB Midi-fader slider controller*it (Behringer).

Peale selle täitsid osalejad ankeedi standardandmetega osaleja kohta (vanus/sünniaasta, sugu, muusikaline taust, tegevusala). Ankeet sisaldas ka lisaküsimusi: kas kuulatud teose helikeel oli tuttav, osaliselt tuttav või võõras ja kas katsete sooritamine tundus lihtne, raskendatud või võimatu. Katses osalejad pidid vastama veel küsimusele, millest pinge nende arvates muusikas kõige enam sõltus. Iga küsimuse juurde oli osalejatel võimalik teha täiendavaid tähelepanekuid ja märkusi.

Enne katse läbiviimist instrueeriti katseisikuid slaideri kasutamise osas (pinge tõusmisel liigatada slaiderit üles, selle alanemisel liigutada slaiderit alla; pinge järsule tõusule või langusele reageerida slaideri kiirema liigutamisega, pinge püsimisel ühel tasandil slaiderit mitte liigutada; arvestada slaideri liikumise ulatusega, s.t. mitte liikuda slaideriga lõpuni enne tunnet, et saabunud on teose peamine kulminatsioon).

Empiiriliste andmete analüüs ja võrdlus vormianalüüsiga

Tajukatse graafide kirjeldamine

Saavutamaks seitsme küllaltki erineva reaktsiooniga katseisiku tulemuste puhul standardiseeritud võrdlusalust (mis võimaldavad üldisemaid järeldusi), on käesoleva artikli jaoks vähendatud andmehulk printsiibil, mis seisneb horisontaal- (s.o. aja-)telje silumises kümnesekundilisteks ajaakendeks ja vertikaaltelje ehk intensiivsustaseme normaliseerimises.[19] Tajukatsete tulemused on esitatud joonisel 2, millel on kujutatud kõigi seitsme osaleja (A–G) normaliseeritud ja ühtlustatud kõverad ja nende aritmeetiline keskmine, mis illustreerib nende üldistatud reaktsiooni ühele ja samale teosele. Mainitud kõverate all on eraldi välja toodud ka veel standardhälbe kõver.[20] Standardhälbe madal väärtus näitab osalejate suhteliselt samalaadset, kõrge väärtus aga suhteliselt erinevat käitumist vastaval ajahetkel. Kõik ühtlustatud[21] andmed on töödeldud ning graafiliselt visualiseeritud tabelarvutusprogrammis MS Office Excel.

Kuigi lokaalses plaanis võivad kõigi seitsme osaleja graafid individuaalselt tugevasti erineda, on neil üldisemas plaanis palju ühist. Nendes kõigis on nähtav kõvera järkjärguline tõus (tõusev trend), kusjuures kõver saavutab enamiku katseisikute puhul oma kõrgpunkti (mille maksimum on normaliseeritud intensiivsusteljel 100) ajavahemikul 06:35–07:15 (välja arvatud kõver D, mis saavutab maksimumi ajahetkel 03:15). Kõikide kõverate puhul on alates 7:25 minutist nähtav võrdlemisi tugev langus, millele enamikus kõverates järgneb mõningane tõus (tipud ajahetkedel 08:15 ja 08:25) (välja arvatud kõver D). Suurem erinevus on kõverates just kuulatud muusikakatkendi esimeses pooles (00:05–05:35), kus kõverate erinevused on kõige suuremad (siiski ühendab kõiki kõveraid tõusev trend). Seda tõusvat trendi illustreerib selgelt keskmist tähistav kõver, mis siin on parema eristuse nimel kujutatud jämedamalt. Silmnähtavalt kalduvad aga trendist kõrvale kõver E ajahetkel 01:05 (lühiajalise üsna kõrge intensiivsustasemega 85), kõver G ajahetkel 02:25 (lühiajalise suhteliselt madala

---

[19] Selle eksperimendi toorandmete põhjal joonistatud (mitte ühtlustatud) kõverad on ära toodud artiklis Lock, Valk-Falk 2008.

[20] Standardhälve on statistilise leviku või varieeruvuse mõõde.

[21] Kõverate intensiivsustaseme normaliseeriseks kasutati aga automatiseerimise eesmärgil taas Hans-Gunter Locki loodud Max/MSP programmi.

intensiivsustasemega 30), kõver B ajahetkel 03:45 (lühiajalise suhteliselt madala intensiivsustasemega 40) ning eelkõige kõver G ajahetkel 04:45 (ootamatult madala intensiivsustasemega 40). Väljapaistvalt madalale langevad vahetult pärast põhikulminatsiooniala (06:35–07:15) kõver G (07:45, 7) ning kõver A (07:55, 4.5). Kõver D seevastu kulgeb sujuvamalt ja erinevalt teistest ei moodusta ajapunktides 08:15 ja 08:25 uut pinge kasvukõverat. Keskmiste näitajate kõveras kajastub siiski vormianalüüsi viimase kontrasti olemasolu, sest valdav enamik kuulajatest reageeris eelmainitud viimastele kulminatsioonidele üsna suure tõusuga (vt. ka standardhälbe kõverat).

Ainult kaks osalejat langetas (nagu eelnevalt kirjeldatud) oma pingenivoo enne viimast kontrasti juba väga madalale. Osalejate A (ekspert) ja huvitaval kombel eelkõige C (mitte-ekspert) joonistatud kõverad on matemaatiliselt tuletatud keskmisele kõverale kõige lähedasemad.

Kokkuvõtvalt võib toodud näite põhjal öelda, et katseisikud olid muusikalise pinge kulgemise suhtes kõige sarnasematel seisukohtadel ajapunktides 05:35–07:05, s.t. vahetult kulminatsioonile eelneva lõigu suhtes. Põhjus võib peituda selles, et katseisikud olid selleks hetkeks muusikalise kontekstiga juba suhteliselt tuttavad, muusikaliste sündmuste kulg siin küll tiheneb, kuid sündmustena on need juba kõik suhteliselt tuttavad (arendus põhineb juba kuuldud motiividel ja faktuuritüüpidel), lokaalsed kulminatsioonid järgnevad üksteisele vahetult ning dünaamiline nivoo on küllaltki kõrge.

Tajukatsete keskmise kõvera võrdlemine vormianalüüsi diagrammiga

Nagu eespool öeldud, on joonisel 1 kujutatud tajukatse osaliste keskmine kõver (tähistatud kõige kõrgemate ja heledamate vertikaaljoontega) kõrvutatuna vormianalüüsis väljatoodud impulsside (im), kulminatsioonide (ku), kontrastide (ko) ja kontrastkulminatsioonidega (kk) tumedamate vertikaaljoontega. Vormianalüüsi tulpade puhul on kõrgemad tulbad omakorda ka jämedamad, mis peaks tulpadega tähistatud muusikaliste sündmuste omavahelise hierarhia visuaalselt paremini esile tooma. Antud graafil on ajaaknad võrrelduna tajugraafiga kaks korda väiksemad (viie sekundi pikkused), demonstreerimaks paremini vormianalüüsis välja toodud väikseimaid liigendusüksusi, impulsse, mida esines küllaltki arvukalt. Vormi liigendamise kujutamine tulpadena on põhjendatud sellega, et mingit lineaarset seost erinevate impulsside, kulminatsioonide, kontrastide ega kontrastkulminatsioonide vahel antud juhul ei analüüsitud. Tehnilisel põhjusel sai seetõttu ka tajukatsete kõverat kujutada ühes ja samas graafis samuti vaid tulpadena (vrd. joonealune märkus 13).

Joonisel 1 esitatud kahe diagrammi – vormianalüüsi ja tajukatse keskmise oma – võrdlemisel lähtusin esmalt vormianalüüsi tulpadest. Kohe esmapilgul on ilmne, et kõik vormi seisukohast olulisemad sündmused (kontrast või kontrastkulminatsioon) kajastuvad ka tajudiagrammis; erandiks on vaid esimene kontrast (00:00), mida katseisikud ei saanudki kontrastina tajuda, sest see kujutas endast nende jaoks lähtepunkti (puudus eelnev muusika, s.o. kontekst, mille põhjal võrrelda). Kontrastid ja kontrastkulminatsioonid ning nendega kaasnevad tajukõverate keskmise kõvera harjad artikuleerivad horisontaalteljel järgmisi ajapunkte (mida siin tinglikult nimetan tajutulpadeks): 02:05, 03:40 (tajutulp 03:35), 05:10 (tajutulp 05:05), 06:45 ja mõnevõrra ebamäärasemalt 07:15 (tajutulp aga juba 07:05). Viimane on põhjustatud sellest, et vormianalüüsis määratleti punkt selles kohas, kus faktuur oli täiesti muutunud, kuulaja hakkas aga selle muutumist tajuma ilmselt juba varem ja ka reageeris sellele vastavalt. Ka kontrast, mis tekib ajapunktis 08:05, on tajudiagrammi keskmise kõveraga teatavas nihkes. Põhjus on ilmselt selles, et kuulaja saab asetleidnud muutustest teadlikuks alles pisut hiljem.

Kommenteerimist vajaksid ilmselt ka teose lõputaktid (taktid 415–421). Keelpillide foonil näivad poognaga taldrikutel mängitud helid (väiksed kulminatsioonid) esmapilgul üsna lokaalsete sündmustena, sest neid tekitatakse ainult dünaamika muutmisega (ei muutu helikõrgus, faktuur jne., nagu oli iseloomulik teose eelnevatele arvukatele kulminatsioonidele). Samas võib aga öelda, et nimetatud taldrikutel mängitavate helide lokaalsete kulminatsioonide tulemusel muutub muusikalise arengu iseloom olulisel määral. Samas on taldrikutel mängitav ka reljeefselt esil, sest keelpillide foon püsib dünaamiliselt *pp* ja *ppp* vahel ning tempo (algselt veerandnoot = 110) on muutunud oluliselt aeglasemaks (veerandnoot = 50). See kõik muudab mainitud muusikalised sündmused laiemas plaanis üsna oluliseks ja seetõttu on antud koht analüüsitud ka kontrastina (vt. vormianalüüsi peatükki). Tähelepanuväärne on ka see, et enamik katsealuseid pidas nimetatud takte muusikalise pinge kujunemisel oluliseks.

Lähtudes aga tajugraafi tulpadest, on näha, et see sisaldab ka mõningaid kõvera harjasid, millele ei vasta vormianalüüsis n.-ö. kõrgema tasandi sündmus – kontrast või kontrastkulminatsioon (vt. eelkõige ajapunkti 01:05, mille puhul see on kõige väljapaistvam). Siin võib olla põhjuseks analüüsitud muusikaliste sündmuste (impulsid ja kulminatsioonid) suhteline hõredus, mille tõttu need tõusevad üksiksündmustena ilmselt enam esile. Sündmuste hõredust rõhutab ka veel suhteliselt madal dünaamiline nivoo (*pp*). Erinevad kuulajad tajusid seda kohta muusikas võrdlemisi erinevalt. Üks kuulajatest (E – mitte-ekspert), intensiivsustasemega 85) fikseerib siin eriliselt kõrge pingenivoo, mida ta seejärel sama järsult langetab. Selles võib leida kinnitust ka Lerdahli väide naiivse kuulaja kohta, kes jälgib muusikalisi sündmusi nende ajalises kulgemises ja otsustab muusikaliste sündmuste olulisuse üle n.-ö. lokaalses plaanis, nende (vahetu) kontrastsuse järgi.

Põhijäreldused

Käesoleva artikli eesmärgiks oli kirjeldada uut, pingedisainil (*tension design*) põhinevat analüüsimeetodit, mis võimaldaks muusikalisest pingest lähtudes analüüsida post- ja vabatonaalset muusikat. Selleks uuriti Erkki Sven-Tüüri ansambliteost „Oxymoron" nii muusikateoreetiliste kui ka empiiriliste vahenditega. Mõlema meetodi rakendamisel saadud ja ühtlustatud andmete kokkuviimine ning visualiseerimine näitas muusikateoreetilise ja empiirilise meetodi kombineerimise otstarbekust.

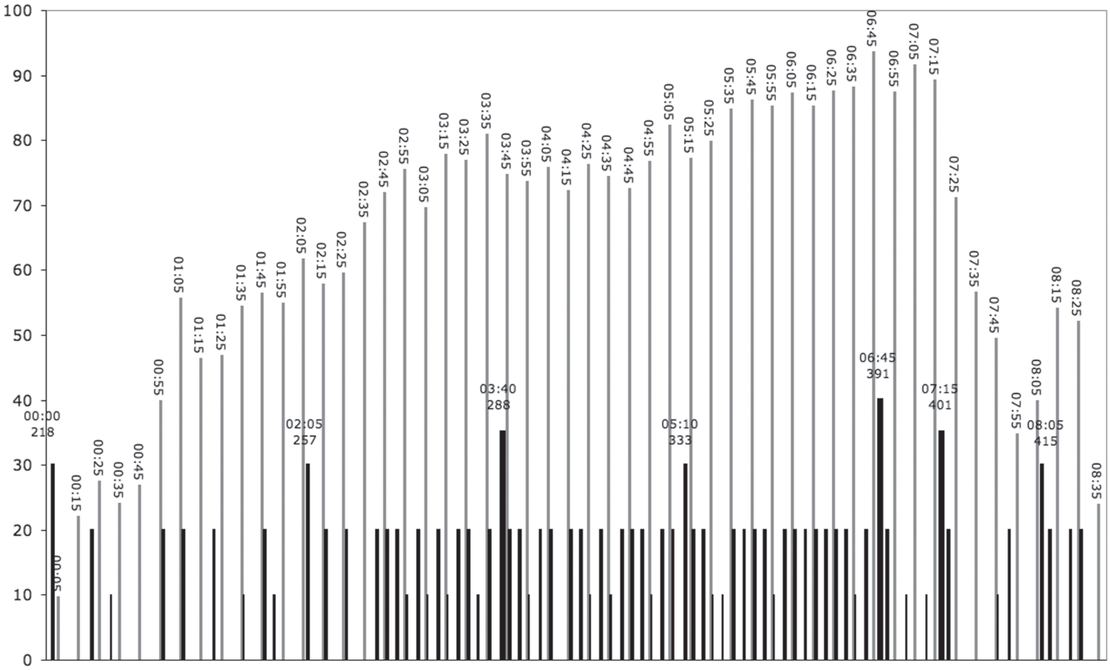
Kokkuvõttena võib välja tuua üldisemat laadi järeldused:

1) Kulminatsioonid mängivad Erkki-Sven Tüüri teoses „Oxymoron" võrdlemisi olulist rolli. Selgelt artikuleeritud tsesuuride puudumise tõttu olid just (lokaalsed) kulminatsioonid (ja nende kõrval ka kontrastsemat laadi lokaalsed muusikalised sündmused, mida käesolevas töös kirjeldati impulssidena) need muusikalised sündmused, mille põhjal oli võimalik teose vormi kirjeldada. Kulminatsioonide kombineerimisel teiste ülalkirjeldatud nähtustega (suurem muutus muusikalises arengus ehk kontrast) oli võimalik kulminatsioone omamoodi hierarhiseerida ja selle kaudu vormi selle eri tasanditel liigendada.
2) Võib üldiselt öelda, et tajukatsete kõverad korreleeruvad omavahel suures plaanis võrdlemisi hästi, s.t. kõigi puhul on täheldatav esmalt n.-ö. tõusev ning peale teose peakulminatsiooni n.-ö. langev trend. Lokaalses plaanis erinesid tajukõverad kohati aga suuresti ja seda eelkõige kõverate esimeses pooles, kus muusikalised sündmused olid olemuselt ambivalentsemad.
3) Vormianalüüsis saadud kõvera ja tajukatsete keskmise vahel eksisteeris selge seos: kõik vormianalüüsis saadud olulisemad punktid kajastusid ka tajukõverates. Siiski sisaldasid tajukõverate keskmised näitajad ka kõvera harjasid, kus vormianalüüsis olulisema tähtsusega sündmus puudus. See oli taas omane eelkõige neile teose kohtadele, mille tõlgendamine muusikalise pinge seisukohast oli ambivalentsem (muusikalise arengu suund või eesmärk ei olnud selge).

Seega võib meie uuringu andmeil väita, et muusikaline pinge ja selle muutumine sõltub posttonaalses muusikas võrdlemisi olulisel määral nähtustest, mida eelnevas tekstis kirjeldati kui impulssi, kulminatsiooni, kontrasti ja kontrastkulminatsiooni. Kuna osalejate arv oli võrdlemisi väike (puudus kvantitatiivsete meetodite mõistes statistiliselt kriitiline mass), ei pruugi tajugraafide keskmised näitajad veel ühtida n.-ö. tegeliku kuulamise pingekõveraga. Selle võrdlemisi suur kokkulangevus vormianalüüsil saadud graafiga on aga üsnagi paljulubav, jätkamaks uuringuid antud valdkonnas.

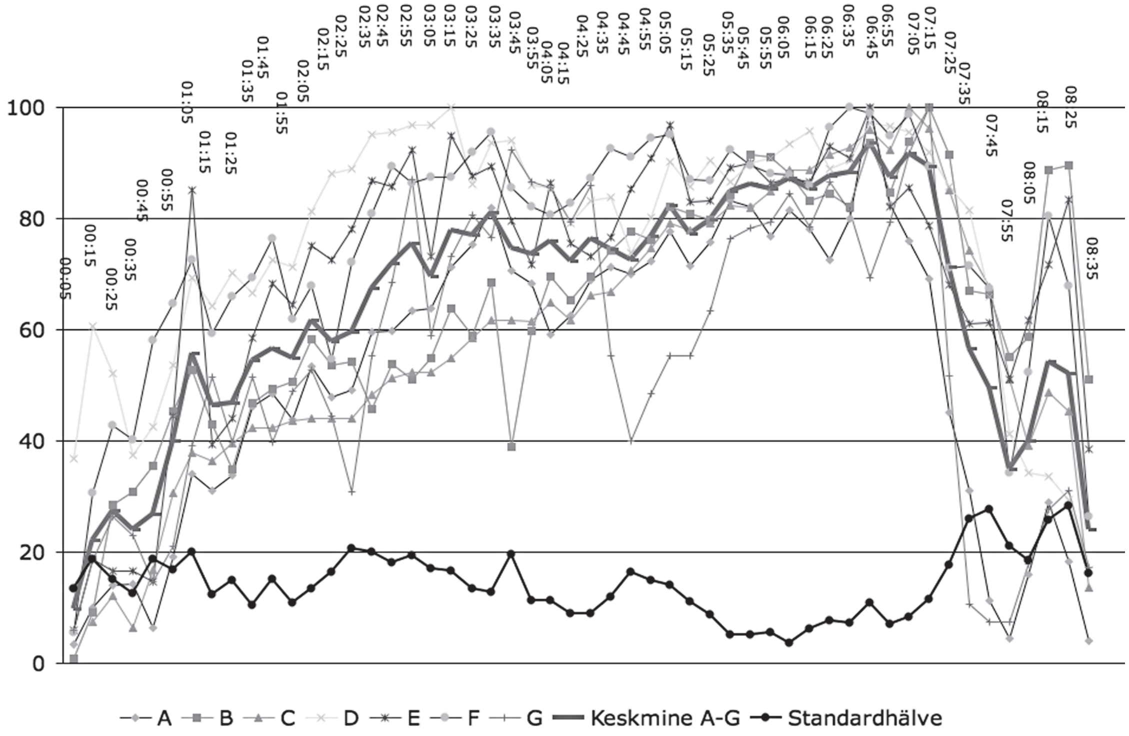
Artiklis kirjeldatud pingedisaini meetodit peaks nii muusikateoreetilises kui ka empiirilises osas saama ka edaspidi edukalt rakendada nii Erkki-Sven Tüüri kui ka sellele sarnase post- ja vabatonaalse nüüdismuusika analüüsimiseks.

LISA



JOONIS 1. „Oxymoroni" vormianalüüs kõrvutatuna tajukatse keskmisega.

Üldiselt kõige pikemad ja heledamad tulbad, mille tipus on ära toodud ajapunktid kümnesekundiliste vahedega, moodustavad tajukatse keskmise. Antud tulbad vastavad tajukatse keskmise kõverale (vt. joonis 2). Erineva pikkusega madalamad, tumedamad ja erineva jämedusega tulbad viitavad teose vormile. Kõige pikem jämedam tulp (väärtusega 40; ajapunkt 06:45; takt 391) ja sellest kaks pisut madalamat tulpa (väärtusega 35; ajapunktid 03:40 ja 07:15; taktid 288 ja 401) kujutavad kontrastkulminatsioone (kk). Tulbad väärtusega 30 (ajapunktid 00:00, 02:05, 05:10 ja 08:05; taktid 218, 257, 333 ja 415) kujutavad kontraste (ko). Tulbad väärtusega 20, mida antud joonises leidub kõige arvukamalt, kujutavad (lokaalseid) kulminatsioone (ku), ning tulbad väärtusega 10 impulsse (im). Kui tajukatsete esitamisel lähtuti kümnesekundilistest, siis vormianalüüsi puhul viiesekundilistest ajaakendest. See selgitab näiteks kulminatsiooni- või impulsitulpade suuremat arvu mõnes ajaaknas. Samuti on parema arusaadavuse huvides asetatud tajukatsele ja vormianalüüsile viitavad tulbad kõrvuti ja mitte kohakuti. Mõnikord võis viiesekundiline ajaaken sisaldada rohkem kui ühte impulssi või kulminatsiooni. Sellisel juhul valiti vastavat ajaakent n.-ö. esindama kõige suurema väärtusega muusikaline sündmus. Vormianalüüsi tulpade pikkus on määratletud tõusval kümnendikprintsiibil. Tajukatse keskmise tulpade vertikaalväärtus aga järgib kasutatud riist- ja tarkvara poolt määratud MIDI-standardi tugevuse skaalat 1–127.



JOONIS 2. Tüüri teose „Oxymoron" kuulamisel saadud kõigi seitsme tajukatse osalise (A–G) ühtlustatud ja normaliseeritud kõverad kõrvutatuna nende keskmise (jämedama) kõvera ja standardhälbe kõveraga. Kõverad A–G viitavad katses osalenud isikute muusikalise pinge suhtelise intensiivsuse tajule.

KIRJANDUS

Agawu, Kofi 1982. *The Structural Highpoint as Determinant of Form in Nineteeth Century Music* (PhD dissertation). Stanford: Stanford University.

Baroni, Mario 2003. The macroform in post-tonal music. Listening and analysis. – *Musicæ Scientiæ*, Vol. VII, 2, pp. 219–240.

Clarke, Eric F. 1989. Mind the gap: formal structures and psychological processes in music. – *Contemporary Music Review*, 3 (1), pp. 1–14.

Cook, Nicholas 1987. *A Guide to Music Analysis*. London: Dent.

Cook, Nicholas 1994. Perception: a perspective from music theory. – *Musical Perceptions*. Eds. Rita Aiello with John Sloboda, Oxford: Oxford University Press.

Deliège, Irene 1989. A perceptual approach to contemporary musical forms. – *Contemporary Music Review: Music and the Cognitive Science*: Volume 4, Proceedings from the „Symposium on Music and the Cognitive Science", 14–18 March 1988. Eds. Stephen McAdams, Irene Deliège, New York: Harwood Academic Publisher, pp. 213–230.

Deliège, Irene, Marc Melen, Diana Stammers, Ian Cross 1996. Musical schemata in real-time listening to a piece of music. – *Music Perception*, 14, pp. 117–160.

Deutsch, Diana, John A. Sloboda et al. 2007. Psychology of music. – *Grove Music Online* [accessed 17.06.2008] <www.grovemusic.com>.

Eitan, Zohar 1997. *Highpoints: A Study of Melodic Peaks*. Philadelphia: University of Pennsylvania Press.

Encyclopaedia Britannica Online. 22 Jun. 2008 <http://www.britannica.com/EBchecked/topic/587516/tension>.

Hindemith, Paul 1940. *Unterweisung im Tonsatz. 1, Theoretischer Teil.* Mainz: B. Schott's Söhne.

Hindemith, Paul 1939. *Unterweisung im Tonsatz. 2, Übungsbuch für den zweistimmigen Satz*. Mainz: B. Schott's Söhne.

Hutcheson, Jere T. 1972. *Musical Form and Analysis. A Programmed Course* (vol. 1). Boston: Allyn and Bacon.

Karkoschka, Erhard 1974. Neue Methoden der musikalischen Analyse und einige Anwendungen auf spätere Instrumentalwerke von Karl Marx (1967). – *Zur musikalischen Analyse.* Hrsg. Gerhard Schumacher, Darmstadt: Wissenschaftliche Buchgesellschaft, S. 644–664.

Kirschbaum, Martin 2001. *Höhepunktbildung und Dramaturgie in Neuer Musik. Die Erstellung „formaler Spannungsverläufe" als Ergänzung der musikalischen Analyse.* Köln: Dohr.

Kotta, Kerri 2008. Zur musikalischen Zeit: Über die formalen Grundlagen der Werke von Erkki-Sven Tüür – ein Beschreibungsversuch. – *Schwarze Milch und bunte Steine. Der Komponist Erkki-Sven Tüür.* Hg. Hans-Klaus Jungheinrich, Frankfurt, Mainz: Schott Music, S. 23–38.

Krumhansl, Carol L. 1985. Music psychology and music theory: Problems and prospects. – *Music Theory Spectrum*, 17, pp. 53–80.

Krumhansl Carol L. 1996. A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension and musical ideas. – *Music Perception*, 13, pp. 401–432.

Lerdahl, Fred, Ray Jackendoff 1983. *A Generative Theory of Tonal Music.* Cambridge, MA, London: MIT Press.

Lerdahl, Fred 2001. *Tonal Pitch Space*. New York: Oxford University Press.

Lock, Gerhard 2005. Muusikalise materjali käsitlusaspekte. – *Tekste Modernismist* [2003. ja 2004. aastal Pärnu Nüüdismuusika Päevade raames peetud konverentside „Modernism" ettekannete kogumik]. Toim. Maris Valk-Falk, Tallinn: Scripta Musicalia, lk. 108–137.

Lock, Gerhard 2008. Muusika ja arhitektuur Erkki-Sven Tüüri sarjas „Arhitektoonika". – *Tekste Modernismist II, Muusika ja Arhitektuur* (2005. ja 2006. a. Pärnu Nüüdismuusika Päevade konverentside ettekanded). Koost. Maris Valk-Falk ja Gerhard Lock, Tallinn: Scripta Musicalia, lk. 83–93.

Lock, Gerhard, Maris Valk-Falk 2008. Investigating structure of contemporary music applying tension design and empirical perception analysis. – *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08).* Eds. Costas Tsougras, Richard Parncutt, Thessaloniki: Aristoteles University [CD-ROM, on-line publication].

Lock, Gerhard, Maris Valk-Falk 2009. Tension design of contemporary orchestral music in different performances of „Sula" by Helena Tulve. – *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009).* Eds. Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, Päivi-Sisko Eerola, Jyväskylä: University of Jyväskylä [on-line publication].

Rothfarb, Lee 2002. Energetics. – *The Cambridge History of Western Music Theory*. Ed. Thomas Christensen, Cambridge, UK, New York: Cambridge University Press.

Thaler, Lotte 1984. *Organische Form in der Musiktheorie des 19. und beginnenden 20. Jahrhunderts.* München, Salzburg: Katzbichler.

Tüür, Erkki-Sven 2003. *Oxymoron (Music for Tirol) for large ensemble* (full score)*.* Frankfurt, Leipzig: Peters.

Tüür, Erkki-Sven 2007. *Oxymoron* (Nyyd-Ensemble, cond. Olari Elts). München: ECM Records (ECM New Series 1919, 476 5778)

Väisälä, Olli 2004. *Prolongation in Early Post-Tonal Music. Analytical Examples and Theoretical Principles.* Studia Musica 23, Helsinki: Sibelius Academy.

On factors influencing musical tension and its perception in Erkki-Sven Tüür's piece "Oxymoron"

The aim of this article is to explore what musical factors influence the perception of musical tension in Erkki-Sven Tüür's "Oxymoron" (2003). The form (as a time dependent structure) of "Oxymoron" was analysed and juxtaposed with graphs derived from perception tests using the same piece as audio stimulus. "Oxymoron" does not consists of a clearly defined form and it was impossible to apply traditional analytical methods of formal analysis. Instead, a set of new terms, "impulse", "culmination", "contrast" and "contrast-culmination", are applied; these are to be understood as significant features or complexes of features appearing during the development of the music over time.

These terms are developed by the author of this article and are ranged on the basis of their relative structural meaning: events called "impulse" take place at the structural low (i.e. local) level, but events called "contrast-culmination" belong to the highest (i.e. global) level. "Impulse" (at a local level) and "contrast" (at a global level) are considered as unprepared events. "Culmination" (local level) and "contrast-culmination" (global level), on the other hand, are so defined because they are in some way prepared events.

The introduced formal events are visualized with bar diagrams in which the structurally lowest event(s) are shown with low black columns and higher events with higher black columns (see figure 1). In perception tests, seven participants with different musical backgrounds were asked to "draw" the tension design of the piece based on their personal perception but focusing on the structural features of the music: move the slider (either slowly or suddenly) up, if tension increased, and down, if tension decreased (see figure 2). The comparison of the bar diagram of formal analysis with the smoothed and normalized graphs and their average graph shows a relatively strong correlation between the formal analysis and the perception results confirming the hypothesis of this article: on a global level and at important points of the formal development of the piece the musical tension is perceived relatively similarly.

A.2 Publication Ia (Lock & Kotta 2012, 5.2)

“Musical tension as a response to musical form” (in English)

Musical tension, musical form, salient features and data reduction – *Oxymoron* and 4th Symphony/Percussion Concerto *Magma*

Musical tension as a response to musical form[1]

Gerhard Lock,[*1] Kerri Kotta [#2]

[*] *Estonian Academy of Music and Theatre, Department of Musicology*

Tallinn University, Institute of Fine Arts, Department of Music, Tallinn/Estonia

[#] *Estonian Academy of Music and Theatre, Department of Musicology*

`[1]gerhard.lock@tlu.ee, [2]kerri.kotta@mail.ee`

ABSTRACT

Musical tension is a complex phenomenon and its comprehensive description should generally include a variety of different approaches. In this study, our goal is to describe the musical tension as a response of a listener to formal patterns by combining perception tests with musical analysis. We hypothesize that the intensity of the perceived musical tension is proportional to the structural (or hierarchical) significance of the corresponding musical event. To ease comparisons of the tension curves obtained from the listening tests and score-based structural analyses, we present the principles of three new analytical methods: 1) Analysis of salient features of music, 2) Analysis of musical "energy", and 3) Reduction and averaging of tension curves. We hope that the results applying those methods will contribute to a better understanding of the formal structure of post-tonal music and the techniques of prolongation.

I. INTRODUCTION

Musical tension is a complex phenomenon and its comprehensive description should generally include a variety of different approaches. In this study, our goal is to describe the musical tension as a response of a listener to formal patterns by combining perception tests with musical analysis.

To the authors of this article, musical form is essentially a hierarchical phenomenon. The main idea behind this study is that perceived musical tension can be seen as being dependant on the hierarchical aspects of form. We hypothesize that the intensity of the perceived musical tension is proportional to the structural (or hierarchical) significance of the corresponding musical event. To ease comparisons of the tension curves obtained from the listening tests and score-based structural analyses, we propose three new analytical methods.

The first method, **analysis of salient features of music**, is based on the discrimination of the relative importance of different types of compound musical events (i.e. impulse and culmination) using analysis of the musical score and cognitive analysis (see Lock 2010). The method and analysis results in Lock (2010) were presented at the ICMPC11 in Seattle (USA) August 23–28, 2010, under the title „Tension design analysis as listening analysis strategy for contemporary music: Perception of tension in contemporary post-tonal orchestral music: a case study".

In the second method, **analysis of musical "energy"**, musical form is treated as a succession of small areas in which the energy of the music determined by rhythm, dynamics, texture, or register, is described with simple terms such as increase, decrease, or sustain (Kotta 2011).

For the third method, **reduction and averaging of tension curves**, listening test curves are reduced to "deeper level" curves as long as they can be compared with the outputs of other types of analysis as described above. Unlike mathematical or Fourier-based smoothing of curves, this method allows for a clearer visualization and structural differentiation of peaks and valleys. Consequently, this method offers improved opportunities to study the perceived musical tension as a response to musical form. In what follows, we describe only the principles of the aforementioned methods, we do not demonstrate their applicability in a specific musical composition entirely.[2]

II. ANALYSIS OF SALIENT FEATURES OF MUSIC

In this section we present a method of score-based cognitive-perceptual analysis of the salient features of post-tonal music. We believe that the form of a post-tonal work can be described as a succession of salient events of varying importance, rather than phrases or larger formal units typically used to describe traditional compositions. By observing the local contexts in which salient events occur, decisions can be made about their varying levels of importance.

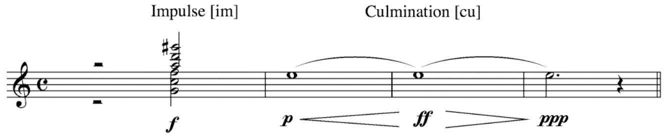


Figure 1. Examples of impulse [im] and culmination [cu].

First, an unprepared salient event may occur (a single tone, interval or chord with different dynamics or timbre, a characteristic motif, etc). We refer to such an event as **impulse [im]**. Second, such event may occur as a result of a previous development. In this case, we refer to it as **culmination [cu]** (see Figure 1). Third, the impulse may be followed by a substantial change in the course of musical development. We refer to such an event as **contrast [co]**. Finally, the contrast and culmination may coincide, in which case they constitute **contrast-culmination [cc]** (see Figure 2).

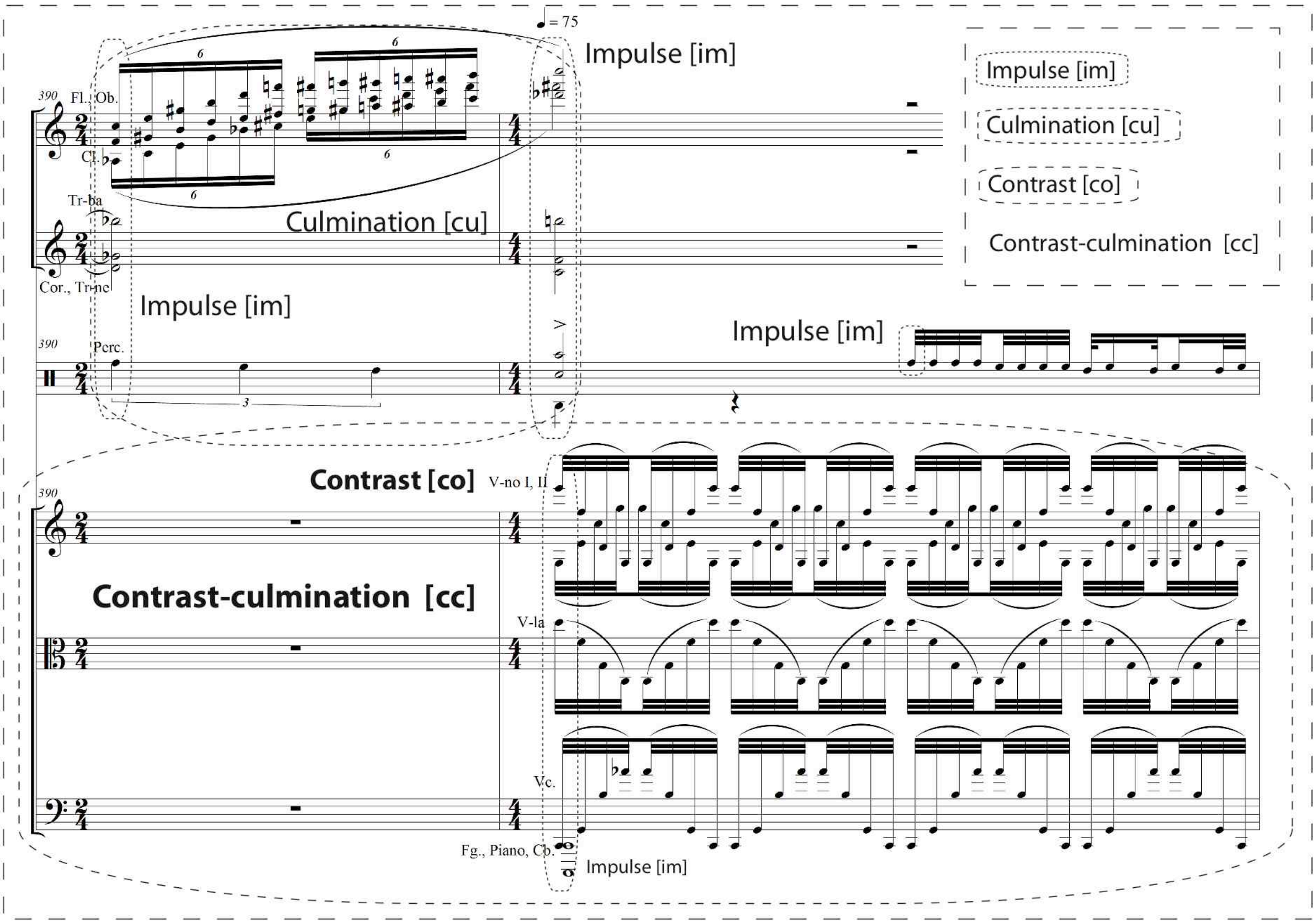


Figure 2. Example of impulse [im], culmination [cu], contrast [co], and contrast-culmination [cc] shown in Erkki-Sven Tüür's ensemble piece *Oxymoron* (2003) in bars 390–391.

The salient features are visualized on the linear time axis (x), and are given fixed values[3] on the vertical linear axis (y) (see Figure 3): impulse [im] = 10, culmination [cu] = 20, and contrast [co] = 30. The contrast-culmination [cc] can be expressed by two different values, 35 and 40, because the musical elements constituting the cc-s can coincide or follow each other in relatively short time but are to be considered as belonging to the event under observation. Measure numbers are included in the graphical representation. The graphical representation of such analysis can be realized as a bar or line diagram (see Figure 3) depending on purpose, i.e. the type of comparison with data obtained from other methods.

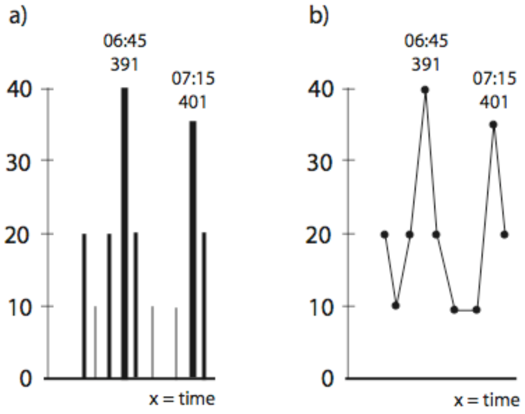


Figure 3. Two possibilities of graphical visualization of the same results of salient feature analysis of impulse [im], culmination [cu], contrast [co], and contrast-culmination [cc] shown in Erkki-Sven Tüür's ensemble piece *Oxymoron* (2003) around the bars 391 (see also Figure 2) and 401. Figure 3a is a bar diagram expressing the varying importance of particular events both in different heights on the y-axis and through the different boldness of columns (used in Lock 2010). Also the columns like the events itself stand alone, not being connected to each other and therefore not implying false connectivity (like in a line diagram, see Figure 3b). Figure 3b expresses the same results in a line diagram simply for the purpose of better comparitivity with line diagram of results of other methods. The time points for both cc's refer to the recording of that piece used for analysis (Tüür 2007).

III. ANALYSIS OF MUSICAL „ENERGY“

The musical process in the works of Erkki-Sven Tüür can be described as a series of **short formal sections** in which the emergence of **musical energy** in each carrier can be described by a single **qualitative state**. Such sections are referred to as **areas**. However, the separate carriers of an area may reveal simultaneously different qualitative states of energy (Figure 4).[4] For example, energy accumulation in dynamics can be accompanied by the energy persistence in the other carriers. In such a case, the area as a whole displays an energy accumulation (Figure 4a). Alternatively, the energy accumulation in register (e.g. rising melodic contour) can be accompanied by the energy loss in rhythm (e.g. shorter rhythmic values are progressively replaced by longer rhythmic values) whereas dynamics display no energy change. In such a case, the area as a whole reveals the energy persistence since the changes in register and rhythm mutually neutralize each other (Figure 4b). The latter case raises the question of the equivalence or compatibility of the different energy carriers which is discussed in more detail in Kotta (2011).

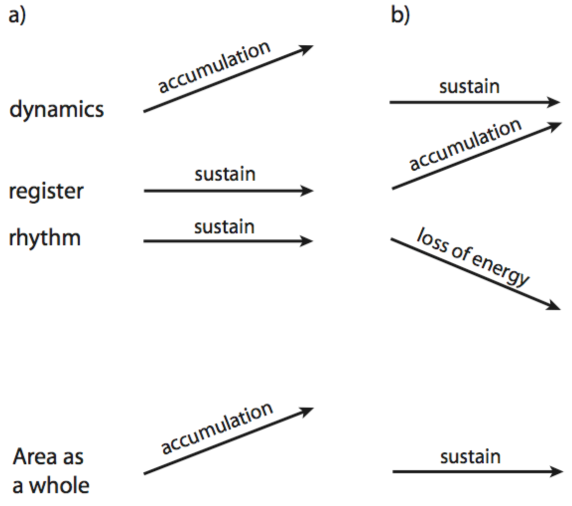


Figure 4. Area and its qualitative state of energy.

Points are places where the qualitative state of energy that characterize a certain carrier are replaced by another qualitative state (in the same carrier). Like areas, points can represent an energy accumulation, loss or sustain (or a tendency to move toward these qualitative states of energy). In order to determine the qualitative state of energy of a point, the energy value of the carriers of the preceding area need to be compared with the energy value of the carriers of the following area (Figure 5).

Next, we describe two possibilities out of many. If an area displays soft dynamics that begins to increase in the following area, and other carriers display no change, the point represents an energy accumulation (Figure 5a). Alternatively, an area may display some rhythmic activity which is then replaced by a longer or more sustained note(s) in the following area. At the same time, however, in terms of dynamics, piano in the preceding area is replaced by crescendo in the following area. The register displays no change. In such a case, the point represents the tendency to approach the energy persistence, since the loss of the rhythmic activity in the second area is offset by the energy accumulation in dynamics which was lacking in the first area, i.e. the energy value of both areas can be considered broadly equal (Figure 5b).

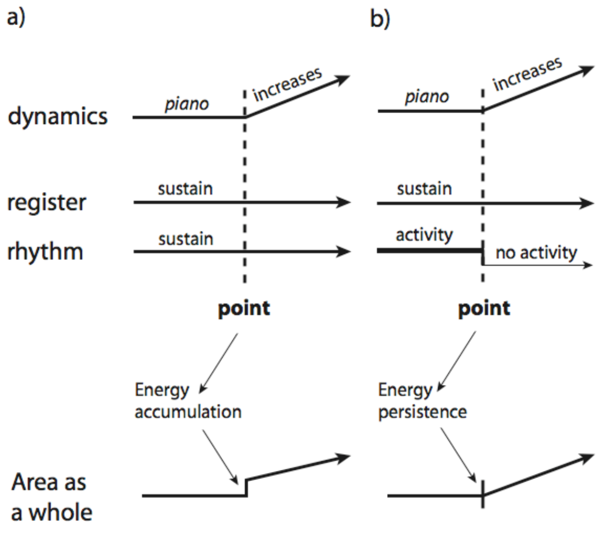


Figure 5. Point and its qualitative state of energy.

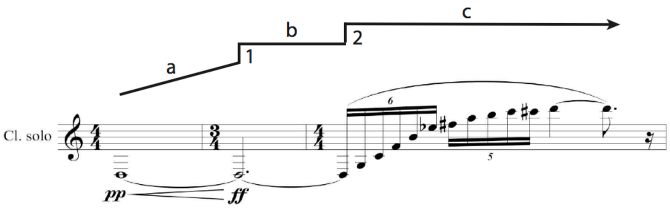


Figure 6. Erkki-Sven Tüürs Concerto for clarinet, violin and orchestra *Noēsis* measure 8–10 (solo clarinet part only).

Figure 6 features the opening gesture of Tüürs Concerto for clarinet, violin and orchestra *Noēsis*. The gesture can be divided into three areas which are referred as a, b, and c respectively. Area a begins *pp* and shows an energy accumulation in dynamics and energy sustain in other carriers (pitch). Measure 8 represents the first point. Area b presents an energy accumulation, because the energy value of area as a whole is higher than the energy value of area a. The area b shows the sustained energy since it displays no change in all energy carriers. The beginning of measure 10 (measure 3 in Figure 6) coincides with the second point. It, too, represents an energy accumulation since the area c, in addition, includes the rhythmic activity which the area b was lacking. Finally, area c represents a tendency toward the energy persistence since the rising melodic contour, i.e. the energy accumulation in register, is balanced by the retardation of rhythm, i.e. the energy loss in rhythm – thus the opposing qualitative states of energy of the two carriers mutually neutralize each other.

Applying the method described above the music of Tüür can be redrawn as an “energy curve” which, can be compared with curves obtained from listening tests and other types of analysis.

IV. REDUCTION AND AVERAGING OF TENSION CURVES

In this section we introduce the basic principles of our method of reducing and averaging tension curves obtained from listening tests using a slider-controller for continuous data capture. The data for this section will be obtained in perception tests and are the basis for further analysis. The perception tests are based on listening to a complete work or a closed part of a piece. The participants are asked to "draw" the tensional development in real-time, using slider controllers following the primary sensations of tension. By moving the slider continuously and/or suddenly, or remain on the same level, if the tension does not change. They are also asked also to suspend the upper region of the slider until the most tensional events would take place in the last third of the section.

A. Reduction principles to analyze tension curves

Here we describe the principles of reduction of tension curves obtained from continuous slider-controller listening test data.[5] First we defined and graphically visualized basic terms such as **Highpoint**, **Lowpoint**, **High Reverse Point**, and **Low Reverse Point** (Figures 7–9). We then demonstrate how these points occur (i.e. how they can be derived from the tension curves) and what role they play in the reduction process (Figures 10–13). In Figure 13 we show how many reductions are needed in order to reach a level in which the reduced curves of different participants can be compared).

A point from which the curve starts to descend we call a **Highpoint**, hereinafter referred to as **HP** (Figure 7a). A point from which a curve starts to ascend we call a **Lowpoint**, hereinafter referred to as **LP** (Figure 7b).

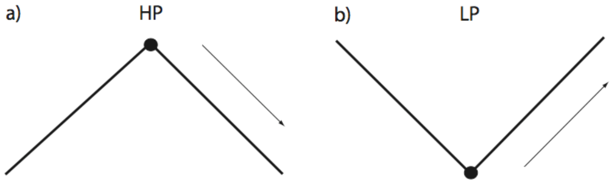


Figure 7. a) Highpoint (HP), b) Lowpoint (LP).

A highpoint which is directly followed by a lower highpoint we call a **High Reverse Point**, hereinafter referred to as **HRP** (Figure 8).

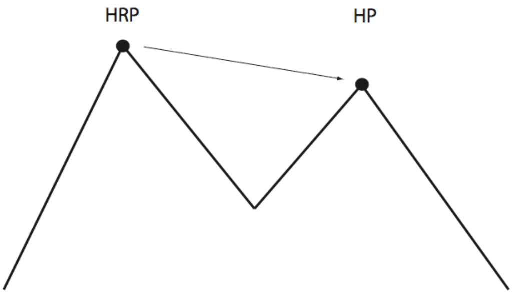


Figure 8. High Reverse Point (HRP) and the following Highpoint (HP).

A lowpoint which is directly followed by a higher lowpoint we call a **Low Reverse Point**, hereinafter referred to as **LRP** (Figure 9).[6]

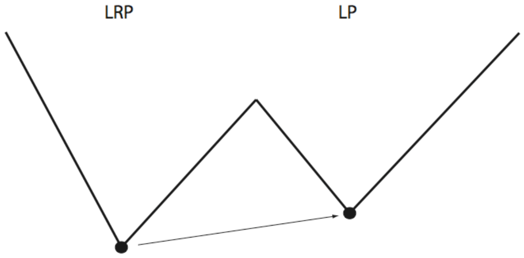


Figure 9. Low Reverse Point (LRP) and the following Lowpoint (LP).

The following examples (Figures 10–13) are hypothetical and fictive curves (read from left to right as a succession of points) in order to demonstrate possible instances which may occur in a reduction process. The Figure 10 represents a hypothetical raw curve (hereinafter referred to as curve A) and its first reduction (hereinafter referred to as curve B). In the process of reduction the high and low reverse points of the raw curve are connected with straight lines.

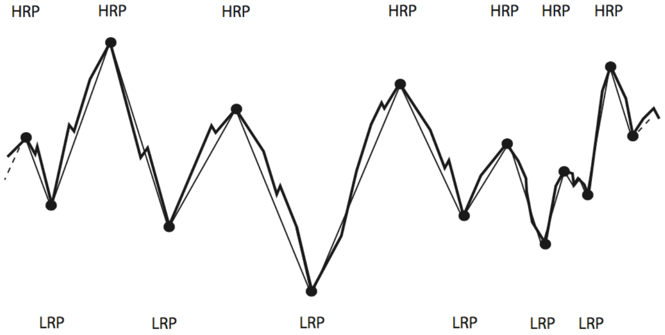


Figure 10. Hypothetical raw-data curve (curve A, thick black line) and its first reduction (curve B, thin black line).

Overlaps of curves and their reductions may occur, as when a curve displays HRPs and LRPs alternately following each other and shows a gradually increasing or decreasing value. Thus, instead of connecting a RP with the next closest RP, we connect it with a RP following the next closest RP. We refer to the points which are "ignored" in the process of reduction as evaded RPs (see the sixth RP in Figure 11).

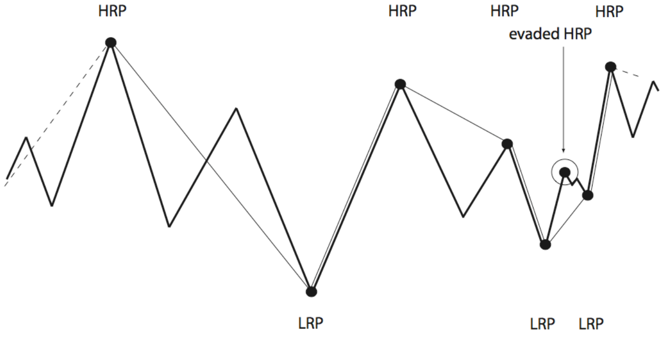


Figure 11. Thick black curve (curve B) and its reduction, thin grey curve (curve C). The figure shows three HRPs and three LRPs. By connecting the consecutive RPs the deeper level curve C emerges. The dashed lines before the first and after the last RP mark the hypothetical course of curve C.

Curve C includes points which cannot be categorized as HPs, LPs, HRPs, or LRPs. These points can be reduced out as shown in Figure 12.

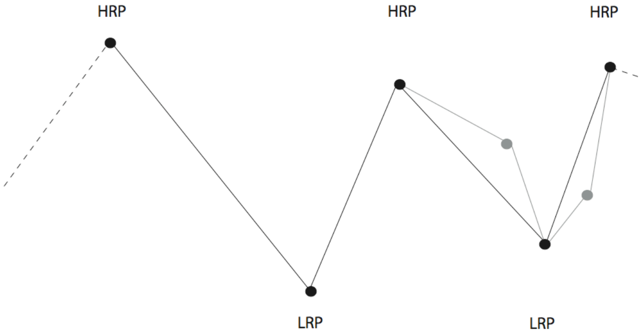


Figure 12. Curve C with points which are not understood as HPs, LPs, HRP-s or LRPs (points 4 and 6, marked with grey) are reduced out.

Analogously to curve A and B, curve C can be similarly reduced. We continued reducing curves until they show no more than three points within a fictive time window of one minute (counting from any point) (see Figure 13).

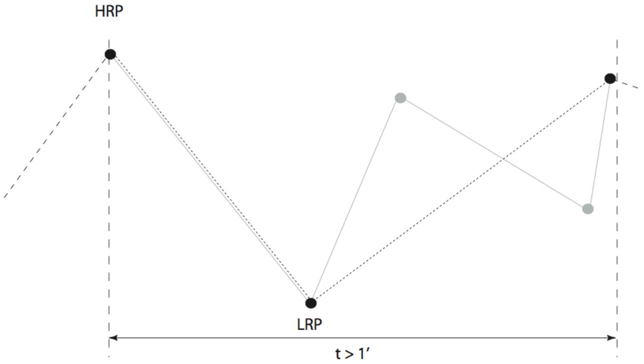


Figure 13. Curve C (grey thin line with partly grey dots) and E (dotted thin black line) within a one-minute time window. Curve C includes more than three points within one-minute time window. Curve E includes only three points within the one-minute time window and will be therefore the last reduction stage which we use in averaging the curves of all participants.

B. Averaging of intensity curves of multiple participants

The data of the reduced curves of selected participants[7] were collected and presented in a diagram with linear time axis (x) and the normalized tension level axis (y, maximum 100) (Figure 14). All points that occurred within 15 seconds were considered as belonging to the same time window. The first time window of 15 seconds began at the Tension Design Point (TDP)[8] 04:47.0 / 1 of the time axis (x) and the next window at the TDP 05:40.0 / 42 that immediately followed the first time window. The third time window began at the TDP 06:18.6 / 99 that immediately followed the end of the second time window etc. (see the grey areas and points in Figure 14). The time windows were converted to time points by averaging the time values of the points within the window under consideration (see the values of the time points next to the diagram of Figure 14).

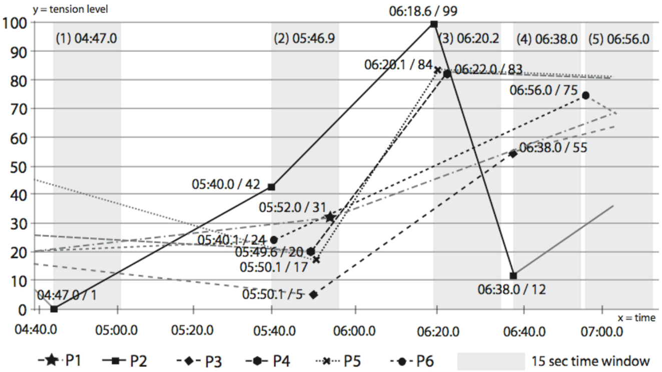


Figure 14. Fragment of the reduced perception test curves of Tüür's Fourth symphony (6 participants marked P1–P6 being distributed into 15 seconds time windows marked with grey areas each labeled with a number in brackets and the averaged time point values are shown in the top of each grey area.

Each time point has a tension value which is the arithmetical average of the tension values of the points making up 15 seconds time window (and which, in turn, is converted into the time point under consideration). However, in order to assess the actual value of a time point, the average reliability of the intensity value of the time point (hereinafter referred to as AR) has to be calculated according to the following equation: $AR = a \cdot n/x$. To calculate the AR of a time point we multiplied the average intensity value of that particular time point (a) with the ratio of the number of points making up the time window (n) and the total number of participants (x). If the time window under consideration includes only a single point we also applied this equation with $n = 1$ (see Figure 15). We considered the curve obtained from the calculated AR of the consecutive time points as the average of the intensity curves of all participants.

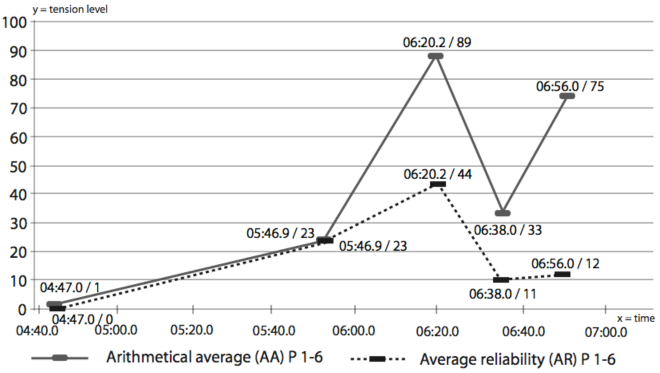


Figure 15. Fragment of the reduced perception test arithmetical average (hereinafter referred to as AA) and average reliability (AR) curves of Tüür's Fourth symphony (participants P 1–6). The grey colored upper curve represents the AA curve of TDP values whereas the fragmented black colored lower curve shows the AR curve of the TDP values.

The reduced and averaged curves are then compared with the graphical outputs of other analytical methods.

V. IMPLICATIONS

Through further research, we will compare the outputs of the three analytical methods described above with a traditional formal analysis of the works of post-tonal music. We are developing a method to explore optimal mappings between outputs of the analysis of form and the analysis of salient features of music, the analysis of musical "energy" and analytical outputs of the listening test participants along with their reductions. The authors of this study hope that the results applying the methods presented will contribute to a better understanding of the formal structure of post-tonal music and the techniques of prolongation.

ACKNOWLEDGMENT

The research behind this article is supported by Estonian Science Foundation (ETF8497) and the European Union sponsored DoRa program (Estonian Archimedes Foundation). We are very thankful to Paul Beaudoin (PhD) from Fitchburg University (USA) for valuable support by reviewing and editing this text.

REFERENCES

Kotta, Kerri (2011). On "Energy Conversion" in the Music of Erkki-Sven Tüür. In: *Abstract book of the Baltic Musics and Musicologies Conference 2011*, Canterbury, England.

Lock, Gerhard (2010). Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses „Oxymoron" [On factors influencing musical tension and its perception in Erkki-Sven Tüür's piece „Oxymoron"]. *Res Musica*, 2, 63–74.

Lock, Gerhard & Valk-Falk, Maris (2009). Tension design of contemporary orchestral music in different performances of "Sula" by Helena Tulve, In: *7th Triennial Conference of European Society for the Cognitive Sciences of Music - ESCOM 2009*, University of Jyväskylä, Finland, August 12–16, 2009 [on-line and memory stick publication].

Lock, Gerhard & Valk-Falk, Maris (2008). Investigating structure of contemporary music applying tension design and empirical perception analysis. In: C. Tsougras, R. Parncutt (Eds.) *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08)*,Thessaloniki, Greece, July 2–6, 2008 [on-line and CD ROM publication].

Tüür, Erkki-Sven (2007). *Oxymoron. Nyyd-Ensemble, cond. Olari Elts*, München: ECM Records (ECM New Series 1919, 476 5778).

---

[1] This article is part of the forthcoming doctoral dissertation of Gerhard Lock at Estonian Academy of Music and Theatre Tallinn (Estonia) under supervision of ass. prof. Kerri Kotta (PhD). It develops further an earlier preliminary approach to analysis of tensional development of post-tonal music on the basis of listening tests and cognitive analysis called Tension Design (TD, Lock & Valk-Falk 2008, 2009).

[2] However, the methods presented here are especially developed for the analysis of the orchestral music of Estonian contemporary composer Erkki-Sven Tüür (b. 1959) and some selected examples are taken from his music as well.

[3] The events are given the values only to express their relative weight in integers. Therefore the numbers do not refer to the absolute values such as frequency, pitch, etc.

[4] An excerpt of a musical score and its analytical explanation is provided in Figure 6.

[5] using Max/MSP software written by Hans-Gunter Lock (Estonian Academy of Music and Theatre, Estonian Academy of Arts, Tallinn, Estonia). In 2011 the apparatus and software were refined enabling the use of up to 16 slider controllers and now called Tension Design Experimental Apparatus (TEDEA), see http://www.schoenberg.ee/tension-design/tension-design.html.

[6] The first reverse point of all participants is always a highpoint, because the first and starting point is the zeropoint (which allows only ascending movement of the slider controller).

[7] The following examples are based on perception tests in 2010 on Erkki-Sven Tüürs Fourth and Sixth symphonies in Estonian Academy of Music and Theatre. The study included musicians, artists and dancers with a blend of expert and non-expert knowledge of music.

[8] By Tension Design Point (TDP) we mean points in which both time and tension level dimensions make up an integral whole.

A.3 Publication II (Lock 2014, 5.3)

“Visualisierende Analyse. Erkki-Sven Tüür: *Oxymoron*” (in German)

Visualizing as intuitive, systematic, system-based and automatized music analysis – *Oxymoron*

Gerhard Lock

Visualisierende Analyse

Erkki-Sven Tüür: *Oxymoron*

Einführung

Ziel des vorliegenden Artikels ist es, eine erste Übersicht über Grundprinzipien visualisierender Analyse zu geben.[1] Dabei können keine einzelnen Methoden im Detail erläutert werden. Zentrale Teile des vorliegenden Textes, darunter das Konzept der intuitiven, systematischen, system-basierenden bzw. automatisierbaren sowie automatischen Analyseverfahren wurden erstmals 2006[2], eine Graphik zur Kategorisierung von Achsen bei graphischen Darstellungsmöglichkeiten wurden als zweidimensionale Graphik erstmals 2009[3] veröffentlicht. Musik und deren Visualisierung scheinen seit der Entwicklung der Verschriftlichung von Musik beinahe unzertrennlich miteinander verbunden zu sein. Schriftliche Fixierung eines musikalischen Kunstwerkes sichert(e) in der Neuzeit das Überdauern und war (bzw. ist) zugleich Ausdruck des hohen Wertes einer Komposition. Komposition als ein »Gebilde mit Kunstanspruch« besitzt nach Klaus-Jürgen Sachs »eine eigentümliche Doppelexistenz von einerseits schriftlich fixiertem und andererseits lebendig vorgetragenem Werk, das erst in gegenseitiger Ergänzung beider Seinsweisen dauerhafte Wirkung und somit seine volle Bestimmung findet.«[4] Nach Erhard Karkoschka ist »die musikalische Aufzeichnung in erster Linie ein Hilfsmittel, um komplexe Musik aufzubauen, zu bewahren und zu vermitteln. Dabei beeinflussen die technischen Möglichkeiten einer Notation aber auch den Kompositionsakt, ja das gesamte musikalische Denken aller Musiker, so dass klangliche und bildhafte Seinsweise eines Musikwerkes in jeder Epoche charakteristisch verbunden bleiben.«[5] Musik und Bildende Kunst werden zwar häufig auf Grund ihrer dominierenden Merkmale als temporäre bzw. nicht-temporäre Kunstform als einander entgegengesetzt betrachtet, aber beide benötigen Raum und Zeit um sich zu entfalten bzw. wahrgenommen werden zu können.[6] Nach Karlheinz Stockhausen hat »die herkömmliche Notenschrift [...] keinen graphischen Eigen-Sinn; der Musiker versteht sie unmittelbar als *räumlichen Symbolismus* für einen *Zeitverlauf* [...]. Sowie aber der Zeitverlauf der Musik ins *Bild* gerinnt, dass zeitliche Bezüge zu räumlichen werden, dass die Qualität aufeinanderfolgender Ereignisse in die Strukturverhältnisse eines optischen Eindrucks umschlagen, gewinnt die Mitteilung von Musik einen gleichsam außermusikalischen Reiz; Zeiterfahrung lässt in Raumerfahrung sich transponieren.«[7] Bezüglich der Raum-Analogie im Zusammenhang mit Form treten nach György Ligeti »in unserer Vorstellungs- und Gedankenwelt Zeit und Raum stets miteinander gekoppelt« auf: »(W)o eine der beiden Kategorien primär vorhanden ist, stellt sich die andere assoziativ sofort ein«.[8] Auch auf struktureller Ebene finden sich bei beiden verschiedene Gemeinsamkeiten, die diese mehr verbinden als trennen.[9] Als Visualisierung von Musik im weitesten Sinne kann man auch koordinierte Bewegung bzw. Gestik zur Musik bzw. deren ›Verkörperlichung‹ – Tanz als gleichermaßen temporäre Kunstform – sowie Musik zu Gestik und Bewegung – u. a. Dirigieren bzw. das Musizieren nach Gesten bzw. einer Tanz-Choreographie – betrachten. Man beachte, dass die musikalische Notation ihren Ursprung in *Cheironomie* (eine Methode, bei der Musiker Handzeichen in der Luft ausführen, die den Klang andeuten) hat, deren nächste Etappe schriftlich fixierbare Neumen waren, aus denen sich dann das Fünf-Linien System entwickelte.

Prinzipien der Analyse und graphischen Darstellung

Wahl der Herangehensweise und das Prinzip der Repräsentation

Die hier vorzustellenden Prinzipien, Herangehensweisen bzw. Methoden gehen sowohl von einem engeren (musiktheoretischen) als auch einem weiteren Verständnis (das bewusste Sich-ins-Verhältnis-setzen-zu bzw. Reagieren-auf-Musik) von Analyse aus. Dazu sollte man sich bewusst machen, dass sowohl Kompositions- als auch Analyseprozesse in ihrem Wesen als kreativ, beide sozusagen als ›vorwärts‹ bzw. ›rückwärts‹ gerichtete Strukturierungsprozesse auf verschiedenen Ebenen betrachtbar (jedoch in der Regel nicht umkehrbar) sind[10] und das Prinzip der Repräsentation als genau zwischen ihnen verortbar. Hierbei ist entscheidend, welche Musik welche Methode ermöglicht und ob man eher verbal-narrativ, mathematisch, empirisch-statistisch oder aber strukturell-graphisch vorgehen sollte, um sowohl das Detail als auch die Gesamtheit (im wahrsten Sinne des Wortes) nicht aus den Augen zu verlieren. Dabei steht auch die Frage nach der ›richtigen‹ Methode im Raum. Letztlich gibt es laut Clemens Kühn »kein richtig oder falsch, sondern nur Grade von Angemessenheit«.[11] Dominik Šedivy fügt in einem Beitrag über den Pluralismus in der Musikanalyse hinzu, dass »der Pluralismus zwar eine einfache Konstruktion in der Theorie, in der Praxis jedoch überaus komplex ist dank Rahmenbedingungen wie Ausgangssituation, Zielvorgabe, dem Begriff der Nachvollziehbarkeit« und der Frage nach Graden der Angemessenheit von Analysen. Ihm zufolge müsse es zunehmend darum gehen, mit dem musikanalytischen Pluralismus überhaupt zu Rande zu kommen, und er schlägt hierfür eine Reihe von Methoden vor.[12] Darüber hinaus hängt die Methodenauswahl auch vom kognitiven Verarbeitungsstil – eher verbal (linke Gehirnhälfte) oder visuell (rechte Gehirnhälfte dominierend) – des Analysierenden ab[13], wobei man auch zwischen Vorlieben für schematische und räumliche Bilder sowie bildlich-piktographische und objektorientierte Visualisierung unterscheidet.[14] Der Mensch nutzt dazu verschiedene Repräsentationssysteme (für die Musikanalyse siehe die sog. *Modes* von Byron Almén[15]), deren Anwendbarkeit wiederum von unterschiedlichen funktionalen Prinzipien abhängen und die meist gemischt verwendet werden. Nach Roberto Donnini existiert Musik, bevor sie hörbar wird, immer »als *etwas anderes*« (wozu auch die Partitur als Form der Visualisierung gehört), und für denjenigen, der Musik hört und seine Visualisierung sieht, öffnen sich dadurch neue Wahrnehmungsaspekte; wer nur die Partitur betrachtet (vorausgesetzt diese als Zeichensystem adäquat zu verstehen, was für Nichtkenner nicht immer unmittelbar intuitiv möglich ist[16]), »*sieht* Musik« in Vermittlung »durch *etwas anderes*«, was als »Spiegelbild der Musik selbst« bezeichnet werden kann.[17] Sowohl schriftliche Fixierung – als Partitur im Allgemeinen aus in ihrer Bedeutung historisch verabredeten, teils auch aus neuen, abstrakten oder symbolischen Zeichen und verbalen Anweisungen bestehend[18] – als auch visualisierende Analyse von Musik beruhen auf solchen funktionalen Repräsentationssystemen und Visualisierungen, die entweder direkte oder indirekte Analogien zwischen auditiven und visuellen Phänomenen oder aber mentale Vorstellungen (innere Bilder) zu Hilfe nehmen, um ein Verständnis von und Denkweisen über Musik mit Hilfe nichtmusikalischer Mittel auszudrücken und sozusagen ›auf einen Blick‹ begreifbar machen zu können. Charles O. Nussbaum unterscheidet zwischen externen – die zu hörende »musikalische Oberfläche« – und internen Repräsentationen; letztere untergliedert er nochmals in den hierarchischen zielorientierten Aufbau der musikalischen Struktur und die mentalen musikalischen Modelle, die auf Basis dieser hierarchisch organisierten Repräsentationen konstruiert sind.[19] Nach Fred Lerdahl hören wir Musik sowohl sequentiell – als linear aufeinander folgende Ereignisse der »musikalischen Oberfläche« – als auch hierarchisch – kognitiven Erwartungs- und Präferenz-Regeln folgend. Er nennt zwei idealisierte Hörertypen: den »naiven« Hörer, der sozusagen nur sequentiell hört und den »erfahrenen« Hörer[20], welcher, erstmals beschrieben in Lerdahl's und Ray Jackendoff's linguistisch fundierter Generativer Theorie der Tonalen Musik (GTTM 1983)[21], entlang der graphisch-hierarchisch dargestellten (baumartigen) Prolongationsstruktur dem Zusammenspiel der wesentlichen strukturellen Aspekte der Musik (darunter Spannung und Entspannung) folgt und dabei die Gesamtheit nicht aus den Augen verliert.

Was ist der Gegenstand visualisierender Musikanalyse?

Der Gegenstand visualisierender Musikanalyse ist jegliche Art von Musik sowohl in schriftlich fixierter Form als Partitur, die sowohl notationstechnischen, semantischen als auch visuellen Wahrnehmungs- und Repräsentationsprinzipien unterliegt, als auch als Klangphänomen, welches Prinzipien der auditiven Wahrnehmung und semantischen Bedeutungen, u.a. auch graphisch darstellbaren mentalen Repräsentationen, unterliegt. Auch das Gestalt-Prinzip (sowohl in Musik als auch in bildender Kunst), die Bedeutung von Archetypen, einfachen Figuren (formalen Elementen) und Motiven sowie verschiedene Hierarchie-Prinzipien spielen hierbei eine Rolle.[22] Ein in beiden Bereichen verwendbares Prinzip ist das der

›Scene Analysis‹, bei der von einer Gesamtheit ausgegangen wird, welche sich jedoch durch unterschiedliche und sich verschiedentlich zu einander positionierende Teilelemente vor dem Auge/Ohr des Betrachters/Hörers erst auf Grund von Konzeptionalisierungsvorgängen etabliert.[23]

Hierarchien findet man sowohl in den oben genannten Baumstrukturen der GTTM als auch mit dem herkömmlichen Konzept des Haupt- bzw. Kernmotivs, welches in all seinen Formen (von der Erstform bis zu deren Ableitungen, Verschleierungen und Abstraktionen) teils auf verschiedenen Ebenen analytisch nachverfolgt werden kann. Analog zu Walter Gieselers drei Dimensionen ›Material‹, ›Struktur‹ und ›Form‹[24] von Musik können wir erweiternd sogar von fünf sprechen[25], auf denen jeweils auch unterschiedlich Arten der Visualisierung von Musik stattfinden kann: (1) ›Nano-Ebene‹: Physik des Klangs im Allgemeinen (Sinuskurve, Hüllkurven, Fourier-Komponenten und deren graphische Repräsentation als Spektren im Sonogramm/Spektrogramm), (2) ›Mikro-Ebene‹: das musikalische Material: Elemente, aus denen eine Komposition besteht (Motive, Phrasen, Themen usw.; graphisch u. a. als Konturen, darstellungsgeometrisch präzise Repräsentation oder Symbole), (3) ›Medio-Ebene‹: Struktur (das Verhältnis der Elemente zu einander, Kompositionstechniken; graphisch u. a. als Netzwerke und Beziehungsgeflechte), (4) ›Makro-Ebene‹ – Form als größerer Komplex und Einheit (gesamtes Werk) sowie Abschnitte in Relation zum gesamten Stück: graphisch u. a. in dynamischen Spannungsbögen bzw. Wellenformen oder symbolisch (proportional oder ikonographisch), (5) ›Meta-Ebene‹ – die Art und Weise, wie wir Musik verstehen und an sie analytisch, historisch, philosophisch, psychologisch, semiotisch usw. herangehen. Eine auf dem Konzept verschiedener aufeinander folgender struktur-hierarchischer Ebenen bzw. Schichten (Vordergrund, Mittelgrund, Hintergrund) beruhende Herangehensweise ist die (auch ›lineare Analyse‹ genannte) graphische Methode Heinrich Schenkers.[26] Die in der Partitur enthaltenen Informationen zu verschiedenen Aspekten von Musik durch das Prinzip des Ineinandergreifens der vertikalen (Tonhöhen, Register) und der horizontalen (Tondauern, zeitliche Organisation, Rhythmus) Dimension[27] sind häufig auch das erste Ziel visualisierender Analyse. Dabei gibt es entweder Methoden, die das visuelle Bild der traditionellen Notation direkt mit graphischen Mitteln analysieren und gefundene Beziehungen einzelner Aspekte überwiegend zeitungebunden ausdeuten und den klanglichen Aspekt meist vernachlässigen (Pitch-Class-Set Theory[28]), oder Methoden, die das traditionelle Notationssystem als Basis für eine graphische Darstellungsweise verwenden, die vor allem die Zeitachse nochmals abstrahiert (Schenker-Methode). Dazu kommen Methoden, die die Musik entweder graphisch mit Hilfe von abbildungsgeometrischen Prinzipien[29], mit auf diesen aufbauendem Sonic Design und ähnlichen Methoden oder in alphanumerischen oder algorithmischen ›Sprachen‹ sozusagen neu notieren, mit dem Zweck Musik digitalisiert universal zu erhalten und somit auch visualisierende Analyse zu ermöglichen (siehe Literaturangaben in der Auswahlbibliographie). Bei Diagrammen wird häufig, ähnlich wie bei der Partitur, die Zeit in die Darstellung mit einbezogen (häufig, aber nicht immer, ebenfalls dargestellt in der Horizontale), aber sie müssen ausgedeutet werden; bei Graphiken dagegen spielt die Zeit (gewöhnlich ebenfalls in der Horizontale) erst dann eine Rolle, wenn man diese gedanklich als eine graphische Partitur interpretiert oder wenn sie als eine solche ausgewiesen ist. Eine Partitur ist (solange sie keine Hörpartitur ist) von ihrem Wesen her vorschreibend (präskriptiv), Diagramme jedoch sind beschreibend (deskriptiv). Graphiken dagegen sind als Gebrauchsgraphik (z. B. als Partitur oder Architekturzeichnung) präskriptiv, als Kunstgraphik dagegen deskriptiv (siehe dazu graphische Partituren von John Cage, Earl Brown oder Morton Feldmann, mit denen auch eine spezielle Aufführungs-Problematik einhergeht[30]).

Nach Henkjan Honing kommt es »bei der Entscheidung für ein spezielles Repräsentationssystem und seinen Funktionseinheiten zu einem Kompromiss [...]; gewisse Informationen werden verdeutlicht auf Kosten anderer, die in den Hintergrund gedrängt werden und die es schwer wird wieder zu finden.«[31] Beispiel 1 veranschaulicht, welche grundsätzlichen Darstellungsmöglichkeiten im x/y/z-Koordinatensystem bei welchem Grad an Definitionsgenauigkeit der Achsen denkbar und auch am weitesten verbreitet sind. Dies beinhaltet auch die Frage, ob die Darstellung eher Phänomene der Makro-Ebene (Form), Medio-Ebene (Struktur), Mikro-Ebene (Material) oder Nano-Ebene (Spektralaufbau) zeigt. Für die Hauptachsen kann man, nach Erhard Karkoschka, bei Notation (und somit auch bei visualisierender Analyse) von folgender Kategorisierung ausgehen: (1) Präzise Notation, (2) Rahmennotation, (3) Verweisende Notation, (4) Musikalische Graphik.[32] Die im Beispiel verwendete Kategorisierung ist präzisiert, d.h. als erste Kategorie ist nun das theoretisch kontinuierliche Spektrum vorangestellt, die präzise Kategorie wurde untergliedert (diskret, linear, logarithmisch; relativ, proportional, nichtproportional) und es wurde eine fünfte Kategorie hinzu gefügt, die innerhalb der musikalischen Graphik zwischen symbolischer Darstellung und gänzlich undefinierten Achsen differenziert.

Die Darstellung musikalischer oder auf Musik bezogener kognitiver Phänomene muss nicht geometrisch, schwarz-weiß graphisch (bzw. in Graustufen) beschränkt bleiben, es können auch verschiedene Farbsysteme zur Anwendung kommen, die entweder synästhetische, symbolische oder strukturelle Bedeutung haben können.[33] Es kann auch vorkommen, dass die Achsen unterschiedlich definiert sind: Z. B. werden bei der traditionellen Partiturdarstellung mehrere Stimmen bzw. Stimmengruppen visuell untereinander aufgereiht notiert, klingen häufig jedoch im gleichen Register, so dass man die entsprechenden Stimmen bzw. Stimmengruppen eigentlich räumlich in Schichten übereinander abbilden müsste, die bei einer darstellungsgeometrischen bzw. Sequenzer-Darstellung übereinander gelagert sind. Dies führt uns zur bildlich-räumlichen bzw. perspektivisch-virtuellen z-Achse (sogenannte ›Tiefendimension‹): bei sogenannten Wasserfalldiagrammen z.B. zeigt die z-Achse Frequenzen/Obertöne und die y-Achse die Amplitude bzw. Lautstärke, die Zeit bleibt auf der x-Achse. Dabei entsteht diese Darstellung, so Dick Raaijmakers, durch das Aneinanderreihen von sich wiederholenden Frequenzschichten (in der Terminologie von Pierre Boulez ein »horizontales Arpeggio«), was als ›akustisches Aggregat‹ bezeichnet wird.[34] Darüber hinaus verliert die »Tiefendimension« in zwei weiteren Fällen überhaupt ihre darstellungsgeometrisch-virtuelle Räumlichkeit: erstens im Zuge der verweisenden Kategorie, wenn die Lautstärke oder Dauer eines Tons durch veränderte zweidimensionale Größenunterschiede dargestellt wird; zweitens wenn bei Spektralanalyse-Bildern die Lautstärke durch Farb- oder Graustufen-Nuancen signalisiert wird.

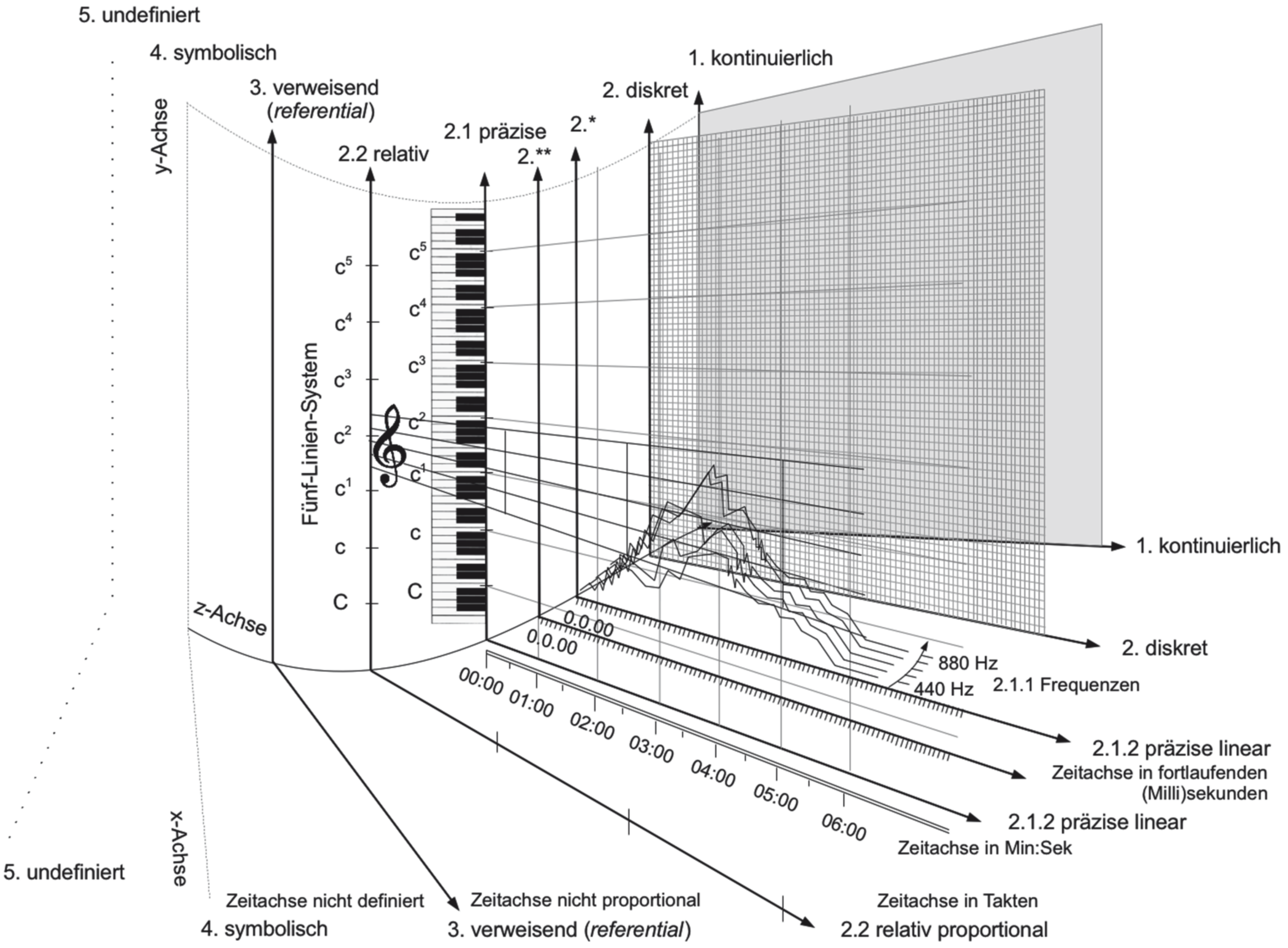


Abbildung 1a

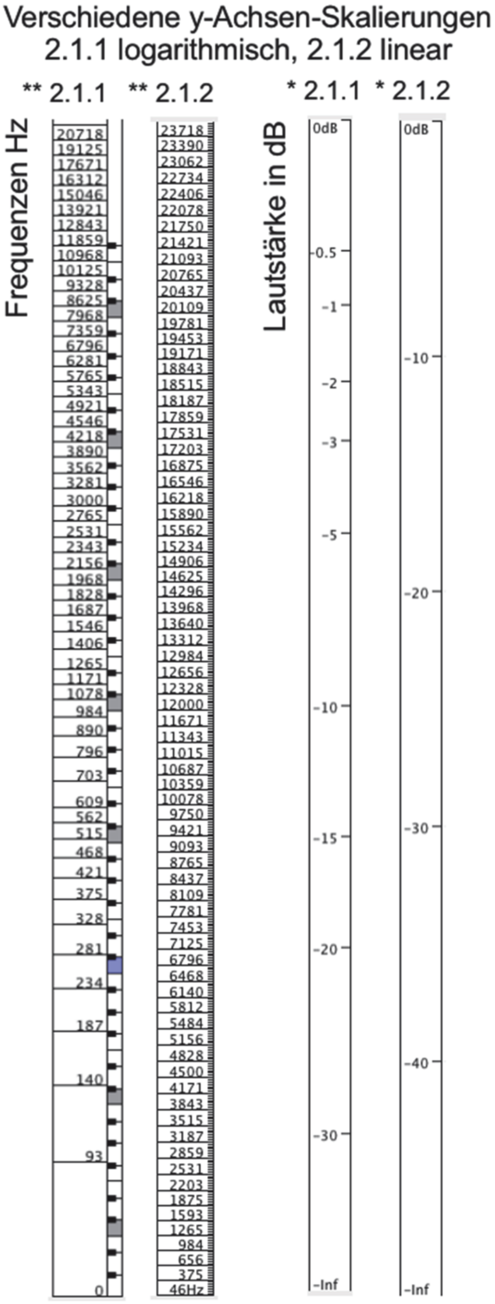


Abbildung 1b

Abbildungen 1a und 1b:
Verschiedene Darstellungsmöglichkeiten im x/y/z-Koordinatensystem. Alle Achsen können wie folgt definiert werden:
1. kontinuierlich,
2. diskret,
2.1 präzise,
2.1.1 nicht linear,
2.1.2 linear,
2.2 relativ,
2.2.1 poportional,
2.2.2 nicht proportional,
3. verweisend (›referential‹),
4. symbolisch,
5. undefiniert.
Rechts sind verschiedene y-Achsen-Skalierungen dargestellt. Die wohl bekannteste ist die sogenannte Klavierrollen-Darstellung (›piano roll‹, siehe mittlere y-Achse der Hauptgraphik des Beispiels), die sich auch bei MIDI-Darstellungen findet und gewissermaßen auch der darstellungsgeometrischen Visualisierung ähnlich ist.

Methoden visualisierender Analyse

Analyse- und Visualisierungsprozesse und das Moment der Automatisierbarkeit

Ein Großteil der Methoden visualisierender Analyse beinhaltet das Moment der Automatisierbarkeit bzw. Automatisierung, welches sich daran zeigt, ob die Analyse- und Visualisierungsprozesse zwischen Analysierendem, Analyse-Objekt und Automatisierungshilfsmittel mittelbar oder unmittelbar ablaufen (siehe Beispiel 2). Hierbei muss unterschieden werden zwischen unterschiedlichen Datenerfassungsmethoden (›analog‹, d.h. ohne Automatisierungshilfsmittel: statische und temporär nicht gebundene Visualisierungsvorgänge; ›elektronisch-digital‹, d.h. durch elektronische Detektoren oder Kontroller: dynamische, temporär nicht gebundene bzw. verzögert temporärere sowie in Echtzeit temporär unmittelbare Vorgänge) und dem Moment der Visualisierung der erfassten Daten dynamisch und in-Echtzeit automatisch.

Die in Beispiel 2 genannten Automatisierungshilfsmittel (AHm) können auch Graphikprogramme sein, mit deren Hilfe der Visualisierungsprozess u. a. durch Kopier- und Modifizierungsmöglichkeiten erleichtert werden kann. Rückkoppelungen zwischen Analysierendem und AHm können aktiv-invasiv und passiv erfolgen, wenn der Analysierende z.B. Einstellungen des Programms (zwecks Abstimmung des Resultats) in Echtzeit verändern kann (aktiv-invasiv) oder bei Performance-Analyse durch Veränderung des Spiels, welches sich auf den Visualisierungsprozess des Programms auswirkt und dann unmittelbar angezeigt wird (passiv). Dabei kann Automatisierung sowohl die Erfassung von Daten als auch die visuelle Darstellung betreffen, wobei bei letzterer unterschieden werden muss zwischen direkter in-Echtzeit-Visualisierung und der nach dem dynamischen Visualisierungsprozess parallel zur Musik ablaufenden visuellen Darstellung in Echtzeit.

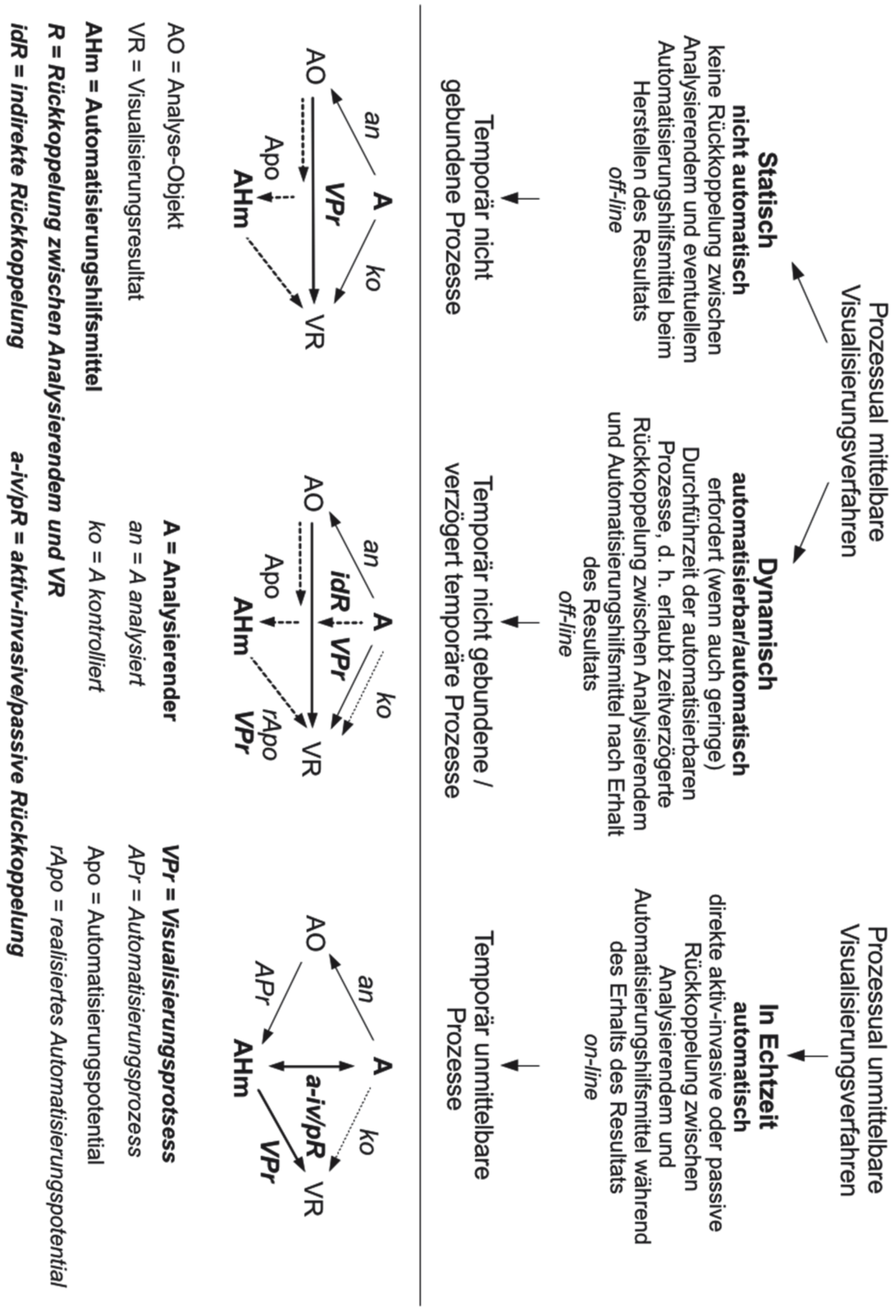


Abbildung 2: Statische, dynamische und Echtzeit-Verfahren und der Anteil von Automatisierungshilfsmitteln (AHm, d. h. Computersoftware) beim Visualisierungsprozess bei temporär gebundenen und nicht gebundenen Analyseprozessen. Schräggedruckte Abkürzungen besagen, dass es sich um Prozesse handelt. Fettgedruckte Strichelungspfeile zeigen an, ob Automatisierungs-/Rückkoppelungspotential (Apo) vorhanden ist, dünngedruckte Strichelungspfeile bedeuten, dass der Analysierende nur unter bestimmten Bedingungen einen Kontrolleinfluss auf das Visualisierungsresultat (VR) hat.

Zur Systematik von Methoden visualisierender Analyse

Im Folgenden werden vier zentrale – intuitive, systematische, systembasierende bzw. automatisierbare sowie automatische – Herangehensweisen bzw. Methoden und Darstellungsverfahren vorgestellt, die sowohl das engere als auch das weitere Verständnis von Analyse einbeziehen. Auf einige spezielle Beispiele wird im Folgenden direkt in Fußnoten verwiesen, zum Teil finden sich Beispiele und Literatur aber auch in der Auswahlbibliographie.

Als ›intuitive Analyseverfahren‹ kann man alle diejenigen Vorgehensweisen bezeichnen, bei denen entweder in der visuellen Darstellung kein vorher fertiges System angewendet wird, mit dem die zu betrachtende Musik in Einklang gebracht werden muss, oder die die spontane und nicht vorher in ein bestimmtes System eingefasste Reaktion des Hörers in Echtzeit aufzeichnen. Es handelt sich hierbei entweder um Höranalyse (bzw. Hörpartitur[35]), oder Partituranalyse, bei der Zusammenhänge auf meist visueller Basis gefunden werden. Häufig geht man dabei von Einzelaspekten des zu analysierenden Stückes und seines Materials aus, um dann auf musiktheoretischer Basis weitere Schlussfolgerungen auf die Struktur und Form der Musik zu ziehen. Ein eigenes Teilgebiet ist auch das kreative Zeichnen und Malen während des Musikhörens, insbesondere auch dessen didaktische Anwendung.[36]

Zu den ›systematischen Analyseverfahren‹ kann man Vorgehensweisen zählen, bei denen im Zusammenhang mit Musik eines bestimmten (zeitgenössischen) Komponisten oder einer bestimmten kompositionstechnisch-stilistisch-ästhetischen Richtung ein Arsenal an Möglichkeiten erarbeitet wird, das man systematisch auf ähnliche Musik anwenden kann. Dabei ist ein solches systematisches Verfahren in der Entwicklungsphase häufig eng mit einer intuitiven Herangehensweise verknüpft und ein System entsteht prozessual mit der kognitiven Erfahrung – es bleibt häufig jedoch entweder auf ein konkretes Werk, eine Werkgruppe oder einen Komponisten (in einer bestimmten Schaffensphase) beschränkt. Auf geometrischen bzw. symbolischen Figuren basierende Darstellungsweisen zur Formanalyse[37] gehören in diese Kategorie genauso wie die Melodiekontur-Analyse (etwa bei Arnold Schönberg[38]) und Herangehensweisen, die dem abbildungsgeometrischen Prinzip bzw. Sonic Design ähnliche sind. Als teils systematisch, teils system-basierend in der graphischen Darstellung und analytisch können auch sogenannte Realisationspartituren angesehen werden, die insbesondere bei (nicht traditionell notierbarer) elektronischer Musik Anwendung finden.

Als ›system-basierende bzw. automatisierbare Verfahren‹ kann man einerseits Methoden zur Analyse dodekaphonischer und serieller Musik bezeichnen, die darauf beruhen, bestimmte Reihen und deren Permutationen und Transpositionen (nicht nur Tonhöhen-, sondern auch Dauern, Dynamik Artikulationsreihen usw.) in der Partitur zu entdecken; dies meistens ausgehend von durch den Komponisten vorgegebenen Grundreihen. In der Pitch-Class-Set Theory wird auf Basis der Tonhöhenstruktur von posttonaler Musik mit Tongruppen (Monaden, Duaden, Trichorden usw.), deren Transpositionen, Umkehrungen, Spiegelungen usw. ein System von Kombinationen entwickelt, das man auf durchaus automatisierte Weise in der Partitur aufsuchen, bestimmen und sowohl innerhalb der Partitur als auch eigenständig graphisch darstellen kann. Dazu können auch Erhard Karkoschkas Klangspannungs-, Intervallschrittgröße, Impulsdichte und Durchschrittender-Raum-Darstellungen (nicht automatisiert) genannt werden. System-basierend und teilweise automatisiert sind auch die Regeln und die hierarchische Darstellung der GTTM[39] sowie teilweise farbige Matrix-Darstellungen der Wahrnehmung musikalischer Grundstrukturen in verschiedenen Schichten.[40]

Andererseits lassen sich hier auch verschiedene auf Algorithmen reduzierbare Methoden hinzuzählen, die man mit Hilfe mathematischer Formeln automatisiert darstellen kann[41] (hierbei können auch mathematische Funktionen musikalisch hörbar gemacht werden[42]), darunter auch verschiedene Aspekte von Entropie und Informationsdynamik sowie Asymmetrie/Symmetrie.[43] Darüber hinaus sollten hier auch visuelle Rückkoppelungssysteme für Performance-Analyse genannt werden.[44]

Als ›automatische Verfahren‹ kann man Methoden bezeichnen, die Musik mit Hilfe von Computerprogrammen (digitalisiert u.a. als MIDI- oder Audio-Dateien o. ä.) graphisch darstellen, darunter die Spektral- bzw. Sonogrammanalyse oder das sogenannte AudioKeySOM[45], welches die harmonischen Felder tonaler Audiodateien in Echtzeit in verschiedenen Farben visualisiert. In Humdrum/Kern-Sprache geschriebene tonale Musik wird u. a. automatisch in ihrer Tonartenstruktur in verschiedenen Farben in einer der GTTM ähnlichen Baumstruktur visualisiert.[46] Dazu zählen auch algorithmische, heutzutage computergestützte, Repräsentationssysteme, u.a. GUIDO, DARMS, MuseData, SCORE, MusiXTEX, LilyPond, Charm-System, MusicKit, MusicXML[47] und Programme, die auf Basis von MIDI-Daten (Musical Instruments Digital Interface, seit Anfang der 1980er-Jahre firmenübergreifender Standard) visualisierende Analyse ermöglichen (u.a. mit Hilfe des Disklaviers[48], RUBATO[49], MAM Player[50], die visualisierende Darstellung erfolgt bei letzterem in Echtzeit) und die auf MIDI-Basis erstellten Darstellungen der musikalischen Zeitanalyse[51], MMA Tools[52], MIDI Toolbox[53], darunter auch *Multitrack* bzw. *Sequenzer*-Programme. Hierbei entstehen die Darstellungen nach Eingabe der Daten aufgrund automatischer Rechenvorgänge ohne Einwirkung des Menschen, der Analysierende muss jedoch die entstandenen Bilder interpretieren und hat (im Rahmen der Möglichkeiten des jeweiligen Programms) oftmals nur eine geringen Einfluss auf das Endergebnis. Das UPIC[54] und andere Programme[55] sind dagegen Systeme zur (dynamischen und z.T. in Echtzeit unmittelbaren) Umwandlung visueller Darstellung in musikalischen Klang (anwendbar im Kompositionsbereich). Ein eigenständiger Bereich ist die auf Video-Motion-Capture oder anderen Techniken beruhende Bewegungsanalyse siehe u.a. die MoCap Toolbox[56] und die *Music Paint Machine*.[57]

Zwischen den systematischen und den sowohl statisch als auch dynamisch und in Echtzeit system-basierenden bzw. automatisierbaren Verfahren stehen sogenannte ›empirische‹ Visualisierungverfahren, bei welchen mit Apparaturen (u. a. Schiebereglern bzw. Kontrollern) die Reaktionsdaten der Beteiligten (die selbst häufig kognitiv-intuitiv entstehen oder bei der erfahrene Teilnehmer auch systematisch vorgehen können) durch bestimmte definierte Mess- oder Aufnahmevorgänge erfasst werden (u. a. der festgelegte Umfang von Schiebereglern bzw. Ausschlagbereich von Kontrollern), die dann meist als Diagramme visualisiert werden und danach eine Ausdeutung oder weitere standardisierte analytische Methoden erfordern (u.a. Mittelwerts- und Trendverlaufberechnungen, Standardabweichungen und Reduktionsverfahren), deren Visualisierung dann in die Kategorie ›automatische Verfahren‹ gehört.

Insbesondere die Analyse des Spannungsverlaufs von Musik bedient sich verschiedener graphischer Darstellungsmöglichkeiten diesen entweder intuitiv, systematisch oder aber system-basieren bzw. automatisierbar realisierend.[58] Ein eigener Bereich ist die sogenannte Kymatik, die die Visualisierung von Klängen und Wellen bezeichnet, z. B. durch eine in Schwingung versetzte Platte, auf der sich Pulver befindet, oder aber durch Lautsprecher in Flüssigkeiten, durch die Muster entstehen. Sie stützt sich u.a. auf die Akustik-Forschungen von Ernst Florens Friedrich Chladni (sogenannten ›Chladni-Figuren‹).[59]

Didaktische Anwendungen und Perspektiven visualisierender Analyse

Visualisierende Analyse findet seit langem Anwendung sowohl im wissenschaftlichen als auch im propädeutisch-didaktischen Kontext. Im Rahmen dieses Artikels ist es nicht möglich, dies auch nur annähernd historiographisch oder umfassender zu dokumentieren. Es können hier stellvertretend nur einige ausgewählte Namen und Publikationen genannt werden, die einen ersten Einstieg ermöglichen sollen, wobei vor allem das didaktische Potential im Fokus steht. So enthält u.a. Hugo Leichtentritts *Musikalische Formenlehre* (1907)[60] aus der Reihe *Handbücher der Musiklehre* (intuitive bzw. systematische) Grafiken, die formale Spannungsverläufe darstellen. In den 1920er-Jahren hat u.a. Hans Mersmann[61] graphische (vornehmlich intuitive) Darstellungstechniken explizit im pädagogischen Kontext verwendet. In der zweiten Hälfte des 20. Jahrhunderts haben insbesondere Hansgeorg Mühe in seiner *Musikanalyse* (1978) und Hubert Wißkirchen im *Arbeitsbuch für den Musikunterricht in der Oberstufe* (1992) umfangreich intuitive, systematische und system-basierende visualisierende Darstellungsmöglichkeiten beschrieben, entwickelt und verwendet. Auch Ulrich Kaisers *Gehörbildung* (1998)[62] enthält Anweisungen zur intuitiven und systematischen graphischen Darstellung von Musik (darunter Hörpartitur-Aufgaben). Genannt werden kann auch eine systematische bzw. system-basierende Methode des Autors des vorliegenden Artikels[63].

Konkrete didaktische Anwendung eines automatischen Verfahrens zur graphischen Analyse der Intonation bei Kinderstimmen bieten die *Musical Micro Analysis Tools* von Stefanie Stadler Elmer und Franz-Josef Elmer (2000), und auch MIDI-basierte automatische Darstellungsweisen von Musik (Malinowski seit 1985) haben einen unmittelbaren didaktischen Wert. Auch bei neuartigen, didaktisch ausgerichteten Notationssystemen, darunter die sog. Figurennotation (deren Ziel es ist, für Kinder unmittelbarer verständlich zu sein[64]) spielen graphische und farbliche Darstellungen eine entscheidende Rolle. Darüber hinaus gibt es eine Reihe von Studien, die die Möglichkeiten graphischer Repräsentation (von intuitiven bis zu automatisierbaren Verfahren) von Musikfragmenten und kurzen Musikstücken durch Kinder und Jugendliche untersuchen[65] und deren Ziel es ist, die damit verbundenen Wahrnehmungsprozesse und Repräsentationsweisen in Korrelation zum Alter und Erfahrungshintergrund der Hörenden zu setzen, um somit das Potential aufzeigen, welches in der visualisierenden Darstellung und Analyse von Musik liegt.

Abbildungen 3a, 3b und 3c:
Erkki-Sven Tüür, *Oxymoron (Musik für Tirol)* für großes Ensemble,
Frankfurt 2003, Particell der Takte 387–391 erstellt von Kerri Kotta.

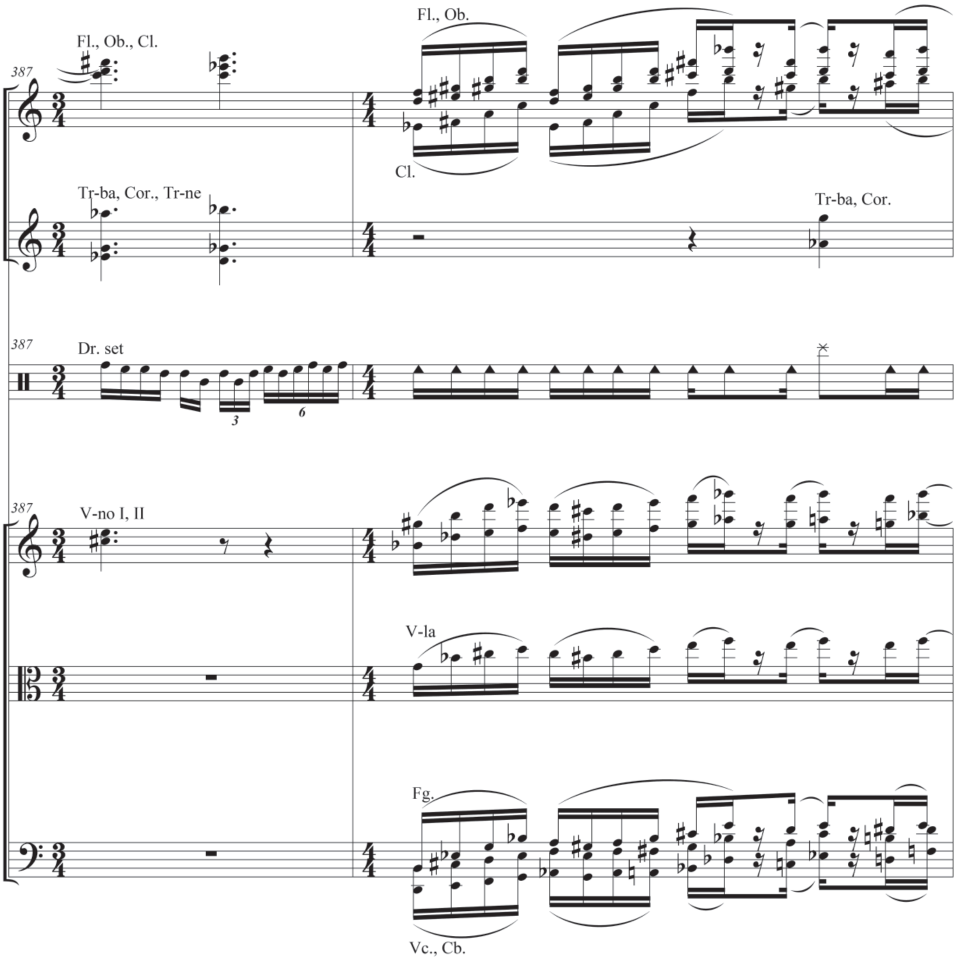



Abbildung 3a

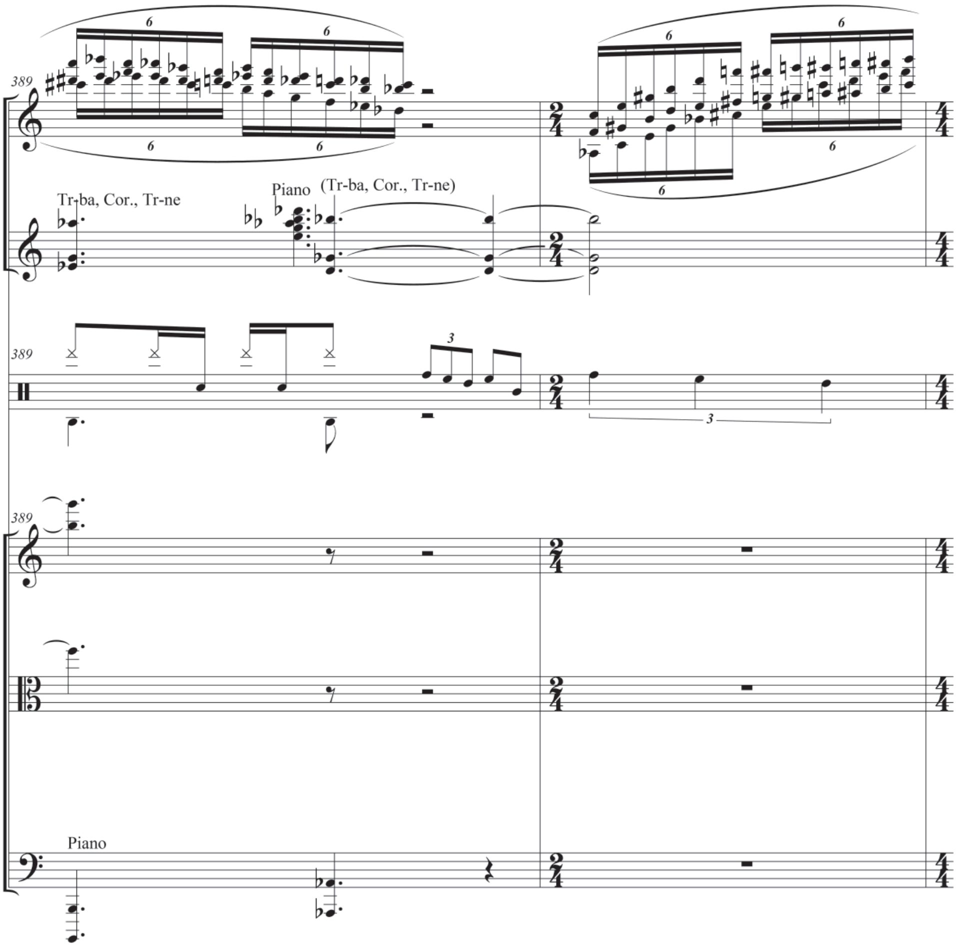



Abbildung 3b

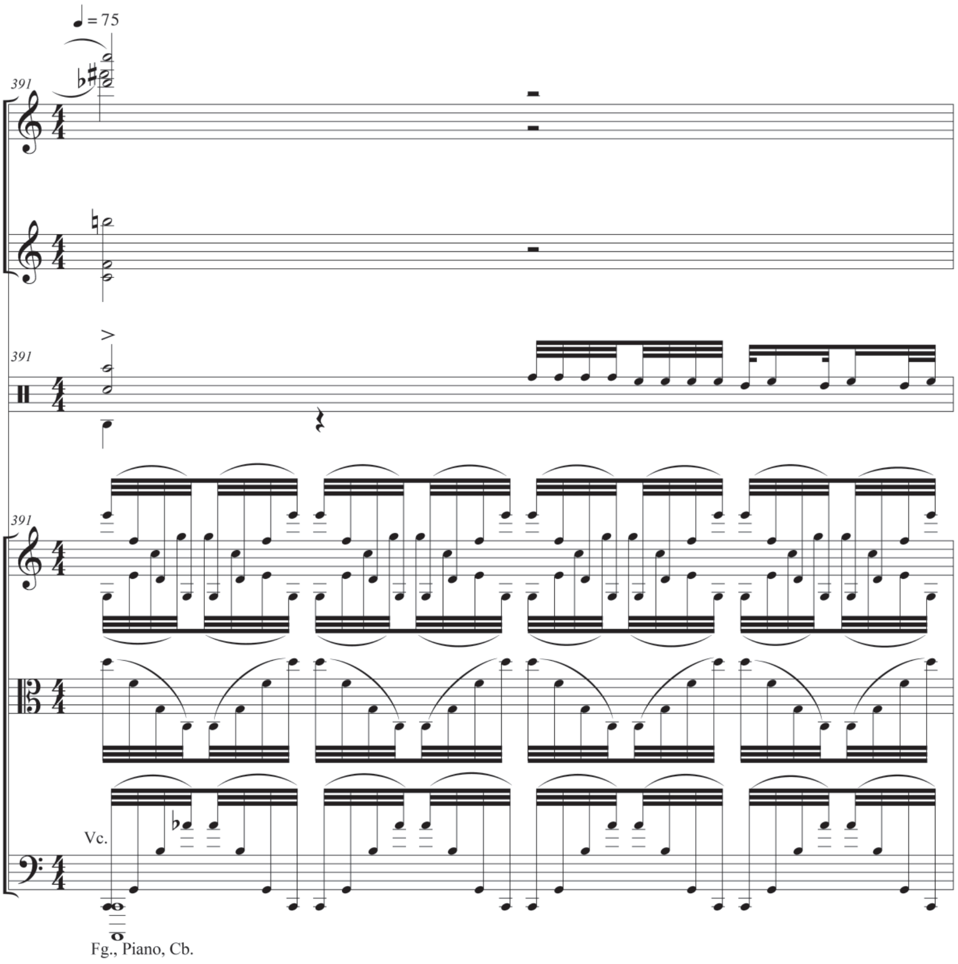


Abbildung 3c

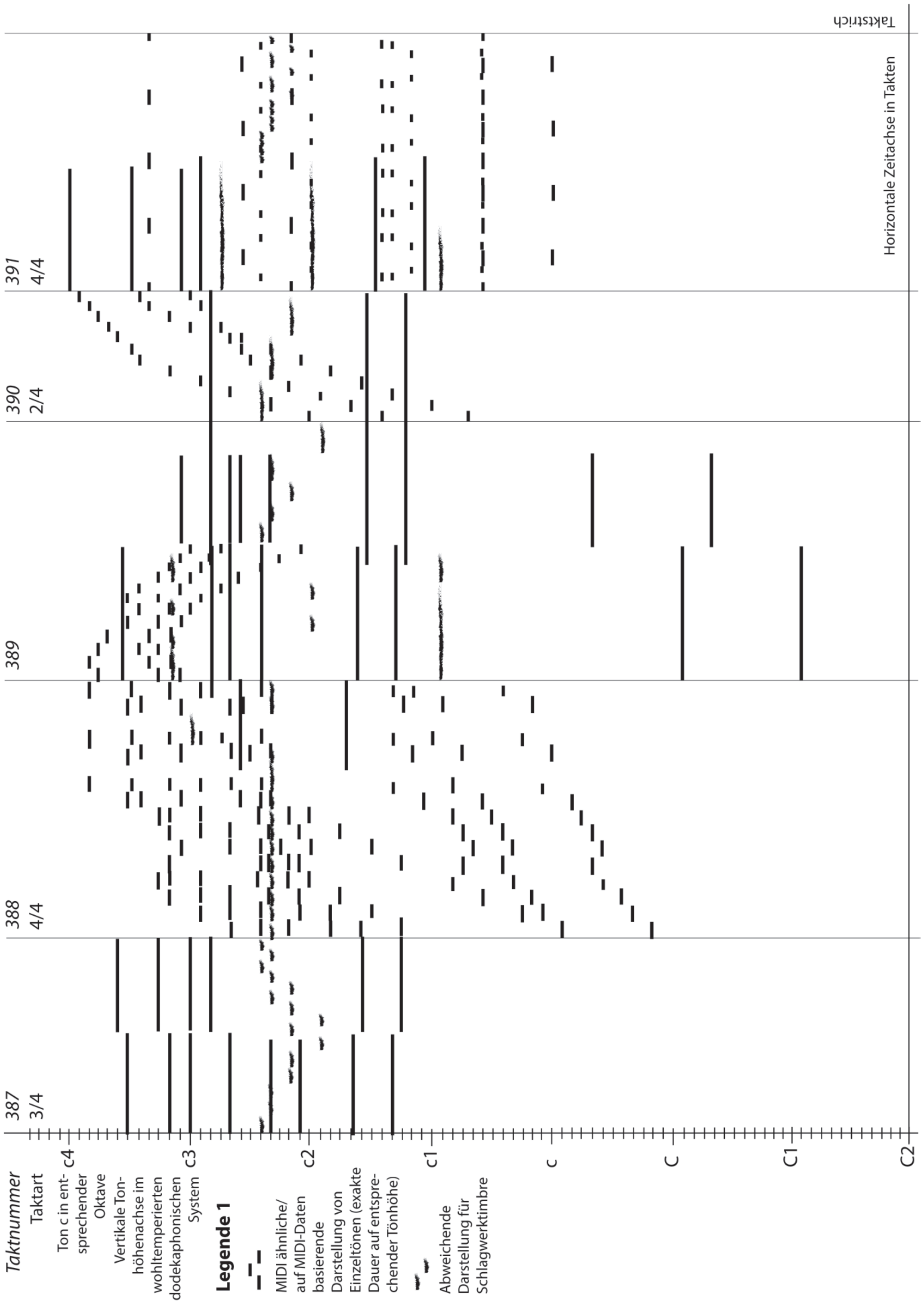


Abbildung 4: Erkki-Sven Tüür, *Oxymoron*, T. 387–391 in MIDI-ähnlicher Visualisierung (G. Lock).

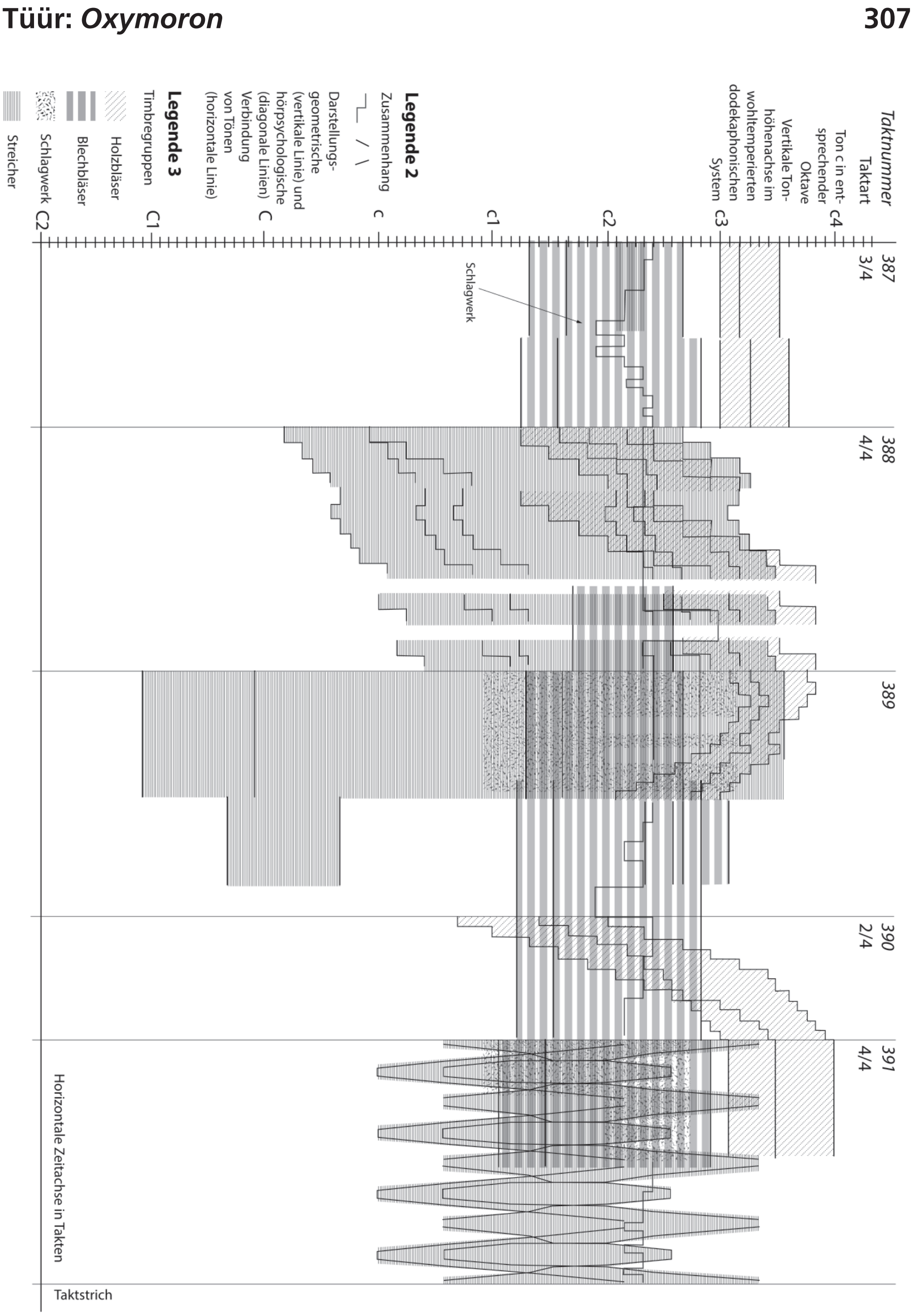


Abbildung 5: Erkki-Sven Tüür, *Oxymoron*, T. 387–391 als Timbregruppen-Analyse, im Hintergrund ist die darstellungsgeometrische Visualisierung der Tonhöhen aller Stimmen zur Orientierung belassen (G. Lock). Die Timbregruppen-Analyse geht auf Ron Weidberg[66] zurück und ermöglicht die visuell-zusammenfassende und räumlich gleichzeitige Darstellung aller in der Partitur (bzw. durch MIDI) sichtbarer Stimmen entsprechend der Timbre- (bzw. Instrumentations-) Gruppen.

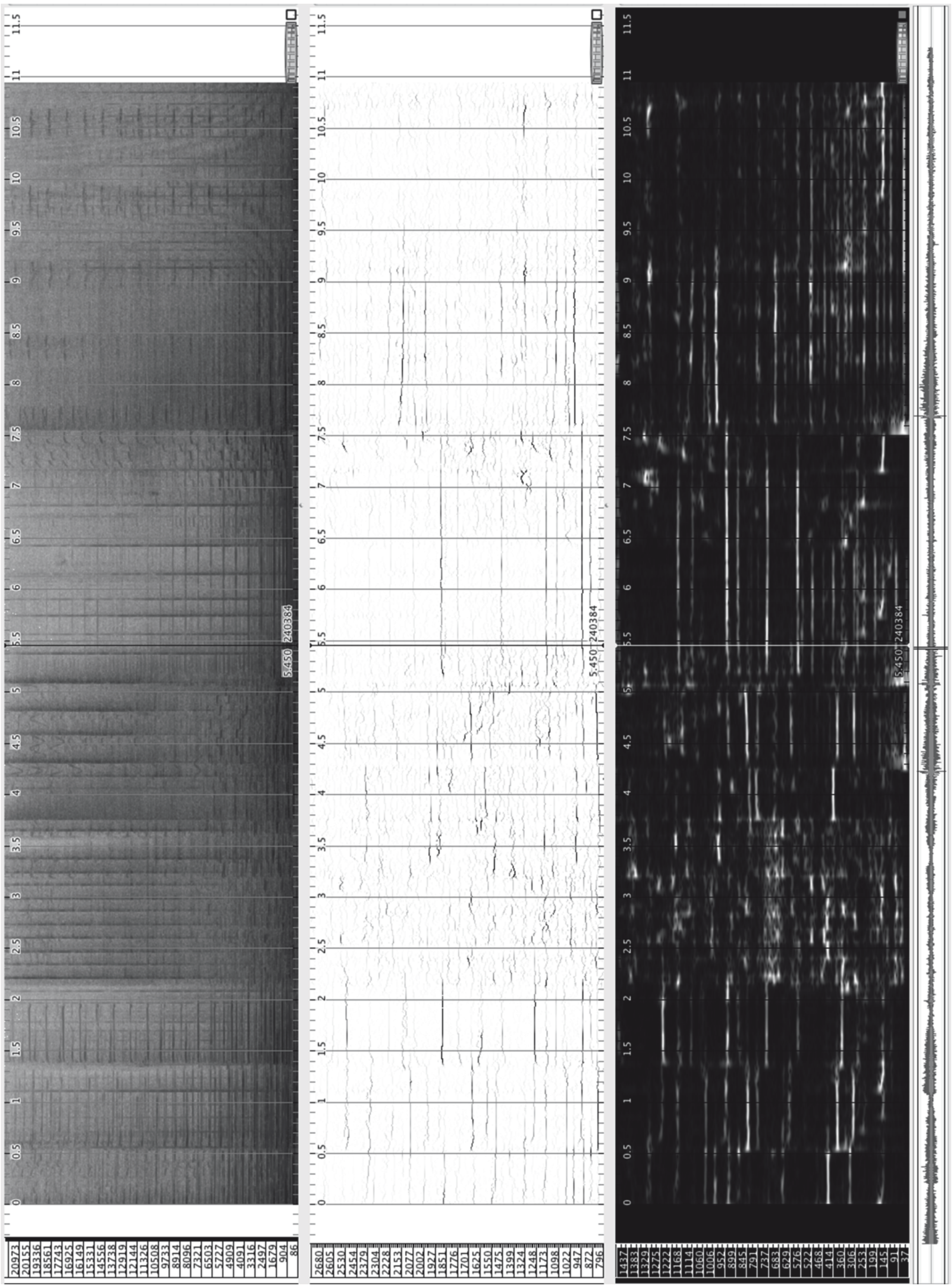

0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
7
7.5
8
8.5
9
9.5
10
10.5
11
11.5
5.450
240384
20973
20155
19336
18561
17743
16925
16149
15331
14556
13738
12919
12144
11326
10508
9733
8914
8096
7321
6503
5727
4909
4091
3316
2497
1679
904
86
2680
2605
2530
2454
2379
2304
2228
2153
2077
2002
1927
1851
1776
1701
1625
1550
1475
1399
1324
1248
1173
1098
1022
947
872
796
1437
1383
1329
1275
1222
1168
1114
1060
1006
952
899
845
791
737
683
629
576
522
468
414
360
306
253
199
145
91
37

Abbildung 6 (links): Erkki-Sven Tüür, *Oxymoron*, T. 387–391, Abschnitte 17:29–17:39 des gesamten Stückes, analysiert mit Sonic Visualizer.[67] Das obere Panel zeigt das vollständige Sonogramm (bzw. die Spektralanalyse) des Abschnitts (schwarz auf weiß, Skalierung in dBV, sichtbarer Bereich normalisiert, Fourier-Fenster-Größe 1024 mit 93,75%, alle Komponenten, linear). Das mittlere Panel zeigt das selbe Zeitfenster in Peak-Frequency (bzw. Spitzenfrequenz-) Darstellung (schwarz auf weiß Skalierung linear, vertikale Achse normalisiert, Fourier-Fenster-Größe 4096 mit 93,75%, Frequence Komponenten, linear). Das untere Panel zeigt den selben Abschnitt als Melodic Range (melodische Komponenten-) Spektrogram (weiß auf schwarz, linear, Fourier-Fenster-Größe 8192 mit 93,75%, alle Komponenten, linear). Ganz unten ist die Wellenform (zusammengefasst in einem Kanal) des Abschnitts abgebildet. Erstellt mit der Aufnahme des Nyyd-Ensemble unter Leitung von Olari Elts (ECM Records, München 2007) von Erkki-Sven Tüür's »Oxymoron«.

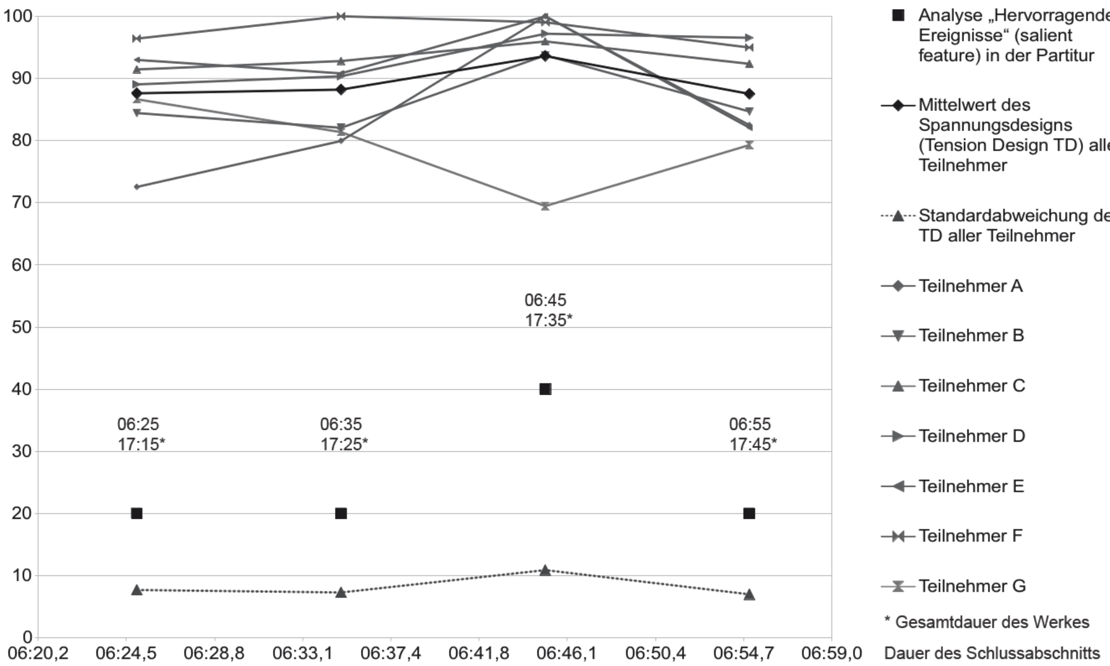


Abbildung 7: Erkki-Sven Tüür, *Oxymoron*, T. 387–391, Abschnitt 17:25–17:45 des gesamten Stückes. Einzelpunkte zeigen die Analyse »hervorragender Ereignisse« (salient features) in der Partitur[68], die durchgängig grauen Diagrammlinien zeigen Ergebnisse von sieben Teilnehmern eines Wahrnehmungsexperiments (normalisiert und geglättet), bei der mit Sliderkontrollern der Spannungsverlauf anhand der Aufnahme in Echtzeit fixiert wurde. Die gestrichelte Diagrammlinie im unteren Teil zeigt die Standardabweichung aller Teilnehmer. Dies zeigt, dass die Teilnehmer des Experiments diesen Hauptkulminationspunkt des Stückes relativ eindeutig (5 von 7 Teilnehmer) als höchsten Punkt anzeigen. Dies fällt mit der Partituranalyse zusammen.[69]

Anmerkungen

1 Der Artikel ist entstanden im Rahmen des von der Estnischen Forschungsagentur (ETAg) finanzierten Forschungsstipendiums »Funktionale Aspekte von Musik« von Kerri Kotta, Mart Humal und Gerhard Lock, 2010–2012. Der Autor dankt Kerri Kotta und Edward Venn (Universität Leeds, GB) für wertvolle Hinweise und Ratschläge.

2 G. Lock, *Aspekte der Visualisierung und graphischen Analyse von Musik. Die Fünfte Sinfonie von Lepo Sumera*, in: *Musical Work: Boundaries and Interpretaions (38th Baltic Musicological Conference)*, hrsg. von A. Žiuraityte, Vilnius 2006, S. 243–269.

3 G. Lock, *Some Preliminary Concepts of Graphical Analysis/Composition for Musicians and Nonmusicians*, in: *Music and Environment*, hrsg. von T. Selke und G. Lock), Tallinn 2009, S. 152–173.

4 K.-J. Sachs, *Komposition*, in: MGG2, Bd. 5 (Sachteil), Kassel u.a. 1996, S. 506–507.

5 E. Karkoschka, *Das Schriftbild der Neuen Musik*, Celle 1966, S. 1.

6 Siehe G. Daurer, *Audiovisuelle Wahrnehmung*, in: *See This Sound – Audiovisuology Compendium*, hrsg. von D. Daniels und S. Naumann, Köln 2009, www.see-this-sound.at/kompendium/abstract/33 [05.01.2013]; D. Melcher und M. Zampini, *The Sight and Sound of Music: Audiovisual Interactions in Science and the Arts«*, in: *Art and the Senses*, hrsg. von F. Bacci und D. Melcher, Oxford 2011, S. 265–292; P. Vergo, *The Music of Painting. Music, Modernism and the Visual Arts from the Romantics to John Cage*, London 2010.

7 K. Stockhausen, *Musik und Graphik*, in: ders., *Texte zur elektronischen und instrumentalen Musik*, Köln 1963, Bd. 1, S. 180–181.

8 G. Ligeti, *Über Form in der Neuen Musik*, in: ders., *Form in der Neuen Musik*, Mainz 1966, S. 23.

9 Lock, *Some preliminary concepts*; J. Jewanski / S. Naumann, *Strukturanalogien zwischen Musik und visuellen Künsten*, in: *See This Sound – Audiovisuology Compendium*, hrsg. von D. Daniels und S. Naumann, Köln 2009, http://www.see-this-sound.at/kompendium/abstract/68 [05.01.2013]; T. Goldberg / G. Schrack, *Computer-Aided Correlation of Musical and Visual Structures*, in: Leonardo 19/1 (1986), S. 11–17.

10 Lock, *Some preliminary concepts*, S. 155; siehe auch ders., *Musical Creativity in the Mirror of Glaveanu's Five Principles of Cultural Psychology*, in: Culture & Psychology 17/1 (2011), S. 121–136, hier S. 125.

11 G. Scholz (Hrsg.), *Pluralismus analytischer Methoden* (Publikationen des Instituts für Musikanalytik Wien Bd. 3), Frankfurt 1996, S. 11.

12 D. Sedivy, *Pluralistlikust Muusikaanalüüsist*, in: Teater.Muusika.Kino 12 (2004), S. 82ff.

13 A. H. Zalanowski, *Music Appreciation and Hemisphere Orientation: Visual versus Verbal Involvment*, in: Journal of Research in Music Education 38/3 (1990), S. 197–205.

14 B. Kollöffel, *Exploring the Relation Between Visualizer-Verbalizer Cognitive Styles and Performance with Visual or Verbal Learning Laterial*, in: Computers & Education 58/2 (2012), S. 697–706; M Kozhevnikov / M. Hegarty / R. E. Mayer, *Revising the Visualizer–Verbalizer Dimension: Evidence for Two Types of Visualizers*, in: Cognition and Instruction 20/1 (2002), S. 47–77.

15 B. Almén, *Modes of Analysis*, in: Theory and Practice. Journal of the Music Theory Society of New York State 31 (2006), S. 1–38.

16 S.-L. Tan / E. M. Wakefield / P. W. Jeffries, *Musically Untrained College Students' Interpretations of Musical Notation: Sound, Silence, Loudness, Duration, and Temporal Order*, in: Psychology of Music 37/1 (2009), S. 5–24.

17 R. Donnini, *The Visualization of Music: Symmetry and Asymmetry*, in: Computers & Mathematics with Applications 12B/1–2 (1986), S. 435.

18 Für eine Übersicht siehe N. Goodman, *Languages of Art: An Approach to a Theory of Symbols*, Indianapolis ^{2}1976.

19 Ch. O. Nussbaum, *The Musical Representation. Meaning, Ontology, and Emotion*, Cambridge, Mass. 2007, S. 21.

20 F. Lerdahl, *Tonal Pitch Space*, New York 2001, S. 143.

21 F. Lerdahl / R. Jackendoff, *A Generative Theory of Tonal Music*, Cambridge, Mass. 1983.

22 Lock, *Some preliminary concepts*, S. 153–154.

23 R. Arnheim, *Art and Visual Perception: A Psychology of The Creative Eye*, Berkeley 1974; A.S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, Cambridge, Mass. 1990.

24 W. Gieseler, *Komposition im 20. Jahrhundert*, Celle 1975, S. 39ff.

25 Lock, *Some preliminary concepts*, S. 154.

26 Siehe dazu den Beitrag von Martin Eybl in diesem Band. Siehe auch O. Schwab-Felisch, *Zur Schichtenlehre Heinrich Schenkers*, in: *Musiktheorie* (Handbuch der Systematischen Musikwissenschaft 2), hrsg. von H. de la Motte-Haber und O. Schwab-Felisch, Laaber, 2005, S. 337–376; D. Neumeyer, *Themes and Lines: On the Question of Hierarchy in the Practice of Linear Analysis*, in: Res Musica 3 (2011), 9–27, www.muusikateadus.ee/resmusica/res-musica-3-2011/#more-610; Ildar D. Khannanov, *Hierarchical Structure in Music Theory before Schenker*, in: Res Musica 3 (2011), S. 92–109, www.muusikateadus.ee/resmusica/res-musica-3-2011/#more-610

27 T. Mäkelä, *Klang und Linie von Pierrot lunaire zu Ionisation. Studien zur funktionalen Wechselwirkung von Spezialensemble, Formfindung und Klangfarbenpolyphonie*, Frankfurt a.M. 2004.

28 Siehe dazu den Beitrag von Stephan Lewandowski in diesem Band.

29 J. Leonhardt / J. Willenbacher, *Abbildungsgeometrische Analyse musikalischer Formen*, in: *Geometrie, Architektur und Musik*, hrsg. von C. Leopold und N. Christmann, Kaiserslautern 2003, S. 63–76.

30 Panos Vlagopoulos, *Is It All Autographic? Samples from the Musical Avant-Garde of the 60s*, in: *Proceedings of the 12th International Conference on Music Perception and Cognition (ICMPC) / 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*, hrsg. von E. Cambouropoulos, C. Tsougras, P. Mavromatis und K. Pastiadis, Thessaloniki 2012, S. 1081–1084, http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/1081_Proc.pdf [05.01.2013].

31 H. Honing, *Issues on the representation of time and structure in music*, in: Contemporary Music Review 9/1–2 (1993), S. 225.

32 Karkoschka, *Das Schriftbild*, S. 19.

33 C. Wilmer, *Color-Encoded Music Scores: What Visual Communication Can Do for Music Reading*, in: Leonardo 28/2 (1995), S. 129–136; M. Kaikkonen / K. Uusitalo, *Figurenotes, Special Music Centre Resonaari*, www.resonaari.fi [05.01.2013].

34 D. Raaijmakers, *Cahier »M«. A Brief Morphology of Electric Sound* (Collected Writings of the Orpheus Institute 3), Ghent 2005, S. 47–50.

35 K. Stockhausen, *Nr. 3 Elektronische Studien. Studie II*, Wien (Universal Edition) 1956.

36 R. Elkoshi, *Interpreting Children's Invented Graphic Notation*, in: Arts and Learning Research 20/1 (2003/2004), S. 61–84.

37 H. Mühe, *Musikanalyse*, Leipzig 1978; L. Thoresen, *Formbuilding Transformations: an Approach to Aural Analysis of Emergent Musical Forms*, in: The Journal of Music and Meaning 4 (2007), www.musicandmeaning.net/issues/showArticle.php?artID=4.3 [05.01.2013].

38 A. Schönberg, *Fundamentals of Musical Composition*, London 1976, S. 113–115.

39 N. Chr. Hansen, *The Legacy of Lerdahl and Jackendoff's* A Generative Theory of Tonal Music. *Bridging a significant event in the history of music theory and recent developments in cognitive music research*, in: Danish Yearbook of Musicology 38 (2010/11), S. 33–55; M Hamanaka / K. Hirata / S. Tojo, *Automatic generation of metrical structure based on GTTM*, in: *Proceedings of the 2005 International Computer Music conference (ICMC)*, Barcelona 2005; dies., *ATTA: Automatic Time-Span Analyzer based on extended GTTM*, in: ebenda, S. 358–65.

40 D. Temperley, *The Cognition of Basic Musical Structures*, Cambridge, Mass. 2001.

41 A. Pinto, *A Dynamical Approach to Music Analysis*, in: Journal of Mathematics and Music 6/3 (2012), S. 207–219.

42 Stefanie Reiter (Anzenhofer), *Musikalische Graphen – Entwicklung eines Verständnisses graphischer Darstellungen im fächerübergreifenden Mathematik- und Musikunterricht*, www.didaktik.mathematik.uni-wuerzburg.de/mitarbeiter/anzenhofer/musikalische_graphen/, http://www.stefanie-reiter.de/index.php?clicked=musikalischegraphen [05.01.2013].

43 S. Abdallah / M. Plumbley, *Information Dynamics: Patterns of Expectation and Surprise in the Perception of Music*, in: Connection Science 21/2–3 (2009), S. 89–117; H. Sabelli / A. Lawandow, *Asymmetry, Symmetry and Beauty*, in: Symmetry 2 (2010), S. 1591–1624.

44 A. Brandmeyer / R. Timmers / M. Sadakata / P- Desain, *Learning Expressive Percussion Performance Under Different Visual Feedback Conditions*, in: Psychological Research 2010, http://link.springer.com/article/10.1007/s00426-010-0291-6/fulltext.html [05.01.2013].

45 P. Toiviainen, *Visualization of Tonal Content in the Symbolic and Audio Domains*, in: *Tonal Theory for the Digital Age* (Computing in Musicology 15), hrsg. von W. B. Hewlett und E. Selfridge-Field, Cambridge, Mass. 2008.

46 D. Huron, *Music Research Using Humdrum: A User's Guide*. Stanford 1999 online: www.musiccog.ohio-state.edu/Humdrum/guide.toc.html, Kern-Scores Webseite: http://kern.ccarh.org [05.01.2013].

47 A. Baratè / G. Haus / L. A. Ludovico, *Music Representation of Score, Sound, MIDI, Structure and Metadata All Integrated in A Single Multilayer Environment Based on XML*, in: *Intelligent Music Information Systems: Tools and Methodologies*, hrsg. von J. Shen, J. Shepherd und L. Liu, IGI Global 2008, S. 298–321; K. Lassfolk, *Music Notation as Objects. An Object-Oriented Analysis of the Common Western Music Notation System* (Acta Semiotica Fennica XIX, Approaches to Musical Semiotics 7, Studia Musicologica Universitatis Helsingiensis XI), Helsinki 2004; K. Renz, *Algorithms and Data Structures for a Music Notation System based on GUIDO Music Notation* (Darmstädter Dissertationen 17), Darmstadt 2002.

48 H. Gottschewski, *Graphische Darstellung musikalischer Zeit*, in: *KlangArt-Kongreß 1993: Neue Musiktechnologie II*, hrsg. von Bernd Enders, Mainz 1996, S. 230–253. Elektronische Fassung, EpOsMusic Osnabrück, 2000. Online-Version http://www.epos.uos.de/music/templates/buch.php?id=13&page=/books/e/endb_96/pages/index.htm [05.01.2013].

49 A. Volk, *Die analytische Interpretation. Schritte zur Erschließung eines Forschungsfeldes am Beispiel der Metrik*, Diss. Berlin, Humboldt-Universität.

Webseite von RUBATO, entwickelt von Guerino Mazzola, www.rubato.org [05.01.2013].

50 S. Malinowski, *Music Animation Machine MIDI Player*, www.musanim.com/player/MAMPlayerUserGuide.pdf [05.01.2013].

51 Gottschewski, *Graphische Darstellung*.

52 S.S. Elmer, F.-J. Elmer, *A New Method for Analysing and Representing Singing*, in: Psychology of Music 28 (2000) S. 23–42, http://mmatools.sourceforge.net [05.01.2013].

53 T. Eerola / P. Toiviainen, *MIDI Toolbox: MATLAB Tools for Music Research*, Kopijyvä und Jyväskylä 2004, https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials [05.01.2013].

54 H. Lohner, *Das UPIC: eine Erfindung von Iannis Xenakis*, in: *Iannis Xenakis* (Musik-Konzepte 54/55), hrsg. von H.-K. Metzger und R. Riehn, München 1987, S. 71–82.

55 J.-B. Thiebaut / P.G.T. Healey / N. B. Kinns, *Drawing Electroacoustic Music*, in: *Proceedings of the International Computer Music Association ICMC*, Belfast 2008.

56 P. Toiviainen / B. Burger, *MoCap Toolbox Manual*, Jyväskylä 2011, http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual. [05.01.2013].

57 L. Nijs, *The Music Paint Machine. An Embodied Constructivist Approach to Technology-Enhanced Instrumental Music Instruction*, Diss., Ghent 2012.

58 Siehe z. B. Karkoschka, *Neue Methoden*; M. Kirschbaum, *Höhepunktbildung und Dramaturgie in Neuer Musik. Die Erstellung »formaler Spannungsverläufe« als Ergänzung der musikalischen Analyse*, Köln 2001; A.R. Addessi / R. Caterina, *Perceptual Musical Analysis: Segmentation and Perception of Tension*, in: Musicae Scientae 4/1 (2000), S. 31–54; G. Lock / M. Valk-Falk, *Investigating Structure of Contemporary Music Applying Tension Design and Empirical Perception Analysis*, in: *Proceedings of the Fourth Conference on Interdisciplinary Musicology (CIM08)*, hrsg. von C. Tsougras und R. Parncutt, Thessaloniki 2008, S. 1–15, online-Version: http://cim08.web.auth.gr/cim08_papers/Lock-Valk-Valk/Lock-ValkFalk.pdf [05.01.2013]; G. Lock / K. Kotta, *Musical tension as a response to musical form*, in: *Proceedings of the 12th International Conference on Music Perception and Cognition (ICMPC) / 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*, hrsg. von E. Cambouropoulos, C. Tsougras, P. Mavromatis und K. Pastiadis, Thessaloniki 2012, S. 612–617, Online-Version: http://icmpc- escom2012.web.auth.gr/sites/default/files/papers/612_Proc.pdf [05.01.2013]; B.W. Vines / R. L. Nuzzo / D. Levitin, *Analyzing Temporal Dynamics in Music: Differential Calculus, Physics, and Functional Data Analysis Techniques*, in: Music Perception 23/2 (2005), S. 137–152.

59 H. Jenny, *Kymatik. Wellenphänomene und Schwingungen*, Baden 2009.

60 H. Leichtentritt, *Musikalische Formenlehre*, Wiesbaden 1907.

61 H. Mersmann, *Angewandte Musikästhetik*, Berlin 1926.

62 U. Kaiser, *Gehörbildung. Satzlehre, Improvisation, Höranalyse*, Kassel 1998.

63 Lock, *Some preliminary concepts*.

64 Kaikonen /Uusitalo, *Figurenotes*.

65 R. Elkoshi, *An Investigation into Children's Responses Through Drawing to Short Musical Fragments and Complete Compositions*, in: Music Education Research 4/2 (2002), S. 199–211; S.-L. Tan / M.E. Kelly, *Graphic Representations of Short Musical Compositions*, in: Psychology of Music 32/2 (2004), S. 191–212; L. Verschaffel / M. Reybrouck / M. Janssens / W. van Dooren, *Using Graphical Notations to Assess Children's Experiencing of Simple and Complex Musical Fragments*, in: Psychology of Music 38/3 (2010), S. 259–284; M. B. Küssner / H.M. Prior / N. E. Gold / D. Leech-Wilkinson, *Getting the Shapes »Right« at the Expense of Creativity? How Musicians' and Non-Musicians' Visualizations of Sound Differ*, in: *Proceedings of the 12th International Conference on Music Perception and Cognition (ICMPC) / 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*, hrsg. von E. Cambouropoulos, C. Tsougras, P. Mavromatis und K. Pastiadis, Thessaloniki 2012, S. 561, online-Version: http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/561_Proc.pdf [05.01.2013].

66 R. Weidberg, *Sonic Design in Jean Sibelius' Orchestral Music*, in: *Sibelius Forum II. Proceedings from The Third International Jean Sibelius Conference, Helsinki, December 7–10, 2000*. Helsinki 2003, S. 216–226.

67 Sonic Visualizer, freie Software der Queen Mary University of London, www.sonicvisualiser.org. Chris Cannam, Christian Landone, and Mark Sandler, Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files, in Proceedings of the ACM Multimedia 2010 International Conference.

68 Siehe Lock/Kotta, *Musical tension*.

69 Siehe G. Lock, *Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses »Oxymoron«* [Aspekte, die die Wahrnehmung musikalischer Spannung in Erkki-Sven Tüür's »Oxymoron« beeinflussen], in: Res musica 2 (2010), S. 63–74.

Kommentierte Auswahlbibliographie

Einführende Publikationen

Lock, Gerhard, *Some Preliminary Concepts of Graphical Analysis/Composition for Musicians and Nonmusicians*, in: *Music and Environment*, hrsg. von T. Selke und G. Lock, Tallinn 2009, S. 152–173.
Enthält neben einer Übersicht zu grundlegenden Hintergründen und Analogien zwischen Musik und Graphik u. a. auch eine auf einfachen graphischen Figuren basierende mit Hilfe des Programms *Coagula* automatisierbare Analyse- und Kompositionsmethode.

Lock, Gerhard, *Aspekte der Visualisierung und graphischen Analyse von Musik. Die Fünfte Sinfonie von Lepo Sumera*, in: *Musical Work: Bounderies and Interpretaions* (38th Baltic Musicological Conference), hrsg. von A. Žiuraityte, Vilnius 2006, S. 243–269.
Grundlage des vorliegenden Artikels. Enthält neben Sonic-Design-ähnlichen Analysen erstmals auch die theoretische Kategorisierung von graphischen Analysen in intuitive, systematische, system-basierende und automatische Methoden.

Daurer, Gerhard, *Audiovisuelle Wahrnehmung*, in: *See This Sound – Audiovisuology Compendium*, hrsg. von D. Daniels und S. Naumann, Köln 2009, www.see-this-sound.at/kompendium/abstract/33 [05.01.2013]
Artikel eines umfangreichen Internetportals zu audiovisuellen Phänomenen. Hier werden u.a. wahrnehmungspsychologische Prinzipien des Hörens und des Sehens im Kontext des Zusammenwirkens und der Singularität der Sinne, wahrnehmungsintensivierende Aspekte sowie Verknüpfungsmöglichkeiten der Ebenen durch synästhetisches Design dargestellt.

Jewanski, Jörg und Sandra Naumann, *Strukturanalogien zwischen Musik und visuellen Künsten*, in: *See This Sound – Audiovisuology Compendium*, hrsg. von D. Daniels und S. Naumann, Köln 2009, http://www.see-this-sound.at/kompendium/abstract/68 [05.01.2013]
Artikel eines umfangreichen Internetportals zu audiovisuellen Phänomenen. Hier werden u. a. visuelle Analogien zur Musik in ihren zeitlichen Abläufen, Beziehungen zwischen den Einzelelementen, in Kompositionstechniken und Improvisation sowie musikalische Analogien zur bildenden Kunst in Bezug auf Räumlichkeit und Farblichkeit dargestellt.

Honing, Henkjan, *Issues on the Representation of Time and Structure in Music*, in: Contemporary Music Review 9/1–2, (1993), S. 221–238.
Grundlegender Artikel zur Repräsentation von Zeit und Struktur von Musik und ihren kognitiven Grundlagen.

Verschiedene Methoden in einer Publikation

Wißkirchen, Hubert, *Arbeitsbuch für den Musikunterricht in der Oberstufe. Band 2: Struktur- und Formanalyse*, Frankfurt a.M. 1992.
Enthält zahlreiche unterschiedliche graphische Analysemethoden, darunter Hörpartitur und Darstellungsgeometrie.

Mühe, Hansgeorg, *Musikanalyse*, Leipzig 1978.
Enthält eine auf einfachen geometrischen Figuren aufgebaute formanalytische Methode, eine dem darstellungsgeometrischen Prinzip ähnliche Methode der Schichtungsanalyse, verschiedene graphische Darstellungen von Melodik (intuitiv-systematisch) und Dynamik (system-basierend).

Karkoschka, Erhard, *Neue Methoden der musikalischen Analyse und einige Anwendungen auf spätere Instrumentalmusik von Karl Marx (1967)*, in: *Zur musikalischen Analyse* (Wege der Forschung 257), hrsg. von G. Schumacher, Darmstadt 1974, S. 644–664.
Enthält Klangspannungs-, Intervallschrittgröße, Impulsdichte und durchschrittener-Raum-Darstellungen.

Mailman, Joshua B., *Temporal Dynamic Form in Music: Atonal, Tonal, and Other*, Diss. University of Rochester 2010, Online: www.joshuabanksmailman.com [05.01.2013].
Enthält zahlreiche Beispiele unterschiedlichster Visualisierungsmöglichkeiten von Musik.

Notation, Höranalyse, Hörpartitur

Goodman, Nelson, *Languages of Art: An Approach to a Theory of Symbols*, Indianapolis ²1976.
Grundlegende Publikation zu Prinzipien von Notation im Kontext von Visualität und Klang. Enthält Kapitel wie *The Sound of Pictures*, *The Theory of Notation*, *Score*, *Sketch* und *Script*.

Schröder, Julia H., ***Grafische Notation und musikalische Grafik*****, in:** ***See This Sound – Audiovisuology Compendium*****, hrsg. von D. Daniels und S. Naumann, Köln 2009, www.see-this-sound.at/kompendium/abstract/78** [05.01.2013]
Artikel eines umfangreichen Internetportals zu audiovisuellen Phänomenen. Er bietet eine Übersicht zur Entwicklung graphischer Notationen und musikalischer Graphiken und definiert sowohl musikalische graphische Notation (als freie, vieldeutige Notationsart, deren Zeichen einer Erklärung bedürfen) als auch musikalische Graphik (ästhetischer Eigenwert als visuelle Kunst).

Treibmann, Karl O., ***Strukturen in neuer Musik. Anregungen zum zeitgenössischen Tonsatz*****, Leipzig 1981.**
Bietet die Hörpartitur als Analysemethode an.

Melodie-Konturen-Analyse

Morris, Robert D., ***New Directions in the Theory and Analysis of Musical Contour*****, in: Music Theory Spectrum 15/ 2 (1993), S. 205–228.**
Umfangreicher Artikel zur Theorie und Analyse von musikalischen Konturen.

Adams, Charles R., ***Melodic Contour Typology*****, in: Ethnomusicology 20/2 (1976), S. 179–215.**
Typologie der Melodie-Konturen-Analyse basierend auf ethnomusikalischen Methoden.

Sonic Design und ähnliche Darstellungsweisen

Cogan, Robert und Pozzi Escot, ***Sonic Design: The Nature of Sound and Music*****, Englewood Cliffs 1976.**
Grundlegendes Lehrbuch zu Sonic Design, welches auf MIDI-ähnlicher Darstellung beruht.

Sciarrino, Salvatore, ***Le figure della musica: da Beethoven a oggi*****, Mailand 1998.**
Sonic-Design/MIDI ähnliche handschriftliche Analysen (auch farbig).

Sonogramm- bzw. Spektralanalyse

Timmers, Renee und Makiko Sadakata, *Training Expressive Performance by Means of Visual Feedback: Existing and Potential Applications of Performance Measurement Techniques*, in: *Expressiveness in Music Performances: Empirical Approaches across Styles and Cultures*, hrsg. von D. Fabian, R. Timmers und E. Schubert, Oxford i.V.
Enthält eine Übersicht zu verschiedenen Spektralanalyseprogrammen.

Clarke, Eric und Nicholas Cook, *Empirical Musicology: Aims, Methods, Prospects*, Oxford 2004.
Beinhaltet neben der Darstellung von Spektralanalyse-Methoden im Kapitel *Analyzing Musical Sound* auch eine Übersicht über verschiedene Visualisierungsmöglichkeiten und eine Fallstudie zu David Hurons HumDrum-Analysemethode im Kapitel *Computational and Comparative Musicology.*

A.4 Publication III (Lock 2017a, 5.4)

“Models and scientific modeling in music context” (in English)

Models and modeling

Problems in Music Pedagogy, Vol. 16(2), 2017, 25- 45

MODELS AND SCIENTIFIC MODELLING IN MUSIC CONTEXT

Gerhard LOCK

Estonian Academy of Music and Theatre, Tallinn University, Estonia
email: gerhard.lock@tlu.ee

Abstract
The object of modelling in this article music can be understood as natural environment (Reybrouck, 2015) and as dynamical system (Burrows, 1997). Scientific models have diverse functions (aims, purposes), but also constraints, they work with different methods. They derive from implicit and explicit knowledge, appear between cognitive experience and implicit and explicit data within the paradox of knowledge as 'world first' vs 'mind first', manifested in two extremes: existence and technique (Varto, 2009).

In the theoretical part, the author introduced his multi-dimensional scheme of modelling and models incorporating models from music theory (analysis) and music psychology (perception and cognition) including the creative approach (musical composition/improvisation). Its aim is to position modelling used in (and for understanding of) music within the context of modelling in science in general. Also, the relation of models to theory, as well as computational and cognitive models in music are discussed briefly.

In the main contribution, the author introduced his Twelve Strategic Steps for modeling/analysis of models (partly inspired by Purwins et al., 2008), presented his enlargement of Justi and Gilbert's (2002) "Model of Modelling" and compared it with the Twelve Strategic Steps. In the implications and conclusion, the author proposed that modelling in science results if an idea can become explained with a concept, which can then be approached systematically, further developed as a method, includes connections/backups to one or more theories, and is consistent (a whole), either representing or exploring a known (but less systematically explained) phenomenon including cross-domain mappings. The author hopes that this article enhances the understanding of modelling and assists in creating models in sciences in the field of music and music education.

Keywords: models in sciences, representation, music analysis and cognition, computation, education

Introduction

In science, the arts, and in education we often use the term 'model'. A common belief is that a scientific model represents something observable in the world (a 'phenomenon') or that a model is a theoretical/mental construct. The word 'model' in education, has come to represent or serve as an example of achievement, or something to replicate. In the arts, a model is a scheme or a description which is used to create a new artwork; also in music outstanding performances can become models informing further interpretations. Models and modelling play a crucial role in science and science education as well. 'Scientific modeling' is defined in the *Encyclopaedia Britannica* as *"generating a physical, conceptual, or mathematical representation of a real phenomenon that is difficult to observe directly. Scientific models are used to explain and predict the behavior of real objects or systems [...]"* (Rogers, 2017). In all these fields mentioned above, a model's purpose, form, and way of application differ. Models and modelling underlie a highly controversial discussion concerning their nature, function, relation to reality and theory, and their ability for scientific explanation, etc. (see more in Frigg & Hartmann, 2012). Interestingly, in the *Stanford Encyclopedia of Philosophy*, Frigg and Hartmann (2012) conclude their article *"Models in Science"* admitting that *"despite the fact that [models] have generated considerable interest among philosophers, there remain significant lacunas in our understanding of what models are and of how they work"*. In practice, however, models are both disbelieved and highly accepted as technical and pedagogical tools to gain knowledge and understand the world around us.

Approaches using models and modelling have been used for a long time in music creation (musical composition/improvisation) and music analysis (music theory) for both – pedagogical and academic – purposes (see Burkholder, 2017), a cross-domain approach (see Zbikowski, 1997, 2002). Since the rise and change of the empiricism paradigm (see Kuhn, 1962/1970; Bhattacherjee, 2012) during the 20th century models and modelling play a major role also in music psychology and cognition research (see Purwins et al., 2008). With the development of powerful computers (since the end of the 20th century) models and modelling are an undoubted part of computational approaches to comprehend music, its perception and creation (see e.g. Giomi & Ligabue, 1995; Marsden, 1995; Tiits, 1995; Vercoe, 1997; Wang, 2007). In education, particularly in the social sciences and science education, models and modelling have been researched (e.g. Lahti, 2013), as well as developed didactically (e.g. Cartier, Rudolph & Stewart, 2001; Justi & Gilbert, 2002; Gilbert, 2010). Here should also be mentioned publications dealing with modelling theory in science in general (e.g. Coll & Lajium 2011; Börner et al., 2012a, 2012b).

The main goal of this article is to present the Twelve Strategic Steps, which I posed as questions, which I have developed to create/analyse scientific models of phenomena or data based on their function and the method of how they work.[1] The additional goal is to enlarge Justi and Gilbert's (2002) "Model of Modelling" based on the Twelve Strategic Steps including 'input' and 'output' data, as well as cross-domain mapping aspects into their model. The pedagogical purpose of this article is to enable the reader to better understand modelling in science providing a practical tool to enhance this understanding. However, questions also appear concerning when or why an idea, concept, systematic approach, a method or a theory may turn into a model and when or why not. The Twelve Strategic Steps to create/analyse a model attempt to assist to answer these questions.

For a faster overview I have subdivided the Twelve Strategic Steps into eight stages/phases (these terms I have defined as follows: stage - rather spatial term, kind of plateau where something is gained, rather stable state or a kind of 'extended' arrival point; a phase - rather temporal term, something which lasts for a certain time, rather dynamic, developmental process or transition towards a certain stage, taking place in gradual stages, 'carrying out' something (Lock, 2016)):

a) definition of the phenomenon and the problem/aim;
b) prospective outcome (with or without hypothesis);
c) discussing 'input' data and retrieval of this data for the model;
d) ability of these data to model/research the phenomenon or its observer, or both;
e) in which stage/phase of the research process (and why) it is possible/necessary to apply modelling and what type of model makes sense to apply;
f) discussing 'output' data of the model;
g) how to treat these data further;
h) comparison achieved with data from other methods (or models), which opens the approach towards cross-domain understanding.

I hereby claim that modeling itself is already a cross-domain approach (see also Zbikowski, 1997, 2002), because of the use of different kinds of analogies (see Duit, 1991; Coll & Lajium, 2011), as well as techniques to visualize data and connections between the data and the phenomena.

---

[1] This article is part of my doctoral dissertation in progress forming a theoretical background for the Cognitive Octagonal Slice Model (COSM) developed to analyze contemporary orchestral music using music theory/analysis as well as empirical and cognitive methods – „Tension design. Methodological principles and analytical examples" (Estonian Academy of Music and Theatre). My acknowledgment for valuable content advice, terminological suggestions and language editing during writing this article goes to my supervisor music theorist, composer and educator prof. dr. Kerri Kotta (EAMT); music theorist, composer and online learning specialist dr. Paul Beaudoin (Brandeis University, USA); my colleague associate prof. of musics didactics dr. Tiina Selke (TLU BFM); the data analyst of Tallinn University School of Digital Technologies Ilkka Kosunen (MSc and PhD candidate at University of Helsinki); and my brother, musicologist, composer and music programming specialist Hans-Gunter Lock (MA, PhD candidate in composition at EAMT), lecturer at EAMT electronic music studio and Head of New Media Lab of the Estonian Academy of Arts (EKA).

Theoretical background

A. Music and models between natural environment and dynamical systems, experience and data, existence, and technique

The main object of this theoretical article on models and modelling is music as a temporal, social and cognitive phenomenon being created and perceived through a variety of human sensory organs. It can be understood and studied as natural (ecological) environment (Reybrouck, 2015), and as dynamical system (Burrows, 1997). The first approach leads us to 'music as a phenomenon', therefore to (representational) 'models of phenomena' (phenomenon as *"umbrella term covering all relatively stable and general features of the world that are interesting from a scientific point of view"* (Frigg & Hartmann, 2012)). The second approach describes the musical 'piece as a whole' that is *"projected from instant to instant throughout the course of a performance by a process of cognitive contextualization"* (Burrows, 1997, 530), which takes place in analogy to dynamical systems that are *"coherent processes constituted of the interaction of two or more components"* (Burrows, 1997, 530). To understand the complexity of such a dynamical system we may use (representational) 'models of data' (see Frigg & Hartmann, 2012).

As a general definition, I propose that models are a way to express *systematic thinking* and *understanding* in science, arts, and education. They are basic tools to be used to interpret *implicit* and *explicit* knowledge (see Dienes & Perner, 1999) gained from experiences**,** and from *implicit* and *explicit data*, to relate them to themselves and among each other, the research object, the social context (to both the individual and the research object) and the natural environment (lifeworld, *Lebenswelt*). The terms 'explicit data' and 'implicit data' can be found also in the computer science in the meaning of data conversion (see e.g. Byham, Hamilton & Guyer, 2017) and the distinguishing meaning in marketing (e.g. Bailey, 2017). The TechTarget website provides the following definition: *"Implicit data is information that is not provided intentionally but gathered from available data streams, either directly or through analysis of explicit data. Explicit data is information that is provided intentionally, for example through surveys and membership registration forms"* (Wigmore, 2017). Implicit data can be retrieved from the context in which explicit data is collected (e.g. as so called metadata in telecommunication or internet traffic). I agree with Hans-Gunter Lock that the generation of implicit data using the context (analyzing explicit data) and statistics works as a model itself.

How problematic knowledge and the understanding of modelling can be in the education field, has been found by Lahti (2013) in a study on cognitive development and students understanding of scientific models, underlining that there is *"a large gap existing between students' initial conceptions about models and science and full scientific conception, it may be ambitious to expect students to move completely to an expert conception in a short period of time such as a one-semester class"* (Lahti, 2013, 178). As to my knowledge, there is no comparable empirical study yet on the application of scientific modelling in music and music education.

For the creation, as well as the perception and understanding processes and conceptualizing of the environment or a dynamical system (constituting of components on different complexity levels), one uses cognitive models, which are not manifestations in themselves but *vehicles* to *imagine*, *verbalize*, *visualize*, *comprehend* and *communicate* ideas or concepts about the world (and music in the case of this article). Mental constructs can be part of these modeling processes (see Justi & Gilbert, 2002; Gilbert, 2010). Theoretical models can be built either as independent constructs based on logic (see Frigg & Hartmann, 2012), or they are the basis for musical composition (creation), or music theoretical systems of different musical styles (as detected by musical analysis). The latter means that a theory or a model (no matter which was first) as a product of the activity of the human mind may be the basis for a creation (musical composition) or its understanding (listening, analysis) and interpretation (performance). Purwins et al. (2008) points out that *"traditional music theory can be used as a guide, but not as a* normative *reference that dictates the kind of processes and structures that operate in the mind of the listener"* (p. 152). They underline, that traditional music theory searches for universal (and normative) laws of music, but these attempts face limitations in the plurality of styles – style-intern rules of one style may not fit rules in another style. For instance, style specific harmony rules in tonal music do not apply in post-tonal music, even if the basic perceptual mechanisms of simultaneously sounding pitches are the same. This means, we may find (and empirically test the perceiving of), for instance, consonant harmonies in atonal (or other post-tonal) music, but this do not fit to the music theoretical explanation of this style. Further, series (12-tone rows) in Schoenberg's (and others') dodecaphonic music, are almost impossible to perceptually trace correctly (as the composer has used them in the music), empirical tests should reveal rather other aspects of this musical style. According to Cross (1998), music analysis and the psychology of music are *"fundamentally incompatible"* (p. 4); Cross discusses here also Cook's (1994) criticism of 'theorism', underlining that *"an a priori reliance on music-theoretic concepts as constituting or corresponding directly to cognitive categories and concepts [...] maybe levelled at specific studies of music cognition, it is not a charge that can be levelled against all studies of music cognition denatura"* (Cross, 1998, 4). Cross calls this understanding of an 'analytical idea' of perception in music analysis also 'folk psychology' referring here to Brunner (1990): *"...a set of more-or-less normative descriptions about how human beings 'tick', what our own and other minds are like"* (p. 35). *"The act of perception is subject to purposive volitional intervention, and categories of experience – qualia – may in the light of the analyst's enquiries be consciously shaped and re-made (for the analyst, at least)"* (Cross, 1998, 4-5).

Because a performed (sounding) musical composition becomes part of our life world (natural environment), even constituting a sounding environment itself or is understood as dynamical system, providing information (experiences turned into knowledge) through perception, a general epistemological question in science/philosophy and the arts (appearing also in education) can be posed: *„Which comes first in knowledge, the mind or the world, and what is their relative significance?"* (Varto, 2009, 8). What follows are, according to Varto (2009: 11), two extremes: *existence* and *technique* (or, in the view of Varto, the *ontic* and the *epistemological*). Varto claims that *„it is the technique that makes the ontic ontological: provides an event with order from some selected point of view"* (Varto, 2009, 11).

This epistemological question appears also in the different understandings of music: either as natural environment – the focus is on its *existence* (including its function), or as dynamical system – the focus is on the *technique* (including the method of how it works). It is mirrored in Cartier, Rudolph and Stewart's (2001) discussion of the explanatory roll of scientific models. They point out that *"models are constituted by* *empirical or theoretical objects and the processes in which they participate"* and *"models are consistently assessed on the basis of empirical and conceptual criteria"* (pp. 4-6).

This is visualized in their figure 3 (Cartier, Rudolph & Stewart, 2001, 6), which I have redrawn and extended based on their claims here in Figure 1. The authors specify the arrows of their scheme as follows: Generally, explanatory models are assessed based upon whether they can (1) explain patterns in data; (2) correctly predict the results of new experiments or observations; and (3) are consistent with other ideas (models, beliefs, and metaphysical commitments).

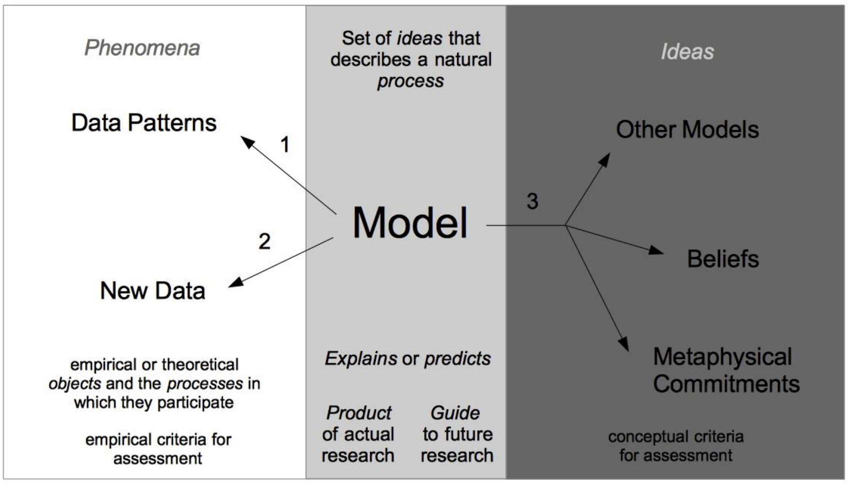


Figure 1. The nature, structure and assessment of explanatory models

To summarize my previous points, I suggest that music can be seen twofold: As empirical object – treated as natural environment, based on its existence how it is experienced and what is its function – Varto's "world-first" view; or as a theoretical object – treated as dynamical system, based on the technique and the method how it is created and/or perceived/cognized/understood – Varto's "mind first" view. But, according to Varto (2009, 8-9), both assumptions remain problematic: the "world first" supposition causes the question *Why is so much fallacy and fantasy in the world?* The "mind first" supposition poses the question *How does the mind come about, and how can it be valid in the world and create intersubjective science?* But, in both cases – especially expressed in modelling – the *processes* in which the music, the music maker, the listener and the analyst/model user or creator are involved, are crucial.

B. Functions and methods of scientific models

Coll and Lajium (2011) point out that models have a key role in scientific enquiry offering four distinct functions (aims, purposes): *discovery*, *development*, *evaluation* and *exposition*. Discovery means the formation/generation of new knowledge or the formulation of a hypothesis – using the model to further develop a hypothesis or discovery – either theoretical or experimental. A model also needs to be tested to evaluate a hypothesis (pro or con). Exposition means to explain hypotheses or theories through a model. Coll and Lajium (2011) describe that models based on analogies are useful in concept generation, a *"personal mental process"*, *"a vital stage of a scientists'* *thinking", "mapping of attributes from the known to the unknown"*, as well as *"generative models are dynamic tools rather than static representations used for understanding" – "individuals seek to find ways of making it easier to explain observed phenomena and so develop their own models to advance their understanding"* (pp. 8-9).

For this article, I have developed a multi-dimensional scheme of modelling and models in music incorporating chosen types of models from music theory (analysis) and music psychology (perception and cognition) including also the creative approach (musical composition/improvisation) as well as the educational functionality (at different school levels), and the way(s) in which each of these types in dependence on the purpose of the model is used or developed for and with learners. These models deal with the phenomena of music and sound (their understanding and perception/cognition) directly, not with sociological opinions, beliefs or paradigms about music and sound. However, models in (developmental) psychology, sociology, identity, personality, motivation, creativity etc. may also fall into the function and method categories introduced below, but their structure, purpose and way of application may differ.

Before I show and explain my multi-dimensional scheme in Figure 2 (see below), I would like to present the basic categories' terms the scheme is developed with. The first category (vertical axis in Figure 2) labels the function (aim, purpose) (I–V) – what a model does: predicting, explaining, exploring, creating, giving strategies. The second category (horizontal axis in Figure 2) defines the method (a–e) models in music creation, analysis and perception/cognition are using to express the function defined before and showing, how and on what level of abstraction the model works. The function category can be generalized as expressing the existence (ontic) aspect, the second category shows the technique (epistemological aspect) which makes the ontic ontological (see Varto, 2009).

The function (aim, purpose) category includes the following types of models: I – prediction/expectation models and II – explanation models, e.g. models on phenomena like musical and tonal tension, and expectation in music (e.g. Margulis, 2005; Huron, 2007; Lerdahl & Krumhansl, 2007; Farbood, 2012).

Type I and II include also computational models of music perception and cognition: auditory expectation in information dynamics of music, Computational Auditory Scene Analysis (CASA) (see Wang, 2007), computational classification, search and retrieval of audio or content-based retrieval of music and audio. Such models try to understand well known phenomena or music (cognition) in general using cognitive, empirical, mathematical, and computational tools. III – exploratory models (including cognitive models, mental models, analogy models, metaphorical models) try to understand not yet formalized phenomena or phenomena from a yet unexplored angle. IV – models in composition/analysis of music (see more in Burkholder, 2017) can be separated into sub-categories IVa (composition) and IVb (analysis). V – teaching/learning models (e.g. strategies (developmental) psychology, sociology, identity, personality, motivation, creativity; individual *vs* group teaching and learning; e- and online learning) can be divided into Va (teaching) and Vb (learning) models.

The method category shows the modeling technique: a) mathematical models (statistics, algorithms, equations, also graphically expressed, including computation), b) representation models (including computational models, also cross-domain models, also graphically expressed), c) cognitive models (including analogy, cross-domain, and rule-based models, computational models, graphical models), d) theoretical models (including rule-based models, graphical models), e) conceptual models (including symbolic and iconic models, graphical models).

The following multi-dimensional scheme of modeling and models in music (Figure 2) considers also creativity aspects as well as how much the field of meaning of the phenomenon (to be observed and represented) and that of the model are in concordance. These aspects are explained below the figure. This rather generalizing overview can be understood as a ground plane scheme (viewer perspective from above), which means that both x and y axis are located on the horizontal plane, the dashed diagonal axes may then go even diagonally upwards to turn this scheme mentally into a 3D scheme.

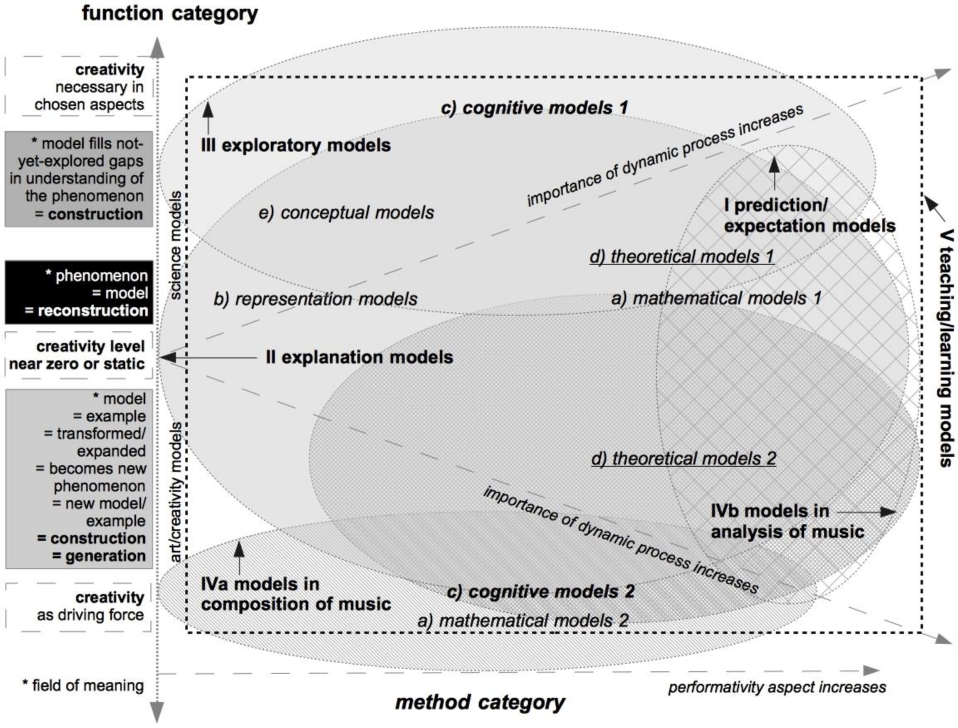


Figure 2. Multi-dimensional scheme of modelling and models incorporating models from music theory (analysis) and music psychology (perception and cognition) including the creative approach (musical composition/improvisation)

To understand Figure 2, one should start on the vertical axis' (left side) central point "creativity level near zero or static" which means that such types of models do not convey much creativity (of course, they may have been created also using creative thinking, but they don't express or perform creativity as such), they either represent or explain something (a phenomenon, source) directly, the analogy between the phenomenon and the model is high. From this central point, models develop in two directions: upwards one will find rather science models, downwards rather art/creativity models. In science models, creativity is necessary in chosen aspects (especially in exploratory models), in art/creativity models like composition (and improvisation) of music creativity is the driving force both in the generation and the use of models. On the left side of the scheme also field-of-meaning aspects of models are shown: in science models the model may either be in analogy with the phenomenon (means reconstruction) or, in exploratory models, not-yet-explored gaps in understanding of the phenomenon are filled (means construction) and need to be verified with experiments. In art/creativity models, the model becomes the example for similar achievements which is transformed and expanded, it becomes a new phenomenon, as such it can become a new model and example etc. (means construction and generation at the same time). The dashed diagonally axes express that the importance of the dynamic process (changing in time) included into the function and method of the models increases. This is in accordance with the dashed horizontal axis that shows the increase of the performativity aspect making models work only in the act of performance in real-time. Both express that models on the left side of the scheme use rather static principles (representing just one single state or situation), models on the right side of the scheme apply more dynamic principles (representing more than one single state, including the flow and development analyzed retrospectively as well as showing their prediction and expectation).

For each of the function category models (with roman numbers I–V) a 'bubble' like shape symbolizes the field of influence they approximately may embrace, and, more important here, how much their fields/bubbles overlap with each other. The 'bubble' of the II explanation models (grey area) is located most centrally and shows the largest extend, most of the models, indeed, explain something. The "bubble" of the III exploratory models (light grey area) is somehow narrower but overlaps to a large amount with that of the explanation models. They, too, explain phenomena, but their real aim is to deal with modeling something not-yet-explored. The 'bubble' of the IVa models in composition of music (tightly hatched diagonally from-left-to-right area), located below the explanation models' 'bubble', is smaller and overlaps both with the explanation models' and the IVb models in analysis of music (tightly cross-hatched area) 'bubble', located rather in the middle of the scheme, showing overlapping with all other model types' 'bubbles'. The I prediction/expectation models' (loosely crossed area) 'bubble' is located on the right side of the scheme, it shows overlapping with most of the other model types 'bubbles', showing the littlest amount with the IVa models in composition of music, because (not mentioning pieces written in a certain style/style copy pieces) creative contemporary musical composing plays with violating rules, models and expectations. I prediction/expectation models, in turn, show developments according to a clear predictability – of course they may show different possibilities and become performed dynamic models. IVb models in analysis of music, on the other hand, may or may not depart from composition models, they explain something already created as well as deal with predictions and expectations. The dotted square framing the scheme symbolizes that most of the previously introduced types of models can be used as V teaching/learning models.

The method category types of modeling (a–e) are put at those places in the scheme where they are most expected (however, different types of function category models use different methods in the same time) or where they can be localized in cross areas between different function types. In addition, c) cognitive models and d) theoretical models are positioned separately in the scheme, because their theoretical background differs: cognitive models 1 include music psychological cognitive approaches, cognitive models 2 include approaches in music theory/analysis (and composition) realms, e.g. treated as individual and folk psychology (see Cross, 1998). Theoretical models 1 are often mathematical or computational approaches. Theoretical models 2 refer to methods used in music theory/analysis which are based on musical scales, intervals, chords and voice leading principles and theories. Sometimes they use also mathematical means, see the position of a) mathematical models 1, in the cross section between I, II, III and IVb. Also in the IVa models in composition of music 'bubble' a) mathematical models 2 can be found, using different computational algorithmic and synthesis type of methods. The e) conceptional models are put in this scheme near the exploratory models, but as cross domain mental models they can be the starting point of several other types of models, too. The b) representation model is mostly a method used for explanation models and exploratory models, using also mathematical and theoretical means.

C. Constraints of scientific models

In this context it is necessary to point out two kinds of constraints scientific models can show (after Frigg & Hartmann, 2012): (1) *'Aristotelian idealization'* through reduction (a kind of imaginary 'stripping away') of aspects and features (properties) not necessarily important to the problem observed in order to concentrate on only those chosen and isolated properties (also called 'abstraction'). This means, in my opinion, pruning the object of observation. This kind of idealization can be found also in models of musical composition. Also, we should consider a constraint called (2) *'Galilean idealization'* involving deliberate distortions, because in reality, there are no idealized situations (ideal limits), a distorted (or even false model) can offer useful information for improvements of both the model and the understanding of the object of observation (and the research method). It was characteristic of Galileo's approach to science to use simplifications of this sort whenever a situation was too complicated to tackle (Frigg & Hartmann, 2012). This means, in my opinion, pruning the structures and relations between objects of observation in order to focus on specific aspects or achieve greater generalization. This kind of idealization can be found also in models of music analysis – the most common example for this in traditional music theory is the reduction of the music under observation to its primary parameters, pitch and rhythm, excluding secondary, but (especially perceptually and emotionally) also influential parameters like dynamics (loudness), articulation, or even instrumentation (timbre) etc.

Simplifications and distortions are useful both in science and (science) education, but these can become problematic – in education the question appears who (the teacher, the curriculum, political agenda) decides how and for what (pedagogical, developmental) reason which parts are simplified and distorted. In science it depends on the purpose of the research and the ability and sensitivity as well as the knowledge and experience of the researcher.

D. Models vs theory

A distinction between models and theory is complicated and can depend on the attitude towards science. One may say that this is 'just a model' or already awarded as 'theory' (Frigg & Hartmann, 2012) or *vice versa* as I think: a theory can have a pejorative meaning (because it may be speculative, incomplete, too complex), but a model (as both an example and consistent part of science) may be regarded as more valid than a theory. Frigg and Hartmann (2012) propose two extremes: a *syntactic* (logical positivistic) and a *semantic* view of theories (a theory is a family of models, models are seen as central units of scientific theorizing (see also Zbikowski, 1997)). They additionally point out that models can be independent of theories as 'autonomous agents' using two aspects: construction and functioning. Models can complement theories or so to speak stepping in when theories are too complex to handle (Frigg & Hartmann, 2012), or they can function as preliminary theories, or even as 'study models' or 'toy models' (without representational function, giving information just about itself, to test theoretical tools for later building of representational models).

Zbikowski (1997) explains that conceptual models are based on image schemata (which are not necessarily visual but mental images), specialized to a specific domain, but *"not in the cognitive processes that it involves"* (Zbikowski, 1997, 195). Conceptual models work also in cross-domain mappings*: "The conceptual models fundamental to theories of music are most typically borrowed from concrete domains and brought to bear on the domain of music through the process of cross-domain mapping. This process relies on establishing correspondences between the image-schematic structure of the source domain and the image-schematic structure of the target domain. Each such mapping will preserve only a portion of the structure of the target domain, and import only as much of the structure of the source domain as is appropriate"* (Zbikowski, 1997, 217).

According to Bhattacherjee (2012), in social science there are (among others like mathematical, network, path models) normative models used *"to guide our activities along commonly accepted norms or practices"* (p. 14). The author further divides models into *static* (represent the situation of a system at a chosen time point) and *dynamic* models (represent an evolution of a system over a certain time). An important aspect in models and modeling is the trajectory – deductive *vs* inductive – it is developed or unfolds, the first starting from theory (theoretical or logical reasons and an initial set of premises), the second from facts and observed (empirical) evidence (Bhattacherjee, 2012).

Lahti (2013) proposes four subdivisions explaining how students relate to scientific theory and modeling: a) student's understanding of theories (through analogy with models), b) the tentative nature of these theories (revision and modification), with possible c) understanding of the role of creativity in science (which variables to include), or d) the scientific method (construction of models, although not a typical experiment, is an appropriate scientific method of investigation, particularly for phenomena that may not be observed directly).

Schwarz et al. (2009) from the MoDeLS (Modeling Designs for Learning Science) project define scientific modeling as *"including the elements of the practice (constructing, using, evaluating and revising scientific models) and the meta-knowledge that guides and motivates the practice (e.g., understanding the nature and purpose of models)"* (p. 632). Their learning progression for scientific modelling includes two dimensions that combine meta-knowledge and elements of practice – scientific models as tools for predicting and explaining, and models change as understanding improves.

E. Computational and cognitive models in music

The computational understanding of modelling is based on the ability to express phenomena, processes (assumptions and consequences) that can be formulated in algorithms and computational representations, also in the neurological domain. In 1997, Vercoe (1997) was convinced that *"acoustic input can be the stimulus for computational models of an entire range of processes involved in the perception and cognition of music"* (p. 325). He continues that the goal should be *"to find how complex musical data is represented and processed by the human auditory-cognitive system. Only then will we understand why music that exploits this capacity has the structure it does"* (p. 325).

Earlier, Marsden (1995) revealed technical problems arising from *computer modeling* (which are still relevant also today) while categorizing the use of computers as tools in research as follows: (1) computers require strictly formal representations of theory; (2) they make possible the testing of theory through the implementation of models embodying that theory. He explains what has been so far (even today) the main approach: *"A computer program or sets of programs is written embodying some music theory, incurring developments in codification and (hopefully) clarification, because of (1) above; then the output of this program is tested by some criterion, e.g. comparison with human musical products, following (2) above, and, on the basis of any deficiencies, further developments of theory are proposed which can be codified and tested in anew cycle and so on"* (Marsden, 1995, 335).

Purwins et al. (2008) ask What is a valuable cognitive model using computational implementation based on findings from several disciplines like neuroscience, psychology, cognitive science, artificial intelligence, and musicology? (p. 154). They summarize the so far known methodology as *"localization of musical processes in the brain, and the flow of cognitive operations involved in turning physical signals into musical symbols, going from the transducers to the memory systems of the brain"* (Purwins et al., 2008, 151), They point out four aspects: a) predictive power, b) generalizing power, c) simplicity, and d) relation to existing theories. Also, they underline that the predicted outcome of the model should be later tested in experiments and without the model this kind of outcome would be unexpected. Furthermore, a model should explain data (from experimental results) *"that have not been taken into consideration in building the model"* (Purwins et al., 2008, 154). Importantly, referring to the fourth aspect (relation to existing theories), a model may have the ability to unify existing theories: *"This may help understanding a phenomenon instead of just reproducing input-output relation inherent in the data"* (p. 154). Purwins et al. (2008) underline that *"assumptions and consequences must be clearly defined and distinguished"* (p. 154). Their questions and claims have inspired me in the first place to develop my own twelve steps of a modelling approach.

Main contribution

A. Twelve Strategic Steps for modelling/analysis of models

In this section, I introduce my Twelve Strategic Steps (posed as questions) developed as a guide to create/analyze models in science (see Figure 3 below). The steps appear as one or more questions and sub-questions in eight developing stages/phases:

A) Initial situation (problem) stage,
B) Prospective outcome (hypothesis) stage,
C) Data developing stage 1,
D) Object of the model/research in relation to the data,
E) Model(ing) application phase,
F) Data developing stage 2,

G) Data analysis phase,
H) Data comparison phase (cross-domain mappings).

The questions presented in Figure 3 inquire necessary aspects of some of these *“selected point[s] of view”* proposed by Varto (2009, 11) in relation to the ‘event with order’ (making the ontic ontological by the technique). The event is the phenomenon observed/represented by the model (e.g. music), the order is the chosen method of approach to the phenomenon with the help of the model (technique). These stages and phases should not be worked through linearly. They should be used in a rather spiral-type of modelling process. Answers to these questions may require their testing (including refinements and modifications) as well as going back and forth between necessary questions, steps, stages and phases depending on their answers and resulting in next follow-up steps. My steps and questions have been partly inspired also by Purwins et al. (2008) four basic questions and other requirements for building cognitive models on the following way: to a lesser amount its predictive power, to a greater amount its generalizing power, to a great amount its simplicity, and, as a basic requirement, its relation to existing theories (if possible, unifying of several theories) as well as experimental testing of its ‘input’ and ‘output’ data.

A	**Initial situation (problem) stage** Q 1. What kind of phenomenon one observes/investigates/explores? **See typologies of models.** Q 2. What is the problem/aim to be researched/explored?
B	**Prospective outcome (hypothesis) stage** Q 3. What may be the postulated outcome to be reached? Q 3a. What kind of theory/theories relate/s to/explains the phenomenon/problem/aim? Q 3b. Is a hypothesis necessary? Quantitative approach: yes Qualitative approach/Artist research: no
C	**Data developing phase 1 - “input” data for the model (IDV – independent variables)** Q 4. What kind of data does the phenomenon provides? Q 5. How to retrieve/record appropriate data, portions of data (e.g. feature analysis)? Q 5a. Is computation of data retrieval possible? Q 6. To what requirements need the data be restricted, portioned, segmented?
D	**Object of the model/research in relation to the data** Q 7. Does the data enable to model/research the phenomenon or its observer, or both?
E	**Model(ing) application phase** Q 8. In which stage of the research process (and why) is it possible/necessary to apply modeling? Q 9. What type (function/method) of model makes sense to apply? **See typologies of models.**
F	**Data developing phase 2 – “output” data of the model (DV – dependent variables)** Q 10. What kind of data are the output of this model? Q 10a. Does it represent the phenomenon well? Q 10b. How generalizable are the results?
G	**Data analysis phase** Q 11. What is the concrete aim of the “output” data? Q 11a. How to treat (deduce, reduce, transpose, analyze) this “output” data further? Q 11b. How to use computational tools for data analysis?
H	**Data comparison phase (cross-domain mappings)** Q 12. How able are the data to be compared with output from other research methods applied on the same phenomenon? Q 12a. How to use computational tools for data comparison?

Figure 3. Twelve Strategic Steps (grouped in eight stages/phases) for modelling/analysis of models

Answers to these questions will lead to a versatile network of information that opens up a growing understanding of the phenomenon the model aims to describe the theory (theories) and scientific paradigm(s) behind the phenomenon, its research field and the model applied or to be developed. Several aspects will be needed to be answered in more than one questions recurrently, but from different perspectives. These mirror the selected points of view and give a multi-faceted insight into the model's function and/or the modelling process. As a useful pedagogical aspect, it may happen that one must correct or change opinions, understandings, functions or the method during the process of modelling or analysis of a model.

After having developed a model using these Twelve Strategic Steps it needs to be tested technically (if it has a computational realization), this also may reveal additional problems or aspects connected with its development, requiring maybe modifications (see also in Justi and Gilbert's *"Model of Modelling"*). If the model works technically well, the next phase is its real application (empirical tests) and the interpretation of the data as well as drawing conclusion in relation to the phenomenon observed or how it explains the mental construction the model represents. If the model goes too far from the observed phenomenon or revolves only around the data retrieved and computationally further developed it should be critically questioned and if needed revised and modified again.

B. Justi and Gilbert's enlarged "Model of Modelling"

Although Justi and Gilbert's (2002) *"Model of Modelling"* includes on the mental and the empirical level (of the research design) the requirement of tests/experiments, they say nothing about the method, how to do this, they do not go into the data topic (explicit *vs* implicit, 'input' vs 'output' data). Therefore, I have re-visualized and enlarged their *"Model of Modelling"* in Figure 4. Interestingly, and different from the distinctive definition shown in the beginning of this article, the authors integrate mental modelling (thought experiments) with the empirical testing requirement, the representation stage of their modelling is located between these both phases. Their first step is to decide on the purpose for a model (1). The second step is selecting the source for the model (2a) and including own experiences (2b). The third step is to produce a mental model (3), the forth step to express this mental model in different forms of representations (4). The fifth step is to conduct thought experiments (5) including modifying in case of fails (5x), the sixth step is to design and perform empirical tests (6) including modifying in case of fails (5x). The fifth and sixth steps may also lead (in case of extreme fails) into rejecting the mental model (6x). If the mode of representation of step four needs data development (a step added by me), one can proceed to (4a). If this results in a certain mode (visualization, verbalization, equation) (6b) which satisfies the modeler, then the purpose of modelling is already fulfilled (6z). Also step six (6) may involve data development (6a) (added by me). In case of successful passing all previous steps, the model has fulfilled its purpose (6z). As a last seventh step, one should consider the scope and limitation of the model (7). I have added step eight, the concept of cross-domain mapping (8) which opens interdisciplinary possibilities to decide on a next purpose (1) to continue the modeling process on a next level. These steps don't include more specific aspects like structure of the source, no mentioning of the target or type of data and originally no data development aspects. The steps are enlarged with an additional data development step (4a) and (6a) as well as step eight cross-domain mappings (8), which opens interdisciplinary possibilities to continue the modelling process on a next level (see Figure 4).

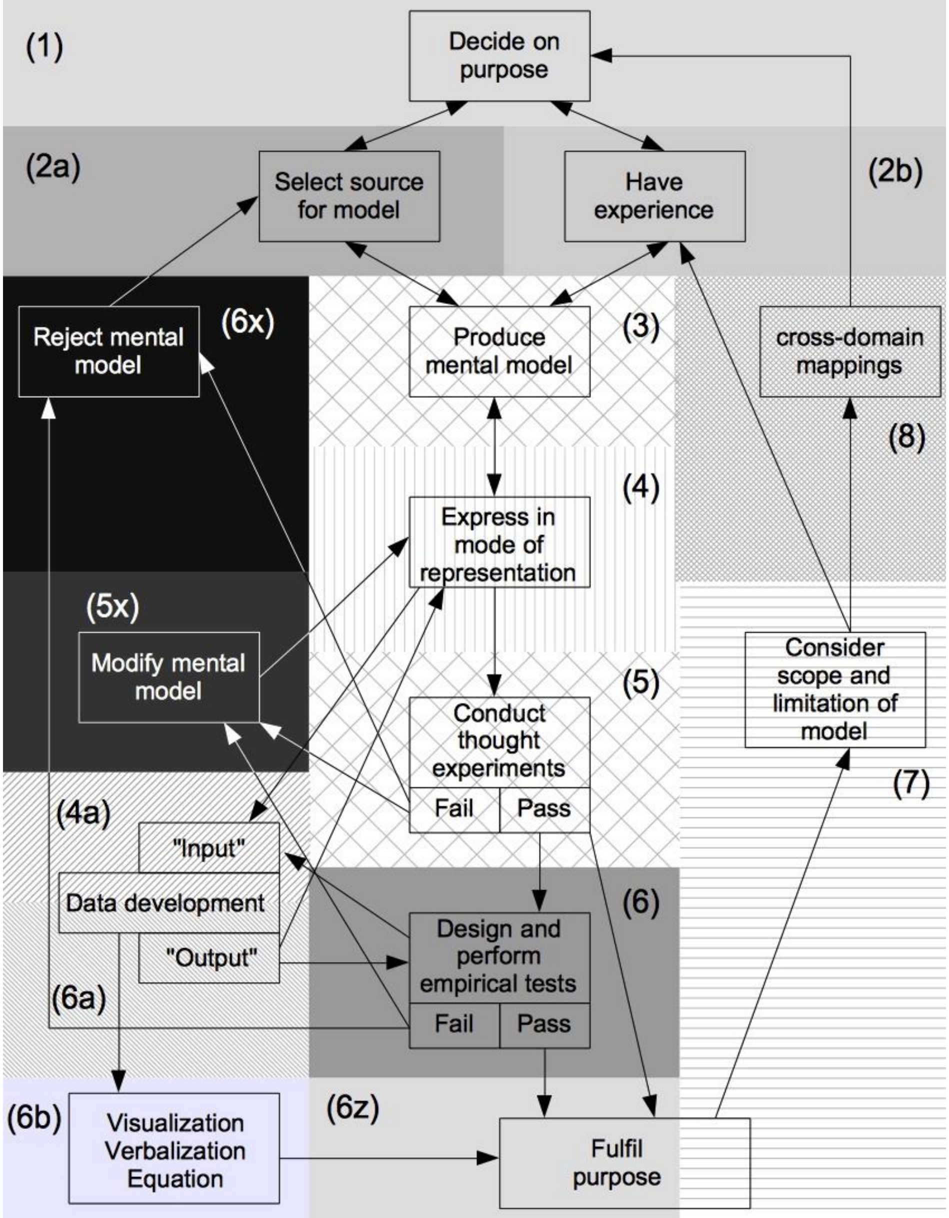


Figure 4. Enlargement of Justi and Gilbert's (2002), Gilbert's (2010) "Model of Modelling" by G. Lock

C. Comparison of Justi and Gilbert's "Model of Modelling" and the Twelve Strategic Steps

Justi and Gilbert's (2002) *"Model of Modelling"* step (1) (decide on purpose) can be compared with stage/phase A (initial situation) and B (prospective outcome) of my Twelve Strategic Steps. Step (2a) (select source for model) is connected to A and E (modeling application phase). Step (2b) (have experience) means the experience with A (the phenomenon one observes/ investigates/ explores) and one comes back after the modelling process including step 7 (consider scope and limitation of the model) which means to having gone through all stages/phases from A to G. Step (3) (produce mental model) is not directly addressed in my strategic steps as separate stage/phase, but it can be part of B (prospective outcome, if mentally run). Step (4) (express in mode of representation) means to visualize step (3) which I have separated info C and F (data development stage 2, data may become an own 'reality'). Having enlarged Justi and Gilbert with a step (4a) the aspect of data, which allows representation, is brought in. Step (5) (conduct thought experiments) also does not have a direct equivalent in my steps, because I operate from beginning on with 'input' and 'output' data which enables also mental aspects to be graspable and visualize-able (6a, 6b) before they can be empirically tested (6). If either the visualization/verbalization/equation derived from steps (3) and (4) or the empirical tests (6) are successful, then the model can be treated as ready (6z) (fulfills purpose), but these steps are in my approach part of other stages/phases during data development (C and F) and data analysis phases (G). Stage/phase H (data comparison/cross-domain mappings) is somehow separated in my approach, I have added it also to Justi and Gilbert's scheme as step (8) (cross-domain mappings) in order to lifting their scheme to a next level of developing a possible new model.

Both approaches are in some aspects, close to each other dealing with purpose, source, experience, experiments, but they are not identical. Justi and Gilbert's (2002) approach can, indeed, be called a *"Model of Modelling"* because in its enlarged version it fulfills most of the theoretical implications I propose below in the implication section (of course, theory is not mentioned, but it could be found in the mental model production step), and it forms a whole: from the starting point through modelling steps back to the same starting point (on a next level). In contrast, I underline that my Twelve Strategic Steps establish not a model, they remain a kind of algorithm to follow, steps from one starting point (initial situation) through different stages/phases until 'output' data analysis (G). Their aim is to develop a basis to compare its outcome with other methods (models). Also, my Twelve Strategic Steps enable the analysis of existing models, their questions can trigger answers concerning when or why an idea, concept, systematic approach, a method or a theory may turn into a model (see also conclusion below). Both approaches should be applied in parallel considering the purpose of the modelling aim and necessity during the research process.

Implications

The implications drawn from the developing of the Twelve Strategic Steps for modelling/analysis of models are the following: a scientific model should be built based on a theory or multiple theories even if the relation between models and theory is 'perplexing' (Frigg & Hartmann, 2012). Models can be also the basis for theories (for general science see Frigg & Hartmann, 2012, for conceptual models and cross-domain mappings in the music field see Zbikowski, 1997, 2002), but in this article, I favor the theory-first approach, because an unexplainable or non-theory-grounded approach goes against the pedagogical purpose of this article. Further, a scientific model must be tested against the phenomenon or theoretical/mental construct it represents, the elements of a model must be always consistent and not functionally empty (providing justified answers to why-questions). A theory, in contrast, may remain incomplete (Frigg & Hartmann, 2012). Scientific models, like creative outputs and comprehensions, derive from explicit and implicit knowledge of the modelling person which is gained either intuitively, systematically, system-based or automatically (see Lock, 2006, 2014), *"models are vehicles for learning about the world"* (Frigg & Hartmann, 2012). A scientific model should gather and synthesize rather well-known, proven or justified elements (see theory requirement in the first implication). The innovations of new models lie often in the surprising way these elements are combined, one can also learn about the model through experiments, thought experiments and simulation (Frigg & Hartmann, 2012). To achieve testability of a scientific model, it itself (and its representation) should operate with explicit or implicit data, which can be expressed as numbers and even computer algorithms to a) enable greater validity and generalizability, b) be applicable in further comparative research (including cross-domain mappings). These data may appear in numerical form (explicit) or may be hidden behind narrative or visual (or nonverbal), rather ambiguous, forms of expression, but it is still existent implicitly as basis for the models' shape, technique and method as both 'input' data (data gained from the phenomenon under observation or the theoretical/mental construct) and 'output' data (data derived from the model itself). To test a theoretical/mental model or a model built and analyzed based on 'non-explicit' (hidden) data, its implicit data need to be made accessible: turned into a logical structure that can be expressed in numerical data. To be clear: the approach I present does not treat commonly called 'data models' in the meaning of mathematical equations or algorithmic procedures, also not 'data models' derived from empirical research using standard procedures of descriptive statistics etc. It therefore does not restrict modelling to an abstract quantitative 'game with numbers' and should not detach it form the phenomenon under observation.

As I showed in the theoretical background section, music can be understood and studied both as (open) natural environment in the sense of world/reality or as dynamical system treated as (closed) theoretical system. But, music evolves only through creation (either technically via composing/performance or mentally interpreting sounds as music), to conceptualize (understand) sounds as music, one uses different kinds of models (either being aware of them or not). Hence, the perception/cognition process is explainable for human beings also through different kinds of models, which are not the phenomenon itself but a vehicle to verbalize, visualize, comprehend and communicate it as idea and concept of a 'real-world reality'. Models therefore, based on Varto's (2009) philosophical background, offer a systematic, visualized and reasonably simplified 'selected point of view' to an 'event' the observer has given an 'order'. The intriguing question concerning knowledge gaining – 'mind first' or 'world first' (Varto 2009) – should be answered as an intertwining dependence of both from each other.

A model, which is theory and data-based, but does not exclude the world in which the phenomenon exists, can have the ability to enhance the awareness, conceptualization, and simplify this intertwining dependence between the world and the mind as well as the existence and the technique having concrete functions (aims, purposes) and using reasonable methods in this model. In other words, the successful functioning of a model depends on the proper method it uses. I also agree with the even more concluding twofold-categorization of models and modelling proposed by Hans-Gunter Lock which merges function and method: 1) analytical or descriptive models – describe a phenomenon of the world; 2) generative models – allow to generate new understanding (explanations) of a phenomenon of the world: a) compare the new explanation with the phenomenon itself (e.g. Chomsky's Generative Grammar, String Theory in physics, b) apply generative models in creative activities (counterpoint rules, algorithmic composition, etc.).

Conclusions

1. To summarize, it makes sense to develop and apply a model under the following conditions and concluding the basic implications proposed in this article:
   a) A scientific model needs to have one or more accepted/proven theories as backbone(s);
   b) The phenomenon observed/modeled must be able to be interpreted, translated and expressed into explicit or implicit 'input' data, the latter should be able to be made explicit, turned into numbers; the explicit 'output' data must be analyzed and interpreted reasonably and simply without losing touch with the phenomenon to be modeled;
   c) The model must be consistent, a whole in itself. Gaps or inconsistencies must be minimized or turned off (eliminated);
   d) As an exploratory or new model, it should explain something generally known, but not yet explained in a similar systematic way from the same angle, from a new viewpoint based on accepted/proven theories, but offering also new or even unexpected aspects using e.g. cross-domain mapping approaches between different fields.
2. Answers to the questions concerning when or why an idea, concept, systematic approach, a method or a theory may turn into a model can be found in my Twelve Strategic Steps introduced above. The A "Initial situation (problem) stage", Q1 inquiring the kind of the phenomenon to be observed/investigated/explored requires in parallel to become familiar with already known typologies of models, Q2 asks about the problem/aim to be researched. The B "Prospective outcome (hypothesis) stage" Q3a reveals what kind of theories can be found behind the phenomenon to be observed/researched. The C "Data developing stage 1" Q4 on the kind of "input" data the phenomenon provides as well as Q5 and Q5a on retrieval/recording and computation possibility reveal explanations towards the applicability of a data-based model. In D "Object of the model/research in relation to the data" Q7 gives insight whom the data actually represents, the phenomenon and/or the observer. The question of whom to model/research seems to be late here and may be taken into account already in earlier stages (A and B). But in my opinion one should first do some kind of data development ('input' data) to understand better what this data may model (an early and to strict determination may lock up/limit the further development of the model as well as the understanding of the phenomenon to be modeled. The answers in E "Model(ing) application phase" to Q8 (In which stage of the research process (and why) is it possible/necessary to apply modeling?) and Q9 (What type (function/method) of model makes sense to apply?) are the most direct questions to detect the usefulness of developing/applying a model. Also the F "Data developing stage 2" enables to find answers to the quality and independence of the "output" data, which with Q10a (Does it represent the

phenomenon well?) Further the G "Data analysis phase" with Q11 (What is the concrete aim of the "output" data?) and Q11a on "output" data treatment and analysis possibilities gives hints together with the final H "Data comparison phase (cross-domain mappings)" Q12 towards detection of the reasonableness of developing an idea, concept, systematic approach, a method or a theory into a model.

3. In short, modelling in science is fruitful, if an idea can become explained with a concept, which can be approached systematically, further developed as a method, includes connections/backups to one or more theories, and is consistent (a whole), either representing or exploring a known (but less systematically explained) phenomenon based on its explicit and implicit knowledge and data including cross-domain mappings.

I hope that my Twelve Strategic Steps for creation/analysis of models and the enlarged Justi and Gilbert *"Models of Modelling"* will enhance the understanding of modelling and will assist creating models in sciences in the field of music and music education.

References

Bailey, M. (2017). Implicit & Explicit Data. *MarketMotive (simplilearn)*. Retrieved 07.10.2017 from http://www.marketmotive.com/training/home/tutorials/email-techniques/implicit-explicit-data

Bhattacherjee, A. (2012). *Social Science Research: Principles, methods, and practices*. Tampa (USA): University of South Florida. Retrieved 07.07.2017 from http://scholarcommons.usf.edu/oa_textbooks/3

Börner, K., Boyack, K. W., Milojević, S. & Morris, S. (2012a). An introduction to modeling science: Basic model types, key definitions, and a general framework for the comparison of process models. In A. Scharnhorst, K. Börner, & P. van den Besselaar (Eds.), *Models of Science Dynamics - Encounters Between Complexity Theory and Information Sciences.* Retrieved 07.10.2017 from http://cns.iu.edu/images/pub/2012-borner-modsci-springer.pdf

Börner, K., Boyack, K. W., Milojević, S. & Morris, S. (2012b). An introduction to modeling science: Basic model types, key definitions, and a general framework for the comparison of process models. In A. Scharnhorst, K. Börner, & P. van den Besselaar (Eds.), *Models of Science Dynamics - Encounters Between Complexity Theory and Information Sciences* (pp. 3-22). Berlin, Heidelberg: Springer.

Bruner, J. (1990). *Acts of Meaning*. London: Harvard University Press.

Burkholder, J. P. (2017). Modelling. In *Oxford Music Online/Grove Music Online*. Retrieved 07.07.2017 from http://www.oxfordmusiconline.com.ezproxy.tlu.ee/subscriber/article/grove/music/53082

Burrows, D. (1997). A dynamical systems perspective on music. *The Journal of Musicology*, 15(4), 529-545.

Byham, R., Hamilton, B. & Guyer, C. (2017). Data type conversion (Database Engine). Microsoft Docs article (23.07.2017). Retrieved 07.10.2017 from https://docs.microsoft.com/en-us/sql/t-sql/data-types/data-type-conversion-database-engine

Cartier, J., Rudolph, J. & Stewart, J. (2001). *The Nature and Structure of Scientific Models.* Madison (USA): The National Center for Improving Student Learning and Achievement in Mathematics and Science (NCISLA) of Wisconsin Center for Education Research, School of Education, University of Wisconsin-Madison. Retrieved 07.07.2017 from https://eric.ed.gov/?id=ED461513

Coll, R. K. & Lajium, D. (2011). Modeling and the future of science learning. In M. S. Khine, & I. M. Saleh (Eds.), *Models and Modeling: Cognitive Tools for Scientific Enquiry* (pp. 3-21). Dordrecht: Springer Netherlands.

Cook, N. (1994). Perception: A perspective from music theory. In R. Aiello, & J. Sloboda (Eds.), *Musical Perceptions*. Oxford: Oxford University Press.

Cross, I. (1998). Music analysis and music perception. *Music Analysis*, 17(1), 3-20.

Duit, R. (1991). On the role of analogies and metaphors in learning science. *Science Education*, 75(6), 649-672.

Dienes, Z. & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences,* 22, 735-808.

Farbood, M. M. (2012). A parametric, temporal model of musical tension. *Music Perception*, 29(4), 387-428.

Frigg, R. & Hartmann, S. (2012). Models in science. *Stanford Encyclopedia of Philosophy.* Retrieved 07.07.2017 from http://plato.stanford.edu/entries/models-science/

Gilbert, J. K. (2010). The role of visual representations in the learning and teaching of science: An introduction. *Asia-Pacific Forum on Science Learning and Teaching*, Volume 11, Issue 1, Foreword (Jun., 2010). Retrieved 07.10.2017 from https://www.researchgate.net/publication/45459005

Giomi, F. & Ligabue, M. (1995). Semiotic bases and computer assisted composition: Towards a cognitive model. In E. Tarasti (Ed.), *Musical Signification. Essays in the semiotic theory and analysis of music* (pp. 355-360). Berlin, New York: Mouton de Gruyter.

Huron, D. (2007). *Sweet Anticipation: Music and the psychology of expectation*. Cambridge MA, London: MIT Press.

Justi, R. & Gilbert, J. K. (2002). Modelling: Teachers' view on the nature of modelling, and implications for the education of modelers. *International Journal of Science Education*, 24(4), 369-387.

Kuhn, T. J. (1962/1970). *The Structure of Scientific Revolutions* (Second Edition, enlarged). Chicago: University of Chicago Press.

Lahti, R. D. (2013). *Does Attainment of Piaget's Formal Operational Level of Cognitive Development Predict Student Understanding of Scientific Models?* (Doctoral dissertation). Missoula: University of Montana.

Lerdahl, F. & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception*, 24(4), 329-366.

Lock, G. (2016). Stages and phases of the creative process of novice composers (paper presentation). Paper presented at the Pedagogical Forum Performing Arts Belgrade *"Building Experience and Meaning in Music Education"*. Belgrade: University of Arts.

Lock, G. (2014). Visualisierende Analyse. In F. Diergarten (Hrsg.), *Musikalische Analyse. Begriffe, Geschichten, Methoden* (ss. 287-319). Laaber: Laaber.

Lock, G. (2004). Aspekte der Visualisierung und graphischen Analyse von Musik. Die Fünfte Sinfonie von Lepo Sumera. *Musical Work: Boundaries and Interpretations (38th Baltic Musicological Conference)*, hrsg. von A. Žiuraityte (pp. 243-269). Vilnius.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, 22(4), 663-714.

Marsden, A. (1995). Musical pragmatics and computer modelling. In E. Tarasti (Ed.), *Musical Signification: Essays in the semiotic theory and analysis of music* (pp. 335-348). Berlin, New York: Mouton de Gruyter.

Purwins, H., Herrera, P., Grachten, M., Hazan, A., Marxer, R. & Serra, X. (2008). Computational models of music perception and cognition I: The perceptual and cognitive processing chain. *Physics of Life Reviews,* 5, 151-168.

Reybrouck, M. (2015). Music as environment: An ecological and biosemiotic approach. *Behavioral Science,* 5, 1–26.

Rogers, K. (2017). Scientific modeling. *Encyclopaedia Britannica*. Retrieved 07.17.2017 from https://www.britannica.com/science/scientific-modeling

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Acher, A., Fortus, D., Shwartz, Y., Hug, B. & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 63-654.

Tiits, K. (1995). Experiments on grammar-based analysis of music. In E. Tarasti (Ed.), *Musical Signification: Essays in the semiotic theory and analysis of music* (pp. 361-369). Berlin, New York: Mouton de Gruyter,

Varto, J. (2009). *Basics of Artistic Research: Ontological, epistemological and historical justifications.* Helsinki: University of Art and Design Helsinki.

Vercoe, B. (1997). Computational auditory pathways to music understanding. In I. Deliège, & J. Sloboda (Eds.), *Perception and Cognition of Music* (pp. 307-326). Hove (UK), New York: Psychology Press.

Wang, D. (2007). Computational scene analysis. In W. Duch, & J. Mandziuk (Eds.), *Challenges for Computational Intelligence* (pp.163-191). Berlin: Springer.

Wigmore, I. (2017). Implicit data (definition). *WhatIs.com /TechTarget.* Retrieved 07.10.2017 from http://whatis.techtarget.com/definition/implicit-data

Zbikowski, L. M. (2002). *Conceptualizing Music: Cognitive structure, theory, and analysis*. Oxford: Oxford University Press.

Zbikowski, L. M. (1997). Conceptual models and cross-domain mapping: New perspectives on theories of music and hierarchy. *Journal of Music Theory*, 41(2), 193-225.

Received 16.07.2017
Accepted 04.12.2017

A.5 Publication IV (Lock 2020, 5.5)

“Salienz, Narrativität und die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung von post-tonaler Orchestermusik [Salience, Narrativity and the Role of Musical Parameters in the Analysis of Musical Tension in Post-Tonal Orchestral Music]” (in German)

Musical tension, salience, narrativity and musical parameters – 4th Symphony/Percussion Concerto *Magma*

Salienz, Narrativität und die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung von post-tonaler Orchestermusik

Gerhard Lock

Dieser Artikel diskutiert Ergebnisse einer Studie zu kognitiven Hörstrategien von post-tonaler Orchestermusik (Erkki-Sven Tüür). Im ersten Teil werden musikalische Spannung und temporale Musikanalyse, musikalisches Ereignis und Salienz sowie kognitive Narratologie und narrative Musikanalyse thematisiert. Im zweiten Teil werden das empirische Experiment-Design (Sample N=26) für das vom Autor entwickelte neue kognitive Scheibenmodell COSM (*Cognitive Octagonal Slice Model*), die Visualisierung und Analyse der COSM-Daten (*Data Mining*) sowie einige Analyse-Ergebnisse bezüglich eines möglichen Zusammenhangs zwischen Tendenzen musikalischer Parameter und musikalischer Spannung diskutiert. Bezüglich der Rolle primärer und sekundärer Parameter für die Wahrnehmung der musikalischen Spannung ergaben sich bis auf ein paar interessante Ausnahmen in der Mitte des untersuchten Werkes keine klaren Muster. Es konnte auch nicht die vermutete Dominanz sekundärer über primäre Parameter festgestellt werden.

This article introduces results of a study about cognitive listening strategies for post-tonal orchestral music (Erkki-Sven Tüür). The first part deals with musical tension and temporal music analysis, musical event and saliency as well as cognitive narratology and narrative music analysis. The second part discusses the empirical Experiment-Design (Sample N=26) of the author's developed new cognitive model COSM (Cognitive Octagonal Slice Model), the visualisation and analysis of the COSM data (data mining) as well as some analytical results concerning a possible relation between tendencies of musical parameters and musical tension. The role of primary and secondary parameters for the perception of musical tension did not show clear patterns, though some exceptions in the middle of the work still appeared valid. Also, the prognosed dominance of secondary parameters over primary parameters could not be detected.

Schlagworte/Keywords: audiovisual perception; audiovisuelle Wahrnehmung; Cognitive music analysis; Kognitive Musikanalyse; modeling; Modellierung; musical parameters; musical tension; musikalische Parameter; musikalische Spannung; saliency; Salienz

> *»Ein einzelnes musikalisches Ereignis kann zum Schlachtfeld für konkurrierende Weltsichten, Methodologien und soziale Einstellungen werden.«*[1]

EINFÜHRUNG

Dieser Artikel[2] diskutiert Ergebnisse einer Studie zu kognitiven Hörstrategien von post-tonaler Orchestermusik. Durch *slider controller* (›kontinuierliche‹ Echtzeit-Daten) wird die ›wahrgenommene‹ strukturelle musikalische Spannung empirisch erfasst. COSM ist

---

[1] Almén 2006, 1.

[2] Der Text ist aus einem Teil der sich im Abschluss befindenden Doktorarbeit des Autors entwickelt, die er unter der Betreuung von Prof. Dr. Kerri Kotta an der Estnischen Akademie für Musik und Theater schreibt.

das Kopfwort des neuen kognitiv-empirischen Modells des Autors[3] und bedeutet ›Kognitives achteckiges Scheibenmodell‹ (*Cognitive Octagonal Slice Model*). Es stellt Möglichkeiten bereit, Hörstrategien auf Basis der audiovisuellen Salienz[4] von Impulsen und musikalischen Parametern empirisch zu untersuchen. COSM kann in den Bereich der temporalen Musikanalyse[5] eingeordnet werden. Im Zuge der Entwicklung von COSM hat der Autor die theoretische Grundlage sowie zwölf strategische Schritte (in acht Phasen/Stadien) zur Entwicklung einer wissenschaftlichen Modellierung formuliert.[6]

Das Gebiet der Musiktheorie wird in diesem Artikel durch die Analyse der musikalischen Spannung sowie primärer und sekundärer musikalischer Parameter berührt.[7] In der kognitiven Musikanalyse wird überwiegend mit Tonaufnahmen gearbeitet, wobei Untersuchungsgegenstand die kognitiven Reaktionen sind, die durch den Hörprozess in Echtzeit hervorgerufen werden. Die kognitive Musikanalyse kann sozusagen als Ausarbeitung (›Strukturierungsprozess‹) von ›relevanten Kategorien‹ angesehen werden.[8] Hierbei kann man das musikalische Ereignis auch als Stimulus, Signal oder Impuls betrachten. Impulse sind – wie auch der Terminus ›Ereignis‹ (*event*) – ebenfalls in der Neurophysiologie verwendete Konzepte zum Verständnis des Wahrnehmungsprozesses.[9]

Musiktheorie und Musikpsychologie verbindende interdisziplinäre Perspektiven zeigen auch Adam Ockelfords »zygonische Theorie«[10] sowie Martin Rohrmeier,[11] letzterer mit seinem als Brückenschlag zwischen musiktheoretischen, kognitiven und computationalen Ansätzen konzipierten Diskussionsansatz.

Ein Beitrag eines internationalen Kongressberichts aus dem Jahr 2016 bestätigt, dass die Forschung weiterhin interessiert daran ist, computergestützte Strukturanalyse anzuwenden, bei welcher man »Musik segmentiert oder zerlegt (dekomponiert) in Muster oder Einheiten, die einige semantische Relevanz besitzen, und des Weiteren diese Einheiten in musikalisch sinnhafte Kategorien gruppiert.«[12]

Als Daten-Analyse wird im weitesten Sinne *Data Mining*[13] verwendet, welches durch Modellentwicklungen, Gruppierungen und Clusteranalyse interessante Muster und Relationen in großen Datensammlungen entdeckt, unter anderem auch bei Zeitskalen-Daten.[14] Auch bei der Musikanalyse entdeckt man Strukturen und deren Relationen zueinander und untersucht, wie diese sich verändern.[15] Bei beiden stehen Prinzipien der Ähnlichkeit und der Segmentation (u. a. Konturanalysemethoden)[16] im Vordergrund, wobei

---

[3] Lock/Kotta 2015.

[4] Salienz bezeichnet in der Wahrnehmungspsychologie das Hervortreten von Merkmalen und kommt der Alltagsbedeutung von ›Auffälligkeit‹ nahe.

[5] Neuwirth 2008, 558, 560, 562–564.

[6] Lock 2017.

[7] Meyer 1989; Snyder 2000.

[8] Chouvel 2014.

[9] Snyder 2000; Abbott 2002; Purwins/Herrera/Grachten/Hazan/Marxer/Serra 2008.

[10] Ockelford 2009, empirisch getestet von Thorpe/Ockelford/Aksentijevic 2012.

[11] Rohrmeier 2013.

[12] Müller/Chew/Bello 2016, 148.

[13] Clifton 2019.

[14] Fu 2011.

[15] Conklin 2006, 349; Tenkanen 2010, 16.

[16] Bor 2009.

Segmentation sowohl bei der Daten-Erfassung für musikalische Spannung in zeitgenössischer Musik[17] als auch bei der strukturellen Analyse zur Anwendung kommt (computergestützte Methoden).[18]

Im Folgenden werden einige methodologische Aspekte der Analyse post-tonaler Musik[19] am Beispiel von Orchestermusik des estnischen Komponisten Erkki-Sven Tüür (geb. 1959) erörtert, von dem sich sagen lässt, dass er in Rang und Bedeutung gleich auf den weltbekannten estnischen Komponisten Arvo Pärt (geb. 1935) folgt. Tüürs Musik verbindet avantgardistische Kompositionstechniken (u. a. Dodekaphonie, Atonalität, Oktatonik sowie die von ihm seit 2002/03 entwickelte Intervallvektor-Technik) mit traditionellem Denken (symphonische Form und »Kulminationsdramaturgie«).[20] Da hierbei traditionelle Analysemethoden nicht mehr gänzlich greifen, die Musik jedoch durch universelle Merkmale wie Wellenförmigkeit, Zielgerichtetheit, Kontraste sowie das Verdichten und Auflockern ihrer Struktur bzw. ihres Klangs für eine breitere Hörerschaft intuitiv verständlich ist, ist sie meiner Ansicht nach als Forschungsfeld für die Entwicklung einer neuen kognitiven Methode geradezu prädestiniert.

Methodologisch liegt der Fokus auf kognitiven Hörstrategien, die die Hörenden zumeist intuitiv anwenden. Musikalische Spannung und die Salienz von Impulsen sind hierbei Gesamtphänome, welche holistische Informationsverarbeitung erfordern und Stimuli »als integrales Ganzes« (»Gruppierung auf Grund von globalen Ähnlichkeiten«) betrachten.[21] Die Salienz von musikalischen Parametern erfordert analytische Informationsverarbeitung, Stimuli werden bezüglich ihrer übereinstimmenden Eigenschaften (Bestandteile) verarbeitet.[22] Dies findet im Verlauf des ›aufmerksamen‹ Hörprozesses[23] statt, der als Narrativ interpretiert bzw. mit einer Erzählung verbunden werden kann. Dies wirft die Frage nach der Rolle bzw. dem gegenseitigen Verhältnis von ›Sender‹ (Musik) und

---

[17] Addessi/Caterina 2000.

[18] Guigue/de Paiva Santana 2018.

[19] Diese Aspekte werden hier nur durch die Analyse eines einzigen Werkes eines konkreten Komponisten veranschaulicht, weshalb die Ergebnisse nicht verallgemeinert werden können; denn um verallgemeinerbare Aussagen zu erhalten, wären ein gesamtes Œuvre oder Beispiele einer Stilperiode empirisch zu untersuchen. Es gibt für post-tonale Musik bisher kaum Korpusstudien. Kirschbaum (2001) thematisiert zwar Werke zehn post-tonaler Komponisten mit unterschiedlichen Personalstilen (u. a. Anton Webern, Karlheinz Stockhausen, György Ligeti, Morton Feldman und Helmut Lachenmann), geht aber streng genommen nicht empirisch vor, da er keine Probanden einbezieht; Addessi und Caterina (2000) vergleichen die Musik von Darius Milhaud, Bruno Maderna und Webern, bei der sich jeweils eine Gradation von ›noch-tonalen‹ bis hin zur Abwesenheit tonaler Strukturen zeigt.

[20] Im Sinne einer Aufeinanderfolge von Höhepunkten als »Zonen besonders gesteigerter Intensität« (auch »Attraktionspunkte« genannt), die »das Erfassen der Form im Gedächtnis unterstützen«. Dies wird von Martin Kirschbaum übergeordnet als Dramaturgie bezeichnet und des Weiteren als gleichbedeutend mit dem musikalischen Spannungsverlauf erachtet. Kirschbaum führt in diesem Zusammenhang die enge Beziehung zwischen Form und Dramaturgie in traditionellen Formen, z. B. der Sonatenform, an. Durch den Wegfall feststehender Formen seit dem 20. Jahrhundert kam es Kirschbaum zufolge zum Verlust einer solchen formimmanenten Dramaturgie, was ihn zur Analyse der Dramaturgie und des »formalen Spannungsverlaufes« in post-tonaler Musik anregte. Vgl. Kirschbaum 2001, 8 f.

[21] Eitan/Granot 2009, 165.

[22] Siehe einige der grundlegenden Publikationen aus dem Feld der kognitiven Musikwissenschaft, die auch explizit die Wahrnehmung von musikalischen Parametern untersuchen, wie z. B. Seifert 1993, 1999 und 2008; Lerdahl/Jackendoff 1983; Purwins/Herrera/Grachten/Hazan/Marxer/Serra 2008; Reybrouck 2010; Pearce/Rohrmeier 2012; Koelsch 2013.

[23] *Attentive listening*; ›aufmerksam‹ bedeutet auch ›achtsam‹ und ›sorgfältig‹; vgl. Deliège/Mélen 1997, 388.

›Empfänger‹ (Hörer) auf – ob Narrative primär durch den Stimulus gegeben oder aber vom Rezipienten konstruiert werden. Die kognitive Musikanalyse befasst sich mit der letzteren Auffassung,[24] und der in diesem Artikel vertretene Forschungsansatz stützt sich auf Daniel Rothbart (2004), demzufolge Narrative wie Modelle »konzeptionelle Konstruktionen sind, die unter der Kontrolle eines Geschichtenerzählers stehen.«[25] Elisa Johanna Känd hat Narrativität in Musikvideos mit Hilfe der musikalischen Spannung (*slider controller*-Experiment) analysiert.[26]

Mit COSM wurde vom Autor dieses Artikels ein kognitives und seit 2017 empirisch anwendbares Modell entwickelt, bei dem die Teilnehmenden des Experiments nicht nur behavioristische Subjekte, sondern Ko-Analysten und »Geschichtenerzähler« sind,[27] die ihre narrativen Strategien somit nonverbal (nicht in linguistischen Satzstrukturen) und zeitlich unmittelbar wiedergeben können. Dem Autor ist keine mit COSM vergleichbare Analysemethode bekannt, allerdings könnte man diese als eine Art Realisierung der Vorschläge von Wilfried Gruhn zur narrativen Musikanalyse betrachten.[28]

Die Daten dieses Experiments sind sowohl qualitativ (individuell) als auch quantitativ (mathematisch-statistisch) analysierbar. Allerdings sind mathematisch-statistische Korrelationen (hier zwischen den Daten von COSM und der musikalischen Spannung) nicht automatisch Nachweise eines direkt begründbaren Zusammenhangs, da bei realer Musik als Stimulus die musikalischen Parameter miteinander verwoben sind, so dass sich nicht eindeutig feststellen lässt, welche Parameter einander gegenseitig bedingen und welche schließlich auf die musikalische Spannungswahrnehmung einwirken. Darüber hinaus sind die psychophysischen Skalen der Datenerfassung und -analyse arbiträr.[29]

Das Problem, dem sich diese Studie widmet, besteht darin, dass die musikalische Spannung als komplexes Phänomen weder absolut noch direkt gemessen oder analysiert werden kann:

i) Automatische Intensitäts-Analysen von Tonaufnahmen sind an die jeweilige Interpretation[30] sowie an Aufnahme- bzw. Mastering-Techniken gebunden.
ii) Aus der Partitur bzw. aus MIDI-Daten kann nur eine begrenzte Zahl von Parametern herausgelesen werden (sofern diese vorher entsprechend programmiert wurden).
iii) Die Parameter sollten nicht ohne Grund linear kumulativ summiert werden.

Dies hat zur Entwicklung von COSM als audiovisuellem Salienz-Modell geführt.[31] Als Grundlage der Analyse von empirischen Daten wurde der DBSCAN-Algorithmus verwendet (*Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with*

---

[24] Chouvel 2014.

[25] Rothbart 2004, xi.

[26] Känd 2016.

[27] Gruhn 1992.

[28] Vgl. ebd.

[29] Pressnitzer/McAdams/Winsberg/Fineberg 2000.

[30] In einer Komposition sind die Werte der Parameter mehr oder weniger genau in der Partitur fixiert (präskriptiv). Vom Interpreten werden diese Werte entweder so exakt wie möglich oder mit improvisatorischer Freiheit zum Klingen gebracht. Die zu untersuchenden Zusammenhänge zwischen der Partitur und dem Erklingenden sind daher meistens approximativ.

[31] Vgl. die Entwicklung interdisziplinärer konzeptueller und kognitiver Modelle bei Zbikowski 1997, 2002.

Noise),[32] mit dem sich mit der Cluster-Bildung der größtmögliche Zusammenfall der Reaktionen der Teilnehmenden in einem definierten Zeitfenster (Radius um jeden Impuls 2 Sekunden, mindestens 10 Reaktionspunkte; zur Visualisierung des Prinzips siehe Abbildung 4) ermitteln lässt und somit eine statistische Grundlage für den Vergleich der Teilnehmenden-Daten gegeben ist.

Das Ziel dieser Studie ist, die wahrgenommene strukturelle musikalische Spannung in zeitgenössischer post-tonaler Orchestermusik (auf globaler Ebene, auch als Form, holistisch) zu verstehen mit Hilfe der auditiven Salienz musikalischer Ereignisse bzw. Impulse (auf lokaler Ebene) sowie der audiovisuellen Salienz musikalisch primärer und sekundärer Parameter von zuvor durch die Teilnehmenden bestimmten musikalischen Impulsen (ebenfalls auf lokaler Ebene). Hierbei werden Prinzipien der Kognitionsforschung (kognitive Analyse, Salienz, Spannung) mit musiktheoretischen Konzepten und Termini (musikalische Parameter) über das Phänomen des ›musikalischen Impulses‹ miteinander verbunden.

Die Kernfrage dieses Artikels ist, welche musikalischen Parameter mit welchen Tendenzen der strukturellen Spannungsentwicklung (steigend, unveränderlich, fallend) in der post-tonalen Musik von Tüür statistisch und narrativ in Verbindung gebracht werden können.

Folgende Fragen und Hypothesen werden aufgeworfen:

1. Sind bei post-tonaler Musik die sekundären Parameter salienter als die primären?[33]
- Hypothese 1: Die Veränderung der musikalischen Spannung zeigt zugleich eine Veränderung bei sekundären Parametern und umgekehrt.

2. Wie lässt sich die Bedeutung der musikalischen Parameter für das Entstehen von struktureller Spannung in post-tonaler Musik narrativ interpretieren?
- Hypothese 2: Saliente Veränderungen im Fluss der Musik (musikalische Spannung), die durch *slider controller* global und lokal sowie mit COSM salienten musikalischen Parametern lokal angezeigt werden, weisen deutliche Regelmäßigkeiten auf.

Hierbei wird der Grad der Regelmäßigkeit als Funktion von Narrativität verstanden: größere Regelmäßigkeit verweist auf klarere Verständlichkeit, somit deutlichere Narrativität. Bei der Datenanalyse (*Data Mining*) erregen gerade solche Regelmäßigkeiten die Aufmerksamkeit der Analysierenden.

In der Musik selbst ist das Hauptkriterium für Regelmäßigkeit die Wiederholung von musikalischen Elementen (Parametern, Motiven, Formabschnitten) in identischer oder variierter Weise (siehe Prinzipien der paradigmatischen Analyse,[34] der »zygonischen Theorie« Ockelfords[35] oder der Ähnlichkeits-Analysen bei *Data Mining*), wobei der Grad der Variation zu einem wahrnehmungspsychologischen Spannungsanstieg oder -abfall führen kann. Bei identischer Wiederholung im Bereich primärer Parameter sowie gleichbleibender sekundärer Parameter (z. B. Dynamik und Tempo) wird generell von einem Spannungsabfall ausgegangen, da nichts Neues hinzukommt, das für die Hörenden salient, somit spannungsmäßig attraktiv wäre. Der steigende Grad der Variativität kann zu

---

[32] Ester/Kriegel/Xu 1996.

[33] Dies wird seit den 1960er Jahren postuliert. Vgl. Bauer 2001, 38.

[34] Wiggins 2010.

[35] Ockelford 2009.

einem Spannungsanstieg führen beim Hervortreten des Neuen (salient, attraktiv) im Vergleich mit dem unmittelbar zuvor erklungenen Bezugselement. Vermutlich spielt die individuelle Hörstrategie bei der Entscheidung darüber, was als salient gelten soll, eine entscheidende Rolle.

THEORETISCHE GRUNDLAGEN DES COSM-MODELLS

Musikalische Ereignisse, Salienz und Impuls

Ein ›musikalisches Ereignis‹ kann definiert werden als ein strukturelles. Es wird somit im Kontext anderer musikalischer Ereignisse zu einem hervorragenden und auffälligen bzw. salienten Ereignis, wenn es sich (aufgrund der unterschiedlichen Werte verschiedener musikalischer Parameter) im Fluss (*flux*, *flow*) anderer musikalischer Ereignisse von seiner unmittelbaren Umgebung[36] unterscheidet bzw. abhebt.

Salienz als Grad der Unterscheidung derjenigen musikalischer Ereignisse von einem Hintergrund, die eine Veränderung im Fluss der Musik indizieren, ist ein bisher selten systematisch untersuchter Aspekt sowohl in der Musikanalyse[37] als auch bei sonstigen Zeitskalenanalysen.[38] Beide Analyseformen beinhalten Segmentations- und Reduktionsprozesse (ähnlich wie bei der Schenker-Analyse), bei denen saliente, sogenannte »wichtige perzeptuelle Punkte«[39] über verschiedene Ebenen hinweg erhalten bleiben.

Als ›salientes musikalisches Ereignis‹ bezeichnet der Autor dieses Artikels den sogenannten »Impuls«[40] in der Definition als »Moment« oder »Ereignis der Veränderung«, das perzeptiv und kognitiv indizierbar ist – unabhängig von seinem Anfang oder Ende. Dieser Impulsbegriff unterscheidet sich von dem der Neurologie und Hirnforschung dadurch, dass es hier zu einer »frühen« Verarbeitung von musikalischen Parametern wie Frequenz (Tonhöhe) und Lautheit (Dynamik) kommt. Der analysierte Impuls als elektrisches Signal im Bereich von Millisekunden bewegt sich also auf einer völlig abweichenden Zeitskala.[41] Saliente musikalische Impulse sind deutlich langsamer. Ein solcher Impuls regt als

---

[36] Auf den Umwelt-Aspekt bei der Wahrnehmung von Musik stützen sich auch Deliège/Mélen/Stammers/Cross 1996; Jensenius 2002 (mit dem Begriff *Soundscape*) und Reybrouck 2015. Dieser Vergleich ist sinnvoll, wenn man die Salienz von Klangparametern und somit den Akt der Analyse durch die Wahrnehmenden als universell-biologisch betrachtet (siehe auch Huron 2006). Wenn Musik und insbesondere der Stil bestimmter Komponistinnen und Komponisten für die Hörenden überwiegend unbekannt ist – so auch die Musik von Erkki-Sven Tüür – wird man beim Hören ähnlich herausgefordert, als wenn man sich in einer unbekannten Umwelt (natürlich, aber auch urban) bewegt und orientieren muss: in beiden Fällen ist man mit bekannten Parametern, aber auch unbekannten Kombinationen von Parametern konfrontiert. Darüber hinaus kann Musikwahrnehmung auch als erweitertes rechnergestütztes kognitives System (*extended computational cognitive system*) im Umweltkontext (*ecological acoustics*) betrachtet werden. Vgl. Kersten 2014.

[37] Jensenius 2002; Lerdahl 1989, 2001.

[38] Fu 2011.

[39] Ebd., 1.

[40] Lock 2010a und b.

[41] Vgl. Koelsch/Siebel 2005, 579; Koelsch 2013, 90; Purwins/Herrera/Grachten/Hazan/Marxer/Serra 2008; Abbott 2002, 12; Snyder 2000.

Stimulus die Aufmerksamkeit der Hörerenden an,[42] leitet ein »mentales Hindeuten« oder den »Akt der fokalen Aufmerksamkeit« ein.[43]

›Salient‹ (hervorragend, bedeutsam, auffällig) können in Musik bestimmte Merkmale sein, darunter primäre (PP) und sekundäre Parameter (SP).[44] Zu den PP zählen Tonhöhen-, Rhythmus- und Harmoniestrukturen, die im Detail eher von erfahrenen Musiker*innen (Expert*innen)[45] analytisch unterschieden werden können. Zu den SP zählen nach Zohar Eitan und Roni Y. Granot[46] die sogenannten extra-musikalischen (›natürlichen‹) auditiven perzeptiven Aspekte wie Lautheit (oder Dynamik), Tonhöhen-Register und Klangfarben, welche für weniger trainierte oder Nicht-Musiker*innen (sogenannte naive Hörer*innen)[47] besser wahrnehmbar, attraktiver sind. Auch Irène Deliège, Marc Mélen, Diana Stammers und Ian Cross[48] unterscheiden zwischen Musiker*innen, deren kognitive Repräsentationen zugänglicher und detaillierter sind als die von Nicht-Musiker*innen.

Musikalische Spannung und temporale Musikanalyse

Wie die Höhepunkt-Analysen von Kofi Agawu für tonale Musik[49] und Martin Kirschbaum für post-tonale Musik[50] gezeigt haben, ist musikalische Spannung ein komplexes Phänomen. Es wird auch als kognitives Gegensatzpaar ›Spannung/Entspannung‹ begriffen, wie dies in Fred Lerdahls und Ray Jackendoffs ›Generativer Theorie der Tonalen Musik‹ (GTTM)[51] mit Bezug auf Salienz auch für atonale Musik unternommen wurde.[52] Ein weiterer verwandter Begriff ist der der »temporalen dynamischen Form«.[53] Auch der Expektanz-Aspekt[54] beinhaltet die musikalische Spannung, u. a. in der ITPRA-Theorie (*Imaginative, Tension*, *event onset, Prediction*, *Reactive*, *Appraisal*) von David Huron.[55]

Die strukturelle musikalische Spannung (definierbar durch das Zusammenwirken strukturbildender musikalischer Parameter) ist bisher hauptsächlich für tonale Musik und deren primäre Parameter, weniger für post-tonale Musik mit deren primären sowie sekundären Parametern untersucht worden. Studien zur musikalischen Spannung in post-tonaler Musik veröffentlichten Anna Rita Addessi und Roberto Caterina (2000, zur musikalischen Form und Segmentation), Lerdahl (2001, zur atonalen Musik und kognitiven

---

[42] Vgl. Deliège/Mélen 1997.

[43] Reybrouck 2010, 2015.

[44] Meyer 1989; Snyder 2000.

[45] Lehrdahl 2001, 143.

[46] Eitan/Granot 2009, 143.

[47] Lehrdahl 2001, 143.

[48] Deliège/Mélen/Stammers/Cross 1996, 128.

[49] Agawu 1982.

[50] Kirschbaum 2001.

[51] Lerdahl/Jackendoff 1983.

[52] Lerdahl 1989, 2001.

[53] Vines/Nuzzo/Levitin 2005; Mailman 2010.

[54] Neuwirth 2008. Zu musikalischer Expektanz und Spannung in klassischer und Jazz-Harmonik vgl. Rohrmeier 2013, zur Erwartungssituation der performativen Analyse tonaler und post-tonaler Musik vgl. Utz 2013.

[55] Huron 2006.

Salienz), Kirschbaum (2001, zur Höhepunktbildung und Dramaturgie), Addessi (2010, zur »Makroform« sowohl empirisch als auch Expert*innen-Analyse sowie zur Wahrnehmung von Spannung und Entspannung), Lock (2010a und b, zur Form des ›Impuls‹ genannten musikalischen Ereignisses als salientem Merkmal), Granot und Eitan (2011, zur Interaktion von dynamischen Hörparametern), Lock und Kotta (2012, zu Form, Impuls und musikalischer ›Energie‹) sowie Yvonne Teo (2020, zur harmonischen Spannung in neoklassischer Musik und zum Verhältnis zwischen theoretischer und wahrgenommener Spannung).[56]

Morwaread M. Farbood,[57] David Pressnitzer, Stephen McAdams, Suzann Winsberg, Joshua Fineberg,[58] Fred Lerdahl und Carol Krumhansl[59] sowie Yvonne Teo[60] haben prognostizierende Modelle entwickelt, die sie mit empirischen Daten vergleichen. Unter den empirischen Studien befinden sich solche, die entweder komplette Werke oder nur Ausschnitte, aber auch speziell komponierte Stimuli verwenden; die Mehrzahl der empirischen Studien arbeitet mit ›kontinuierlichen‹ Selbst-Report-Technologien, u. a. durch *slider controller*.[61]

Musikalische Spannung als Gegenstand empirisch-psychologischer Untersuchung ist auch Teil der Emotionsforschung.[62] Dieser Bereich wird im vorliegenden Artikel jedoch ausgeklammert, da das Hauptaugenmerk des COSM-Modells auf der Analyse von Musik und nicht etwa auf der Analyse des durch sie hervorgebrachten emotionalen Zustands der Gespanntheit bei den Teilnehmenden liegt. Allerdings wird hier auch von der Unterscheidung zwischen ›gefühlter‹ (*felt*) und ›wahrgenommener‹ Spannung (*perceived tension*) ausgegangen, die Moritz Lehne und Stefan Koelsch[63] bei ihrem generalisierten Modell von psychologischer Spannung und Erwartungsungewissheit (im Kontext der Analyse von Musik, Film, Literatur und alltäglichen Lebenstätigkeiten) voraussetzen. Eine analoge Unterscheidung, wie sie zunächst Alf Gabrielson (2002) für gefühlte (›subjektive‹) und wahrgenommene (›objektive‹) Emotion vorschlug, wurde von Kari Kallinen (2006) sowie Kallinen und Niklas Ravaja (2006) experimentell untersucht.

In musiktheoretischen Kontexten werden die Hörenden oftmals als erfahrene und kompetente »ideale Hörer« aufgefasst. Bezüglich des »erstmaligen Hörens«, wie es bei traditionellen Musikanalysen vorausgesetzt und in der klassischen empirischen Forschung gefordert wird, formuliert Markus Neuwirth die Auffassung, ein idealer Hörer sei »eine logische Unmöglichkeit«: »Bestimmte Aspekte der Musik, die die Analyse unterstellt, kann ein Hörer prinzipiell nicht nachvollziehen, ohne einen atemporalen, quasi-architektonischen Standpunkt einzunehmen«[64] (ähnlich wie Carl Dahlhaus' Auffassung

---

[56] Eine Literaturübersicht bietet Teo 2020, 61–64.

[57] Farbood 2006, 2012.

[58] Pressnitzer/McAdams/Winsberg/Fineberg 2000 vor allem über Orchester-Timbres und psychoakustische Rauheit in post-tonaler Musik. Zur Rolle von Rauheit und Inharmonizität in der neurodynamischen Modellierung musikalischer Spannung s. Hadrava/Hlinka 2020.

[59] Lerdahl/Krumhansl 2007.

[60] Teo 2020.

[61] Literaturübersichten siehe Granot/Eitan 2011, 219–221 und Farbood 2012, 387–391.

[62] Juslin/Soboda 2010; Lehne/Koelsch 2015.

[63] Lehne/Koelsch 2015, 3, 7.

[64] Neuwirth 2008, 564.

von »Wagners dramatisch-musikalischem Formbegriff«[65] oder die retrospektive Betrachtung musikalischer Form als Abstraktion von Musik im Sinne räumlicher Analogie bei Ligeti).[66]

Daraus folgend, wird in dieser Studie anstelle des nach Neuwirth[67] sozusagen ›logisch unmöglichen idealen Hörers‹, welcher zugleich eine atemporale Position einnehmen müsste, ein sogenannter kontextspezifischer, im ständig sich verändernden Jetzt-Moment der »psychologischen Gegenwart«[68] sich bewegender sozusagen ›realistischer Hörer‹ angenommen, welcher im Hörprozess das Gehörte und zu Hörende (ähnlich wie auch sich in der Umwelt befindliche auditive Eindrücke) entweder stetig neu bewertet (oberflächlich, resourcenaufwendig) oder eine kognitive Strategie entwickelt, nach der einige Parameter im Verlauf des Hörens als salienter betrachtet, andere dagegen ausgeblendet werden, was schließlich zur Ausbildung eines Narrativs führt.

Kognitive Narratologie und narrative Musikanalyse

Nach Achim Saupe und Felix Wiedemann[69] ist der narrative Konstruktivismus (ein »Zusammensetzen der Geschehnisse«) nicht unumstritten, gleichwohl lässt sich das Wesen einer Geschichte »in der narrativen Strukturierung und Artikulation« erkennen.[70] Die kognitive Narratologie[71] ist ein Untergebiet der narrativen Analyse der Literaturwissenschaft. Hierbei wurden u. a. empirische Studien zur Untersuchung von Korrelationen zwischen Merkmalen in Texten und prozessualen Strategien, die diese auslösen, durchgeführt.[72] Ähnlich funktioniert auch COSM (siehe dessen audiovisuelle Prinzipien sowie die Visualisierung und Analyse der Daten).

Narrative Methoden hält Gruhn für vielversprechend, zugleich sieht er die Notwendigkeit der Systematisierung und Verfeinerung der Methodologie.[73] Sein Analyseverfahren sieht vor, assoziative Verbindungen zwischen den »reflektierenden Hörgeschichten« statistisch zu erfassen, zu quantifizieren und in Netzstrukturen abzubilden.[74] Mit COSM ist die Möglichkeit gegeben, Hörstrategien als Narrative nonverbal zu erfassen und statistisch quantifiziert darzustellen.

Byron Alméns fünf Punkte des musikalischen Narrativs[75] treffen auch auf Prinzipien des Kognitiven Achteckigen Scheibenmodells COSM zu und werden im folgenden Abschnitt in Bezug auf das Modell sowie auf die Musik Erkki-Sven Tüürs erläutert:

---

[65] Dahlhaus 2004, 326.

[66] Ligeti 1966, 23 f.

[67] Neuwirth 2008, 562 f.

[68] Ebd.

[69] Saupe/Wiedemann 2015.

[70] Ebd.

[71] Herman 2009.

[72] Ebd., 79.

[73] Gruhn 1992.

[74] Ebd., 49.

[75] Almén 2003, 12 f.

1. Interaktion von musikalischen Elementen.
2. Begrenzte Zahl von narrativen Archetypen.
3. Muster-Organisation durch Konflikt zwischen zwei oder mehr hierarchisch auf einander bezogenen Elementen innerhalb eines Systems mit Hilfe der Neubewertung konstituierender Elemente.
4. Die Muster weisen dem Medium entsprechende einzigartige Oberflächen-Merkmale auf, so auch in der Musik.
5. Die Muster müssen psychologisch bedeutsam sein; ihr häufiges Erscheinen als prägendes Prinzip in den Künsten basiert nicht auf logischer Notwendigkeit oder Zufall, sondern auf der Signifikanz dieser Muster für die Wahrnehmenden auf verschiedenen Ebenen.

Die interagierenden musikalischen Elemente (1) sind bei COSM 8 musikalische Parameter und Kontrast. Narrative Archetypen (2) finden sich auf der syntaktischen Ebene: Sonatenform, Sonaten-Zyklus, Anfang-Mitte-Ende, Kulminationsdramaturgie mit Hauptkulmination, und speziell für Tüürs Musik auf parametrischer Ebene: »schnelle«, »langsame«, »kontinuierliche« Zustände der Zeit,[76] die durch die entsprechende Behandlung der Parameter Tonhöhenbewegung, Rhythmus, Tempo erreicht wird. Der Grundkonflikt bei der Muster-Organisation (3) in Tüürs Musik besteht in Kontrasten innerhalb des musikalischen Materials und zwischen unterschiedlichen Kompositionstechniken; in seinen nach 2002 entstandenen Werken nehmen Übergänge zwischen diesen Kontrasten – speziell zwischen den drei temporalen Zuständen der Musik – immer mehr Raum ein. Bezüglich einzelner Parameter erscheinen in diesen Werken die Kontraste für die Hörenden individuell, konstituierende Parameter werden zu Beginn durch die Hörenden festgelegt, und zwar diejenigen mit dem höchsten Salienz-Grad. Die einzigartigen Oberflächen-Merkmale der Muster (4) sind in Tüürs Musik die im vorigen Punkt genannten Kontraste sowie nach 2002 die ebenfalls bereits erwähnten Übergänge zwischen den drei temporalen Zuständen der Musik. Die psychologische Bedeutsamkeit dieser Muster (5) wird in COSM durch das Konzept der audiovisuellen Salienz erfasst.

ZUR EMPIRISCHEN METHODIK DIESER STUDIE

Analyse-Gegenstand, Teilnehmende und Prozedur

Gegenstand der Studie ist eine Einspielung der 4. Sinfonie bzw. des Schlagzeugkonzerts *Magma* von Erkki-Sven Tüür (2002, Aufführungsdauer 30:58 Min.,[77] Estnisches Nationalorchester ERSO, Dirigent: Paavo Järvi). Seine Musik erscheint fließend ohne klare Grenzen, wellenförmig, sich ständig verändernd, zugleich kontrastiv und kulminationsdramaturgisch ausgerichtet. Nach Kerri Kotta[78] kann sie mit Hilfe der Prinzipien der »schnellen«

---

[76] Kotta 2008, 2011.

[77] Zu möglichen Ermüdungseffekten der Teilnehmenden aufgrund der langen Dauer sowohl der Musik wie auch des Experiments ist Folgendes festzuhalten: Da die Dauer des Experiments vom Teilnehmer selbst abhängt – je nachdem, wie viele Impulse indiziert werden und wie detailliert die Parameter analysiert werden –, können keine generellen Vorgaben festgelegt werden. Es ist davon auszugehen, dass die Teilnehmenden entsprechend ihren persönlichen Dispositionen (z. B. Big-Five-Modell) mehr oder weniger intensiv reagieren und somit in der Lage sind, ihre Kräfte einzuteilen.

[78] Kotta 2008, 2011.

(aktiver Rhythmus), »langsamen« (klare Linear- und Harmoniestruktur) und »kontinuierlichen« (andauernde Klang-Faktur) Zustände der musikalischen Zeit analysiert werden. Bis 2002 überwiegen bei Tüür Kontraste innerhalb des musikalischen Materials und zwischen unterschiedlichen Kompositionstechniken (u. a. tonal versus frei-atonal versus dodekaphon), nach 2002 nehmen Übergänge zwischen diesen Kontrasten, speziell zwischen den drei temporalen Zuständen der Musik immer mehr Raum ein. Dies führt u. a. dazu, dass Formgrenzen hauptsächlich bei kulminativen Höhe- und Tiefpunkten anzusiedeln sind, da es keine klassisch klaren Abschnitte mehr gibt.[79]

Die empirischen Experimente fanden im Herbst/Winter 2017 an der Tallinner Universität statt; mit zwei von insgesamt vier Expertinnen und Experten fanden die Experimente außerhalb Tallinns statt. Die Teilnehmenden des ersten Experiments (N=26) waren Bachelor-Studierende zum einen des Studiengangs »Integrative Kunst, Musik und Multimedia« aus drei Jahrgängen, zum anderen des auslaufenden Musikstudiengangs sowie eine Gymnasiastin mit musikalischer Ausbildung (n=22; Altersdurchschnitt: 24,3 Jahre, Altersspanne: 17–37 Jahre). Unter den Expert*innen (n=4; professionelle Komponist*innen, Dirigent*innen und Musiker*innen) befand sich ein Magisterstudent mit Bachelor-Abschluss, drei Personen mit (teils mehreren) Master-Abschlüssen, darunter zwei Doktorand*innen (Altersdurchschnitt: 38,8 Jahre, Altersspanne: 31–50 Jahre).

Während des ersten Experiments zeigten die Teilnehmenden durch *slider controller* die ›wahrgenommene‹ strukturelle musikalische Spannung an. Sie wurden darum gebeten, sich beim Hören auf die Analyse der Musik zu konzentrieren, nicht auf den sich dadurch eventuell einstellenden eigenen emotionalen Spannungszustand. Diese Vorgabe stützt sich auf Lehne und Koelsch[80] und beruht auf der Unterscheidung zwischen ›wahrgenommener‹ (›objektiver‹) und ›gefühlter‹ (›subjektiver‹) Emotion.

Im zweiten Experiment analysierten Teilnehmende aus demselben Feld (N=14, darunter Bachelor-Musikstudierende n=10; Altersdurchschnitt: 23,3 Jahre, Altersspanne: 19–37 Jahre; ferner dieselben Experten n=4) mit Hilfe von COSM zunächst die Salienz von Impulsen (Momenten der Veränderung). Anschließend fixierten sie quantitativ die Salienz von acht vorgegebenen musikalischen Parametern[81] zum Zeitpunkt der zuvor individuell angezeigten Impulse.

Empirisches Experiment-Design

Das erste Experiment dieser Studie durch *slider controller*[82] ergibt ebenso wie bei Farbood (2006, 2012) kontinuierliche Echtzeit-Daten. Das zweite Experiment (COSM) ergibt in der ersten Phase (i) diskrete Echtzeit-Daten, in der zweiten Phase (ii) diskrete retrospektive Bewertungen.

---

[79] Vgl. Lock 2010a, 65.

[80] Lehne/Koelsch 2015, 3.

[81] Primäre Parameter (PP): Tonhöhe, Rhythmus, Harmonie; Sekundäre Parameter (SP): Dynamik, Tempo, Instrumentation/Timbre, Faktur, Effekte.

[82] Der TEDEA (Tension Design Apparatus) wurde 2011 von Gerhard Lock (Konzeption) und Hans-Gunter Lock (Software *Max/MSP*, Arduino, Hardware) entwickelt. http://www.schoenberg.ee/tension-design/tension-design.html (11.12.2020)

1. ›Holistisches‹[83] und ›aufmerksames‹ Hör-Experiment: ›kontinuierliche‹[84] Erfassung von Veränderungen in der Musik als Ausdruck der ›wahrgenommenen‹ musikalischen Spannung durch *slider controller* über das ganze Werk hinweg (ohne Unterbrechung).
2. ›Holistisches‹ Hörexperiment: diskrete Erfassung von Veränderungen in der Musik:
   i. ›aufmerksames Hören‹ (*attentive listening*):[85] Anzeigen von holistischen musikalischen Ereignissen (Impulsen) als Indikatoren für musikalische Spannung über das ganze Werk hinweg (ohne Unterbrechung).
   ii. ›aufmerksames Hören‹ (*attentive listening*) und ›analytisches Hören‹:[86] die individuellen Impulse (mit ihrer Umgebung -/+7 Sekunden), die beim ›holistischen‹ Experiment angezeigt wurden, werden als Fragmente (*chunks*) (Gesamtdauer 14 Sekunden, siehe Abbildung 1) für wiederholtes Hören durch jeden Teilnehmenden separat benutzt, um die Salienz der gegebenen musikalischen Parameter zu bestimmen, die durch die COSM-Scheibe die Analogie der Salienz sowohl im visuellen als auch im auditiven Bereich hinzufügen.

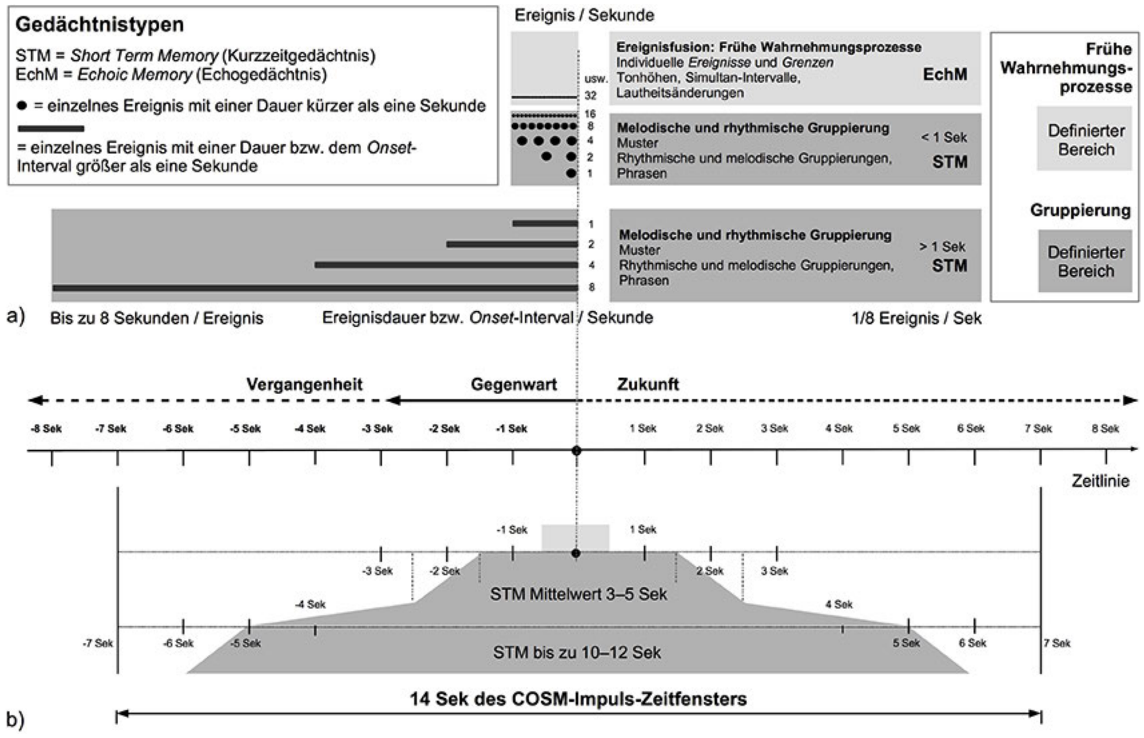


Abbildung 1: Gedächtnis-Aspekte bei der Wahrnehmung von a) Ereignissen und primären musikalischen Parametern, b) der temporalen Dimension, Gedächtnis-Mittelwerte (nach Snyder 2000) zur Begründung des -/+7=14 Sekunden COSM-Impuls-Zeitfensters.

---

[83] Eitan/Granot 2009.

[84] Der Ausdruck ›kontinuierlich‹ ist technisch nicht ganz korrekt, da das *Max/MSP*-Programm *slider controller*-Daten nur alle 500 ms erfasst, und gleichbleibende Datenpunkte nicht mitgeschrieben werden aus Gründen der ökonomischeren Datenweiterverarbeitung.

[85] Deliège/Mélen 1997, 388.

[86] ›Analytisches Hören‹ nach Eitan/Granot 2009 entspricht dem ›synthetischen Hören‹ nach Purwins/Herrera/Grachten/Hazan/Marxer/Serra 2008.

Zwei Phasen von COSM

In der ersten Phase von COSM werden musikalische Ereignisse analysiert, die wahrnehmbar sind, indem sich im als Impuls bezeichneten Moment für die Hörenden etwas ändert in der Musik. Was genau diese Änderung hervorgerufen hat, bleibt in der ersten Phase des Experiments noch verborgen. Die Impulse werden von den Teilnehmenden des Experiments im Audioverarbeitungsprogramm *Reaper* mit der M-Taste in Echtzeit während des ununterbrochenen Musik-Hörens ausgewählt und gespeichert. Hierbei sehen die Teilnehmenden im leeren *Track* (der Aufnahme) weder den Maus-Cursor noch die Wellenform mit laufender Zeitmarkierungs-Positionsline, um nicht von dieser visuellen Information beeinflusst zu werden. Die Impulse werden anschließend als Zeitcodes aus *Reaper* exportiert und in der zweiten Phase in das Programm *Max/MSP* importiert, mit der die COSM-Software von Hans-Gunter Lock geschrieben wurde.

In der zweiten Phase ist es das Ziel der Modellierung, den ›Inhalt‹ dieser Impulse greifbar zu machen, d. h. zu ermitteln, welche Merkmale/Eigenschaften (*features*, *properties*) im Fluss der Musik auftauchen (*emergent*) und sich als auffällig salient erweisen. Hierbei ist es das Ziel, jene musikalischen Parameter, die die in der ersten Phase im Impuls-Moment wahrgenommene Änderung hervorgerufen haben, in genau diesem Moment zu bestimmen und zu quantifizieren. Dabei wird einkalkuliert, dass speziell in komplexer Orchestermusik alle wesentlichen Parameter[87] miteinander verwoben sind und ineinanderwirken – auch dann, wenn sie (aus welchen Gründen auch immer) nicht einzeln wahrgenommen werden (können).

Audiovisuelle Prinzipien von COSM

Für die Modellierung wurde der Umriss eines Achtecks (Oktogon) als geometrische Figur gewählt (siehe Abbildungen 2 und 3), dessen Ecken durch die quasi 3D-Darstellung leicht seitlich von vorn sowie ungefähr im 45-Grad-Winkel von oben unterschiedlich hervortretend (salient) sind. In Analogie zu dieser visuellen Salienz sind in jedem vom jeweiligen Experiment-Teilnehmenden zuvor ausgewählten Impuls alle acht Parameter immer vorhanden, aber sie sind für verschiedene Hörende unterschiedlich salient, mehr oder weniger auditiv hervortretend. Ihre Salienz hängt davon ab, welche Parameter für welche Hörenden in welchem Umfang und in welcher Kombination perzeptiv überwiegen bzw. attraktiv sind. Die Hörenden entwickeln dabei im Laufe des Experiments in Abhängigkeit von der Musik und ihren eigenen Vorlieben, aber auch von ihren jeweiligen perzeptiven und kognitiven Fähigkeiten und Erfahrungen eine eigene Strategie, die qualitativ sowohl über die Musik als auch über die Hörenden etwas aussagt. Daraus lässt sich wiederum ein analytischer Ansatz entwickeln, da für die zweite Phase sämtliche Impulse innerhalb eines -/+ 7 Sekunden dauernden Fragments, eines sogenannten *chunks* mit der Gesamtdauer von 14 Sekunden, so oft wie nötig angehört und im Detail analysiert werden können. Ähnlich geht

---

[87] Es sind acht Parameter – PP: Tonhöhe, Rhythmus, Harmonie; SP: Dynamik, Tempo, Instrumentation/Klangfarbe, Faktur und Effekte –, die beinahe mit denen von Kirschbaum (2001) zusammenfallen. Allerdings betrachtet der Autor des vorliegenden Beitrags im Gegensatz zu Kirschbaum die Faktur von Instrumentation und Klangfarbe getrennt und unterteilt den Tonhöhen-Parameter nicht in die Unterkategorien Bewegung und Register, da diese im Kontext der Musikwahrnehmung kaum voneinander zu trennen sind. Eine solche Unterscheidung vermögen weder Amateur*innen noch Nicht-Musiker*innen zu leisten, sondern allein Expert*innen. Das Modell konzentriert sich daher eher auf die Unterschiedlichkeit der gewählten Parameter. Über Kontrast als einen separaten, zusammengesetzten Aspekt äußert sich ebenfalls Kirschbaum.

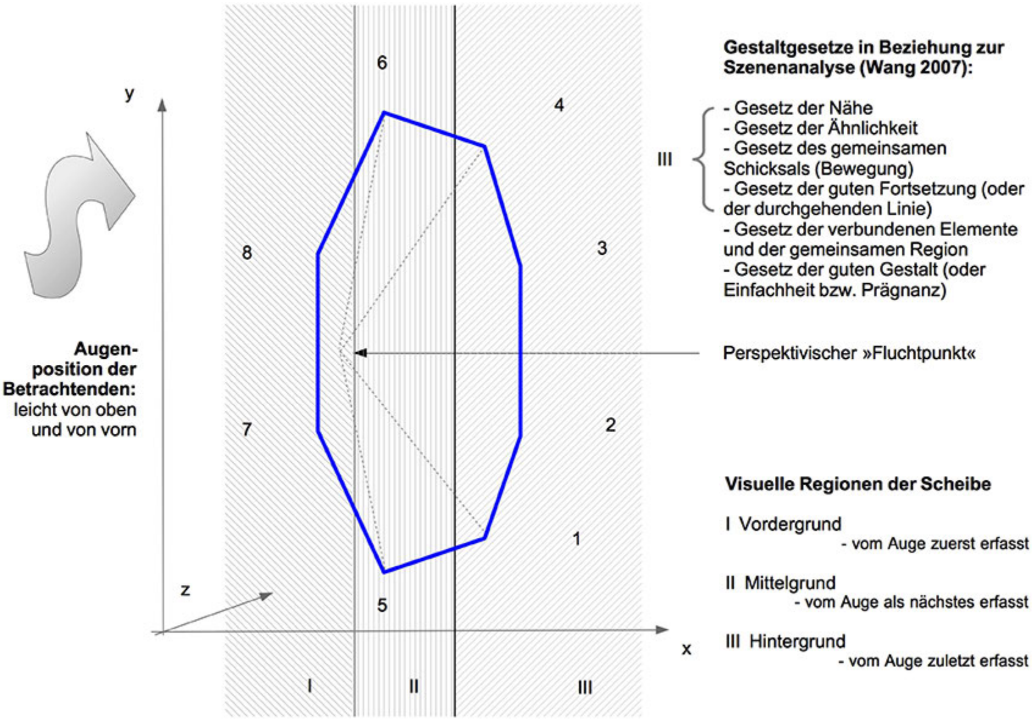


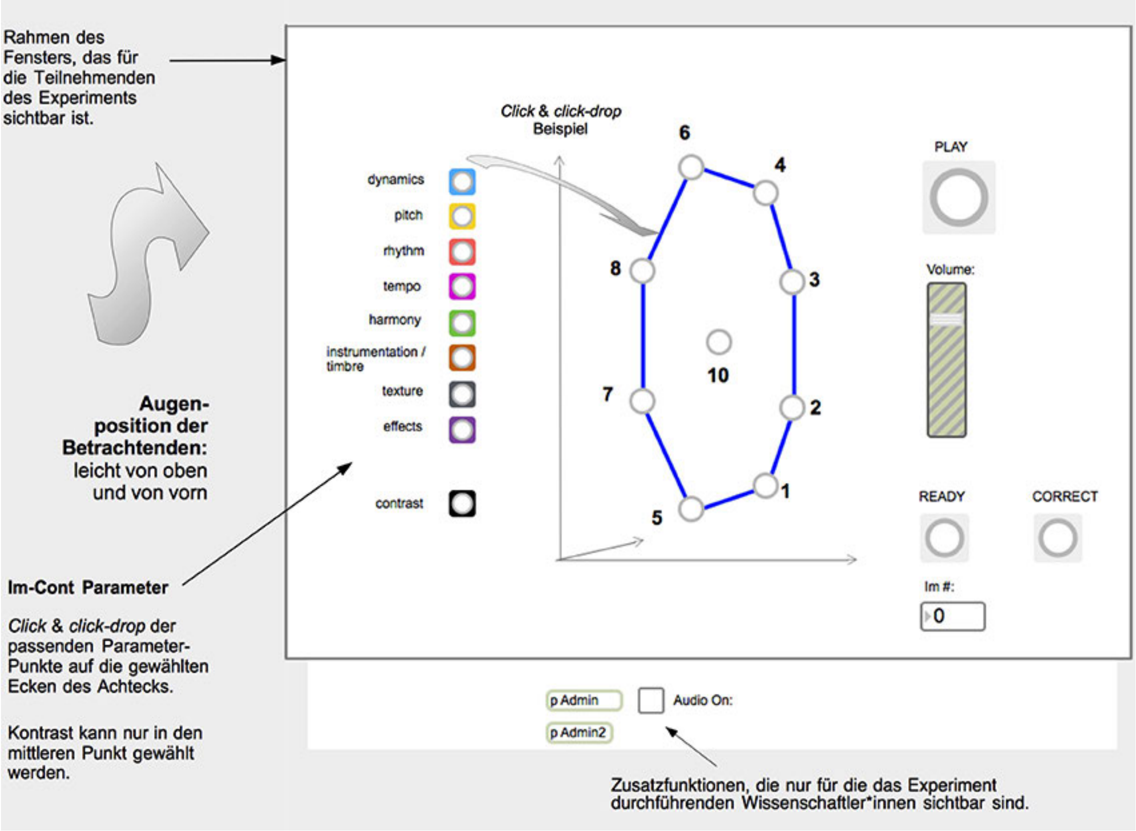
Abbildung 2: Die zentrale geometrische Figur eines Achtecks (Oktogon) für die COSM-Scheibe. Die Ecken des Oktogons befinden sich in mit den römischen Ziffern I–III bezeichneten visuellen Regionen, wobei sie in den Regionen I und II unter dem Gesichtspunkt der visuellen Salienz betrachtet werden. Die Region III funktioniert nach den Gestaltgesetzen, da diese Punkte perspektivisch näher beieinander liegen. Die arabischen Ziffern 8–1 in absteigender Reihenfolge und somit abnehmender Salienz an den Ecken des Oktogons sind die zugeteilten Werte, mit deren Hilfe die jeweiligen Ecken quantifiziert werden können.

auch die *Progressive Exposure Method* (PEM) von Joshua D. Albrecht vor, bei der die Musik in einzelnen Fragmenten (*chunks*) mehrfach angehört werden darf.[88] Allerdings wählen die Teilnehmenden die Fragmente bei COSM in der ersten Phase des Experiments durch Impuls-Erfassung selbst (mit statischer Dauer der oben genannten 14 Sekunden) und analysieren diese chronologisch in der Originalreihenfolge.

Abbildung 3 zeigt die COSM-Scheibe (Oktogon) und die musikalischen Parameter zusammen mit dem Extra-Parameter Kontrast als Experiment-Oberfläche in demjenigen Fenster, das für die Teilnehmenden des Experiments am Bildschirm sichtbar ist, sowie die Zusatzfunktionen, die nur von Personen eingesehen werden können, die das Experiment wissenschaftlich betreuen. Der Richtungspfeil *Click & click-drop* zeigt, dass der Parameter Dynamik mit zwei Mausklicks der gewählten Oktogon-Ecke zugeordnet wird. Der Hintergrund der Dynamik-Schaltfläche (hellblau) verschwindet dann beim Menü und erscheint als Hintergrund der Schaltfläche der jeweiligen Ecke. Das Fenster zeigt alle Funktionen, die die Teilnehmenden während des Experiments wählen können und müssen, um fortfahren zu können: *Play* startet das Fragment, *Volume* ermöglicht eine Justierung der Lautstärke des Fragments, *Correct* erlaubt den Teilnehmenden, ihre vorherige Auswahl neu zu treffen (der Vorgang für den jeweiligen Impuls beginnt von vorn), *Ready* wird gewählt, wenn die Teilnehmenden mit ihrer Auswahl zufrieden sind und den nächsten Impuls analysieren möchten. ›Im#‹ zeigt die Nummer des jeweils zu analysierenden Impulses und ändert sich automatisch, wenn die Teilnehmenden mit der Erarbeitung eines Impulses fertig geworden sind. Die Ziffern an den Ecken des Oktogons sind die zugeteilten Werte, mit deren Hilfe die jeweiligen Ecken quantifiziert werden können.

---

[88] Albrecht 2012.



Rahmen des Fensters, das für die Teilnehmenden des Experiments sichtbar ist.

Augenposition der Betrachtenden: leicht von oben und von vorn

Im-Cont Parameter

Click & *click-drop* der passenden Parameter-Punkte auf die gewählten Ecken des Achtecks.

Kontrast kann nur in den mittleren Punkt gewählt werden.

Zusatzfunktionen, die nur für die das Experiment durchführenden Wissenschaftler*innen sichtbar sind.

Abbildung 3: COSM-Scheibe (Oktogon) und musikalische Parameter als Fenster für Teilnehmende der zweiten Phase des Experiments mit dem Ziel, die Salienz der Parameter für den jeweiligen Impuls zu erfassen

Funktionsprinzip und Visualisierung des DBSCAN-Algorithmus

Durch den DBSCAN-Algorithmus werden die Impulse der am Experiment Teilnehmenden (Expertinnen und Experten n=4, Musikstudierende n=10, zusammen N=14) in 48 Cluster nach folgenden Kriterien aufgeteilt: 10 und mehr Reaktionspunkte (Impulse) mit einem Radius (Epsilon) von 2 Sekunden (bezogen auf die jeweiligen Punkte) sind Zentralpunkte (*core points*); liegen weniger als 10 Reaktionspunkte (Impulse) innerhalb dieses Radius (Epsilon), so werden sie als Grenzpunkte (*border points*) bezeichnet. Die gewählte Mindestanzahl von 10 Punkten bezieht sich auf die Teilnehmenden-Anzahl N=14, weniger als 10 Punkte sind nicht hinreichend signifikant. Das Generierungsprinzip der DBSCAN-Cluster-Punkte ist in Abbildung 4 am Beispiel des Clusters 21 mit Kreisen visualisiert. Die Daten werden am Beispiel der Cluster 21–23 im Audiobearbeitungsprogramm *Audacity* zur Wellenform der Tonaufnahme ins Verhältnis gesetzt (Abbildung 4). Neben den Zeitangaben zu den Impulsen bzw. Reaktionspunkten zeigt das Bildschirmfoto auch die Dauer der dichtesten Kumulation von Impulsen innerhalb eines Clusters, die Dauer innerhalb von Cluster-Paaren sowie die Dauer zwischen einzelnen Clustern/Cluster-Paaren.

Abbildung 4 zeigt, dass die Reaktionen der am Experiment Teilnehmenden relativ genau mit den musikalischen Ereignissen übereinstimmen. Verzögerungen liegen zwischen 100–500 ms, welches die Reaktionszeit-Verspätung von 200–300 ms[89] sowie die sensorische Schwelle des möglichen physiologischen Zeitverzugs[90] (ca. 500 ms) einschließt. Für die Zeit, in der eine Bedeutung im Gehirn sozusagen ›auftaucht‹, geben Stefan Koelsch und Walter A. Siebel 250–550 ms an;[91] für das auditiv-sensorische Gedächtnis zeigen sie 100–200 ms, für die Strukturbildung 180–400 ms; die Reanalyse und Revision finden wesentlich später erst bei 600–900 ms statt.

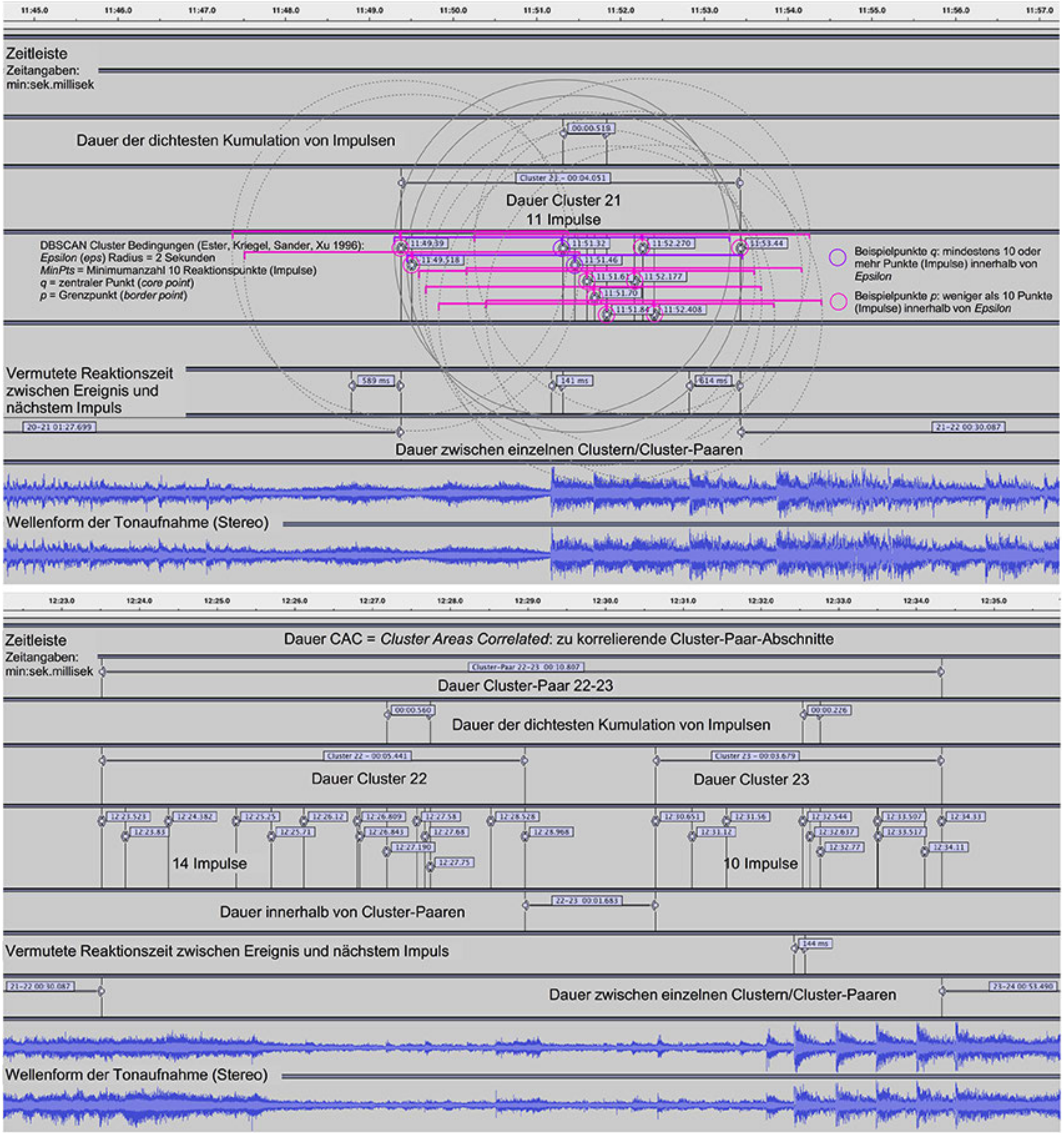


Abbildung 4: Cluster 21 und Cluster-Paar 22–23: Darstellung und Analyse mit Zeitangaben der Impulse (Teilnehmende N=14), sowie bei C21 der Visualisierung des Prinzips des DBSCAN-Algorithmus

---

[89] Vgl. Jensen 1979 und Libet 2004, 54 f.

[90] Libet 2004.

[91] Vgl. Koelsch/Siebel 2005, 570 sowie Koelsch 2013, 90.

Ergebnisse und Analyse

Lokal-Analyse der Cluster 19/21–23/26

Die Auswahl der zu analysierenden Cluster 19–26 erfolgt auf Grundlage der summierten Tendenz der COSM primären (PP) und sekundären Parameter (SP) sowie der Mittelwerte der *slider controller*-Ergebnisse per COSM-Cluster, Cluster-Paare und zu korrelierender Cluster-Paar-Abschnitte (*CAC = Cluster Areas Correlated*). Bis auf die Übergangscluster 17 und 18 umfassen C19–26 im Prinzip den scherzoartigen B-Teil des Werkes (zur Formanalyse siehe Abbildung 10a–f).

Ausgehend von den in der Matrix in Abbildung 9 dargestellten Ergebnissen, die die COSM-Parameter sowie die *slider controller*-Tendenzen für jeden einzelnen Cluster bezüglich des Ausgangs- und Endpunkt-Niveaus auf der Basis der jeweiligen arbiträren psychophysischen Salienzskala vergleichen, ergibt sich, dass insbesondere der einzelne C21 sowie das Cluster-Paar C22–23 die regelmäßigsten Datenmuster für die Parameter zeigen.

An den Clustern 21–23 lässt sich beispielhaft zeigen, wie Parameter aus der Partitur formalisiert werden können (Tabelle 1). Die vereinfachte Partitur (Notation des Autors, Abbildung 5) zeigt die Cluster(-Paare) 21 und 22–23. C21: S. 51–52, Takt 225–226, Zeitdauer 11:49.390–11:53.440 (00:04.051); C22–23: S. 55–56, Takt 238–239 und 240–241, Zeitdauern (C22:) 12:23.528–12:28.968 (00:05.441), (zwischen C22 und 23:) 00:01.683, (C23:) 12:30.651–12:34.330 (00:03.679), (C22–23:) 12:23.528–12:34.330 (00:10.807).

C-Nr.	Primäre Parameter PP (S → E)	Sekundäre Parameter SP (S → E)
C21	**Tonhöhe:** hoch (3) → hoch (3)/mittel (2)/tief (1) rotierend/absteigend (–0,5) **Rhythmus:** kontinuierlich (1) → schnell/rotierend (–0,5)(3)/reliefhaft (1,5)/langsam (2)/ absteigend (–1) **Harmonie:** konsonanzreich (1)/mit Unregelmäßigkeiten (2,5) → chromatisch (3)/ konsonanzreich (1)/mit Unregelmäßigkeiten (2,5)	**Dynamik:** sehr laut (4) **Tempo:** kontinuierlich (0) → schnell (3) (ro-tierend) (–0,5)/langsam (1) **Intrumentation/Timbre:** hell/scharf (3) → hell/mittel/dunkel (3/2/1) **Faktur:** mittel (2) → dicht (3) **Effekte:** Schlagzeug Becken Glockenspiel (hängende Stäbe) (3) → Trommel (1)
C22	**Tonhöhe:** hoch (3)/mittel (2) → mittel (2) **Rhythmus:** kontinuierlich (1)/langsam (2)/ Triller (4) → langsam (2)/schnell (3) **Harmonie:** konsonanzreich (1)/mit Unregelmäßigkeiten (2,5) → konsonanzreich (1)/mit Unregelmäßigkeiten (2,5)	**Dynamik:** laut (3)/ sehr laut (4) → laut (3) **Tempo:** kontinuierlich (0)/langsam (1) → schnell (3)/langsam (1) **Intrumentation/Timbre:** hell/scharf (3)/mittel (2) → mittel/weich (1,5) **Faktur:** mittel (2) → mittel (2) **Effekte:** Streicher glissando (0,5)/Trompeten/ Posaunen scharf (3)/Schlagzeug tremolo (2) → Schlagzeug schelle Punktierung (0,5)
C23	**Tonhöhe:** mittel (2) → ansteigend (1)/mittel/ (2)/tief (1) **Rhythmus:** schnell (3)/mittel (2) → schnell (3)/mittel (2) **Harmonie:** konsonanzreich (1)/mit Unregelmäßigkeiten (2,5) → konsonanzreich (1)/mit Unregelmäßigkeiten (2,5)	**Dynamik:** laut (3) → laut (3) **Tempo:** schnell (3)/mittel (2) kontinuierlich (0) → schnell (3)/mittel (2)/ leicht beschleu-nigend (0,5) **Intrumentation/Timbre:** mittel/weich (1,5) → mittel/weich (1,5)/dunkel (1) Schlagzeug **Faktur:** mittel (2) → dicht (3) **Effekte:** Becken/scharf (2) → Trommel/relief-haft (3)

Tabelle 1: Cluster 21–23, Partitur-Parameter-Analyse (Start → Ende per Cluster)

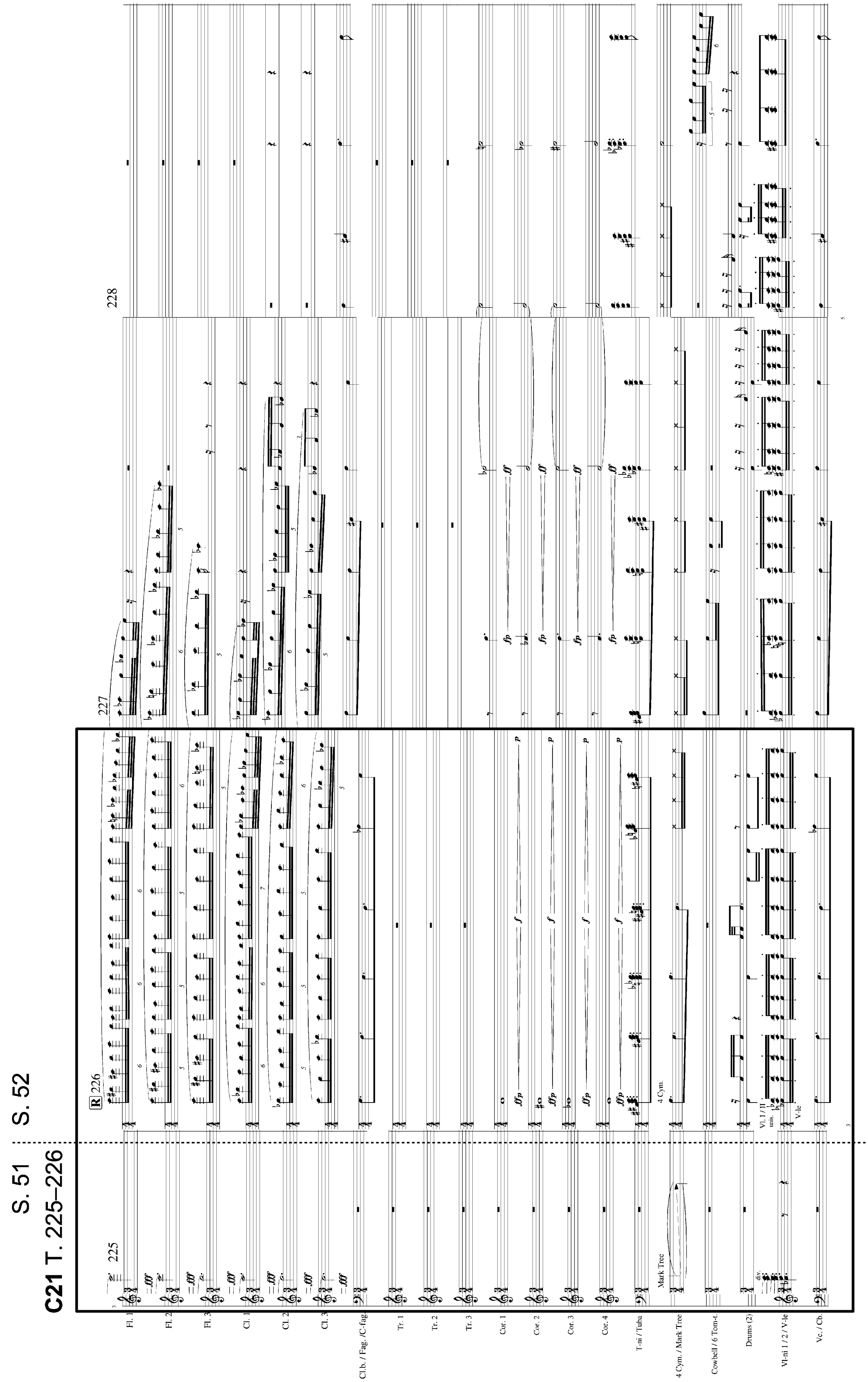


Abbildung 5: Cluster(-Paare) 21 und 22–23 (Reduktion der Partitur) (Fortsetzung auf der nächsten Seite)

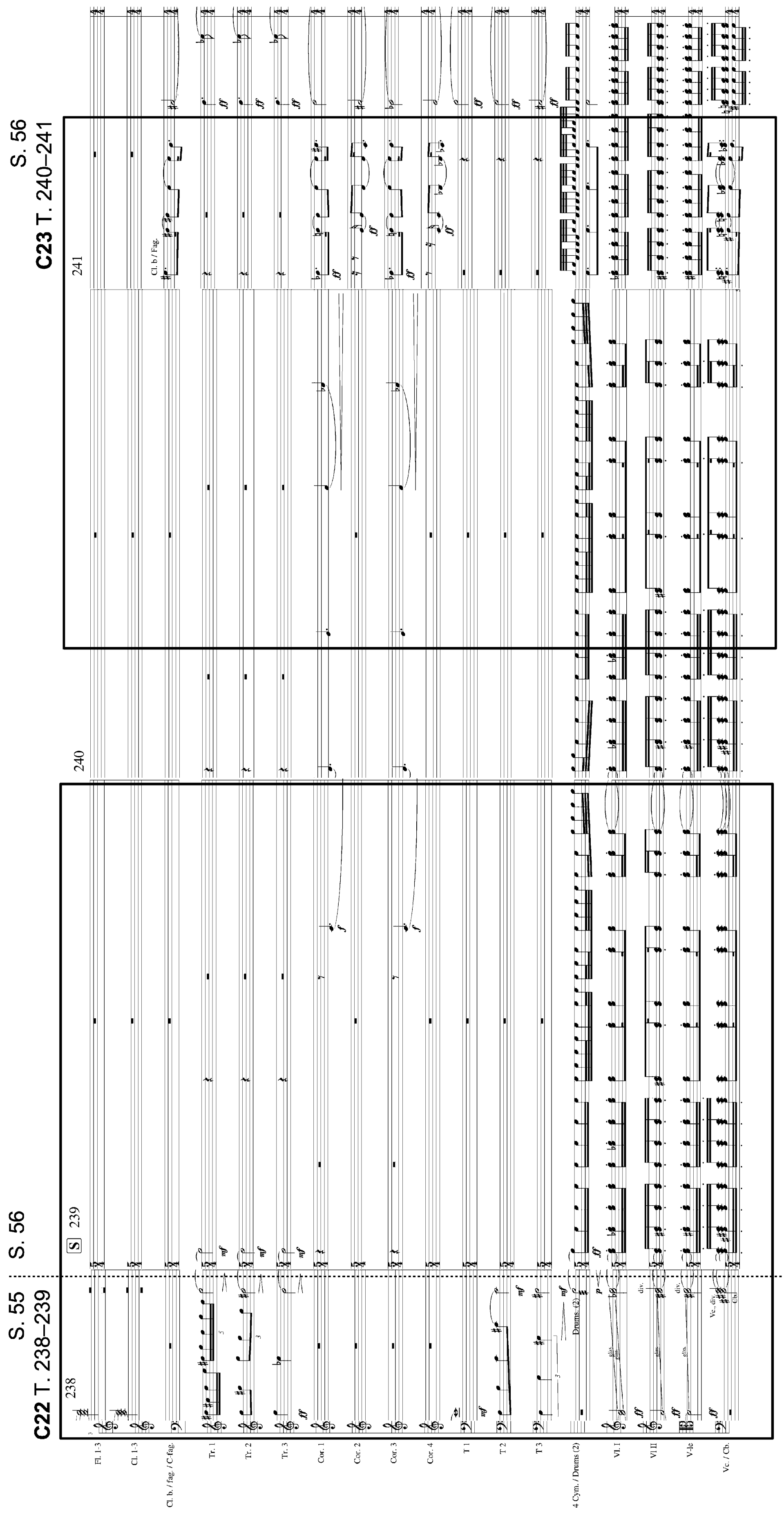


Abbildung 5 (Fortsetzung von vorangehender Seite)

Wie die Abbildungen 7c sowie 10c und d zeigen, geht die musikalisch zu erwartende Spannungsentwicklung des B-Teils für musiktheoretisch geschulte Hörende in Richtung der Kadenz (Cluster 27). Es wäre also ein musikalischer Spannungshöhepunkt für C26 oder C27 zu erwarten. Interessanterweise zeigt jedoch schon C21 den höchsten musikalischen Spannungswert bei den Expert*innen (zwischen 70 und 80 Punkten auf der arbiträren psychophysischen Skala [bei absolutem Maximum 100], in der Tendenz fallend) für die erste Hälfte (bis zur Kadenz) sowie den dritthöchsten für das gesamte Werk. Der Spannungswert bei den Musik- und Kunststudierenden (ca. 60) ist hingegen in C21 niedriger als im Cluster-Paar 17–18 (60–70), er zeigt steigende Tendenz und einen (in der Tendenz fallenden) Maximalwert erst im folgenden C22 (60–70). C23 zeigt die hohe gehaltene Spannung der Expert*innen (70), die Studierenden fallen hier aber deutlich ab (40–60). Mit den Clustern 24–26 ergibt sich der zu erwartende erneute Spannungsanstieg von niedrigerem Niveau aus (40–60), wobei die Musikstudierenden in C25 eine fallende Tendenz annehmen, während die anderen Teilnehmergruppen in der Tendenz unverändert bleiben. Bei allen drei Gruppen zeigt sich ein deutlicher Spannungsanstieg in C26 (50–60). Die Kadenz in C27 erregt nur bei den Expert*innen das zu erwartende Spannungsmaximum (65), bei den Musikstudierenden zeigt sich hier ein merklicher Rückgang (unter 30), bei den Studierenden der Bildenden Künste bewegen sich die Werte im mittleren Spannungsniveau (über 50). Entsprechend der Abbildung 6 zeigt die Impuls- und Zeitdauern-Statistik für die Cluster-(Paare) 19–26 einen Anstieg der Impulszahlen für C26 (19, darunter die der dichtesten Kumulation 9). Auch in C22 liegt dieser Wert mit 14 höher als bei den umliegenden Clustern, C21 weist nur 11 Impulse auf. Das Cluster-Paar 25–26 hat zusammen 30 Impulse. Die Dauer der Cluster liegt zwischen 3–6 Sekunden, die Cluster 19, 22 und 26 sind hierbei länger, die Cluster 20, 21, 23–25 kürzer.

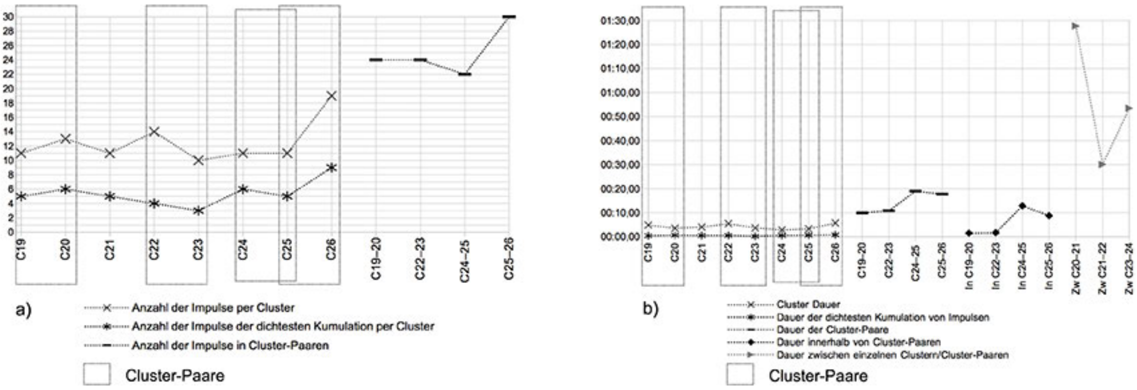


Abbildung 6: Cluster 19–26, a) Impuls-Statistik, b) Zeitdauern-Statistik

Abbildung 7 zeigt eine Gegenüberstellung der positiven Korrelation am Beispiel der Parameter-Daten für die Cluster 19–26 aus der Partitur mit COSM-Parameter-Tendenzen (steigend, fallend, unverändert) und Spannungsresultaten per Teilnehmenden-Gruppen (Expert*innen, Musik- und Kunststudierende). Der globale Vergleich der Spannungs-Gruppen mit den Parameter-Gruppen für die gesamte Dauer des Werkes (C01–48) findet sich in Abbildung 10f. In C19 und C22 (fallend), in C24 und C26 (steigend) weisen alle Teilnehmenden-Gruppen (E, M, K) dieselbe Tendenz innerhalb des jeweiligen Clusters auf. Mit Ausnahme von C25 zeigt sich bei Hör-Protokollen für alle Cluster bei den Musik- und Kunststudierenden eine einheitliche Tendenz, die sich allerdings in den Protokollen der Expert*innen so nicht zeigt.

Es fällt auf, dass darüber hinaus bei beiden Teilnehmenden-Gruppen die Parameter-Gruppen-Statistik identisch ist. In C25 entsteht sowohl bei den Expert*innen als auch den Kunststudierenden keine Korrelation zwischen Spannungsmittelwert, COSM und Partitur-Parameter-Gruppen.

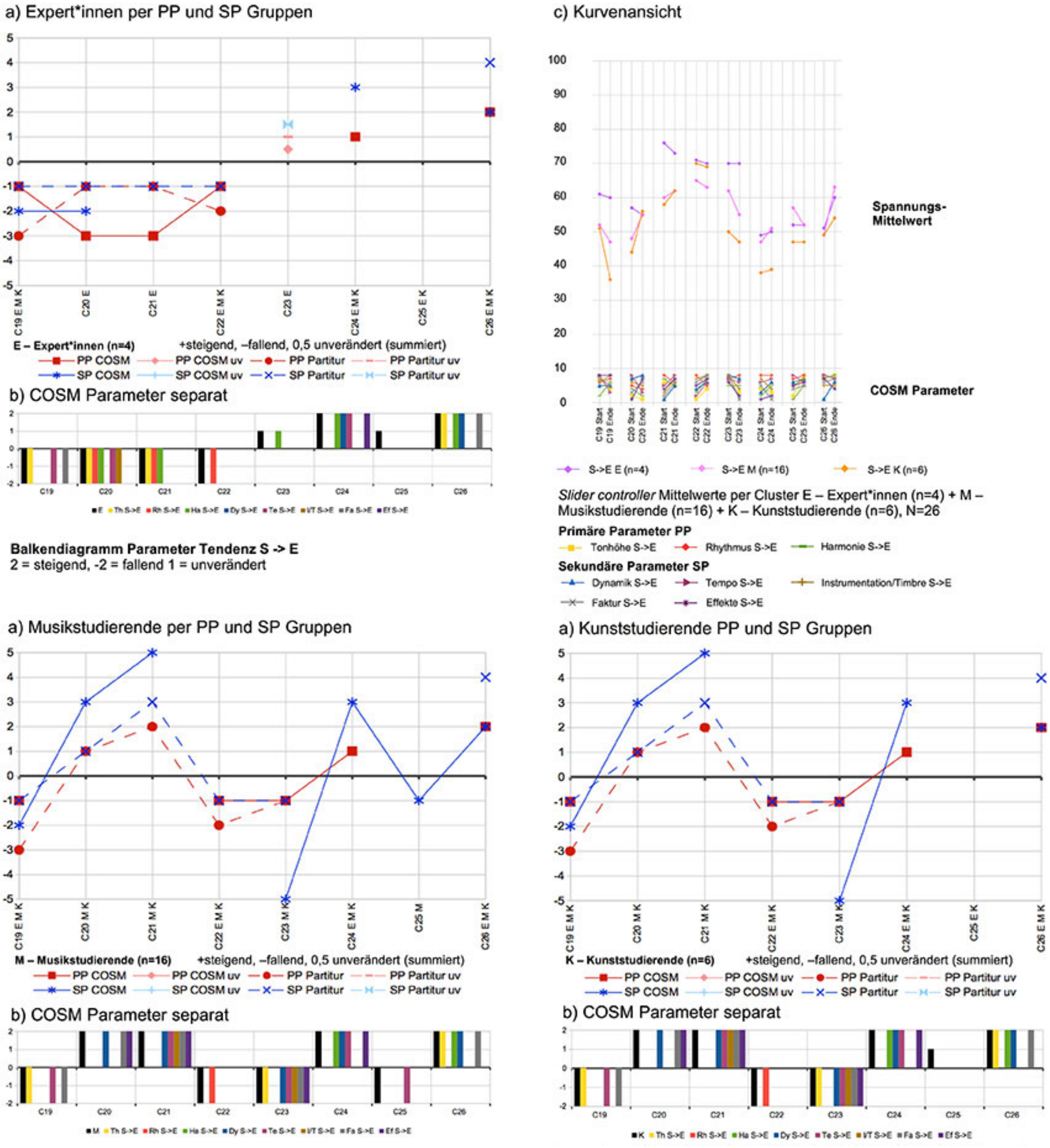


Abbildung 7: Cluster 19–26, Positive Korrelation a) per PP und SP Gruppe: COSM (n=14) vs. Spannungsmittelwert von E, M, K mit Parameter-Statistik der Partitur (Start → Ende + steigend, - fallend, uv unverändert); b) Parameter separat; c) Kurvenansicht.

Ferner fällt in Abbildung 7 auf, dass die Expert*innen grundsätzlich anders zu reagieren scheinen als die Musik- und Kunststudierenden. In den Clustern 19–22 der Expert*innen befinden sich viele Parameter (COSM und Partitur) im korrelierenden Abwärtstrend (keine sind steigend), bei den PP verlaufen COSM und Partitur-Daten spiegelbildlich – die PP sind abwechselnd salienter,[92] die SP verlaufen jedoch parallel, die COSM-Parameter sind in C19 und 20 salienter: in den Clustern C20 und 21 ist die Wahrnehmung (COSM) klarer als dies vermutlich der Fall gewesen wäre, wenn die Teilnehmenden des Experiments *prima vista* mit der Partitur konfrontiert worden wären. In C23 sind die COSM und Partitur-Parameter unverändert, in C24 und 26 sind die meisten Parameter in der Tendenz steigend sowie die SP salienter. Bei den Musik- und Kunststudierenden zeigt sich ein überwiegend paralleler Parameter-Verlauf zwischen PP und SP in COSM und der Partitur-Analyse. Wenn in C19 und C22 die COSM-PP dominieren, so sind es in C20 und 21 sowie C23 und 24 die COSM-SP. Im Cluster-Paar 22–23 fallen COSM und Partitur-Daten zusammen. Eine Schlussfolgerung daraus ist, dass für die Nicht-Expert*innen die SP (in C20, C21, C23) tatsächlich salienter erscheinen, im Gegensatz zu den PP, die für die Expert*innen salienter sind (in C20, C21). Allerdings zeigt sich bei allen Teilnehmer-Gruppen eine klare Abgrenzung um die Cluster 22 und 23 herum. Diese deutliche Regelmäßigkeit in den Daten visualisieren auch die Abbildungen 9 (Matrix) sowie die Abbildungen 10e und f für das gesamte Werk.

In Abbildung 8 werden die musikalischen Parameter-Ergebnisse von COSM der Cluster 21–23 nochmals detailliert abgebildet, nicht jedoch die zeitlichen Positionen der Impulse in jedem Cluster proportional zueinander. Um die Proportionalität der Impulse untereinander besser einordnen zu können, kann man die Abbildungen 8 und 4 zusammen betrachten. Es fällt auf, dass nicht für jeden Impuls alle Parameter angezeigt wurden. Die Salienzwert-Summen sind in Abbildung 8d dargestellt und zeigen, dass der Parameter Rhythmus in allen drei Clustern am häufigsten gewählt wurde, gefolgt von Tempo (C21) sowie von Instrumentation/Timbre (C22) und Instrumentation/Timbre sowie Faktur (C22). Weitere Parameter sind, besonders in C23, weniger salient. Darüber hinaus gestaltet sich die Verteilung der Parameter innerhalb der Cluster augenscheinlich variativ: es gibt keine temporal durchgehend salient-dominierenden Parameter, und bei genauer Betrachtung können sich lokale Muster für jeden Parameter ergeben.

---

[92] Salienter bedeutet hier und des Weiteren, dass auf Basis der Ergebnisse mehr Parameter per Cluster auf der vertikalen Skala der Grafiken als steigend oder fallend dargestellt werden können.

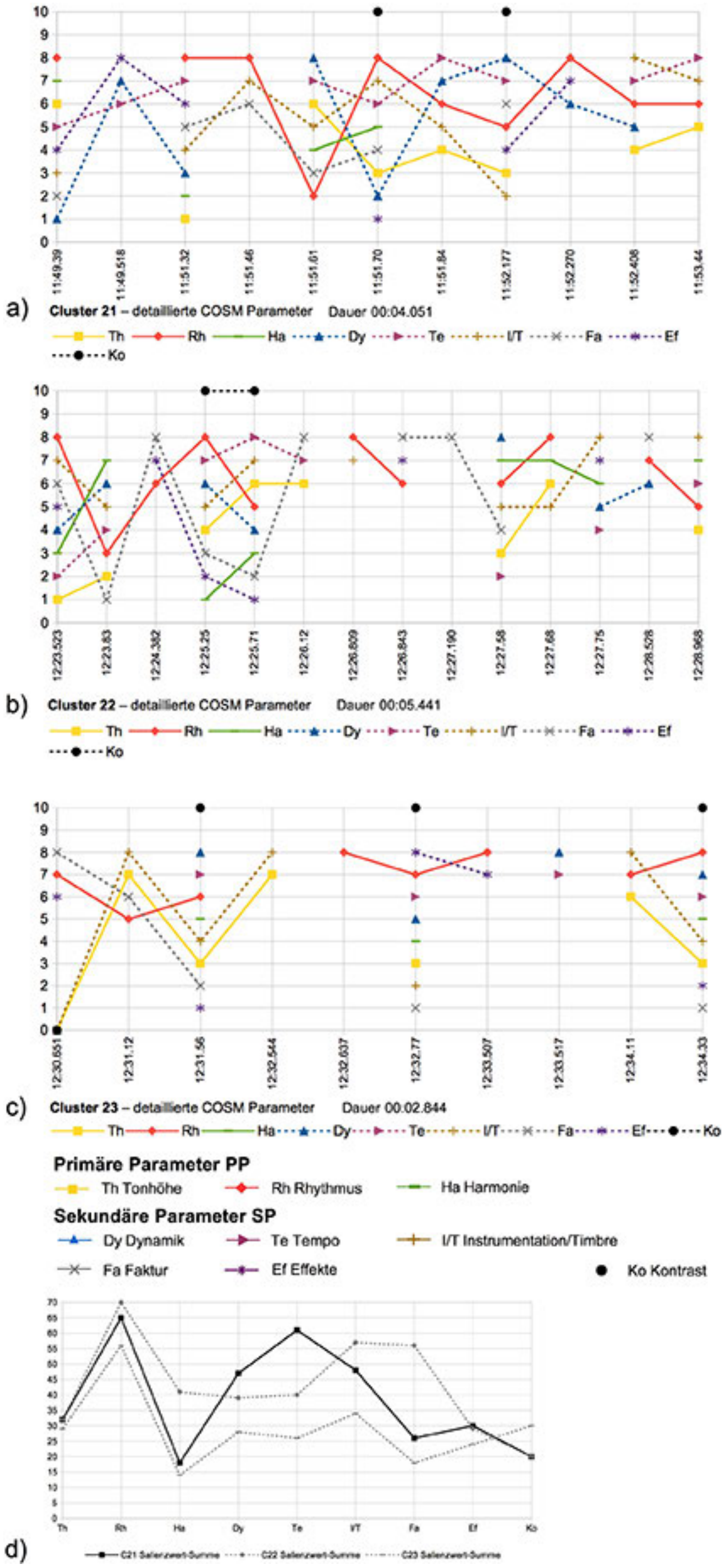


Abbildung 8: Cluster 21–23, (von oben a–c) detaillierte COSM-Parameter-Statistik, Salienzwert-Summenvergleich (d).

Höranalyse der Cluster 19–26

Verbale Beschreibung der musikalischen Ereignisse in C19–26

Cluster-Paar 19–21(Dauer 00:09.884)

C19 (Dauer 00:04.844, 10:11.806–10:16.65) beginnt mit schnellem Tremolo in den Holzbläsern, diese gehen in zwei aufeinanderfolgende schnelle Abwärtspassagen (Quasi-Glissando) über, beim zweiten Mal steigt das Quasi-Glissando erneut an. Das Schlagzeug verwendet den aus C19 kommenden Hi-Hat Schlag – bis zu Beginn von C20 wird der Schlagzeugbesen leise und als Effekt eingesetzt.

Mit C20 (Dauer 00:03.55, 10:18.14–10:21.69) beginnt das Solo-Schlagzeug fast ununterbrochen mit Rockmusik-Rhythmen (Floortom, Tomtom, Große und Kleine Trommel). Eine einzelne Unterbrechung dieser Faktur ist aber nicht von einem Cluster erfasst worden (einzelne Teilnehmende haben aber darauf reagiert). In C20 ertönen auch (erneut) dunkle Blechbläserakkordsäulen (chromatische Bewegung *d–cis*), die schon einige Zeit vorher für Spannung gesorgt haben.

Zu Beginn von C21 (Dauer 00:04.051, 11:49.39–11:53.44) endet die auf klangfarblich-dynamische (geräuschhafte) Kulminationseffekte ausgelegte Schlagzeugpartie (Ridebecken). In der zweiten Hälfte erklingen kontrastive Holzbläser in hohem Register (am Ende des Clusters ein Abwärts-Quasi-Glissando) sowie erneut die Rockmusik-Rhythmen des Solo-Schlagzeugs (Floortom, Tomtom, Große und Kleine Trommel).

Cluster-Paar 22–23 (Dauer 00:10.807)

In C22 (Dauer 00:05.441, 12:23.523–12:28.968) enden zunächst schnell (chromatisch) ansteigende und sich verlangsamende Blechbläserakkorde (in mittlerem bis hohem Register). In der zweiten Hälfte beginnt das Schlagzeug schnell zu pulsieren, wobei sich erneut die charakteristische punktierte Rhythmusfigur herauskristallisiert. In C23 (Dauer 00:03.679, 12:30.651–12:34.33) endet die zwischen beiden Clustern begonnene, modal ansteigende, diesmal aber sich beschleunigende Passage der Blechbläserakkorde (leise und mit leichtem *crescendo*). Diese endet direkt nach dem Cluster in einem hohen langen Ton. In der zweiten Hälfte kommen zudem schnelle Schläge der großen Trommel kontrastiv zum Einsatz, die erneut einen Spannungsanstieg hervorrufen.

Cluster-Paar 24–25 (Dauer 00:18.982)

In C24 (Dauer 00:02.844, 13:27.82–13:30.46) endet die bis zu diesem Zeitpunkt kulminierend verdichtete Steigerung des Kuhglockengeläuts. In der zweiten Hälfte begegnen abermals die Rockmusik-Rhythmen, die durch schnelles akzentuiertes Spiel auf der Großen Trommel eine Kontrastwirkung evozieren. Die Trommelschläge treten nicht zuletzt durch die Parallelführung mit den Holzbläsern im hohen Register (Motiv *a-h-gis-a-h-c*) klangfarblich deutlich hervor. Diese Schläge enden unmittelbar nach dem Cluster. Bis zu C25 übernehmen die Streicher das Motiv in hoher Lage und mit relativ geringer Lautstärke. Im Hintergrund von C24 erklingen lang gehaltene Basstöne sowie kurze schnelle Holzbläsermotive in hohem Register. Eine kulminierende Entwicklung vor C25, die jedoch weniger auf eine kontrastierende Wirkung hin angelegt ist, lässt sich nicht innerhalb eines Clusters erfassen. Allerdings beginnt C25 (Dauer 00:03.272, 13:43.53–13:46.802) mit dem klaren und verlangsamten Ende dieser kulminierenden Entwicklung durch das

Schlagzeug (im Hintergrund die schon aus C24 bekannten langen Basstöne), und es erklingen kurze schnelle Blechblasmotive (erst vier, dann sechs Schläge mit gleichbleibender Tonhöhe) zusammen mit solistischen lauten, punktierten Schlagzeug-Motiven, die zu einem erneuten Spannungsanstieg führen, zugleich jedoch eine Art Abschluss der bisherigen Spannungsentwicklung bedeuten.

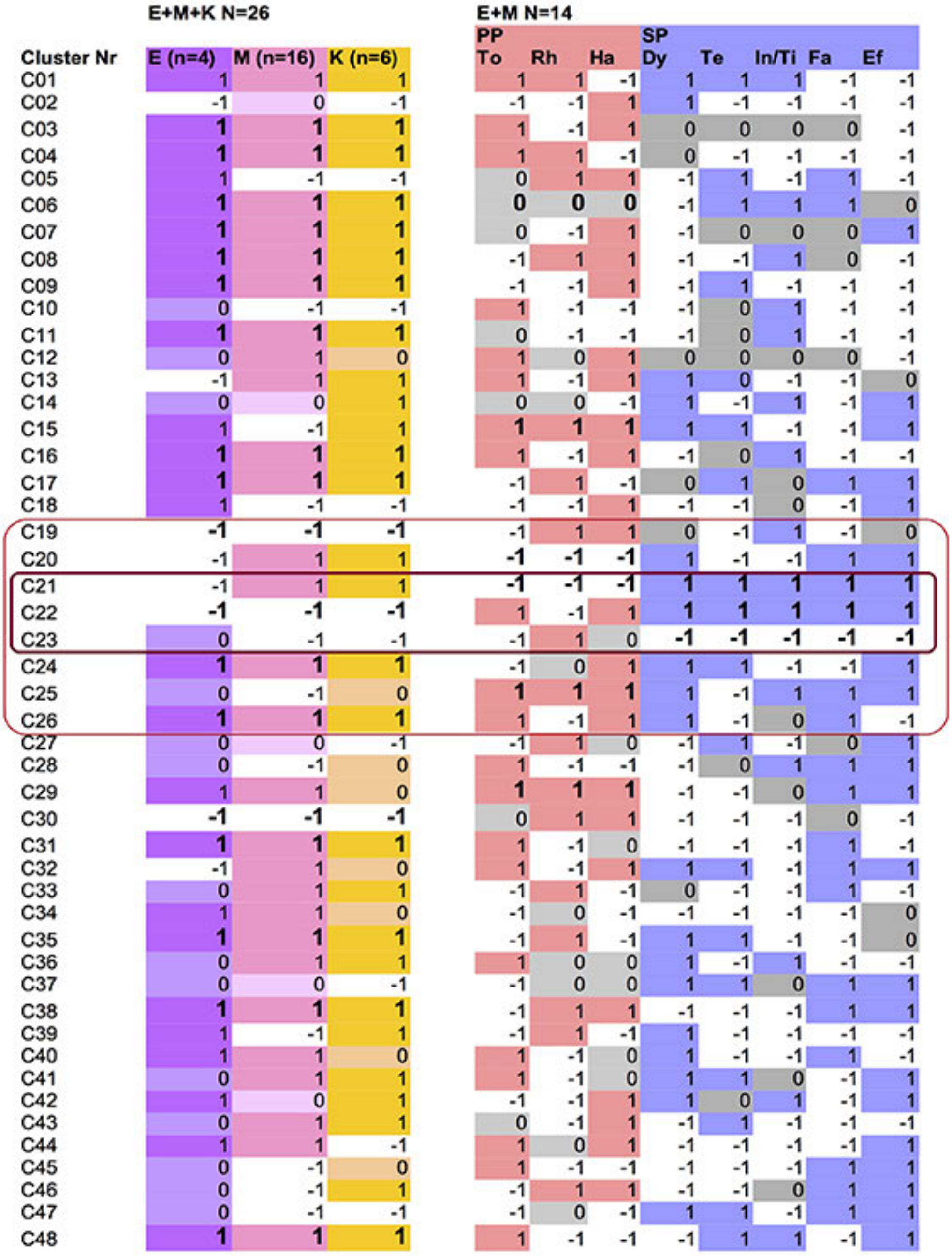
Cluster-Paar 25–26 (Dauer 00:17.751)

Im Cluster-Paar C25–26 ist der Endpunkt der bisherigen kulminierenden Entwicklung der Musik innerhalb des zweiten Drittels des gesamten Werkes – sogar noch vor Erreichen der zweiten Hälfte des zweiten Drittels – erreicht. Dieser Endpunkt mündet in die Solo-Kadenz des Schlagzeugs, die durch C27 erfasst ist. In C25 endet das ständige Solo-Schlagzeug. Bis zu C26 vollzieht sich eine Verlangsamung der Entwicklung. Zu Beginn von C26 (Dauer 00:05.713, 13:55.567–14:01.28) musizieren die Streicher *rallentando* in hohem Register (dritte Oktave: *cis-d-e-f-d-f-d*) Dann bilden die Blechbläser mit einem sukzessive aufwärtssteigenden übermäßigen Dreiklang (zweite Oktave: *a-cis-eis-a*) und dem darauf einsetzenden, die harmonische Spannung des *a* ins *b* auflösenden Tutti den vorläufigen Höhepunkt. Der Akkord, der mehr ein Cluster ist als ein Dreiklang, steht nicht zuletzt wegen seiner dunklen Klangfarbe in auffälligem Kontrast zu den Klängen der Streicher und Blechbläser.

In den Clustern C19–24 endet jeweils die dort kulminierende Spannungsentwicklung, und es baut sich eine neue auf. Diese Cluster sind zweigeteilt durch eine Zäsur in ihren Spannungsentwicklungen ungefähr auf der Hälfte ihrer Dauer. Es zeigt sich ein deutlicher Kontrast mit dem unmittelbar anschließenden Beginn einer neuen Phase mit zunehmender Spannung. C25 lässt als hybrider Cluster zwar das Auslaufen der bisherigen kulminierenden Entwicklung erklingen, die Spannung bleibt gleichwohl erhalten. In C26 ist die in ihrem Spannungsverlauf wieder ansteigende Abschlussphase des B-Teils zu hören, die in eine erneute, zwar langsame, aber nach wie vor spannungsgeladene Entwicklung mündet. Dies alles bereitet die Schlagzeug-Solokadenz vor, die sich exakt in der Mitte des 30:50 Min. dauernden Werkes ereignet, nämlich zwischen 14:26 und 16:17 Min.

Integrative Analyse

Aus der folgenden integrativen Analyse – einem Vergleich der Formanalyse mit den COSM-Parametern und den *slider controller-*/Spannungsmittelwerten – ergeben sich die Regelmäßigkeiten, die zuvor in der Lokal-Analyse der Cluster 19/21–23/26 schon im Detail zu erkennen waren. Die vertikal-chronologische Matrix der Abbildung 9 stellt die detaillierten Tendenzen für die Spannungsmittelwerte (nach Teilnehmenden-Gruppen) den Parameter-Gruppen (PP und SP) gegenüber und zeigt hingegen eine überwiegend unregelmäßige Struktur auf. Einzig in den Clustern 21–23 bzw. 19–26 lassen sich einige Regelmäßigkeiten erkennen, indem alle Spannungs-Teilnehmenden-Gruppen sowie sämtliche PP- und SP-Gruppen dieselben steigenden oder fallenden Tendenzen wahrnehmen.



	E+M+K N=26			E+M N=14							
				PP			SP				
Cluster Nr	E (n=4)	M (n=16)	K (n=6)	To	Rh	Ha	Dy	Te	In/Ti	Fa	Ef
C01	1	1	1	1	1	-1	1	1	1	-1	-1
C02	-1	0	-1	-1	-1	1	1	-1	-1	-1	-1
C03	1	1	1	1	-1	1	0	0	0	0	-1
C04	1	1	1	1	1	-1	0	-1	-1	-1	-1
C05	1	-1	-1	0	1	1	-1	1	-1	1	-1
C06	1	1	1	0	0	0	-1	1	1	1	0
C07	1	1	1	0	-1	1	-1	0	0	0	1
C08	1	1	1	-1	1	1	-1	-1	1	0	-1
C09	1	1	1	-1	-1	1	-1	1	-1	-1	-1
C10	0	-1	-1	1	-1	-1	-1	0	1	-1	-1
C11	1	1	1	0	-1	-1	-1	0	1	-1	-1
C12	0	1	0	1	0	1	0	0	0	0	-1
C13	-1	1	1	1	-1	1	1	0	-1	-1	0
C14	0	0	1	0	0	-1	1	-1	1	-1	1
C15	1	-1	1	1	1	1	1	1	-1	-1	1
C16	1	1	1	1	-1	1	-1	0	1	-1	-1
C17	1	1	1	-1	1	-1	0	1	0	1	1
C18	1	-1	-1	-1	-1	1	-1	-1	0	-1	1
C19	-1	-1	-1	-1	1	1	0	-1	1	-1	0
C20	-1	1	1	-1	-1	-1	1	-1	-1	1	1
C21	-1	1	1	-1	-1	-1	1	1	1	1	1
C22	-1	-1	-1	1	-1	1	1	1	1	1	1
C23	0	-1	-1	-1	1	0	-1	-1	-1	-1	-1
C24	1	1	1	-1	0	1	1	1	-1	-1	1
C25	0	-1	0	1	1	1	1	-1	1	1	1
C26	1	1	1	1	-1	1	1	-1	0	1	-1
C27	0	0	-1	-1	1	0	-1	1	-1	0	1
C28	0	-1	0	1	-1	-1	-1	0	1	1	1
C29	1	1	0	1	1	1	-1	-1	0	1	1
C30	-1	-1	-1	0	1	1	-1	-1	-1	0	-1
C31	1	1	1	1	-1	0	-1	-1	-1	1	-1
C32	-1	1	0	1	-1	1	1	1	-1	1	1
C33	0	1	1	-1	1	-1	0	-1	-1	1	-1
C34	1	1	0	-1	0	-1	-1	-1	-1	-1	0
C35	1	1	1	-1	1	-1	1	1	-1	-1	0
C36	0	1	1	1	0	0	1	-1	1	-1	-1
C37	0	0	-1	-1	0	0	1	1	0	1	1
C38	1	1	1	-1	1	1	-1	-1	-1	1	1
C39	1	-1	1	-1	1	-1	1	-1	-1	-1	-1
C40	1	1	0	1	-1	0	1	-1	-1	1	-1
C41	0	1	1	1	-1	0	1	1	0	-1	1
C42	1	0	1	-1	-1	1	1	0	1	-1	1
C43	0	1	1	0	-1	1	-1	1	-1	-1	-1
C44	1	1	-1	1	0	1	-1	-1	-1	-1	1
C45	0	-1	0	1	-1	-1	-1	-1	-1	1	1
C46	0	-1	1	-1	1	1	-1	-1	0	1	1
C47	0	-1	-1	-1	0	-1	1	1	-1	1	1
C48	1	1	1	1	-1	-1	-1	1	1	-1	1

PP = Primäre Parameter, SP = Sekundäre Parameter

steigend = 1, fallend = -1, unverändert = 0

***Slider controller*: Tendenzen der Mittelwerte per Cluster**
E – Expert*innen (n=4) + M – Musikstudierende (n=16)
+ K – Kunststudierende (n=6), N=26

COSM: Tendenzen innerhalb der Cluster per Parameter
E – Expert*innen (n=4) + M – Musikstudierende (n=10), N=14

Abbildung 9: Matrix: *Slider controller* und COSM-PP+SP, detaillierte Tendenz der Mittelwerte per COSM-Cluster Start (S) → Ende (E)

Im Folgenden werden die horizontal-chronologischen Grafiken der Abbildungen 10a–e beschrieben. In Abbildung 10a wird die musikalische Form einer Analyse unterzogen nach Kriterien der Sonatenform sowie der klassischen Rhetorik. In 10b werden die durch den DBSCAN generierten COSM-Daten-Cluster und -Cluster-Paare dargestellt, die sich gut in die Formanalyse einordnen lassen. Sie unterteilen den A-Teil in A1 und A2, was auch der Spannungsmittelwert mehr (10d, die exakten ›Start → Ende‹-Daten für jeden Cluster) oder weniger (10c, die Mittelwerte aus ›Start → Ende‹-Daten für jeden Cluster) bestätigt. 10e zeigt die höchste Formalisierung (positive Korrelation) der Daten von COSM- (Parameter-Tendenzen) und *slider controller*-Tendenzen.

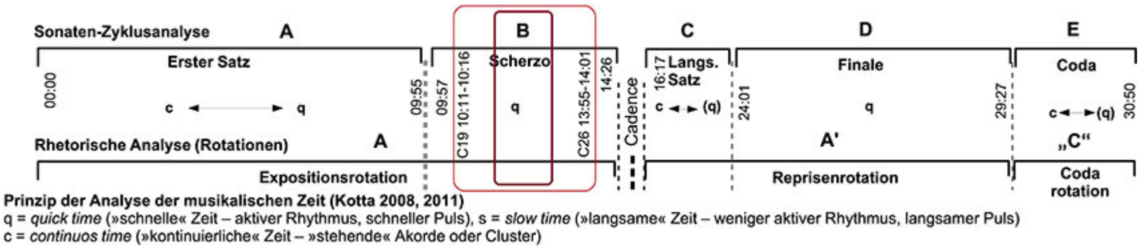


Abbildung 10a: Musikalische Formanalyse (Kotta 2011) nach Hepokoskis und Darcys (2006) rhetorischer Formanalyse

Abbildung 10a vergleicht zwei Formschemata der wellenförmig fließenden, sich ständig verändernden kulminativen und kontrastiven Musik – hier 4. Sinfonie / Schlagzeug-Konzert (2002) von Tüür –, die unter der Kategorisierung »langsame«, »schnelle« und »kontinuierliche« Zeit[93] analysiert worden ist – im ersten Satz geht die Entwicklung von der »kontinuierlichen« zur »schnellen« Zeit über, im Scherzo und im Finale überwiegt die »schnelle« Zeit, im langsamen Satz und in der Coda findet eine bedingt ähnliche Transformation wie im ersten Satz statt. Das musikalische Material wird rhetorisch rotierend entwickelt.[94] Klassisch klare thematische Anfangs- oder Endpunkte als formale Segmentationspunkte scheint es kaum zu geben. Dies hat den Autor dieses Artikels dazu veranlasst, kulminative Höhe- und Tiefpunkte als formbildende Segmentationspunkte zu bezeichnen.[95]

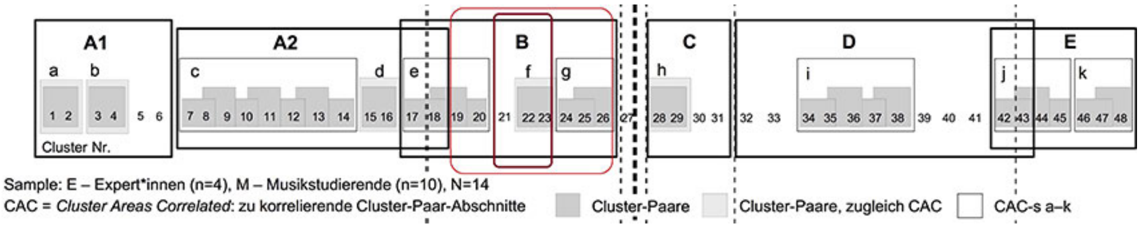


Abbildung 10b: DBSCAN, COSM-Cluster-Paare und CACs

Abbildung 10b zeigt, wie die durch den DBSCAN-Algorithmus ausgewählten 48 Cluster der COSM-Daten in Cluster-Paare und korrelierende Cluster-Paar-Abschnitte (CACs) eingeteilt werden können (A–E). Die Grundlage hierfür ist die zeitliche Nähe/Ferne der Solo-Cluster/Cluster-Paare zueinander. Innerhalb der Cluster-Paare liegt diese für C19–21 zwi-

---

[93] Vgl. Kotta 2008, 2011.

[94] Vgl. Kotta 2011, 104–108 und die dort präsentierten Modelle rhetorischer Formanalyse. Zum Prinzip der rhetorischen Rotation in der Sonatenform vgl. Hepokoski/Darcy 2006, 611.

[95] Vgl. Lock 2010a, 2010b.

schen 1 und 13 Sekunden: C19–20 (00:01.49) (siehe Abbildung 4), C22–23 (00:01.683), C24–25 (00:12.866), C25–26 (00:08.766). Zwischen den Cluster-Paaren/Solo-Clustern beträgt sie zwischen 30 Sekunden und 1.30 Minuten: C20 und 21 (01:27.699), C21 und C22 (00:30.087), C23 und C24 (00:53.419). Diese Daten lassen sich in Einklang bringen mit der Sonaten-Zyklusanalyse im Abbildung 10a, nur der erste Satz unterteilt sich nochmals in zwei Abschnitte (A1, A2), welche deutlich durch eine Zäsur in der Musik und durch die Spannungsergebnisse erklärbar sind (siehe Abbildung 10c–d).

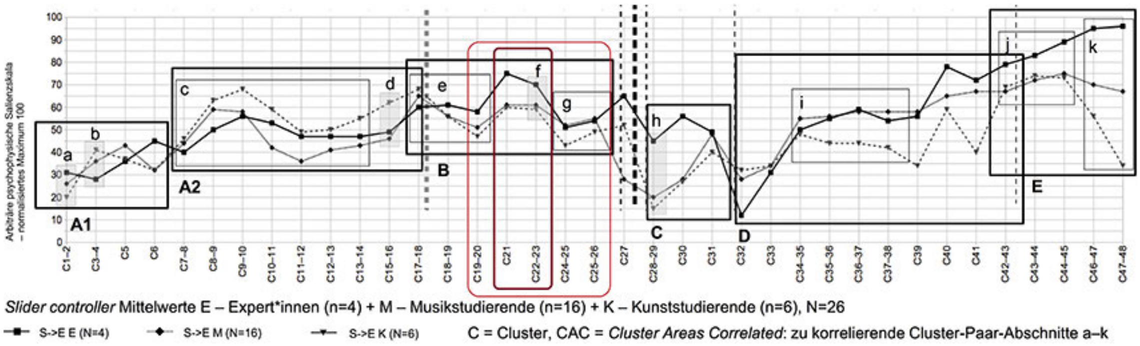


Abbildung 10c: *Slider controller*-(Spannungs-)Mittelwerte per COSM-Cluster-Paar und CACs, Start (S) → Ende (E)

Abbildung 10c zeigt, wie die Spannungsmittelwerte (Expert*innen, Musikstudierende und Kunststudierende jeweils separat) per COSM-Cluster-Paar und CACs sich ebenfalls in die Formanalyse/Cluster-Paar/CAC-Analyse einfügen (A–E). Während Daten der Expert*innen (E) in den Teilen B–E höhere Mittelwerte aufweisen, sind es in Teil A2 die der Nicht-Musiker*innen. In A1 zeigen die Gruppen alternierende Ergebnisse. Die Daten der Musikstudierenden (M) nähern sich abwechselnd mehr denen der Expert*innen oder denen der Kunststudierenden (K) an. Dies zeigt sich in Teil B: in den Clustern 19–23 liegen die Daten der M und K nahe beieinander, in den Clustern 24–26 die der M und E. Deutlich gehen die Mittelwerte in Teil C sowie in Teil D (ab i) auseinander, sie unterscheiden sich am klarsten in Teil E (k): die Daten der Expert*innen zeigen eine weiterhin hohe Spannung bis ins Maximum, die der Musikstudierenden fallen ins mittlere Niveau ab, die Daten der Kunststudierenden fallen deutlich ab.

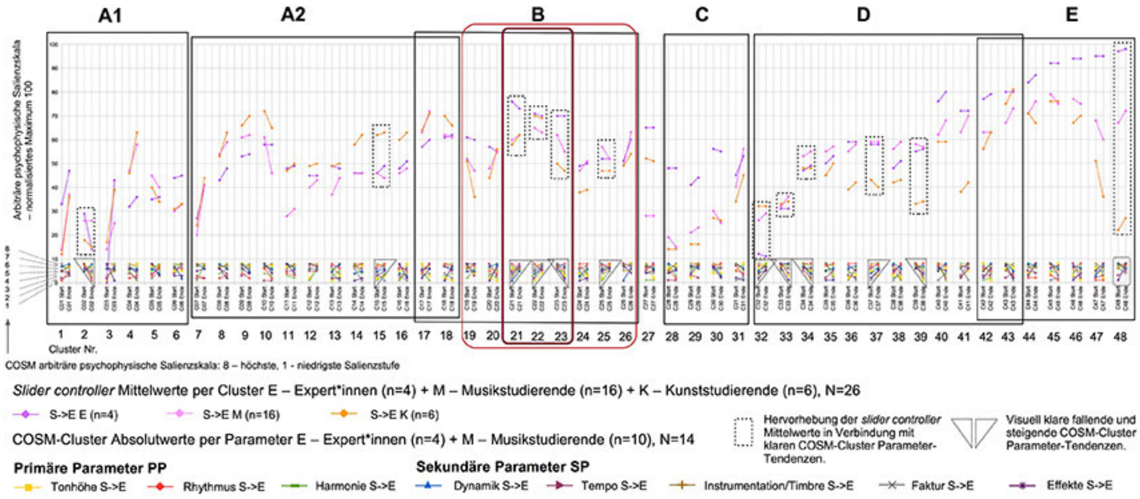


Abbildung 10d: *Slider controller*-(Spannungs-)Mittelwerte per COSM-Cluster und Absolutwerte per Parameter, Start (S) → Ende (E)

Abbildung 10d zeigt, wie sich die Spannungsmittelwerte detailliert per Cluster (Start → Ende) und Gruppen (Expert*innen, Musik- und Kunststudierenden separat) darstellen im Vergleich zu den Absolutwerten per Parameter. Mit gestrichelten Umrahmungen sind visuell auffallende ›S→E‹-Ergebnisse der Parameter hervorgehoben (steigend, fallend), die mit den *slider controller*-Werten verglichen werden können. Diese Daten sind in Abbildung 10e in ihrer Tendenz vergleichend summarisch generalisiert (per PP- und SP-Parameter-Gruppen).

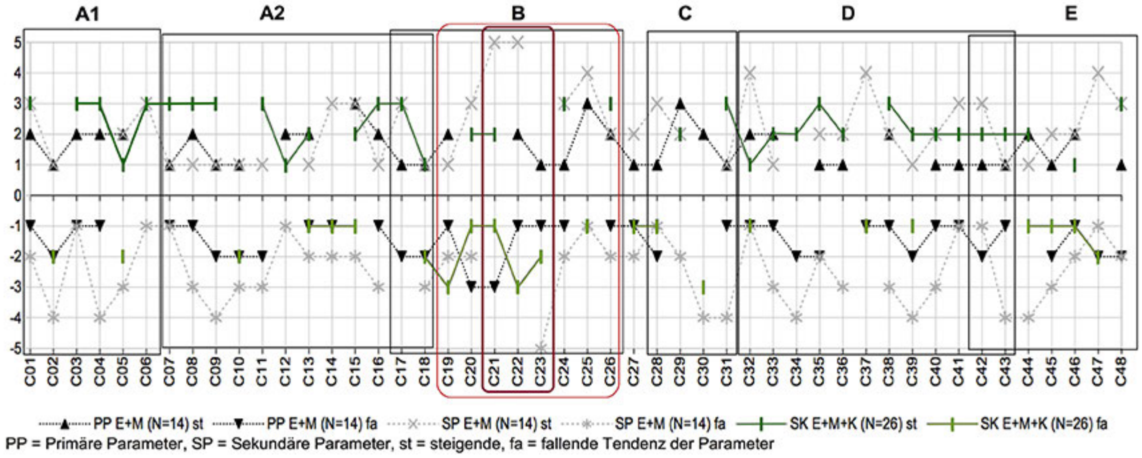


Abbildung 10e: *Slider controller*-(Spannungs-)Mittelwerte und COSM-PP+SP, summierte Tendenz der Mittelwerte per COSM-Cluster, Start (S) → Ende (E)

Abbildung 10f zeigt die Spannungs-, Parameter- und COSM-Mittelwert-Tendenzen per PP- und SP-Gruppen. Wenn eine oder mehrere Gruppen keine Tendenz zeigen, sind sie unverändert und in der Grafik nicht vertreten. Per Cluster wird auf der Vertikalachse neben der positiven (steigenden) und negativen (fallenden) Tendenz auch die Anzahl der beteiligten Parameter abgebildet. Es fällt auf, dass in Teil B bei Cluster 21–23 einmalig die Extreme erreicht werden. In C21 und C22 eignet allen SP steigende, in C23 dagegen fallende Tendenz. In C21 weisen ebenfalls sämtliche PP fallende Tendenz auf, in C22 jedoch nur noch zwei (Tonhöhe und Harmonie); in C23 zeigt lediglich ein PP (Tonhöhe) fallende und ein weiterer (Rhythmus) steigende Tendenz, während ein dritter (Harmonie) unverändert bleibt. In unerwarteter Weise korrelieren die Spannungs-Tendenzen: in C21 zeigen allein die Resultate zweier Gruppen (M und K) steigende, die der Expert*innen dagegen fallende Tendenz, ausgehend allerdings von einem Maximum. In C22 weisen alle, in C23 nur zwei Gruppen (M und K) fallende Tendenz auf, während sich die Ergebnisse bei den Expert*innen auf unverändert hohem Niveau bewegen. In den Clustern 19–20 sowie 21–22, 22–23 und 24–25 vollzieht sich eine gegenläufige Entwicklung zwischen Spannungs- und Parameter-Tendenzen innerhalb der Cluster – ein Phänomen, das sich auch in anderen Teilen findet. Daher kann nicht eindeutig von einer Korrelation der Tendenzen innerhalb der Cluster über das gesamte Werk hinweg gesprochen werden.

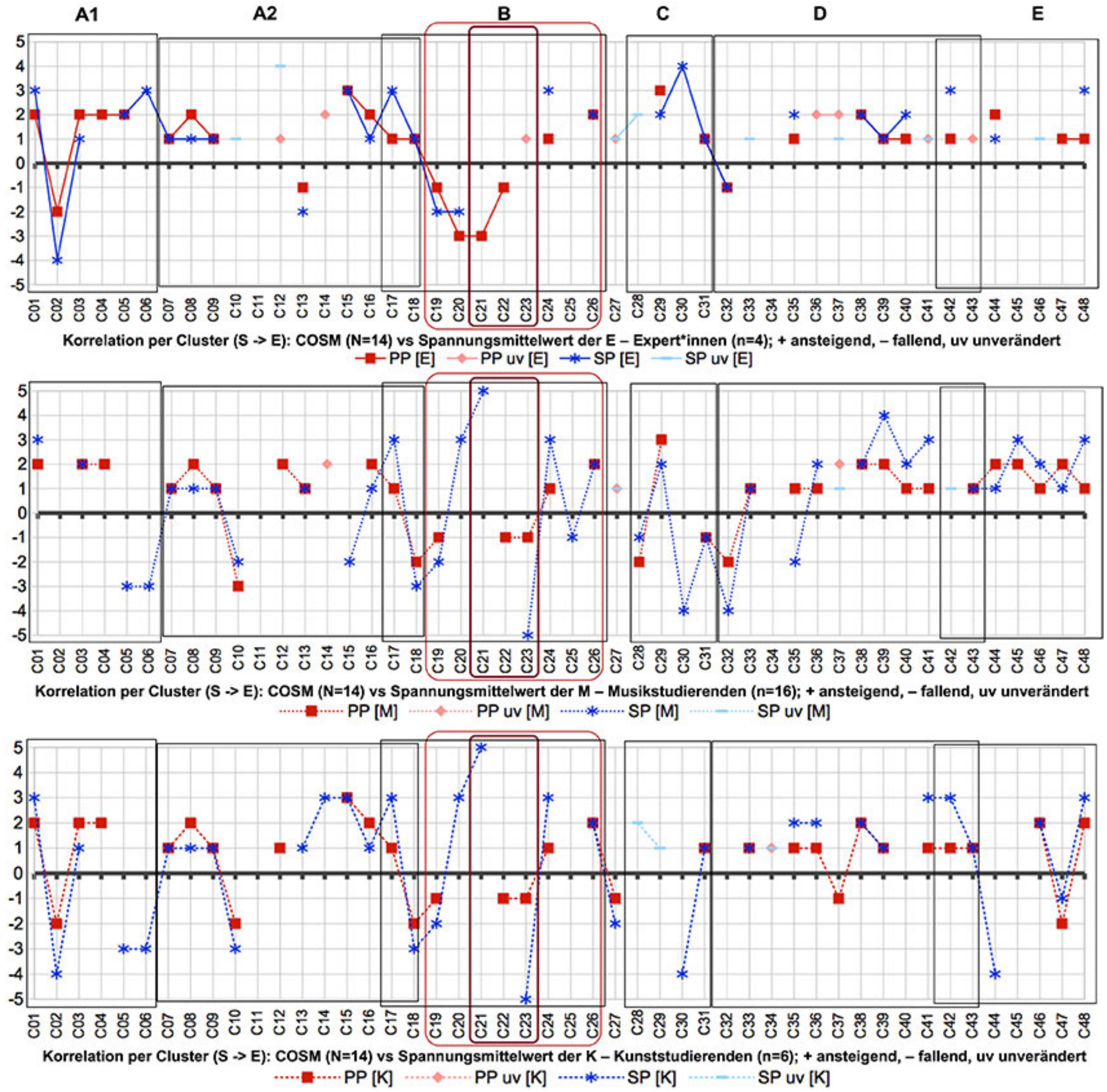


Abbildung 10f: *Slider controller*-(Spannungs-) und COSM-PP+SP-Tendenz der Mittelwerte per COSM-Cluster, Start (S) → Ende (E), per Teilnehmenden-Gruppen: E – Expert*innen, M – Musikstudierende, K – Kunststudierende

SCHLUSSFOLGERUNGEN UND DISKUSSION

Ausgehend von Abbildung 9 (Matrix) muss die erste eingangs gestellte Frage negativ beantwortet werden: SP sind bei post-tonaler Musik im Stil Tüürs nicht salienter als PP. Nur in den Clustern 21 und 22 verlaufen alle SP in steigenden Tendenzen, in C15, 20, 25 und 29 zeigen alle PP steigende Tendenzen. Einzig in C22 zeigen alle SP fallende Tendenzen. In allen anderen Clustern ergeben sich über das Werk hinweg kaum regelmäßige Muster. Die Hypothese 1 muss im Vergleich mit den *slider controller*-Spannungsmittelwerten negativ beantwortet werden, da die Veränderung der musikalischen Spannung nicht zugleich eine Veränderung bei den SP und umgekehrt nach sich zieht.

Ausgehend von den Abbildungen 7, 8, 9, 10e und f ergeben sich für die Beantwortung der zweiten Frage zur narrativen Rolle der musikalischen Parameter für die strukturelle Spannung von post-tonaler Musik erwartete und unerwartete Kombinationen von steigenden und fallenden Tendenzen sowohl innerhalb der Untergruppen der Teilnehmenden (Expert*innen, Musik- und Kunststudierende) als auch innerhalb von Untergruppen der musikalischen Parameter (PP, SP) sowie zwischen Spannungs-Mittelwert- und PP/SP-Tendenzen.

Die beiden möglichen Antworten auf die erste Frage sind:

a. Es ergibt sich ein systematisch reguläres Bild über das ganze Werk hinweg, welches es ermöglicht, die Rolle der PP/SP eindeutig zu den Spannungsmittelwerten in Bezug zu setzen, als Bestätigung der Hypothese 1.

b. Es ergibt sich kein systematisch reguläres Bild in den Darstellungen; unerwartete Zusammenhänge werden erkannt und interpretiert. Einzelne Cluster, Clusterpaare, korrelierende Abschnitte oder größere Formabschnitte können allerdings interessante Muster aufweisen, was hier auch der Fall ist. Hypothese 1 ist daher teilweise bestätigt für einzelne Cluster und Teilnehmenden-Gruppen – Musik- und Kunststudierende zeigen dies in den näher untersuchten Clustern.

Die Hypothese 2, nach der saliente Veränderungen im Fluss der Zeit durch Spannungsmittelwerte global und lokal sowie durch musikalische Parameter nur lokal angezeigt werden und deutliche Regelmäßigkeiten aufzeigen, kann, wie die Abbildungen 10c–f zeigen, für den ersten Teil (Spannungsmittelwerte) bestätigt werden. Für den zweiten Teil (musikalische Parameter im COSM-Experiment) bestätigt sie sich hingegen nicht, da im Zuge der Höranalyse keine deutlichen Regelmäßigkeiten erkannt wurden. Bei sämtlichen Abbildungen handelt es sich um generalisierende und mehr oder weniger detaillierte ›Hörgeschichten‹.

Ausgehend von Abbildung 9 und Abbildung 10e kann festgestellt werden, dass innerhalb der Cluster die Tendenzen der Spannungsmittelwerte in Teil A1, dem Beginn von A2 und Teil D eher steigend sind. Im weiteren Verlauf von A2, besonders in B und E, sind sie überwiegend fallend. Da in B und E die höchsten Spannungsmaxima-Mittelwerte erreicht werden (Abbildungen 10c und d), kann die fallende Tendenz damit erklärt werden, dass zum Zeitpunkt der Reaktion der Teilnehmenden (Anzeigen von Impulsen im COSM-Experiment) die Zeitverschiebung zwischen Ereignis und Reaktion sehr deutlich ist, weshalb der eigentliche Anstieg in den Clustern nicht erkennbar wird. Um dies zu überprüfen, wären die Spannungsmittelwerte um die Cluster herum separat zu untersuchen. Die fallende Tendenz könnte allerdings auch der Tatsache geschuldet sein, dass die im Experiment ermittelten Spannungsmittelwerte nur Start- und Endpunkte per Cluster berücksichtigen, nicht aber Maxima und Minima, die auch innerhalb der Cluster größere Extreme erreichen können. Letztere sind jedoch systematisch komplizierter zu erfassen. Schließlich könnte sie auch darauf zurückzuführen sein, dass viele Cluster (insbesondere die hier genauer betrachteten Cluster 19–26) zweiteilig sind: in der ersten Hälfte ein Spannungsabfall, dann ein kontrastives musikalisches Geschehen, das zu einem erneuten Spannungsanstieg führt. Hinzu kommt, dass sich der Spannungsverlauf in Cluster-Paaren oder in noch längeren Einheiten vollzieht, so dass die Tendenz nur einen Ausschnitt abbildet. Darüber hinaus ist zu betonen, dass es sich bei den Spannungsdaten nur um Mittelwerte handelt, die sich aus den individuellen Verläufen aller Teilnehmenden ergeben. Folglich wären auch individuelle Spannungsverläufe näher zu untersuchen.

Die Rolle der musikalischen Parameter bezüglich der musikalischen Spannungsmittelwerte in C21 und C22–23 kann anhand der zuvor gegebenen verbalen Beschreibung als durchaus korrelierend angesehen werden: in C21 sind die SP salienter durch klangfarblich-dynamische (geräuschhafte) Kulminations-Effekte des Schlagzeugs, und auch die Holzbläser erklingen kontrastiv, in hohem Register und mit Effekten wie dem Quasi-Glissando. Vermutlich drängen sie daher die PP (Tonhöhe, Rhythmus, Harmonie) in den Hintergrund. Überdies verlieren hier die Rockmusik-Rhythmen des Solo-Schlagzeugs durch Wiederholung an Salienz. Auch in C22 sind sämtliche SP mit steigender Tendenz vertreten. Erneut erklingen die sich verlangsamenden Blechbläserakkorde in mittlerem bis hohem Register, was vermutlich zu einer höheren Salienz der PP ›Tonhöhe‹ und ›Harmonie‹ führt. Hier scheint die charakteristische punktierte Rhythmusfigur, die sich aus der schnellen Pulsierung des Schlagzeugs herauskristallisiert, ebenfalls durch Wiederholung weniger salient zu sein. In C23 endet die sich beschleunigende Akkordpassage des Blechs (leise und mit leichtem *crescendo*), daher ist PP ›Harmonie‹ unverändert, während sich der PP ›Rhythmus‹ steigend und der PP ›Tonhöhe‹ fallend darstellen. Interessanterweise sind hier alle SP fallend, auch wenn in der zweiten Hälfte des Clusters schnelle Schläge der Großen Trommel kontrastierend einen vom Autor verbal beschriebenen erneuten Spannungsanstieg verursachen könnten, was sich aber nicht in den Spannungsmittelwerten bestätigt.

Bei empirisch-statistischen Mittelwert-Ergebnissen treten individuelle Hörstrategien in den Hintergrund. Es kommen nur besonders kumulierende Phänomene zum Tragen. Wenn im Extremfall die musikalische Struktur nur Neues bietet, also keine Variationszusammenhänge zum vorher Erklungenen erkennen lässt und die Frequenz von salienten Ereignissen zu dicht wird, hängt die Annahme eines Spannungsanstiegs bzw. -abfalls von den Voraussetzungen ab, die die Hörenden mitbringen: musikalisch erfahrene Personen finden vermutlich Metastrukturen, die ihrerseits Regelmäßigkeiten bieten können, wodurch sich die Interpretationsmöglichkeit als Narrativ erhöht. Musikalisch Unerfahrene hingegen geben an diesem Punkt vermutlich auf und zeigen einen Spannungsabfall an, so dass die Interpretationsmöglichkeit als Narrativ abnimmt. Bezüglich musikalischer Parameter kommt es hier vermutlich auf die Wiederholung (identisch oder variiert) von Salienz-Kombinationen im Vergleich zwischen den Clustern an, welche sich bei den Hörenden im Verlauf des Experiments als Strategie mehr oder weniger herausbilden können. Wenn Hörer*innen eines Teilnehmenden-Samples sich nur auf bestimmte Parameter konzentrieren, fallen die anderen Parameter aus den Ergebnissen heraus, was jedoch nicht bedeutet, dass die anderen Parameter nicht wirksam gewesen wären (siehe Theorieaspekte des COSM-Modells). Die Ergebnisse dieser Studie werden daher als Grad der Narrativität interpretiert, nicht als direkte (bzw. naturwissenschaftlich nachweisbare) Spannungs-Parameter-Korrelationen oder gar Ursache-Wirkungs-Phänomene.

ZUSAMMENFASSUNG

Resümierend lässt sich sagen, dass die Spannungsmittelwerte der unterschiedlichen Gruppen (Expert*innen, Musik- und Kunststudierende: E, M, K) in der globalen Tendenz ähnlich verlaufen. Hinsichtlich der verschiedenen Formabschnitte des Werkes ergeben sich jedoch auf der relativen vertikalen psychophysischen Intensitätsskala wechselnde signifikante Unterschiede, insbesondere gegen Ende des Werkes. Offenkundig sind Expert*innen sowie Musikstudierende mehr daran gewöhnt, Musik bis zum Ende mit Spannung zu verfolgen, als die meisten Kunststudierenden. Bezüglich der Parameter-Gruppierungen (PP, SP) sowie einer möglichen Rolle der musikalischen Parameter für die

Wahrnehmung der musikalischen Spannung ergeben sich bis auf die Cluster 21–23 vor der Solokadenz in der Mitte des Werkes keine klaren Muster. Am deutlichsten stellt sich die Situation der musikalischen Parameter– PP fallend, SP steigend – im höchsten Spannungsmaximum (C21) vor der Solokadenz des Schlagzeugs dar. Das darauffolgende Cluster-Paar 22–23 ist bezüglich der SP kontrastiv: in C22 sind sämtliche SP steigend (allerdings negativ korrellierend mit der fallenden Tendenz der Spannungsmittelwerte aller Teilnehmenden-Gruppen), in C23 fallend (diesmal korrelierend mit der Tendenz der Spannungsmittelwerte von zwei Teilnehmenden-Gruppen, siehe Abbildung 10e). Es ergeben sich in anderen Clustern keine klaren positiven Korrelationen zwischen Spannungsmittelwert- und musikalischen Parameter-Tendenzen. Dies spricht für die tatsächliche Variativität von Tüürs Musik und zeigt, dass ihr gegenüber unterschiedliche Hörstrategien möglich sind, welche im Rahmen dieser Untersuchung durch Mittelwerte eher künstlich im Sinne einer vermeintlich einheitlichen Tendenz zusammengezwungen wurden.

Nach Wilfried Gruhn[96] besteht eine Verbindung zwischen dem Rezeptionsvorgang – zwischen Wahrnehmen als physiologischer Perzeption und Verstehen als Kognition im Sinne eines bedeutungsgebenden, interpretatorischen Akts des Denkens, Verarbeitens, (Wieder-)Erkennens, Erinnerns und Verstehens – sowie konnektionistischen Modellen (selbstorganisierender dynamischer neuronaler Netze) mit Hilfe narrativer Methoden (Erfinden einer Hörgeschichte) beim Finden von Zugängen zu musikalischen Hörerlebnissen. Gruhns Verbalisierung der Erkenntnisakte bei der Perzeption im Sinne von ›*von* etwas‹ und der Kognition im Sinne von ›*als* etwas‹ beinhaltet eine narrative Dimension. Es wird hierbei erkannt, dass etwas *ist* (als Ereignis, Impuls), dieses wird benannt als etwas (basierend auf schon vorhandenen Mustern und Erfahrungen). Mit COSM kann die Aufeinanderfolge solcher Ereignisse formalisiert und grafisch dargestellt werden.

Der holistische Ansatz der neuen kognitiv-empirischen Analyse-Methode COSM befindet sich in der ersten Phase, wenn die Musik ohne Unterbrechung bis zu Ende analysiert wird (Form als Ganzheit). Hierbei wird von der Ganzheitlichkeit des Systems Musik (Klang) bzw. der ›Musik als Umwelt‹ (System/Umwelt als Ganzheit) ausgegangen. In der zweiten Phase führt das vermutete ständige Zusammenwirken aller (hierbei ausgewählten acht) Parameter in jedem Moment (musikalisches Material als Ganzheit) zu einer holistischen Perzeption.

Da jede Grafik in der Wissenschaft auch verbal erklärt, erzählt und begründet wird und signifikante bzw. ausgewählte Elemente enthält, gelangt man unweigerlich auf die narrative Ebene. Auch wenn man bei musikalischer Spannung von Kulminationsdramaturgie spricht,[97] unterstützt dies die Analogie zur Narrativität im Sinne eines Plots, der sich in der musikalischen Form widerspiegelt. Wenn, wie in diesem Artikel, die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung post-tonaler Musik untersucht wird, ist dies *qua* Terminus (interpretiert als Rolle) eine narrative Herangehensweise, die auf den Erwartungen der Hörenden (Hörstrategien) und Analysierenden (Hypothesen) z. B. bezüglich der primären und sekundären Parameter beruht.

Auch wenn COSM hauptsächlich nonverbal und über audiovisuelle Analogien funktioniert, wird es durch die Fokussierung auf ausgewählte musikalische Parameter im zweiten Schritt ebenfalls auf die verbale Ebene gehoben. Die Teilnehmenden des Experiments sollten eine Vorstellung von den musikalischen Parametern besitzen, die sie auf

---

[96] Gruhn 1992, 44 und 49.

[97] U. a. Kirschbaum 2001.

die wahrzunehmende Musik im Rahmen eines Wiedererkennungsprozesses anwenden können. Ein solcher Prozess führt seinerseits zu einer den selbstorganisierenden Lernprozess einschließenden kognitiven Strategie. Die musikalischen Parameter wurden im Verlauf des COSM-Experiments quantifiziert (in Daten umgewandelt). Danach musste das grafisch darstellbare Ergebnis verbalisiert werden, und es entstand eine entweder individuelle oder eine empirisch verallgemeinerbare Hörgeschichte einer Anzahl von Teilnehmenden. Somit kann das COSM als eine formalisierte und empirische Realisierung der Gruhn'schen narrativen Methode zum Verstehen des Hörerlebnisses verstanden werden.

Literatur

Agawu, Victor Kofi (1982), *The Structural Highpoint as Determinant of Form in Nineteenth Century Music*, Stanford: Stanford University Press.

Abbott, Alison (2002), »Music, Maestro, Please!«, *Nature*, Heft 416, 12–14.

Addessi, Anna Rita / Roberto Caterina (2000), »Perceptual Musical Analysis: Segmentation and Perception of Tension«, *Musicae Scientiae* 4/1, 31–54.

Addessi, Anna Rita (2010), »Auditive Analysis of the Quartetto per Archi in Due Tempi (1955) by Bruno Maderna«, *Understanding Musical Structure and Form: Papers in Honour of Irène Deliège*, 225–249.

Albrecht, Joshua (2012), *Affective Analysis of Music Using the Progressive Exposure Method: The Influence of Bottom-Up Features on Perceived Musical Affect*, Ph.D., The Ohio State University Columbus.

Almén, Byron (2003), »Narrative Archetypes: A Critique, Theory, and Method of Narrative Analysis«, *Journal of Music Theory 47/1*, 1–39.

Almén, Byron (2006), »Modes of Analysis«, *Theory and Practice. Journal of the Music Theory Society of New York State 31*, 1–38.

Bauer, Amy (2001), »›Composing the Sound Itself‹: Secondary Parameters and Structure in the Music of Ligeti«, *Indiana Theory Review 22/1*, 37–64.

Bor, Mustafa (2009), *Contour Reduction Algorithms: A Theory of Pitch and Duration Hierarchies for Post-Tonal Music*, Ph.D., The University of British Columbia Vancouver.

Burrows, David (1997), »A Dynamical Systems Perspective on Music«, *The Journal of Musicology 15/4*, 529–545.

Chouvel, Jean-Marc (2014), »Categories and Representation in Musical Analysis«, *Sonus 35/1*, 17–35.

Clifton, Christopher (2019), »Data Mining«, in: *Encyclopaedia Britannica*. https://www.britannica.com/technology/data-mining (11.12.2020)

Conklin, Darell (2006), »Melodic Analysis with Segment Classes«, *Machine Learning 65/2–3*, 349–360.

Dahlhaus, Carl (2004), »Wagners dramatisch-musikalischer Formbegriff«, in: ders., *Gesammelte Schriften*, hg. von Hermann Danuser in Verbindung mit Hans-Joachim Hinrichsen und Tobias Plebuch, Bd. 7, Laaber: Laaber, 322–332.

Deliège, Irène / Marc Mélen (1997), »Cue Abstraction in the Representation of Musical Form«, in: *Perception and Cognition of Music*, hg. von Irène Deliège und John Sloboda, New York: Psychology Press, 387–412.

Deliège, Irène / Marc Mélen / Diana Stammers / Ian Cross (1996), »Musical Schemata in Real-Time Listening to a Piece of Music«, *Music Perception 14/2*, 117–159.

Eitan, Zohar / Roni Y. Granot (2009), »Primary Versus Secondary Musical Parameters and the Classification of Melodic Motives«, *Musicae Scientiae Discussion Forum 48*, 139–179.

Ester, Martin / Hans-Peter Kriegel / Jörg Sander / Xiaowei Xu (1996), »A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise«, *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, Palo Alto: AAAI Press, 226–231.

Farbood, Morwaread M. (2006), *A Quantitative, Parametric Model of Musical Tension*, Ph.D., Cambridge (MA): MIT Press.

Farbood, Morwaread M. (2012), »A Parametric, Temporal Model of Musical Tension«, *Music Perception 29/4*, 387–428. https://doi.org/10.1525/mp.2012.29.4.387 (11.12.2020)

Fu, Tak-chung (2011), »A Review on Time Series Data Mining«, *Engineering Applications of Artificial Intelligence 24*, 164–181.

Gabrielson, Alf (2002), »Emotion Perceived and Emotion Felt: Same or Different?«, *Musicae Scientiae 5/1*, 123–147.

Granot, Roni Y. / Zohar Eitan (2011), »Musical Tension and the Interaction of Dynamic Auditory Parameters«, *Music Perception 28/3*, 219–245. https://doi.org/10.1525/mp.2011.28.3.219 (11.12.2020)

Gruhn, Wilfried (1992), »Wahrnehmen und Verstehen. Kognitive Grundlagen der Repräsentation musikalischer Elemente und Strukturen. Entwurf zu einem Forschungsprogramm«, in: *Musikalische Erfahrung. Wahrnehmen, Erkennen, Aneignen*, hg. von Hermann J. Kaiser, Essen: Die Blaue Eule.

Guigue, Didier / Charles de Paiva Santana (2018), »The Structural Function of Musical Texture: Towards a Computer-Assisted Analysis of Orchestration«, *Journées d'Informatique Musicale (JIM 2018)*, 97–103. https://hal.archives-ouvertes.fr/hal-01791398 (11.12.2020)

Hadrava, Michal / Jaroslav Hlinka (2020), »A Dynamical Systems Approach to Spectral Music: Modeling the Role of Roughness and Inharmonicity in Perception of Musical Tension«, in: *Frontiers in Applied Mathematics and Statistics 6*, 1–14, https://doi.org/10.3389/fams.2020.00018 (20.12.2020)

Herman, David (2009), »Cognitive Approaches to Narrative Analysis«, in: *Cognitive Poetics: Goals, Gains and Gaps*, hg. von Geert Brône und Jeroen Vandaele, Berlin: Mouton de Gruyter, 79–118.

Hepokoski, James / Warren Darcy (2006), *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth-Century Sonata*, Oxford: Oxford University Press.

Huron, David (2006), *Sweet Anticipation. Music and the Psychology of Expectation*, Cambridge (MA): MIT Press.

Jensen, Arthur (1979), »›g‹: Outmoded Theory of Unconquered Frontier«, *Creative Science and Technology 2*, 16–29.

Jensenius, Alexander Refsum (2002), *How Do We Recognize a Song in One Second. The Importance of Salience and Sound in Music Perception*, Master of Arts, Music Technology Department of Music and Theatre University of Oslo.

Juslin, Patrik N. / John A. Sloboda (2010), *Handbook of Music and Emotion: Theory, Research, Applications*, Oxford: Oxford University Press.

Kallinen, Kari (2006), *Towards a Comprehensive Theory of Musical Emotions. A Multidimensional Research Approach and Some Empirical Findings*, Ph.D, University of Jyväskylä, Jyväskylä: University of Jyväskylä.

Kallinen, Kari / Niklas Ravaja (2006), »Emotion Perceived and Emotion Felt: Same and Different«, *Musicae Scientiae 10*, 191–213.

Kersten, Luke (2014), »Music and Cognitive Extension«, *Empirical Musicology Review 9*, 193–202.

Kirschbaum, Martin (2001), *Höhepunktbildung und Dramaturgie in Neuer Musik. Die Erstellung »formaler Spannungsverläufe « als Ergänzung der musikalischen Analyse*, Köln: Dohr.

Koelsch, Stefan (2013), *Brain and Music*, Oxford: Willey-Blackwill.

Koelsch, Stefan / Walter A. Siebel (2005), »Towards a Neural Basis of Music Perception«, *Trends in Cognitive Sciences 9*, 578–584. https://doi.org/10.1016/j.tics.2005.10.001 (11.12.2020)

Kotta, Kerri (2008), »Zur musikalischen Zeit: Über die formalen Grundlagen der Werke von Erkki-Sven Tüür – ein Beschreibungsversuch«, in: *Schwarze Milch und bunte Steine. Der Komponist Erkki-Sven Tüür*, hg. von Hans-Klaus Jungheinrich, Mainz: Schott, 23–38.

Kotta, Kerri (2011), »Classical Formal Archetypes in Selected Works by Erkki-Sven Tüür«, *Lietuvos muzikologija [Lituanian musicology] 12*, 102–110.

Känd, Elisa Johanna (2016), *Narratiivsuse ja pinge taju valitud »Valtari Mystery Film Experiment« audiovisuaalteostes (Perception of Narrativity and Tension in Chosen Audiovisual Works of the »Valtari Mystery Film Experiment«)*, Audiovisual Media Bachelor, Tallinn University.

Lehne, Moritz / Stefan Koelsch (2015), »Towards a General Psychological Model of Tension and Suspense«, *Frontiers in Psychology* 6, 1–11. https://doi.org/10.3389/fpsyg.2015.00079 (11.12.2020)

Lerdahl, Fred (1989), »Atonal Prologational Structure«, *Contemporary Music Review 4*, 65–87.

Lerdahl, Fred (2001), *Tonal Pitch Space*, Oxford: Oxford University Press.

Lerdahl, Fred / Carol Krumhansl (2007), »Modelling Tonal Tension«, *Music Perception 24/4*, 329–366.

Lerdahl, Fred / Ray Jackendoff (1983), *A Generative Theory of Tonal Music*, Cambridge (MA): MIT Press.

Libet, Benjamin (2004), *Mind Time. The Temporal Factor in Consciousness*, Cambridge (MA): Harvard University Press.

Libet, Benjamin / Elwood W. Wright Jr. / Bertram Feinstein / Dennis K. Pearl (1979), »Subjective Referral of the Timing for a Conscious Sensory Experience – a Functional Role for the Somatosensory Specific Projection System in Man«, *Brain 102*, 193–224. https://doi.org/10.1093/brain/102.1.193 (11.12.2020)

Ligeti, György (1966), »Über Form in der Neuen Musik«, in: *Form in der Neuen Musik*, hg. von Ernst Thomas, Darmstadt: Schott's Söhne, 23–35.

Lock, Gerhard (2010a), »Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses ›Oxymoron‹« [On Factors Influencing Musical Tension and its Perception in Erkki-Sven Tüür's piece »Oxymoron«], *Res musica 2*, 63–74.

Lock, Gerhard (2010b), »Tension Design Analysis as Listening Analysis Strategy for Contemporary Music: Perception of Tension in Contemporary Post-Tonal Orchestral Music: a Case Study«, Poster-Präsentation bei der 11th International Conference on Music Perception and Cognition (ICMPC11), 23.–27. August 2011, Seattle (WA). http://depts.washington.edu/icmpc11/ (11.12.2020)

Lock, Gerhard (2017), »Models and Scientific Modelling in Music Context«, *Problems in Music Pedagogy 16/2*, 25–45. http://pmp.du.lv/wp-content/uploads/2016/08/PMP_2017_Vol-16-2.pdf (11.12.2020)

Lock, Gerhard / Kerri Kotta (2012), »Musical Tension as a Response to Musical Form«, in: *Proceedings of the »12th International Conference on Music Perception and Cognition« (ICMPC) / »8th Triennial Conference of the European Society for the Cognitive Sciences of Music« (ESCOM)*, Thessaloniki: Aristotle University of Thessaloniki, 612–617. http://icmpc-escom2012.web.auth.gr/ (11.12.2020)

Lock, Gerhard / Kerri Kotta (2015), »A Preliminary Cognitive Model of Musical Tension in Contemporary Orchestral Music«, Paper-Präsentation bei der 12th international Music Analysis and Theory Conference, 24.–27. September 2015, Istituto Superiore di Studi Musicali »G. Lettimi«, Rimini.

Mailman, Joshua B. (2010), *Temporal Dynamic Form in Music: Atonal, Tonal, and Other*, Ph.D., University of Rochester.

Meyer, Leonard B. (1989), *Style and Music. Theory, History, and Ideology*, Chicago (IL): University of Chicago Press.

Müller, Meinard / Elaine Chew / Juan Pablo Bello (2016), »Computational Music Structure Analysis«, *Dagstuhl Reports 6/2*, 147–190. http://www.dagstuhl.de/16092 (11.12.2020)

Neuwirth, Markus (2008), »Das Konzept der Expektanz in der musikalischen Analyse: Möglichkeiten und Probleme einer kognitiv orientierten Musikanalyse«, in: *Musiktheorie im Kontext. 5. Kongress der Gesellschaft für Musiktheorie Hamburg 2005*, hg. von Jan Philipp Sprick, Reinhard Bahr und Michael von Troschke, Berlin: Weidler, 557–573.

Ockelford, Adam (2009), »Zygonic Theory: Introduction, Scope, and Prospects«, *Zeitschrift der Gesellschaft für Musiktheorie 6/1*, 91–172.

Pearce, Marcus / Martin Rohrmeier (2012), »Music Cognition and the Cognitive Sciences«, *Topics in Cognitive Science 4*, 468–484. https://doi.org/10.1111/j.1756-8765.2012.01226.x (11.12.2020)

Pressnitzer, Daniel / Stephen McAdams / Suzann Winsberg / Joshua Fineberg (2000), »Perception of Musical Tension for Nontonal Orchestral Timbres and its Relation to Psychoacoustic Roughness«, *Perception & Psychophysics 22*, 66–80.

Purwins, Hendrik / Perfecto Herrera / Maarten Grachten / Amaury Hazan / Ricard Marxer / Xavier Serra (2008), »Computational Models of Music Perception and Cognition I: The Perceptual and Cognitive Processing Chain«, *Physics of Life Reviews 5*, 151–168. https://doi.org/10.1016/j.plrev.2008.03.004 (11.12.2020)

Reybrouck, Mark (2010), »Music Cognition and Real-Time Listening: Denotation, Cue Abstraction, Route Description and Cognitive Maps«, *Musicae Scientiae 14/2*, 187–202.

Reybrouck, Mark (2015), »Music as Environment: An Ecological and Biosemiotic Approach«, *Behavioral Sciences 5*, 1–26. https://doi.org/10.3390/bs5010001 (11.12.2020)

Rohrmeier, Martin (2013), »Musical Expectancy – Bridging Music Theory, Cognitive and Computational Approaches«, *Zeitschrift der Gesellschaft für Musiktheorie 10/2*, 343–371. https://doi.org/10.31751/724 (11.12.2020)

Rothbart, Daniel (Hg.) (2004), *Modeling: Gateway to the Unknown. A Work by Rom Harré*, Amsterdam: Elsevier.

Saupe, Achim / Felix Wiedemann (2015), »Narration und Narratologie. Erzähltheorien in der Geschichtswissenschaft«, in: *Docupedia-Zeitgeschichte*, hg. vom Zentrum für zeithistorische Forschung Potsdam, Institut der Leibniz-Gemeinschaft, Version 1.0. http://dx.doi.org/10.14765/zzf.dok.2.580.v1 (11.12.2020)

Seifert, Uwe (1993), *Systematische Musiktheorie und Kognitionswissenschaft – Zur Grundlegung der kognitiven Musikwissenschaft*, Bonn: Verlag für systematische Musikwissenschaft.

Seifert, Uwe (1999), *Systematische Musiktheorie als Kognitionswissenschaft – Modellbildung und Computersimulation kognitiver Prozesse in der Geisteswissenschaft*, Habil., Universität Hamburg.

Seifert, Uwe (2008), »Kognitive Musikwissenschaft, Systematische Musiktheorie und die Frage ›Quid sit musica?‹«, in: *Musiktheorie im Kontext. 5. Kongress der Gesellschaft für Musiktheorie Hamburg 2005*, hg. von Jan Philipp Sprick, Reinhard Bahr und Michael von Troschke, Berlin: Weidler, 509–526.

Snyder, Bob (2000), *Music and Memory. An Introdruction*, Cambridge (MA): MIT Press.

Tenkanen, Atte (2010), *Comparison Structure Analysis*, Ph.D., University of Turku.

Teo, Yvonne (2020), »Measuring Harmonic Tension in Post-Tonal Repertoire«, *Empirical Musicology Review 15*, 61–104. https://doi.org/10.18061/emr.v15i1-2.6994 (11.12.2020)

Thorpe, Michael / Adam Ockelford / Aleksandar Aksentijevic (2012), »An Empirical Exploration of the Zygonic Model of Expectation in Music«, *Psychology of Music 40/4*, 429–470. https://doi.org/10.1177/0305735610392103 (11.12.2020)

Utz, Christian (2013), »Das zweifelnde Gehör. Erwartungssituationen als Module im Rahmen einer performativen Analyse tonaler und posttonaler Musik«, *Zeitschrift der Gesellschaft für Musiktheorie 10/2*, 225–257. https://doi.org/10.31751/720 (11.12.2020)

Vines, Bradley W. / Regina L. Nuzzo / Daniel J. Levitin (2005), »Analyzing Temporal Dynamics in Music: Differential Calculus, Physics, and Functional Data Analysis Techniques«, *Music Perception 23*, 137–152.

Wiggins, Geraint A. (2010), »Cue Abstraction, Paradigmatic Analysis and Information Dynamics: Towards Music Analysis by Cognitive Model«, *Musicae Scientiae 14/2*, 307–331.

Zbikowski, Lawrence M. (1997), »Conceptual Models and Cross-Domain Mapping: New Perspectives on Theories of Music and Hierarchy«, *Journal of Music Theory 41/2*, 193–225.

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

Lock, Gerhard (2020), »Salienz, Narrativität und die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung von post-tonaler Orchestermusik« [Salience, Narrativity and the Role of Musical Parameters in the Analysis of Musical Tension in Post-Tonal Orchestral Music], *Zeitschrift der Gesellschaft für Musiktheorie* 17/2, 311–349. https://doi.org/10.31751/1074

Tallinn University/ Estonian Academy of Music and Theatre [Universität Tallinn/ Estnische Akademie für Musik und Theater]

eingereicht / submitted: 10/08/2020

angenommen / accepted: 11/08/2020

veröffentlicht / first published: 23/12/2020

zuletzt geändert / last updated: 17/01/2021

A.6 Publication V (Lock 2024, 5.6)

"Musical tension: Methodological dilemmas" (in English)

Scientific Dilemmas, Research Questions (RQs), Statement (STs) and Answers

MUSICAL TENSION: METHODOLOGICAL DILEMMAS

Gerhard Lock
Estonian Academy of Music and Theatre (EAMT),
Tallinn; Tallinn University (TLU) Baltic Film, Media
and Arts School (BFM), Estonia
gerhard.lock@gmail.com

RESUME

Cet article associe l'analyse musicale et la psychologie musicale à une approche cognitive pour mieux comprendre la musique orchestrale post-tonale contemporaine. L'objectif est de formaliser cinq dilemmes méthodologiques de recherche cruciaux pour une analyse empirique de la tension musicale en tant que phénomène ondulatoire dynamique temporel (TDWP). Alors que les questions théoriques et philosophico-méthodologiques sont considérées comme « l'expérience » et « l'intuition », une compréhension kantienne élargie de la perception et de la cognition, la saillance audiovisuelle, le « changement relatif » et les principes de mesure psychophysiques. Sur la base de ces dilemmes, cinq questions et énoncés de recherche sont présentés et leurs réponses sont fournies. Dans cet article, l'auteur propose des stratégies de «résolution» qui constituent la base méthodologique pour le développement et l'application d'un modèle d'écoute cognitive (COSM) dont il est suggéré que les données de résultats (en combinaison avec d'autres méthodes) soient analysables via la narrativité afin de comprendre l'écoute cognitive. Tension Musicale (CMT).

This article intertwines music analysis and music psychology with a cognitive approach to better comprehend contemporary post-tonal orchestral music. The aim is to formalize five research-methodological Dilemmas crucial for an empirical analysis of musical tension as a temporal dynamic wavelike phenomenon (TDWP). As theoretical and philosophical-methodological issues are considered 'experience' and 'intuition', an enlarged Kantian understanding of perception and cognition, audiovisual salience, 'relative change' and psychophysical measurement principles. Based on these Dilemmas five Research Questions and Statements are presented and their Answers are provided. In this article the author proposes "solving" strategies that are the methodological basis for the development and application of a cognitive listening model (COSM) which's outcome data (in combination with other methods) is suggested to be analyzable via narrativity in order to comprehend Cognitive Musical Tension (CMT).

1. INTRODUCTION

The aim of this article is to formalize **five research Dilemmas**[1] I have encountered during my doctoral studies. These Dilemmas are **methodological in nature** – musical tension can be researched only with reservations: somehow with exact physical measurement tools, applying music theoretical procedures (verbal/ narrative/ visualizing means), and empirically. One has to deal with limitations and simplifications both on the object-of-research (accessibility/affordance), the data-gaining and data-analysis/-presentation/interpretation-side (matters of scientific modeling [31]), and needs to choose between alternatives. The basic dilemma-type problem is the following: aiming at a goal or to achieve an outcome one has to waive certain aspects because of limitations and simplifications to be faced with.

The basic goal of this article and my dissertation in progress[2] on **musical tension** is to intertwine music analytical (theoretical) and music psychological (empirical) thinking with a cross-domain and generally a cognitive approach aiming to better comprehending contemporary post-tonal [44] orchestral music.

As general theoretical background I assume that (after [3] and [20]) Sound and "Music" may appear in several domains [54], but Music, based on human perception and cognition via the auditory (perceived) domain, therefore via **listening**, is best comprehended as **cognitive** [54] (see Figure 1 below). Furthermore, Jean-Claude Risset has expressed that "music is flux, movement, it is not an abstract form, it exists only in its

---

[1] To look at dilemmas was proposed by prof. dr. Mauri Kaipainen and I am thankful to him for his thorough review and suggestions to improve the cognitive part of this article (see Footnote 2).

[2] Working title "Methodological contributions to a cognitive analysis of perceived structural musical tension in contemporary post-tonal orchestral music". Supervised by head of musicology, curriculum coordinator and professor of music theory dr. Kerri Kotta (Estonian Academy of Music and Theatre, Tallinn, Estonia) and advised by cognitive musicology and media technology em. professor dr. Mauri Kaipainen (Södertörn University, Sweden) – currently adjunct professor in cognitive science (University of Helsinki, Finland), and member of the Enactive Virtuality Lab research group at Tallinn University BFM.

incarnation in time and in sound [...]" ([43], cf. translation in [7] p 16).

Music has been analyzed cognitively, e.g. with principles of "immanent musical analysis" during listening as elaborating/structuring process of "relevant categories" [6] and modeled computationally as an ecologist cognition/knowledge of music based on of the individual's mind's cognitive dynamics [18]. Cue abstraction and information dynamics have been dealt with while modelling cognitive music analysis [55]. For an overview of computational modeling of music cognition and the process of musical composition see [56].

In the context of a three-angles view on how we can understand the world in relation to technology and AI [19][3] powerful supportive possibilities of music analysis in the **physicalist** machine-like-brain metaphor context still have limits related to their aims, data and decoding matters [47], [2], [38]. Integrative Musicology suggests that a so-called **representational barrier** seems nowadays crossable via combining cognitive science, digital signal processing, hardware and software engineering – a clear musicological agenda is needed to coordinate and realize this development [9]. For an insight into the past and future of the "science-music borderlands" see [35]. However, taking into account the second view – **the mind outside physicalism existence** [19] – and assuming that **music** and **tension** take place in the mind being approachable best via listening as cognitive, enables to not being stuck to the so-called explanatory or semantic gap [54].

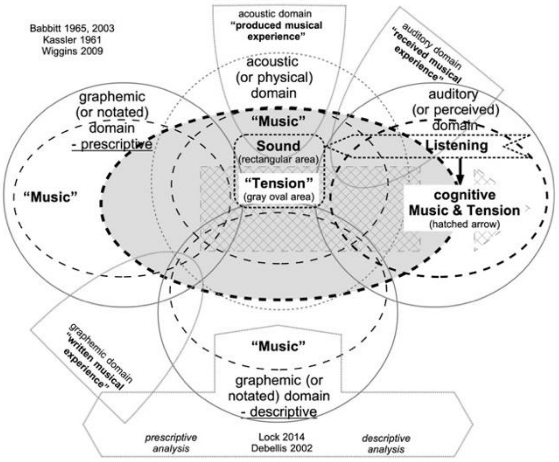


Figure 1. Representational domains of Sound/Music & Tension and the comprehension of Music & Tension in the context of analysis, perception and experience.

The domains in Figure 1 are visualized based on [54] after [3] (four full circles) and [20] (three half-circles). The large hatched left-to-right arrow in the middle expresses that, although music (dashed circles) and tension (gray oval area) are located between all these domains, listening in the auditory (perceived) domain is crucial for identifying music and tension as cognitive. The large three-part arrow at the bottom of the figure shows the pre- and descriptive aspects that can be attributed also to analysis [10], not only to the graphemic (notated) domain [30].

As shown in Figure 1 above, it seems that the non-physical cognitive phenomenon Music (existing only in the mind [54]), represented in extended four domains[4], can only be grasped adequately via **Listening** (auditory, perceived domain) based on **experience as cognition** (musical sense-making and the 'in-time/outside-of-time' dichotomy [42]). The same, I claim, can be applied to **Tension in Music** – I furthermore call it **Cognitive Musical Tension (CMT)**.

As philosophical-methodological background I have taken the Kantian understanding of cognition based on a variety of sensible intuitions that are crucial for judgement-based knowledge as "the discovery of the one in many" sensible representations [57].[5]

Kantian *objective* cognition requires experience based on *a priori* transcendental principles and "something mind-independent that affects the senses" [25]. Furthermore, I advocate an enlarged understanding of Kantian extensive magnitudes that enables also numerical quantification of intensities objectively if using "external, intersubjectively shared standards that have to be spatial" – a forerunner to later psychophysics [25] and quantitative (representational) psychophysical measurement (a psychological sensation between the physical stimulus and the human response) [48], [50], [36], [52].

I have chosen the perceptual mechanism of **salience** ([27], [17]) to firstly developing a system of musical-score-based affirmative listening involving **salient features[6]** while comparing this music theoretically driven rather speculative [40] outcome to slider-controller empirical experiment results [28], [29], [33]. Later, while abandoning the musical score (graphemic domain), I focused on just one promising (auditory domain) salient feature **Impulse** (Im) that appeared to be applicable also empirically. I applied it because its mediative audiovisual abilities [32].

The object of my research, **contemporary post-tonal orchestral music**, represented by two works of the Estonian composer Erkki-Sven Tüür (b. 1959)[7] differs from clear classical/tonal structures (bordered motifs,

---

[3] 1) physicalism – from the mind-as-machine/brain-as-machine metaphor to a machine-like-brain metaphor, artificial neural networks and deep learning, 2) a reality outside the physical world in contrast to physicalism – see mind-body/world problem, 3) a holistic view – connection of biological and technological approaches.

[4] Taking also into account the distinction of the graphemic domain into pre- and descriptive domains [30]. Also analysis can be understood as pre- and desriptive [10].

[5] Kantian *sensibility* as a basis for music theoretical thinking has been proposed to more adequately comprehend Schenkerian analysis [22], see an actual discussion in [39].

[6] Detectable in relation to musical form: Impulse (Im) and Culmination (Cu) at local level (structure), Contrast (Co) and Contrast Culmination (CC) at global level (form) – see [28], [29].

[7] "Oxymoron" (2003) for large ensemble, 4th Symphony/Percussion Concerto "Magma" (2002).

themes, harmonic schemes). The composer merges several contrasting musical materials, temporal organization types (“continuous”, “slow” and “quick time” [23]), styles and composition techniques[8] while creating a dense symphonic wavelike texture (including outstanding high- (maximum) and low-(minimum) points as formal borders [28], [29][9] that still inherits classical formal principles [24]. This, together with a stylistic cross-over between avant-garde and rock music, makes his music more easily comprehensible by a larger audience and is just proper for developing methodological research towards a wider applicability to a similar repertory furthermore.

Why put the focus on this kind of music? Because it is complex and rather complicated to analyze, and there is a strong need to include computational analysis and neuroscientific modeling (e.g. [12], [13], [37]). Therefore, the **general aim of my dissertation** in progress (the present article presents key problems and “solving” strategies) is to i) manage the complexity of contemporary orchestral music, ii) advance analytical methods (using empirical and partly computational means), and iii) advance systematic methods for musical tension analysis. Computational means are applied in the process of data retrieval (using hard- and software) as well as data analysis and visualization (e.g. data mining and simple common statistical means).[10]

2. THE FIVE DILEMMAS

This section introduces the five Dilemmas that need to be addressed while developing methodological tools for analyzing **cognitive musical tension (CMT)**. The Dilemmas are presented with an inherent progression concerning suggested solvability of questions posed in previous Dilemmas by a following Dilemma. Dilemmas 2–5 are visualized with hypothetical schemes that are interrelated in their content and symbols’ meanings. The Dilemmas are the base for the Research Questions (RQ's) and Statements (ST's) presented in section 3. The RQ's and ST's are answered in section 4 based on my earlier dissertation-related publications. For crucial methodological aspects of representation and visualization of music and analysis/research data (intuitive, systematic, system-based and automatized strategies, matters of its automatization) see [30].

2.1. First Dilemma

While assuming that tension in music resides in the mind and not only in just one domain, it is compound and cognitive – like music itself (see Figure 1). Furthermore, **tension** and suspense can be understood as universal psychological phenomena (especially in the arts) beyond emotions [23]. It is perceivable consistently by listeners with whatever background or previous familiarity with the music [15]. Segmentation, tension and relaxation function as separate “common perceptual processes” based on temporal dynamic memory and are independent of grammar or syntax elements [1]. However, **cognitive musical tension (CMT)** is often verbally (and less systematically) described e.g. in its similarity to drama theory[11] or as a cross-domain everyday life (real-world) analogy of a (e.g. natural, physical, electromechanical, psychological or metaphorical) wavelike phenomenon. What follows from the previous is that it is well commonly experienced, but complicated to describe and formalize – while being cognitive it cannot be examined directly.

2.2. Second Dilemma

Intensity and energy levels in sound and music can be detected based on defined parameters in separate domains (mainly in the acoustic/physical domain, also in the auditory/perceived domain, e.g. on dB scales) at single time point values using technical tools and in relation to a technical absolute zero (see Figure 2). But intensity and energy measured in sound and music do not necessarily equal cognitive musical tension (CMT). The question arises, what are other relevant features and parameters and how to measure them based on the idea of ‘relative change’. Instead, measurement based on ‘relative change’ and salience seems to be appropriate; see next Dilemmas.

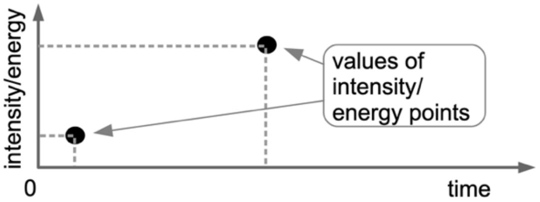


Figure 2. Second Dilemma. Single hypothetical values of intensity or energy of music in the acoustic/physical domain measured in relation to a technical absolute zero.

2.3. Third Dilemma

The musical events in form of high points (maxima, short duration) or *plateaux* (culminations, longer duration) as part of a generalizable culmination “arch”

---

[8] Tüür applied his “vectorial” composition technique first in “Oxymoron” (2003).

[9] Based on empirical results retrieved via analysis (conducted by musicians and non-musicans) of music by Milhaud, Webern and Maderna (that more or less in- or excludes traditional tonal aspects) segmentation, tension and relaxation built a separate “common perceptual process” with constant and variable elements based on “sameness” and “difference” at segmentation borders that doesn’t depend on grammar or syntax [1].

[10] However, the data are neither retrieved nor analyzed biometrically or at a neuronal level.

[11] Although the analogy with drama theory is highly relevant as a narrative domain, this dissertation will not deal with comparisons in this direction, because it deals with narrativity on a more abstract level.

of wavelike cognitive musical tension (CMT) can be detected well[12] via salience (as salient features) in contemporary post-tonal orchestral music, but not the beginning or ending points of such a tensional development – despite the fact that a culmination/high point/maximum is generally defined by three points ("start point" of increase of tension, maximum, decrease of tension towards an "end point")[13]. The main complications are found in the detection of such beginning/ending points in relation to the structure and form of the music.

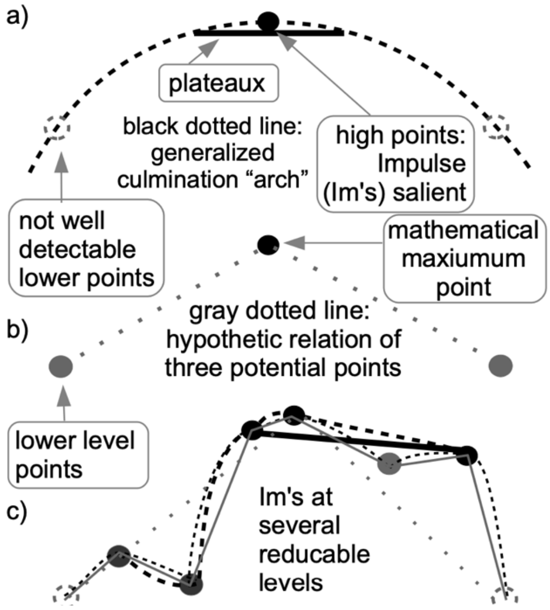


Figure 3. Third Dilemma. Musical events (a) as high points or *plateaux* as part of a culmination "arch"; (b) as maximum defined by its lower neighboring points; (c) with some additional hypothetical Im's at several reducible levels. The meaning of the graphical symbols is the same in all panels.

2.4. Fourth Dilemma

Cognitive musical tension (CMT) can be comprehended as temporal dynamic (wavelike) phenomenon (TDWP) – a development that is experienced as taking place between two time points (musical events/Impulses, Im's) as the 'relative change' of to-be-defined musical parameters (MP's). However, it remains unclear how their "perceived intensity" can be measured to allow to comprehend cognitive musical tension (CMT). One has to figure out what MP's are the 'content' of the Im's. This cannot be achieved methodologically via single-point detection unless extending the Im's temporally into short temporal areas (chunks, adding seconds before and after) (see Figure 4 below).

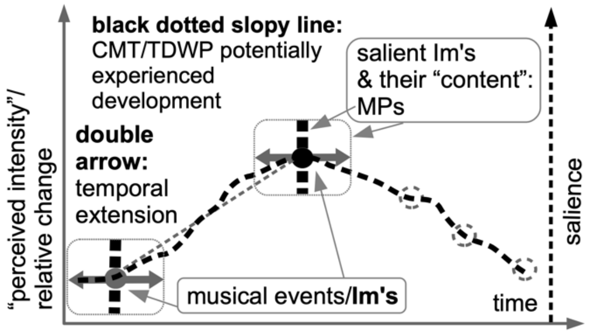
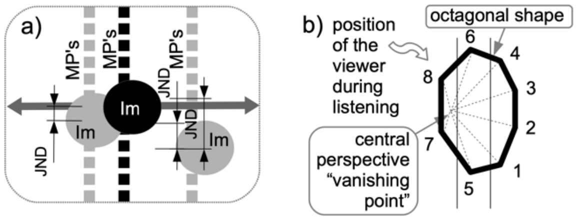


Figure 4. Fourth Dilemma. CMT/TDWP potentially experienced development. The methodologically suggested temporal extension (double arrow box) and the "content" (musical parameters, MP's) of the salient Impulses (Im's) is also depicted (see fifth Dilemma below).

2.5. Fifth Dilemma

While assuming that the 'relative change' of salient musical parameters[14] and their interplay creating "perceived intensity" cause the 'relative change' of cognitive musical tension (CMT) between two or more musical events (Impulses, Im's), the following problems remain: i) how to gain insight into the Im's, more specifically how to observe their "content" – MP's, ii) how the MP's can be weighted in their salience between each other. Furthermore, the role of primary (PP) and/or secondary (SP) musical parameters (MP's) remains unclear in contemporary post-tonal orchestral music, also how their 'relative change' can be detected (e.g. collecting empirical data) and represented/visualized.



temporal extension: time window

8 – visually most salient point, 7–1 decreasing salience

Figure 5. Fifth Dilemma. More detailed excerpt from Figure 4. Impulses (Im's) together with their musical parameters (MP's) (a). The octagonal shape and its visual salience mechanism enable the weighting of the auditive salience of MP's (b) (see more in section 4.4 and 4.5).

---

[12] So called low points (minima) in musical tension are rather ambiguous: technically they should be of low tension/intensity/energy, but such a low point or *plateau* of material energy can be really intense psychologically, therefore loaded with tension and suspense [21], [26].

[13] "Start point" and "end point" are put in quotation marks, because they might not always be definable well, even in classical/tonal music.

[14] Primary musical parameters (PP) are pitch, rhythm, harmony. Secondary musical parameters (SP) are dynamics, tempo, instrumentation/timbre, texture and effects. This list of PP's and SP's is modified after [21] and applied here [32].

The psychophysical measurement principle “just noticeable differences” (JND) (Figure 5a) is applicable between 2 (or more) Im's (gray & black dots) as well as between musical parameters (MP's) (gray & black dotted vertical line). The black Im's/MP's are in their JND more salient than the gray Im's/MP's. The listener applies this intuitively (being rather unaware of this mechanism) while focusing on audiovisual salience aspects. The problem remains how the MP's can be weighted in their salience; see more in section 4.5.

3. THE RESEARCH QUESTIONS (RQ'S) AND STATEMENTS (ST'S)

This section presents the Research Questions (RQ's) and Statements (ST's) in the role of hypotheses relevant for my dissertation. They are based on the Dilemmas presented in section 2. The RQ's and ST's are answered in section 4 concerning methodological “solving” strategies offered in my previous dissertation-related publications.

3.1. RQ and ST 1

RQ1: What are general principles to examine and comprehend cognitive musical tension (CMT)?

ST1: Cognitive musical tension (CMT) can be examined similarly to everyday life (real-world) temporal dynamic wavelike phenomena (TDWP) via analogies of cross-domain modeling and principles of psychophysical measurement, and via a yet-to-be-specified application of narrativity.

3.2. RQ and ST 2

RQ2: From Dilemma 1/RQ/ST1 follow-up questions arise: what are relevant features or parameters beside intensity/energy of sound/music, where they are located, how they can be detected/measured to enable to examine and comprehend cognitive musical tension (CMT)?

ST2: Instead of absolute acoustic/physical domain measurements of intensity/energy levels of sound/music the examination of cognitive musical tension (CMT) can be conducted using the principle of salience. Salient features in relation to musical events [28], [29], [33] have been proposed applying the principle of ‘relative change’.

3.3. RQ and ST 3

RQ3: What is an appropriate way of examining/detecting/interpreting/comprehending wave-like cognitive musical tension (CMT) in contemporary post-tonal orchestral music, if salient features of a culmination “arch” cannot reliably be detected/measured in important aspects (start/end points: formal segmentation problems caused by the absence of classical clear-cut formal units) in respect to musical structure and form?

ST3: If start/end points of culmination “arches” cannot be detected, the problem of formal segmentation of contemporary post-tonal orchestral music remains actual and does not help to comprehend cognitive musical tension (CMT). Instead, the focus is directed towards Impulses (Im's) that can be either high points/maxima (short duration), part of *plateaux* (areas of high tension, longer duration), or other musical events that trigger salience. From the proposed salient features ([28], [29], [33]) only the Impulse (Im) is methodologically reasonable to be applied [32] because of its “one-dimensionality” (only one time point on the horizontal axis, no vertical axis value). Other salient features are way too complex in their multidimensionality.

3.4. RQ and ST 4

RQ4: Based on what principles can cognitive musical tension (CMT) – experienced and detected as a temporal dynamic (wavelike) phenomenon (TDWP) taking place between two or more time points (salient musical events/Impulses) based on the ‘relative change’ of musical parameters (MP's) – be detected/measured and comprehended?

ST4: Concepts enabling CMT's temporal comprehension are: attentive listening [11] or music listening as ongoing (real-time) process of sense-making via focal attention [41], intuitive experience [8], perceptual experience and perceptual justification [49], microgenesis of brief time spans [45], [4]; temporal dynamic forms [34], dynamic forms of vitality [51]; narrativity [32]. These are based on Kantian perception/cognition mechanisms [57] including intensive quantities measurement [25] with COSM functioning as external, spatial and intersubjectively agreeable standard – a spatiotemporal object to connect intensive and extensive magnitudes’ measuring based on Impulses and their salience as relative change.

3.5. RQ and ST 5

RQ5: Which musical parameters (MP's) – their salience, relative change and interplay that trigger perception – enable the examination/detection/measuring and comprehension of cognitive musical tension CMT) based on what kind of salience principle?

ST5: Musical parameters (MP's) constitute the musical structure, musical events that are perceivability as Impulses (Im's) that cause a change in perception based on their relative salience. In contemporary post-tonal orchestral music mostly secondary musical parameters are hypothesized to trigger Impulses (Im's), therefore enabling comprehension of cognitive musical tension.

4. ANSWERS TO THE RQ'S AND ST'S OFFERING METHODOLOGICAL SOLUTIONS

This section presents answers to the RQ's and ST's offering methodological solutions proposed by the author's dissertation-related empirical-data-including publications [28], [29], [33], [32]. To "solve" these Dilemmas methodologically, I have developed **COSM: Cognitive Octagonal Slice Model** for this dissertation, applying my **Twelve Strategic Steps for modeling/analyzing scientific models** [31].

4.1. Answer to RQ and ST 1

While treating musical tension as a psychological sensation between the physical stimulus and the human response, principles of psychophysical measurement can be applied to examine it. Its main principle, "just noticeable difference" (JND) detection, is structurally analogous to a structurally/abstractly understood narrativity: a "story" being "told" based on a perceived (significant or just sufficient noticeable difference) development from event A to B is residing between the physical stimulus (the acoustic domain) and the human response to it (via the auditory domain). Modeling itself is considered as working in the "cognitive space of ideas" and as "storytelling" [46]. While the most common simplification of an abstract conceptualization of narrativity concerns just "one" dimension (movement from event A to B) in respect to time (on the horizontal axis), the interpretation (modeling) as a wavelike phenomenon adds a second dimension (on the vertical axis) that reveals the nature of the events (less/more loaded with tension) and the movement between them in respect to its "intensity". This "loadedness with tension" or "intensity" needs a method to examine or detect it by a listener. I have chosen the concept of salience; the model developed (COSM) includes especially audiovisual salience as a cross-domain aid. Salience is measurable by the listener intuitively (and in a rather unaware manner) while searching for "just noticeable differences" (JDN's, psychophysical measurement) between incoming sense data – see also Dilemma/RQ/ST/Answer 5.

4.2. Answer to RQ and ST 2

Salient features constitute the structure (on local level) and form (on global level) of the music [28], [29], [33]. The principle of relative change is crucial, because cognitive musical tension (CMT) does not reside in one domain only, especially not in the acoustic/physical domain alone. The follow-up question remains: what are appropriate features or parameters relatively changing; see next Dilemmas, RQ's, ST's and Answers to them.

4.3. Answer to RQ and ST 3

Because start/end points of local culmination "arches" cannot be easily detected due to formal segmentation problems of contemporary post-tonal orchestral music, the "time-related" and "one-dimensional" salient feature **Impulse (Im)** offers the best capability of pairing up with the **musical event**. However, its "time-unrelated" "multidimensionality" requires the exploration of its "content": musical parameters. Also (like the musical event itself) the Impulse is a compound phenomenon, its "multidimensionality" still needs to be explored to examine responsible musical parameters (MP's) that are the "content" of an Impulse.

4.4. Answer to RQ and ST 4

Neither the absolute nor the relative states of "intensity" of musical parameters (MP's) can reliably be measured in an instant to serve the comprehension of cognitive musical tension (CMT). It needs a time span, even to be relatively brief (see microgenesis [45], [4]) and the consideration of time delays in perception (no directness in human perception is possible [4]) to become conscious in mind, that makes it possible to examine, detect, measure and comprehend cognitive musical tension (CMT). The 'relative change' of to-be-defined musical parameters, their interplay can have larger or, like in psychophysical measurement, just noticeable differences (JND's) between two or more time points (musical events/Impulses, Im's). These 'relative changes' are perceived as *pre- (or semi-)categorial objects** in inner perception [25] and their interplay based on tension and relaxation of musical parameters (MP's) in outer sensation can be seen as *protonarratives* based on dynamic forms of vitality [16][15] creating such a "perceived intensity".

4.5. Answer to RQ and ST 5

The audiovisual-salience-based **COSM: Cognitive Octagonal Slice Model** for empirical data retrieval [32] aims to "solve" the weighting problem of the musical parameters' salience. Musical parameters (MP's) have been generally categorized into primary (PP: pitch, rhythm, harmony) and secondary parameters (SP: dynamics, tempo, instrumentation/timbre, texture, effects). Primary parameters are generally extensive, many secondary parameters are intensive in their nature. Extensive magnitudes are additive successions, intensive magnitudes are degree-based. While applying an experience-based approach in this dissertation in progress, the psychophysical measurement of intensive magnitudes is related to extensive magnitudes, because it needs an external spatial intersubjectively shared standard (e.g. an extensive object) which makes possible the numerical quantification of intensities. COSM has

---

[15] Developed based on earlier Publications; see in [51].

the role of such an external spatial intersubjectively shared standard (e.g. an extensive object). However, my latest study [32] on musical tension, salience, musical parameters and narrativity has revealed that the postulation of the prevalence of secondary MP's in post-tonal music [5] doesn't apply (fully) to contemporary post-tonal orchestral music in the late the style of Tüür. Secondary parameters (SP's) appeared dominating only in the high culmination area in the middle of the music before the drum-set solo cadence and created a clear narrative (based on principles of clarity of structure/in the empirical data as clear narrativity applied also in data mining).

As Figure 6 shows MP's in COSM must be attached by the participants per their own previously detected and enlarged Impulses (+/–7=14 seconds) to the corners of the octagonal shape (see also Figure 5) in accordance to their audiovisual salience [32]. Data analysis includes clustering (based on all Impulses' data to gain significant time windows; 48 clusters for Tüür's 30 min duration 4th Symphony/Percussion Concerto) with the data mining algorithm DBSCAN [14], several common data analysis and visualization tools (averaging, smoothing, summing up, simplification operations like minimum-maximum or start-endpoint extraction, number coding, diagrams and matrix visualization).

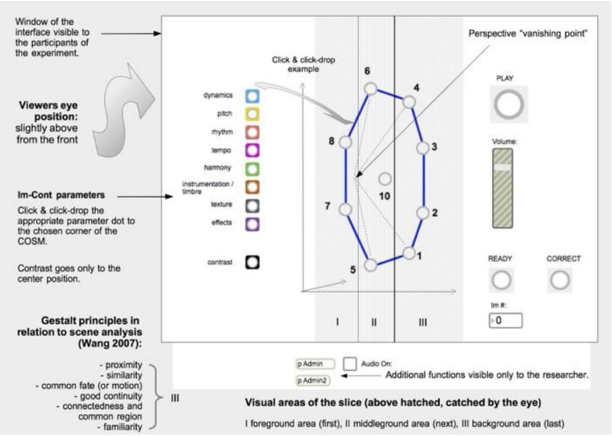


Figure 6. Principles of the visual interface of the COSM slice based on gestalt principles. This Figure combines the earlier issued separate Figures: principle of the octagon with gestalt principles (Figure 2 in [32]), COSM slice Max/MSP interface used by the experiment participants (Figure 3 in [32]). Im = Impulse, Cont = 'content' = musical parameters. Gestalt and scene analysis principles see [53].

5. IMPLICATIONS AND CONCLUSIONS

In this article five Dilemmas concerning the research of cognitive musical tension (CMT) have been introduced together with my dissertation-in-progress-related Research Question (RQ's), Statements (ST's) and Answers.

Musical tension is a temporal dynamic wavelike phenomenon (TDWP), but its examination as a compound and multidimensional phenomenon is not easily possible. In contemporary post-tonal orchestral music the start and end points of maximums/culmination "arches" are rather impossible to define clearly. So, one cannot examine successfully longer durational salient features. Instead, Impulses (Im's) as single short durational phenomena (musical events) can be examined. One then needs to extend these Impulses in time again (as a time window) to achieve a reasonable duration within which the triggers of these Impulses (Im's) and their "content", musical parameters (MP's), can be examined, too. This time-window is needed for single participants as well as empirical data based on those time windows analyzable with the DBSCAN algorithm clusters. Both salient Im's and MP's relative changes are similar to the psychophysical measurement principle „just noticeable differences" (JND's). However, the listener applies this intuitively (being rather unaware of this mechanism). Based on this one should be able to compare two or more Im's (and their "contents", MP's) towards a reasonable interpretation of the tensional development of contemporary post-tonal orchestral music (as succession of musical events/Impulses) as a **narrative**.

To better understand this narrative approach, I present the following definition of **'Perceived structural musical tension'**[16] **(PSMT)**:

PSMT can be comprehended based on general cognitive phenomena (musical events as impulses) and processes (focal attention and listening, mental pointing) via a cognitive mechanism (salience) as a cross-domain concept (narrativity) connecting perception and cognition as a listening story based on structural elements/musical parameters.

PSMT (like narrativity) is a dynamic principle with teleological impulses, musical events/structural patterns/a piece are actions in flux and a journey, a "real-time" discovery process/algorithm (requiring attentive listening, mental pointing) of musical events/parameters that built limited archetypical patterns that are in conflict on various hierarchical levels causing constant revaluation of these elements, assuming that the unique surface features (cues, salient features) and effects of these patterns (saliency principle) should be psychologically meaningful.

In contemporary post-tonal orchestral music (represented by the late music of Tüür) it is the analogy i) of such music to a wavelike phenomenon (via modeling) that enables to focus on the detectable Impulses (Im's) (either high points/maxima/culmination areas, or other salience triggering musical events; ii) of wavelike phenomena and cognitive musical tension

---

[16] This phrase – perceived structural musical tension – I have also used in my latest empirical study [32], but its first definition there is still based on my earlier definition in relation to musical form [33]. However, a considerable part of the theoretical background on music and narrativity underlying the above presented new definition can already be found in [32]. Therefore it is here now formulated in relation to cognition and narrativity, and the next step towards a cross-domain solution my dissertation in progress aims to present.

(CMT) to narrativity that enables to comprehend/interpret cognitive musical tension (CMT) with cross-domain/modeling/visualization tools. Hereby it was necessary to abandon the detection of segmentation points like in classical/tonal musical form as well as the durational salient features.

To conclude, the key concepts enabling me to propose my "solving" strategies for these research-methodological Dilemmas in this article and my dissertation in progress via answering the research questions and statements are the following: intuitive salient Impulse detection, audiovisual salience principles of an octagonal shape that functions in COSM as external, spatial and intersubjectively agreeable standard (enlarged Kantian principle) necessary in psychophysical measurement of extensive and intensive magnitudes. These concepts are based on 'relative change' as "just noticeable differences" (JND) principles.

6. REFERENCES

1. Addessi, A. R. & Caterina, R. «Perceptual Musical Analysis: Segmentation and Perception of Tension», *Musicae Scientiae* 4/1, 31–54, 2000. https://doi.org/10.1177/10298649000040

2. Andersen, J. S. How to Think Music with Data. Translating from Audio Content Analysis to Music Analysis (Ph.D dissertation). University of Copenhagen, Copenhagen, 2017.

3. Babbitt, M. «The Use of Computers in Musicological Research (1965)», S. Peles, S. Dembski, A: Mead, J. N. Straus (eds), *The Collected Essays of Milton Babbitt* (202–212), Princeton University Press, Princeton, 2003.

4. Bachmann, T. Microgenetic Approach to the Conscious Mind. John Benjamins, Amsterdam, Philadelphia, 2000.

5. Bauer, A. «‹Composing the Sound Itself›: Secondary Parameters and Structure in the Music of Ligeti», *Indiana Theory Review* 22/1, 37–64, 2001.

6. Chouvel, J.-M. «Categories and Representation in Musical Analysis», *Sonus* 35/1, 17–35, 2014.

7. Christensen, E. The Musical Timespace. A Theory of Music Listening. Aalborg University Press, Aalborg, 1996.

8. Clifton, T. «Some Comparisons between Intuitive and Scientific Descriptions of Music», *Journal of Music Theory* 19/1 66–110, 1975.

9. Dahan, K. «Of Music, Mind, and Machines: Towards an Integrative Musicology». *Music, Mind, and Invention Workshop* (The College of New Jersey, 31.03.2012) (1–4). College of New Jersey, Ewing Township, 2012. http://www.tcnj.edu/~mmi/

10. Debellis, M. «Musical Analysis as Articulation», *The Journal of Aesthetics and Art Criticism*, 60/2, 119–135, 2002.

11. Deliège, I. & Melen, M. «Cue Abstraction in the Representation of Musical Form», I. Deliège & J. Sloboda (eds), *Perception and Cognition of Music* (387–412), Psychology Press, New York, 1997.

12. Dubnov, S., McAdams, S. & Reynolds, R. «Structural and affective aspects of music from statistical audio signal analysis», *Journal of the American Society for Information Science and Technology* (Special Issue on Style) 57/11, 1526–1536, 2006. https://doi.org/10.1002/asi.20429

13. de Paiva Santana, C. & Guigue, D. «A Computer Model for the Analysis of Diversity and Coordination in Orchestration», *Proceedings of EVA 2020 Conference* (AI and the Arts: Artificial Imagination), 120–127, 2020. http://dx.doi.org/10.14236/ewic/EVA2020.21

14. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. «A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise», *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)* (226–231), AAAI Press, Palo Alto, 1996.

15. Farbood, M. «A Parametric, Temporal Model of Musical Tension», *Music Perception* 29/4, 387–428, 2012. https://doi.org/10.1525/mp.2012.29.4.387

16. Imberty, M. «The Question of Innate Competencies in Musical Communication», N. L. Wallin, B. Merker, S. Brown (eds), *The Origins of Music* (449–462), Cambridge, MA, MIT Press, 449–462, 2000.

17. Jensenius, A. R. How Do We Recognize a Song in One Second. The Importance of Salience and Sound in Music Perception (Master of Arts). Music Technology Department of Music and Theatre University of Oslo, Oslo, 2002.

18. Kaipainen, M. Dynamics of Musical Knowledge Ecology. Knowing-What and Knowing-How in the World of Sounds (Ph.D dissertation). University of Helsinki, Helsinki, 1994.

19. Kaipainen, M. & Tikka, P. «Artificial Intelligence and Art. What's Left After the Hype?» R. Kelomees, V. Guljajeva, O. Laas (eds), *The Meaning of Creativity in the Age of AI* (123–149). Estonian Academy of Arts, Tallinn, 2022.

20. Kassler, M. The Decision of Arnold Schoenberg's Twelve-Note-Class System and Related Systems. Princeton University Press, Princeton, 1961.

21. Kirschbaum, M. Höhepunktbildung und Dramaturgie in Neuer Musik. Die Erstellung «formaler Spannungsverläufe» als Ergänzung der musikalischen Analyse. Dohr, Köln, 2001.

22. Korsyn, K. «Schenker and Kantian Epistemology», *Theoria: Historical Aspects of Music Theory*, 3, 1–58, 1988.

23. Kotta, K. «Zur musikalischen Zeit: Über die formalen Grundlagen der Werke von Erkki-Sven Tüür – ein Beschreibungsversuch», H.-K. Jungheinrich (Hrsg), *Schwarze Milch und bunte Steine. Der Komponist Erkki-Sven Tüür* (23–38), Schott, Mainz, 2008.

24. Kotta, K. «Classical Formal Archetypes in Selected Works by Erkki-Sven Tüür», *Lietuvos muzikologija [Lituanian musicology]* 12, 102–110, 2011.

25. Kraus, K. T. «Quantifying Inner Experience? – Kant's Mathematical Principles in the Context of Empirical Psychology», *European Journal of Philosophy* 24/2, 331–357, 2013/2016. https://doi.org/10.1111/ejop.12068

26. Lehne, M. & Koelsch, S. «Towards a General Psychological Model of Tension and Suspense», *Frontiers in Psychology* 6, 1–11, 2015. https://doi.org/10.3389/fpsyg.2015.00079

27. Lerdahl, F. «Atonal Prologational Structure», *Contemporary Music Review* 4, 65–87, 1989.

28. Lock, G. «Muusikalist pinget mõjutavad tegurid ja nende tajumine Erkki-Sven Tüüri teoses ‹Oxymoron›» [On Factors Influencing Musical Tension and its Perception in Erkki-Sven Tüür's piece »Oxymoron«], *Res musica* 2, 63−74, 2010. https://resmusica.ee/wp-content/uploads/2017/04/rm2_2010_63-74_Lock.pdf

29. Lock, G. «Tension Design Analysis as Listening Analysis Strategy for Contemporary Music: Perception of Tension in Contemporary Post-Tonal Orchestral Music: a Case Study» (Poster-Presentation), *11th International Conference on Music Perception and Cognition (ICMPC11)*, Seattle, WA, 2010. http://depts.washington.edu/icmpc11/

30. Lock, G. «Visualisierende Musikanalyse. Erkki-Sven Tüür: *Oxymoron*», F. Diergarten (Hrsg.), *Musikalische Analyse. Begriffe, Geschichten, Methoden* (287−319), Laaber, Laaber, 2014. (Grundlagen der Musik 8, Hrsg. F. Diergarten & M. Gervink).

31. Lock, G. «Models and scientific modeling in music context», *Problems in Music Pedagogy* 16/2, 25−45, 2017. http://pmp.du.lv/wp-content/uploads/2016/08/PMP_2017_Vol-16-2.pdf

32. Lock, G. «Salienz, Narrativität und die Rolle musikalischer Parameter bei der Analyse musikalischer Spannung von post-tonaler Orchestermusik», *ZGMTH: Zeitschrift der Gesellschaft für Musiktheorie* 17/2, 311–349, 2020. https://doi.org/10.31751/10742020

33. Lock, G. & Kotta, K. «Musical Tension as a Response to Musical Form», *Proceedings of the 12th International Conference on Music Perception and Cognition« (ICMPC) / 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)* (612–617), Aristotle University of Thessaloniki, Thessaloniki, 2012. http://icmpc-escom2012.web.auth.gr/proceedings.html

34. Mailman, J. B. Temporal Dynamic Form in Music: Atonal, Tonal, and Other (Ph.D dissertation). University of Rochester, Rochester, 2010. http://hdl.handle.net/1802/12513

35. Margulis, E. H., Loui, P. & Loughridge, D. The Science-Music Borderlands. Reckoning with the Past and Imagining the Future. The MIT Press, Cambridge, MA, 2023.

36. Michell, J. «Representational measurement theory: Is its number up?», *Theory & Psychology* 31/1 3–23, 2020. https://doi.org/10.1177/09593543209308

37. Noble, J., Thoret, E., Henry, M. & McAdams, S. «Semantic Dimensions of Sound Mass Music: Mappings between Perceptual and Acoustic Domains», *Music Perception* 38/2, 214–242, 2020. https://doi.org/10.1525/mp.2020.38.2.214

38. Onwuegbuna, I. E. «Diginalysis: The Man-Machine Collaboration in Music Analysis», S. M. Brito (Ed.), *Toward Super-Creativity – Improving Creativity in Humans, Machines, and Human-Machine Collaborations* (online version), IntechOpen, London, 2020. https://doi.org/10.5772/intechopen.84355

39. Parkhurst, B. «Completing the Triad: Schenker and Kantian Practical Philosophy», B. Parkhurst, J. Swinkin (eds), *Perspectives on Contemporary Music Theory* (260–329), Routledge, New York, 2024. http://dx.doi.org/10.4324/9781003357797-13

40. Pearce, M. T. The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition (Ph.D dissertation). City University London, London, 2005.

41. Reybrouck, M. «Music Cognition and Real-Time Listening: Denotation, Cue Abstraction, Route Description and Cognitive Maps», *Musicae Scientiae*

14/2, 187–202, 2010. https://doi.org/10.1177/10298649100140S211

42. Reybrouck, M. «Experience as cognition: musical sense-making and the ‘in-time/outside-of-time’ dichotomy», *Interdisciplinary Studies in Musicology* 19, 53–80, 2019. https://doi.org/10.14746/ism.2019.19.4

43. Risset, J.-C. «Timbre et synthèse des sons», *Analyse Musicale*, 2, 9–20, 1986.

44. Roig-Francolí, M. A. Understanding Post-Tonal Music. McGraw-Hill, Boston, 2008. https://archive.org/details/understandingpos0000roig

45. Rosenthal, V. «Microgenesis, immediate experience and visual processes in reading», A. Carsetti (ed), *Seeing, Thinking and Knowing. Meaning and Self-Organisation in Visual Cognition and Thought* (221–244), Kluwer Academic Publishers, New York, 2004.

46. Rothbart, D. (ed). «General introduction», *Modeling: Gateway to the Unknown. A Work by Rom Harré* (vii–xiii), Amsterdam, Elsevier, 2004.

47. Schüler, N. «Towards a General Methodological Classification of Computer-Assisted Music Analysis», *Proceedings of the International Computer Music Conference, ICMC 2000* (27.08.–01.09.2000, Berlin), Michigan Publishing, Ann Arbor, 2000. https://researchr.org/publication/icmc-2000

48. Shepard, R. N. «On the Status of «Direct» Psychophysical Measurement», C. W. Savage (ed), *Perception and Cognition. Issues in the Foundation of Psychology* (441–490), University of Minnesota Press, Minneapolis, 1978.

49. Silins, N. «Perceptual Experience and Perceptual Justification», *The Stanford Encyclopedia of Philosophy*, 2021. https://plato.stanford.edu/archives/win2021/entries/perception-justification/

50. Steingrimsson, R. «Subjective intensity: Behavioral laws, numerical representations, and behavioral predictions in Luce’s model of global psychophysics», *Journal of Mathematical Psychology* 75, 205–217, 2016. https://doi.org/10.1016/j.jmp.2016.06.004

51. Stern, D. N. Forms of Vitality. Exploring Dynamic Experience in Psychology, Arts, Psychotherapy, and Development. Oxford University Press, Oxford, 2010.

52. Vessonen, E. «Representation in measurement», *European Journal for Philosophy of Science* 11/76, 1–23, 2021. https://doi.org/10.1007/s13194-021-00365-6

53. Wang, D. «Computational Scene Analysis», W. Duch and J. Mandziuk (eds), *Challenges for Computational Intelligence* (163–191), Springer, Berlin, 2007.

54. Wiggins, G. A. «Semantic Gap?? Schemantic Schmap!! Methodological Considerations in the Scientific Study of Music», *Proceedings of the 11th IEEE International Symposium on Multimedia*, 477–482, 2009. https://doi.org/10.1109/ISM.2009.36

55. Wiggins, G. A. «Cue Abstraction, Paradigmatic Analysis and Information Dynamics: Towards Music Analysis by Cognitive Model», *Musicæ Scientiæ*, 14/2 Special issue 2010, 307–331. https://doi.org/10.1177/10298649100140S217

56. Wiggins, G. A., Pearce, M. T. & Müllensiefen, D. «Computational Modelling of Music Cognition and Musical Creativity», *The Oxford Handbook of Computer Music* (383–420). Oxford University Press, Oxford, 2012. https://doi.org/10.1093/oxfordhb/9780199792030.013.0019

57. Wolff, R. P. Kant's Theory of Mental Activity: A Commentary on the Transcendental Analytic of the Critique of Pure Reason. Harvard University Press, Cambridge, MA, 1963. https://archive.org/details/kantstheoryofmen0000wolf

B. Technical steps to apply the TEDEA/COSM experiments

Short list of the experiments including the participants' and researcher's activities

Participants' activity:
Experiment I – TEDEA
TEDEA part 1: "intuitive" perceived musical tension detection process with slider-controllers (behavioristic and "intuitive" consciousness approach – term inspired by Clifton 1975)
Experiment II – COSM
COSM part 1: Impulse detection – "primary categorization" (term inspired by Rosenthal 2004: 224)
COSM part 2: Impulse 'content' (musical parameters) detection – "local discrimination" (term inspired by Rosenthal 2004: 224)

Researcher's activity:
Data visualization and analysis
TEDEA part 2: Data visualization and analysis
COSM part 3: Data visualization and analysis
COSM and TEDEA data comparison: visualization and analysis
COSM part 4 & TEDEA: Cross-domain data comparison
Narrativity in relation to perception and cognition
Spectral analysis, personality traits, bio-feedback data

Explanation of abbreviations that includes a short-list of the aims, activities and agents conducting them:
PP = Data recording software preparation process for TEDEA & COSM
[Researcher activity]
DR = Data recording for TEDEA & COSM
[Researcher activity]
IT = "Intuitive" perceived musical tension detection process with slider-controllers
– TEDEA [Participant activity]
PC = "Primary categorization" process: musical event Impulse (moment of change) detection
– COSM part 1 [Participant activity]
LD = "Local discrimination" process: Impulse 'content' musical parameters' salience detection
– COSM part 2 [Participant activity]
RP = Representation process
– TEDEA & COSM [Researcher activity]
CP = Calculation process
– TEDEA & COSM [Researcher activity]
CD = Cross-domain mapping/data analysis process
– TEDEA & COSM [Researcher activity]

No./ process type	**Aim**	**Tools**	**Activity description** [Researcher activity] [Participant activity]
Experiment I – TEDEA			
PP	**Data recording software preparation process for TEDEA** [Researcher activity] [Participant activity]		
PP	Preparing the TEDEA (continuous data capture technology) hardware	TEDEA box with Arduino 10 slider-controllers (built by Hans-Gunter Lock) Cables to connect: 2 stage box XLR cables (8x8), extending XLR cables	Distribute 10 slider-controllers to 10 desks for 10 participants, connect all cables to the stage box cables, connect the TEDEA box with the stage box. [Researcher activity]
PP	Preparing the TEDEA software	Max/MSP software (written Hans-Gunter Lock)	Test the sound file playback to the loudspeaker system in the room. Test the incoming signal of each slider-controller. [Researcher activity]
DR	Preparatory questionnaire	Paper, pencil	Participants answer a questionnaire before the experiment based on pre-knowledge of musical tension, style and composer style: [Participant activity] 1) What means musical tension to you? 2) What musical parameters influence the creation of musical tension (composer's view) – give a ranking. 3) What musical parameters influence the comprehension of musical tension (listener's view) – give a ranking. 4) Rank the 8 pre-given, but visually randomly presented musical parameters* in accordance to the participants opinion concerning their influence on musical tension (listener's view) 5) Likability of the music/style under observation on a four-point scale (No, Rather no, Rather yes, Yes) * (1) dynamics (loudness), (2) pitch (hight/melodic contour; register, gradual movement, (3) rhythm, (4) tempo (4), (5) harmony, (6) instrumentation/timbre, (7) texture, (8) effects (e.g. special playing techniques)
IT test	"Intuitive" perceived musical tension detection process with TEDEA slider-controllers [Participant activity]		

1.	Starting a test file: style of music and slider-controller test	Max/MSP software (written Hans-Gunter Lock)	Play a test file 1–2 minutes and check, which's slider-controllers does not work. The participants are required to test their slider-controller in relation to the music to become familiar with the style of music and the slider-controller mechanism. The data will not be saved. [Participant activity]
DR	**Start the recording of the piece to analyse in Max/MSP** [Researcher activity]		
IT	**TEDEA part 1: "Intuitive" perceived musical tension detection process with slider controllers** (behavioristic and "intuitive" consciousness approach – term inspired by Clifton 1975) [Participant activity]		
PP	Explaining constraints of the slider-controller range	Slider-controller	Explain to the participants to spare around 1/3rd or 1/4th of the slider-controller range for cases that the musical tension may rise higher than expected before [Researcher activity]
2.	Moving in real-time a slider-controller up and down - start from zero	Max/MSP software	Response to changes of the musical tension: [Participant activity] (1) If you think the tension is **increasing**, move the slider-controller button **upwards**, if the tension is **decreasing**, move the button **downwards**. (2) If you think the musical tension is **steady** (remains generally unchanged) **do not move** the slider button. (3) With the **tempo of the slider-controller button movement** you indicate the **relative speed of the change** of the musical tension as you perceive it. If you think it changes relatively fast, move the slider-controller button up or down rather abruptly, if it changes slowly, move the button rather smoothly.
DR	After-experiment questionnaire	Paper, pencil	After the experiment the participants are asked to judge (partly again): 1) the feasibility of the experiment on a three-point scale (No, Rather yes, Yes), 2) likability of the music on a four-point scale (No, Rather no, Rather yes, Yes) 3) Rank the 8 pre-given musical parameters in accordance to the participants opinion concerning their influence on musical tension (listener's view) [Participant activity]

Experiment II – COSM			
PP	**Data recording software preparation process for COSM** [Researcher activity]		
PP	Preparing the Impulse detection software	Reaper DAW	(1) Turn **timeline** invisible: Actions → Section: Main → Theme Development: Show theme tweak/configuration window Timeline Foreground Timeline Background → change both into same color (2) Turn sound file **waveform** invisible: Options → Preferences → Appearance → Peaks/Waveforms → uncheck Display peaks for media items (3) Turn left-side toolbar invisible: Drag with the mouse the border between toolbar and track waveform area to the left to hide the toolbar (4) Turn transport bar (below) invisible: View → Transport (5) Change Time unit View → Time unit for ruler → Seconds (6) Turn grid (button in toolbar) invisible.
PC	**COSM part 1: Impulse detection – "Primary categorization"** (term inspired by Rosenthal 2004: 224) [Participant activity]		
PC	**"Primary categorization" process: musical event Impulses (moments of change) detection** [Participant activity]		
1.	Listening to the beginning of a test file while detecting Impulses (moments of change).	Reaper DAW: timeline, waveform of the recording and toolbar of the program are made invisible for the experiment participant.	Listen and use M-key to mark the time points of each Impulse. The data will be erased before the real experiments. The test file is the same style as the observed piece to become familiar with the procedure.
DR	**Turn on Quicktime screen recording software, collect data** [Researcher activity] Turn on: Show mouse clicks in recording Reason for recording: 1) How often the participant changes each Impulse (marker) position? 2) Output: get to know what Impulses are more complicated to detect. 3) In case text-file data is missing or show errors, it can be extracted from the video.		

2.	Detecting Impulses (moments of change):	Reaper DAW: timeline, waveform of the recording and toolbar of the program are made invisible for the experiment participant.	Listen and use M-key to mark the timepoints of each Impulse. [Participant activity]
DR	**Turn off Quicktime screen recording software, save data** [Researcher activity]		
3.	Export marker data	Reaper DAW	View → Region/Marker Manager With right mouse click choose from the menu Export regions/markers and export it as .cvs file. Later rename this file in .txt
LD	**COSM part 2: Impulse 'content' (musical parameters) detection – "Local discrimination"** (term inspired by Rosenthal 2004: 224) [Participant activity]		
PP	**Import phase I data and the sound file for each participant separately into COSM** [Researcher activity]		
MD	**Turn on Quicktime screen recording software to collect video data** [Researcher activity] Turn on: Show mouse clicks in recording How often the participant changes Im-Cont parameters for each Impulse position? Aim: do get to know what Impulses are more complicated to detect.		
LD	**"Local discrimination" process: Impulse 'content' musical parameters' salience detection using audiovisual (AV) salience mechanism** [Participant activity]		
1.	Listening to the Impulse chunks/windows repeatedly to detect the salience of the musical (Im-Cont) parameters visualized in the 8-corner slice model (COSM) responsible for the moment of change in each Impulse The chunks/windows principle is similar to Albrecht's (2012: 17–20) Progressive Exposure Method (PEM).	COSM (Max/MSP) visual interface	Start listening to every Impulse (Im) separately using the PLAY button. The Im you choose sounds as 14 sec (+/– 7 sec) chunks/windows. Which Im-Cont parameters are most and less salient? Mark the parameter chosen for the fitting corner of the model as follows: click the Im-Cont parameter in the left side menu, then click the appropriate corner in the 8-corner slice – the parameter will be now visualized as used and inactive in the menu, and is active in the corner chosen. If necessary repeat this listening activity around each Im as much as necessary to become sure of the decision. The choices can be corrected using the Correction button. All choices will be erased and the whole process for this Im must be repeated from the beginning. If the choice is firm, press the Ready button. The next Im will be ready to be played automatically. To orientate in the order of the Impulses its order number appears in

			the Im# number box in the right corner of the model. While defining the visual salience of the eight corner shape model please depart from the following gestalt principles: corner 8 and 7 are visually most salient, corners 6 and 5 are visually less salient, corner 4–1 are visually the least salient (in decreasing order). NB! You need not to insert all parameters for every Impulse. It is sufficient to choose those parameters that you are really able to detect and decide to be salient for you. If the change in an Impulse (separately from the musical parameters) is especially strong you can choose as an additional aspect also the contrast (gray circle in the black box). The contrast button fits only to the circle in the middle of the model.
MD	**Turn off Quicktime screen recording software, save metadata** [Researcher activity]		
RP, CP	**Data visualization and analysis**		
RP, CP	**TEDEA part 2: Data visualization and analysis** [Researcher activity]		
RP, CP	**Output data preparation and visualization: TEDEA slider-controller data** [Researcher activity]		
1.	Visualizing the raw output data (each participant separately).	Microsoft MS Excel/Open Office OO Spreadsheet	Visualize the raw data in XY Scatter diagram
1.1.	Visualizing the summed output data (each participant separately).	MS Excel/OO Spreadsheet	Visualize the raw data in XY Scatter diagram
2.	Visualizing the averaged output data (each participant separately).	MS Excel/OO Spreadsheet	Visualize the raw data in XY Scatter diagram
2.1	Visualizing the normalized averaged output data (each participant separately).	MS Excel/OO Spreadsheet	Visualize the raw data in XY Scatter diagram
RP, CP	**COSM part 3: Data visualization and analysis** [Researcher activity]		

1.	Automatic process of data recording and exporting! COSM data export in the fitting format for visualizing and further analysis in 2D diagram form (Excel or other data analysis programs).	COSM (Max/MSP) visual interface and recording function	COSM data export for each participant separately.
RP, **CP**	**Output data preparation and visualization: COSM part 1 data** [Researcher activity]		
2.	Automatic process of data analysis! DBSCAN application for part 1 data (Impulse detection)	Cluster finding: data mining programs (Phyton, Math-Lab etc) DBSCAN algorithm: Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise (Ester *et al* 1996)	Choose the density and cluster size parameters in accordance to the object of analysis, number of participants (significance) and input data structure, e.g. N=14, 10 participants are significant Cluster size 2000 ms = 2 seconds Trigger cluster building process
2.1	DBSCAN output data manual cleaning	Txt/MS Excel/OO Spreadsheet	Check for errors in the DBSCAN output data
RP, **CP**	**Output data preparation and visualization: COSM part 2 data** [Researcher activity]		
3.	Summing up the values for each Im-Cont parameter per DBSCAN cluster (each participant separately).	MS Excel/OO Spreadsheet	Sum up the data of each Im-Cont parameter per DBSCAN cluster and attach it to the correct time points of COSM part 1 data per DBSCAN cluster. Use a defined color-code for each musical parameter.
3.1	Averaging of the values of Im-Cont parameter per DBSCAN cluster (each participant separately).	MS Excel/OO Spreadsheet	Average all Im-Cont parameters per DBSCAN cluster: sum up all Im-Cont parameters and divide them with the total number of the slices per DBSCAN cluster. Use a defined color-code for each musical parameter.
RP, **CP**	**Output data representation: visualization and analysis of COSM part 1 and 2 data** [Researcher activity]		
4.	Visualizing COSM part 1 and 2 data together	MS Excel/OO Spreadsheet	Visualize the data in a XY Scatter diagram. Use a defined color-code for each musical parameter.

4.1	Detecting increasing/decreasing trends	MS Excel/OO Spreadsheet	Analyze the data in a XY Scatter diagram – mark increasing/decreasing trends. Use a defined color-code for each musical parameter.
RP, **CP**	**COSM and TEDEA data comparison: visualization and analysis** [Researcher activity]		
1.	Application of the DBSCAN generated clusters to TEDEA slider controller data	Max/MSP software (written by Hans-Gunter Lock)	Produce a text file of the normalized TEDEA data of all participants and of chosen samples to detect the DBSCAN cluster time-points. The Max/MSP software generates the trends for the tension graphs: starting → ending maximums and minimums. These can be compared to the starting → ending maximums and minimums of the parameter data of COSM part 2.
2.	Visualization and analysis of TEDEA and COSM data based on DBSCAN clusters	MS Excel/OO Spreadsheet	Visualize and analyze the TEDEA and COSM data based on DBSCAN clusters. Discriminate different samples, primary and secondary parameters etc.
CD	**COSM part 4 & TEDEA: Cross-domain data comparison** [Researcher activity]		
RP, **CP**	**Narrativity in relation to perception and cognition**		
1.	Applying Byron Almén's (2003: 12–13) five aspects of a musical narrative (Lock 2020: 319) together with Kramer’s (1991) definition of narrativity (dynamic principle and teleological, impulse): arriving at the cross-domain definition of ‘Perceived structural musical tension’ (PSMT) (see 3.5). Interpretation of the data based on Gruhn’s (1992) suggestions for quantifying and analysis of listening stories (Lock 2020).	MS Excel/OO Spreadsheet Text file	Treat the degree of regularity as a function of narrativity claiming that more clear regularity means a more clear comprehensibility, therefore more clear narrativity (Lock 2020: 315). This is derived from data mining principles, where such clear regularities gain special attention. Narrativity is interpreted in the combined data derived from the COSM first step (salience of impulses) based on cumulative patterns (clusters). Narrativity can be also interpreted in the slider-controller results (musical tension) and the COSM second step (salience of musical parameters). Here the more usual principle of a teleological narrative, up/down trends, in the meaning of Aristotle’s drama theory (Freytag 1900/1863) are the base for interpretation. Results see Lock (2020: 315, 340–342).

2.	Interpretation of the data based on the perception/ cognition/ narrativity combining ‘Perceived structural musical tension’ (PSMT) definition. Take into account Brower's (2000) structural aspects of cognitive narrativity.	MS Excel/OO Spreadsheet Text file	Analyze, visualize and describe the data based on PSMT (see this dissertation, 3.5).
3.	Interpretation of the data based on Imberty’s (2000) suggestion to apply Stern’s (2010) dynamic vitality forms, supported by Mailman's (2010) temporal dynamic form approach	MS Excel/OO Spreadsheet Text file	Analyze, visualize and describe dynamic vitality forms in the context of temporal dynamic form.
4.	Interpretation of the data based on Conceptual Blending Theory (CBT, Antovic 2018, 2022) and its audiovisual approaches (Athanasopoulos & Antović 2018). Take into account Brower’s (2000) CBT-related aspects of cognitive narrativity.	MS Excel/OO Spreadsheet Text file	Analyze and visualize the data in the context of CBT, take into account similar processes of audiovisual cognition.
RP, **CP**	**Spectral analysis, personality traits, bio-feedback data**		
5.	Cross-domain comparison with Spectral analysis data of the stimulus (recorded music).	Cluster finding: data mining programs (Phyton, Math-Lab etc)	Normalizing, averaging, correlations, trend/regression etc.
6.	Cross-domain comparison with Personality traits data of the experiment participants.	Questionnaire	Apply quantitative and qualitative data retrieval and analysis methods from social sciences and psychology
7.	Cross-domain comparison with Bio-feedback data with the same experiment participants.	Bio-feedback measurement	Normalizing, averaging, correlations, trend/regression etc.

C. Methodological questions during the development of the dissertation

C.1 Questions after publication I (Lock 2010a)

Research assumptions, aims and questions:

- musical tension is both a structural and a perception/psychology category: this leads to a combined methodological approach
- there is no widely accepted analysis method for contemporary post-tonal musical yet
- offers a combined method (music theoretical and psychological/empirical approach) that describes and integrates chosen structural aspects that cause musical tension
- investigates the relation of main culmination points with the musical material and its form

Q1: What is the role of the culminations in the development of the musical form?
Q2: How clearly the empirical curves correlate* among each other?
*However, correlation is not measured as statistical results, it appears visually by comparing curves.
Q3: What is the relation of the analysis of the form (score- and listening analysis based on salient features) to the empirical curves and its average curve?
Hypothesis: Listeners perceive the development of the musical tension both generally and in relation to important points of the musical form as developing in correlation, this means that the tensional in- and decreasing are relatively objective phenomena and, based on them, it is possible to develop a so called tension design analysis method.
Object of analysis:

- the final section of the contemporary post-tonal work *Oxymoron* for large ensemble by Erkki-Sven Tüür (b 1959)
- the work has no clear (classical form-like) boundaries, although it can be analyzed based on the three types of musical time (continuous, slow, quick) characteristic für Tüür's music either as sonata cycle movements or detecting sonata allegro formal section features based on their rhetorical and rotational structure (Kotta 2011).
- characteristic for Tüür's music since 2002/2003: transitions between different states/textures/temporal organization of the musical material are dominating instead of contrasting them like in earlier music by Tüür

(A) results in relation to musical form:

- Q1: Formal boundaries appear rather in maxima and minima points of musical tension

(B) perceived structural musical tension is a complex/compound phenomenon

- slider-controller listening experiment: tension increasing, decreasing, remaining

Data representation/visualization method: curves are normalized, smoothed, averaged, reduced, 10 seconds static (overlapping) time windows

- Q2: Single curves are rather individual, but develop in correlation at important culmination points.

(C) salient features describing musical form based on chosen aspects related to musical tension

- speculative music theoretical analysis (score and listening; intuitive listening approach)
- complex salient features: “impulse”, “culmination” (local) and “contrast”, “contrast-

culmination” (global level/dimension)
- Q3: Single curves are rather individual, but develop in correlation at important culmination points and at average with important points of the salient-features-based musical form.

Answer to the hypothesis: The hypothesis is confirmed partly based on these coinciding areas and points. The perceived structural musical tension and the salient-features-based musical form develop partly in correlation. The averaged curve is technically objective, but due to the low number of participants (N=7) the result is quantitatively not generalizable, further aim to develop a tension design (TD) method.

Further developing aspects:
- **Impulse** as general cognitive feature remains further relevant
- it is a model (although without using that term), because it is consistent (a whole) juxtaposing two contrasting phenomena related to its temporal characteristics (short singular moment *vs* longer lasting teleological development) and levels/dimensions

Model explanation based on Lock (2017a):
- Function category (type): II explanation, III exploratory, IVb music analysis model
- Method category: cognitive

Data representation/visualization method: arbitrarily scaled based on moment *vs* development and local *vs* global (increasing arbitrary values 10, 20, 30, 35/40)
- explanation: explains the relation of musical form to musical tension
- exploratory: tries to understand a phenomena from a yet unexplored angle – perceived structural musical tension and its related salient features (moment *vs* development, local *vs* global)
- music analysis model: analyses particular music based on musical relevant, but not exclusively musical aspects (salient features)

Further research questions:
Detecting the general (and hypothesized) correlation of perceived structural musical tension and musical form and admitting that single curves are often individual, but still correlate in important points of the musical form, causes the following follow-up questions concerning further developing the topic:

1. What kind of more detailed aspects cause this correlational development?
1.1 How the salient features function in detail – beyond speculative subjective analysis (score and recording, intuitive listening approach)?
1.2 How concretely the salient features are definable to develop an automatized analysis (score or recording) or an empirical method (that is still an intuitive listening approach)?
1.3 What are the more concrete aspects according to which the listening experiment participants decided to show an in- or decrease, or remaining state of perceived structural musical tension?
1.4 What are the reasons for both the individual differences and the correlating results in important culmination areas/points of the musical form?
1.5 How meaningful is it to arrive through a speculative and subjective analysis (intuitive listening approach) that assumes the main culmination area to be rather in the second part of a work to a pre-suggested (therefore biased) result? See Aristoteles’ drama theory in Freytag 1900/1863. Compare the problem with the Schenkerian *Ursatz*/*Urlinie* pre-known result.
1.6 How to approach and solve **representation/visualization issues** for the analysis of more detailed aspects?

2. How effectively the salient-features method worked so far manually and will work especially analyzing full-length works?
2.1 Does an increased amount of input-data (score and recording) enable a scientifically sufficient analysis in detail and analytical precision?
2.2 Is an automatization of the salient-features method possible to develop?
2.3 How to approach and solve **representation/visualization issues** connected to an increased amount of input-data, its details and precision in analysis and automatization?

3. How effectively will the data treatment of slider-controller curves be, especially analyzing full-length works (mass-data)?
3.1 Does an increased amount of output-data (curves) enable mathematically-statistically meaningful data-analysis, including the meaningfulness problem of different statistical methods while comparing single curves?
3.2 Is qualitative data analysis possible and meaningful?
3.3 How meaningful is it to arrive through a reduction of maximums to a simple two-phase main-culmination graph (including a global in- and decrease phase) for the whole piece? Compare the problem with the Schenkerian *Ursatz*/*Urlinie* pre-known result.
3.4 How to approach and solve **representation/visualization issues** for large data-sets gained from full-length works (quantitative method) as well as for qualitative data?

4. How to approach technical and meaning problems related to the comparison of results (data) from different analysis methods especially analyzing full-length works and large output-data-sets?
4.1 On what level output-data of different methods need to be compatible (comparable) to each other to enable meaningful results?
4.2 What does the 4.1 question mean for the gaining of input-data for both methods?
4.3 What kind of cross-domain/meta-domain interpretations may enhance the understanding of the meaning of such compared data?
4.4 How to approach and solve **representation/visualization issues** connected to the analysis of such compared data?

C.2 Questions after publication Ia (Lock & Kotta 2012)

Hypothesis

The intensity of the perceived musical tension is proportional to the structural (or hierarchical) significance of the corresponding musical event (Lock & Kotta 2012: 612).

Similar methods

The proposed data analysis (tension curves' reduction) algorithm is, in principle, similar to the melody contour reduction method of Morris (1993) (suggested by Joshua Mailman during ICMPC/ESCOM 2012), see also overview of contour reduction algorithms by Bor (2009). Significant data points are reduced out or remain on consecutive levels: foreground, middleground, background like in Schenkerian analysis, possible mathematical methods of time series data treatment see Fu (2011).

Automatization attempt

In 2012 Toby Gifford (Queensland Conservatorium, Griffith University, Australia) started to write a software (Java with Max/MSP interface), but its development took several years and it remained to have both software bugs and methodological questions – either high- or low-points where favored by the reduction algorithm causing the *why*-question and leaving the analytical explanations concerning the choice of the reduction level to compare curves open. Therefore, at this moment, this software is abandoned.

Further outlook

The moving-window-principle shows similar features as the later applied DBSCAN (Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, Ester *et al* 1996) automatized algorithm: collecting together close-by data-points into clusters.

D. Some overall results of the COSM experiment

In this Appendix I present some overall results of the COSM experiment **part 2** on the salience of musical parameters that did not fit into publication IV (Lock 2020). These results accompany the statistical overview of the DBSCAN clusters presented in 4.6 (see Figures 5 and 6) and are based on the more outstanding events/Impulses/responses. In Figure 7 (below) the salience of the musical parameters is summed up per cluster and visualized in chronological order. On the vertical axis (summed amount) the clusters 7, 26 and 38 stand out showing the highest summation of musical parameters. This coincides with the event/Impulse/responses count statistics shown in Figure 5 and the more detailed cluster/common second length statistics in Figure 6.

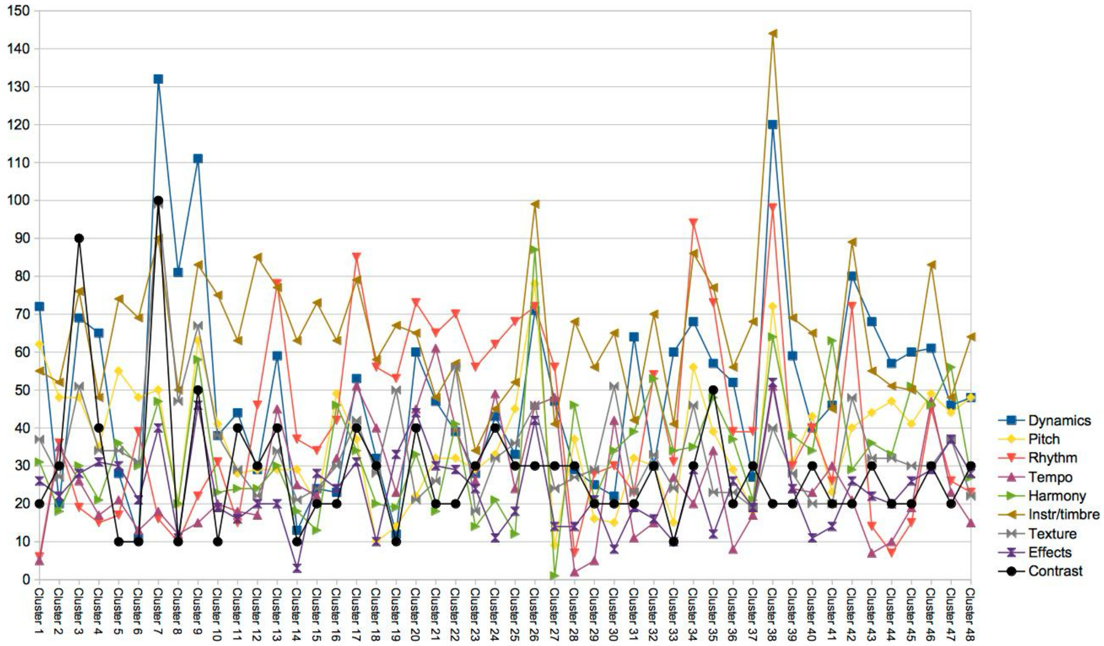


Figure 7. Summed COSM parameters' salience statistics. Visualized are all eight parameters (and contrast) per clusters (48) of Tüür's complete Fourth Symphony/Percussion Concerto (N=14). Primary parameters PP: pitch, rhythm, harmony; Secondary parameters SP: dynamics, tempo, instrumentation/timbre, texture, effects.

The information density of Figure 7 is high. One can discover e.g. extreme changes in summed values between neighboring clusters: for instance the higher salience values of all parameters from cluster 26 is dropping in cluster 27 – e.g. for harmony to a significant minimum. This can be interpreted in relation to the rhetorical/rotational form analysis by Kotta (2011) (see Figure 10 in publication IV) to become meaningful: cluster 26 is the last cluster at the end of the Scherzo “movement” B (being the “end” of the “first half” of the symphony). Cluster 27 is the only cluster detected in the Cadence of the percussion soloist where harmony plays now role. Of course, in cluster 26 the summed salience values for all musical parameters has significantly increased in comparison with the earlier cluster 25 summed values. Furthermore within cluster 26 the trend of all slider-controller averaged results is increasing, too (see Figure 10 in publication IV).

The following Figures 8 and 9 make an attempt to detangle the information density of Figure 7 analyzing the distribution of the summed parameters in accordance to their appearance based on a partitioning of the vertical summation value axis. It seemed reasonable to partition the vertical axis in three levels: basic salience level 0–50, higher salience level 50–100 and highest salience level 100–150 points. This revealed two categories of musical parameters being more or less salient over the course of the music. Five parameters’ (PP pitch, harmony; SP tempo, texture, effects) salience has been detected to be rather low, staying mainly in the basic level 0–50. These can be called Basic salience level (BSL) parameters or short-hand Basic level (BL) parameters. Three parameters (PP rhythm; SP dynamics, instr/timbre) show more extreme changes (meandering) between all three levels. These can be called All salience levels (ASL) parameters or short-hand All level (AL) parameters. However, also the parameters texture (SP), pitch (PP) and harmony (PP) appear sometimes at the higher salience level, but with a significantly lesser degree of salience.

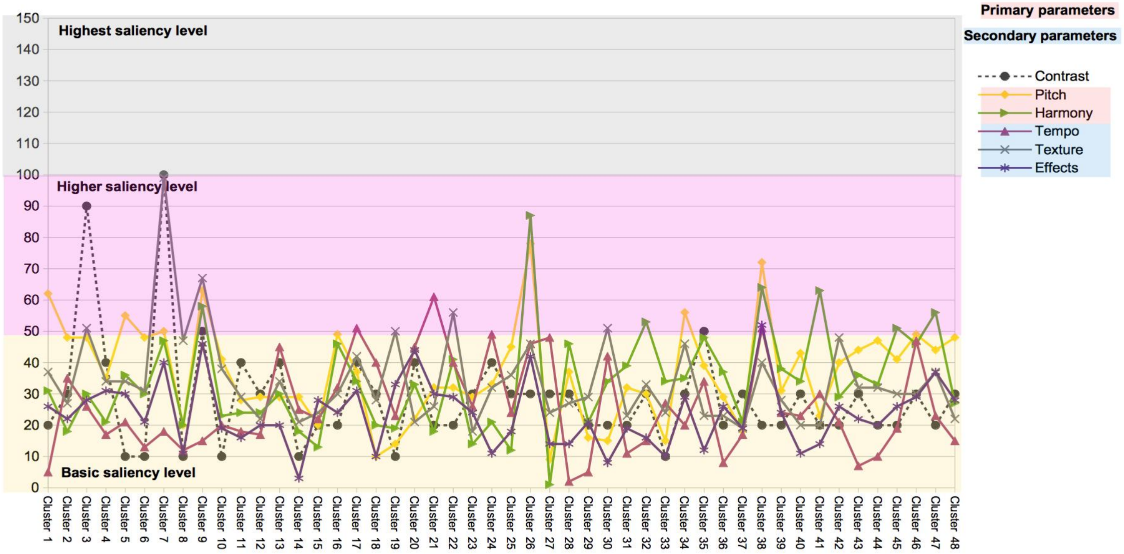


Figure 8. Summed COSM parameters' salience results. Basic salience level or Basic level (BL) parameters per clusters (48) of Tüür's Fourth Symphony/Percussion Concerto (N=14). Primary parameters PP: pitch, rhythm, harmony; Secondary parameters SP: dynamics, tempo, instrumentation/timbre, texture, effects. This figure was first presented in Lock (2019, 2023a).

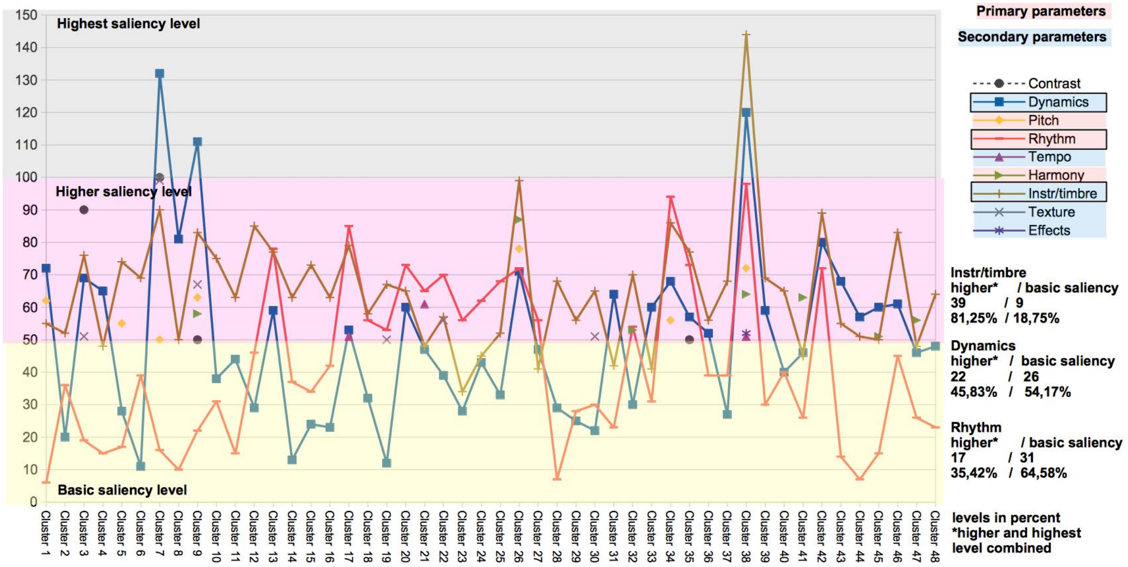


Figure 9. Summed COSM parameters' salience results. All salience levels or All levels (AL) parameters per clusters (48) of Tüür's 4th Symphony/Percussion Concerto (N=14). This figure was first presented in Lock (2019, 2023a).