Proceedings of the
12th Art of Record Production Conference
Mono: Stereo: Multi
Photo A. From the general meeting Saturday 2nd December 2017 when Simon Zagorski-Thomas announced that he would step down as Director of the Art of Record Production Conference. (Photo credit: Eleonor Gislason-Ferrari)
Editor’s Preface

The 12th Art of Record Production Conference (ARP2017) - Mono: Stereo: Multi was held at the Royal College of Music, Stockholm, Sweden, December 1 - 3, 2017. ARP2017 was organized in co-operation between the Association for the Study of the Art of Record Production and the Royal College of Music.

But why Mono: Stereo: Multi? Drawing on the both literal and metaphorical connections with the history of audio formats, the 2017 Art of Record Production conference explored a range of ways in which singularity, dichotomy and multiplicity can be contrasted. Whether it was the monocultural versus the multi-cultural, the mono-media versus the multi-media, a mono-disciplinary versus a multi-disciplinary approach or the mono-phonic versus the multi-phonic, we had long and short format paper presentations, poster presentations and recorded music playback presentations in the following contexts:

A: Research vs. Practical Education. How does the singular versus the multi-faceted relate to the various dichotomies faced in the juxtaposition of research and practical education? What are the differences between singular and multi-disciplinary approaches? How broad or narrow should the scope be in our approach to education or in the development of new knowledge? What does it mean to be a specialist in the rapidly changing world of recorded music and how broad a range of skills and knowledge can an individual successfully assimilate?

B: Production vs. Consumption. In what ways are production and consumption systems related to our engagement with and immersion in recorded music? How are distribution systems that situate control of the mix process in the hands of consumers affecting the idea of authorship and ownership in production and consumption of recorded music? How are new and varied systems of production and consumption affecting our understanding and control of production and/or consumption?

C: Aesthetics vs. Technology. Whether it be Schaffer’s concept of reduced listening or the complexities of multi-modal perception, there is a complex interaction between the affordances that technological developments offer and the way that aesthetics and cultural values drive their usage. How are these socio-cultural and technological changes related to the sounds of recorded music and the way we interpret them?

D: Traditions vs. Innovations. How have histories and developments in the practices of recorded music making related to these broad interpretations of the terms mono, stereo and multi? This stream ranges from the multiple formats of recorded media to the variety of musical and technical practices and from ways of thinking about sound to the nature of our academic disci-
pline. When and how do technologies become old and what differentiates a smooth progression from a discrete innovation?

E: Centre vs. Peripheral. How do notions such as ‘mainstream’ and ‘alternative’ relate to recorded music and, indeed, to our academic field? This might be related to geography, economics or various forms of identity, and might explore the idea of the ‘normal versus the ‘other’ or might be a more literal interpretation of the concept of centrality. Indeed, in spatial terms, why do we need the centre to be in front of us?

At the conference, we had 107 registered participants from 14 countries. We had 64 oral presentations and of the presentations, 20 are documented as papers in these proceedings. All submitted abstracts are also published (Appendix 1). We also had five keynote speakers: Bernard Löh: From Mono Music to the Future; Göran Folkestad: The Ecology of Music Production; Bill Brunson: Space for Imagination: The Sound Dome at KMH; Linda Portnoff & Thomas Florén: The Music Industry in Scandinavia & Swedish Music Export and a celebrity interview with Benny Andersson, interviewed by Göran Folkestad. Furthermore we had two panels: Producers on Mono – Stereo – Multi and Methodologies in Record Production Research-Panell/Workshop. We rounded off the first conference day with a special event at the Abba Museum, and also had a pre-celebration of Saint Lucy's Day by a group of young choral singers from Adolf Fredrik's Music School.

The conference committee organized the selection of the 64 Oral presentations at the conference and the 20 papers in this publication, with help from 26 expert reviewers from three continents. The editorial team of this publication, Jan-Olof Gullö, Per-Henrik Holgersson, Bo Westman, Katia Isakoff & Shara Rambarran, would like to thank the all reviewers for reviewing all the abstracts and papers. We give special thanks to those who helped us with multiple reviews. Both the editors and the authors highly appreciate the input and dedication of all our reviewers. Many thanks.

Jan-Olof Gullö, May, 2019 on behalf of the Editorial Team
Table of contents

**Tuomas Auvinen**: Differences and Similarities in the Role and Creative Agency of Producers in Pop, Rock and Classical...................... 1

**Andrew Bourbon**: Hit Hard or Go Home - An Exploration of Distortion on the Perceived Impact Of Sound On A Mix ................. 19

**Toivo Burlin**: The creation (and re-creation) of virtual spatiality: Surround sound mixing in King Crimson’s back catalog box sets..... 37

**Carlos Andrés Caballero Parra**: The recording and sound aesthetics of tropical music in Colombia.............................................. 55

**Mark Durham**: Multi-Channel Sound Design: Instruments for 360-Degree Composition.......................................................... 71

**Michail Exarchos (aka Stereo Mike) & Glenn Skinner**: Bass | The Wider Frontier: Low-end Stereo Placement for Headphone Listening ........................................................................................................... 87

**Steve Gamble**: Listening to virtual space in recorded popular music ........................................................................................................ 105

**Lachlan Goold & Phil Graham**: The Uncertain Future of the Large-Format Recording Studio............................................................... 119

**Tor Halmrast**: Sam Phillips’ Slap Back Echo; Luckily in Mono ... 137

**Kirsten Hermes**: Enhancing creativity through automatic mixing research: testing spectral clarity predictors in the mix process ....... 155

**Marco Antonio Juan de Dios Cuartas**: The incursion of stereo into Spanish popular music: the English influence in the definition of a local sound through the professional exchanges between Madrid, Barcelona and London......................................................................... 171

**Martin K. Koszolko**: The Tactile Evolution — Electronic Music Production and Affordances of iOS Apps ........................................... 187

**Paul Novotny**: Creating an Immersive Fold-Out – Look Ahead..... 205
Daniel Pratt & Andrew Bourbon: Life In Between Phase: Understanding and Manipulating Microphone Relationships With Visualisation Tools ................................................................. 225


Méi-Ra St-Laurent: "It's kind of in the middle": The ‘Mid-Fi’ Aesthetic: Toward a New Designation of Black Metal Aesthetic of Recording. The Case of the Québec Black Metal Scene of recording ................................................................. 267

Paul Thompson & Phil Harding: A ‘Service’ Model of Creativity in Commercial Pop Music at P&E Studios in the 1990s ................................................................. 287

Elizabeth Navarra Varnado: The sampling aesthetic of Bon Iver’s “33 ‘God’” ........................................................................................................................................... 303

Alan Williams: Surround Sound Auteurs and the Fragmenting of Genre ........................................................................................................................................... 319

Jason Woolley: The cultural politics of using technology to support the aesthetic in jazz record production ........................................................................................... 329

Appendix: Abstracts - Parallel sessions The 12th Art of Record Production Conference Mono: Stereo: Multi (ARP2017) Royal College of Music, Stockholm 1-3 December 2017 ................................................................. 347
Tuomas Auvinen: Differences and Similarities in the Role and Creative Agency of Producers in Pop, Rock and Classical

Abstract: Differences and similarities in the creative agency of the producer in the production process of urban pop music produced in a home studio, rock music produced in a conventional studio facility and classical concert hall music produced in a concert hall setting is explore in this paper. Starting from the premise of record production being a collaborative effort, I approach agency as the capacity to make and effect decisions within a structure or even to alter it to some extent, and creativity as contributing to the domain of existing works through exercising aesthetic decision-making. Based on these understandings of agency and creativity, I will examine how different cultures in different production settings and different studios conceived as cultural spaces affect the construction of the producer’s agency within creative communities in the production process. Furthermore, I will discuss how differences in understandings of the ontology of the music contribute to the level of creativity, i.e. the contribution to the domain of existing works, that a producer agent can possess. I base my presentation on extensive ethnographic fieldwork of three case studies on production processes, which took place in the course of 2015-2017.

Introduction

The producer as the central figure of record production has often been the object of scholarly studies on the production process. Studies on single producers (e.g. Warner: 2003; Martin: 2014; Hennion: 1989) and more comprehensive historical or canonical accounts (e.g. Moorefield: 2005) lack direct comparisons between producers of different genres, although exceptions do exist (e.g. Burgess: 2013). In this article, my aim is to compare and contrast the roles and agencies of producers in pop, rock and classical by answering the following research questions:

[1] What differences in the producer’s role and agency does the comparison of ethnographic case studies of a pop production, a rock production and a classical production reveal?
What similarities can be found between the role and agency of the producer in a pop production, a rock production and a classical production and what do they infer about the role and agency of the producer in general?

What can the comparison of differences and similarities of the producer’s role and agency between the different case studies yield about the underlying philosophical underpinnings of different production settings?

Based on issues arising from my research material, the questions are framed though the conceptual binaries of production/consumption (e.g. Théberge: 1997), innovation/tradition (e.g. Moorefield: 2005) and technology/aesthetics (e.g. Warner: 2003; Martin: 2014) which frequently occur in the research field.

Here, I share some results of my forthcoming PhD dissertation, titled “Producer as Creative Agent: Studio Production, Technology and Cultural in Three Case Studies”. Next, I will shortly describe my cases. Then, I will discuss my methodology, my materials and explain my theoretical concepts. I will then discuss relevant earlier research. Thereafter, I will explore similarities and differences in giving feedback, editing, formation of authority and the influences of different recording spaces on agency. I will end this article by some concluding thoughts.

Description of Case Studies

The first case study I conducted was on the young Finnish pop producer Mikke Vepsäläinen (b. 1992) and his work on the song ‘Kunhan muut ei tiedä’ (Eng. As Long as Others Don’t Know) (Ida Paul: 2016) with the singer Ida Paul in a home studio. My second case study was on the classical producer Seppo Siirala (b. 1952) and his production work with the orchestra Tapiola Sinfonietta as they recorded the Estonian composer Erkki-Sven Tüür’s Symphony No. 8 (Tüür: 2010) at the Tapiola-hall in Espoo, Finland. My third and final case study was on the producer Jonas Olsson (b. 1981) and his work on the production of the songs ‘Can’t Hold Us’ (Blind Channel: 2017a) and ‘Alone Against All’ (Blind Channel: 2017b) with the rock band Blind Channel at Olsson’s studio InkFish in Vallila, Finland. The cases were conducted during the course of two years starting in the mid-spring of 2015 and ending in the early spring of 2017. I have already published two articles on my first case study on the producer Mikke Vepsäläinen (Auvinen: 2016; Auvinen: 2017) to which I refer in this article. I have not yet published anything on the other cases and on their part, I will be referring to the original research material.
Methodology and Research Materials

This study is rooted in the traditions of ethnomusicology and cultural musicology. Here, I have applied ethnographic methodologies such as interviews (Blind Channel: 2016; Mäemets: 2016; Olsson: 2017a; Siirala: 2015; Siirala: 2015c; Vepsäläinen: 2016), field observations written in field diaries and photos (Photo 1–4) during studio production. Furthermore, I have conducted music analysis on music (Blind Channel: 2017; Ida Paul: 2016; Tüür: 2010) under production to the degree in which they have illustrated the similarities and differences of the producers. In my analysis, I refer to my field diary with the letters “FD” and a date in the European form “(date.month.year)”. For example, the reference “(FD 1.1.2017)” would mean January the 1st in the year 2017.

In my analysis, I have applied the principles of data triangulation (e.g. Bennett: 2011) to establish a multi-faceted understanding of the object of research, as my material includes different kinds of materials. In the interpretation of my material I have applied the principles of the hermeneutic circle (e.g. Rice: 2008, p. 58) and, to some extent, interpretative phenomenological analysis (e.g. Martin: 2014). I have translated the interviews into English from the original Finnish and aimed at maintaining the original meaning of what the interviewees have said.

One could also ask how comparable these producers are, as they were all very different not only in terms of the musical style but also in terms of age and experience. Vepsäläinen could be described as a member of “the new generation” (Auvinen: 2016; 2017) and Siirala again retired in the beginning of the year 2016 continuing his work on a case-by-case basis. On the other hand, at least in Finland, a “young classical record producer” is an oxymoron in itself, as there only exist three full-time professional classical producers in the country; the numbers of options for possible participants were very few. Also, selecting case producers was not exactly hard. Very few responded with a willingness to take part and most producers I approached did not reply (FD 11.4.2014; 2.10.2016; 10.10.2016).

As I have explained earlier (Auvinen: 2016; Auvinen: 2017), I knew Vepsäläinen from before, which might have been the reason I was allowed to study his work in the first place; it seems to be difficult to find creators who would allow their work to be studies in an ethnographic manner (Bennett: 2011). Siirala was formerly unknown to me; he was recommended by another producer who refused to take part in my study (FD 11.4.2014). Olsson was somewhat familiar to me as I had interviewed him for my master’s thesis in 2012. The artists in each case study were unknown to me.

My position as a researcher could be described as somewhere between emic and etic in ethnographic terms. I did some production work in my early 20’s so I had an idea on of the process. Nevertheless, I have never worked as a full-time producer and nine years have passed since I last acted as a pro-
producer. Therefore, I have been able to remain a more etic researcher. All participants agreed to take part with their own identities. I have offered all the participants the opportunity to read through this article before publishing. This way I have made sure that my research will not reflect negatively on my research subjects.

**Theoretical Approach**

Here, the concept of *creative agency* is central and consists of two separate concepts: *creativity* and *agency*. I use them as operative concepts through which I interpret my research materials through the hermeneutic process. Although the two are often used as an entity in the context of creative processes like record (music) production, I find some degree of separate definition necessary to grasp their essence.

The American social scientist Timothy D. Taylor (2001, p. 35) defines agency as “an individual actor’s or collective capacity to move within a structure, even alter it to some extent”. This definition works well for collective processes like that of record production. Even more simply, agency could be defined as the “ability to make choices” (Taylor, p. 2017). Furthermore, agency refers “…not to the intentions people have in doing things but to their capability of doing those things in the first place” as the British sociologist Anthony Giddens (1984, p. 9) puts it. This would separate agency from role which refers more to how the producer ends up acting.

For creativity, I draw from the theories of the Hungarian psychologist Mihaly Csikszentmihalyi (1997, p. 28). For him, “Creativity occurs when a person, using the symbols of a given domain such as music [-] has a new idea or sees a new pattern, and when this novelty is selected by the appropriate field for inclusion in to the relevant domain”. This approach has been widely used in the field of the study of the art of record production. Examples include work by Philip McIntyre (2008) and Robert W. Taylor (2017). Similarly, the British sociologist Jason Toynbee (2000: 35), who refers to (popular) music makers as “creators [original italics], that is agents who make musical differences in the form of texts, performances and sounds”. He (ibid.) adds that this understanding of creativity includes “all stages of music-making from ‘writing’ through ‘performance’ to ‘production’”. These understandings of agency and creativity formulate a good conceptual base here and create a solid connection to the relevant research field.

**Earlier Research**

Despite (or perhaps, because of) the seemingly obvious differences of the producer’s role between classical music and genres of popular music, there exists very little directly comparative ethnographic research. To the best of my knowledge, the field of the study on the art of record production lacks a
comparative study, in which the role and agency of producers in EDM-based pop-music, rock music and classical music are compared through ethnographic methods. The lack of such research might stem from the idea that the prevailing differences, which result from tradition and the differences in the ontology of the work, are too obvious and an inquiry on the topic would simply be dwelling on facts that everyone already knows. However, similarities between the roles of classical producers and producers of popular music are perhaps not as self-evident and to discuss similarities requires exploring differences. Perhaps the self-evident surface differences are not as clear when they are properly examined and rigorously scrutinized from multiple perspectives.

Comparisons between producers in different genres so far limit to shorter accounts based on for instance inquiries on music criticism. The British musicologist Simon Frith offers one such perspective:

In this respect rock record producers are seen as both more significant for rock as an art form than producers in jazz, folk or classical music, but less important for rock as a cultural project than producers in pop or dance music. The producer was both obdurately present in the music and readily ignored in the way that music was discussed. (Frith, 2012, p. 221)

Frith’s statement can perhaps be seen as a summary of the reception of the differences in the producer’s role between different genres. It does not, however, necessarily say much about what actually happens in the studio process.

The shortage of direct comparisons leaves the option of comparing individual studies of popular music producers, which comprises the vast majority, and classical producers. Earlier research on the producer of popular music has emphasized her/his creative and artistic agency and the ways in which developments in music technology contribute to the producer’s authority making the producer essentially a creative agent from the 1950’s onwards (e.g. Warner: 2003; Horning: 2013, p. 204; Moorefield: 2005). This was obviously a part of a larger transformation of popular music production from craft to art (Heinonen: 2015, p. 34). Scholarly accounts on the creative process of classical record production are scarce, which might stem from the understanding summarized by the American musicologist Arved Ashby:

… recording has had less an aesthetic influence on classical-musical practices than an ontological effect. In other words, it has helped shape and define the sort of thing that music is. (Ashby, 2010, p. 22)

Furthermore, as Ashby (p. 226) explains, a classical record has often been considered “just a picture of the score taken by somebody”. Similarly, the Australian musicologist Colin Symes argues that while the body of research
dealing with “the ways in which the phonograph has transformed the conditions of listening to [classical] music

...the nature of recorded sound and its underlying discourses have not been subjected to the same degree of analysis. The neglect is particularly pronounced in the area of classical music, which has been insulated from those developments in cultural studies that have shed light on the way recording has transformed the nature of popular music. (Symes, 2004, p. 60)

According to Symes, the reason behind this neglect lies in the cultural ethos embedded in the analysis of classical music, which has avoided "contextual questions relating to music and the technology in its reproduction" (ibid.). Consequently, scholarly writing on classical record production has concentrated more on the philosophical considerations on the concept of the work (e.g. Benjamin: 1936) than on the creative capacities on the agents involved in the process. Exceptions, especially when it comes to the recordings of Glenn Could, do exist (e.g. Mantere: 2006).

Earlier accounts on the classical producer have often emphasized the producer’s mediating capacity of the "relationship between the score, the performing artists and the processes and technologies of recording" (Blake: 2012, p. 195) and the conservative attitudes towards using music production technologies to enhance the final product (Burlin: 2008). Even if the Swedish musicologist Toivo Burlin’s account is more historical and it is limited to Sweden, I see his arguments as valid in the contemporary situation. In summary: while popular music producers have largely been thought of more or less as artists and creative forces behind the music, classical producers have been conceptualized more in terms of craftsmanship and servant to the artists and the work. My contribution here is to provide detailed ethnographic analysis on how the roles and agencies of producers between different genres differ and, perhaps more importantly, how they do not differ.

Giving Feedback

Giving feedback was a key element of the producer’s activities in all cases and aimed at getting the artists to perform their best in each performance. In Vepsäläinen’s case, the aim of recording vocals was to record as many takes as possible and get different Takes with different feelings and interpretations. The aim was also to strive towards what possible audiences would “dig”.

Giving feedback also served the purpose of recording a vast amount of takes to be used as raw material in the editing process. (Auvinen: 2016, p. 17–18.) What is also important to note here is the way feedback worked both ways; the singer Ida Paul also comments on Vepsäläinen’s activities and had a final say on for example his edits (Auvinen: 2016, p. 18).
Olsson’s work with the band Blind Channel followed very similar lines. What stood out especially in Olsson’s case was the overwhelmingly positive feedback he didn’t hesitate to give to the members of the band Blind Channel whenever it was appropriate (FD 29.11.2016). This created a good overall feel during the sessions, which also came up in interviews with the band. According to Niko Moilanen, the singer/rapper of Blind Channel:

We always expect good sessions. I also expect that we are going to have fun. The producer is important also from a humor-perspective. We work long days and it is important that the atmosphere stays good. (Blind Channel: 2016)

Here, the good feel of sessions connected to editing practices as well. In the classical case of Siirala and his work with the orchestra Tapiola Sinfonietta, the producer’s feedback, to me, seemed somewhat technical. In between takes, Siirala gave statements like “measure number 47 is [rhythmically] inaccurate” or “the string section has problems at measures 48 and 49” (FD 12.3.2016). Or, he would give extremely accurate feedback like “stay [rhythmically] together at measures 200 and 300” (FD 12.3.2016), which has a difficult unified rhythm pattern (Tüür: 2010, measure 21). On other occasions Siirala’s feedback was more figurative. He gave comments like "the phrase" in take number two "didn't speak to me" (FD 12.3.2016). In Siirala’s
case the soothing effect of the feedback seemed of great importance. Siirala gave his feedback from an instrument-storage room-turned-into-a control room through a talkback microphone that was connected to a loudspeaker situated next to the conductor onstage in the concert hall (Photo 1).

Siirala always gave his feedback in a very soft and calm voice despite the fact that the schedule was tight and that the recordings perhaps didn’t always progress as fluently as planned. (FD 12.3.2016.) This illuminates how important the producer is from a social perspective in classical record production; the producer has to make people feel that everything is going well and according to plan even if he himself feels that things could be going more fluently. Furthermore, s/he must stay calm and make others feel that there is no rush, even if time is running out. The social aspects of the producer’s agency thus become a part of the creative agency of the producer, as her/his social skills, character and feedback-giving capabilities influence the creative activities taking place.

On the surface, the difference between Siirala’s feedback in classical and Vepsäläinen’s and Olsson’s feedback in pop and rock seemed to be in its quality. While Vepsäläinen and Olsson concentrated on what to me sounded like issues concerning interpretation and the “feel” of a performance, Siirala’s feedback in classical music, to me, sounded technical. Siirala, however, offered a contradicting perspective. He elaborates:

First of all, what is interpretation? That needs to be resolved first. I think that everything is interpretation. If there's a wrong note, it is a bad interpretation from the perspective of the recording. You can't separate the right notes and how accurately the musicians are playing together from the interpretation of the work. They are always related…everything is connected to everything. I always aim at everything being correct, the way it is written in the music [in the score]. (Siirala, 2016c)

This statement brings forth an important aspect into how music resists atomistic analysis. Furthermore, it is a reminder of how my interpretation of a situation might completely differ from that of the people whose work I study.

Reasons for the quality of feedback between the classical and the popular music cases can be several. The greatest difference was in the number of artists the feedback was addressed to. In the classical case the feedback was often addressed to an instrument group, like for instance the violins with 16 musicians, or to the whole orchestra and the conductor. Therefore, the feedback couldn’t touch deep on the performances of an individual musician. Even with smaller instrument groups, like for example the cornets consisting of two musicians, the producer was not able to give very detailed feedback; tight recording schedules along with union breaks wouldn’t allow it. Also, the threshold of giving harsh feedback to an individual musician in front of an entire orchestra might have been high. In the cases of pop and rock, how-
ever, the producer’s feedback was pointed towards one musician at a time. Consequently, the possibility of getting deeper into the performance of an individual was therefore possible. Another factor behind the difference in the aims of feedback might lie in the differences of the structuration of producer compensation. Vepsäläinen’s compensation was heavily based in his shares in the copyright of the song (cf. Burgess). Thus, his earnings were heavily based on future streaming numbers of the song creating a strong incentive to think of what act as the “ear of the audience” (Hennion: 1983, p. 161). Síirala again worked on a fixed compensation. He did not have to worry about the sales of the record, not to mention the nominal sales of contemporary classical records released by Finnish record companies to begin with. Sales would not affect his earnings in the same direct way as in Vepsäläinen’s case.

Editing

Intensive editing was a key activity for all producers here. However, editing in the different cases served different purposes. In Vepsäläinen’s work, editing became essentially a mode of composing or arranging. Consider for example the first post-chorus of the song ‘Kunhan muut ei tiedä’ (Ida Paul: 2016), which starts at 1:05. By heavy chopping and editing, Vepsäläinen has used the last vowel of the punch line that ends the chorus, “ei tiedä”, as a rhythmic element of the arrangement. It could even be understood as a counter-melody for the pentatonic post-chorus melody, which sounds like a mallet instrument. In Vepsäläinen’s case, editing was also important in the way that the musical background was composed as most of it was essentially digitally programmed or coded through MIDI.

A similar example can be found in the band Blind Channel’s cover of the Macklemore song ‘Can’t Hold Us’ (Blind Channel 2017a), produced by Jonas Olsson. At 0:58, one can hear an “aah”-sounding riser consisting of overdubbed male vocals, which form a choir-like sound. This riser starts at the note G and glides into A. The riser works as a sort of a marker leading into the first pre-chorus. The riser appeared to be a spontaneous idea by the producer Jonas Olsson during a recording session. The idea sprung from the Backstreet Boys song ‘Everybody’ (Backstreet Boys: 1997), in which there is a similar choir-like “aah”-sound at 1:00. This sound functions essentially as a lead-in to the first chorus. During a recording session, Olsson first listen to the song ‘Everybody’ and then has Blind Channel’s lead vocalist go into the recording booth to record several takes and overdubs. Olsson then edits everything together. (FD 19.12.2016.) This is another example of how digital editing as a technological practice is important from the perspective of constructing the arrangement. In Olsson’s case, editing during small breaks in the recording sessions also served the purpose of making the musicians feel better. Olsson explains:
It’s nicer for the band to get some finished results. It’s very uninspiring to record some takes into some tube or into a kind of a black hole and often the artist’s perception of how well s/he has played is very different from the truth. (Olsson: 2017)

Additionally, this exemplifies the important practice of listening to reference material, which was an important aspect also in the case of Vepsäläinen (Auvinen: 2016, p. 25–26; Auvinen: 2017). The producer Olsson sees listening to reference material as an important part of producing. He elaborates:

The Internet is an important tool. We get to listen to all the music in the world as reference when we want. In the earlier days, we had to cycle to the library. Then they’re like “yeah, we don’t have that record here, come back in two weeks. We should be getting it by then”. (Olsson: 2017)

In addition to this practice as such, Olsson’s statement highlights the importance of new technologies in making this practice easier and broader. Listening to reference material can be understood as a form of consumption integrated into production (e.g. Théberge 1997); the producer consumes the music of others through digital technologies (streaming services) for aesthetic and creative ideas to be brought into the music s/he is working on. This practice did not occur in the classical case study. The classical producer’s point of reference was the score (Tüür: 2010) of the work under production and a recording of its premier performance. He didn’t draw influences from other compositions.

Editing in service of arranging was, however, more prevalent in the case of Vepsäläinen as opposed to the case of Olsson. This can be through understanding the producer’s agency in contemporary pop as “tracker” (Auvinen 2016; 2017, Hiltunen 2016). As tracker, Vepsäläinen wrote the song together with Paul and the arrangement was an intrinsic part of the composition; it guided the compositional process as a template for genre apposite creative decisions” (Bennett 2011). In the case of Olsson, however, editing had less to do with arranging and more with enhancing the performances of the musicians. The arrangements were mainly constructed by the band prior to the studio sessions and Olsson even used demo versions of the songs as templates when recording. Olsson’s role as an arranger was smaller and mainly restricted to nuances like the one I described above, although seemingly small nuances might bear convey meanings to the listener especially when sound is in question (see e.g. Lacasse: 2000). This suggests that the rock producer is less present in the music than the pop producer also in the real situation of the production process, not just “in the way music is discussed”, like Frith (2012, p. 221) suggests. Still, the rock producer remains “more important for rock as an art form than producers in jazz or classical” (ibid.).

In the classical case study, Siirala’s editing essentially aimed at constructing the final edit of the performance from the 88 separate takes recorded
during the recording sessions (Tüür: 2016). None of these takes were recordings of a performance of the entire piece played from beginning to end but were mostly rather short in duration. Despite the fact that the finished recording was a technologically built ideal performance, Siirala’s take on technological enhancements of music was comparably conservative. Siirala states:

We can use equalizers, but if we start to mess with pitch, we are operating in a grey area. –– I have to confess that I have committed a sin like this in the past, but in these cases, it has been the only option to remedy the situation and the other option would have been not to release it [the record]. (Siirala 2016c)

These notions reflect conservative attitudes towards music production technologies. The aim is to produce the ideal performance (e.g. Blake: 2012) of a score and what the listener hears should be something that has actually been played by the musicians. I see this as a value related to authenticity that arises from the philosophical and historical background of classical music. The other complication with respect to sound manipulation techniques in classical recordings is the recording technique. The orchestra is recorded as a unit and sound sources are not isolated (Photo 1). When all sound sources bleed to all microphones, manipulating one sound source would change the entire soundscape (Siirala 2016c). Even if this may seem like a purely technical issue, it is also an aesthetic one; recording all at once creates a desired orchestral sound, which could be very hard to achieve by recording everything in isolation. Having the space as an element in the final sonic product is obviously also an aesthetic choice stemming from the conventions and tradition of classical music (e.g. Klein 2012). In Siirala’s case, however, the fact that the recording space was the home venue of the orchestra, created a significant impediment on his agency. For example, Siirala wished to change the seating arrangements of the orchestra for the recordings. The musicians, however, didn’t agree to this because they wouldn’t “find each other musically” if arranged otherwise (FD 9.3.2016). Similar considerations of the importance of space in production from an aesthetic viewpoint didn’t come up with Olsson and Vepsäläinen.

Still, for me, a question regarding sound manipulation in classical post-production arises: if it is possible, why not use technologies to manipulate the recorded sounds if it makes the end product better? This also begs the question of whether or not editing in the sense of constructing an entire performance from numerous short takes should be called editing. To me, the word editing refers to making changes and editing a recorded performance. Perhaps editing in this form should be relabeled for example performance construction.
Formation of Authority

A key difference in how the producer’s agency was constructed was related to how authority and credibility are formed. The classical producer Siirala emphasizes the importance of professional classical musicianship:

I think it is a requirement that in this trade [classical producer] the producer has personal experience of making music, playing, singing or conducting —— This [giving feedback] is the essential issue in the producer's work. In the recording situation, you have to be able to give credible feedback to the artists so that a [relationship of] trust is established and that you really have an opinion upon which the artists can base their own decisions. (Siirala: 2015)

Siirala’s career as a producer started by him replying to a job advertisement in a newspaper and he went on to produce a record on his very first day on the job (Siirala: 2015) with a monthly salary without previous major chart success. The rock producer Olsson on the other hand emphasized production experience on the job. He discusses the beginnings of his career:

I expressed my opinions many times but people just shrugged their shoulders and continued the old way and didn’t want to see the problem. Or then they saw the problem but it was an ego-issue to them, in a way that a 19-year-old can’t know and we know better, we’ve played these songs. (Olsson: 2017a)

This comment insinuates that the authority of the producer in rock music comes from prior chart success and experience. Also, the band Blind Channel who Olsson worked with discussed how they chose Olsson as their producer for the very reason that he is a “top producer” (Blind Channel: 2016). One cannot be a top producer without prior success. The construction of authority through achievements as a producer is also reflected in the studio design. In his studio InkFish, Olsson had hung his gold and platinum records on the walls (FD 27.11.2016; Photo 2; Photo 3). This lets all his customers know about his achievements as a producer.
The key difference here is that the authority of the producer in classical music seems to stem more from the role or job description. The producer has the authority because s/he is assigned as the producer independent of the specific achievements. This perhaps reflects the nature of classical music as a more institutionalized art form. Projects are initiated from the top down; the record label wants to do a recording of a work, selects performers, engineers and producers. There exists a collective trust that whoever has been selected to do a specific job, in this case the job of the producer, has endured the necessary scrutiny or jumped through the required hoops to be capable of performing well in the role s/he is acting in. In classical music, the institution in a way works for the individuals and does the screening whereas in popular music individuals themselves have to be aware of the achievements of the people they want to work with. This also leads to the notion that the producer in classical music perhaps does not have to constantly prove her/himself to maintain authority. In popular music again, at least in the cases here, the process starts from the bottom down; artists and producers make contact and offer the package to the record label.

Finally, an important similarity between all cases was that the producer’s responsibilities ended after the editing process was finished. Mixing and mastering in each case study was conducted by another engineer. Reasons ranged from what could be perceived as limitations of one’s studio (Auvinen: 2016; 2017), to the acknowledgement of the limitations of one’s personal skills (Olsson: 2017) and to a self-evident and rigorous differentiation between specialized duties in classical music (Mäemets 2016), which was also manifested in the fact that, unlike Vepsäläinen and Olsson, Siirala had with him at the recordings the engineer Enno Mäemets to do the actual technical part of the recording. Siirala himself was able to concentrate on giving feedback and, by the powers granted by the red button (Photo 4), mediating communications between the control room and the concert hall where the orchestra recorded. (FD 12.3.2016.) This supports Gibson’s (2005: 205) finding that despite the development and digitalization of music technology, which has given more creative power to musicians (e.g. Williams 2012) and producers alike, “high-level mastering and post-production
facilities” as specialized services have still survived. My findings would suggest that this tends to be the case with several very different genres.

Concluding Thoughts

The producer’s role and agency are strongly formed through digital technological practices, independent of genre. Warner’s (2003, p. 33) argument that the relationship between technology and pop music is embodied in the producer extends into classical music as well. Here, Mike Howlett’s (2012) idea of the producer as ‘nexus’ comes closest to a common description of the producer’s role and agency across all genres, styles and production settings. However, as it is difficult to generalize on such a small number of cases, further comparative research is needed. I nevertheless contend that my findings have some relevance outside of these cases especially as earlier research has reached some of the same conclusions.

Differences in conventions and philosophical traditions between popular music and classical music shape the producer’s agency and role in the production process despite similarities in working practices. Despite the opportunities provided by digital technologies, the aesthetic aim in the production of classical musical records has remained to stay invisible to the listener.

Photo 4. Tapiola hall, Espoo Cultural Center.


Tuomas Auvinen: Differences and Similarities in the Role and Creative Agency of Producers in Pop, Rock and Classical (Klein: 2015). This reflects to the role and the creative agency of the producer.

With regards to the relationship between aesthetic aims and technological practice, all producers regardless of genre did not demonstrate high interest in technologies themselves. For all producers, technology was subordinate to aesthetic values and creative ideas (see: Martin: 2014, p. 232). This was different from for example the engineer in the classical case study, who was very specific about his technologies and equipment (FD 23.3.2016).

The biggest similarity between the three cases was the way the producer gave feedback to the artist(s). Feedback in all cases aimed at increasing the quality of the artists performance. Also, heavy editing was common to all producers but the difference was the degree to which editing was a compositional or arranging (popular music) practice or a performance construction (classical music) practice. The nature of editing here made me question the essence of the term “editing”, which to me means making changes to a recorded performance. Here, editing seemed to be a much broader practice.

Differences between technological practices of producers exemplified the different conventions and aims behind production, which frame the agency of the producer. In popular music, producers used all and any technological means to make the music better and music production technologies, much through the agent of the producer, are in the service of new practices and ideas. This came up the strongest in in the way the phases of composition, arranging, pre-production, recording and post-production editing got purposefully entangled and mixed during the process. Recording and sound manipulation technologies can therefore not be isolated as tools that belong only to the recording process to realize a performance on a record but are an intrinsic part of the entire creative process from composition to mastering. This partially connects to how agents work “directly with sound”, which constantly evolves (Warner: 2003, p. 18–19; Théberge: 1997, p. 192). In classical music, however, contemporary technological practices seem to be in the service of traditional ideas and aims and goals; reproducing the ideal performance of a score into a recording medium, an idea that prevailed in the beginning of record production in all genres. At play here is the conceptual binary innovation/tradition. All producers here were technologically current in how they conducted their work by using the latest digital technologies. However, in popular music, contrary to classical, innovation also expanded into ideas of what should be done. The conceptual binary of production/consumption offers an interesting aspect into the differences of the guiding philosophical underpinnings of classical and popular music. The consumption practice of listening to reference material reveals that whereas in popular music it is natural to be influenced by and take ideas from other pieces of music, a classical work is viewed as more self-sufficient as the point of reference is the work under production itself.
Finally, rather than differentiating between different genres of music stylistically, a more fruitful distinction might arise from the way the producing agents conceptualize the aims of record production. Perhaps a continuum could be made between performance music and production music, a dichotomy that was introduced to me by the Canadian engineer and producer Paul Novotny at the 12th Art of Record Production conference.

Acknowledgements

I would like to express my gratitude to my supervisors, Professor John Richardson (University of Turku), Professor Johannes Brusila (Åbo Akademi), and Dr. Susanna Välimäki (University of Turku) for their support and constructive comments in the various phases of my research. I also want to thank the three peer-reviewers for their constructive criticism and comments for the betterment of this text. Furthermore, I would like to thank Professor Alan Williams, Professor Paul Novotny and the independent producer and engineer Dr. Phil Harding for their helpful comments and questions on my presentation at the 12th Art of Record Production conference in Stockholm, Sweden. Finally, I'd like to thank the editors of this issue of the Journal on the Art of Record Production.

Bibliography


Tuomas Autinen: Differences and Similarities in the Role and Creative Agency of Producers in Pop, Rock and Classical Music


Discography


17

Scores

Sonic Materials

Interviews


Photos
Photo 1. March the 12th 2016. Photo taken by Tuomas Auvinen at Tapiola hall, Espoo Cultural Center. Photo in the possession of the author.

Photo 2. November the 27th 2016. Photo taken by Tuomas Auvinen at InkFish studio. Photo in the possession of the author.

Photo 3. November the 28th 2016. Photo taken by Tuomas Auvinen at InkFish studio. Photo in the possession of the author.

Photo 4 March the 12th 2016. Photo taken by Tuomas Auvinen at Tapiola hall, Espoo Cultural Center.
**Andrew Bourbon**: Hit Hard or Go Home - An Exploration of Distortion on the Perceived Impact Of Sound On A Mix

*Introduction*

Distortion is a powerful tool used by mixengineers to enhance the presentation of audio through manipulation of an existing sound. Distortion can take on many forms, with common distortion tools either emulating an existing hardware tools or providing new interfaces for the enhancement of audio, adding new material to manipulate the energy delivery of a chosen sound. This enhancement has the potential to not only draw attention of the listener to particular components of a sound, but also changes the fundamental delivery energy of that sound.

As mix practice has moved into the Digital Audio Workstation (DAW) environment, plugin developers have increasingly moved to emulating classic hardware, with ever more accurate representation of the signal paths of the target processing tools. Recent developments in plugin design have seen the provision of user control for the manipulation of the distortion component of the processing in a channel strip or console input channel, with opportunities to either reduce the impact of distortion or indeed to simulate circuit drive to a higher level than would have been possible in the analogue circuit without increases of gain. In addition to the provision of emulations of classic hardware, there have been developments in channel strip designs that do not offer direct emulation of a particular hardware unit, instead choosing to offer flexible direct control over parameters such as distortion, providing drive controls and the ability to focus the distortion, commonly on either even or odd low order harmonics.

This paper will explore usage of distortion in the contemporary DAW environment, focusing on the musical impact of low order harmonic distortion as applied in music mixing using a range of audio processing tools. The selected audio plugins are taken from a range of manufacturers, demonstrating both emulations of existing hardware and tools designed specifically for use in the DAW environment. The selection of plugins made does not represent
endorsement or quality judgment on the tools themselves, rather offering a representative case study of low order harmonic distortion in mixing. Selection of tools is derived from mix and production process, exploring processing utilized in the traditional analogue chain and now emulated in the digital domain. Digital only tools developed for use in the DAW have also been chosen for exploration. Through technical measurement, listening analysis, and audio examples various low order distortion approaches will be explored in relation to the manipulation of performance gesture and energy in music mixing.

**Audio Files**

The link provided includes a folder of audio examples giving musical context to the measurements undertaken through this paper https://www.dropbox.com/sh/l4x0t3i6uzmigfc/AADGBnvIND90YaBP4MuJbo3Ea?dl=0.

**Exciting Stuff**

Distortion is commonly used in music mixing to manipulate the perception of a particular sound or sounds, with the properties of various processing tools known to engineers and sought out for their particular sonic affordances (Clarke, 2005). The sonic impact of various large format consoles is “instantly recognizable” to experienced engineers (Massy, 2016), with specific sounds sought after for their association with existing records. Massy goes on to describe the process of “creaming the mic pres”, driving classic microphone preamps into desirable distortion that is pleasing to the listener. By driving the mic pre a sense of excitement is added to the sound, due to combinations of reduction in crest factor (ratio of peak to RMS level) through natural volume compression and the addition of harmonically related material. Case (2007) also refers to the excitement afforded by distortion, stating that

> When a device is overloaded, something exciting must be happening. Someone is misbehaving. Rules are being bent or broken. Distortion in rock and roll is as natural as salsa in Mexican food. It is the caffeine of music (Case, 2007, p.97)

During recording and mixing it is common to drive the input circuitry of tools to add excitement to an audio signal. Contemporary tools allow increased range of distortion, with for example the Brainworx console series of plugins allowing greater depth of distortion from the same input signal through a variable increase in total harmonic distortion (THD). The impact
of maximum THD is shown in figure 1, with the addition of third harmonic material to the sine wave being processed by the Console E emulation.

![Figure 1. Additional distortion from maximum THD addition.](image)

It is also however clear that the impact of driving the mic input of a console results in a very different distortion profile to that achieved through re-amping via a guitar amplifier or high distortion tube interface such as a Thermionic Culture Culture Vulture. Low order harmonics are “dramatically less audible” (Whitlock, 2015, p.378) with significantly different affordances to higher order harmonics.

**Adding Odd Harmonics- Transformers**

The channel strips of the classic consoles all offer specific harmonic enhancement in their analogue circuitry. Recognized recording consoles include those manufactured by Neve and API, each of which has differing sonic characteristics desirable for particular approaches to record making. The transformers in the audio circuits are a significant contributor to audio performance, with an impact on both the harmonic content added to the signal and in envelope shaping in response to transient material in the musical content. The transformer in the audio circuit performs a range of functions, being used in signal balancing and noise rejection and as a gain-providing element. As level increases, transformers also exhibit increasing frequency specific phase shift and non-linearities in distortion, making them dynamic devices that can effectively be ‘dialed in’ through the end user driving signal into the transformer to match the specific sonic requirements of a record. The harmonic impact of a transformer is defined by the materials found within the transformer core, the size of the transformer, and the nature of the transformer windings (Whitlock, 2015). The bigger the transformer is, the deeper the low frequency response and the higher the level of signal that can be handled by the transformer before saturation takes place (Winer, 2012). It is worth noting however, that increase in inductance in larger transformers does compromise high frequency response, making larger transformers unsuitable for audio usage. Selection and design of audio transformers has historically been undertaken with the aim of providing the optimal frequency
response and amplification characteristics at a given impedance. The harmonic content that is now craved by engineers was an unfortunate by-product of the available technology at the time. It is interesting that modern designs have often embraced active integrated operational amplifier (opamp) technology to improve linear behavior and minimize distortion and noise beyond the measured performance offered by the audio transformer. Despite the improved measured performance of an opamp based circuit, the technical flaws of older transformer designs are still incredibly desirable from a musical perspective.

Distortion found in audio transformers is predominantly third order in nature, resulting in for example new material at 300Hz added to an input signal at 100Hz. The percentage of Nickel in an audio transformer directly correlates to the level of distortion provided by the core, with higher percentage of Nickel resulting in a less distorted output. Three common core materials are 84% Nickel, 49% Nickel and M6 steel, with increasing distortion as the Nickel content reduces and the highest distortion provided by the M6 Steel core (Whitlock, 2015, p378). The lower the frequency hitting the transformer the higher the distortion that is added by the transformer, with the high nickel content transformers exhibiting negligible harmonic distortion above 50Hz. The cutoff frequency continues to rise as the percentage of nickel in the core drops, however, all core materials exhibit lower distortion as frequency increases. Shadow Hills Industries offer variable transformer control to the end user, with the Gamma preamp for example offering Nickel or Steel settings, as well as a transformerless discrete option to help shape the musical impact.

One desirable characteristic of audio transformers is the resistance to intermodulation distortion. Intermodulation distortion is both musically undesirable and easily heard, and as such, much effort is made in the design stage to minimize intermodulation distortion. As discussed, transformers exhibit third harmonic distortion, with very little measured intermodulation distortion. Whitlock (2015) provides an example of the performance of a Jensen audio transformer in comparison to an amplifier circuit with a similar distortion profile;

The Jensen JT-10KB-D line input transformer has a THD of about 0.03% for a +26dBu input at 60Hz. But, at an equivalent level, its SMPTE IM distortion is only about 0.01% or about a tenth of what it would be for an amplifier having the same THD. (Whitlock, 2015, p.379)

The lack of undesirable intermodulation distortion with the musically pleasing third harmonic enhancement makes transformers an important aspect for consideration in the musical application of distortion.
Comparative Measurement Methodology

In order to begin the process of comparing the musical impact of distortion from transformers, two emulations of the input transformers from a vintage Neve and an API console have been explored. Kush Audio has developed a hybrid mic preamp solution; with a 500 series preamp working in concert with the Omega transformer plugin to replicate the harmonic enhancement provided by either a Neve or API input circuit. It is important to note that the impact of the channel processing and mix bus in the consoles is not replicated in this process, with the plugin simply replicating the transformer itself. In order to demonstrate the harmonic impact of the emulated transformer a 100Hz sine wave is passed through the input, with the output measured to compare. Figure 2 shows the initial measurement with the plugin bypassed, resulting in a pure 100Hz sine wave at the output.

![Figure 2. Sine Wave Measurement](image)

Figure 3 shows the same 100Hz input signal, but this time run through the Omega A API transformer emulation.
The measurement shows significant presence of the third harmonic, with some additional fifth harmonic content present. As well as these odd harmonics, there is also some presence of second harmonic distortion present in the signal. By contrast when measured the Omega N plugin, emulating a Neve input transformer, figure 4 demonstrates a pure third harmonic distortion at a lower relative intensity to that found in the API transformer.

Figure 4. Neve Transformer

Though it is difficult to calibrate these plugins, it is clear that at this level of drive the characteristic of the API transformer features a higher level of harmonic complexity to the pure third harmonic demonstrated in the Neve. Increasing drive on the Neve (figure 5) sees pure odd harmonic distortion,
with third and fifth harmonic present, and each harmonic reducing in intensity. Driving the API (figure 6) sees predominantly odd harmonic distortion, but with some second and fourth harmonic present at a low level. The distribution of energy in the harmonics is also very different in the API, with the fifth, seventh and ninth harmonic all contributing significant energy relative to the fundamental.

Figure 5. Neve Drive

Figure 6. API Drive
Even Harmonics

As demonstrated, transformers are commonly associated with low order odd harmonic distortion, with higher order distortion only manifesting as the core begins to saturate. Other magnetic sources such as tape exhibit similar harmonic characteristics, with a well-maintained tape machine exhibiting low levels of third harmonic content that increase as signal is driven into the tape. Tubes offer a different distortion profile to transformers and tape, with a predominant association with second harmonic distortion. Figure 7 demonstrates the warm tube setting from the Fabfilter Saturn plugin, with a measured response showing dominant second harmonic and very low levels of third harmonic distortion.

![Figure 7. Emulated Tube Distortion](image)

As the drive level is increased, the warm tube profile sees an increase in second harmonic distortion, before then showing odd and even harmonics as the drive moves towards maximum, showing the increased harmonic complexity associated with high levels of drive into tube circuitry.

Contemporary plugin design has seen manufacturers embracing manipulation over odd and even harmonics distortion. One such example of this is in the Waves Scheps Omni channel (Figure 8), which provides variable odd or even distortion. Both odd and even exhibit similar behavior, with the odd harmonic adding increasing third harmonic distortion in the first half of the drive range, and then adding fifth when driving beyond fifty percent. The even mode simply swaps the third and fifth harmonic for second and fourth harmonic distortion. Unlike the tube and transformer distortion devices measured, the Omni Channel does not offer high order saturation at the highest drive levels, instead adding either fifth or sixth harmonics depending on the selected mode.
Another example of a channel distortion tool that offers variation between second and third harmonic is the SSL X-Saturator. Figure 9 shows the range of controls, with the option to blend between second and third harmonic distortion. When set to second harmonic, there is still some measured third harmonic content, with higher order distortion also prevalent when the drive and depth controls are increased. The X-Saturator also features a unique shape control, allowing a choice between hard and smooth edged distortion.

Figure 8. Waves Scheps Omni Channel Interface

There are a number of tools available from a range of manufactures that offer similar low order odd and/or even harmonic enhancement, as found in classic console strips. Having identified the nature of this distortion it is important to evaluate the impact of this distortion on the musical delivery of sounds in a mix.

How Does It Sound? Drawing the Sonic Cartoon

Zagorski-Thomas (2014, 2016) proposes the concept of the sonic cartoon, which provides a useful method for considering the impact of distortion on a
musical event. In the example cited by Zagorski-Thomas, a guiro scrape is used to increase the sense of angst in the Britney Spears vocal performance, manipulating the listener’s interpretation. Distortion is another tool available to engineers to manipulate the listener’s interpretation of performance, giving the ability to manipulate the perceived gesture in the performance of the sound. Distortion is one of a number of tools that allow engineers to manipulate the perception of the original recorded sound (Bourbon & Zagorski-Thomas, 2016), resulting in a sense of increased impact, heaviness, energy and effort subject to the source and the nature of distortion employed. As an analytical tool *sonic cartoons* provides a useful framework for the study of the affordances of distortion in a mixing context.

![Figure 9. SSL X Saturator](image)

In considering drums the impact of low order distortion is most obvious in the perception of the gesture of hitting the drum. When distortion is engaged using the Waves Omni Channel there is a clear sense of the drum being hit harder, with increased drive leading to a more dramatic sense of the drum being hit. The impact is less dramatic than that provided by the transformer emulation tools, with the additional transient shaping and increased compression created by the processing also creating the sense of effort associated with distortion and compression. There is however a clear sense of change in
Andrew Bourbon: Hit Hard or Go Home - An Exploration of Distortion on the Perceived Impact Of Sound On A Mix

impact, with the distorted versions offering a much greater sense of conviction in the hit of the drums. There are 3 sets of audio files provided as drum examples, with microphone output processed only with the distortion section of the Waves Omni channel. Both odd and even distortion afford a sense of increased impact and punch, but with changes in the nature of that impact and the staging of the drums depending on the distortion setting.

The even harmonic distortion gives a much stronger sense of the drum being hit hard, with a strong feeling of solidity and drive. The kick drum in particular feels tighter and drier, with the lower midrange in the snare feeling considerably more solid. The presentation of the overheads is also significantly manipulated, with a real sense of detail in the stick hitting the cymbals and an overall more coherent presentation of the image of the drums. The sense of a single drum kit, rather than a collection of drums that is provided by this distortion processing is significant, creating a very different picture of the drummer and the performance to the listener.

When exploring the balance of the example featuring odd harmonic distortion there is again a sense of increased impact. Compared to that of the second harmonic distortion example however, there is a very different sense of groove. Whilst the second harmonic provides a real sense of hart hitting and coherent drive, there is a sense of lift in the third harmonic example, creating a sense of lightness and bounce. There is an increased sense of air around the cymbals, and overall a greater sense of the space that the drums exist in. The second harmonic distortion example has a greater sense of focus, whilst the third offers coherency but through air and bounce. Placed in context with a track the impact of the distortion changes again, showing the importance of evaluating distortion in context when undertaking a mix. If we were to compare a different drummer playing a different part, we would see a different result again- the aim here is not to create a typology for distortion, but merely to recognize that the impact of distortion is to change the perceived presentation of the element being distorted. The decision to provide a more focused driven sound, or a lighter more open groove is entirely dependent on the context of the recording, with choices made through the production and mixing process in response to a musical language.

A similar impact can be heard when working with orchestral sample parts. When engaging the SSL X-Saturator plugin on programmed strings, it is possible to manipulate the playing gesture, increasing the sense of the bow dragging on the string. Adding depth creates a sense of friction, bringing out detail and a sense of reality in the performance audible in sample. By moving from second to third harmonic it is possible to move the focus, with the second harmonic creating a sense of richness and warmth, and the third harmonic offering a sense of edge and friction. Again the appropriate choice is entirely dependent on the context- in this case the composer was keen to add intensity and a sense of aggression, which was supported by the choice to embrace third harmonic distortion on the string sections. The brass parts also
received third harmonic distortion with significant drive, adding not only a sense of the instrument being played harder, but also allowing the instrument to cut through the mix.

**Distortion and Masking**

In addition to manipulating the performance gesture, distortion can also be a powerful tool in managing the impact of frequency masking on the mix. As the harmonic content is enhanced, separation between elements can also be enhanced, with harmonic content clearly identified as belonging to a particular element. In an example of this, after hearing a saturated piano part in a recent mix, the client asked what had happened to the bass? The bass line had clearly become more pronounced, with detail in the bass line significantly enhanced. This improvement in bass detail and perceived increase in performance solidity came solely from distortion of the piano line, which also moved from sounding insecure to sounding resolute. When exploring these tools it is important to always work in context, as the harmonic interaction of elements in a mix are only understood when all elements are exposed. It is also important to work in context as low order, low energy harmonic distortion is considerably subtler than higher order harmonics, and can often only be heard in context. A vocal with light distortion in a mix may see emotional enhancement that matches the requirements of the song, but in isolation may be difficult to detect.

**Distorting the Delivery- Envelope Manipulation**

As well as providing harmonic coloration, transformers can also distort the envelope of musical material, particularly when transients are processed. Audio transformers are inherently slower to react to an input signal (Whitlock, 2015 p.380), resulting in the rounding of edges of a square wave when presented into the transformer subject to the way the transformer is integrated into the circuit. Improper damping of the transformer can see audible ringing in the transformer, with a network of resistors used to customize the response of the transformer to audio transients. The result of this is that transients take longer to pass through the transformer, rounding the transient in addition to adding the discussed harmonic distortion. On a drum this can be seen as being similar to comparing a contemporary bearing edge with a vintage leading edge. The vintage drum has a softer edge, creating a longer contact time with the drumhead and passing the signal less quickly in to the drum than the modern cut drum edge. In a modern drum the transfer into the shell is faster, with less damping effect on the drumhead through the transfer of energy. This increased transfer time creates a sense of increased impact time, which from a hitting and sonic cartoon perspective could be described as a punch to the stomach rather than a slap to the face. Both approaches
increase the sense of being hit, but with a significantly different gestural impact.

As level sent into the transformer increases, the core moves towards a point of saturation, with a resultant increase in distortion and average level. Combined with the rounding of transients, the impact of transformer distortion moves beyond simple harmonic enhancement. As the core cannot cope with increased level the peak output level increases at a slower rate than the input level, creating a decrease in the ratio of average level and peak level. This natural compression results in a changed envelope, and also a change in the overall spectrum with the core saturation leading to frequency specific dynamic changes subject to the musical material, and the nature of the transformer. In addition to this we also see significant phase shift present in some transformers- the Neve transformer for example showing 40 - 60 degrees of phase shift above 10kHz. These characteristics come together to create a multi faceted distortion, which can significantly impact on the perceived contribution of the processed sound to the overall musical presentation.

**Compressors and Distortion**

Compressors are another audio processing device that add distortion to signals, and as with transformers tend to offer frequency dependent distortion with a more aggressive distortion profile often encountered when compressing low frequency with a fast attack time. It is also the case that many hardware compressors offer two stages of distortion, the first being in the audio path of the compressor regardless of the presence of compression, and the second being the distortion added by the gain reduction cell. Figure 10 shows a Universal Audio 1176 LN Revision E audio path distortion response, with figure 11 showing the same compressor in the act of compression. It is clear that the distortion added by this device is entirely dependent on the amplitude of the signal arriving at the sidechain.

![Figure 10. Out of compression](image1)

![Figure 11. In compression](image2)

By setting the point of onset of compression it is possible to dynamically manipulate a performance, with the louder sections receiving the combina-
tion of level change and harmonic enhancement. The harmonic profile of compression varies from compressor to compressor, with the Teletronix LA-2A performance presented in figures 12 and 13.

![Out of Compression](image1)

![In Compression](image2)

**Figure 12. Out of Compression**  
**Figure 13. In Compression**

The above figures show that the LA-2A has prominent second harmonic distortion in the audio path, with third and fifth harmonic added during compression. The end result of the presented distortion and compression profiles is a sense of hair and air in the compression action of the 1176, with increasing aggression at higher ratios and with greater gain reduction. As a semantic descriptor hair is perhaps a controversial term, however the sense of detail and growl added to a vocal supports the use of this terminology. Faster attack enhances the sense of hair, with a genuine sense of air added as release times are decreased. The compressor is also fast enough to add significant low-end distortion, acting as a pseudo fuzz box on a bass guitar if set aggressively. The LA-2A by comparison has a sense of warmth in the audio path, but with a sense of openness, lightness and detail added during compression, adding a smooth yet open sound to a vocal and adding a sense of detail in the impact between the stick and the cymbal on a pair of overhead microphones. It is clear that low order harmonic distortion in a compressor plays a huge part in the manipulation of performance energy and gesture, in concert with attack and release characteristics to give engineers a multi-faceted tool for the manipulation of perceived performance.

*From Light To Heavy Distortion - Changing The Role Of Distortion*

The focus of this paper has been on low order harmonic distortion used at a relatively low level to manipulate the perceived performance energy of a musical sound. As channel strips are driven in level, considerably higher distortion can be exhibited. Figure 14 demonstrates the Waves NLS Nevo channel, emulating a Neve 5116 channel input section. As has been demonstrated in previous Neve consoles, the distortion is mainly third harmonic in nature with similar sonic impact to that experienced with the other Neve
channel. Figure 15 is the same Neve channel, but running through the mic pre emulation, adding significant gain and distortion.

![Figure 14. Neve Line Drive](image14)

![Figure 15. Neve Mic Drive](image15)

The higher order harmonic distortion present in the mic drive option results in a very different sonic impact to the lower order harmonics discussed in this paper. The NLS Nevo on a vocal creates a sense of midrange weight and air, effectively creating a sound with more conviction in the performance and an increased sense of quality and luxury in the vocal chain. On engaging the mic mode, the vocal is instantly transformed, sounding over-
driven and buzzy. Rather than creating a sense of quality and expense, the vocal chain instead sounds broken, with the vocal breaking up in the distortion. There is a sense of heightened emotion in the delivery after the aggressive distortion, but not in a way that could be used in the context of the chosen song to support a musical direction.

When placed on a parallel processing path, the high order distortion pictured in Figure 15 creates a sense of energy and heightened emotion in the delivery. The breakup in signal is particularly noticeable on louder words, adding impact and a sense of theatre to the vocal delivery. There is a significant increase in density of the sound, creating a vocal that dominates the center of the image and adds a sense of weight to the lyrics of the song. The impact on the lyric is particularly interesting, with the song feeling less like the artist is reflecting on ‘how to fix a broken heart’, and instead asking the listener directly. This transition from an internalized thought process to an externalized question posed to the listener represents a subtle change sonically, but a significant and indeed powerful change from the perspective of performance and communication. As the level of the parallel distortion is increased, the more prominent aggressive distortion creates a sense of fire and emotion in the vocal delivery, but with the distracting presence of high order harmonics detracting from the emotion of the song. Utilizing the pentode tube emulation of Soundtoys Decapitator (Figure 16) still provides the feeling of fire and emotion in the vocal, but with a pleasing less edgy quality that retains musicality despite high levels of drive.

![Figure 16. Soundtoys Decapitator Pentode](image)
When drums are added to the same distortion parallel the sense of intensity increases across the mix, with a sense of desperation in the vocals supported by the now significantly distorted drums. The distortion parallel brings the drums and vocals together, adding both drive and depth, as well as a sense of vintage tone to the overall presentation. Significant parallel distortion of vocals, and indeed other elements is a technique that is explored by a number of top mix engineers, with mix engineers such as Tom Elmhirst regularly creating a sense of emotional energy and depth in vocals, drums and keyboards through what would be traditionally considered as over distortion. One example of this can be heard in Elmhirst’s mix of ‘Hello’ (2015), where the distorted vocal can be heard with the main vocal, adding emotional weight to core lyrics and enhancing performance. As the song moves into the chorus the distortion becomes more prominent, which both heightens the emotional impact of the vocal and also creates a sense of size as the distortion interacts with the space and the vocal layering and delay processing. As the drums are exposed through the track the presence of distortion also adds to the sense of delivery and grand energy. As the filtering on the drum opens through the final chorus the distortion on the drums is further exposed, adding a sense of the entire track calling out to the listener. Distortion is an essential tool employed in this mix in order to manipulate performance energy, emotion and gesture communicated to the listener, enhancing the iconic sound of Adele and the associated staging of the record.

**Conclusions - Distortion In Modern Workflow**

The DAW has provided engineers with an environment where distortion can be used in a variety of ways to enhance the performance energy perceived by the listener. From low-level harmonic distortion, adding a sense of size and weight to delivery through to heavy distortion creating a sense of aggression or destruction, distortion plays an important role in manipulating the performance gesture and emotion. It is important to note however that the prevalence of distortion in contemporary mixing practice does not come without its cost. As the loudness wars have seen a reduction in perceived loudness, mixes have increased in distortion as a by-product of reduced crest factor and significant limiting. Modern DAW gain staging requirements see practically infinite headroom within the DAW, with the only consequence of heavily clipped channels being the requirement to turn down the master fader before bouncing the track. The high order distortion in many contemporary mixes is also brought out in the mastering process, leading to brighter and more abrasive mixes that lead to the listener reaching to turn down rather than turning up the music.
Further Study

Whilst this paper has focused on low order distortion on channels, there is potential for continued research into higher order distortion and its use in contemporary mix practice. The impact of high order distortion, and low order distortion of high frequencies also provides a subject for further consideration, particularly with the identified trend for brighter more aggressive mixes. The next area for the continuation of this research is in the study of distortion on mix busses, and in mastering.

Bibliography


Discography

Toivo Burlin: The creation (and re-creation) of virtual spatiality: Surround sound mixing in King Crimson’s back catalog box sets

Abstract: The article analyzes surround mixing of King Crimson’s analog recordings. “Lizard,” “Larks’ Tongues in Aspic (Part 1),” “The Night Watch,” “Trio,” and “One Time” are analyzed using the theoretical perspectives of the music recording as representation and spatial design. There is a difference between live and studio recordings in surround: live recordings tend to be mixed according to the direct/ambient principle, while studio recordings are mixed according to the principle of direct-sound all around. Live recordings in surround can also often be understood as examples of documentarism, and studio recordings as examples of hyper realism. The spatial design of the surround mix tends to emphasize aspects of spatiality in the broad sense: where the recording can, for example, be interpreted as a landscape in sound.

Introduction

This essay discusses some tendencies in the emerging conventions in contemporary surround mixing of popular music. The subject is music somewhat outside the pop mainstream, but quite popular in a musical sub-culture context. Specifically, the essay looks at studio and live recordings of multi-layered (both musically and in terms of recording technology), so-called progressive rock music. The progressive band King Crimson—a group that encompassed many musicians, instruments, setups, and musical styles over a timespan of fifty years—recorded and produced its output using a range of technologies. Its first professional recordings in the 1960s were on 8-track reel-to-reel machines; in the 1970s it used 16 tracks, and in the 1990s 48

1 The band was formed in London in 1969 by Robert Fripp, Ian McDonald, Michael Giles, Greg Lake and Peter Sinfield.
tracks. The band also frequently was documented in many live and studio sessions in soundboard recordings on e.g. cassette, DAT, and ADAT tape. In the 2000s, a great many recordings of the band, in its various incarnations, were systematically digitized, mixed for the first time or remixed to new stereo mixes or in modern surround formats that were mostly not available at the time of the original recordings. The analog multitrack tape recordings—originally mostly designed for stereo—were transformed into high resolution digital audio, and mixed and stored (at the suggestion of Steven Wilson) in surround formats such as (uncompressed) LPCM 5.1 and (compressed) DTS HD surround. The transformation raises questions about the status of the recordings: what happens to them as musical representations, and how is this status affected when remixed to surround? It also raises questions about spatiality in recorded music. I want to address both issues here.

Specifically, I will look at King Crimson’s ambitious box sets of their back catalog from the late 1960s to the 1990s: In the Court of the Crimson King (2010), Larks’ Tongues in Aspic: Complete Recordings (2012), The Road to Red (2013), and Starless (2014), all covering their mid-1970s material; Thrak Box: Live and Studio Recordings 1994–1997 (2015), covering a huge part of the 1990s; On (and Off) the Road (2016), covering the 1980s; and finally Sailors’ Tales 1970–1972 (2017), covering the studio and live recordings of that early period.

Box sets are conceptualized, produced, and packaged in the alternate sphere of the more musician-driven music industry, for the most part outside of Spotify and its streaming compatriots, and with their own distribution channels. King Crimson’s own record company, Discipline Global Mobile (DGM), was established by Robert Fripp and David Singleton in 1992; Fripp and King Crimson now have full control over the master recordings. DGM has shown that it is possible to define alternate ways to release and distribute music—perhaps with more emphasis on the needs of musicians and fans rather than the industry—and simultaneously to mostly release music on discs (LP, CD, DVD, Blu-ray) rather than for streaming. Consequently, they have released much recorded material—even sub-standard bootleg live recordings, mostly of historical interest for the most dedicated fans—that more conventional record companies probably would not have, owing to the lack of immediate commercial potential.

In the following, however, I will analyze only a strictly limited set of audio material. I will discuss in detail a few example recordings: new surround

---

4 DGM offers streaming and downloads from its homepage.
mixes of originally analog multitrack tape recordings taken from the box sets. The box sets in their entirety as sources are also taken into consideration in the analysis.

Transforming analog popular music recordings to digital surround

The history of surround sound is a fascinating one, and here I will outline just a few aspects relevant to the subject of this article. It is a history of technological development that also clearly has important connections to both the history of music in the largest sense (for example, in its spatiализing function, to music as a multidimensional spatio-temporal phenomenon connected both to the universe and to earthly, human “territorializing” functions) and to the history of recorded music, sound, and cinema. The short background given here draws on a variety of audio sources and on Tomlinson Holman’s *Surround Sound: Up and Running* (2nd ed., 2008). (Holman is the inventor of the THX cinema system and was involved in the development of 5.1 surround and 10.2 systems.)

Surround sound in recorded music has roots both in music composition itself and in sound recording technology: partly for purely audio music, partly for cinema. The early roots of surround reach back to prehistorical, even mythological, times. Within a more historical timeframe, surround as the idea of a planned spatial organisation of instruments, vocals, and sounds as an integrated parameter of a musical composition dates back at least to the antiphonal singing of Adrian Willaert in St. Mark’s Basilica, and the “all-around” polyphony, or “spaced-antiphonal singing”, of Giovanni Gabrieli, the first to use precise, not just left-right orientation directions for musicians, in sixteenth-century Venice. From there it can be traced to Hector Berlioz’ *Symphonie Fantastique* with its brass players playing from behind in 1830, and Berlioz’ 1837 *Requiem*, with its four discrete orchestras called North, East, West and South. Composers such as Richard Wagner and Gustav Mahler also recognized the importance of spatiality in music and tried to write it into their music, later realized on recordings (Patmore & Clarke, 2007, Holman, 2008, p. 2–3; Burlin, 2008).

The long technological development of surround sound includes such movie milestones as *Fantasia* (1940), with its Fantasound system, partly designed by Walt Disney and Leopold Stokowski. Consisting of three front channels and two surround channels located at the back of the cinema—a precursor to the 5.1 standard—it also included advanced optical recording.

---

The Cinerama system invented by Fred Waller and Hazard Reeves in 1952 consisted of a widescreen film with seven channels of sound: five in the front and two surrounds in the back. A more financially successful incarnation of Cinerama was the Cinemascope system, with four channels; and Star Wars (1977) used a system consisting of three channels in the front, one surround, and one low-frequency channel, for enhancing explosions in space (Holman, 2008, p. 2–9).

With a precursor in quadraphony in the 1970s (when much classical and popular music was issued in quad on both open-reel or 8-track cartridge and on three specially designed, but incompatible, vinyl formats for home music listening), surround in digital audio for the consumer market dates back to the early 1990s. Surround sound in movies, in a lineage descending from Star Wars, made an impression on the home market, and in 1987 the 5.1 standard for film was established by the Society of Motion Picture and Television Engineers (SMPTE). It was followed by a series of digital audio industrial systems (and related codecs) including Dolby Digital, Digital Theater Systems (DTS) and Sony Dynamic Digital Sound (SDDS) in the 1990s. All of these systems used low bit-rate coding, due to lack of space on media such as the CD (which, like the later Blu-ray format, used uncompressed linear PCM coding) and CD-ROM. With the advent of new audio media such as SACD in 1999 and DVDA in 2000, it became possible to release high-definition pure music mixed in surround for the consumer market (Holman, 2008, p. 10–22).

From around 2000 onwards there were occasional re-releases of older mainstream pop and rock music in surround 5.1. Examples include Pink Floyd’s Dark Side Of The Moon SACD 5.1 remix of 2002; Mark Linnett’s technically advanced remix to surround of the Beach Boys’ Pet Sounds in 2003; and Genesis’ critically acclaimed SACD/DVDA box set Genesis 1970–1975, released in 2008, with surround mixes by Nick Davis. King Crimson’s re-released box sets began in 2009 with the 40th anniversary series In the Court of the Crimson King and continued with most of the surround mixes produced by Steven Wilson on the earlier material and Jakko Jakszyk on the newer material, who each took quite different approaches to surround mixing.7 Both Wilson and Jakszyk also remixed recordings in surround for other 1970s acts, including Emerson, Lake & Palmer (Trilogy and Brain Salad Surgery), Gentle Giant (Octopus and The Power and the Glory) and Jethro Tull (Aqualung, Thick as a Brick, A Passion Play and Songs From the Wood). Other notable releases in the same field are Pink Floyd’s “immersion” editions in 2011–2012, including Wish You Were Here in 5.1, remixed by James Guthrie, and the Beatles’ 50th anniversary editions of Sgt. Pepper’s Lonely Hearts Club Band and the Beatles, with both a new stereo

7 Steven Wilson has remixed recordings by, for example, Gentle Giant and Jethro Tull, Jakko Jakszyk for Jethro Tull and ELP.
mix and surround mix produced by Giles Martin, released in the summer of 2017 and autumn of 2018, respectively. All of these modern surround mixes of older material offer different solutions to transforming the original media (often analog tape recordings) into a digital remix. Often the process includes—at digitization of the analog material—a painstaking rebuilding of the recording track by track, to find the bits and pieces bounced down in an early stage of the recording process, as in *Pet Sounds*, and separate them from other audio sources in order to be able to make them discrete in the mix. We might also say that the job involves not only transforming the analog multitrack to a chosen digital codec, but simultaneously managing the transformation of the original compositional idea, designed in a specific time-bound media technology, into something sonically different that still retains a common identity as a work—a musical representation—with the original recording. *Pet Sounds*, again, is a good example. It was originally conceived, written, recorded, designed, and produced as a mono recording, and is famous as such. Transformed to surround, it became a very different “work of phonography,” to use a common term. Something similar could be said of the surround version of *Sgt. Pepper*, which is a transformation of the original idea of the record as a mono mix. The richly orchestrated King Crimson album *Lizard* from 1970 is on another level: according to Steven Wilson, who remixed it for 5.1, *Lizard* was “too big for stereo to contain,” and therefore was better realized as music in a surround mix. But this raises questions about the recording as a representation of an original musical idea and of musical spatiality. This is a fundamental problem that deserves discussion, especially in the case of the King Crimson surround remixes.

**Some stylistic traits in the music of King Crimson**

When King Crimson produced its debut LP *In the Court of the Crimson King* in 1969, it established an advanced and eclectic freeform style with virtuosity and surrealistic lyrics that set the musical framework and standards for progressive rock in the 1970s. Genre-wise, this music is not easily categorized by style and deviates from the progressive rock mainstream. Its points

---

8 General information about the mentioned releases can be found on the Internet.
9 The work of phonography concept has generated an extensive discourse, stemming from Evan Eisenberg (1986) and including e.g. Lee B. Brown’s (1998) general definition: “sound-constructs created by the use of recording machinery”.
11 Contemporary groups like Genesis, Yes, Gentle Giant, Van Der Graaf Generator, and Black Sabbath were all much inspired by the high level of musicianship, heaviness, dynamics, sophistication, and the dark and surrealistic lyrics and atmosphere on *In the Court of the Crimson King*: probably one of the most influential debut albums by a completely new and unknown band in the history of rock music.
of departure included the advanced rock music by the Beatles, the Moody Blues and Jimi Hendrix, as well as modern jazz (bebop, free form, cool), influencing the albums *In the Wake of Poseidon*, 1970, *Lizard*, 1970 and *Islands*, 1971. Serious rhythmic influences in the form of irregular time signatures and metric patterns came from the modernist art music of Béla Bartók, Gustav Holst and Igor Stravinsky—most obviously in their adaptation of Holst’s “Mars, Bringer of War” as “The Devil’s Triangle,” but also identifiable in much of King Crimson’s later music. The band's overall stylistic development also included many personnel changes, with guitarist Robert Fripp as the sole constant, and a varied orchestration with unusual instrumental setups. Besides the standard rock setup of drums, bass guitar, vocals, electric and acoustic guitars, the ensemble sound also included, at various times, woodwinds, synthesizers, mellotron, piano, violin, percussions, stick and frippertronics. The instrumental aspects of the music were often prioritized over the vocal parts.

King Crimson’s music developed in the mid-1970s into a hard rock style completely devoid of blues but with an art music, free jazz and fusion attitude. It built on sharp and harsh contrasts between percussion and pastoral elements on violin, flute, and acoustic guitar with heavy, rhythmically complex, asymmetrical (sometimes atonal) electric guitar riffs. The albums *Larks’ Tongues in Aspic* (1973), *Starless and Bible Black* (1974), and *Red* (1974) became highly influential.

The band re-formed in 1981 with a complex, new wave- and minimalism-influenced style and within a short span of time (1981–1984) released the albums *Discipline*, *Beat* and *Three of a Perfect Pair*. It re-formed once more in 1994 as the so-called “double trio” (two each on guitar, drums, and bass), releasing the stylistically diverse CD *Thrak* in 1995. Over the years, King Crimson performed and recorded at consistently very high levels of musicianship and audio technology, while also exhibiting high productivity both live and in the studio—a feat admired by many musicians.

The surround mixes in the box sets: Methodology

The King Crimson box sets are rich in recorded and additional material and could well be used for many different and detailed studies: of the music itself, of studio recording practice, of performance practice, etc. All of the box sets include both studio and live recordings. The recordings include original mixes as well as remixed versions and alternate takes and mixes.

---

12 On the *In the wake of Poseidon* LP (1970).
13 Many elite UK and US musicians have been members of the band over the years.
14 The band has since been active in periods, releasing the studio albums *The ConstruKction of Light* (2000) and *The Power to Believe* (2003) and many live recordings. King Crimson has been a world touring band since 2014.
and use different sound codecs—mixed in mono, stereo, and surround. Assorted quadrophonic mixes from the 1970s are also included, possible to listen to in a 5.1 setup. Essential for this discussion are the surround mixes in LPCM 5.1 and other codecs, both uncompressed and compressed, including e.g. DTS HD surround. However, I will principally turn to the mix using the best, i.e. uncompressed, codec available, and will not compare the same recording or mix in different codecs. The surround mixes I will discuss are a few representative examples, highlighting tendencies only. The method used in this investigation was a careful listening to the selected 5.1 mixes in an adequate, calibrated 5.1 setup, balanced primarily for music listening (Holman, 2008, p. 36–45). The listening experience is subjective but informed by my knowledge of the band’s repertoire, my hearing and ear training skills, and my personal experience of surround mixing. The focus is on the balance and placement of the instruments and sounds in the surround field in the selected recordings. The latter part of this essay presents the results of my listening, more so than the process itself.

**Aim and research questions**

My aim is to analyze the mixing strategies and the conventions of surround mixing of this older, originally analog recorded musical material from both a technological and a musical perspective. I want to show how emerging conventions for surround mixing are connected to production ideals of *documentarism* and *hyper realism*. Specifically, I ask: How is the music mixed, panned, and spatialized in the surround field? And how do surround mix conventions for studio and live recordings differ? In particular, five songs on six recordings mixed in surround will be analyzed, presented here in the order they were first released:


15 Compression is just one variable of sound quality in different codecs. The quality of the master is obviously very important.

16 I have a hearing range of about 8 Hz to 19700 Hz.
• “The Night Watch.” From the “pseudo” studio LP *Starless and Bible Black*, released in March 1974. Major parts were recorded live on tape with studio overdubs. The surround mix was produced in 2014 in LPCM 5.1 and DTS-HD Master audio 5.1 and released on the *Starless* box set, 2014.


**Theory: Representation and spatiality**

Here I would like to introduce two useful analytical models: the first for analyzing recorded music as representations, and the other for analyzing spatiality in music from a meta-perspective.

The representation model (Burlin, 2008, 2012, 2015) develops and combines previous theoretical positions regarding the ontological status of music recording (Bennett, 1983; Tagg, 1979; Brown, 2000). It postulates that a work of phonography (Brown, 2000; Simonsen, 2008; Burlin, 2008; Ternhag, 2009) can be ontologically understood as a representation (or parallel representations) of historical, aesthetic, musical, and descriptive/prescriptive parameters in the recording. These aspects fall into four main categories, numbered I–IV: historical representation, aesthetic representation, musical representation, and hyper notation (Burlin 2008, 2012, pp. 239–272). Of immediate relevance here are II, III, IV.

1. Historical representation.

II. Aesthetic/idea representation. The recording has been created according to a production ideal that may or may not be especially pronounced: either documentarism or hyper realism. Documentarism is

17 Historical representation. The recording represents something that has happened: a documentation of one or more musical events, at one or several moments, under specific circumstances (e.g. in a certain kind of spatial environment and with certain equipment).
an ideal of auditive realism in recording, with minimal editing and processing of the sound. The goal is for each instrument and voice to sound as it does “live.” This ideal has been normative in classical music, folk music, and jazz. The documentarist recording should be a true auditory reflection of the musical repertoire, the style, and the musicians’ technical ability. Hyper realism (hyper = above, beyond), means using all available technical tools and methods to produce a satisfying recording, even if this results in music that is harder to realize live. The ideal was established in 1960s pop, with the Beatles’ *Sgt. Pepper* album as a prototype. Hyper realistic recordings are produced in relation to a wider field of recorded music, which has gradually convinced audiences to perceive highly produced music as “realistic.” Hyper realism is associated with rock and pop music, but both documentarism and hyper realism are production ideals that cross genre boundaries.

III. Musical representation. A number of musical parameters (like beat, harmony, melody line etc; see Burlin, 2012, p. 247–248) can be represented on music recordings. Two important parameters are technical sound, which can be clearly associated with the technology employed, including mixing conventions, and spatiality, which is explained below.

IV. Hyper notation. Works of phonography—in any genre—can represent realized ideal versions of a scored or unscored work. The work of phonography will be assigned this status by composers, musicians, producers, or listeners. As an ideal version of a work it becomes something more than “music.” It becomes a normative, prescriptive instruction: a hyper notation (i.e. for how the work will be realized in future performances and recordings).18

The *spatiality model* (Burlin, 2012, p. 133–134, Burlin, 2015) complements category III, musical representation, as a way to analyze spatiality—or spatial design—in (recorded) music. Spatiality has been recognized as an enormously important aspect of music (see for example Moylan, 1992, 2015; Moore, 1993, 2001; Gibson, 1997, 2005; Doyle, 2005; Brøvig-Hanssen & Danielsen, 2013; Kraugerud, 2017; Vad, 2017). The spatiality model is largely independent of these discussions, but has some elements in common

18 The representation of hyper notation becomes particularly clear in notated music, as the difference in information content between the score and the musical realization on the recording appears clear (i.e. the meaning of interpretation). Likewise, the higher the degree of hyper realism, the greater the possibility that the work of phonography will work as hyper notation, as the recording tends to be a unique representation of the work. However, just because the work of phonography is characterized by hyper realism does not guarantee that it also functions as hyper notation, and a work of phonography characterized by documentarism can also function as hyper notation, if given the status as an good example of recorded music, prescriptive or normative.
with them. In this model, there are (or can be) four main recorded “spatialities” in a chosen piece of recorded music—overlapping, but principally distinct. Spatiality is understood as a symbolically fundamental aspect of all music. As such, it consists of much more than room acoustics; it also encompasses distances and relationships between instruments as well as metaphorical spatiality, which may signify inner (that is, mental states) or outer (that is, “the world”) spatialities in human life in the broadest sense. The spatial conditions and parameters may be sorted into the following scheme:


c. External spatiality: metaphorical signifiers of the world outside the music (voices, water, birds, cities, caves, etc.)

d. Internal spatiality: sounding aspects of the recording, signifying introversion and human mental states, including “unnatural” reverb or effects.

Some observations on the 5.1 mixes

With respect to the theories and models above, I would like to make some observations and draw a few conclusions regarding my examples: “Lizard,” "Larks’ Tongues in Aspic, Part 1,” “The Night Watch,” “Trio,” and “One Time.”

According to Holman (and in line with production common sense) there are two basic approaches to surround mixing: direct/ambient and direct-sound all around (Holman, 2008, p. 107). The two approaches are vaguely related to “live” vs. ”studio” as recording environments and concepts, spaciousness vs. envelopment (Holman, 2008, p. 187), and also to the production ideals of documentarism and hyper realism. Recordings with a live—and simultaneously “documentaristic”—production ideal, such as surround recordings of orchestral classical music, are often made and mixed using a direct/ambient approach. The sonic impression is of sitting and listening in a concert hall or a church, with reflected sound coming from behind. The King Crimson surround live recordings are mainly mixed in this way, but some of them are mixed more in the manner of surround studio recordings, with direct-sound all around, therefore displaying a tendency toward hyper realism. In general, surround mixing in rock and pop music is related to the established conventions of the stereo “diagonal mix” (Dockwray & Moore, 2010). The King Crimson surround mixes, both live and studio recordings, are no exception. That is, the drums are located somewhere in the middle of the surround field, the bass guitar somewhere near the drums, guitars are sharply panned to the left or right, pianos are also often panned to the left or right,
and finally the voice or voices are located in the center (speaker)—or on the opposite side in the surround field, coming from behind the listener.

Let me turn first to the live recordings. My first live example tracks (all with the same lineup of Robert Fripp on electric guitar, John Wetton on bass guitar and vocals, Bill Bruford on drums, and David Cross on violin and mellotron) are from a sonically superb live recording from Mainz Elser Hof on March 30, 1974, which includes versions of “The Night Watch” and “Trio.” This surround mix was created by Steven Wilson by combining the board recording with a live bootleg from the same concert, “resulting in a prime seat about ten rows back at one of the best shows this band ever played,” according to engineer David Singleton.\(^\text{19}\) The mix of the whole composite recording seems to come very close to the original live event: the direct/ambient sound approach is used for the surround mixing and it is therefore also clearly documentaristic. The listener is virtually placed in the audience, in front of the stage, probably close to the position of the bootlegger. However, in the surround mixes of the live versions of “The Night Watch” and “Trio” from Mainz Elser Hof the tendency towards hyper realism is strong. “Trio” as a live improvisation encircles the listener, who virtually flies, placed somewhere between the stage and in the concert arena. This effect is technologically dependent: with two analog recordings from the same concert, double tracked together in the digital studio environment, it was possible to create the track as a piece of documentaristic hyper reality. It is therefore possible to experience the recording as a real concert—nearly as if it were happening now.

Let us compare this to another recorded live concert, this one from the Amsterdam Concertgebouw on November 23, 1973 (mix by Steven Wilson). This is a well-recorded professional multitrack recording, including versions of “The Night Watch” and “Trio”, which was broadcast and also mixed in quad by the engineer George Chkiantz.\(^\text{20}\) Here, however, the surround mixing approach tends even more towards direct-sound all around. The listener is virtually situated in the middle of the stage, surrounded by the musicians, fully enveloped by the music. The aesthetic ideal of hyper realism is more apparent here. The other available mixes of the concert presents different perspectives (there exist at least two different stereo mixes in 16/44.1 and 24/48), perhaps tending more towards documentarism.

\textit{Starless and Bible Black}, King Crimson’s 1974 LP, is an interesting production that links together live and studio recordings in a subtle and integrated way. This LP was conceptualized as a “studio recording,” but was mainly recorded live on tour, with audience sounds edited out.\(^\text{21}\) Although not mar-

---

\(^{19}\) Singleton, David (2014) \textit{The tale of the tapes Continued}. Starless box set.

\(^{20}\) Ibid.

\(^{21}\) Compare with the King Crimson LP \textit{USA}, 1975, which was recorded live but had studio overdubs and was marketed as a live LP.
keted as a live LP, it was no secret that most of it was in fact recorded live, on the same tours as the Amsterdam Concertgebouw concert—many of the recordings for *Starless and Bible Black* came from that particular concert—and Mainz Elser Hof. *Starless and Bible Black* was made up partly of pre-written songs, partly of freeform improvisations such as “Trio.” The latter was an improvisation recorded in concert—the location is not known, but there are similarities to the versions recorded at Concertgebouw and Elser Hof—and released as-is, minus audience sounds, on the record.\(^{22}\) Out of eight tracks on *Starless and Bible Black* only two were recorded primarily in the studio. And at least one of these two tracks—the Rembrandt-inspired single “The Night Watch”—had an introduction taken from a live recording, maybe Concertgebouw, although the principal part of the song was recorded in the studio (George Martin’s AIR). As a surround mix, it uses the direct-sound all around approach and it is enveloping: the listener is almost part of the band. Thus on *Starless and Bible Black*, the band seems to strive in its stereo mix to create a recording somewhere between documentarism and hyper realism: a record that captures the best of live performances and presents them as studio recordings. The surround mix of “The Night Watch,” especially, underscores the listening position in middle of the band—the representation of the drums is very realistic—with a direct and sharp sound that makes the tendency towards hyper realism more pronounced.

The three studio recordings in surround here—“Lizard” from the *Lizard* LP (1970), “Larks’ Tongues in Aspic, Part 1” from the *Larks’ Tongues in Aspic* LP (1973), and “One Time” from the *Thrak* CD (1995)—have both similarities and differences as surround mixes.

Of these, “Lizard” is the longest at 23:35, sonically probably the most complex, and one of the first side-long progressive suites.\(^{23}\) It was recorded with a large and unstable lineup (Robert Fripp, guitar, mellotron, electric keyboards and devices; Andy McCulloch, drums; Gordon Haskell, bass guitar and vocals; Jon Anderson, vocals; Mel Collins, saxes and flutes; Robin Miller, oboe and cor anglais; Mark Charig, cornet; Nick Evans, trombone; and Keith Tippett, piano). The track is an extraordinary fusion of free jazz and progressive rock, characterized by contrasts between strict pre-written arrangements and free jazz piano and wind instruments improvisations. The extraordinarily creative surround mix was produced by Steven Wilson in 2009.\(^{24}\) In the difficult process of remixing the multitrack tapes to new stereo and surround mixes, he found that the 16 tracks originally used for the recording of “Lizard” were in practice over used: one track could have many

---


\(^{23}\) Jon Anderson, who was the guest singer on *Lizard*, took the ambition and ideas back to to his group Yes with “Close to the Edge” (1972), and it is probable that both ELP’s “Tarkus” suite (1971) and Genesis’ “Supper’s Ready” (1972) owe a debt to *Lizard.*

separate takes of different instruments, followed by bounces of used or unused takes, etc., and it was not always obvious what should be included in the remix. He had to use the original mix from 1970 as a template, or, with the vocabulary used here, as a hyper notation for the new mixes (Smith, 2017).

In Wilson’s surround mix the instruments do not have fixed locations in the surround image. Instead, they have great mobility within it. The instruments move around in the surround field between the four parts of the song and also within each part. For example, the piano, which is mostly located in the left and center speakers, later moves back to the surround rear speakers. These are used as much as the two front speakers and the center speaker, and not only to reproduce reflected sound. In “Lizard” the surround speakers are on an even footing with the front speakers, and in practice used to great effect as the primary location for some of the instruments: for example, double-tracked drums and woodwinds. The drums in the front are partly in parallel with the back drums, which perform accentuations. The main vocals—bright and shimmering—are located in the center speaker with processed voices sounding from the rear speakers and creating a spacious effect. The center speaker is also used for purely instrumental parts, for example the cornet. High-frequency mellotrons—a trademark of the band—are spread across the surround field and create an almost breathtaking ambient feel. The wind instruments throughout the track are often divided between the front and the rear speakers. Together they produce a full surround music experience: an impression of being inside the music. The mobility of the instruments is even more pronounced at the end of the track, in the “Big Top” section, when the electric guitar—which is more or less absent from most of “Lizard”—plays a solo that starts in the rear speakers and then circles clockwise around the surround field with the bass drum and bass guitar playing low bass notes in the left and center speakers. The concluding “circus” track is wildly panned around, perhaps like a carousel: the melody, played on the mellotron and a piano, moves from the left to the right speaker and then circles clockwise. At the end of “Lizard” all the instruments, even the drums, are in motion around the surround field before they fade out and disappear.

“Larks’ Tongues in Aspic, Part 1” was recorded when King Crimson was a five-piece in 1972—1973 (Robert Fripp, guitar; Bill Bruford, drums; Jamie Muir, percussion; John Wetton, bass guitar and vocals; David Cross, violin). The track is a collective composition and it marked a new King Crimson fusion: compared to “Lizard,” it has a completely different timbre, characterized by violin, up-front electric guitar riffs, and sophisticated percussion in a multi-sectional form. It is an extremely dynamic, instrumental work. The multitrack recording includes a huge palette of non-musical sounds as well as voices and some percussive sounds, reminiscent of nature. “Larks’ Tongues in Aspic, Part 1” is a mind-blowing recording in stereo and even
more so in surround, with a perfectly balanced mix: all of the surround field, direct-sound all around, is used to great effect. The many percussion instruments are panned mainly behind in the two rear speakers. The solo violin parts are placed in the front/middle of the surround field, between the center speaker and the rear speakers, and the most subtle percussive parts—evoking chilling bird and wind sounds but belonging to Jamie Muir’s advanced palette of percussion instruments—sound alongside whispering and speaking voices. It opens the door to a different musical-metaphorical world: a mental landscape of sounds. Through the creative process of remixing to surround, “Larks’ Tongues in Aspic, Part 1” becomes a more “real” hyper realistic representation of the music than the stereo mix. Perhaps it is also possible to hear it as a hyper notation: a truly prescriptive sonic notation for the work of phonography that is “Larks’ Tongues in Aspic, Part 1.”

“One Time” was recorded in 1994 with a six-piece setup, the so-called “double trio” (Robert Fripp, guitar; Bill Bruford, drums; Pat Mastelotto, drums; Tony Levin, bass guitar and stick; Trey Gunn, bass guitar and stick; Adrian Belew, guitar and vocals). The album Thrak was recorded on 48 tracks at Real World Studios. The album unmistakably has the sound of its time, with lots of digital processing on all the instruments: even the guitars and basses were recorded or processed in stereo. This, according to Jakko Jakszyk, who re-mixed the master tapes, “all added to the mush and lack of clarity in the original stereo, so in surround I made these either very narrow or used just one side” (Kelman 2015). Jakszyk produced the new surround remix with the aim of separating the instruments and creating more distance between them. He also added parts that had been recorded but not included in the original mix, and were now heard with more clarity. He strictly limited the use of spatiality, with more emphasis on mono reverbs, imitating the mixing and production practice of the 1970s. The new stereo mix was produced with the surround mix as the sonic template, hyper notation (Kelman, 2015). Compared to the Larks’ band and track—with its great dynamics in the music, from the utterly pastoral violin parts to heavy metal riffs—the Thrak instrumental setup was more limited and more severe, with its double “rock trio” instrumentation. However, “One Time” is a richly spacious recording and mix, an orchestrated, ambient track in the double trio format. With a simple drum pattern mixed behind the sweet spot, in the surrounds, and with an added, distinct, longer mono echo of the drums right in front of the sweet spot, it produces a particular spaciousness of marked sonic beauty. One of the two bass lines is in the front center speaker while the second is behind the listener in the rear speakers, both contrapuntally played. The clear frequency distribution between bass and treble, for example, the two bass guitars and the high treble guitar lines—strongly contributes to the spacious feel. The vocals are in the front center, with guitars left front and right front with spacious delays, creating an ambient, floating feel. “One Time” is mixed using surround principles partly similar to “Larks’ Tongues”—
inspired by that time—and with its direct-sound all around mix, it is distinctly hyper realistic. The two tracks share a similarity of approach in that drums and percussion are much important in the music and are clarified in the mix through distinct separation. The two drummers and percussionists play with and against each other, in cross-rhythms, polyrhythmically. This is distinctly spatialized in the surround mix of “One Time” and the whole surround mix of Thrak, so it becomes obvious that there are two percussionists playing together in the same room, in the same “acoustic spatiality” (i.e. the studio).

It is relevant to ask why the spatial reinterpretation of the King Crimson stereo recordings was produced: why surround in recorded music? In reasoning inspired by Gilles Deleuze’s and Félix Guattari’s theory of a complementary relationship between music and spatiality, Peter Doyle argues that the spatial aspect of recordings must be related to music’s function as art and communication. Music is part of a process of continuous creation and dissolution of space, or “territory.” Musical spatiality is thus a “territorializing” expression of individual or group communication and struggles for space, a perspective that I find reasonable (Doyle, 2005, p. 17).

Here I want to draw a connection to the spatiality model. In “Lizard”, for example, the dynamic surround mix interprets and comment on the lyrics, creating a three-dimensional sonic world where the listener is situated in a fantasy landscape. The distribution of the instruments, such as the high-frequency mellotrons spread widely across the speakers, creates the sonic impression of an open landscape, or perhaps a battlefield. The acoustic and musical spatiality together evoke an external spatiality. The spinning circus melody at the end is another signal of the environment, another “external spatiality.” The creative use of different percussive sounds in “Larks’ Tongues in Aspic, Part 1”, which obviously reference sounds from the natural world, also clearly generates an “external spatiality,” signifying the world outside the music. The track is a spacious recording and the surround mix underscores inherent sonic and spatial qualities in the original recording not fully realized in the original stereo mix—instead re-creating spatiality and creating a new hyper notation. In the surround mix, it is possible to simultaneously hear many layers of meaning: the music performed with an acoustic, a musical, and an external spatiality. Maybe it is the internal spatiality we hear in the spoken voices: surround as an advanced sound technology whose principal function is to realize the spatiality inherent in music, to open up the analog tape recordings for a truer form of communication.

It is thus probably clear that new conventions for surround mixing of "studio" or "live" recordings of music are established slowly through the digitized re-interpretation of analog recordings, and it is possible to hear this process in real time in the remixed King Crimson recordings. The "studio" and "live" surround conventions examined here—related to the direct-sound all around and direct/ambient approaches, respectively—are principally developed from the diagonal stereo mix convention of popular music, but are
freer, especially in the longer studio tracks. The live tracks, on the other hand, are conceptualized as part of the outside world, with external spati-ality. Therefore, the music of the band is mainly panned in the front. The sounds of spaciousness, reflections, and the voices of the audience envelop the listener from behind.

The mixing conventions whose emergence is studied here have much in common with surround mixing of other recordings in the same genre or related genres, as in the examples mentioned above, but are more articulated: therefore, it is possible to listen to the King Crimson surround mixes as something like sonic templates, *hyper notations*, for surround in music.

**Acknowledgments**

Thanks to Robin Blanton and the peer reviewers for constructive comments on the text.

**Bibliography**


Toivo Burlin: The creation (and re-creation) of virtual spatiality: Surround sound mixing in King Crimson’s back catalog box sets


Discography

King Crimson, In the Court of the Crimson King DGM/Panegyric, 2010.
King Crimson, “Larks’ Tongues in Aspic (part 1)” Larks’ Tongues in Aspic: Complete Recordings DGM/Panegyric, 2012.
King Crimson, the Road to Red DGM/Panegyric, 2013.
King Crimson, “The Night Watch”, “Trio” Starless DGM/Panegyric, 2014.
King Crimson, On (and off) the Road DGM/Panegyric, 2016.
Photo B. The Erling Persson Atrium at KMH during the ARP2017-conference.
(Photo credit: Eleonor Gislason-Ferrari)
Carlos Andrés Caballero Parra: The recording and sound aesthetics of tropical music in Colombia

Abstract

The record production made in Colombia during the boom of danceable tropical music between the 1960s and the 1970s had a prominent transcendence in the historical development of the Colombian music industry. On the other hand, the phenomenon of tropical urban music of today is similar to that of more than 40 years ago, considering that the two moments take place mostly in the city of Medellín, and are more related to sociocultural situations than to the purely musicological ones. Thus, this paper aims to show the points of coincidence of both eras from the study of their record productions versus the commercial and sociocultural circumstances of these genres.

Introduction

The city of Medellín distinguished itself for being the epicenter of the music and record production in Colombia between the mid-fifties and the end of the seventies of the twentieth century. Such production reached its peak in the 1960s and 1970s with the so-called “paisa sound” (Wade: 2000, p.189), characterized by youthful and irreverent groups [1] who mixed rock with the traditional rhythms of Colombia’s Atlantic coast; during this period, a whole musical and social sound revolution was created that revolved not only around music, but represented a universal feeling of change.

Thus, this era was characterized by common elements of encounter around the activities related to record production, which marked a particular sound aesthetics given not only by the technological scope of the sound recording equipment, but by its social characteristics, that allowed a third world country to position a sound that crossed the borders with its phonographic productions and seized the whole planet with its artistic manifestations through its most recognized tropical music groups.

The sound aesthetics of the music that is currently generated in this same epicenter (Medellín) has elements similar to the previous period that influence its outcome from both the technological and social points of view, alt-
hough, without doubt, with very different aesthetic meanings. However, the cultural phenomenon is very similar: young people who make music that is internationally accepted as an innovative commercial offer and to which popular artists of great commercial success have joined, but that, for traditional artists, folklorists, academic musicians and cultural researchers causes more animosity than admiration.

Taking into account the above, it is possible to conduct a cultural parallel between the phenomena of the Antioquian tropical music (the Colombian department whose capital is Medellín) of the sixties and seventies and that of current urban music, and demonstrate the influence of technology in the aesthetics of the resulting sound of the productions. Additionally, the perceptions that society of both epochs has had of the artists and the characterization of the similarities must be pointed out.

To answer these questions it is necessary to make an analysis of the aspects that stood out in each of the two periods.

*The tropical music in Colombia in the 1960s and 1970s*

This music appeared after the rise of the music of the same genre played in the style of the American big bands that was performed by the orchestras of composers and arrangers such as Pacho Galán, Edmundo Arias and Maestro Lucho Bermúdez, “The most influential musician [...] not only in the coastal region, but, even more, in the cities of the interior [of Colombia]” (Wade: 2000, p.111). Maestro Bermúdez was hired in 1946 by RCA Victor to record some sixty of his creations –*cumbias, porros* and *gaitas*– in the city of Buenos Aires, with the purpose of boosting the Spanish-speaking market.

Colombian record industry, which had begun in the 1930s in Cartagena de Indias, on the Atlantic coast (Jaramillo: 2017), showed an exponential growth at the end of the 1950s in Medellín, a city that was glimpsed at those times as the country’s industrial capital.

Thus, Maestro Bermúdez penetrated Colombia’s high society with music inspired by the Caribbean coast folklore, and he and his fellow composers – Pacho Galán, Edmundo Arias and others – put her to dance.

With the arrival of “*costeño*” music [2] to social clubs and live radio stations between 1950 and 1960, the three big Colombian record companies, Discos Fuentes, Sonolux and Zeida, were created in Medellín [3]. Each of them, in its own way, was the protagonist of the tropical sound revolution of the sixties and seventies.

The civil unrest of May of 68 in France marked an important change in the way of making music: irreverent young musicians influenced by the Anglo-Saxon rock and roll and the twist replaced the big bands by groups of small format that included electric instruments such as the guitar and the organ, taking advantage of the technological advances.
Thus, it was how they, with the influence of this new sound, were the protagonists of the spectrum of the musical industry from the end of the sixties until the beginning of the eighties, and became a sonic and cultural imaginary that is still enjoyed in Colombia and in other Latin American countries.

Groups of purely urban essence like Los Teen Agers, Los Golden Boys, and New Star Club in the sixties, and Afrosound, Los Hispanos, Los Graduados, and Fruko y sus Tesos in the seventies [4], with their tropical danceable sound, revolutionized the Colombian musical industry and took it to its heyday.

The tropical urban genres of the 21st century

Between the end of the 20th century and the beginning of the 21st, another urban musical phenomenon was started in Puerto Rico: reggaeton (Carballo Villagra: 2006, p.33); this movement had varied musical influences: from the Caribbean region, with genres such as the Jamaican dancehall reggae; from the American rap of the nineties (Marchal, Rivera and Pacini: 2010, p.3); and from the Panamanian sound of the eighties. Additionally, it was closely linked to the new technological advances and the era of digital audio through the use of samplers and loops with which the rhythmic bases were created. As of this paper, this genre –reggaeton [5]– and its derivations will be named “urban tropical music”.

Medellín was also the protagonist of this revolution; the great radio communication media turned to this genre and changed their sound spectrum. Thus, a large percentage of the stations began to program reggaeton only, and, as a consequence, the rest of the musical genres, with the exception of vallenato, as well as the traditional groups, were relegated to bygone hits [6].

Artists such as Daddy Yankee, Wisin y Yandel, Don Omar, and Tego Calderón, among others, who were the great exponents of this genre at the beginning of the first decade of the 21st century, established a close relationship with Medellín by turning it into one of their centers of operation; there, they not only had a great popular acceptance through the support of the media, especially the radio, but also found a society cultured by the influence of the drug trafficking mafias, very similar to the gang and prisoner movements that were part of the sociocultural roots that gave rise to this genre (Carballo Villagra: 2006, p.91).

In the current decade, artists from Medellin have become the main protagonists of the urban tropical musical; they tour around the world and are very much coveted by the Latin artistic community to make collaborative performances. Their success was evident in both the Latin Grammy Awards in Las Vegas (2017) and the Anglo Grammy Awards in New York (2018), where J Balvin and Maluma were awarded.
Methodology

Despite the time span that separates Medellin’s production of urban tropical musical, the two periods surveyed have had, for the music production and creation, an influence not only from the commercial, social and cultural points of view, but also from the technological one. For this reason, it is proposed to make an analysis of the points of view related to the characteristics of music production.

To better understand why a city like Medellin has become a benchmark for record production in Latin America, it is necessary to establish some common parameters and meeting points. For this, it is proposed to make an analysis whose starting point is the revision of general aspects such as the message and content of the songs, the influence of sex and dance, the condition of the record industry, and a musicological examination, followed by an analysis of record production from two aspects: audio technologies, and recording methods and music production. Thus, a sample of representative works of six artists of each one of the two periods is taken, complemented with some extracts of the interviews made to the protagonists of the first period.

Analysis between the two epochs

The groups selected from the first period are the following: Los Teen Agers (1956), Los Golden Boys (1961), Los Hispanos (1964), Los Graduados (1969), Fruko y sus Tesos (1970), and Afrosound (1973). And from the second period, the Puerto Ricans Daddy Yankee (2002), Don Omar (2003), Wisin y Yandel (2000), and Tego Calderon (2003), of great acceptance and influence in Medellin, and the locals J Balvin (2013) and Maluma (2015) [7].

The message and the content

Although there are similarities between the message and the content of the songs, and in the message that the artists send through their staging and artistic concept in both periods, there are also profound differences.

The lyrics of the tropical groups of the seventies have a somewhat romantic and naive message, but also a lot of innuendo, with great sexual content, although without falling into vulgarity or explicitness, something very typical of the Medellin’s culture; in fact, they reflect a historical moment of change in the thinking of humanity (the counterculture), which allowed to say in a song what in plain conversation was more difficult to do. As for the lyrics of the current tropical urban genres, these are characterized by their high sexual content and the explicitness of their message, although some artists occasionally opt for romantic situations.
Costa Rican researcher Priscilla Carballo notes in this regard:

Therefore, the vision of women as a ‘sexualizable’ object is reproduced, since their place in music production can only be a means of producing pleasure. This ‘hipersexualazing’ tendency that raises this rhythm is understood more clearly if it is placed within the market of symbolic capitals, where the body is the fundamental referent, and the female body in particular is eroticized to capitalize it. (Carballo: 2006, p.38)

It is clear that this situation is not alien to the cultural and social reality of the artists; in their lyrics they reflect themselves and are not telling anything beyond of what they have experienced. In this way, they become witnesses and speakers of the harsh reality of the society to which they belong, in which machismo, lack of opportunities and crime mark the lines that inspire their lyrics.

On the other hand, the image that each period presents from its artistic approach through its iconographic representations and staging is quite similar in the background, but differs in its form. Thus, the groups of the first period were characterized by having in their record jackets, in addition to jocular scenes that sometimes narrated surrealist stories, a high sexual component marked by the representation of women as objects of desire; in this way they aroused interest in a predominantly conservative public.

With respect to the current period, although artists have a marked tendency towards sexuality, women no longer appear in the booklets of their CD’s and in the video clips; this, by no means, implies that they are not even more misogynistic than the former; on the contrary: for them, women are already subjugated and are always seen as an object of beautification, accompanied, what else, by jewelry, luxurious cars and stunning mansions.

“The artistic power of suggestion, where the refined audiovisual treatment of eroticism resides, is lost to show female body parts: breasts, buttocks and even pornographic nudity” (Lavielle-Pullés, 2014, p.123). In the graphic material of the first period, women were represented in a time of feminine liberation; and in the current one, they, in artistic and aesthetic terms, have already been subjugated and reified.

The influence of sex and dance

Besides the sexual content in the lyrics of the songs and in the graphic material of the two periods analyzed, body and dance have had a great preponderance. It is known that the latter has ancestral roots and that it is a manifestation that has been present in most cultures, in a particularly marked way in the ones of Latin America located in the tropical zones. The movement of the body linked to reggaeton “translates into a particularly sensual dance called perreo, which imitates erotic positions and has come to be considered the equivalent of the lambada of the eighties” (Galucci: 2008, p.86).
It is seen, then, how the relationship between dance, body and sex are intimately linked and are a manifestation of youthful irreverence in two different eras, spun by a common imaginary.

**The recording industry and its changes**

In regard to the state of the record industry in the first period, the big corporations maintained the monopoly of the business for a long time. Novice artists had to wait for a talent hunter or an artistic producer to watch them in a gig and propose them to sign for his company. Although today this modality still exists—the artist being hired by a record company—anyone with an unpretentious video clip on a streaming platform can achieve great popularity, without having to rely on corporate support. Even if he so desires it, can get a contract.

On the other hand, new protagonists of the music industry have shown up: the post-digital intermediaries and the digital distribution companies, which, in addition to the posting of the music on the platforms, offer services such as publishing, collection and management of the platforms income, mailing lists with positioning in social networks, and merchandising sale.

All this, added to the already complex situation of the entertainment industry, has generated the rejection of some prestigious groups, accustomed to the traditional system of record companies, publishers, and royalty management associations, which have seen their income diminished by these concepts and have concentrated on the sale of concerts. Naturally, there are demonstrations in favor of the new scheme on the part of young artists who grew up with these platforms and distribution systems, have not taken a slice from the traditional cake of the collection of royalties, and simply prefer to place their names online to generate income through their live acts. To this should be added the issues of intellectual property and copyright derived from the reproductions and the visualizations on the platforms.

It is worth mentioning the importance that the record industry has had in relation to the successful musical genres, which has always been linked to the trends of cultural consumption and the influence that this phenomenon has had on society. Whatever the technology that mediates between the consumer and the product be, the latter will be linked to the popular, to people’s tastes; in this way, the sound of a particular period becomes a combination that mixes those of each tradition with the sonority of the modern ones. In this regard, the researcher on cultural issues Ligia Lavielle-Pullès notes the following:

> The large-scale social expansion of musical consumption is strengthened thanks to media action and technological supports. This leads to the location of music in the nucleus of communicative and socializing networks, physical and virtual, especially young. In this way, musical cultures are created and
developed within the same national space, which, in turn, dialogue and feed on the popular fabric, traditional or not, which constitutes the musical culture of the people. In this way, musical consumption is catapulted where, to a greater or lesser extent, the sound tradition is mixed with the thriving sounds of modernity. (Lavielle-Pullés: 2014, p.114-115)

A brief musicological analysis

The completion of a complete musicological analysis of the two periods examined would merit a new research in another space; however, some points of the musical content of each period can be highlighted.

Music for dancing: the two periods are characterized by producing commercial dance music. Latin American people, particularly those located around the Caribbean Sea, have in their blood a strong relationship with dance and with body movement; any celebration turns out into a reason to dance, and even when there is no reason for it they dance. It is simply a part of their genetics: they’re always dancing. The researcher and philosopher Juan Parra Valencia notes in this regard the following:

This is how the forms of measurement and marking in dance movements integrate the abstract (mental) spaces of aesthetic liberation. [...] that is to say, that the ritualization and symbolization of the pulses discretizes the physiological background from which they come, so that the composition is almost always valued in melodic and harmonic terms (or what is the same, in intellectual terms). (Parra Valencia: 2014, p.124)

For this reason, in the compositions of both musical genres –tropical and urban–, the relationship between the bass –built from rhythmic motifs similar to heart beats– and the base of percussion –performed by instruments of African origin– encourages movement; for Caribbean cultures, it is impossible to listen to this music and not feel the urge to step to dancing.

Instrumental differences: by being separated 40 years in time, it is evident that the instruments used in the tropical dance music of the sixties and seventies are very different from the current ones. At the former epoch, acoustic percussion instruments, clarinets and saxophones were combined with electric ones such as the so-called “solo vox”, the organ, and the electric piano and bass; the latter was replaced by the electric contrabass (called “baby”). Given the influence of masterful performers like Jimmy Hendrix and Carlos Santana, the electric guitar was also a protagonist, to the point of producing instrumental recordings with it as a soloist.
In relation to the instruments used for the creation of 21st century urban tropical music, initially these were reduced to one: the computer; in it, the organology of the genre was summarized, and the sequences, the cues, the rhythmic patterns and the bass lines were assembled. In recent years, however, some music producers have incorporated acoustic and electric instruments.

Taking up the issue of collaborative recordings, Colombian artists of the tropical urban genre have joined with their peers in the pop genre such as Pharrell Williams, Ariana Grande, Justin Bieber, Ricky Martin and Shakira; in these productions, the use of high technology is mandatory.

The handling of the voice: as with the instrumentation, the difference is abysmal. In the first period, the voice was embodied by powerful singers, mostly tenors, with nasal characteristics typical of the phonetics of the Caribbean coast; and in the second, by rap or MC (master of ceremony) voices [8], which highlight the recited text.

Tempo, duration and structure of the songs: cumbia – from the first period – and reggaeton – from the second – share a very similar tempo: between 80 and 90 BPM. In other danceable tropical genres such as the paseo, the paseaito and the cumbión [10], tempi vary and are faster. In general, and for the analysis made of the works of the selected artists, the base tempo is very similar.
In regard to length, it ranges from two and a half to a little more than three minutes in both periods, a time span that compels to the demands imposed by radio stations; in the musical genres of *salsa*, *merengue* and *ballad*, the duration can be a little longer: around four minutes. At this point, it is important to highlight the creation of musical compilations established in 1960 by Discos Fuentes, whose particularity was given by the skill of its engineers in introducing fourteen songs –seven instead of six– on each side of a standard vinyl record (LP); this innovation, edited under the name of “14 Cañonazos bailables” (see Image 1), imposed a way of commercializing the hit songs that sounded on the radio stations during a given year; the “Cañonazos” (gunshots) are still distributed and are part of the collective imaginary and the intangible heritage of Colombian music and the record industry.

With the arrival of *reggaetón* and tropical urban music of the 21\(^{\text{st}}\) century, the length did not vary much: between a little less than three and a little more than four minutes. Significantly, these times allow programmers of radio stations to transmit a substantial number of songs between the advertising strips and get more profit from the “payola” (pay off) they generally demand [11].

Table 1 shows an average of the lengths of the most representative and popular songs by artist and period.

*Table 1. Average length of the most representative and popular songs by artist and era*

<table>
<thead>
<tr>
<th>Lengths of the songs in the two periods</th>
<th>Period</th>
<th>Artist</th>
<th>Length [min: sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First period</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1960-1970</td>
<td>Tropical music</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Los Teen Agers</td>
<td>2:45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Los Golden Boys</td>
<td>2:45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Los Hispanos</td>
<td>2:52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Los Graduados</td>
<td>2:59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fruko y sus Tesos</td>
<td>4:14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Afrosound</td>
<td>3:04</td>
</tr>
<tr>
<td></td>
<td>Second period</td>
<td>Urban tropical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>21(^{\text{st}}) century</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dady Yankee</td>
<td>3:16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Don Ómar</td>
<td>3:40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wisin &amp; Yandel</td>
<td>3:27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tego Calderón</td>
<td>3:17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>J. Balvin</td>
<td>3:26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maluma</td>
<td>3:22</td>
</tr>
</tbody>
</table>

Source: author’s elaboration.

The songs lengths are linked to their structure. In the first era they were shorter, the music arrangements included instrumental sections, and the tempo was a little faster. In contrast, *reggaetón* and tropical urban music of the 21\(^{\text{st}}\) century have fewer sections; in the last section, which is instrumental, it
is usual for the artist to greet his followers or to send offensive messages to
his competitors, a very common practice in rap or hip-hop.

Table 2 shows the musical structure of the songs by era.

Table 2. Songs structure by era.

<table>
<thead>
<tr>
<th>Song structure</th>
<th>60’s and 70’s</th>
<th>Today</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intro</td>
<td>Intro</td>
<td></td>
</tr>
<tr>
<td>1st verse</td>
<td>1st verse</td>
<td></td>
</tr>
<tr>
<td>Refrain</td>
<td>Refrain</td>
<td></td>
</tr>
<tr>
<td>Instrumental bridge</td>
<td>2nd verse</td>
<td></td>
</tr>
<tr>
<td>2nd verse</td>
<td>Instrumental section (end)</td>
<td></td>
</tr>
<tr>
<td>Refrain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd instrumental bridge (mambo)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refrain – Pregon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final mambo</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: author’s elaboration.

The record production

In addition to the sociocultural aspects that surround any musical genre, for
this analysis it is essential to establish a common axis in relation to the audio
technology used in the two periods, which originated a particular sound aes-
thetics from the use of electronic musical instruments and equipment. Al-
though in consideration of the technological advances that emerged in four
decades the idea of establishing a common line of the elements used in both
periods would be unproductive, it is possible to establish a coincidence in the
direct relationship that the sound and its particular aesthetics have had with
the technologies of every moment.

Audio technologies

The first was the age of analog audio. The record companies of Medellín
worked with the most advanced equipment available, and retained, together
with the business conglomerates of the entertainment, the monopoly of mu-
sic production. In this way, the head of the label, arrangers, producers and
engineers determined the final sound. This is how the music recording of the
city’s tropical music in the 1960s and 1970s functioned on the three largest
companies in the country: Discos Fuentes, Sonolux and Zeida.

The implementation of the first multichannel consoles, the microphones
and preamplifiers of vacuum tubes, the magnetic tape, and the electric in-
struments, in combination with those of percussion of African origin marked
a clear difference with the music of the big bands of the fifties, and created a sound that still lasts in the collective imaginary of many Latin American cultures.

The most common brand of tape recorders employed by the three companies was Ampex: single-channel at the end of the 1950s and early 1960s, and two-, three- and four-channel later. The microphones were the Neumann U47 and U67 and the RCA 44A and 77D. The consoles and the cutting machines of the master disks differed from company to company; for the latter, Zeida—which changed its name to Codiscos—used the German record cutting lathe VMS80 from Georg Neumann, and Discos Fuentes, the American record cutting lathe from Scully, with Westrex amplifiers. The operators of these cutting machines were very appreciated for their fine and specialized work; so much so, that they used to stamp their signature on each record, right on the vinyl strip between the grooves of the last song and the center label.

Among the musical instruments, similar in the three companies, were the amplification equipment for the guitar—the classic Fender Twin Reverb—and the bass—the Ampeg B15N--; and the unique Hammond solo vox organ, which played the second voice for the clarinet and even had interventions as a soloist.

The studios of this early period had a large design: ample acoustic spaces with high ceilings and isolation rooms that allowed block recording, a technique with a resulting sound very different from that achieved by overdubbing.

The second period makes its entrance, and digital audio technology commences to generate the majority of the instrumental tracks that support the vocal executions of the MCs (masters of ceremonies) using the resource of the multi-channel software of reproduction of pre-recorded sounds known as samplers or loops. Through this technique, which had been used by DJs, the new “musicians”—people with little preparation in the field of academic music—have at their disposal a composition tool based on the repetition of rhythmic-harmonic patterns, sequences in the bass, licks in the synthesizers, and a rhythmic base derived from Caribbean music. Together, digital instrumentation and voice in the style of American rap generated an aesthetic that identifies this musical genre in the world.

**The methods of music recording and production**

Although the goal of musical recording and production is always the same: to create successful commercial songs, their methods are not. In the first period, the start-up of a project was the responsibility of the artistic director of the label—who was in charge of proposing projects, artists, songs and repertoires—, the producer and the arranger—who chose the interpreters—, and the recording and mixing engineers.
According to testimonies gathered in some of the interviews, this was a time of great industrial zeal; the engineers were forbidden to enter the studios of the rival companies, because none of them wanted the engineers to see and copy their equipment. It was not until the 1990s that the recording technicians of the three major record companies that remained in the city (Discos Fuentes, Codiscos and Discos Victoria) –Sonolux had moved to Bogotá– established the first section of the Audio Engineering Society (AES) in Medellin.

The recording methods were very similar in all the companies. Their sequence was established indistinctly by the musical director, the arranger or the engineer, depending on the kind of group and the musical genre; in the first period, Los Hispanos, Los Graduados, and Fruko y sus Tesos, for example, had a musical arrangement written in scores that were given to session musicians, who, by that time, belonged to most of the traditional tropical orchestras of the city.

The recording session operated as follows: the rhythmic-harmonic basis, that is, the drums, the Latin percussion (congas, bongos, güira, timbal and cowbell) separated by gobos, the piano and the bass, plus the wind section (trumpets, trombones, clarinets and saxophones), also separated by gobos or in isolation rooms, were recorded into the four-channel Ampex machine. Then, the four tracks were returned to the console that, in turn, send cues to the vocal soloist and backup singers and recorded all the instruments and voices into the two-channel Ampex machine, while the one-channel Ampex...
machine simultaneously fed a tape delay for selected instruments or voices. That is how the final mix of the song was extracted. Image 2 shows the control room of the company Discos Fuentes in 1960.

For smaller groups such as Los Teen Agers and Afrosound, which only had drums, conga or bongos, bass, guitar, organ and vocals, the recording was also done in session. Sometimes the artists came to the studio with the songs already learned; in others, like with Afrosound, the songs were created in the same studio, from the musical ideas that sprung out from its members.

Finally, and unless it had been necessary to resort to the risky edition with a blade—applied through an oblique cut to the very master tape to correct mistakes—, the last technical-artistic step was the cutting of the master disk in the lathe.

In comparison with the recording methods of the current era, this technique was much more crafted, and allowed more demanding musical processes: the precise, round and deep sound of the bass; that of the electric guitar, without effects or distortion; that of the solo vox in duet with the clarinet or the saxophone; the grand piano and the electric double bass, in short, a sound aesthetics that still persist in the majority of celebrations of Latin American people.

Turning to the recording techniques of the second epoch, in its beginnings it was limited to the use of the computer, controlled by a team that rarely exceeded the composer, the producer and the interpreter. The convenience of digital audio allows producing tropical urban music in small spaces or in home studios. Of course, to the extent that this genre has become sophisticated, the music productions of artists such as J Balvin, Maluma or the aforementioned Puerto Ricans have reached great complexity in both the recording and the production techniques, and increasingly include new elements that facilitate the interaction between audio technology, musical instruments and performers; the constant sub bass and the automatic voice tuners are some examples.

Conclusions

From the aforementioned, it can be said that there are enough elements that allowed us to draw a parallel for the analysis proposed. To make an analysis of the record production in these two periods, it is necessary to know what happened at a sociocultural level and to understand why these movements gestated important trends and reached international musical relevance.

Today, at a Latin American level, Colombia is a benchmark for record production, and Medellín in particular for tropical urban genres. A commercial success like Despacito, a song that has had more than 4,700 million
visits on its YouTube page since January 2017, was produced by two Colombians: Mauricio Rengifo and Andrés Torres.

Regarding the technological influence, it is clear that in both periods music production was influenced by the evolution of the recording equipment and the instruments; the two moments pick up the latest technological advances and involve them in production; this fact marks its characteristic sound aesthetics and makes it recognizable and time enduring. From the warm sound of analogue tapes to the premeditated sequenced loops, technology has been transcendental in the evolution of each era, and we, composers, producers, arrangers and interpreters, are the result of the different technological advances; without them we would not be professionals of the record production.

Finally, in society there have always been anti-cultural and counterculture movements; this is part of the evolution of humankind, and both in the sixties and seventies, with the social revolutions that marked that moment, as it is happening now, society is still full of prejudices and traditionalisms. It is very difficult for a classical and academic musician to move from his comfort zone, but it is easier to judge the new musical expressions from that very same place. It is not a matter of becoming artists and producers to the new sound paradigms whenever significant changes occur in commercial music, but to see and understand each of the processes, review their background and take what we consider valuable can be worth the effort. On the other hand, to carry out a complete analysis of the prejudices that the society in general of each epoch has had with these sound and musical revolutions would be beneficial for another research in the areas of modern sociology and musicology.

Acknowledgments

In the writing of this article, the testimonies of some the people who participated in the development of the first era of the record industry boom in Medellin were taken. They are the following: Mario Rincón: recording engineer and master disk cutter; Hernán Darío Usquiano: promoter and artistic director; Mariano Sepúlveda: composer and guitar player; Rafael Mejía: AR manager; Luis Carlos Montoya: arranger and violin player; Julio Ernesto Estrada: recording engineer, performer and composer; Tony Peñarredonda: manager of Discos Fuentes and recording engineer; Dario Valenzuela: recording engineer and master disk cutter; Álvaro Rojas: saxophone player; Humberto Moreno: artistic director and AR manager; Fernando López: artistic promoter and manager of Codiscos; Humberto Chaparro: recording engineer and master disk cutter; Juan Escobar: recording engineer; Juancho Vargas: piano player, composer and artistic director. Finally, I must thank Juan José Arango Escobar for his help in the style correction and the English translation of this paper.
Notes

[1] The expression “irreverent young people” refers to the phenomenon coming from Anglo-American cultures through rock and roll and twist, which entered Colombia in the late 1950s and early 1960s, and which was enforced by the social upheavals of May 68 in France.

[2] The word “costeño” is used to call people and everything related to the Colombian Caribbean coast; although Colombia has two oceans, only those who live or are born on the Atlantic coast, in the north of the country, receive this name.

[3] Although Discos Fuentes was established in Cartagena in the 1930s, it moved to Medellín in the 1950s. As for Sonolux, it moved to Bogotá after the economic opening of the 90's and was absorbed by the Ardila Lülle Holding. The Zeida label would become Codiscos, acronym of Compañía Colombiana de Discos. Another important record label of the time, but that closed businesses, was Discos Victoria. 

[4] At that time it was customary to name musical groups in English, something that was fashionable in most Latin American countries.

[5] Although the Royal Academy of the Spanish Language has adopted the linguistic term from its phonetics in Castilian as “reguetón”, in the media and in the popular collective the spelling used is reggaeton.

[6] Of all the traditional tropical dance genres, vallenato continued to grow between the 80's and 90's until our days, where it competes in popularity with urban genres in Colombia.

[7] The date of each artist was determined by the first published record work.

[8] The term MC (Master of Ceremonies) refers to the person whose function was the rhythmic recitation of improvised or prepared texts accompanied by the beats or musical tracks made by the DJ’s.

[9] The most outstanding voices of tropical dance music from the 60’s and 70’s generally came from the Colombian Caribbean coast; however, within the paisa sound that characterized the music of this time there were voices from other regions, among them the one of the singer Gustavo "El loco" Quintero (1939-2016), who sang in Los Teen Agers (1956), Los Hispanos (1964) and Los Graduados (1967).

[10] These traditional rhythms of the paseo, the paseaito and the cumbión, although are derivations of the cumbia from the morphological, organological and even ethnographic point of view, the tempo between them differ from the original cumbia: the paseo and the paseaito, around of 100 BPM; the cumbión, around 125 BPM; and the cumbia, around 90 BPM.

[11] "Payola" is the term used in Colombia to refer to the phenomenon by means of which artists, record companies or managers pay a fee, tax or bribe to the musical programmers of the radio stations; in this way they could ensure a minimum of broadcasting daily of some of their songs and position it on the radio.

Bibliography


Bennett, S. (2012). Endless analogue: Situating vintage technologies in the contemporary recording & production workplace. *Journal on the Art of Record Production*, 07. Available at:


Galucci, María J. (2008). Análisis de la imagen de la mujer en el discurso del reggaetón. *Opción*, 24(55), pp. 84-100. ISSN 1012-1587


Mark Durham: Multi-Channel Sound Design: Instruments for 360-Degree Composition

Abstract

The continuing development and industry uptake of multi-channel audio is creating new potential for sound designers. This paper presents research that provides a new approach to designing sound for spatial audio applications, by investigating the potential of combining sound creation and spatialisation through performance. The research uses a practice-based approach, involving the design, development and testing of a software-based instrument that combines gestural control, multi-voice sound generation and an Ambisonic spatialisation system. The focus of the research is to prototype an instrument that is easy to learn and intuitive to use.

Introduction

Sound Design is now a complex term to define succinctly. Its origin stems from the post-production audio sector, with the term initially used as a credit for Walter Murch on Apocalypse Now (Coppola: 1979). From this beginning, the use of the term has expanded and changed, and is now used both in its original context and by musicians in a newer one, to refer to the process of creating sounds through a design process. Often, this involves using techniques that employ synthesis, recording, effects processing or different processes in combination.

Post-production approaches to sound spatialisation have changed within recent years due to the development of new forms of content delivery. Ambisonics is currently the industry standard format for interactive applications that form part of virtual or augmented reality work. Dolby Atmos [1] and Auro 3D [2] are systems for both cinema and home, and these both allow for sound positioning within three-dimensional space. Outside of music and sound for audio-visual media, there is also a growing interest in music composed specifically for high speaker count spatial audio systems such as 4D Sound System [3], Envelop [4] and Dolby Atmos for nightclubs.

The technique of producing sound assets in spatial audio formats, such as Ambisonics, is gaining in popularity amongst sound designers, especially
those working in interactive media. This trend has however, not extended beyond this, to the techniques and tools used as a means to generate new sounds or modify existing ones, one notable exception being Sound Particles [11]. The majority of current approaches involve designing sound assets in mono or stereo, then spatialising these within a larger three-dimensional mix.

Due to the uptake of spatial sound formats and the growing interest of musicians in multi-channel sound reproduction, there is now potential for new instruments to be developed specifically for the creation of multi-channel sound. This research looks to question current practices and workflow in the aforementioned areas by asking if new tools can be used to facilitate both the design and mix of sound assets that are inherently multi-channel, existing in the same spatial form - from design through to mix.

These developments in delivery formats have the potential to further bring together music and sound disciplines, providing opportunities for the development of new working practices and collaborative approaches. Alongside this, comes the potential for the creation of sound design tools that are more flexible and intuitive, with greater accessibility (i.e., they can be utilised by sound designers without programming experience). Some areas of industry standardisation further this, for example in delivery formats or loudspeaker types and arrangement, enabling material to more accurately cross from one system to another. Transferring material accurately between formats is currently possible, for example using the Harpex plugin [5]. Some formats previously focused on music reproduction, such as Ambisonics, can be converted to cinema formats such as 5.1 or Dolby Atmos 7.1 beds. More uniformity now also exists in playback systems, with loudspeakers capable of producing near full-range audio in surround positions.

This research looks to investigate this crossover area, where the lines between sound design and music meet, and spatialisation begins. The core aim of this research is to produce a prototype instrument that is suitable for designing and spatialising multi-channel sound assets in an intuitive way for both expert and non-expert users.

**Objectives**

The research objectives are to:

- Develop a sound generation system that can function in the context of a multi-channel instrument, and that is flexible enough to generate a range of sounds.

- Integrate approaches to DMI design into the development towards a user-oriented design, aiming to reach the creative and technical requirements of sound designers.
Explore options for gestural control of the instrument, before integrating a control system that leverages the affordances of the multi-speaker environment.

Link the sound generator and controller with a mapping system that is both flexible and encourages easy experimentation.

Develop the instrument through a series of iterations, documenting the user experience and potential use cases.

**Background and Related Work**

There is a large body of cross-disciplinary research covering approaches to multi-speaker sound, ranging from those using Ambisonics (Lossius & Anderson: 2014), (Schacher: 2010), to more industry focused analysis [12], along with comparisons of the various benefits and drawbacks of different systems (Kostadinov et al: 2010), (Satongar et al: 2013), (Pulkki & Hirvonen: 2005). Examples of work that couple spatialisation and sound synthesis are of the most relevance here, these include research focused on ambisonic granular dispersion (Mariette: 2009), (Wilson: 2008), control of synthesis parameters through gestural control (Wanderley & Depalle: 2004), (Schacher: 2007), and live diffusion of sounds using gestural control (Di Donato & Bullock: 2015), (Cannon & Favilla: 2010).

The uniqueness of this research is the combination of synthesis methods, spatialisation and focus on performance during the sound production stage, embracing approaches to software design from the research area Human Computer Interaction (HCI), more specifically the development of Digital Musical Instruments (DMIs). A commonly used definition of a DMI is provided by Wanderley & Depalle (2004), “An instrument that includes a separate gestural interface (or gestural controller unit) from a sound generation unit. Both units are independent and related by mapping strategies.” Figure 1 demonstrates the connections between the various components of a DMI.

Dividing up the components of the DMI, research focused on the control of electronic or digital instruments through gesture includes understanding...
the requirements of controllers (Wanderley: 2001), (Wanderley & Depalle: 2004), (Schacher: 2007), nd analysis of playability and leveraging the potential for musical expression (Poepel: 2005), (Dobrian & Koppleman: 2006). Also within this field a large amount of research has been conducted into the mapping of control data to instrument parameters. This is accepted as being crucial to digital instruments, especially in enabling expressivity (Dobrian & Koppleman: 2006), (Rovan et al: 1997), (Wanderley & Battier: 2000). More complex mapping or recognition of gestures is one area that can potentially enhance this, where control of sound needs to be both “intimate (finely detailed) and complex (diverse, not overly simplistic).” (Dobrian & Koppleman: 2006, p.278). Rovan et al (1997) further defined a system for categorizing mapping into three categories:

- One-to One Mapping: A single control signal is mapped to a single parameter on an instrument
- Divergent Mapping: A single control signal is mapped to multiple instrument parameters
- Convergent Mapping: Multiple control signals are combined to modify a single control parameter

*Trends in Gestural Control of Music* (Wanderley & Battier: 2000) conducted a comprehensive round table discussion titled ‘Electronic Controllers in Music Performance and Composition’, sending questions to several composers and instrument designers. Machover (in Wanderley & Battier: 2000) suggests that “part of the interest in new controllers is to extend the range of what is manipulated, whether in the density of sound textures or the complexity of musical structures.”

Waisvisz (in Wanderley & Battier: 2000, p.422) also describes a feedback loop between performer and instrument, highlighting why a fast response is a key factor in DMI expressivity, the components of which are illustrated in Fig. 3.

Figure 2: Convergent and Divergent mapping strategies (Wanderley & Battier: 2000).
**Methodology**

Max/MSP [17] was chosen as a development environment, chosen for its balance between flexibility and ease of use for rapid prototyping.

**Gestural Control**

The primary requirements of the control system were ease of use and the expressive potential of controlling both sound generation and spatialisation parameters. Initially two controllers were tested within the system: The Leap Motion controller [6] and the MYO arm-band [7]. These were attractive options as they are both capable of producing accurate hand position data in three dimensions, and have a proven background as alternative controllers in a DMI (Di Donato & Bullock: 2015), (Nymoen et al: 2015). Later in the development of the project a third controller, the MacBook Pro trackpad, was introduced to gauge the benefit of a three-dimensional controller that did not require open-air gestural input. The Leap Motion controller is capable of skeletal tracking of both hands in three-dimensional space, alongside inbuilt sensing of various gestures built in to the Leap Motion SDK V2 Skeletal Tracking Beta. This provides a range of data that is potentially usable for controlling a musical instrument. This implementation uses the Leapmotion external developed by Jules Fran-
coise [8] to connect to the Leap and make the data available within Max/MSP for mapping to instrument parameters.

The Myo controller is a wireless armband that is worn by the user. It provides a range of sensor data from a 3D gyroscope, 3D accelerometer and eight electromagnetic (EMG) sensors that measure muscle actuations. The Myo implementation uses the Myo for Max/MSP external [9] for obtaining raw data from the Myo armband.

Figure 5: Myo armband [9].

Figure 6: Granulator user interface.
Sound Generation

In an attempt to meet the requirements for an instrument with the widest timbral range possible, the development focused on a granular synthesis solution. This approach was attractive for several reasons that aligned with the research objectives. Firstly, it is attractive due to relative ease with which timbrally-rich sounds can be synthesized, providing there is suitable source material available. Secondly the process is flexible through its ability to decouple parameters of pitch and playback speed (Roads, 2004). Many of the control parameters required by a granular system are also relatively intuitive and lend themselves well to direct mapping. Finally, the process is relatively computationally inexpensive, enabling multiple oscillators to present a solution towards a spatial sound instrument.

Figure 6 shows the design of the granulator user interface, intended to provide clear visual feedback to the user.

The sound generator was created in Max/MSP specifically for implementation in this instrument. At its core, the device generates up to eight mono streams of grains using a synchronous granular technique as described by Roads (2004).

Global Controls

These controls set the grain generation parameters across all granular oscillators. Position sets the base starting point for each grain before any modulation is applied. The control unit choice is a floating-point percentage value through the sound file loaded into the instrument. Pitch controls the global pitch of all of the grain streams simultaneously; within the Max/MSP patch this control is affecting the speed of the master phasor~ object that drives playback across the entire instrument. The control is set as semitones and cents, providing four octaves of pitch control in both positive and negative directions. Scan Range sets the extent of any position modulation input into the device. Streams sets the density of grains present in a single cycle of each oscillator, from a single pair of phase offset grain generators to up to four overlapping grains per-stream. Towards the bottom of the interface it is possible to set the Grain Envelope and to also define whether the playback is linear or non-linear with the Grain Phase Distortion control, which distorts the phasor~ as it reads over the audio buffer.

Randomisation

The randomisation controls also affect sound generation at a per-grain level, introducing user definable amounts of fluctuation into all grain-streams. Volume introduces a varying level of volume reduction, defined at the start of grain generation. Pitch introduces a controllable range of pitch variation.
up to one octave, either positive or negative. *Position Variation* introduces a varying level of fluctuation in the grain start position after the current global value. In terms of implementation, each value is derived from scaling a white noise source. This is applied at the level of individual grains, to ensure fully random values across the instrument.

At the right-hand side of the interface there are controls for setting the pitch of each grain-stream. *Pitch (Rate)* adjusts the pitch by increasing or decreasing the speed of the phasor–ramp, whilst *Pitch (Length)* adjusts the oscillator pitch by varying the size of the buffer area being sampled by each grain.

**Mapping**

The mapping approach was intended to be flexible, with a focus on usability. Sonami (in Wanderley and Battier: 2000) suggests a flexible mapping system is a vital part of any DMI, encouraging experimentation and faster development between instrument and performer. With this in mind a modulation matrix was implemented to route controller data to parameters of the instrument. Alongside basic mapping, values can be scaled, offset and inverted to provide more user control. Visual feedback of mappings and their current value is provided in the instrument user interface shown in Figure 7.

![Figure 7: Modulation matrix user interface.](image)

Potentially useful parameters from the controllers were made available to the instrument mapping system throughout the prototyping phase of the project.

**Spatialisation**

The spatialisation approach uses third order Ambisonics, chosen primarily as it is capable of accurate spatial positioning, but also because of the expandability and scalability of the system to a range of other formats (Lossius & Anderson: 2014).

The Max/MSP implementation uses the ICST Ambisonics library (Schacher: 2010). Third order was selected as the most appropriate scale, following guidelines advice in the ICST package that there should be as many speakers as components in the B-Format.
Ambisonic Panning

The initial implementation focused first on creating an Ambisonic panning system. To critically judge the effectiveness of the system a 14-speaker Ambisonic array comprising of an upper quad, lower quad and ear height hexagon of speakers was used. A basic mono source was manually panned around the space with sufficient spatial accuracy.

An early research objective was to implement motion controlled panning within the system using a 1-1 mapping strategy on all X,Y,Z axis. This was completed using the Leap Motion controller, with the controller position in the middle of the interaction area being paralleled by the position of the user in the mixing space (see Figure 8). In effect this allowed the user to intuitively pan towards any point in the room, simply by moving their hand around the Leap.

Spatialised Grains

Connecting the granular synthesizer to the Ambisonic panner allowed for individual mono grain streams to be positioned anywhere within the Ambisonic soundfield. Control over the parameters of individual streams within the synthesiser allows the user to build a soundfield with sonic variation in three dimensions. Useful approaches to this include varying the pitch of each grain stream, through either changing the grain length or rate of grain playback, along with the position of the stream in the soundfield.

Additional visual analysis of the spread and intensity of each stream was evident using the Harpex Ambisonic plugin.

Example output demonstrating analysis [13].
**Gestural Control of Timbre**

To experiment with the control of timbral parameters through motion, each grain stream was panned to fixed positions, evenly placed around the user. This created a very spatially alive sound, but shifted the emphasis away from panning, allowing for an exploration of gestural mappings from the Leap Motion to the synthesiser and effects parameters.

Initially the following mappings were made:

- Palm Position X to Grain Start Position
- Palm Position Y to Grain Start Position Variation Amount
- Palm Position Z to Oscillator Level

The result of this mapping approach creates the following effects:

- Palm position X-axis to Grain start position: Moving the hand from left to right distributes the grain start position along the sample loaded. As the pitch remains constant (through the implementation of a granular oscillator), this effectively selects an area of the sample to granulate.
- Palm position Y-axis to Grain start position variation amount: Moving the hand vertically adjusts the level of randomisation added to the grain start position. The perceived effect of this action adds fluctuation to the grain stream, effectively increasing the variation between grains – depending on the sound source used.
- Palm position Z-axis to Oscillator level: Moving the hand along the depth axis increases the volume of all oscillators linearly.

A second variant applied a more complex set of mappings between gestural controller and synthesis engine, with the aim of exploring mapping strategies that go beyond a 1-1 approach. An additional 8-channel filterbank and 8-channel convolution reverb were added to the system as effects.

The following mappings were made:

- Palm Position X-axis to Grain Start Position (as in 3.1)
- Palm Position Y-axis to Grain Start Position Variation Amount (as in 3.2)
- Palm Position Z-axis to Volume and Reverb Mix
- Hand Rotation to Filtering

Additional effects included:

- Palm position Z-axis to Volume and Reverb Mix: This combination replicates a common technique used in audio production to move sounds backwards in the sound stage. The process increases the wet mix of reverb, whilst also reducing the overall volume.
- Wrist rotation to Filtering: This implemented a combined filter that sweeps from 20hz – 20khz as lowpass through the first half of the
range, then 20hz – 20khz highpass for the second half of the range with adjustable resonance. The effect of the combined control is to tilt the equalisation from a stronger bass response to treble response.

- Grab strength to Volume: Using gesture recognition within the Leap API, this parameter reduces volume when the grabstrength value increases. The effect allows the user to effectively make a fist with their hand to lower the volume to zero.

Example output of this stage of development [14]

**Gestural Control of Timbre and Panning**

Many conventional instruments divide physical input between different limbs of the body, eg pitch and rhythm with the bass guitar, or position and mix source with turntables and mixer. Hunt and Kirk (2000) describe this as the user “injecting energy” into the system (that is the instrument). The example of the violin is shown in Figure 10. This concept formed the basis for the next step in implementation, with three-dimensional panning mapped to one hand, and timbral controls mapped to the other.

The resulting implementation takes XYZ position data from the five fingertip positions on the user’s left hand, then maps these to the XYZ position data inputs to the Ambimonitor object. Figure 11 shows the approach taken within Max/MSP, here thumb position data (as an X,Y,Z list) is split, scaled and mapped to the inputs of Ambimonitor. Figure 12 further demonstrates the result of the approach: with an illustration of two separate gestures through a photograph of actual hand position, the interpreted hand position by the Leapmotion object, and finally the resulting panning position in Ambimonitor.

---

**Fig 10: “Human energy input and control”**

As the Leap Motion is capable of sensing two hands, a logical progression for including both timbral and position control would be to split the two duties between hands, mapping data from a single Leap to each target area. This option was avoided to preserve the intuitive 1-1 mapping approach around the Leap Motion controller (as described in 5.1: Ambisonic Panning). As an alternative, the next arrangement uses the left hand for panning position and the right hand for timbral control using the Myo controller. Orientation from the Myo for Max object was converted into Euler angles, and gyroscope data was summed to create a parameter that measured acceleration in any direction.

In terms of mapping, the approach builds on the previous implementations, with data from the Myo X-axis position mapped to the grain start position in the sample. This approach provided a direct connection between horizontal arm position and the area of the sample being granulated, effectively allowing the user to ‘scan’ over the sample with left to right arm movements. To provide further control over the sound, an additional destination control was added to control the pitch of the grain streams independently of the grain position. This parameter was then mapped to the Y-axis output from the Myo armband, effectively allowing the user to raise and lower their arm to set the pitch of the instrument. Data from rotation along the Z-axis was mapped to the reverb mix, effectively allowing the user to move from fully dry to fully wet reverb mix by rotating their hand clockwise.

Example output using Leap Motion with Myo [15].
As an alternative to using two ‘open-air’ controllers such as the Leap Motion and Myo combination, a third alternative controller, a MacBook Pro trackpad, was added to the system to allow for additional evaluation. The implementation used the Fingerpinger Max/MSP external [10], to capture data for use within Max/MSP. The parameters used for control are X-axis position, Y-axis position and size of finger (effectively similar to pressure). The X-Axis trackpad position is mapped to grain start position, effectively allowing the user to ‘scan’ over the sound using positional movements on the trackpad. Y-axis trackpad position is mapped to global pitch, effectively allowing vertical movements up and down the trackpad to control pitch accordingly. The trackpad is incapable of reading finger pressure, but can read finger size on its surface. As pushing the finger harder into the trackpad increases the size due to compression of the fingertip, this functions in a similar way to a pressure or a Z-axis parameter. In this way finger size was mapped to volume, allowing the user to press on the trackpad and raise the sound level, whilst releasing the finger fades the level down to zero.

Fig 12 (a,b&c): Position mapping comparison - These images demonstrate the mapping between hand and sound position. Photograph (top), 3D rendering (middle) and Ambisonic soundfield position (bottom).

Fig 13: Using Leap in combination with the MacBook trackpad.
Conclusions and further work

The granular sound generator is capable of producing Ambisonic soundfields that are spatially active and constantly fluctuating. This is especially true when adjustments are made to individual oscillator pitch controls and grain start position randomisation. A spatial phenomenon is created by the effect of the individual oscillators running simultaneously and at different pitches; rhythmic cycles across the spatial soundfield vary between noticeably periodic to seemingly random and imperceptible. Source material with percussive attacks provide more clues for the listener to localise sources in this way. At lower grain rates sounds are perceived as coming from their panning location, but at faster rates the sound is perceived as one mass, positioned perceptually by fluctuating interaural time delays (ITDs) and interaural level differences (ILDs) as described by Goldstein (2010).

The combination of gestural controller and flexible mapping system provides a range of control for the user over the sound output. Poepel (2005) suggests musical expression can be coded into performance using “tempo, sound level, timing, intonation, articulation, timbre, vibrato, tone attacks, tone decays and pauses” (Poepel: 2005, p.229). Of these parameters sound level, slow vibrato and tempo (through grain rate) are controllable through the instrument, alongside pitch. Timbral changes can also be programmed into the instrument, due to the way the granular engine handles position within the soundfile. By first designing a performable sound object that moves through the desired timbral range, a sound designer can create a morph that approaches a parametisation of the sonic continuum introduced by Wishart (2002).

The development process of the instrument included informal demonstrations of the system to a set of sound designers, whose feedback was assimilated into the development process. Much of the positive feedback of the system centred around overall ease of use, with all participants able to understand the connection between gesture and sound output after a short introduction to the control system. There was also a consensus that the controls were intuitive and playful, and that the combinations of mappings encouraged rapid development of sounds through use and experimentation.
Areas where usability could be improved centred around response rate and control complexity with some mappings. Improving responsiveness could potentially be achieved through an increase in computer processing power, further optimisations in the Max/MSP patch or by moving to a complete signal driven panning system as an alternative to the ICST Ambisonics implementation used.

Acknowledgements

The Max/MSP objects Leapmotion and Myo were created by Jules Francoise [8, 9]. The Fingerpinger external was created by Michael and Max/MSP Egger [10]. The ICST Ambisonics externals are by Jan C. Schacher (2010).

I would like to thank Mike Exarchos and Savraj Matharu, whose time, support and guidance has been invaluable throughout this project. This work would not have been possible without the generous support of Ravensbourne University London.

Notes

[14] Example output of Leap Motion control. Available at: https://youtu.be/u02h9aNHV30
[16] Example of position and timbre control using Leap Motion and a MacBook Pro trackpad: https://youtu.be/HmLdJ7m-9Yg

References


Abstract
The placement of low-end spectra in popular mixes has ranged from dogmatic centrality to extreme staging, but current developments in production workflows, stylistic evolution and increasingly personalised modes of consumption have resulted in alternative—if less polarised—approaches to low-end production. The authors question how contemporary practitioners ‘sculpt’ bass in their mixes, the strategies that are deployed to secure translation, and the relationship between innovation in the bass spectrum and sensible mix architecture. The aim is to investigate the aesthetic and pragmatic causality behind the phenomenon, but also to examine and theorise on the creative possibilities for future bass staging.

Introduction
The positioning and role of bass elements in a mix has seen significant development in the last sixteen years. Driven by changes in delivery, technological innovation and cultural and musical imperatives, mixers are now faced with an extensive range of options. Where previously bass staging was proscribed by the delivery methods, current trends suggest that there are no rules any more. Limitations imposed by both vinyl playback and loudspeaker response meant that, historically, bass tended to be dealt with quite conservatively in regard to placement. The move to digital playback in the 1980s, followed by the launch in 2001 of iTunes, signalled a shift in listening delivery. Though the Sony Walkman (Du Gay et al: 2013) and other mobile devices had initiated the phenomenon, within a short period the widespread adoption of the iPod heralded a migration from speakers to headphones. This has meant a paradigm shift in the role of the bass. In this paper we will examine the new role of low-end staging, investigate the history of bass placement, create a typology of current approaches, and then illustrate some of
these approaches through two case studies. The aim will be to explore not only how low bass can go, but also how far and how wide.

**Practice**

Initially most early pop and rock and roll recordings were released in mono. Phil Spector famously made his “little symphonies for the kids” (Daley: 2002) using two or three bass players, such as Carol Kaye and Ray Pohlman, who would tune one bass up a fifth and play the same part at the same time. Additionally, the bass would be layered with a baritone sax. This can be heard on many of Spector’s productions such as ‘Da Doo Ron Ron’ (1963) by The Crystals and ‘Why Do Lovers Break Each Other’s Heart’ (1963) by Darlene Love (Daley: 2002). As they were mixed to mono, there was little problem with reproduction.

As stereo vinyl started to become the norm in the early 1960s, producers and engineers began to experiment with bass staging in stereo. ‘Paperback Writer’ (1966) by The Beatles features the bass panned to one side, and the vocals and drums to the other. Pink Floyd employed an alternating left/right bass pulse on ‘Welcome To The Machine’ (1975), while Sly Stone split the output from Larry Graham’s Vox bass into both inputs of the Mu-Tron Bi-Phase on ‘Poet’ (1971) (Rotondi: 2012). Tonto’s Expanding Head Band used a number of stereo bass patches from a Moog Modular, and Oberheim SEM and ARP 2600 synthesisers on their seminal mid-1970s recordings such as ‘Tama’ (1975) (*The home of T.O.N.T.O.*, no date). However, vinyl mastering imposed certain restrictions on bass reproduction. Stereo bass frequencies can cause styli to skip due to large vertical movement and fears of complaints from the public meant most records from the 1960s and the 1970s restrain the bass to the centre (see the ‘Literature’ section below for further information on vinyl implications, including a respective discussion of Zagorki-Thomas’s (2010) ‘The Stadium in your Bedroom’ article). Talking about the track ‘Chameleon’ (1973) in his recent autobiography *Possibilities*, Herbie Hancock provides an excellent account of stereo bass innovation in the face of technical restrictions:

> After we recorded the piece, I started to think that the bass line I’d laid down wasn’t punchy enough, so I decided to rerecord it, programming the Mini-moog so the notes were shorter. The next day I took the tape to [producer] David Rubinson, but when he played that new bass line as an overdub, he forgot to turn off the original track. “Hold on, David,” I told him. “Both tracks are playing.” We didn’t need both those bass lines…but they actually sounded kind of cool together.

> Normally it’s impossible to have stereo bass, because the lower the frequency, the more difficulty the ear has in perceiving direction. That’s why you can put subwoofers in any part of the room: Your ear can’t tell where
low-frequency sounds are coming from. It’s only when pitch gets higher that your ear starts to perceive directionality more clearly. Consequently, nobody ever records stereo bass.

But listening to those two tracks, I got an idea. “Hey, David, since those bass lines weren’t recorded together, they won’t be exactly the same sound,” I said. Couldn’t that function like stereo bass?” (Hancock and Dickey: 2014, p.180).

During the 1970s improvements in home Hi-Fi speakers and amps meant that low frequencies could be heard clearly. Consequently, producers were encouraged to feature the bass more prominently, and bass parts would begin to assume a more melodic and less functional role. Tracks like ‘Good Times’ (1979) by Chic and ‘I Wish’ (1976) by Stevie Wonder feature the bass as more of a lead instrument (albeit with substantial low register and low-end frequency content), but this musical invention could not be matched sonically due to the previously mentioned limitations of the vinyl medium. With the emergence of digital playback, producers began to employ a larger palette of processing to the low end.

Hip Hop producers using track samples often face the problem of adding a more powerful bass element to a historic loop containing a bass part. The resulting muddiness is problematic. Kanye West solves this by applying mid/side processing to the sample, thus creating ultra wide stereo with a significant dip in low end frequencies in the middle of the image. Into this he places low bass, often only occupying the sub-bass spectrum. A good example of this is ‘Sunshine’ (2004) by Mos Def: the careful application of the mid/side processing allows for acceptable mono reproduction, whereas simply reversing the phase of one side would not achieve the same result. One of the most typical contemporary bass placement techniques has indeed grown out of hip-hop. With the multi-layering of beats in the genre, kick drums have become much more active and genres such as Trap and EDM have adopted this methodology. ‘New Rules’ (2017) by Dua Lipa exemplifies this: the trap-style hats and multiple kicks fill the groove space leaving the bass to be a sustaining sub-frequency element. For this stylisation to succeed, it is imperative for the bass to obtain considerable stereo width with some modulation from side to side, in order to retain interest in the low end. However, in this application the bass is back to adopting a functionary role. Listeners feel the bass rather than hear it.

As will be discussed in the “Head-phonography” section below (a term we are coining here to describe the current era, where the predominant mode of music consumption is via headphones), the employment of sub-bass synthesis plugins such as Waves’ MaxxBass and Renaissance Bass can somewhat ameliorate the physical limitations of bass reproduction on small transducers. ‘Fake Friends’ (2017) by the Norwegian singer Sigrid makes good use of this technique. She utilises the same single-note-per-chord tactic as Dua Lipa on the chorus, but with extreme sub-bass synthesis as well as a
wide stereo image. In the verse, the production adopts a different technique, with a shimmersing ping-pong type delay applied to a subsonic throb that barely registers a note. Then when the chorus drops, the power of the low drone kicks the mix up a gear.

‘Electric Pow Wow Drum’ (2012) by A Tribe Called Red goes further by re-triggering a mono synth bass line twice and staging it to extreme panned tracks. The modulation wheel filtering is slightly different on each, creating an almost out-of-phase width. The bass is one of the main features of this instrumental track, so the subtleties of the cross modulation are not hidden by a lead vocal. ‘Dancing (Again!)’ by Eats Everything ft. Tiga vs. Audion combines a number of the above techniques. It has a twin-triggered signature ‘wobble’, panned hard left/right, a stereo drone and a conventional mono bass synth in the middle. This track, a hybrid dubstep/house track shows the sonic complexity of current bass production. There is little or no harmonic progression and the interest comes from a combination of bass production techniques intersected with polyrhythmic drum patterns. Though a long way from ‘Good Times’ and ‘I Wish’ the bass is still a main protagonist that commands attention.

**Literature**

Beyond a practice-based exploration of the position of low-end spectra in popular music mixes, it is also important to consider the theoretical context informing this analysis. A number of scholars have theorised on the placement of musical elements within the space of a popular music mix, and the concept of ‘staging’ has emerged as a useful theoretical notion. In essence, it suggests conceptualising a music mix as a ‘stage’ where the placement—but also the dynamic movement and manipulation—of musical elements (mediation) has thematic and narrative implications (meaning) for both listeners and producers. The concept was first introduced by Moylan (1992) with a focus on the spatial implications of mediation possible within a mix, while Lacasse (2000) explored it further, investigating the effect of textural and dynamic manipulation on the voice in rock production. Zagorski-Thomas has extended the definition to include “functional” (Zagorski-Thomas: 2006) and “media-based” (Zagorski-Thomas: 2009) staging, respectively taking into account “the function to which the recorded output will be put” but also the effect of “technological mediation” (Zagorki-Thomas: 2010, p.252).

Holland (2013) has expanded the concept to include the use of acoustic spaces captured in tracking as a form of staging mediation. Liu-Rosenbaum (2012) has responded to both Zak (2001) and Zagorski-Thomas (2010), tracing musical (and narrative) meaning in recording studio aesthetics and offering “an analytical example of an expanded notion of staging which applies not only to the voice, but also to instruments.” Although Liu-Rosenbaum
(2012) does consider the role and function of the (electric) bass—and the kick drum somewhat—in this context, he admits that its function is supportive rather than ‘protagonistic’; analysing Led Zeppelin’s ‘When The Levee Breaks’ (1971), he explains: “The bass guitar, by virtue of its low register in lock-step with the rest of the section is more felt than heard, but when the levee breaks, it makes its presence known”. He also maps the increase of low register in the mix to an observed tendency for low-end prominence in songs dealing with dystopian themes (Tagg and Collins: 2001); finally, he observes how the unusual for the time (farther kick) drum recording technique has allowed for a ‘bigger’ sound that supports the threatening (river) implications of the ‘pounding’ rhythm section (ibid.).

But beyond these level/amplitude-based observations there has not been a systematic study focusing on bass or low-end staging. Nevertheless, Zagorski-Thomas (2010) does make a number of important observations relating to bass from the perspective of ‘function’ and output ‘media’ in his article ‘The Stadium in your Bedroom’. Key to this investigation is that he notes “a gradual increase of bass on records” in the 1960s and 1970s due to “the thickness of the vinyl and the depth of the cut” used, but also—for the disco era and clubs—due to “the introduction of the 12” single, which allowed deeper grooves with wider spaces between them” (2010, p.253). Furthermore, he informs us that in their study on stereo staging of rock music “Moore and Dockwray (2009) … point out that … the central positioning of the bass that became normative in the 1970s was, in all likelihood, related to the technical needs of stereo groove cutting rather than the positioning of musicians on a stage” (ibid., p.255). He also notes the implications of club acoustics on creative decisions relating to low-end in mixes (function), impacting on both recording techniques and mix processing choices attempting to minimise low-end ambience through close-miking, gating and spatially separating the kick drum (dry) from the snare (wet) (ibid., pp.253-254). Finally, he maps culturally constructed notions of high-fidelity with the production of both mixes and technology that can deliver expanded frequency curves (and ranges); a practice that is pertinent to the current headphone-consumption era, creating both aesthetic expectations and pragmatic challenges for contemporary music producers and mixers:

(E)arly hi-fi enthusiasm … added features to the culturally constructed notion of a high-quality recording that go beyond the technical specifications of dynamic and frequency range to include: exaggerated bass and treble frequencies to enhance the experience of a full range of frequencies by making the extreme ends of the audible spectrum more noticeable (Zagorski-Thomas: 2010, pp.261-262).

To this observation we should add the impact of sound system culture on the elevated focus on—and prominence of—bass spectra in both the aesthetic evolution of particular genres (the remix, EDM) and the hand-in-hand
development of consumption technologies that were capable of reproducing them.

**The Era of Head-phonography?**

As noted above, the launch of iTunes on the 9th of January 2001 signified the first of a number of inflection points that marked a significant shift in listening modes of music consumption (see Figure 1). Although, at the time, iTunes represented a tectonic shift in music consumption modes by signaling the end of physical phonographic products (for many) and the beginning of digital music ownership, seen from the very perspective of 'ownership', perhaps a more significant shift followed; that of a form of services that replaced downloading with streaming. This was exemplified by the launch of YouTube (media streaming) on the 14th of February 2005, and Spotify (music streaming) on the 7th of October 2008, until Apple Music eventually embraced the streaming-and-social-media-and-downloading model on the 30th of June 2015. As Spotify licensing director Guillaume Arth remarked in conversation with the author at the time: “It’s now about access, not ownership” (Arth: 2015). What this meant for audiences, however, was that although headphone listening was not new, ‘access’ to (new) music would from now on be mostly acquired via personalised modes of consumption such as smartphones, and tablets. Zagorski-Thomas (2010, p.255) delineates ideal (live concert) from popular (living room playback of LPs) forms of consumption for rock music in the 1960s and 1970s, highlighting the “idealised or stylised” implication this has had for rock production and mixing (particularly in relation to mimicking large ambiences and staging spatial effects). Equally, an era preoccupied with headphone listening as the predominant form of consumption has its own set of implications for the production of music. If seen in conjunction with the elevation of ‘bass’ to a focal point in the music (and the mix) as a result of both technological affordances and

![Figure 1. A timeline of key inflection points related to contemporary consumption forms of music.](image)
stylistic evolution, the ‘new music’ (requiring a prominent low end for most popular music styles) and the ‘new listening technology’ (headphones) create a dynamic that puts producers, consumers and manufactures in a noteworthy cause-and-effect triangle.

This is exemplified by the attempts of (and collaboration between) producers such as Dr. Dre and Jimmy Iovine, and manufacturers such as Monster Cable (Beats’ partner from 2008-2012), to create headphones that enhance the bass. Effectively, letting you “hear your music the way the artist would play it back—specifically, the way hip-hop and rap artists, like Dre, would want to play it back: with a lot of bass…” (Sharp 2014). There is a lot of speculation about the technology that actually enables this, ranging from questioning the published detail on designer Robert Brunner’s cryptic patents, to equating the ‘Beats Audio profile’ technology to just a “fancy name for (an) equalizer setting” (Kiger 2012). A technical analysis is beyond the scope of this paper, nevertheless the technical challenge of reproducing enhanced (loud) bass on small speakers has remained a focal point for manufacturers for a number of years, and Beats headphones represent only one such highly popular alternative for a generation of headphones that have managed the feat. As early as 1999, in an AES convention paper entitled ‘The Effect of MaxxBass Psychoacoustic Bass Enhancement on Loudspeaker Design’, authors Ben-Tzur and Colloms (1999) discussed the effective application of the MaxxBass signal processing algorithm as a means of extending the perceived range of low frequencies against physical property limitations. Citing Beranek (1993), they explain how “bass notes are supplied physiologically or psychologically because several of their harmonics are present in the signal”, and suggest applications in loudspeaker design exploiting the phenomenon of the “missing fundamental” (Ben-Tzur and Colloms 1999: p.3). What is interesting about technologies such as MaxxBass is that they represent both a technical solution to enhance bass perception in loudspeaker design and a software plug-in that can be used in production.

Of course, different kinds of headphone design—circumaural (around ear), supra-aural (on-ear), earbuds or in-ear—take advantage of a range of further physiological affordances to produce the sensation of low or sub frequencies adequately: from bone conduction for the around/on-ear designs (Buroojy: 2009), to various degrees of coupling directly onto the ear canal for the earbud/in-ear variety (Maya: 2017). This brings us face to face with the production conundrum of when and how (much) to apply such bass-enhancement processing. In a more general sense, mixers are tasked with creating low-end in their outputs that will not only work effectively on a range of systems, but will also translate well on technologies (headphones, speakers) that already utilise bass enhancement technology, physiological resonances or equalisation processing. Although it is still the norm to master differently for vinyl, for most other purposes a producer and mixer will at-
tempt to create a single sonic product (file) that represents the music vision successfully on a range of systems and consumption scenarios: from bass-enhancing headphones connected to a streaming device, to a club with digital-file playback facilities, to a living room speaker-dock connected to a consumer’s smartphone (physically or wirelessly). Below, we will attempt to describe a number of approaches to working with contemporary ‘bass’ with these productions aims in mind, looking at a number of representative examples from recent discography, and subsequently analysing case studies from the authors’ own professional practice.

*Theorising | Typology*

By way of analytically systematising approaches to producing bass for contemporary styles and forms of listening, we suggest that the practice could be initially explored from the perspectives of low-end source *creation* (capturing and synthesising) and then mix/master *processing* (timbral, dynamic, spatial). The range of examples offered in the practice-based background section above, on the one hand, portray a number of approaches related to providing low-end source content: from the capture of electric and acoustic basses (and kick drums) with microphone and DI techniques (e.g. ‘Poet’); to the sampling of previously captured live content (e.g. ‘Sunshine’); to cutting edge synthesis (e.g. ‘Electric Pow Wow Drum’). On the other hand, the (mix/master) processing approaches that can be deducted from aural analysis of these works suggest harmonic, dynamic and spatial enhancement techniques, which require further triangulation from practitioner testimonials (interviews/ethnography) and creative practice research (auto-ethnography).

For this initial research into the topic, we deploy the latter method so as to objectify the aural findings, but we recognise the need for further ethnographic work to extend the investigation into contemporary low-end production and enrich the proposed typology. Figure 2 below provides a schematic representation of contemporary bass creation and processing approaches, delineating further categories related to these respectively, such as: microphone, recording, tracking, arrangement, synthesis and sampling techniques (in terms of *sources*); and equalisation, filtering, phase, panoramic, stereo enhancement and spatial effects techniques (in terms of *processing*). The following practical case studies will map findings to these classifications, and it is important to note that—in practice—a number of these techniques can be deployed in parallel to achieve the desired results. As such, we have opted for examples that demonstrate a range of approaches simultaneously.
Case Study 1—Glenn’s Pop: ‘Freak For You’

‘Freak For You’ is a song recorded and mixed in November 2017 for a University of Westminster student band called “The Bench”. They are a five-piece live outfit featuring drums, bass, guitar, keys and vocal. As an unsigned act, it is envisioned that nearly all of their exposure will come through streaming and listening on headphones, small Bluetooth speakers or via laptop speakers. The Bench have tight, well-arranged songs, but are relatively new to recording as a band. In order to keep the live feel and groove that exemplifies them, it was decided, in consultation with the band, to record the track in a traditional way. As a result, the tempo, keys, arrangement and parts were refined in rehearsals, with the aim of developing the live-performance arrangement to one more suitable for a recording. The parts were then cut live using the university’s Solid State Logic Duality mixer via
Orpheus Prism converters. Using Pro Tools HD, the Fender Precision electric bass part was recorded using a DI, a Neumann U 67 microphone (approximately six inches away and off-axis from the speaker), and a Sontronics Sigma ribbon microphone (around twelve inches away from the speaker and perpendicular to the U 67); the bass was amplified using an Ampeg BA-115 featuring a single fifteen-inch Celestion driver (see Figure 3 below). The mono ribbon source was recorded onto two tracks, which were panned hard left and right with the left channel phase-reversed.


Figure 3. Neumann U 67 and Sontronics Sigma ribbon microphones positioned in front of an Ampeg BA-115 bass amplifier with a single Celestion speaker.

Stylistically the song is reminiscent of 1970s disco—in the manner of Chic or Kool & The Gang—and, as such, one of the key components of the track is the bass line, which includes both fingerstyle and thumb techniques. The acquired bass performance gave the recording an appropriate stylistic signature, which could be further enhanced dynamically by using mid/side and other stereo enhancement techniques on the thumb parts, but leaving the finger parts centred.

The bass rig was setup in an isolation booth to avoid drum spill, while an assistant adjusted the distance of the ribbon microphone to obtain the best phase relationship with the U 67 (see Figure 3 above). Both microphones went through Neve 1073 pre-amps, followed by the separate sides of a Tube
Tech LCA 2B compressor (with a medium threshold setting and a ratio of 3:1) to counteract the noticeable dynamic disparity between the fingerstyle and thumb parts. The DI was recorded without processing, through the SSL pre-amps. Once the tracking had been completed, a composite master part was edited from a small number of takes, before the comp was divided and copied to separate DAW tracks for the fingerstyle and thumb sections. For the fingerstyle parts, the ribbon microphone tracks were not deployed, and both the U 67 and DI tracks were set to the centre aiming at optimum balance. Conversely, for the thumb parts, the U 67 was centred and adjusted in level in relation to the ribbon tracks to obtain the required stereo image, while the DI tracks were muted. To further enhance the stereo field, the left side of the ribbon recording was equalised with some low-pass filtering, removing most frequencies above 2KHz. This had the effect of positioning the thumb parts more to the right of the stereo image, which allowed complimentary panning for the busy rhythm guitar slightly to the left. Though stylistically incompatible with this track, a similar approach could be taken with an upright bass, splitting the part onto different tracks, and placing the high-end elements in the out-of-phase setup whilst maintaining mono integrity for the low parts.

To cater for the envisioned streaming potential of the track, the Waves Rennaisance Bass plugin was used in parallel for both parts as a way of enhancing (synthesising further) the harmonic content of the bass, but as the production of the chorus arrangement became more busy, it became difficult to retain some sense of staging for the verse bass parts. The Waves PS22 Mono to Stereo enhancer—a psychoacoustic spatial enhancement plugin—was thus inserted on the DI recording (opting for an LF spread greater than 2.00 and a frequency of approximately 10 sweeps) to acquire an artificial bass spread, which would, however, fold down to mono effectively without significant artefacts (despite the considerable use of phase and small delays in the plugin). Finally, the Dear Reality dearVR plugin was inserted in the middle-eight section of the song, placing the composite bass part around five (virtual) metres back from the centre sweet spot, and creating simultaneously a sense of both depth and pseudo-panoramic staging. Throughout the bass processing experiments, the aim was to accentuate the bass as a focal element in the track, illustrate some of the typologies identified and highlight their use in strategic staging concepts. To ensure that headphone monitoring, speaker listening and mono playback were compatible, a range of monitoring scenarios were deployed as references, including Beyerdynamic DT 200 headphones, Yamaha NS 10 M speakers with a Mackie bass sub, and a mono Avantone MixCube. The isolated composite bass part and full instrumental mix for the track can be heard using the links below:

- Isolated bass: https://soundcloud.com/user-562957398/freak-for-you-bass/ZED3x
– Instrumental mix: https://soundcloud.com/user-562957398/freak-for-you-instrumental/s-Kt6Q1

Figure 4. software plug-ins utilised on a Spector Euro 4lx doug wimbish electric bass GUITAR Di RECORDING.

**Case Study 2—Stereo Mike’s Hip-hop/Electronica: ‘Psakse’**

As part of a wider research project exploring the interaction of vintage production techniques and contemporary sample-based hip-hop, a number of pieces have been constructed starting from the creation (writing, performing, recording and engineering) of original music referring to older music forms (blues, soul, funk, rock), which are then sampled and manipulated further as part of a hip-hop compositional/production process; this often entails the addition of further sample-based or synthesised sources. The work functions as both work-in-progress content for the author’s independent current album-offering—following solo releases with EMI Music Greece under his Stereo Mike alias—and as practice-led output for ongoing doctoral research (a video blog of the developmental process entitled ‘Hip Hop Time Machine’ can be found at https://goo.gl/N2CblN). For this example, the live performance phase has involved the recording of a 30-minute semi-structured improvisation of acoustic drums, electric bass, electric guitar, keyboards and vocals in a loosely industrial rock aesthetic, which was then sampled in smaller sections and brought into an Akai MPC (X) for further ‘chopping’ and manipulation.

The Spector Euro 4LX Doug Wimbish electric bass was DI’d through a Universal Audio Apollo Twin interface in mono and run through distortion
Michail Exarchos (aka Stereo Mike) & Glenn Skinner: Bass | The Wider Frontier: Low-end Stereo Placement for Headphone Listening

(Bermuda), amplifier (Ampeg B15N) and compressor (dbx 160) UAD software emulations (see Figure 4 above). The aim for the bass was to mimic an analogue tracking aesthetic, infusing it with a sonic footprint that will be important to consider in terms of its interaction with further low-end elements in the subsequent hip-hop production phase. The same applies to the recording of the acoustic kick drum, which was captured on a Prism Orpheus interface with an AKG D 112 microphone through a Focusrite Platinum VoiceMaster pre-amp with some compression applied during tracking. One of the keyboard parts featured made use of a TB-303 arpeggiator emulation on a Korg RK-100S keytar run through an Electro-Harmonix Bass Big Muff distortion pedal, also DI’d through the Apollo interface. All recordings were made using the Ableton Live 9 Suite DAW. The hip-hop production elements relating to low-end spectra involved Roland TR-808 kick drum samples triggered from the Akai MPC and analogue synthesiser bass (courtesy of a Make Noise 0-Coast semi-modular synthesiser). The synth bass was sequenced on the MPC but recorded as audio using its ‘looper’ function, to allow for both control-voltage automation and live manipulation of its modifier and control-signal source parameters (see Figure 5 below).

Throughout the process, the bass layering of and interaction between 808 kick samples, acoustic kick drum, and electric and synth basses, were monitored on studio monitors (active Blue Sky ‘Sky System One 2.1’, Focal CMS 40 speakers and Avantone Mixcube mono speaker) but also on a range of headphones (Audio-Technica ATH M50 and Beats by Dre urBeats3). Parallel to the stylistically-driven juxtaposition of rock/electric versus trap/synthetic bass lines (and Roland TR-808 versus sampled/acoustic kick drums), the aim had been to make the numerous low-end elements sonically co-exist in the mix, while providing an enhanced headphone experience that can also work on speakers. Particular attention was paid to how stereo imaging translated from any layering achieved on headphones (and dampened control rooms), to untreated or common environments. Referencing focused on previously unidentified bass resonances, blurring of definition in the bass/kick layers, and merging of panned bass information farther from the speakers (the track had been performed in Stockholm’s Royal College of Music ‘Lilla Salen’, but also played as a work-in-progress mix out of a variety of consumer systems in multiple home environments). To achieve this, a number of tasks where undertaken. Firstly the acoustic drums were chopped as ‘individual hit’ samples (kick drum, snare drum and hi-hat) on the MPC and re-programmed mimicking the performed patterns. This was done in order to retain some of the live feel but avoid ‘flamming’ with the programmed/808 beats, also allowing for further micro-temporal phase alignment between—most crucially—acoustic and synthetic kick drums. As such, the acoustic kick drum would function as a sampled layer for the more predominant 808 kicks.
Furthermore, the distorted electric bass was spread in stereo using an onboard flanger insert effect on the MPC, before being sent—rather heavy-handedly—to a short stereo auxiliary reverb. Excessive low/sub frequencies were avoided by virtue of the distortion and flanger effects, allowing sufficient space in the frequency spectrum for the electronic elements to come (this could be secured further with high-pass filtering, which in this case was not deemed necessary). The aim here was to both widen the electric bass but also stage it further ‘away’ on the depth axis as a sampled element that would provide ‘space’ (both panoramic and ambient) for the added, synthetic basses.

Finally, arguably the most striking bass feature for the track, came from the synthesis and then manipulation of multiple takes of the pattern triggered from the 0-Coast bass synthesiser; on the one hand, aggressively filtering the overtone content of the designed patch and, on the other, triggering rhythmic LFO-style modulations by altering—in real-time—the shape of a cycled envelope. The live manipulation aimed at achieving varying degrees of rhythmic modulation over variable base sonic spectra to provide numerous ‘takes’, some of which would function as extreme versions of the pattern placed hard left and right in the panorama. The synthesis of the base spectra...
was aimed at an aggressive type of lead bass, without much sub-sonic information, which sits well with the prominent and sub-heavy 808 kick drums (this can be heard in many current ‘bass’ music styles such as Trap, Dubstep and various hybrids thereof)—but also as a complimentary layer to the electric bass. The example illustrates a number of techniques identified in the contemporary examples above and highlighted in the typology (double/multiple-tracking, two/more different parts/instruments, room treatment, hard-panning, time-based effects such as phasing or flanging, and spatial effects such as reverb and delays), providing insight into the inner workings of the practice. It also highlights a mixing-whilst-producing/composing workflow, which is congruent to the type of low-end sounds that are created for contemporary styles, and headphone consumption. The complete track as performed live at ARP 2017 (Bourbon and Exarchos: 2017) can be heard using the link below:


Conclusion

The paper has provided a brief historical synopsis of stylistic and technological developments affecting the evolution of bass creation and production in popular music mixes, before contextualising theoretical problematics about bass placement against literature dealing with the notion of staging in the musicology of record production. Our proposed typology for the analytical study of bass production for contemporary music styles and forms of listening suggests an initial delineation between low-end source creation, and timbral, dynamic and spatial mix processing. These have been further categorised—respectively—into bass recording, sampling and synthesis techniques on the one hand, and harmonic, dynamic and spatial mix enhancement techniques on the other. We recognise that our mixed-methods (predominantly auto-ethnographic) approach to this investigation, consisting primarily of creative practice and intertextual and phonographic analysis, requires further triangulation from other perspectives. The proposed typology and theoretical findings can thus be enriched by further practitioner testimonials, interviews and ethnographic work, as well as other contemporary practitioners’ own reflexive research. As such, much needed practice-based perspectives can be brought to the musicological study of contemporary bass creation, which seem especially necessary at a time when the low-end has become an inter-stylistic focal point across all popular music production. Nevertheless, our research demonstrates that contemporary forms of music consumption and an interrelated shift in stylistic foci have opened up ample creative opportunity in bass staging techniques. These strategies necessitate a heightened awareness of—and balanced approach to—the relationship

101
between technical considerations and expressive sculpting of low-end spectra for music producers. The evolving delivery landscape means continuing reappraisal of low-end reproduction is likely. Far from either a historically conservative or polarised route to bass production in the stereo panorama, low-end sources can now be enhanced or conceived of in dynamic, subtle or uniquely expressive fashion; a notion that we only begin to explore theoretically given the new opportunities in immersive sound and the increasing stylistic domination of Bass Music worldwide.

Bibliography

Discography

Photo B. The “Dome” in Lilla salen, KMH. (Photo credit: Eleonor Gislason-Ferrari)

Photo C. Celebrity interview. Benny Andersson interviewed by Göran Folkestad. (Photo credit: Eleonor Gislason-Ferrari)
Steve Gamble: Listening to virtual space in recorded popular music

Abstract

Research on virtual recorded space manifests a division between production- and reception-based approaches. I address a number of issues which complicate discourse across perspectives and outline why a convergence may be beneficial to research in various disciplines. I consolidate previous models of listening, including Moore (2012) and Zagorski-Thomas (2014), and argue that the ecological approach to perception and research on embodied cognition may provide a useful theoretical framework for bridging this divide. This is exemplified by music analysis and interpretation of Karnivool’s (2009) ‘Goliath’. I discuss the virtual recorded space that the track affords me and consider how listeners may narrativise the track’s personic environment according to ecological/embodied principles.

Introduction

In recent years, popular music scholars have increasingly turned their attention towards the virtual spaces of recorded music. The emergence of ASARP as an institution, along with its conference and journal, particularly stimulates further research on this topic. Researchers of varying disciplinary perspectives have diverse aims for such work, from providing guides for mix spatialisation to developing theoretical models of listener experience. Kraugerud (2017), in a recent review of approaches to virtual space, notes that a close relationship between space and meaning is often theorised in academic work. Indeed, it has been widely observed that the imagined spatial layout of any given track appears to shape that track’s meanings for its listeners (Moore: 2005, 2012; Clarke: 2013; Zagorski-Thomas: 2014). Investigations into virtual musical space may benefit from addressing how listeners experience environments more generally (Windsor and de Bézenac: 2012; Brøvig-Hanssen and Danielsen: 2013). Furthermore, the underlying question of this research combines issues of creativity, technology, perception, and cognition: how does virtual space mean?
There are two dominant perspectives for investigating space and meaning in recorded music. It is often tempting to portray them as binary oppositions: the 2017 ASARP conference, where I presented a formative version of this article, titled the theme ‘aesthetics versus technology’. Related characterisations of this dualism include reception versus production (Zagorski-Thomas: 2014, pp.2–4), and interpretation versus expression (Brackett: 2000, p.15). This dichotomous structure seems to emerge principally from disagreement between two overarching disciplines. In day-to-day practice, the seasoned recordist depends upon their understanding of spatial techniques in order to produce tracks which appease artists and labels. They may claim, therefore, that specific production or mix decisions affect a given musical meaning. By contrast, the music studies academic may stress the potential for any particular track to mean differently for different listeners. This responds to a concern for subjectivity, cultural context, and the body, following poststructuralist thought in the new musicological turn. The gulf between the two perspectives is crossed in ASARP meetings to productive ends, but remains an unavoidable consequence of addressing research foci which span so many disciplines. Each provides valuable contributions to the analysis of virtual recorded space, but uniting varied approaches under a single theoretical framework has proven a challenge (Vad: 2017). Conceiving of these positions ‘versus’ one another particularly illuminates the difficulty of combining these historically distinct perspectives, practices, and ways of musicking.

As it was developed over the last two decades, study of the art of record production deemphasised some dominant musicological concerns in order to prioritise research on creative recording practices in the studio (Cottrell: 2010). Although some work grouped under the ARP label contemplates record listening, the clue is, of course, in the name: record production. Zagorski-Thomas (2014, pp.1–46) examines the disconnect between established research traditions and the new ground that ARP attempts to cover, particularly the creation of recorded music. Many publications which address recorded space principally inform practice, varying from technical manuals to more musicologically situated investigations. To give a representative range of examples: Hodgson (2010), Rumsey and McCormick (2014), Huber and Runstein (2018) are technical guides, closely associated with audio engineering and the Audio Engineering Society; Lacasse (2000), Young (2015), and Mynett (2017) are somewhat more theoretically grounded; Chanan (1995), Théberge (1997), Zak (2001), Doyle (2005), Greene and Porcello (2005), Moorefield (2005), Katz (2010), Schmidt Horning (2013), and Papenburg and Schulze (2015) are historically, culturally, or ethnographically oriented works. Most importantly, ARP called for the recognition of record production as a form of artistic creation, further challenging established notions of high and low art in music studies. Research consequently focuses upon the conditions, technology, and interpersonal negotiations of studio creativity.
Meaning in virtual recorded space

In this record production work, recordists relate to the meanings of recorded virtual space in a very different manner to the privileged position that some musicologists (and general listening publics) provide them. The former relationship is exemplified clearly by Moylan (2015, p.60), who claims that “the listener’s ability to correctly interpret the sounds of the musical message impacts their understanding of the intended artistic meaning of the music”, his work also addressing “the factors that limit the listener’s ability to effectively interpret the artistic elements and understand the intended musical message (or meaning of the music)”. This intentionalist or authorial view of musical meaning presupposes a very specific ideological configuration: musical creators intentionally place messages into an artwork, where they inhere, and listener interpretation is correct insofar as it focuses upon those meanings. Fales (2005, p.173) would reject such an account, noting that a composer “cannot enter the perceived world of his [sic] listener”. Of course, the author has long been dead in critical approaches to interpretation (Wimsatt and Beardsley: 1954; Barthes: 1977). It is understandable to privilege the authorial position in record production, where recordists’ work is evaluated on the basis of successfully interpreting the desires and intentions of performing artists (at least in traditional recording practices, whereby producers/engineers and artists are separate labour roles). Discussions regarding song meaning may therefore provide guiding principles in arranging, engineering, and mixing processes: the spatial layout of the track may be designed with certain interpretations in mind. The resulting approach to music listening is ultimately a pragmatic consideration rather than a theoretically informed account of how listeners understand recorded tracks. A relative of this perspective can be found in Bennett (2015), where the classification of the Sex Pistols’ Never Mind the Bollocks… as rock, not punk, depends upon the successful realisation of its producers’ intentions. Intent does not necessarily materialise in meaning for the listener, however. It may potentially be factored into a listener’s wider understanding of a work, but this involves making a second-order interpretation (Moore: 2012, pp.163–165) when it may be “more prudent and productive to turn primarily to its final arbiters” (Tagg: 2012, p.198), listeners. The view of meaning suggested by production guidebooks may overlook music psychology, aesthetics, hermeneutics, phenomenology, and cognitive science, each of which offers models for understanding music’s potential to affect us as listeners. It also assumes that each listening experience takes place in the same neutral environment with relatively accurate reproduction technology and full listener attention (as in the studio where the producer typically works on tracks). But the meanings of virtual space are not static or monovalent (Negus 2012). To assume that they are is to deny the imagination of the listener, the very same imagination that the recordist relies upon to make their own subjectively grounded decisions.
The account from record production may therefore seem to oppose any musicology of listening. How might this perspective be reconciled with approaches to recorded music listening? We might also ask what there is to gain from such a confluence. ARP research focuses upon recordist practice in order to legitimise the new field and destabilise historical divisions between composer, performer, and the work (Zak: 2007). Music studies may simultaneously develop theoretically and empirically grounded perspectives on music listening. However, this narrative of separate progress overlooks the potential, mutual benefits of further combining the two. After all, a theory of popular music listening must account for historical changes in recording practice, a whole host of culturally grounded transformations which have been well addressed by ARP research (Cottrell: 2010; Zagorski-Thomas: 2014). There are at least two further contributions from production-oriented studies with particular significance for accounts of listening: prioritising the recording as the primary text for investigation, and tracing the increasing technological potential of composition. In return, ARP work may gain from studies of listening a greater awareness of the diverse contexts for reception, and a reconsideration of wider musical understanding. Foremost of these is the reminder that recordists, too, are listeners. The decision-making of a recordist crucially depends upon their listening and understanding of the track as it materialises. In other words, there is no process of record production independent of listening. The maxim ‘mix/edit with your ears, not your eyes’ (Burgess: 2013, p.155–6) reminds engineers to rely upon their listening skills rather than focus upon the visual feedback of digital audio workstations. It is the purely aural outcome of recordists’ decision-making which matters to listeners, who experience the track with no (or an entirely different) visual source. Attending to the theoretical models of reception-based approaches may therefore help recordists to understand the cyclical, mutually dependent processes of listening and decision-making in their creative practice. For instance, Moore’s (2012) view of musical style usefully accounts for the assumptions that individuals bring to their listening experiences. Expectations regarding spatial placement and source motion in a mix depend deeply upon contextual and personal factors as individuals negotiate style conventions. So too does an awareness of our individual capacities: Zagorski-Thomas (2014, p.28) notes how his trumpet lessons as a child inform his personal, embodied experiences of hearing trumpets, a theory relying upon neurological processes of mirroring. Encouraging recordists to critically examine their own listening and embodied understandings of music-making may inform their work in the studio. Ultimately, they may become more aware of their biases, priorities, and preferences. There is, of course, no objective stance from which to make decisions, no “perspectiveless perspective” (Moore: 2012, p.330). Considering how listeners creatively develop meaning in their experiences of recordings encourages us to reflect on our compositional practices as producers.
An ecological-embodied framework

There have been a number of previous endeavours to bridge this divide. Several models of virtual space and listening draw from an ecological approach to perception and/or research on embodied cognition. The application of these perspectives to recorded music is gaining traction within popular musicology in particular. Such work usefully addresses a range of issues, challenging traditional theories of musical meaning (Clarke: 2005; Cox: 2016), providing a methodology for analysing popular song (Moore: 2012), or studying the work of a single artist in detail (Dibben: 2009; Osborn: 2017). These perspectives may be used to combine the distinct foci on record production and listening, and to stop seeing them as fundamentally opposed in the first place. While a full introduction to ecological perception (as it is often abbreviated, i.e. applying ecological psychology to musical perception) and embodied cognition is not possible here, I consider how each approach provides tools for analysing and interpreting virtual musical space. Combining the two research paradigms may require further justification. I therefore review and consolidate a number of previous models, recognising the potential of a combined framework to approach listening, record production, and wider studies of musical meaning. In this discussion, I draw particular attention to the abilities of these models to overcome subject/object divides. This may avoid upholding dichotomous views towards aesthetics/technology and reception/production. I subsequently demonstrate the use of this framework for music interpretation, employing it as a basis for an analysis of Karnivool’s (2009) ‘Goliath’.

Recently, the term ‘affordance’ has found particular use in addressing listening experience within music studies. Originally conceived in the ecological psychology of Gibson (e.g. 1986), the theory of affordances acknowledges that the potential for action and meaningful experience emerges in the mutualism of an organism and their environment (Windsor and de Bézenac: 2012, p.104). Applying this to music listening, a recorded track is an environmental stimulus, and our individual listener the organism. What the track affords will depend upon the capacities of the listener and their perception of the music in each listening circumstance. Whereas ecological perception may characterise environmental information as “structure that is already out there in the environment” (Clarke: 2005, p.15), my specific approach emphasises the listener’s interpretive capabilities. It is affordances, not stimuli, that are perceived (Michaels and Carello 1981: 42). This view aligns more closely with the experiential perspective of embodied cognition, developed particularly by Lakoff and Johnson (1980; 1999). The view from embodied cognition argues that human reasoning processes are embodied and largely metaphorical. Concepts or image schemata, which are specific precognitive gestalts, are understood to underlie human thought, depending deeply upon the peculiarities of the individual body. Thus the perceiver’s relation to their
ecological environment – and by extension the emergence of meaning – depends not only upon the structure afforded by the environment but the adaptive, embodied capacities of the individual. As recorded music can be an environmental stimulus, these accounts can be fruitfully applied to music listening. Whereas traditional perspectives may view musical meaning as abstract, disembodied, and representational, the embodied and ecological approaches characterise individuals’ understandings of music as conceptual, embodied, and metaphorical.

On this basis, Moore (2012) analyses recorded virtual space using the soundbox, which is an imaginative model of the listener’s spatial experience. Following ecological and embodied principles, we may use it to analyse how tracks afford locating certain sound sources in a virtual environment. It shares a concern for meaningful understandings of recorded music with Lacasse’s (2000) conception of staging. While Lacasse’s (2000) approach suggests the objective arrangement of sound sources, however, it may be argued that any physically verifiable spatial placement is not what generates meaning for listeners. What matters for the perceiver is the virtual space that they understand (even if visualisation software might suggest they are hearing it, in some sense, incorrectly). Recorded audio is objectively verifiable data, but listening experiences are not, and we are in deeply interpretive territory as soon as a listener has conceived of a given source as music. I propose further combining the ecological and embodied approaches into a coherent framework of listening in order to address this activity. Meaning is taken to arise in the interaction between the object and its perceiver, instead of emerging from immanent properties of the environmental object. In other words, the soundbox models the spaces that listeners understand, rather than a spatial environment “existing as physical space” (Brøvig-Hanssen and Danielsen: 2013, p.72). Such understandings can be extremely imaginative and creative, as listeners narrativise the virtual spaces that they hear (Harden: 2016). It is valuable to recognise this creative activity when discussing the meanings of virtual musical space.

Indeed, the interpretive complement to the soundbox that Moore (2005; 2012) provides is the personic environment, which addresses the relationship perceived between the lead voice and everything else heard in a track. The listener may imagine the lead voice to indicate a persona, and the remaining elements of the track form their environment. This approach is broadly compatible with other theories of the persona developed in popular music studies. Auslander (2009) and Frith (1996) provide complementary models of musical personas for studying star personalities and fandom beyond an emphasis upon music listening. These are consistent with embodied and ecological perspectives which address everyday interaction with other individuals. In the analytical example I present, the persona is defined only as the character which the listener may imagine in the listening experience. This focused conception of the personic environment is framed by ecological and embod-
ied principles: as listeners, we may ask where the persona is in relation to their environment, and where the track affords positioning ourselves. When listening to recorded music – just as we do in our real environment – we attempt to understand our surroundings by means of asking what is afforded by the sources around us. Just as we turn our heads to face the speaker when we hear a voice day-to-day, so that the sound reaches each ear at approximately the same time, we may comparably focus upon the voice in a recorded space. The spatial placement of the voice in the centre of the stereo field has become normative, at least in rock (Moore and Dockwray: 2008), and there may be an embodied basis for this convention: as binaural animals, humans prefer balance, and hearing a voice stuck to one side of the stereo field (whether listening on speakers or headphones) may prove disorienting. The recordist may have conscious intent to create disorientation or discomfort using such a spatial layout, but they are unable to prescribe the listener’s experience of the track. Clarke (2013, pp.108–9) elaborates that what any given “sounds afford for any individual listener cannot be definitively stated, since it arises out of complex interdependencies between detailed attributes of the sounds themselves and the particular sensitivities and circumstances of each listener”. Zagorski-Thomas’ (2014, p.244) approach to sonic cartoons comparably attends to interpretations of recorded space, stating conclusively that “record production doesn’t create meaning for listeners; neither does music. Listeners engage in the process of interpretation. The meaning is in the people, not in the sound”. Although this perspective might privilege subjectivity over the conventional Gibsonian approach, the debate regarding whether affordances should be seen as properties or as relations is still unresolved in ecological psychology (Sanders: 1997; Michaels: 2003).

Nonetheless, it is the interaction of people with their environments that both the ecological and embodies approaches take seriously (Lakoff, 1987: pp.215–6). I offer a new characterisation of their relationship to this end, which principally allows for interpretive discussion of recorded virtual space without privileging any single response. We may nonetheless suggest how some readings are more immediately or widely afforded, while some affordances may emerge for only a small number of people. The track (a particular recorded instance of song) is a suitably stable, widely accessible text for analysis. It provides a reliable substitute for the real interpretative focus, which is the listening experience. This experience might be repeatable to some extent, although listening closely to the track at home on headphones will differ dramatically from hardly hearing the track from a shopping centre tannoy speaker (Kennett: 2003; Negus: 2012). Moreover, we may be able to consider an average or ideal experience of a track, although this may change based upon our listening context and understanding of the style, artist, and ourselves. We can really only speak of our own experience, after all, even if we may expect some intersubjective agreement of others. This account does not require adopting an extreme subjectivist theory of meaning, because the
most convincing interpretations can be based upon specific understandings of virtual recorded spaces. The key argument I prioritise is that the individual matters: what is heard, and what the track affords them, might not necessarily correlate with the recordist’s focused hearing in a treated environment. As such, any advice to mix, edit, or spatialise a certain way in order to produce a particular effect may be misleading.

I refer to the theoretical basis for understanding recorded musical experience which I outline here as an ecological-embodied framework. The perspectives subsumed under this label have been variously employed in the literature I have discussed, but this term and the specific characterisation are new. I suggest that it is most beneficial to approach this composite pragmatically: depending upon what we are investigating, we may draw more from either paradigm, but the underlying principle remains the same. At its most fundamental, the ecological-embodied framework posits that all of our experiences and understandings of the world – and of recorded music – depend upon our being the types of bodies we are, as we interact with our environments. Importantly, this model also responds to new findings in psychology and cognitive science, and can be adapted as new paradigms emerge. It is worth acknowledging that the study of the embodied mind is still in its infancy, relatively speaking, but it is beneficial to model musical experience based upon up-to-date empirical evidence (Zagorski-Thomas: 2014, p.30).

Embodied cognition particularly bemoans the erasure of the body in established perspectives, and ecological perception attempts to define the environment ecologically rather than physically, which sidesteps the objective/subjective binary. This combined view adopts an experientialist account of meaning (Lakoff and Johnson 1980), which depends deeply upon embodied and cultural relativity in relation to environmental stimuli. Consequently, the distinction between nature and culture that traditional perspectives uphold is subverted from the outset. Cultural understanding forms part of our experience, based upon the perceptual and cognitive learning that we, as individuals, undertake (Windsor and de Bézenac: 2012, p.105). In other words, culture is part of our cognition.

The ecological-embodied framework may also help us to consider that the opposing accounts with which I introduced this paper – one research strand versus the other – may not ultimately be at odds with one another. They provide different insights from a range of perspectives that may be channelled back through the framework. For instance, the recordist’s belief in their ability to create a meaningful effect may be reinterpreted as a heuristic which their work affords, one which helps them to make satisfactory decisions. Lacasse’s (2000) staging, Moore’s (2012) soundbox, Brøvig-Hanssen and Danielsen’s (2013) music sound stage, Zagorski-Thomas’ (2014) sonic cartoons, Moylan’s (2015) sound stage, and related approaches all provide conceptual and metaphorical tools for understanding virtual space. As Brøvig-Hanssen and Danielsen (2013, p.75) note, “Gibson’s notion that the same
environment might afford different things to different people rings especially true when using the sound stage as a metaphor to describing sound”. Recall that the view from embodied cognition suggests that metaphor is not merely a linguistic, poetic function, but a principle underlying thought processes. We may use the framework in this way to usefully address the ways that a range of listeners understand virtual recorded space.

Following this development of the theoretical framework, I now exemplify its use in music listening and interpretation. This is surely not its only use, as other scholars have pointed out (Zagorski-Thomas: 2014, pp.37–46): many types of musical engagement may be addressed using this framing for experience, including approaches to studio creativity, recordist practice, and audience research. In this paper, I provide the track analysis as a salient application of the framework, which provides a certain degree of stabilisation to the ‘versus’ debates in record production research. This analysis additionally contributes to scholarship on popular music analysis, hermeneutics, and metal music studies. I discuss the potential for meaningful experiences of virtual space to arise in listening to Karnivool’s (2009) ‘Goliath’. This track, taken from the Australian band’s 2009 album Sound Awake, is an appropriate example for this contribution, due to its critically acclaimed production (Barkan: 2010) and its interesting relationship to production norms. In some ways, it abides by modern rock/metal standards (Mynett: 2017), but in others it varies idiosyncratically, affording me a rich narrative interpretation. While I cannot embed the track here, it is widely available on the usual listening platforms. It may help to listen to it now before engaging with my understanding of the track.

**Karnivool – ‘Goliath’**

The spatial environment of the track is introduced piece by piece. First, there is a metrically ambiguous and highly distorted electric guitar monotone, situated just right of centre. Another guitar that I hear towards the left ear plays call-and-response with this, notably unbalanced by anything to the right. We may expect something present on the right to provide balance – complying with mix conventions (Moore and Dockwray: 2008) – depending upon what we are used to hearing. In ‘Goliath’, however, the right side of the stereo image is rather vacant, and the listener might feel slightly off balance if they imagine themselves within this environment: the sound powerfully engages only one side of the body. When the drum kit and bass enter at subsequent repeats of the riff pattern, they are confined rather narrowly to the centre. Although the drums are forceful and the bass highly abrasive, their central stereo location may make the distorted guitar on the left feel even more remote, and the space on the right even emptier. Shortly after the kit and bass enter, the voice appears in the centre. The persona that I imagine is
slightly more distant than the other elements: I feel his declarations hit me from afar, then dissipate upward and outward. His presence is rather unobtrusive, seemingly floating above the menacing high-frequency content of the distorted bass. It is as if he is trying to call over this mass of sound, which works against him, obscures him. After a few vocal lines, the stereo space offers some balance, as a new guitar joins towards the right. This guitar part is not in synchrony as we may expect in rock or metal, however, instead repeating a flurry of higher notes. It may provide an unsettling contrast to the gritty tonic pedal of the left-panned guitar. At 1’02”, the persona’s voice spreads to the mid-left and mid-right rather than occupying the centre, perhaps providing a little stereo stability, or otherwise dividing disconcertingly. The right-side guitar flurry settles into harmonic filler with a cleaner tone. The persona presently returns to the narrow centre, and the right side soon empties altogether, the guitar departing by ringing out a slightly dissonant B over C#. Later, the persona is underlined by a new voice doubling the octave below. Although it feels somewhat closer to me, and perhaps more urgent, it does not seem able to fully reach through the thick guitar walls either.

The environment generally alternates between these spatial layouts up to 1’50”, where the main rhythm guitar is doubled to both ears for the first time. At this point, the unsettled lead guitar semiquavers sit just right of centre, closer than before, and perhaps alarmingly so. The dense, homophonic motion of the doubled, wide-panned guitars and airy crash cymbals in this section more clearly resemble a typical metal mix (Mynett: 2017). The rests between the unison guitars’ polyrhythmic downstrokes cause the stereo space to suddenly open up, which we may understand to enact the persona’s lyrical demand to awaken. The listener might imagine this to be a sudden blinking open of the eyes, a gasp for air between the opaque, oppressive guitar slides. The persona now seems rather trapped – confined to the centre between these thick timbral walls – and unheard, because he keeps repeating himself. More specifically, I conceive of the persona trying to shake himself awake when trapped in a nightmare, as if experiencing sleep paralysis.

The worst of the dream passes for a short while, and a new riff is introduced by the left-side guitar from 3’19”. This virtual environment is again rather imbalanced, save for a subtle palm-muted tonic pedal on the right. Little textural equivalence is provided here, and I am unsettled by the thick timbral wall of the left and the thin strand of the right, especially when the latter rises a perfect fifth. The guitars and kit become more animated throughout this section, still based on one harmonic centre and one side of the soundbox, even into the riff proper, where the hard-left guitar begins leaping frantically. On the far right, the semiquaver flurry returns, now more frenetic and mostly chromatic. This guitar may seem rather separate from the drums and other guitars, spatially and harmonically, barely aligned in time. In my imagination, this virtual environment enacts both a troubled sleeper
and their nightmare: the right-side guitar acts out the torment of being trapped in the dream, and the punchy kit and heavy guitar on the left create the sleeping human violently shaking (perhaps trying to break free) in the night. Such potentially horrifying scenes of nightmare and/or sleep paralysis have been frequently depicted in the visual arts, associated with cultural beliefs regarding the supernatural (Jalal and Hinton: 2013). As the sleeper in ‘Goliath’ tries harder to wake up, the persona calling out again at 4’20”, the right-side guitar seems to answer ‘no’. This guitar synchronises with the sinister riff of the left side guitar, and both drop down a half-step to C as the nightmare prevails. This is a potent vertical shift, as if the entire environment has been suddenly tugged downward, particularly as the track has wallowed in a C# tonic centre for its entire duration up to this point. We may also conceptualise such a change stylistically, associating the track first with a rock style, before the wide-panned guitars finally offer the equivalence and lower pitch typical of metal.

**Conclusion**

This interpretation of ‘Goliath’ is based upon the virtual space that I hear on a typical experience of the track. It may more closely resemble the kind of narratives which fans often report in their listening. My interpretation draws from analysis of the virtual music space, but generally avoids the specific language of music production. After all, many listeners do not relate their experience to the precise placement of sound sources or the use of particular compression styles. The ecological-embodied framework acknowledges that individuals may instead relate to the track’s virtual musical space more metaphorically, as appropriate for their interpretation. They may hear shapes (e.g. the guitar ‘walls’) relationships (the distant persona), and the feelings (unsettled, threatened, horrified) that these are believed to produce. On the experiential view, listeners themselves are the producers of such embodied understandings (Zagorski-Thomas: 2014, p.244), but the potential of music to evoke affective responses remains a prevailing conception of music’s power. The listener responds first and foremost to what they understand a track to mean by listening, whether or not this is supplemented by knowledge of production processes or recordist commentary. In the analysis I provide, the emphasis is placed on my listening and interpretation free of any authorial sentiment. Specificities of the recording remain central to my understanding, however. I did not approach the track simply thinking about nightmares at first encounter, but formed this interpretation by experiencing characteristics of the virtual space: the voice constricted to the centre, seemingly distancing the persona, who panics about waking; the wide, animated guitar walls and an abrasive, threatening bass; an unbalanced stereo spectrum, with empty spaces, then sudden bursts of stereo synchrony. These at-
tributes may be alarming in our everyday environment, and they are when listening to ‘Goliath’, too. This is perhaps the most significant contribution of the ecological-embodied framework: investigating our experiences of the world, of our environments, and of our music, as we understand them.

Bibliography


**Discography**

Lachlan Goold & Phil Graham: The Uncertain Future of the Large-Format Recording Studio

Abstract

This paper investigates new spatial relationships in music production triggered by the proliferation of low-cost digital music production tools and how they have changed factors of time and creativity for the record producer. We address the problem through a focused set of production projects and the peer review of those projects by seven well regarded commercial producers. Put simply, we are asking whether those producers can determine whether the recordings were made in large-format or DIY “home” studio environments in blind listening tests.

Introduction

This paper investigates new spatial relationships in music production triggered by the proliferation of low-cost digital music production tools and how they have changed factors of time and creativity for the record producer (cf. Bennett: 2012, p. 8; Théberge: 2012, pp. 89-90). We address the problem through a focused set of production projects and the peer review of those projects by seven well regarded commercial producers. Put simply, we are asking in what ways do DIY recording spaces change the experience of recording for the producer and whether those producers can determine if the recordings were made in large-format or DIY “home” studio environments. The recordings were created in both DIY recording spaces, large-format recording studios, and a combination of both environments. The producers conducted blind listening tests. The producers took a questionnaire (see appendix) about their attitudes toward DIY recording, the effect on the producer’s creativity in DIY situations, and their opinion on the fidelity of the finished recordings from those environments.

We argue that the DIY paradigm in recording is a domesticated form of a formerly industrialised production process. As a result, the large-format, industrial era studio has become increasingly unviable, with many iconic facilities closing down throughout the world, especially over the last decade. We define the large-format recording studio as consisting of a console of 24
channels or more; the ability to record 24 or more inputs at the same time; an
acoustically treated and isolated control room, and multiple acoustically
isolated recording spaces of varying size. We define DIY (or home) record-
ing as the practice of recording in an acoustically untreated domestic space,
or any space not specifically designed for recording, which uses limited
portable equipment based primarily around a computer and software. Our
definition of a “hybrid” process describes a process in which initial record-
ings were made in a large-format studio with overdubbing done in a DIY
studio.

Since the late 19th century, the quality, fidelity, and sophistication of au-
dio production have progressed along with technology (Leyshon: 2009, p.
1319; Pras, Guastavino, & Lavoie: 2013, p. 612). However, this trajectory
seems to be in question as the quality and use of new technologies, advanced
as they are, are claimed by many to be detrimental, socially, musically, and
concerning audio fidelity (e.g. Burkeman: 2006, p. 1; Williams: 2015, p. 3).
There is little research to either reinforce or refute whether recordings have
changed in respect of any of those factors.

Quality of space and creativity are commonly linked. There is, for exa-
ample, a common-sense assumption that the more comfortable an artist is, the
better they will perform. Howlett (2009) writes of spending a whole day in
the large-format studio during the 1980s to acclimatise the artists to the arti-
ficial environment, reduce alienation, and promote better performances (p.
49). That is a generally unaffordable scenario in the current era of self-
funded and small-label recording budgets. The large-format studio has been
in financial decline since the first Digital Audio Workstation (DAW) began
to compete in terms of quality with the analogue 2” tape machine early in the
twenty-first century. Industry and academics discuss the trend in mostly
woeful terms (Leyshon: 2009, pp. 1326-1327; Théberge: 2012, p. 82; We-
denbaum: 2012). However, the evidence we present here demonstrates that
in a DIY context, the advantages of new digital technologies combined with
an experienced producer can enable at least equivalently creative musical
performances to those in a large-format studio by removing creative barriers
built into the architecture of the large-format recording studio.

Csikszentmihalyi (1996) argues that it is ‘easier to enhance creativity by
changing conditions in the environment than by trying to make people more
creative’ (p. 1). The large-format studio has long been considered a creative
space and regarded as another instrument and compositional tool (Eno: 2013,
p. 129; Kealy: 1990, p. 179). By merging recording with the place of com-
position (for example, the home or a representation thereof), the recording studio
becomes another musical instrument without the performer having to
confront the institutional relations of power built into the large-format studio.
The recording studio is a space in which musicians’ creativity combine
with the recording environment and the producer’s ability to enable inspiring
performances (Gibson: 2005, p. 192). Our focus here is to know how profes-
sional producers understand the large-format recording studio given the current recording downward budget trends, and whether its ability to capture a creative performance is vulnerable now that professional quality DIY home recordings can be made.

There are indications that the laboratory-like recording space of the large-format recording studio can affect a recording negatively, particularly when professional results can be achieved “in the comfort of your own home” or in another DIY scenario more conducive to creativity (Byrne: 2012, p. 174; see Gibson: 2005, p. 196; Hennion: 1989, p. 406; and Leyshon: 2009, p. 1319 on the 'laboratory'). Given the rapid uptake of DIY recording, this topic has significant ramifications across the music and recording industries and brings into question the relative usefulness of large-format studios as creative recording spaces. This paper aims to evaluate producer attitudes toward exactly what, if anything, is lost as the industry moves away from large-format studio recording spaces into DIY recording scenarios.

The changing space of record production

The pre-electric (mechanical or acoustic era) recording studio developed a culture of ‘getting it done’, and records were meant to document performances with artistic considerations being secondary (Horning: 2013, p. 2). By the late twentieth century, the electrical era changed the creative character of the studio radically, but the basic design principles of the studio remained unchanged from the earliest days of the electrical era (Williams: 2007, p. 2). The post-war period saw great changes in recording technologies yet the primary purpose of the studio remained entirely utilitarian: to capture a live performance with the best frequency response and fidelity (Kealy: 1990, p. 174). As fidelity increased, so did the demands on the recording engineer’s skill with the main goal being ‘concert hall realism’ (p. 174).

The development of magnetic tape recording in early 1949 added to the factory-like impetus of the recording studio, adding enhancements to fidelity, factorising asynchronous production processes, and offering economic benefits to recording with tape being much cheaper than direct-to-disc techniques (Kealy: 1990, p. 176; Toynbee: 2000, p. 80). The further development of multi-track recording and overdubbing enabled the studio to become a compositional tool and a place of creativity in which well-funded major label artists could “set up shop” for months and sometimes years at a time (Théberge: 1997, pp. 215-216; Williams: 2006, pp. 209-211). Extended time periods in the studio facilitated composition and experimentation (Byrne: 2012, p. 205; Leyshon: 2009, p. 1316; Théberge: 1997, p. 216). Technological developments from the late 60s through to the 90s, changed recording production cultures with the implementation of MIDI and digital recording technologies. In particular, during the 80s, production styles moved further
from a performance capture paradigm with the extensive use of programming and synthesisers, often with results never intended for live performance (Bennett: 2009, p. 2).

Even over the short time that recording audio has been possible, the dominance of large-format recording studio between the mid-70s and the mid-90s is a brief period of anomaly. The progression of recording techniques now relies less on the sound and architecture of the performance space because recording has well and truly departed from the “performance capture” paradigm toward the digital manipulation of pure signal. The advance of technology has put the large-format recording studio’s economic viability into question and continues to influence changes in the current era (Watson: 2013, p. 330).

**DIY recording culture: domesticating technologies**

The recording studio has had a history that is akin to the industrial zeitgeist at any given time. At some stages, it was like a laboratory. At others, it became like a production line factory (see Hull, et al., 2011, p. 221 on the ‘factory’). At others, it became the creative equivalent of a project office. Through those periods, the recording studio progressed from an uncomfortable space with limited creativity to an instrument integral to the construction of a successful and creative recording. The growth of low-cost music production tools concomitant with the closure of large-format facilities has dramatically changed the current era of recording practice (Bennett: 2012, p. 8).

Homer (2009) argues that DIY music production began to thrive from the early 70s, although it is possible to establish those origins in early examples of experimental overdubbing techniques known as sound-on-sound (pp. 86-87). In 1941, Sidney Bechet released a jazz recording called “one-man band” in which he used sound-on-sound techniques to record all of the instruments, including drums, bass, saxophone, clarinet, and piano (Théberge: 1997, p. 217). More famously, Les Paul innovatively recorded an instrumental using sound-on-sound techniques in his shed, in 1947 (Milner: 2009, p. 126). Despite Paul’s success in using innovative recording techniques, the conservatism of the recording industry restricted the multitrack’s integration into the laboratory model of recording architecture based on an institutional separation between control and performance as identified in the term “control room” (Toynbee: 2000, p. 81). Sporadic experiments in DIY continued after Les Paul’s recordings, outside the majors, sometimes with incredible success. Sam Phillips’ early work at Sun Studios is a well-known example, as are Berry Gordy’s early Motown recordings in Detroit’s Hitsville studios,
and Bill Putnam’s work at United Recording Electronic Industries (UREI) (cf. Bisel: 2015; Milner: 2009, p. 151; Sutheim: 1989, p. 729). Compensating for the shortfalls in their archaic DIY recording environments in the late 60s and 70s, producers Lee ‘Scratch’ Perry and King Tubby creatively used and experimented with reverb and echo to create new sounds (Gibson: 2005, p. 198).

Portable four-track recorders appeared in the early 1970s and gave artists a glimpse of what it may be like to record at their convenience (albeit in a non-professional capacity) while preparing and experimenting with their demo recordings (Milner: 2009, p. 174; Théberge, 1997, p. 221). Those new technologies were not comparable to the fidelity of the large-format studio, but as Théberge suggests, new musical genres emerged around the use of this equipment (Filipetti as cited in Massey: 2000, p. 194; Théberge: 1997, p. 195). A genre known as lo-fi (low fidelity) emerged in the 1980s and gained prominence in the early 90s, established itself using cheap domestic recording equipment to purposefully sound as if it wasn’t professionally recorded (Carew: 2017; Jones: 2014, pp. 42-48). Bennett and Peterson (2004) define lo-fi as ‘a form of DIY recording that was typically done by amateur musicians on cheap, domestic recording appliances, such as cassette or reel-to-reel tape machines, in home settings such as the bedroom or living room’ (p. 218). In an effort to subvert the ‘80s rock polish’ of the commercialised music industry, a new lo-fi genre emerged in which the participants recorded on cassette 4-track recorders and established an underground network of DIY culture (Spencer: 2005, pp. 273-274).

The 1980s is also the time at which professional-level home recording became possible, with the arrival of MIDI, sequencing, and affordable mixers. As Bennett (2009) notes:

The technological acceleration that began in the early 1980s had, by the decade’s end, culminated in a wealth of time-saving, space-saving, not to mention money-saving systems, bringing with them new working practices and redefining the roles of the producer and engineer (p. 2).

By the late 1990s, the perceived quality of the home studio increased with further storage medium improvements of digital recording, and low-cost microphones that performed close to the most expensive high-level microphones (Filipetti as cited in Massey: 2000, p. 194).

Much of the literature on the DIY recording phenomenon concerns solo-performed electronic-based music (cf. Homer: 2009; Knowles & Hewitt:

---

1 There is ongoing conjecture over who made the first sound on sound recording. Most literature has focused on Les Paul, however Bill Putnam claims that he used sound on sound before Paul (as cited in Sutheim, 1989, pp. 728, 729), despite Sidney Bechet’s “one-man band” single. Regardless, sound on sound recording emerged in 1947 in popular music in the work of both Bill Putnam and Les Paul.
In contrast, the focus of this research is popular song recordings; guitar- or keyboard-driven music that is usually recorded in large studios due to the loud volume of acoustic drums and other electrified musical instruments. This research does not address orchestral recording, small string ensembles, Jazz, repertoire, musical theatre cast recordings, field recording, spoken word recording, or EDM/dance music production. Although some of the recordings used in the experiment we present here were created in a DIY manner, our intentions were to not sound lo-fi; rather, the music consisted of new and original works by independent artists, written by the artists, where innovation and novelty are desirable approaches in the recording process.

Watson, et al. note in 2009 that ‘professional quality recordings can be produced by individual musicians and producers in modest recording facilities and home studios, enabling artists to control more aspects of the production process’ (p. 867). Which is also to say that before then it was not possible to achieve a production comparable to large-format studios using home recording technologies (Homer: 2009, p. 92). Homer (2009) bemoans the development: ‘While there seems to be significant economic advantages for artists to record at home, it is important to draw attention to what exactly is lost when the musician dispenses with the professional studio’ (p. 90). In DIY spaces, not only are the large acoustically treated rooms missing, so too is the state of the art equipment, and the experience and knowledge of highly trained engineers and producers (Jenkins as cited in, Homer: 2009, p. 90; Pras, et al., 2013, p. 623). Without the need and expense for an engineer or producer Homer (2009) claims that:

> With home recording technologies potentially providing a democratisation of the music making process through their affordability, it is important to explore what this greater access means for the creation of music. Through its ease of use, the proliferation of digital technologies has arguably enabled greater experimentation in the music making process. (pp. 90-91).

As major record businesses faced severe downturns from 2004 onwards, it has become commonplace to record part of an album in a large-format recording studio and finish the overdubs and editing in a DIY context, thus saving money and leaving more time for creativity (Slater & Martin: 2012, p. 69).

Recording at home requires reorganisation to accommodate the studio. Théberge (1997) argues that substantial inconveniences created by bringing the recording studio into private spaces and its effect on the household (p. 215). Home studios are usually hidden away within the household so as to not disrupt the normal flow of the house with a tangle of wires and equipment, making home recording a private act (p. 234). Auvinen (2016) argues that the domestic space is not ideal for production even if the studio is tech-
nologically current and the space comfortable because the home is not always conducive to work and it presents challenges acoustically in both recording and monitoring playback (p. 6).

Digital technology and the respatialisation of recording

Digital technology has given the recording studio a kind of freedom, but not without a cost to the industry. World-renowned studios have closed, including Olympic in London (Homer: 2009), Cello in Los Angeles (Milner: 2009, p. 347), Sony Music Studios (Théberge: 2012), The Magic Shop (Welsh: 2016), A&R Recording and the Hit Factory in New York (Rose: 2009), Studio Victor (CBS News Montreal: 2015) and Le Studio in Montreal (Kovac: 2015), and York St in New Zealand in recent times due to the combination of industry downturns and technological pressures.

The Internet, file sharing, and streaming significantly reduced the income potential of the once lucrative recording business (Hull, et al.: 2011, p. 29), returning only after major platforms like Spotify devised often-maligned deals with the majors in the development of a new business model that has yet to turn a profit for the platforms or generate significant returns to artists. Théberge (2012) asks ‘whether the recording studio is in fact threatened with extinction or whether we are witnessing another stage in its evolution, a reconfiguration of the studio as a technology, a means of production and a form of musical practice?’ (p. 79). In his audit of recording studio numbers in the US, Théberge (2012, p. 89) found that the number of studios is in fact growing and not shrinking, just as they are in Australia (Graham: 2013, p. 23). The closure of many larger facilities has led to more opportunities for the autonomous producer/owner/operator (Watson: 2013, p. 331). Homer (2009) adds that many more people are now making recorded music due to the lowered entry barriers afforded by digital technologies (p. 88). As the major budgets shrank in the early 2000s, recordists still needed to find a way to work. An engineer could continue to work in a large-format studio for a reduced rate, or they could buy their own DAW and not charge for studio services, thus retaining their usual studio fee (Dye as cited in, Milner: 2009, p. 339). This trend is certainly a crisis for the large-format studio and Théberge notes that ‘while the overall number of recording studios may be expanding, the status of the large, high-end studio is in question’, referring to the large-format’s place at the pinnacle of recording technology (2012, p. 89). Gibson (2005) maintains that while home recording cannot challenge the mythical status of the high-end studio, it can still become an effective emotional and creative space, interacting with its built environment, in reference to the cultural significance of the recording studio (p. 205).
Methodology

This paper is based on a comparative series of music production projects that were conducted either by recording them in a large professional recording studio (Queensland University of Technology’s recording facilities); by recording in situations dictated by a set of limitations common to smaller budget projects (DIY scenario); and by recording with a hybrid approach of both methods. This paper analyses four recording projects with three different artists, within which each project incorporated a mix of the three distinct approaches, dispersed as evenly as possible during the creation of new original recordings. The research design is as naturalistic as possible. Naturalistic design attempts to undertake research in conditions that are ‘committed to the primacy of natural context’ (Lincoln and Guba: 1985, p. 226). Every recording session involved the creation of new work. To re-record any works in an opposing scenario would not be naturalistic and would reveal a contrived result.

In the large-format studio setting, time spent recording was limited to 1.5 days per song (depending on the project and style of the artist) which, going by experience, is a typical amount of time to record a song. There was a greater time budget for the DIY studio, which is indicative of typical DIY processes. Each song was limited to no more than three days, which allowed time to experiment and capture the desired sound in those spaces. Using the hybrid approach, we recorded as much as practicable in a one or two-day session in QUT’s studio and then overdubbed on these sessions in the DIY studio.

QUT’s studios meet our definition of a large-format recording studio, consisting of a large-format console, numerous selections of outboard equipment, and a separation of control rooms and recording spaces. The DIY recording equipment for the DIY recording scenarios was sourced from QUT’s hire department within the Creative Industries Faculty along with additions from a personal collection. A UAD Apollo system and a laptop are the centre of the DIY setup, which was portable and representative of typical small-budget DIY recordings. No microphone worth more than $600 AUD was used and the entire DIY studio had a value of around $10,000, keeping the budget in line with a typical DIY setup. The cheaper microphones of the DIY setup have a reputation of being lo-fi and not able to capture audio as efficiently as high-end microphones. All recordings were completed in a manner that suited that particular song creatively. Some songs were all overdubbed, one part at a time, and some are mostly performance as a group ensemble, depending on the desired aesthetic for that particular song’s production. Most of the editing took place in that particular studio’s environment, as is our preferred practice. A complete list of the equipment used in the recordings can be found in appendix D.
Multiple professionals assessed the completed recordings so as to achieve maximum reliability on an assessment of quality (Denzin: 1973, p. 297). The completed recordings were sent to recording professionals for critical evaluation of production and sonic quality. The recordings were de-identified and respondents were asked to determine the provenance of the recording scenarios. All producers involved have at least 20 years of experience during which they have gained artistic and critical acclaim. Draper (2013) points out that critical listening is motivated by ‘highly personalised practice-based drivers’ (p. 4). Using non-professionals, or even those who work in the music industry more generally, would have revealed ‘many interpretations and approaches to listening to recorded music’, giving results that would be ‘quite different from many external forms of evaluation schema theorised’ and would therefore be too imprecise to be of use to the research (p. 4). Experienced producers can perceive minute details in the completed mixes and thus may be able to determine in which recording space a recording was made. The producers completed further closed questions to convey their attitudes toward the current circumstances in the recording studio. Those data collections allowed me to draw credible conclusions about relative quality.

**Producer Attitudes**

Twelve producers showed interest in the research and seven completed the survey. They were: Andy Baldwin (Brooklyn); Mike Howlett (UK); Jeff Lovejoy (Brisbane); Jimi Maroudas (Melbourne); Paul McKercher (Sydney); Steve Schram (Melbourne); and Tim Whitten (Sydney). The producers were asked to determine the spatial provenance of the de-identified recordings. They were also asked to answer five structured questions regarding their attitudes toward DIY recording techniques and the current state of the large-format studio. Those questions and the de-identified audio files can be found in Appendix B. The producer responses can be found in Appendix C. Paul McKercher explains how he evaluated the recordings:

Things I listened for as clues to the recordings’ provenance were the depth of the kick drum, the smoothness of cymbals and the ambience around the drums, the noise floor, the quality of the top end in vocal mics, especially sibilants and the amount of clarity and detail in the sounds generally which I assumed was a reflection of the converters used. (personal communication, 27/03/17)

None of the others gave detail about how they listened to the material or about the basis of their judgements. Figure 1 is a screenshot of the results of the listening tests. A green square represents a correct assessment and a red square is an incorrect assessment. The best score was 7 out of a possible 17
(40.18%) with the mean score being 5.14 (30%) out of 17. All respondents picked the song Don’t Suffer Kate by the Oyster Murders as a DIY recording. That track involved found sounds in place of a traditional drum kit and was more easily identified as DIY in origin simply because it would be difficult to achieve in a large-format studio without literally bringing in the kitchen sink. If we removed this track from the list, the highest score would have been 6 out of 16 (37.5%) and the mean 4.28 (26.79%) out of 16.

Figure 1: Producer results.

Five producers correctly determined the provenance of Blossom by Tylea, in which I chose not to remove the sound of crickets from the preferred vocal take. The crickets are clearly audible in the breakdown verse of the song. No other tracks generated any consistency with producer responses. Even though it is a small sample, it is a sample of highly qualified listeners. It is unlikely that untrained listeners would be able to determine whether the recordings were made in a large professional setting or on a laptop in DIY situation based on listening to the final mixes alone whereas this acknowledged industry panel would have the highest likelihood of being able to identify hallmarks of quality recording in multiple contexts. A link to the mixes can be found in Appendix A.

Producer attitudes toward DIY recording

The producers were asked five questions were to gauge their attitudes about the DIY movement. McKercher was the only producer who overtly avoided DIY recording. He argued that artists who want ‘to work DIY are happy to operate their own equipment and produce their own music’ (McKercher: interviewed, 27/3/17). He prefers the immediacy of capturing things quickly in the large-format studio and finds the restrictions of the DIY studio prohibitive stating: ‘Achieving this [a quality production] in a DIY environment can be challenging due to the space and equipment constraints’ (McKercher: interviewed, 27/3/17).

The other 6 producers were positive toward DIY recording, including Steve Schram who enthusiastically stated:

I would not have a career as a mixer without the DIY artists or enjoy myself as much as a producer if forced to only work within the boundaries of a stu-
Mistakes and limitations are crucial to discovery and innovation and present themselves more often in the world of DIY. (interviewed, 22/3/17)

Not only has he drawn links between DIY recording and his livelihood, but also to the abundance of novelty created in the DIY scenario. He comments positively on the spatial effects on creativity in DIY contexts:

A place which may feel homely, or a place removed from the comforts of home. Away from distraction. A producer looking for a space which nobody has recorded in before thereby sounding like no other record made before. Budget restrictions may offer more time for an artist to explore the process. (interviewed, 22/3/17)

Lovejoy added that creativity is more important than fidelity.

The relaxed atmosphere adds to creativity that in my opinion is more important than sound fidelity. There are also more location opportunities like the bedroom or the rehearsal space etc., where musicians and producers can record themselves without pressure of expensive recording budgets of the large format studio. (interviewed, 26/3/17)

Producer Maroudas said that DIY is

… a means to resolving a budget constraint, other times it’s a means of creating structural closure, for instance getting out of town and recording in a holiday house by the ocean. Often, it’s a means of using the budgetary constraint to our creative advantage by working in a unique space that we may otherwise not have. For instance, I have recorded a large portion in an empty two-story holiday house that had unique reverb that created a unique and interesting sonic identity for the album. (interviewed, 2/4/17)

Mike Howlett explains his motivation for working in his home studio: ‘I can control when I work and for how long’ (interviewed, 29/3/17). Producer Baldwin had a similar attitude: ‘[I] like having the artist stay/cook/eat/sleep together, focused, removed from distraction. As far as sonics are concerned, I LOVE the sound of my house’ (original emphasis, Baldwin: interviewed, 16/3/17). Producer Whitten adds to the positivity of a DIY scenario: ‘[s]ometimes cost and sometimes being in an unusual place can have positive/creative effects on people’. (interviewed, 26/3/17)

**Studio necessities**

The respondents were asked what were the most important elements to make a good recording for an artist. Schram identified ‘[t]he skill and personalities of the people they choose to work with’ (interviewed, 22/3/17). Howlett simply added: ‘Comfortable ambience, good headphones’ (interviewed,
29/3/17). Responses highlighted the importance of comfort and communication in the studio. Both those elements are independent of spatial concerns and are available in DIY or large-format environments.

Whitten drew attention to the importance of the space in which you record, adding that ‘a stimulating atmosphere with a varied choice of equipment are the two most important for me. The balance of these two elements changes depending on the project’ (interviewed, 26/3/17). That brings into focus what a stimulating environment should be like. The definition of a creative environment is different for different people but is nonetheless a crucial ingredient in the recording studio. As Csikszentmihalyi states, ‘choosing the wrong environment will probably hinder the unfolding of creativity’ (1996, p. 133).

Other respondents provided a list. The most common items included:

**A comfortable and creative space:** McKercher specifies ‘quality air and light, a cohesive and pleasing interior design that shows some thought and artistry and that inspires artists to do their best work’ with ‘a comfortable and spacious recreation space’ (interviewed, 27/3/17). Maroudas said the space should be ‘isolated and removed from external pressures’ and must also have ‘the ability to easily facilitate the recording of a band’ (interviewed, 2/4/17). Baldwin likes ‘inspiration, energy, comfort and safety (especially for female artists).’ (interviewed, 16/3/17)

**Acoustic requirements:** McKercher likes ‘musically pleasing acoustics, a number of isolation spaces’ with ‘good sight lines’ and ‘a low noise floor’ (interviewed, 27/3/17). Lovejoy identifies ‘a variety of good acoustical spaces with separation options’ and ‘high-quality monitoring with an even frequency response’ (interviewed, 26/3/17) and Baldwin prefers wood and other natural surfaces. (interviewed, 16/3/17)

**Technical requirements:** All specify well maintained, and a good variety of equipment. Baldwin also adds that ‘an inspiring environment is way more important.’ (interviewed, 16/3/17). Many of these specifications could, of course, describe a typical large-format studio but they can also be features of a well-chosen DIY setting.

**Producer attitudes toward large-format recording**

The producers’ responses show that the spatial provenance of a recording is difficult to identify in a final mix recorded by a competent engineer-producer. This research began with questions about the way in which large-format recording studios might be reconceptualised given the current recording studio era. Lovejoy distinguishes between acoustic band performances and purely electronic production approaches.
In my opinion, large-format studios will always be the better setting for recording well-rehearsed organic music played live by humans! Whereas cut and paste music such as electronic dance music has little to no benefit to the large format and would better be suited to DIY recording, as most of the sound manipulation happens inside the computer. (interviewed, 26/3/17)

Most of the producers mirror the thoughts of McKercher regarding the future of the large-format studio:

I see large-format studios as continuing to be useful for certain artists and production requirements, though the demand for, and the number of large-format studios will probably decrease. Large ensembles and orchestras for soundtrack recording will continue to require more spacious recording rooms with low ambient noise floors and the extensive mic kits and number of mic preamps afforded by large studios. Bands that prefer to perform and record simultaneously, to capture the energy they conjure as an ensemble, will need larger spaces to accommodate them along with multiple isolation spaces for their amplifiers. (interviewed, 27/3/17)

While we found the DIY scenario capable for most of the recording requirements for the chosen artists, it was hard to record quieter sounds, particularly autoharp and some vocals. Baldwin adds that

Music has changed so much that they [large-format studios] are becoming less relevant. All that said, as an engineer, it is always such a pleasure working in a beautifully equipped, beautiful sounding studio. That will never change! (interviewed, 16/3/17)

Whitten argued that:

There will always be a need for a “large acoustically designed space or spaces” that allow for musicians to perform together. Perhaps large consoles are not a requirement or a large list of outboard. I think those items are now more about personal choice than need. (interviewed, 26/3/17)

The results of the research show that recording spaces don’t need to be ‘acoustically designed and isolated’ for many types of contemporary music to sound good. However a large space is still an asset. I may have been lucky when sourcing venues for our DIY recording, but it is relatively easy to find an appropriate space through the current Airbnb sharing economy. Although some of the DIY spaces were not large, I still achieved results that made it difficult for professional producers to identify any difference.

Whitten also acknowledges the improvements in DIY technology in such developments. He says that ‘large consoles’ and ‘a large list of outboard gear’ is just matter of personal choice rather than the technical necessity it once was considered to be. Sonically, the research has shown the difference
in audio fidelity among the different recording scenarios is small. However, workflow issues regularly appear in field notes as issues in the DIY space. Producer Schram is of the opinion that the large-format console is less necessary than the recording space but that large-format recording space will endure.

The LFS [large-format studio] will exist for a while to come due to passion of the studio owner but restricted to what a 1 or 2-person operation can offer. The full service of owner, reception, in-house tech, engineer and assistant is no longer sustainable. The large-format console will disappear before the large-format studio does. The space is more critical to multiple genres than the equipment. A larger good sounding space will be in more demand than a large console. Any recording from any format will only sound as good as the song allows and the skill of the mix engineer. (interviewed, 22/3/17)

One implication of the research is that producers will need to become more creative when selecting spaces for recording and more mobile with their equipment. Producer Maroudas takes a distinctly creative angle. However, in describing features of a large-format studio, he could easily be describing those of a DIY space. What’s important, says Maroudos, is the ease for artists being able to create a sonic landscape and for them to be able to be ushered away from the world and into a creative environment where all of the creative tools are ready at hand, make it a relevant, and a necessary ingredient in the production of commercial release music. However, the dynamics of the music industry mean that the economics necessary to be able to build, maintain and keep a large format studio running are presently difficult to sustain. (interviewed, 2/4/17)

McKercher speaks of the efficiencies of the large-format studio:

While the daily cost of larger studios is higher, recording ensembles in a large studio simultaneously can be considerably more time efficient than tracking each member successively on a more restricted setup, making studio recording budgets comparable with those associated with longer schedules in lower daily cost DIY recording environments. (interviewed, 27/3/17)

Of the respondents who didn’t choose the producer as the most integral part of the recording process said that time in the studio was the most crucial part of the process contra to McKercher’s comments about time.

Conclusion

This research has shown that DIY spatial scenarios are today sonically indistinguishable from those of large-format recording spaces. The seven producers who responded to the listening tests were able to pick the provenance of
the recording space 30% of the time, with only 2 songs out of 17 having been identified with any consistency across all responses. This demonstrates that the sound of a large-format studio can be produced in DIY spaces. Most of the producers were positive about the shift to DIY recording and remarked on the abundance of novelty and creativity such situations produce. The impetus for recording in DIY situations tends to be largely financial. But it was also a way to use budgetary restrictions to advantage, allowing more time for artistic exploration, more time to use the studio as the musical instrument it is, or at least can be. Many of the producers listed spatial requirements they felt necessary to facilitate a successful recording. They typically describe a large-format space as most desirable. Many of those requirements can be found in well-chosen DIY settings. Although the producers agree that large-format studios will continue to decline in number, some studios will remain to cater for specific genres and ensemble recording tasks. The need for a large-format console in those studios will be a matter of personal choice rather than necessity.

The recording space does not affect the fidelity of the final recording, at least in the genres explored here. However, the space affected the mood, performance ability, creativity, and productivity of the artists involved. Successful DIY recording requires the right equipment, productive (not necessarily good) relationships, and a skilled technician to achieve good results. The term “good” here refers to sonic and artistic qualities both of which may or may not be required for commercial success.

The recording space is not incidental to recording. It is part of every recording, whether professional or domestic, and so the choice of space needs careful consideration regardless of whether it is a DIY or large format environment. A suitable environment is now as much a part of the recording budget as recording equipment because it is a key factor in achieving maximum creativity from artists. DIY recording is not as technically efficient as large-format studio spaces which are hard-wired and permanently set up to record. The preparation for DIY recording can be immense and, once setup, the lack of time restrictions can lead to spending more time on a recording than might otherwise be the case. Financial necessity drives the use of the DIY studio to some large degree, and considering there are no sonic advantages to recording in large-format studios, attitudes will continue to change, particularly given the current trend towards largely self-funded recordings. Hybrid recording does address some of the shortcomings of DIY discussed here, but even short amounts of time in the large-format studio is financially unfeasible for many self-funded artists.

Having enough time to record is crucial to the success of a recording and the comfort of the artists. However, given lower incomes overall for musicians, time is even more scarce in the current era. The producers saw the future of the large-format studio to continue its decline as a central infrastructure for the recording industry. For the owner of a recording studio an
ability to adapt to smaller, but more frequent jobs, as much as the passion of the owner, will be critical the ongoing operation of any such facility. The novelty of DIY recording can be a catalyst for creativity. However, DIY recording does not suit everyone or every purpose. Repertoire based genres do not require large amounts of creativity and time in the recording studio. The producers propose that orchestras, small string ensembles, and jazz, among many other forms, still require large ambient spaces with a low-noise floor and good separation using isolated rooms.

The era of spending extended periods of time in the large-format studio is limited to very successful artists and is over for the emerging musician unless they enrol in a tertiary education institution that has such equipment. The conservative character of the record industry has once again asserted itself in resisting change and brought the large-format studio to the edge of extinction. Producers have few options other than to adapt their recording approaches to the rise of DIY technologies and spaces.

In the “factory” era of recording, large-format recording studios relied on a healthy music business to pay for studio time, usually as an advance against the artist’s future royalties. As recording revenues have declined, DIY technologies have increased in quality, such arrangements are rare. The current economic state of the recording sector has forced producers to embrace the DIY studio creating further growth in portable recording technologies and new recording paradigms more generally. The high level of creative labour required for recording quality music demands some degree of interplay with the DIY studio. Even when the intention is to record entirely in the large-format studio, artists almost invariably include tracks from a demo, or other elements recorded “at home” on the artists’ own equipment, in the final recording. The DIY studio represents the domestication of an industrialised production process and, therefore, the resurgence of a “cottage industry” approach to production, even while distribution becomes an entirely global and instantaneous affair.

References

Tape Op.
Journal on the Art of Record Production. 3. Available at: http://arpjournal.com/producer-compensation-
pp. 192-207.
Hennion, A. (1989) An intermediary between production and consumption - The producer of popular-
Homer, M. (2009) Beyond the studio: The impact of home recording technologies on music creation and 
Horning, S. S. (2013) Chasing sound: Technology, culture, and the art of studio recording from Edison to 
the LP. Baltimore, Maryland: Johns Hopkins University Press.
Howlett, M. J. G. (2009) The record producer as nexus: Creative inspiration, technology and the record-
Hracs, B. J. (2015) Cultural intermediaries in the digital age: The case of independent musicians and 
Doctor of Philosophy. The University of North Carolina at Chapel Hill, ProQuest Dissertations 
Publishing.
207-220
Art of Record Production, 6. Available at: http://arpjournal.com/performance-recordivity-studio-
Leyshon, A. (2009) The software slump?: Digital music, the democratisation of technology, and the 
decline of the recording studio sector within the musical economy. Environment and Planning A., 41, 
6, pp. 1309-1331
Francisco, Milwaukee, WI: Miller Freeman Books.
Straus and Giroux.
In E. Pekkilä, D. Neumeyer & R. Littlefield (Eds.), Music, meaning, and media. Helsinki: Interna-
tional Semiotics Institute, pp. 174-191.
Slater, M., & Martin, A. (2012) A conceptual foundation for understanding musico-technological creativi-
ty. Journal of music, technology & education. 5, 1, pp. 59-76.


Discography
https://www.dropbox.com/sh/i36488qqio599kf/AAB1pD0CnkJsw6HyVoVt_uia?dl=0

Appendices

Appendix A: Ethics approval forms - https://www.dropbox.com/s/3hwi5v248c9d6er/ETH_Info-Consent_Interview_20170303.pdf?dl=0

Appendix B: Producer’s survey - https://www.dropbox.com/s/3f1jpeal0h6h8t2/Producers%20survey%20final.docx?dl=0

Appendix C: Producer's response - https://www.dropbox.com/s/1ace092yl9h72c/Producer%20answers.docx?dl=0


Acknowledgements
We’d like to acknowledge the Queensland University of Technology who funded our participation in the 2017 ARP conference at the Royal College of Music in Stockholm, Sweden.
Tor Halmrast: Sam Phillips’ Slap Back Echo;

Luckily in Mono

Abstract

‘Slap back echo’ was created by Sam Phillips for Elvis Presley’s Sun recordings and later simulated by RCA. Using cepstrum and autocorrelation, we find that the tape delay used in Sun Studios was 134 - 137 ms, which is so long that the echo is perceived mainly as a distinct echo in the time domain, more than a coloration of timbre in the frequency domain. Even though the delay time is long, the echo is still perceived as rather “intimate”, because the echo is in mono. Panned in stereo, the feeling of being inside a quite small room would disappear. The simulations by RCA, using a hallway instead of tape echo, shows somewhat shorter and less pronounced delays.

Introduction

Elvis Presley’s first recordings were done by Sam Philips in Sun Studios in Memphis. He created a tape echo by using two Ampex 350 tape recorders. This method was used for, amongst others, Tryin’ to Get to You, and Baby Let’s Play House from 1955. The story goes that when Elvis changed to the bigger company RCA Victor’s studio on McGavock St., Nashville, “RCA was anxious to recreate the ‘slapback’ echo effect that Sam Phillips had created at Sun. To add them to Elvis’ vocals Chet [Atkins] and engineer Bob Farris created a pseudo ‘echo chamber’ by setting up a speaker at one end of a long hallway and a microphone at the other end and recording the echo live”1. This was done for the recording of Heartbreak Hotel (Jan. 1956). A third type of echo/reverb analysed, was for the RCA Hollywood recordings of amongst others Long Tall Sally (Sept. 1956).

The delay time of the slap back echo used by Sam Philips has been discussed, mostly on web sites. Steven Trent2 argues that “The delay time should be between 60 and 120 milliseconds”, and that: “At 60 Milliseconds, you don’t hear much doubling, but everything sounds a little thicker and

1 http://www.history-of-rock.com/sam_phillips_sun_records.htm
2 https://steventrent.com/omblog/blog/what-is-slapback-echo
fatter”, and that “At 120 Milliseconds you are really starting to hear the distinct doubling of the sound”. Elvispresleymusic.com discusses a suggestion of 530 ms, but find that to be “too long”, and argues that 163 ms “seemed to sound about right”. The Wikipedia page on Slapback Echo mentions “75-250 ms, with little or no feedback”.

Pictures and info from the studios indicate that Sun Studios has acoustic tiles on most walls, probably high-/mid-frequency absorbers, and the room is described as so dry that “When you speak, you can feel the air pressure in the room”. Some bass resonance problems have been reported, (as such tiles seldom absorb lower frequencies), but this should not have influenced the frequency range of Elvis’ voice too much. Pictures from RCA Nashville, from the web site show thick curtains and what seems to be mid-/high frequency wall absorbers, so the conclusion must be that there should be no possibility of getting a natural, distinct, separate, late echo for a vocal inside these rooms.

Overview

Searching for recordings with Elvis unaccompanied ‘a cappella’ sections, we found two examples from Sun Records (Baby Let’s Play House, and Tryin’ to Get to You), and two from the early RCA period (Heartbreak Hotel and Long Tall Sally). For details about the recordings, see References. We investigated the presence of echo in the recordings by analysing Power Cepstrum and Autocorrelation. Some additional analyses included Linear Predictive Coding and reverberation times (or more correctly: decay times, as the decay is far from linear). The analyses were performed in Praat, Pure Data, MIR Toolbox (MatLab), Librosa (Python) and Arta.

The paper starts with the investigations of the unaccompanied ‘a cappella’ parts of Elvis’ Sun recordings, and here we find that they have one distinct, single reflection with a delay time of 134 - 137 ms. For the RCA recordings analysed, the decay is more smooth (not giving such distinct echo) and the delay times for the strongest reflections are somewhat shorter; 82 ms in Nashville and 117 ms in Hollywood. More technical information regarding the analysis methods; Comb Filters, Cepstrum and Autocorrelation is given in a special section.

Even if the delay times are very long, the recordings are still perceived as being quite ‘intimate’, so, in the next section we investigate if the measured time delays could provide what is defined as comb filter coloration, and we

4 http://www.history-of-rock.com/sam_phillips_sun_records.htm
5 http://www.scottymoore.net/studios.html
6 http://www.history-of-rock.com/sam_phillips_sun_records.htm

138
find that the delay for Sun Studios is so long that the ‘boxy’ sounding timbre ‘Box-Klangfarbe’ as defined by Halmrast (1999 and 2000) is not perceived. (Klangfarbe is German for Timbre). Also the delay times for RCA are too long to provide such Box-Klangfarbe, but for these recordings, the overall reverberation decay somewhat masks the measured delayed ‘reflection’, so that a clear, distinct echo is not as easily perceived as for the Sun recordings. Even if the delay times are so long that we do not get the kind of ‘close-ness’/‘intimacy’ typical for Box-Klangfarbe comb filter coloration, Elvis’ voice is clearly perceived as being rather close, or at least inside a small room. In the final part we discuss that the reason for this is that the recordings are in mono, not stereo, which has been verified in a simulation/listening test.

**Sun recordings**

Two recordings from Sun Studio with audible slap back echoes were analysed: *Tryin’ To Get To You* and *Baby Let’s Play House*.

*Tryin’ to Get to You*

The Power Cepstrum of the ‘a cappella’ opening of *Tryin’ to Get To You* is shown in figure 1. We see that the delay time for the echo is 0.137 s = 137 ms, which is confirmed also in the Autocorrelation analysis in the lower pane of figure 1. The x-axis of the Power Cepstrum is shown in seconds, but as explained later, quefrency along the x-axis in figure 1, upper pane, is a ‘sort of time axis’. In the theoretical part we explain that also pitch provides results in cepstrum analysis, so we have to check that the pitch is separated from our detection of echo in figure 1. In figure 2 we see that the fundamental pitch is mainly in the region \( f_o = 250-300 \text{ Hz} \), corresponding to quefrencies \( (1/f_o) \) between \( 1/250 = 0.004 \text{ s} = 4 \text{ ms} \) and \( 1/300 = 0.0033 \text{ s} = \)
3.3 ms. From the upper pane in figure 1 we see that the ‘voice’ thus occupies the lower part of the Power Cepstrum.

A delay time of 137 ms corresponds to an excess travelling path for the reflected sound of 343 x 0.137 = 47 m, assuming a speed of sound of 343 m/s. If you imagine standing in front of a wall, listening to the echo back to yourself, such a delay time means that the wall must be 47/2 = 23.5 m away.

**Baby Let’s Play House**

The pitch analysis to the left in figure 3 is inaccurate, due to the very short syllables in the “a capella” opening of *Baby Let’s Play House*. We could have performed additional analysis with a shorter window for the Fast Fourier Transform, but then the uncertainty in frequency would increase, and, as long as we know the main pitches of Elvis voice in this part, we have enough information to know that we have separation of the voice and reflections in the cepstrum analysis. From the panes to the right we see that the delay time for the slap back echo is 134 ms. We find indications of 134 ms also by looking at the Acoustical Energy Decay for the 250 Hz octave, which is shown in figure 4, together with the corresponding Schroeder curve, (backwards integration of the decay), but this is not a clear method for analysing echoes. Reverberation time for the decay of a voice burst including an echo is not really defined, but for this 1/1 octave band around 250 Hz (which includes the fundamental frequency of this sound file), T30 is app. 0.4 seconds. (If such a reverberation time was measured in a real room, it would be categorised as something between a well-furnished living room and a dampened studio). The small difference between 134 ms for this recording and 137 ms for *Tryin’ to Get to You* is might be due to measurements uncertainty, possible small stretch of tape etc.
RCA recordings

**RCA Nashville, ‘Heartbreak Hotel’**

From figure 5, upper pane, we see that the fundamental pitch is about 330 Hz (the note E, a fourth below the Chamber Tone A=440 Hz). In the lower pane, we see a “reflection” at 82 ms for *Heartbreak Hotel*, which is shorter than the measured tape delay from Sun Studios. If we assume that the information about the ‘pseudo echo chamber’ mentioned in Part 1 is correct, this measured delay time could indicate a reflection path in the hallway of 343 x 0.082 = 28 meters, assuming a speed of sound of 343 m/s. (This might appear to be somewhat long for a common hallway, but the actual set-up in the hallway is not known. It could, theoretically, also be that both the microphone and the loudspeaker were positioned at the same end of a hallway with half of this length). The perceived reverberation for this recording provides less impression of...
a distinct echo than the slap back echo from Sun Studios. This is probably due to many diffuse reflections also from the side walls of the hallway that arrive earlier than the main ‘echo’ from the end of the hallway.

**RCA Hollywood, “Long Tall Sally”**

From the middle pane of figure 6 we find that for *Long Tall Sally*, there is no strong indication of a distinct echo, but we see a weaker reflection at 117 ms. From the zoom-in of the Power Cepstrum in the lower pane, we could speculate if there is a short reflection at 9.7 ms. Such a short delay could have been due to a floor reflection, but, as discussed in the Appendix, this is interesting, but somewhat questionable.

**Analysis Method: Comb filters; Cepstrum; Autocorrelation**

**Comb Filters**

General information and theory on comb filtering due to an added, delayed reflection (echo, delay) is given in Halmrast (1999 and 2000). A short summary is that when a delayed reflection with the same strength is added to a signal, some frequencies will sum up in-phase, giving +6 dB. Exactly in-between these frequencies, the delayed reflection and the original signal will be out of phase, and there will be sharp cancellation to (theoretically) minus infinity dB. Using a linear frequency scale, the resulting spectrum looks like a comb, and is called a comb filter. In Halmrast (1999 and 2000) CBTB (Comb-Between-Teeth-Bandwidth) is defined as the distance (in Hz) between the succeeding teeth (and also between the succeeding peaks, but they are often not as easily detected). A short reflection provides very broad combs, with very large CBTB, and could be looked upon as a simple bass boost. A very long delay for the reflection/echo provides very small CBTB, so the teeth in the comb filter will be so close that we do not perceive the changing in timbre, but the reflection is perceived as a distinct echo in the time domain.
We will start with discussing a click signal with an added a delayed reflection/echo. Figure 7 shows the signal (or what we could regard as an Impulse Response). The delay is chosen to be 137 ms (as for Elvis’s Sun recording in 3.1), and, for simplicity, the reflection is equally strong as the direct sound.

```
<table>
<thead>
<tr>
<th>0</th>
<th>1000</th>
<th>1010</th>
<th>1020</th>
<th>1030</th>
<th>1040</th>
<th>1050</th>
<th>1060</th>
<th>1070</th>
<th>1080</th>
<th>1090</th>
<th>1100</th>
</tr>
</thead>
<tbody>
<tr>
<td>dB/Hz</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>
```

**Figure 7:** Upper pane: A pulse (dirac pulse) with a repetition (echo, delay) of 137 ms. Lower pane: Zoom in of the frequency response of the pulse with reflection, showing the comb filter.

In the lower pane of figure 9 we see the comb filter shape. The figure is zoomed in to frequencies between 1000 Hz and 1100 Hz, but the shape of this curve is repeated through the whole frequency range, from 0 dB. We see that the distance between the dips is CBTB = 129.3-122.0 = 7.3 ms. (The ‘distance between two peaks’ is also 7.3 ms, but the peaks are not that easily detected). The ‘distance’ is CBTB = 1/Δt [Hz], where Δt is the difference in time (in seconds) between the delayed echo and the direct sound.

The CBTB is dependent only on the time delay, but the amplitude of the peak and also the depth of the dips depend on how strong the reflection is. For the general understanding of comb filters as a result of adding reflection(s), it could be mentioned that if a comb filter has some kind of repetition/feedback, the peaks gradually becomes sharper and the dips wider in frequency so that the comb shape will be gradually be transformed into upside-down. The feedback could be electronic or natural. Blowing a flute is such a feedback system of repetitive reflections between the blowing point and the flute’s end/tone hole. The fundamental and harmonic partials are
shown like regular, sharp peaks in frequency and this shape is actually the comb shape from figure 7 upside-down, (often with gradually falling amplitude).

Cepstrum
The word Cepstrum comes from reversing the first four letters of ‘spectrum’, and is chosen because we analyse the ‘rhythmic behaviour’ of the spectrum, or, the somewhat un-academic: ‘spectrum of the spectrum’. As we see from the lower pane of figure 7, the shape of the frequency spectrum repeats itself every $1029.3 - 1022 = 7.3$ Hz. This is what we find by cepstrum analysis. Cepstrum is the inverse (Discrete) Fourier transform (IDFT) of the log magnitude spectrum, and a float diagram looks like figure 8.

![Figure 8: Float diagram of Cepstrum analysis.](image)

There are several versions of cepstrum: complex cepstrum, real cepstrum, phase cepstrum and power cepstrum, and the result can be shown as a plot of cepstrum over time, example: Power Cepstrogram. The version we used for our analysis is the Power Cepstrum, which is the squared magnitude of the inverse Fourier transform of the logarithm of the squared magnitude of the Fourier transform of a signal, which can be written as equation 1.

$$Power \text{ Cepstrum} = |\mathcal{F}^{-1}[\log (|\mathcal{F}[f(t)]|^2)]|^2 \quad (1)$$

where $\mathcal{F}$ is Fourier the transform (DFT), and $\mathcal{F}^{-1}$ is the inverse Fourier transform (IDFT). The x-axis in cepstrum analysis is called Quefrency. It is shown in seconds, but is a sort of “illusionary” time scale. Power Cepstrum of the pulse with added delayed reflection (echo) at 137 ms from figure 8 is shown in figure 9.

![Figure 9: Power Cepstrum of a pulse with a reflection at 137 ms](image)

We see the peak at 137 ms, which we know (because we introduced this delay), so this peak provides the time delay correctly for this example. But
we also see a (smaller) peak at the double value, and we could find smaller and smaller peaks also at triple value etc. This is because we do not really measure ‘time’, but how often the spectrum (in fig.7) ‘repeats itself’, and from figure 9 we can see also ‘repetitions’ between every other or every third etc. peak (or dip) in the spectrum in figure 7.

For this test, with a just pulse as signal, the result from the cepstrum analysis was quite clear, but for a more common (musical) signal, the quefrency value found by cepstrum analysis provides information about every kind of ‘repeating behaviour’ in the spectrum of the signal, not only comb filters due to echoes etc. Another type of signal that provides ‘repetitions’ in the spectrum is a spoken or sung voice that provides the fundamental pitch and the overtones as multiples of the fundamental. This will also appear in the cepstrum analysis of the voice, but luckily, in a somewhat separate region along the quefrency axis, as pointed out in figure 1. Figure 10 shows the Power Cepstrum (The evolution of Power Cepstrum over time) for Elvis’ Sun recording, zoomed in to the range where one can use Cepstrum for analysis of pitch (in order to provide results like in fig. 2).

Room resonances, formants and flutter echo will also appear in the cepstrum. Figure 11 shows a typical overview, with a “zoom-in” in the lower pane.

We have marked the lower limit for echo as 50 ms, which is a typical limit for echo for speech given in the literature (possibly mostly for western lan-
The limit for detection of distinct echo depends on type of signal, and clicks or transient instruments (marimba etc.) have shorter limits. (see Halmrast 2000). As an example of how the human ear works, we can notice that a repetition of 50 ms corresponds to a frequency of $1/0.050=20$ Hz, which is known to be the lower limit for humans to hear sounds as “tones”. That means that quefrencies shorter than 0.05 is (might be) perceived as pitch, and that the smaller the quefrequency is, the higher the pitch.

**Autocorrelation**

Correlation between two signals provides information about “how alike” they are (cross-correlation). Autocorrelation shows to what degree a signal ‘looks like itself’ after different times. Both Autocorrelation and Cepstrum can be used for analysing pitch of complex, harmonic signals, and both also provides information regarding echoes. The differences in analysing methods and results are not discussed further here. Technically, Autocorrelation provides values also for negative time, but for sound signals, negative time is, of course, skipped.

**Timbre/Comb Filter coloration**

Investigations in rooms show that if a distinct, separate reflection provides a comb filter with a CBTB in the order of critical bandwidth along the basilar membrane in the cochlea, the added reflection is perceived as if the source was placed in a small ‘boxy’ room, and such influence of the timbre is called ‘Bok-Klangfarbe’-coloration in Halmrast (1999 and 2000). This is indicated in figure 14 as Box-Klang-Zone.

On the x-axis in figure 12 we have the time delay, $\Delta t$ in milliseconds, (shown in negative direction for a logarithmic scale along the x-axis) and below that, the corresponding excess pathway for the delayed echo, shown in the drawings below as typical sound pathways from a source, reflected back to sender from surface at different typical distances. The bottom x-axis shows the corresponding centre frequencies of the Comb Between Teeth Band (CBTB = $1/\Delta t$ in Hz, also in a logarithmic scale, but since it is the inverse of the time axis, this axis is in the positive direction). The y-axis shows bandwidths in

![Figure 12: Comb filter from reflection, compared with critical bandwidth. Box-Klang-zone indicated.](image-url)
Hz. Two curves show the Critical Bandwidth along the basilar membrane in the cochlea for different centre frequencies, one curve from measurements in acoustic literature, and the other for the more ‘mathematical’ ERB; Equivalent Rect. Band. The difference between these data for Critical Bandwidth of our hearing is not very important for our study. Both these two ‘psychoacoustic’ curves for our Critical Bandwidth are compared with the curve for CBTB (mirroring the lower x-axis, but remember that the y-axis is linear). The drawings to the left of the y-axis shows typical shapes of the comb filter corresponding to the time delays. (Such shapes of comb filters due to reflections with different time delays will be shown more in detail in fig. 13).

In Halmrast (1999 and 2002) it was found that when listening in ‘real rooms’, a ‘boxy’ timbre was perceived when a distinct, separate reflection (delay) with a CBTB comparable in size with the Critical Bandwidth is added to a broad banded signal (like a singing voice). This region where the curves cross in figure 8 is called the ‘Box-Klang-Zone’. (The curves marked A, B and C are results from a psychoacoustic investigation for noise signals with added reflections, see Salomons (1995), and will not be discussed further).

Do the echoes provide “closeness” due to comb filter coloration? Figure 13 shows the typical, simplified comb filters created by the different time delays measured from the recordings. (Ca. 130 ms for Sun and 80 ms for RCA). The figure shows spectrum (linear frequency) from 0 to 200 Hz, and this pattern is of course repeated all the way up in frequency.

In general, long delays provide very small CBTB, so that the “combs” are not perceived in the frequency domain, and the reflection is perceived as a distinct echo in the time domain. For very short delays, the reflection provides a rise for the bass (as a low pass filter). For a medium delay time (example: the 9.7 ms shown in the lower pane of fig. 13), CBTB is in the order of critical bandwidth. Figure 14 shows the delays measured from the recordings, compared with the results for Box-Klangfarbe Coloration from figure 12, cfr. Halmrast (1999 and 2000).
We see that we must have shorter delay times than measured for Sun Studios, (and also for RCA), in order to get a clear Box-Klangfarbe coloration due to comb filtering. As an example, an imaginary, strong, distinct ‘floor reflection’ (9.7 ms, shown in the black dotted rectangle in fig. 14) would be in the ‘Box-Klang-Zone’). As shown in Appendix, such a floor reflection is interesting, but questionable, and in any case not strong enough to provide ‘intimacy’, but it is included here, in order to show what kind of short delays would be necessary in order to get Box-Klangfarbe.

Tape recorders and mono/stereo

Creative use of tape recorders

The modern, magnetic tape recorder was introduced by BASF/IG Farben in Germany in the 1930’s, and was intended to be just a machine for recording sounds. But creative people soon started to experiment using these machines for musical composition and sound effects. Pierre Schaeffer (1910-1995) made his Etude aux chemins de Fer in Radio France 1948, one of a series of short ‘studies’, which used the effect of playing recorded material at different speeds. Later, at about same the time as Sam Philips’ recordings of Elvis, Karlheinz Stockhausen (1928-2007) experimented in Germany, using tape loops played back at different speeds and also re-recorded in a very reverberant room. However, his most important composition, Gesang der Jünglinge (1955–56) was composed after Elvis’ Sun recordings. We cannot suspect that Sam Philips was aware of these European ‘contemporary’ music experiments. It is more likely that Sam Philips could have heard the creative use of tape recorders in more ‘popular’ styles of music, like how Les Paul (1915-2009) and others used the tape machine for ‘sound on sound’, for instance Lover (When You’re Near Me) from 1949.
Mono-Stereo

All these examples of creative use of the tape recorder were in mono. It is true that Alan Dower Blumlein at EMI/Abbey Road patented ‘stereo’ already in 1931, but at that time it was only used in a few very special cinemas etc. For commercial recordings, stereo was not much used before Abbey Road/EMI ‘rediscovered’ it in 1956, so mono was the ‘only choice’ at the time of the early Elvis’ recordings.

Adding more comb filters

The slap back echo is often described as giving Elvis more ‘power’ (as for a ‘King’), but on the other side, the echo still provides some kind of ‘intimacy’, as being inside a room. As long as a mono signal is played back thru just one loudspeaker, the only dimension possible to increase is the depth, by adding an echo ‘behind’ the signal/singer. Such an echo arriving from the same angle as the direct sound is not very common in real life, other than in fact; in a corridor, behind the talker/singer, just like RCA tried. Such echoes are not much discussed in the literature, but a corridor could, at same time, be considered to be both small/intimate and long/big.

The perception of a comb filter (a delay) depends mostly on the time delay: Short reflections give coloration in the frequency domain and long reflections give echo in the time domain. The strength of the comb filter effect depends on the relative strength of the reflection compared to the direct sound. Comb filtering is the result of ‘mixing’ a signal with a reflection. A comb filtering occurs no matter where this ‘mixing’ appears: a) on the site of the recording, b) in the mixing console/tape machine or c) in the listening room. This means that a listener might hear several comb filter effects overlapping, depending on the loudspeaker set up etc. When several comb filters overlap, they might: 1) increase the comb filter effect (only in special situations where the delay times are multiples of each other, for instance a flutter echo), or, more common: 2) somewhat cancel each other out (if there are several reflections with more ‘randomly’ distributed delay times). For the playback of a mono signal with an added reflection (like Elvis’ Sun records) thru one central positioned, single loudspeaker, the comb filtering/echo on the record will of course be the present for all listening positions in the room. If this mono record was played back thru two (or more) loudspeakers, there will be added comb filtering(s) due to the time difference between the arrival of the signal from speaker 1 and speaker 2, (if the listener is not sitting ‘dead centre’). This might, in some positions, lead to a decrease of the perception of the slap back echo on the record. Moving the second loudspeakers very far away (137 ms equals 47 m assuming a speed of sound of 343 m/s) and to the side (while keeping the level at the listener constant!), the delay will be perceived more and more like a distinct slap back echo in the time domain,
but arriving from a totally different direction than the direct sound from the first loudspeaker. We can imagine that this set up will not give any kind of intimacy.

In addition to such comb filtering (echoes) created by the listening room/loudspeaker positions, we have some additional effect due to anatomy of our head. When the signal from one loudspeaker in a room reaches the listener, the arrival time will be slightly different at the two ears, typically between 0 ms when the source is right in the front and 0.5 ms directly from the side. We will not go into details about spatial/binaural hearing, but this, and other effects provide what is called ‘binaural decorrelation’ (see Salomons 1995).

The conclusion is that the perceived effect of comb filter coloration and probably also slap back echo is somewhat reduced when listening to a mono signal played back through several loudspeakers and especially for the situation when the set up resembles stereo.

‘Artificial stereo’

After stereo had been introduced, some companies wanted to ‘renew’ their old mono recordings. Such methods of creating ‘artificial stereo’ were of course introduced later than the Slap Back Echo recordings by Sam Philips, but the method is included here in order to show that the perceived effect of a delayed reflection panned in stereo is highly dependent on the time delay. Very short delay comb filtering, typically <15 ms, was a common method for ‘Mono Electronically Remixed to Stereo’ or ‘Duphonics’ introduced by Capitol Records in 1961:

- splitting the mono signal into two channels,
- delaying the left and the right signals
- desynchronizing the two channels by fractions of a second,
- cutting the bass frequencies in one channel with a high-pass filter,
- cutting the treble frequencies in the other channel with a low-pass filter.
- In some cases, the effect was enhanced with reverberation and other technical tricks”.

This explanation about the filtering is probably not quite correct. Based on what we now know from the theory about comb filters, the resulting comb filtering is somewhat more complex than the discussed high-pass/low-pass filtering a common method was to mix a short delayed ‘reflection’ with the direct sound in-phase for left channel and out-of-phase for the right channel. The result will be that the comb filter for the out-of-phase channel will be shifted $\frac{1}{2}$ CBTB. This is probably the method introduced by RCA as Elec-

---

tronic Stereo Reprocessing System (ESP) in 1961, and used for the processed stereo versions of Elvis first albums. They sound quite weird, and almost all ‘intimacy’ is lost. For some tracks there is almost a high/low filtering of Left/Right, as indicated for Capitol above, but that is probably due to setting the time delay of the comb filter too low, so that the first teeth in the comb filters moves towards higher frequencies, and the result is almost like a low pass filter in the left channel and a high pass filter in the right, but in the high treble, the next peaks of the comb are audible also in the left channel. The main idea is that comb filters are the result both for artificial stereo and Slap Back Echo in mono. The main difference regarding the comb filtering is that the time delay is much longer for Slap Back Echo.

Research on echoes
Echoes in rooms have been investigated by numerous researchers, especially in order to avoid them in concert halls. The most common echo criterion was introduced by Dietsch (1986) but this does not give full answer regarding the perception of an echo, and includes very little discussion about the angle of arrival of the echo. A strong echo with a delay time as long as 134-137 ms is clearly perceived as a distinct echo, so the question about the possible difference in perception for different angles of arrival is not discussed in common literature. For the design of concert halls, Barron (2010) states the importance of lateral (sideways) reflections in order to give ‘envelopment’ for acoustic music. The results are mainly given for lateral reflection arriving from sidewalls and thus from a different angle than the direct sound, but a delay time of 134-137 ms is way out of the region of time delay discussed for beneficial lateral reflections in concert halls.

Simulation of slap back echo panned in ‘real’ stereo
All the recordings of Elvis in Sun Studios were in mono. To be absolutely sure that no ‘Electronically Re-channeled for stereo’-like effect had been added later, this was checked by comparing the autocorrelation of one channel (Left) and the cross correlation of both channels (Left and Right), see figure 15 as an example (for Heartbreak Hotel), and we see that the curves are identical.
As there is not much in literature about the difference in perception of long echoes in mono/stereo it was necessary to perform our own simulation test. We gave a short extract of Elvis’s voice (without any clear slap back echo) a single, distinct echo hard panned in stereo. From listening tests and the audience’s reaction when these files were played at the ARP conference in 2017, it is clear that a simulated time delay of 134-137 ms in the other channel makes the whole ‘room’ disappear, sounding like a distinct open air echo from one single building positioned to the side of the listener.

Summarised we find that adding a late delayed reflection was a most convenient way to “fatten” the sound around 1955, and that Sam Philips Slap Back Echo only works in mono.

**Conclusions**

Our measurements of Slap Back Echo from Sun Studios (*Baby Let’s Play House, Tryin’ to Get to You*) show that this is a single tape echo in mono with a time delay of 134-137 milliseconds. The delay for the RCA, Nashville recording (*Heartbreak Hotel*) is less distinct and shorter, app. 82 ms, which might correspond to a reflection path in a hallway. In RCA Hollywood there is a reflection after 111 ms, but this is even more masked by some general reverb.

Because the slap back echo from Sun Studios is in mono, the delayed reflection is perceived as somewhat ‘intimate’, even if the delay is so long that they do not provide comb filter coloration (Box-Klangfarbe) in the sense defined in Halmrast (1999 and 2000). If a slap back echo was panned in stereo, the intimate feeling of being in a room would have disappeared.

**References**


Dr. thesis Technical University, Delft, Netherlands


[https://steventrent.com/omblog/blog/what-is-slapback-echo](https://steventrent.com/omblog/blog/what-is-slapback-echo) (last visited 20180312)

[https://steventrent.com/omblog/blog/what-is-slapback-echo](https://steventrent.com/omblog/blog/what-is-slapback-echo) (last visited 20180312)


Discography
Presley, E. (1955) *Tryin’ to Get to You* and *Baby Let’s Play House* Sun Studios, Memphis
Presley, E. (1956) *Heartbreak Hotel*, RCA Victor’s studio, McGavock St., Nashville

For recording dates, see: [http://www.keithflynn.com/recording-sessions/50_index.html](http://www.keithflynn.com/recording-sessions/50_index.html)

All recordings taken from The Complete Elvis Presley Masters 30 CD set

**APPENDIX**

**IS THE SPIKE AT 9.7 MS IN FIGURE 7 A WEAK FLOOR REFLECTION?**

We have found that the slap back echoes have a too long decay time to provide comb filter coloration (Box-Klangfarbe). In the search of any shorter delays in the cepstrum curves: Could the small spike at 9.7 ms in figure 6 be a floor reflection? If so, the distance from the singer’s mouth down to the floor and up to the microphone must have been $343 \times 0.0097 = 3.33$ m. (assuming a speed of sound of 343 m/s). This is (theoretically) possible, as Elvis was 1.82 m tall, so that his mouth (and the microphone) could be positioned 1.67 m above the floor. Such a delay is indicated with a dotted black rectangle in figure 9. A 9.7 ms reflection might of course also have arrived from the ceiling or a wall positioned 1.67 m from the singer’s mouth/microphone, but as mentioned in Part 1, this is unlikely in the actual rooms, as the walls are absorbent in the photos.

![Power Cepstrrogram Long Tall Sally](image)

*Figure 16: Power Cepstrrogram of “Long Tall Sally”. Quefrencies corresponding to the fundamental pitch indicated as the lowest number to the right.*
If we investigate this further, we see that such a short delay time is almost in
the order of the period of the fundamental frequencies of the male voice and
we need to distinguish between the ‘voice’ and the ‘reflection’ in the
Cepstrum analysis. From figure 16 we see that a mean value for the funda-
mental pitch in this short clip of *Long Tall Sally* is around 330 Hz, which
corresponds to a period of 1/330 = 3 ms. This is also detected in the Power
Cepstrogram (Power Cepstrum over time) shown in figure 16. The lowest
number on the right hand side of figure 16 indicates the same result, a funda-
mental pitch of 1/0.003 = 330 Hz.

![Autocorrelation](image)

*Figure 17: Autocorrelation of Linear Predictive Coding of “Long Tall Sally” with
white noise as excitation.*

The fact that we do not see an additional, strong spike at 9.7 ms is a first
indication that there might be no strong, distinct floor reflection with a delay
of 9.7 ms. (also because 9.7 is close to 3 x 3 = 9 from the voice). In order to
try to eliminate both the fundamental and overtones of the voice, Linear
Predictive Coding (LPC) was performed; exchanging Elvis’ voice chords
with white noise as excitation. The result sounds like Elvis whispering
“Gon’na Tell Aunt Mary”. Power Cepstrum of this sound file (unfortunate-
ly) gave no clear indication of a strong, distinct reflection after 9.7 ms be-
cause the LPC transform smooth the attacks (and thus the echo). Therefore
we instead performed autocorrelation of this sound file. From figure 17 we
might see that there is a somewhat higher value for app. 9.7 ms, but this is
not very strong and also ‘smeared out’ in time. The result is by no way clear,
but indicates that if this was a floor reflection, the singer must have moved
his head somewhat during the recording, so that the distance from the floor
changed over time,….a highly reasonable assumption for Elvis!
Abstract
As established in a prior research study, spectral clarity is an important parameter of high-quality mixes. Two predictors for this are the harmonic centroid (a weighted centre mass of energy of a sound spectrum) and spectral inconsistencies related to sharp peaks roughly in the middle of the frequency spectrum (Hermes et al., 2017). The impact of these predictors on the creative process is tested and starting points for further research are established in this paper.

Introduction
Mixing music is a complicated process where several tracks of recorded audio are combined to an overall piece. Difficulties can arise from e.g. time constraints or lack of expertise. Research in automatic mixing seeks to work towards the development of perceptually informed assistive or fully automated mix tools. In a previous PhD research study (Hermes et al.: 2017), two predictors were developed for one important parameter of music mixes, i.e. single sound spectral clarity. As the author is also a creative practitioner (artist and producer), the current paper includes discussion on testing the two spectral clarity predictors in the mix process of an electronica production. The aim here is to assess the usefulness of the predictors and to develop starting points for follow-on research. Another aim is to discuss different research methods that can be employed in furthering the understanding of music mix parameters. The author argues that an interdisciplinary approach, drawing on both scientific and creative knowledge, can yield particularly useful results.

The structure of this paper is as follows. Section 1 introduces the field of automatic mixing, to provide readers that are less familiar with this area of research with an overview of aims and methodologies. Section 2 is a sum-
mary of the author’s PhD findings. In section 3, two predictors of single sound spectral clarity are tested in the mix process and findings are related to relevant literature. Section 4 is a discussion with suggestions for further research.

What is automatic mixing?

The democratisation of audio technology and advancements in internet file sharing have resulted in the delocalization of professional recording studios and the decline of traditional record companies (Pras et al.: 2013, pp. 612–626). Almost anyone can create musical outputs and share these online, hence many musical artists are now self-produced (Bell, 2014, pp. 295–312). The large number (982) of music related degree courses offered in the UK indicates that there are many such artists (What Uni: 2014). Not only music artists but also filmmakers and entrepreneurs use media tools which they may not be experts in. Music mixing is equally important in the context of live music. Audience attendance at UK live music events has recently increased by 12% (Ellis-Peterson: 2017) and live sound mix engineers often work under tight time constraints (Biederman and Pattison: 2014).

Mixing music is a complicated process and traditionally requires extensive ear training and an in-depth understanding of specialized tools and techniques. Most mix tools relate to physical parameters of sound, making it difficult for the novice user to understand the connection with perceptual parameters (e.g. compression vs. loudness). When musical artists undertake the entire creative process alone, this can also result in a lack of objective feedback. Recordings taken under less than ideal conditions, such as in “bedroom” studios, can contain unwanted artefacts and spectral problem areas, complicating the mix process further (De Man and Reiss: 2017).

Since all successful mixes seem to have certain qualities in common, it is possible to automate parts of the mix process and to develop powerful, perceptually informed, artificially intelligent (AI) mix tools. The term “automatic mixing” was first used by Dugan (1975) in the context of automatic microphone gain handling for speech. Today, automatic mix tools are commercially available, including Izotope’s assistive mix tool Neutron and the online mastering platform Landr. Existing mix tools range from completely autonomous mixing systems to more assistive, workflow-enhancing tools and perceptually enhanced interfaces (De Man and Reiss: 2017).

Many disciplines, including signal processing, music cognition, machine learning and human computer interaction contribute to automatic mixing research (Scott: 2014). Different approaches exist to solving this complex problem. The most common is knowledge engineering (De Man and Reiss: 2013), where informally known rules for creating high quality mixes are implemented in technology. These mixing rules are derived from prosumer
mixing guides and the expertise of mix engineers. The second approach is grounded theory, which was first presented by Glaser and Strauss (1967) in the context of social research. The authors propose that conclusions that are grounded in data can be more reliable than conclusions based on existing theories. Hence, grounded theory is the discovery of theory from data systematically obtained through research. De Man and Reiss (2013) relate the grounded theory approach to the field of automatic mixing. Here, basic knowledge about high quality mixes is acquired first and subsequently transferred to an intelligent system. In this approach, psychoacoustic studies are undertaken to define mix attributes, and perceptual audio evaluation (i.e. listener-based experimentation) is employed to determine listener preference for mix approaches (Bech and Zacharov: 2006). The grounded theory approach can be slow and resource intensive. De Man and Reiss (2013) argue that therefore, it is too limited to constitute a sufficient knowledge base for the implementation of an overall system. Knowledge engineering is a less formalized approach (Scott: 2014) and many commonly accepted rules in mixing do not hold true in formalized studies, for example the notion that most elements should be high pass filtered above their fundamental frequency (Pestana and Reiss: 2014a). Hence, both approaches have advantages and disadvantages.

De Man and Reiss (2017) provide a useful overview over existing studies that seek to automate parts of the mix process. Initial studies contributed to the development of tools for the automatic adjustment of e.g. stereo panning (Gonzales and Reiss: 2010). Further studies have focussed on automating mix parameters such as level (e.g. Wilson and Fazenda: 2016a), reverb (e.g. Benito and Reiss: 2017), panning (e.g. Pestana and Reiss: 2014), EQ (spectral equalization, e.g. Hafezi and Reiss: 2015, pp. 312-323) and compression (e.g. Ma et al., 2014, pp. 412-426).

Some studies do not directly develop prototype automatic mix tools but instead help further the understanding of important perceptual parameters that can feed into the development of such tools (grounded theory). For example, Fenton and Wakefield measure perceived punch and clarity in produced music (Fenton and Wakefield: 2012). Pestana et al. (2013) investigate average spectra of commercially recorded pop songs. Wilson and Fazenda (2016b) investigate the perception of audio quality in productions of popular music. Research in spatial quality perception can also be useful in the field of automatic mixing (e.g. Conetta et al.: 2015, pp. 847–860). The author took a similar approach in her PhD, as summarized in section 2.

Motivation and summary of PhD research findings

The author’s motivation for pursuing a PhD in the field of automatic mixing was to investigate whether the quality of music mixes could be measured
objectively. There can be disagreement as to what constitutes quality in any creative product. Generally agreed quality parameters can help guide this discussion, which can be especially useful in an educational context. The author is also fascinated with the way in which scientific research and creativity can enhance each other. In this case, findings in psychoacoustics and auditory perception can be used to explain preferences in recorded music.

During her PhD, the author and her PhD supervisors worked towards measuring and modelling the perceived quality of music mixes, taking a grounded theory approach. Key findings summarized in the current section are all based on the author’s PhD thesis (Hermes et al., 2017). First, relevant high-level descriptive mix quality criteria were established through a search of scientific and creative literature. These are “clarity and separation”, “balance”, “impact and interest” and “freedom from technical faults”, alongside context-specific parameters. Clarity and separation is the extent to which individual components can be heard in a mix. Balance is an even distribution of energy in the spatial and frequency domains. Three sub-categories of balance are horizontal or stereo balance, depth and tonal balance. Horizontal or stereo balance is the extent to which sound energy is distributed symmetrically and evenly between the left and right channels (within any given frequency range). Depth is a sense of perspective in a mix, where sound sources can be placed at various distances from the listener and inside a fictional, reverberant space of a certain size and shape. Tonal balance is the extent to which sound energy is distributed evenly across the frequency spectrum. Impact and interest is the extent to which the mix grabs the listener’s attention. Freedom from technical faults is the absence of e.g. unwanted recording artefacts or clipping. Lastly, context specific characteristics are the extent to which the mix fits current trends, fashions and norms, complements artistic purpose and supports the musical content. The latter category relates to mix quality parameters that are difficult to generalize, whereas the other categories can be measured automatically. An overview of all parameters is shown in figure 1.

![Diagram showing parameters of high quality music mixes](image.png)

*Fig. 1: The parameters of high quality music mixes.*
Clarity and separation were deemed particularly important for music mix quality and were therefore investigated further. As established in a literature review, clarity and separation in general depend on spectral, spatial and intensity factors and temporal changes in these factors and this is also likely to be the case in music mixes. Spectral factors play a particularly important role across all areas of literature consulted, i.e. timbral clarity, clarity in concert halls, masking, loudness, auditory scene analysis and speech intelligibility. Hence, the impact of spectral equalization (EQ) on spectral clarity was investigated in a series of listener-based experiments. The focus here was on changes in the spectral clarity of single, isolated sounds to keep the complexity low. Single sound spectral clarity is the extent to which the spectral shape of a sound allows all the important components of its natural timbre to be heard. It was established that the clarity of naturally-occurring sounds can be increased when low-Q EQ is applied to boost the less-audible higher frequency regions. If EQ exaggerates or introduces timbrally unpleasant spectral inconsistencies, then these can mask or distract from other sonic components and lead to a clarity reduction (Hermes et al., 2017).

Based on these findings, two predictors of relative changes in single sound spectral clarity were established. These are the harmonic centroid (HC, a weighted centre mass of energy of a sound spectrum, Hermes et al., 2016) and mid-range spectral peakiness (measuring spectral inconsistencies related to sharp peaks roughly in the middle of the frequency spectrum). The HC is a weighted mean of a sound’s spectrum, indicating the harmonic at which the centre mass of energy is situated. It is defined in Equation 1.

\[
HC = \frac{\sum_{k=0}^{K-1} f(k)X(k)}{F \cdot \sum_{k=0}^{K-1} X(k)}
\]

(Equation 1)

\(X(k)\) is the magnitude of frequency bin number \(f(k)\) is the centre frequency (Hz) of \(k\), \(K\) is the number of bins output from a discrete Fourier transform of the sound, and \(F\) is the sound’s median fundamental frequency. \(F\) is defined as the pitch directly in the middle between the highest and lowest note played (Hermes et al. 2017). The HC needs to be raised above around 1.5 harmonics before clarity increases. Mid-range spectral peakiness is calculated by measuring the height of sharp peaks in the middle of stimulus long-term average spectra (LTAS). A computational model was developed that fits a curve to the LTAS, such that potentially unpleasant-sounding peaks lie above it, while the remaining frequency areas lie below it (Hermes et al., 2017). The number of data points above the curve are to estimate relative changes in mid-range spectral peakiness. The computational model is used in
section 3. More information about the model, including a MATLAB download link can be found in Hermes et al. (2017).

As mentioned earlier, the above findings were mainly tested for isolated sounds in order to keep the complexity low. Sounds in mixes were only briefly investigated, in the context of one additional listening test (Hermes et al., 2017). It was concluded that the predictors are still somewhat useful for measuring spectral clarity changes for individual sounds in mixes but the presence of the backing track (rest of the mix) means that complex masking and fusion phenomena need to be considered additionally. Following the analysis of listening test data, it appeared that the more that peaks in the target protrude through the backing track, the clearer the target is perceived (peak audibility). Hence, interestingly, the contribution of spectral peaks to relative changes in clarity appears to depend on the context: when EQ is used to increase peaks on the target sound itself, clarity is reduced. However, if the audibility of peaks is increased by cutting this area in the surrounding backing track, clarity is increased. Therefore, spectral peaks on target sounds appear to contribute to clarity in a complex way and further research needs to be carried out to investigate this. Spectral peaks will be the focus of the next section.

Testing spectral clarity predictors in the mix process

In the previous section, PhD findings on the parameters of high quality mixes and the spectral clarity of sounds were summarized. The aim of the current section is to make informed suggestions for further research by applying the findings to the creative process. As an electronic artist (Nyokee), the author has been writing, producing, performing, mixing and mastering original songs for approximately ten years. Since the entire creative process is undertaken by one person, external, objective feedback on the mix process can be useful. Such feedback could be provided through automatic mix tools. Therefore, as an initial step towards furthering the understanding of spectral clarity in mixes, an autoethnographic study is carried out where the above findings are used to mix a track. Like all research methods, autoethnography has strengths and weaknesses.

Ellis et al. (2011) present some of the criticism that autoethnography has received as a research method. It is occasionally described as insufficiently rigorous, theoretical, and analytical, as conclusions may be based on biased data. However, the authors point out that these criticisms erroneously position art and science at odds with each other. Similarly, Dwyer investigates the question as to whether qualitative researchers should be members of the population they are studying. She comes to the conclusion that the dichotomy of “insider versus outsider” is misplaced here and that we should instead explore the complexity and richness of “the space between entrenched per-
spectives”. As mentioned in section 2, many disciplines are involved in the field of automatic mixing. Therefore, the use of autoethnography may help develop a more holistic understanding of spectral clarity and help guide follow-on research. The remainder of this section is structured as follows. In section 3.1, the predictors are tested in the mix process and findings are presented. In section 3.2, the role of peaks in the natural character of a sound is discussed. In section 3.3, the potential impact of phase issues is presented and in section 3.4 the influence of masking and auditory scene analysis phenomena on clarity is discussed. Lastly, section 3.5. argues that a more holistic understanding of mix quality may be necessary.

**Testing the spectral clarity predictors in the mix process — observations**

The previously established predictors of single sound spectral clarity were applied to the mix process of a vocal in an electronica production (“Serendipity”) in order to assess whether they may be able to help improve lead sound clarity in this type of production. In particular, the contribution of spectral peaks to clarity is investigated further. Feedback was also informally gathered from a group of additional audio professionals with no previous knowledge of the predictors. The finished track, “Serendipity” can be auditioned online (https://soundcloud.com/kirsten-hermes/serendipity). Conclusions in this paper are based on one mix only, and, as mentioned above, the author is both the scientist and subject. While this can be seen as a limitation of the study, the aim here is not to develop a universal model for sound clarity in mixes but rather to explore whether EQ-related clarity changes may be related to harmonic centroid changes and mid-range spectral peakiness for sounds in mixes. For a more holistic understanding of vocal clarity, a larger collection of mixes will need to be investigated. Vocal clarity in the current mix is also more formally tested in a publication under review (Hermes: 2018). Here, ten participants compare versions of the mix in terms of clarity in a custom GUI. The participants are experienced in critical listening and in verbalising sensations of timbre and have no previous knowledge of the predictors.

Having completed an arrangement of synthesizers and electronic sounds, the author recorded her vocal into Apple Logic Pro X in an acoustically untreated home studio, using an SE2200 microphone, an SSL channel strip and an Focusrite Saffire Pro 24 audio interface. The recording lacked clarity and had unpleasant spectral peaks which made it a useful starting point for testing the predictors. Three vocal mixes were created for the first (30s) verse, as follows. First, EQ, compression, deEssing and reverb were applied without explicitly consulting the predictors (version 1). Previous knowledge of the predictors still influenced the EQ process: low frequencies were cut and high frequencies were boosted in order to raise the HC and it was attempted
to reduce some of the unwanted spectral peaks. Since the increased audibility of spectral peaks of target sounds in mixes appears to contribute positively to clarity (section 2), it was attempted to remove energy near peaks in the vocal from surrounding instruments with EQ. However, since there were many small adjacent peaks, this was not feasible in practice.

A second version of the mix was created, guided by the clarity predictors (version 2). The HC of this version was increased by another 6 harmonics. Using the computational model of mid-range spectral peakiness introduced in section 2 (Hermes et al., 2017), it was attempted to eliminate peakiness entirely. This process required a large number of additional fine EQ adjustments. The resulting version was spectrally much flatter and free from mid-range spectral peakiness. The LTAS and mid-range spectral peakiness for versions 1 and 2 are shown in figures 2 and 3. A third version was also created where all EQ was removed entirely for comparison (default version). In this version, the HC was lower than in version 1 by 6 harmonics. Mid-range spectral peakiness was greater than in version 1 but lower than in version 2. For comparison, the LTAS and peakiness for this version is shown in fig. 4.

Since the author’s previous knowledge of the predictors had
influenced the EQ of the first version, the difference between versions 1 and 2 was small. The author felt that version 2 sounded somewhat smoother and clearer than version 1. Both versions sounded much clearer than the default version, which appears to indicate that the HC is a useful clarity predictor in this case: the HC was considerably lower in the default version than in versions 1 and 2. In the following, each version is described in greater detail.

When played in isolation, the default version sounded fairly natural but had excessive energy in the 300Hz area. This gave it a ‘dull’, ‘muffled’ and ‘muddy’ timbre. There also seemed to be some unwanted noise and distortion as a result of the recording process, leading to a ‘fuzzy’ quality. In the mix, the default version sounded particularly unclear and not separate enough from other sounds.

Version 1 sounded much clearer than the default version, both in isolation and in the mix. The increased high frequency to low frequency balance made it appear more ‘present’ and ‘thin’. Version 1 did sound somewhat less natural than the default version when played in isolation but clarity and separation was much improved in the mix. However, by boosting high frequencies and cutting low frequencies, an unpleasant peak in the high frequency area was made more obvious. The author struggled to identify the exact location of this peak. This lead to a degree of ‘harshness’ and the overall timbre did not appear to be tonally balanced. Additional boosts and cuts seemed to either increase this harshness or make the timbre ‘duller’ like in the default version.

The use of the clarity predictors in version 2 allowed the author to locate the aforementioned peak in the 3—5kHz area. It was possible to flatten the peak without affecting other areas, as the model had correctly identified it as contributing to unpleasant mid-range spectral peakiness. Hence, the peak can be seen in the LTAS for version 1 (figure 3) but not in the LTAS for version 2 (figure 4), where it had been removed. Additional, smaller peaks could also be treated. In the author’s opinion, the resulting sound was smoother, clearer, and more present than in the other two versions with reduced sharpness. The noise and distortion in the default version had been altered to a pleasant ‘sizzle’. Despite this improvement, version 2 sounded less natural and more processed than version 1. In the mix, the difference between versions 1 and 2 was small.

Informal discussions with other audio professionals (students, lecturing staff and sound technicians at the University of Westminster with no previous knowledge of the predictors) revealed some disagreement. While the general consensus was that both equalized versions were much clearer than the unequalized version, some listeners felt that version 1 was clearer than version 2, others the other way around. While some listeners did also describe version 2 as smoother, they also commented on the fact that it sounded overprocessed and unnatural. It seems that some listeners perceive naturalness as an important aspect of clarity, while others do not (including the
author). It appears that the HC is a suitable clarity predictor, since the HC in versions 1 and 2 was notably higher than in the default version. Peakiness still seems relevant, albeit to a lesser degree, since the version containing no peaks (version 2) was not perceived as clearer by all audio professionals consulted. It is concluded that the HC would be useful in an overall clarity model. The contribution of spectral peaks to naturalness, and therefore spectral clarity, however, needs to be investigated further. In the following sections, starting points for this investigation are presented.

When is a spectral peak part of natural character of sound?

It is possible that some spectral peaks are perceived to be part of the natural character of a sound while others are considered unpleasant. If naturalness is important for clarity, it is possible that an entirely flat spectrum resulting from fine EQ adjustments can reduce clarity much in the same way as a peaky spectrum can in other cases. Therefore, it is possible that some spectral peaks appear to increase clarity while others reduce it. However, what constitutes a natural timbre may be difficult to measure, as discussed in the following part of the paper.

The impact of spectral peaks on clarity may also, to some degree, depend on the instrument. While some instruments tend have fewer peaks (e.g. cello), others feature natural, strong resonances, such as the Erhu (Chinese violin) (Hermes et al., 2017). It is possible that recordings of acoustic instruments and voices sound clear when their spectra resemble their natural, unrecorded spectra. However, spectra can only be measured by recording sound and the recording process always introduces spectral distortions.

Spectral clarity may be even more difficult to establish for newly created timbres such as electronic synths. Zagorski-Thomas (2007) relates musical elements to “physical manifestations of emotions, gesture and being in space”, describing music recordings as sonic metaphors for physiologically and culturally determined gestures and morphologies. Further, Zagorski-Thomas (2017) states that multi-track and electronic music can be interpreted as something impossible and yet understandable. Similarly, Théberge (1997) links adjectives used to describe low-level mix parameters to bodily sensations. It is possible that listeners might still agree on the clarity of newly synthesized sounds, since there might be a shared understanding of what constitutes a ‘believable’ timbre. In the author’s experience, combinations of contrasting timbres may result in greater separation in electronic music mixes. The author tends to prefer ‘edgy’, ‘hard’ timbres for her own productions, some of which have strong peaks (e.g. chiptune and 8-bit timbres). For that reason, it would be useful to establish whether there are spectral areas that always contribute to clarity in the same way for these sounds.
Phase issues

The LTAS of sounds can be useful for measuring timbral attributes and is used in a multitude of research studies (a literature review on clarity can be found in Hermes et al., 2017). However, this measurement method ignores the relative phases of spectral components and short-term spectral fluctuations. Bregman (2007) states that sounds with the same frequency content but differing phases can sound different, which can also influence separation in sound mixtures. Laitinen et al. (2013) confirm that humans can perceive differences in the phase spectrum of otherwise identical sounds and that the phase spectrum affects the perceived timbre, especially in sounds with lower fundamental frequencies. Toulson (2008) argues that therefore, it can be difficult to fix spectral problem areas in the mix. He points out that an over-use or incorrect implementation of EQ can be detrimental to the sound quality due to the resulting phase distortion. Several listeners confirmed that the second version of the “Serendipity” vocal sounded over-processed, which may have been due to distortions in the phase spectrum resulting from the many fine EQ adjustments.

The recorded vocal take was 30 seconds long, hence the predictors would have ignored any large spectral fluctuations throughout. It could be argued that in order to keep the complexity low, it is useful to fully understand in which way features of steady-state spectra contribute to clarity before considering temporal factors and phase spectra. However, a fully functional clarity model would most likely be more accurate if these parameters were also considered.

Sounds in isolation and in the mix

As indicated in section 2, the interaction of the target sound (in this case the vocal) with the backing track (rest of the mix) also needs to be considered, due to masking and fusion phenomena. Parts of the target sound spectrum may become masked by the backing track. According to the American National Standards Institute (ANSI/ASA: 2013), masking is the process by which the threshold of audibility for one sound is raised by the presence of another, masking sound. Instruments occupying similar frequency regions in music mixes are likely to mask each other. Partial masking can reduce the loudness of the target in the mix (Ma et al.: 2014) and is therefore likely to lead to a reduction in clarity. Pestana and Reiss (2014) point out that in music mixes, EQ should be applied to ensure that no element masks any of the frequency content of lead sounds. Overall, it is apparent that masking can reduce clarity but it is not clear whether there may be any frequency areas that particularly should be unmasked and how this relates to the audibility of spectral peaks in the target sound. As mentioned in section 2, it is possible that natural peaks in the target sound should be unmasked. Phenomena such
as upwards masking, where lower frequencies in the masker can mask higher frequencies in the target and temporal phenomena like forward and backward masking add to the complexity (Moore: 2012).

The audibility of target sounds can also be compromised when not enough separation exists from the backing track, even when the target sound is unmasked. Fusion and separation phenomena are assessed in the field of auditory scene analysis. Auditory scene analysis (ASA) is the process of forming mental representations of individual sound sources from the summed waveforms that reach the ears. The ASA process consists of the following two conceptual stages (Bregman: 2007). First, the auditory system divides the input into its constituent atomic units, i.e. packages of acoustic evidence (segmentation). Following segmentation, any packages that appear to have arisen from the same source are either grouped (to form a stream for a given source) or segregated (to form separate streams for different sources). Elements that fall in the same auditory stream are perceived as stemming from the same sound source. Grouping and segregation are related to the perception of separation in music mixes. The factors that determine whether sounds are fused or separated are complex but spectral similarity between target and backing track is particularly important.

Woszczyk and Bregman (2005, pp. 13–25) state that the ear is more easily able to follow a sound in a mix if it has a unique timbre, leading to greater separation from other sounds. They state that unique timbres usually have obvious features that the listener can track over time. Strong spectral peaks may constitute an obvious feature, increasing clarity in mixes. Bregman (1990) provides the example of classically trained singers who can enlarge their pharynx cavity and produce a strong resonance in the mid-frequency area, an area that is usually not occupied by a lot of other instruments in the orchestra. This resonance allows e.g. opera singers to be heard over their accompaniment, even without the availability of amplification.

A potentially similar technique appears to be prevalent in pop singing, i.e. singing with ‘twang’. ‘Twang’ is a vocal timbre produced through increased subglottal pressure, leading to increased energy in the first two formants, decreased energy in formants 3 and 5 and an overall higher sound pressure level (Sundberg and Thalén: 2010). It is possible that ‘twang’ leads to a spectral peak around the first two formants that can increase audibility and separation of vocals in mixes. If this peak were to be removed with EQ, clarity may decrease as a result. Therefore, spectral flatness in lead sounds such as vocals may decrease spectral clarity in mixes. It would be interesting to establish whether there are specific, generalizable frequency areas that contribute to the clarity of lead sounds. Izhaki (2008, p.251) suggests that the spectral area for vocal clarity area lies around 2kHz – 9kHz. An increase in energy in this area is likely to correlate with an increase in the HC.
Holistic perception of clarity in music mixes

It is possible that target clarity in mixes is not an isolated concept but ties in with the holistic perception of the overall mix. The clarity of each sound in the mix may bias the perception of the clarity of each other sound. As established in section 2, overall tonal balance is another important parameter of music mixes. Pestana et al. (2013) explain that spectra of professionally produced commercial recordings show consistent trends, which can roughly be described as a linearly decaying distribution of around 5 dB per octave between 100Hz and 4000Hz, becoming gradually steeper with higher frequencies, and a severe low-cut around 60Hz. It is possible that overall tonal balance influences the perception of clarity of each sound therein.

Similarly, so far, only spectral parameters of clarity have been investigated. For a complete model of clarity in music mixes, spatial and intensity related factors should also be considered, alongside temporal changes in all these factors.

Discussion and suggestions for further research

As established in the last section, spectral clarity is likely to be a multifaceted, complex concept. The two predictors of single sound spectral clarity still appear to be important, in particular the HC. Mid-range spectral peakiness can help measure the influence of resonances on the clarity of isolated sounds. In music mixes, the influence of peaks on clarity appears to be more complex. For a better overall spectral clarity model, further research needs to be carried out. First, it would need to be established which spectral peaks constitute a part of the natural character of a sound. This is especially difficult for newly synthesized timbres, where no natural reference exists. Naturalness is likely to influence clarity; however, it should be established how strongly these two attributes correlate. Second, it would be useful to assess how spectral parameters that are not measured by the LTAS contribute to spectral clarity. Phase spectra appear to be particularly important in this context, since phase distortions, as introduced by EQ, can be detrimental to sound quality. Third, the impact of the complex relationship between the target and backing track spectra on clarity needs to be understood further. Not just masking phenomena, but also fusion and separation between target and backing track need to be taken into consideration. It would therefore be useful to reassess spectral target clarity in mixes using computational auditory scene analysis (CASA) models. The clarity of sounds appears to depend on fusion phenomena in a complex way. Fourth, it would be useful to assess the impact of non-spectral parameters on clarity, such as spatial, intensity related and temporal factors.
Lastly, in order to measure mix quality successfully, it would be necessary to measure all high-level parameters of mixes, that is, not only clarity and separation, but also balance, impact and interest and the freedom from technical faults. Some context-specific parameters could be measured through comparison to a reference, e.g. mixes of a similar fashion or style. Since the perception of each of these mix parameters may not be isolated, their influence on each other should be assessed.

When spectral clarity is more fully understood, the implications of the findings on other research areas should also be explored. For instance, if specific resonances do increase clarity in singing, it would be useful if singers could integrate this knowledge into their training. Similarly, sound synthesis tools could be based on more perceptually informed parameters, such as a spectral clarity control. Since all research methods presented in this paper (grounded theory, knowledge engineering and autoethnography) have limitations, ultimately, an interdisciplinary approach may yield the most useful results. Therefore, in the author’s opinion, scientists and creatives should continue to collaborate in furthering the understanding of music mix quality, since this is likely to lead to a rich, holistic understanding of the subject.

The aim of the current study was to test whether the HC and mid-range spectral peakiness may be able predict clarity for sounds in mixes and to develop starting points for follow-on research. To conclude, the HC appears to be a strong predictor of spectral clarity in music mixes. Mid-range spectral peakiness also seems useful but should be supported by a metric for naturalness. Additionally, metrics of phase and masking, as well as a CASA model should be included in an overall spectral clarity model. Follow-on research should not only consider spectral clarity but also consider temporal, spatial and intensity related factors. An interdisciplinary approach is likely to produce useful results.

References


De Man, B. and Reiss, J. D. (2013) 'A Semantic Approach To Autonomous Mixing.' In: 2013 Art of Record Production Conference, Québec, Canada, July 12th–14th.


Marco Antonio Juan de Dios Cuartas: The incursion of stereo into Spanish popular music: the English influence in the definition of a local sound through the professional exchanges between Madrid, Barcelona and London

Abstract

In Spain, the role of the engineer and the music producer emerges at the beginning of the 60s from the “imitation” of the Anglo-Saxon model. The English professional experience of producers and engineers such as Alain Milhaud and Pepe Loeches, helps to import some of the technical methodologies used in English recording studios into Spain. Production equipment from Abbey Road is sent to the EMI studio in Barcelona, allowing stereo recordings to be made with mixing consoles previously used in The Beatles’ productions. The introduction of stereo to Spanish productions does not seem to have come much later than in other markets like the UK or the US.

The incursion of stereo into Spanish record productions of the 60s

The evolution of recording studios must necessarily be linked to certain processes of globalization. Ulrich Beck defines globalization as a “process that creates links and transnational social spaces, revalues local cultures and brings third cultures to the foreground” (Beck: 1998, p. 30). We can thus speak of an emerging global culture and the global cities in which it develops (Featherstone: 1990, King: 1991, Hannerz: 1991). Perhaps one of the main consequences of this “sonic globalization” is, without a doubt, the loss of hegemony in the record production processes of the major capitals of the music industry, from the UK-US centre of reference to the periphery represented by other countries which, as in the case of Spain, will end up assuming the production methodologies of their centres of reference, creating an infrastructure of recording studios adapted to the new sounds. But the production of Spanish popular music, at the beginning of the sixties, witnesses a
professional migratory phenomenon towards England, seeking in its place of origin the authenticity of the sonority of the British popular music productions. This fact was determined mainly by the lack of confidence of musical producers towards the human and technical resources of the Spanish recording studios of the time: the Spanish recording industry in these early years did not yet have an adequate infrastructure to face the new sounds of pop and rock. From a historical perspective, from the 1950s standards of sound deriving from the creation and implementation of certain pieces of equipment in recording studios -including certain mixing techniques related to the use of stereo panning- meant that a particular English or American sound could be “distinguished”. Both of these countries decisively influenced the musical productions made in other Western countries that, as in the case of Spain, did not have an industrial fabric capable of accommodating the needs of music producers. In regard to music production, this would force Spain to become a country that imported audio technology, absorbing the “sonic imperialism” exercised mainly by the English due to geographical proximity. Only those societies that could afford the necessary technical requirements were able to adopt the aesthetic canons proposed in English and American productions:

[...] the more music aesthetics are defined by sound, and hence by technology, the more those who own the means of production and the expertise to operate them will be also able to control the music market and the musical discourse on a global level. Music production technology and the discourse on technology act as gatekeepers: those who lack full access to technology will not be able to participate in the definition of musical aesthetics. (Nardi: 2015)

The socio-economic situation of a country must therefore be considered an influential element in the sound aesthetics of record productions, and this should be taken into account when analysing and historically contextualising these productions. Popular music in Spain often imitated English or American record productions, which also entailed using the same recording devices. In this respect, the audio engineer Pepe Loeches (1979) raised what has always been one of the main difficulties when acquiring and upgrading the devices that are involved in the recording process: being a country that was fundamentally an “importer”, the increase in cost and the difficulty in accessing certain pieces of equipment were added problems when competing with the international recording industry:

One of the biggest problems we have to contend with is the acquisition of recording equipment — as there are no local manufacturers it all has to be imported. Consequently, arranging orders and shipping of equipment can cause considerable problems and high import taxes can often double, and some-
times triple the cost of the equipment. Nevertheless, the money, time and patience are usually found, to thoroughly equip studios.

The importation of English equipment such as Neve, which had a great impact on Spanish recording studios during the 1970s, together with the application of acoustic designs under international standards (Westlake-Eastlake acoustic design), progressively reduced the differences between Spanish recording studios and those of their English or American counterparts. To this, it also contributed the establishment of multinational record companies in Madrid or Barcelona that built their own recording studios mirroring the designs of others located in London or the USA –as it is the case of EMI in Barcelona or RCA in Madrid. To cite a relevant example, Neve mixing consoles, whose importance to music production during the 1970s is unquestionable, appeared in Spanish studios at the same time as in recording studios in the Anglo-Saxon world. To this effect, the American magazine dB published the following report in February 1969:

The largest music-recording control console yet built by Rupert Neve and Company Limited, has been installed in the New York studios of the Vanguard Recording Society, Inc. The Cambridge, England-based firm’s console embodies twenty-four input channels, sixteen output groups, four echo groups, and two foldback groups. The console also has comprehensive four-speaker monitoring and re-mix for sixteen tracks. This is their second installation for Vanguard, the first being a sixteen-channel mastering console. Neve has recently completed two consoles installed in Spain and is currently working on units for re-recording desks for Pye Records, Ltd., and Associated British Pathe Ltd., both of England.

The recording industry focused on the two main cities of the country for one simple reason: it was in these cities that the offices of the major record companies were located, and especially, the headquarters of the international companies. During the 1970s, small studios were proliferating in other cities, but always with an inferiority of resources that, in general, turned them into “second-rate” studios.

EMI had its own recording studios in Madrid, but the studios located in Barcelona were especially important. Although the recording studios housed at the headquarters of international record companies usually boasted competitive technology, they were fitted out with equipment that had been updated in their studios of origin. Ramón Arcusa and Manuel de la Calva tell how material from Abbey Road found its way to the EMI studios in Barcelona, allowing them to make their first recordings in stereo:

---

1 Interview of Pepe Loeches in Studio Sound magazine in a special issue published in the autumn of 1979 titled International Recording.

In Barcelona we were lucky enough to record with a good company that was EMI and also to have a great sound engineer, and great musicians. Perhaps in that sense Barcelona was a bit more advanced than Madrid, but Madrid caught up very quickly. Initially we recorded the early songs, until 1964, in mono. Then, suddenly, The Beatles had an eight-track [machine] instead of a four-track in London, and the same company sent the four-track that they left behind to Barcelona and there we recorded several songs. We recorded the first stereo in 1964 on a console with which The Beatles had recorded in their time. Love Me Do had been done there.3

The incursion of stereo into Spanish commercial recordings—it must be remembered that research into stereo began in the 1930s—does not seem to have come much later than in other markets like England or the United States. In an article about the Hispavox studios in Spain, the March 1961 issues of the journal Audio Engineering noted the experimentation with multiple microphones and the evolution of the early techniques of stereo recording:

How would you like to take on the assignment of recording one hundred guitars in stereo? Someone on the staff of Hispavox Records in Spain was recently handed just such a job. The group includes sections that play the laud (a relative of the lute) and a contingent of bandurrias. These three main sections are further subdivided since all instruments are employed in both large and small sizes. All told, eight types of guitars are represented in this orchestra which consists entirely of blind musicians. During the recording session, the instructions of the conductor were relayed to each individual player by means of earphones. Half the selections were composed by the conductor, Rafael Albert, indicating that a somewhat specialized guitar repertory is required by a group of this size. Serenades and dances figure heavily in the choice of tunes. The sound, when heard on a full-fledged stereo system, is difficult to describe in terms of everyday comparison because I’ve never experienced anything like it. Multiple miking must have been used in a ball of exceptional liveness in order to capture so effectively the full impact of the total ensemble without losing the fine detail in the sound of each section.

Both the recordings made during the early 1960s in the Hispavox studios in Madrid and those produced in Barcelona, a consequence of the technological trade-off with Abbey Road studios in London, demonstrate that stereo recordings were already a reality during the first half of this decade. Although the techniques of stereo recording began to be generally used during this period, the most experimental uses of the stereo panorama were mainly developed in popular music recordings than in other types of productions

---

3 Interview broadcast in 2014 in the programme Música Ligerísim a of TVE Spanish television network.
related to traditional music (eg Flamenco⁴) or classical music (eg Symphonic Music or Zarzuela).

The professional exchanges that took place during this period between Madrid, Barcelona and London seem to have been constant, such as the above-mentioned case of Pepe Loeches, of compulsory study in the history of record production in Spain. After working as an apprentice at the Hispavox studios, Pepe Loeches moved to England in 1969 to work at the Pye Records studios in London, where he remained until 1975. Loeches’ return to Spain somehow implies the importation of some of the Anglo-Saxon work methodologies into Spanish production, exerting an important influence over other professionals and recording studios in which he subsequently worked: Kirios, Eurosonic and Musigrama.

Despite some of the professional experiences of Spanish engineers were able to transmit greater confidence to producers and record companies, in the Spanish recording industry it was common practice, at least in an early stage, to carry out the music productions with the biggest budget in recording studios in London. In order to obtain an international product, with the characteristic sound of the English productions, the safest thing to do was to travel to a London studio and work with its engineers and session musicians.

A case study: the record productions of Alan Milhaud

Although Spain has a remarkable industry linked to record production prior to the 60’s, there is no "physical infrastructure" capable of "emulating" the "English sound" related to the incipient rock, nor have the rules inherent in a "practice of recording", methodologically different from that carried out in Spanish recording studios up to that time, been assimilated. For René de Coupaud, Spanish musician and producer who began his professional career in the 70s, the process of professional adaptation to new production methods related to the new sounds of popular music came slowly:

In the studios, there was no experience of recording that kind of music, neither the musicians nor the arrangers were ready. The most successful at the beginning was Milhaud, because he did know where he could turn to. In London, with some specialised session musicians, it was all easier, so even in the mid-70s -although in Spain we already had great professionals- some producers were still choosing to record in English studios. Recording abroad

⁴ In Flamenco, the use of reverb acquires a special relevance in order to achieve a greater sense of space and amplitude in the stereo mix. In this case, the reverb encourages the ritual character of a music that seeks to immerse the listener in the space in which the performance is developed.
was logically more expensive, but as it took less time to do it and the results were also guaranteed, it was chosen to do so.\(^5\)

Alain Milhaud, a musician of French origin—who initially began his career in Barcelona at the beginning of the 1960s, and subsequently moved to Madrid—was the first music producer in Spain. Milhaud’s professional development will be conditioned by his relationship with the record industry of London, Madrid and Barcelona, maintaining a continuous exchange through his projects, whose influence on the subsequent Spanish discographic production is indisputable. Milhaud interrupted his Economics and Politics studies to fully concentrate on conducting. Curiously, this sparked his interest in the world of recording as he began “to note the enormous difference between the sound mass that he received from the conductor’s podium and that which came from the speakers in the sound booth” (Domínguez: 2002, p. 268).

What Milhaud initially pursued in his productions is in line with Ed Ward’s description of the Rolling Stones’ record productions (1969) in an article cited by Frith (2012) analysing the role of the producer in the discourse of rock: “They spend a lot of time remixing and overdubbing, but the end result is always credible—one can imagine little Stones performing in the speakers” (in Frith: 2012, p. 211). But the reality of the recording studios in Barcelona at the beginning of the 1960s was dramatically different: the enormous difference between what was recorded and what was definitively heard through the speakers minimized the credibility that Ward mentioned, discouraging Milhaud, who intended to use his practical experience in the portable studio of the Belter\(^6\) record label to acquire the skills necessary to allow him to practise the profession of producer as he conceived it. During these three years at Belter, Milhaud had the opportunity to produce bands including Los Gatos Negros, Latin Combo and Los Tres Sudamericanos, as well as many flamenco recordings such as Antología Flamenca by Dolores Abril and Juanito Valderrama.

After his stage at Belter, Alain Milhaud arrived at EMI-Odeón as manager of the international catalogue—receiving samples of the records released by EMI International and selecting those he considered could be launched in Spain—\(^7\) and as national A&R. In Milhaud’s own opinion, during this period there was only one serious studio in Barcelona, which was La Voz de su Amo (HMV: His Master’s Voice) and which formed part of the business fabric of the British firm EMI: “They had a studio set up by the English, the

\(^5\) Taken from the interview made to him on 11 August, 2014 by the author of this article.

\(^6\) A recording company created in Barcelona in 1954 that remained in operation until 1984.

\(^7\) The case of The Beatles deserves an analysis of its own. In Spain, the distribution of their music was limited on the express order of the company’s sales manager, who prioritised the launch of the Spanish version of the band’s hits by the group Los Mustang over the original songs.
material was fairly basic, but the best that there was at the time”\(^8\). But Milhaud’s time at EMI was very short—only three months—after which he moved to Madrid to lay the foundations in the history of Spanish record production. Milhaud’s concept of recording studios in Spain during this period was very negative:

> Recording studios in London and studios in Spain were completely different, both in regard to technical equipment and the engineers. I’ll give you an example: the first time I attended a recording in London, half an hour before the recording, the engineer was already there in the studio adjusting the “azimuth” of the heads of the tape recorders.\(^9\)

The gap Milhaud refers to, regarding the training of the Spanish recording-studio engineers with respect to those of other countries like England, justifies the fact that he chose to make his recordings outside Spain during this period, despite belonging to a company (Spanish Columbia) that had its own recording studios in Madrid: “As an artistic director and producer, the most logical decision in pursuit of that professional quality was to leave”.

Alain Milhaud’s professional experience in England helped some of the standard work methodologies in English recording studios to be imported into Spain. This is the case with the song “Black is Black” by the Spanish band Los Bravos. For this 1966 production, even though he had the company’s own studios in Spain at his disposal, the song’s producer Alan Milhaud decided to go to Decca Studios in London.

“Black is Black” was recorded at Decca Studios in London by session musicians, the only exception being the recording of the lead vocal by Mike V. Kogel and the backing vocals by Toni Martínez. Although it could be considered far from the “authenticity” of a rock production, this work methodology was quite common at the time and was part of the decisions the music producer had to make to obtain an optimal production, both at a musical and technical level. One might think that using session musicians and foregoing part or all of the band’s musicians, goes against “authenticity”, which was perceived as a paradigm of the rock genre, a paradigm that was under construction during the 1960s. Some producers consider other practices, such as overdubbing or editing with the intention of correction, as actions that make rock music less “authentic”.

The recording of “Black is Black” at Decca Studios was carried out by a young Bill Price, an English audio engineer (at this time a “balance engineer”, a term used to describe the engineers in charge of “balancing” the

---

\(^8\) All statements by Alain Milhaud are taken from the interview made to him in 2014 by the author of this article.

\(^9\) In a magnetic recording system, the “azimuth” defines the direction of the head of the recording and/or reproduction with respect to the tape and forms part of the necessary adjustments for a correct maintenance of the sound recorder, enabling optimum recording and reproduction.
level of the different mixed instruments through the console faders) and sub-
sequently a producer as well, who was at the beginning of his professional
career at the time. Price would go on to work with artists such as Tom Jones,
Eric Clapton, The Moody Blues, Mott the Hoople, Pink Floyd, Paul McCa-
Milhaud described the recording session in the following terms:

It was recorded on four live tracks. It was recorded with everyone in the same
room, and with the singer in a booth. Mike sang live but later he effectively
sang it again, doubling his voice himself and Toni doubling. The whole
rhythm section was on one track, the brass on another track, the voices on the
third and the rhythm section was divided; the bass was on a track as well.
Panels were used inside the room to try to separate the instruments, and
headphones for the musicians as well. Mike’s voice was in a separate booth,
there was no leakage, the voice was clean and the channel was clean.

The session musicians would have used the first recording of the solo
voice as a guide for playing the song, performing a second time for the inde-
pendent and isolated recording of the voice in relation to the rest of the in-
struments. The presence of the reverb from the echo chambers in Mike V.
Kogel’s voice is clearly noticeable in the mix, adding spatiality to a lead
voice that is situated in the centre of the stereo field. The focus of the sound
source in the foreground is diffused in the “soundbox” – applying the meth-
odology proposed by Alan F. Moore (2012) – by the reverb action, blending
the main voice with the rest of the elements in the mix (Fig.1).

The musical arrangement is not based on a staggered polyphonic begin-
ning in which the different melodic lines that are introduced at different
pitches are interwoven, but in this case the sound block is created through
the introduction of the different elements in unison. The real contribu-
tion of the progressive increase in the instrumenta-
tion derives, on the one hand, from enrichment at
the timbral level with the progressive incorporation
of elements into the record-
ing, which reaches a
climax with the harmon-
ics that are generated as a
result of the reflections
caused by the lead voice

Figure 1: Location of the lead voice inside the
“soundbox” and projection of the reverb in the lead
voice of the song “Black is Black” (1966) (Drawn up
by the author).
and, on the other hand, from the progressive increase in the “stereo width”, playing with the phased incorporation of instruments to the left and right of the panorama.

These characteristics can be clearly seen in a comparison of the harmonic content generated by the spectrogram in each part of the introduction prior to the entry of the lead voice. The evolution of the harmonic content must be connected to a conception of the stereo mix in accordance with what Moore (2012) terms a “triangular” mix, placing the voice in the centre and grouping the rest of the elements that make up the different layers of the mix to the left or right. This is the main reason why there is a significant difference between the amplitude of the left (L) and right (R) channels of the phonogram during the introduction (Fig. 2 & 3).

If the levels of both channels during the introduction are observed on a RMS meter, it can be seen how there is a difference of about 10 dB calculated on the RMS level of the master. This initial decompensation of the channels is balanced by the progressive incorporation of the instrumentation during the introduction. As Alan Moore (2012) points out, the extreme panning of the elements in the mix, including the voice (“cluster mix”), is part of the configurations used before the consolidation of the “diagonal mix” and in which the solo voice, snare drum and bass are located in the centre on a slight diagonal (Fig. 4).
The presence of low frequencies in both channels of the stereo, as pointed out by Moore (2012), does not become a standard until the 1970s. The distribution of the elements in stereo must follow certain rules, for example, the bass drum and bass go in the middle because the surface of two loudspeaker membranes, which have to simultaneously transmit a low sound, is better than an individual loudspeaker with half the surface. The fact that the bass is situated on the left causes a poor distribution of frequency energy. A comparison of the spectrogram generated by the left channel –to which the “drum set” and the bass are panned – and the right channel –with a minimal presence of what Owsinski (2013) calls base or “foundation”– provides a clear idea of the harmonic content generated by this layer of the mix during the first four bars of the theme (Fig. 5 & 6).

Figure 5: Spectrogram of the left channel (L) of the first four bars of “Black is Black” (1966).

Figure 6: Spectrogram of the right channel (R) of the first four bars of “Black is Black” (1966).

The energy of the attack of the “drum set” snare generates harmonics in the left channel (L) that virtually reach 20kHz and musically represent the rhythmic motor of the theme, while the frequency presence of these elements in the right channel (R) is limited to reproducing the reflections that the base or “foundation” generates in a large room like the one in which the song was recorded. The predominance of the left channel over the right in these first bars can be clearly seen on a Goniometer or Lissajous stereo field analyser (Fig. 7).

Figure 7: Stereo image from the introduction to “Black is Black” (1966).
Although mono recording and mixing evidently required some rules of positioning and angulation of the microphones, the development of binaural recording techniques will require that the distances between the microphones placed before the sound source are adequate for the sum to occur correctly and without cancellations. Regarding the recording process of Black is Black, Milhaud pointed out the following:

An ambient stereo mic [overheads] was placed. The kick had a special dynamic mic, then the snare, the toms, and so on. The toms were individually miked. All that was put into a group -the mixing console had the possibility of working with individual channels and groups- and afterwards, the balance of the drums was made for the group. Then the lead voice was added and it was balanced with the rhythm section (drums and bass). Finally, the rest of the harmonic instruments were also added.

It is more than likely that Milhaud refers to "stereo pair" (two "mono microphones" that represent the left and right channels and that are positioned following a certain microphone technique based on their location, distance and angle with respect to the sound source) instead of "stereo microphone" (a single device with the ability to pick up two left and right signals simultaneously). The Decca Company maintained a constant concern for the application of new microphonic techniques aimed at capturing and expanding the stereo width within the mix. In 1954, Arthur Haddy, Roy Wallace and Kenneth Wilkinson developed the "Decca Tree" technique, a stereo microphone pickup system for large orchestras. The "Decca Tree" consists of a set of three microphones with an omnidirectional polar pattern -sensitive to pickup in all directions- spaced according to the desired width and stereo amplitude. The search for mono compatibility in stereo mixes, interpreting this as the result of the sum of both channels (Left and Right), has been part of the work methodology of the mixing engineer from the consolidation of the stereo mix to the present. All mixing consoles have a "switch" type commutator that allows L+R to be heard. This enables technicians to monitor the possible phase cancellations that occur when performing this action and to anticipate the possible results of listening to the phonogram on a mono device—a turntable with a single loudspeaker, for example— or in a "mono" radio broadcast, which was common during the second half of the 1960s. The use of a correlation meter shows how during these first four bars of the song, the meter registers negative values, therefore implying the disappearance of the mono compatibility of the stereo mix (Fig. 8).
The October 19, 1968 issue of *Billboard* magazine published an interview with Milhaud: “Alain Milhaud was one of the first record producers in Europe to recognize that to compete with British and American product it was essential to produce an Anglo-American sound and to get his artists to record in English”. Milhaud’s objective was to obtain a product with an “international” sound, which implied a significant qualitative leap in comparison with his first recordings in Barcelona, only four years earlier. His ultimate goal was to be successful on the English and American lists, competing with the local bands: his strategy was based on proposing recording projects, bridging the gap as much as possible between his language as a producer of Spanish bands and the Anglo-Saxon bands whose sonic language was his “mother tongue”.

A comparative analysis of the recording of the song “Speedy Gonzales” by Los Gatos Negros (1962) at a macro-timbral level –recorded at the Casino de la Alizanza in Barcelona with the portable studio of the record company Belter– and “Black is Black” by Los Bravos (1966) –recorded at Decca Studios in London– presents a very significant picture. A comparison between the spectrogram that each phonogram generates shows how the spectral content of the song by Los Gatos Negros, recorded with very limited resources, is mainly represented by midrange frequencies: the harmonic content is restricted up to 6 kHz. There is a complete absence of high frequencies and, as a consequence of the limited number of microphones and channels, no key elements can be detected in the lower area of the frequency spectrum of a pop-rock song such as, for example, the bass drum (Fig. 9 & 10).

![Figure 9: Harmonic content of the song “Speedy Gonzales” by Los Gatos Negros (1962).](image)

![Figure 10: Harmonic content of the song “Black is Black by Los Bravos (1966).](image)

In terms of the horizontal dimension (panorama), a comparison between the treatment of the stereo field in the mix of the songs by Los Gatos Negros and Los Bravos reveals a considerable widening of the stereo width in “Black is Black” with respect to “Speedy Gonzales”. The latter has a much narrower stereo field, an almost mono mix with the single opening on the
horizontal axis that creates the reverb reflections applied to the voice (Fig. 11 & 12).

![Figure 11: Stereo image of the song “Black is Black” by Los Bravos (1966).](image1)

![Figure 12: Stereo image of the song “Speedy Gonzales” by Los Gatos Negros (1962).](image2)

The representation of the stereo image on the Lissajous Meter corresponds to a moment in the song when all the elements of the production have already been introduced. There is also a greater presence of the frequency axis within the three-dimensional conception of the mix in comparison with the poor stereo image of the song “Speedy Gonzales”.

An analysis of the correlation of the song “Speedy Gonzales” yields a phasometer with measurements close to +1, that is, the information in the left and right channel is very similar: the monaural compatibility is complete because the mix has a poor stereo image.

**Conclusion**

The incursion of stereo into Spanish music production does not seem to have taken place much later than in the recordings made in the Anglo-Saxon world. During the first half of the 1960s, the Spanish recording industry presented a mixed picture in regard to the quality of its technical and human resources, with considerable differences with respect to the results obtained from one recording studio to another. While some studios, such as EMI Internacional in Barcelona—with the transference of technology from London—or the Hispavox company in Madrid, offered some guarantees of quality that led their artists to record there, some producers like Alain Milhaud decided to move to studios outside Spain in order to achieve the English sound “in situ” and sought-after international projection, in light of their distrust of the technological possibilities and human resources in the recording companies’ local studios. But the fact that Milhaud began to work on his Spanish recording productions (Los Bravos, Los Canarios, Los Pop Tops) with the engineer Bill Price at Decca and with Adrian Kerridge at Lansdowne Studios, implied, on the other hand, the importation of certain work methodologies that were progressively established in the new business fabric in Spain associated with the emerging recording studios from the second half of the 1960s. Lon-
don was not only the Mecca of pop-rock, the city that had catapulted bands of reference for Spanish artists onto the international stage, but the nerve centre of the great recording studios: sounding like the English bands necessarily implied travelling to make recordings at the same studios in which these had been made. As part of certain work dynamics in which it was common for a band to record a demo at the record company’s studios and subsequently go to London to make the definitive recording before returning to Madrid, it was inevitable that a progressive incursion into the work methodologies of the British studios would take place in Spain: microphone techniques, the use of reverb, stereo panning of the elements of the mix, fitting out of equipment, organisation of rehearsals with session musicians, etc. Without entering into the debate as to whether or not recording engineers in Spain during this period were technically qualified to take charge of the recordings of the emerging pop and rock groups, it is important to point out that apart from technical knowledge, their main shortcoming was a considerable ignorance about the timbral characteristics of a new music that was built around a new concept of the recording studio. From this point on, the recording studio became a creative space through techniques derived from multi-track recording, microphone placement, and the effect of reverb or distortion, which now became one more artistic parameter. While the 1962 song by Los Gatos Negros analysed here and recorded during Milhaud’s Barcelona period had already completely assimilated the stereo format, the work methodology inherited from mono sound was still in force and was reflected through the scarce use of panorama in the horizontal axis of the mix. Although the use of stereo now allowed a particular sound space to be simulated, the opening up of the stereo field was modest, with the majority of the elements of the mix still located in the centre.

The necessary “adaptation to the genre” of technical staff accustomed to making live recording of symphony orchestras, stars of copla and other popular or traditional music genres, gave rise to a lack of trust on the part of some producers. In some instances, as in Milhaud’s case, Spanish producers of this period decided to record in studios outside Spain (for example, Maryni Callejo chose to make his recordings in Italian studios). But the social impact of pop-rock music, and the definitive commitment of recording companies in response to the commercial viability of the genre, led to an essential upgrade in the technical resources of some Spanish recording studios. Many evolved substantially during the second half of the 1960s, and grew exponentially over the next decade, with the appearance of a significant number of independent studios at the service of the large recording companies.

The visual representation of the stereo mix, based on the theoretical approaches of authors including Allan Moore (2012), have assisted this analysis of the distribution of the elements of the mix in the stereo band, as well as the progressive incorporation of musical instruments from left to right, using
the stereo contrast as another technique available to the artistic producer. On the other hand, visual monitorisation instruments, such as the spectrum analyser or Lissajous Meter, indeed reflect the timid approaches to stereo in the early recordings of rock music in Spain, a consequence of the technical limitations and the professional inability of recording engineers to adapt to the new musical genre in equal proportions. The qualitative leap in the work carried out in British studios represented a point of reference for Spanish audio engineers who, as in the case of Pepe Loeches, decided to supplement his professional experience in London. These professional experiences, together with the influence exerted by the presence of recording studios belonging to delegations of international recording companies in Spain, as in the case of EMI, would lay the foundations for the development of stereo mixing techniques in Spanish music recording.

References

Bibliography


Audiovisual resources
RTVE.es, (2014) Música Ligerísima. Program N.2. [video] Available at:

Discography
Los
Martin K. Koszolko: The Tactile Evolution —
Electronic Music Production and Affordances of iOS Apps

Abstract

In this paper, grounded in my music production practice, I investigate the various affordances of iOS apps in the field of electronic music production. Drawing on affordance theory and notions of portability and performativity, I consider how contemporary technologies have generated renewed approaches to electronic music production and performance. My analysis shows that portable, touchscreen technologies offer new ways of interaction and musical expression, and facilitate experimentation, which empower producers to redefine their methods of music creation in the contemporary digital space.

Introduction

A growing trend among the community of music producers is a return to the tactile experience of music creation tools and a focus on hardware that reduces the reliance on desktop systems. The different feel of mobile technologies impact upon the way producers choose to engage with particular formats. This paper explores an area of the contemporary digital landscape that is significantly changing the creative practice of music producers around the world. I investigate the various affordances of Apple iOS-based mobile music technologies in the field of amateur as well as commercial electronic music production. In doing so, I examine selected aspects of iOS music apps and discuss various ways in which they impact on practices of contemporary music producers. I also discuss how selected music production apps facilitate particular methods of performance and production.

The launch of the iOS App Store in 2008 signalled a significant moment for enthusiasts of mobile music creation as it allowed access to a growing range of tactile music making and sound production apps. The store features apps that are priced significantly lower than typical plugins or Digital Audio Workstations (DAWs) for macOS or Windows and can be utilised on mobile devices such as iPads, iPhones and iPods. Relatively low prices of iOS apps encourage musicians to experiment with various approaches to sound pro-
duction inherent to the platform and deepen the process of democratising music creation.

In my creative practice, over the last five years I have been incorporating two models of iPads (3 and Air 2) as well as an iPhone 5s. This included my live performances as well as studio recordings where iOS software has been used standalone or alongside desktop computer programs. My practice is the key aspect informing this paper and the selection of discussed affordances. My work has been documented in the form of recordings, blog posts (Philosophy Of Sound 2018), screenshots and interviews given to international media outlets, such as radio and press. I drew on this documentation in gathering evidence for this paper.

In addition to grounding in my creative practice, my research also incorporates ethnographic methodology and draws upon interviews and data gathered through my participation in several online communities, of which the most important is the iPad Musician Facebook group with the current membership of over 9000 users. I used group posts to ask what affordances are important to the members as well as what specific setups they use. A selection of these setups is visible throughout the paper. In addition to discussion in group posts, I followed up with private Facebook messages, email and in one instance, a face-to-face interview (Haq 2018). Overall, I received responses from approximately 50 group members. My interaction with the group, over the five years of my memberships, indicates that a large number of members are mature age users with a wide range of musical experiences: beginners as well as substantially experienced musicians who work across electronic and non-electronic music styles and play variety of instruments.

Most of the participants interviewed for the purpose of this study can be classified as bedroom music producers. The producers discussed here undertake multiple roles reflected in Burgess’s understanding of a producer as an auteur, a self-produced artist leading the creative process (2013). While there is a significant body of work that explores the impact of new technologies, there is little work on the impact of music production apps. As such, this paper fills a gap and explores one area of the contemporary digital landscape that is significantly changing the creative practice of music producers around the world. Drawing on notions of mobility and performativity, I consider how contemporary technologies have generated renewed approaches to electronic music production and performance. I do not intend to present a comprehensive overview of all available workflows and key apps, but rather, a point of view based on my own experience of producing music with iPads with a reference to discussions within the above-mentioned community of iOS musicians. The key affordances of mobile music production tools that I would like to highlight, concern the following five areas:

- iOS-specific workflow, which allows accessibility and flexibility but also introduces certain constraints
- **Performance**, which enhances *playability* and *improvisation*
- **Acquisition of knowledge**
- **Sequencing**
- **Mobility**

The relationship of affordances listed in Figure 1 is demonstrated in their position in the graph and will be further explained in the remaining sections of the paper.

Figure 1: The key affordances of iOS mobile music production tools. (Image credit: Martin K. Koszolko)

Incorporation of iOS devices in one’s music production practice requires challenging certain assertions which deny that such devices can be serious contenders to laptop/desktop computers (Sinofsky 2016). On the one hand, the popular press provides reports such as the one on how the song ‘PRIDE.’ from Kendrick Lamar’s album ‘DAMN.’ (2017) was produced by Steve Lacy with the use of just an iPhone 6 and an iRig interface used to connect guitars directly into the iPhone (Hein 2017). On the other hand, reports such as the one by Dobrev (2017) show that certain challenges of the platform can be difficult to overcome for some users. Interestingly, the comments section underneath Dobrev’s article presents an opposite point of view expressed by music producers, who praise the positives of the iOS-based workflow.
Theoretical background

The theory of affordances stems from Gibson’s assertion that “[t]he possibilities of the environment and the way of life of the animal go together inseparably” (1986, 143). Building on Gibson’s ecological psychology approach, Norman proposed that in the realm of information technologies and product design, the use of affordances relates to cultural, physical, logical and cultural constraints and conventions (1999). “Affordances per se are independent of perception” (Gaver: 1991), however, as noted by Gibson, “affordances are properties taken with reference to the observer” (Gibson: 1986, 143). Other scholars have also highlighted the subjective properties of affordances, which relate to our experiences and perceptions (Capulet and Zagorski-Thomas: 2017). Volkoff and Strong stated that affordances “represent the potential for action rather than action itself” (2013) and that there are different ways in which an affordance can be actualised. Following on this assertion, my discussion concentrates on affordances that can be empirically observed. Moreover, my analysis of affordances of the iOS apps confirms Volkoff and Strong’s statement that “multiple affordances exist and may interact with each other” (2013).

The theoretical background for my brief discussion on creativity is drawn from a selection of scholarly texts on this topic (Csikszentmihalyi 2014, Kaufman and Sternberg 2010, Russ 2013, Sawyer 2003) as well as texts analysing the creative process of music producers (Bennett 2011, Hugill 2012, McIntyre 2008). I use the term ‘creativity’ to refer to outputs and ideas that are novel as well as acknowledged by experts in the field as possessing value. As asserted by McIntyre (2008) similar definitions are found in a broad range of texts on the subject. Authors, such as Blanning (2016), Mumaw (2013) and Stokes (2006), point out that creativity needs restrictions to thrive. As Stokes (2006) demonstrated in her research, constraints lead problem solvers to more variable and more creative outcomes. My statements on creativity are also linked to the notion of constraints in the workflow of producers using iOS apps.

Workflow

The workflow of a producer using portable, touchscreen devices with iOS apps is often noticeably different from the modes of work afforded by desktop DAWs, where the majority of music production still takes place nowadays. To characterise the workflow involving iOS apps, I will discuss its three facets: accessibility, flexibility and constraints. These facets are intertwined with other affordances that overlap with ‘workflow’ in Figure 1.

Working with iOS devices is characterised by the speed of access of desired tools, since tablets and phones are typically always on. Musical ideas
can be captured quickly with minimal setup time, particularly when using the gear standalone without the need to interface it with other equipment. The setup can be as quick as waking the device up and launching a chosen app. Because of that, touchscreen devices are very helpful in quickly capturing and sketching ideas. I tend to use these devices as portable recording machines to capture location sounds for further productions as well as for improvised live performances as KOshowKO or Iubar Project, where I record my audiences and then manipulate the recordings live while creating improvised compositions. For the latter, I tend to use two apps ‘iDensity’ by apeSoft and ‘Yellofier’ by Boris Blank. The first allows me to apply granual synthesis to sampled material and the second allows for quick live sequencing of samples recorded during my live shows.

The small size of touchscreen devices increases the accessibility factor. This is further intensified by a possible lack of reliance on additional gear, such as studio monitors, MIDI controllers and associated cables. As I will discuss further, additional equipment is often used in conjunction with mobile devices, although users have the choice of using their devices standalone. Limiting oneself to just the touchscreen control means that music producers interact with music apps with tactility that is aligned with specific features of the chosen apps.

The workflow associated with touchscreen tools imposes certain constraints, for example a limited file system within iOS or, on occasions, limited inter-app connectivity. Furthermore, it is common that producers using iOS work with just one device. In my experience, the restrictions imposed on the amount of equipment and options available to a music producer and performer can help to streamline the creative process via increased accessibility. A limited setup allows for faster music creation and with a lesser scope for technical problems to interrupt the creative process. A practical example taken from my work is that I tend to restrict the number of apps that I use on a single device. I prefer to run apps in a standalone mode outputting their signal directly to a venue’s mixing desk (when playing live) or my DAW (when working in a studio). With this setup, I manage to remove the reliance on mixer apps such as ‘Audiobus’ by Audiobus Pty Ltd and ‘AUM’ by Kymatica, and therefore, can concentrate on the creative work first and foremost, without having to do MIDI mapping for additional hardware or setting up processing chains within mixer apps.

There are also some external constraints that touch screen devices allow to overcome. For example, constraints of desktop computer systems which enforce the reliance on using a computer mouse—a device that has not been designed with music creation in mind and can be detrimental to a creative workflow. To quote Alexander Randon, who builds music apps as Alexandernaut:
I’m convinced that using a touchscreen — as opposed to using a mouse — has a HUGE impact on workflow. After my experience with iPads over the past few years, I now notice how constrained I feel when using a mouse. So much so, that it often breaks my musical flow. (Cited in Wilson: 2016)

An additional aspect of iOS music production workflow is the use of touchscreen devices in studio or live settings alongside other gear, such as hardware MIDI controllers and hardware synths or desktop DAWs. While iOS based ‘in the box’ music production is often pursued and can be particularly successful in relation to experimental styles of music, there are technologies, such as ‘musicIO: Audio and MIDI over USB’ by Power Trio that allow music apps to become a part of a larger ecosystem of instruments that are being controlled via a desktop DAW. Examining how additional hardware equipment is used alongside iOS devices illustrates the flexibility of available setups. Figure 2 provides an example of a setup used in a live performance context by one of iPad Musician group members—Mathias Hellquist.

![Figure 2: Live performance setup of Mathias Hellquist from Enköping, Sweden. (Image credit: Mathias Hellquist)](image)

The workflow of iOS musicians frequently involves studio or stage performance and for that reason, these areas overlap in Figure 1. Hellquist uses an iPad and iPhone in conjunction with Alesis MIDI keyboard, Roland’s FC-100 footswitch pedal and EV-5 Expression Pedal. All other names listed in
Figure 2 refers to iOS apps. Hellquist describes the construction of his setup as follows:

I was trying to find the best setup where I could play live in a preconfigured environment, and still be able to cover almost any need. Another aspect was that I wanted it to be controllable from my MIDI keyboard (Alesis vi49 + footpedals), with only minor screen interaction. Thirdly, I wanted to ‘stress test’ my iPad Pro (12.9”; 1st generation) to see if it could handle it all. For me it was also important to be able to record both MIDI and audio, but I was trying to avoid using a regular iOS DAW (such as Auria, Cubasis etc). These tests were also carried out before apps like LayR, BeatMaker 3 and Xequence had been released. (Email interview: January 15 2018)

Over the last five years, despite the occasionally problematic cycle of annual updates of the iOS, several technologies have emerged that simplify the process of playing multiple apps synchronously. Of these technologies, the most notable examples include apps such as ‘Audiobus’ and ‘AUM’. However, a pivotal change arrived with the introduction of a protocol established by Ableton and called ‘Link’, which enables electronic instruments to be synchronised over a wireless network and is being implemented in a growing number of apps.

My own workflow tends to bypass iOS DAWs and focus on playing individual synthesizer apps in a live performance context or recording into Ableton Live installed on my home studio computer. An example of the latter is audible in ‘Untold’ by Iubar Project vs Modus Op (2017). I have, however, produced some songs entirely on the iPad and sometimes even with the use of a single app as it was the case with ‘Now You’re Talking Baby’ for KOshowKO (2015), which was created with the ‘Yellofier’ app. Many iOS musicians aim for iPad-only based workflows centred on either a standalone DAW such as ‘KORG Gadget’ by KORG INC, or commonly, a mixer app such as ‘Audiobus’ and/or ‘AUM’ (see Figure 2). The mixer app scenario incorporates several other apps being con-

connected to the mixer which can distribute the signal to other destinations such as iOS DAWs or standalone recording apps such as ‘AudioShare’ by Kymatica. This type of workflow often incorporates live sequencing which I will discuss in a further section of this paper. Figures 2 and 3 illustrate that dependence on desktop DAWs is not mandatory and iOS producers increasingly treat their iOS devices, particularly iPads, as the central recording and sequencing element of a studio, replacing desktop computers altogether.

Retaining the touch-device centred approach while limiting additional gear has the added benefit of overcoming the constraint of a static studio space. This notion has been in the past analysed by Theberge (2004) in relation to the concept of remote, networked studio and in my own research on remote music collaboration software (Koszolko 2015, Koszolko and Montano 2016). A portable setup, with an iPad replacing the need for a desktop DAW can be also seen in Figure 4 and its benefits will be further discussed in the section on Mobility.

The above summary of workflows possible with the use of iOS-based devices paints a picture of a flexible environment where a traditional studio approach is possible but not necessary. Similarly, in the context of live performance, users can choose between standalone, iOS-only gear or incorporate additional equipment. My experience as a live performer indicates that standalone operation intensifies the exploration of new ways of interaction with musical instruments in the form of iOS apps.

**Performance**

The affordance of performance is implied primarily as playability, which can be incorporated in a studio setting or on stage. When using iOS apps, the key enhancer is, naturally, a touchscreen interface and also features such as the gyroscope and accelerometer that allow for more fluid and new playing styles in comparison to what can be accomplished on traditional hardware. This is characterised by the graphical user interface (GUI) that allows for touching, sliding on and wiggling the virtual keys or buttons in order to modulate the sound with parameters such as note on and off, vibrato, pitch bend and velocity. The ability to perform these actions is of course depend-
ent on a specific instrument design. Examples of synthesesers allowing for powerful gestural modulation include ‘TC-11’ by Bit Shape and ‘Mazetools Soniface’ by Ectoplastic UG (Figure 5). The former is a modular synthesiser and the latter is described by the manufacturer as a visual synthesiser.

Several apps allow manipulation of multiple sonic parameters with various finger gestures. For example, in ‘Mazetools Soniface’, enlarging the distance on the screen between two fingertips defines the radius in which notes are played. Doing the same with three fingers controls the level of distortion, while turning three fingers clockwise adds the vibrato effect. As stated by Haq in one of his video guides (2015), learning how to program sophisticated apps such as ‘TC-11’ can be daunting and requires that users invest time into learning how to harness the power of the app. This example illustrates that app designers can utilise possibilities of touchscreen devices and create complex instruments that demand users’ attention and devotion of time in order to master the operation and customisation options. Similar to what is expected from a musician learning to play an acoustic instrument.

Music production tools that I have used facilitate experimentation. New ways of playing various iOS apps go beyond gestures described above and include direct interaction with audio content by “touching visual representations of the actual waveforms” (Johnston: 2015, 27) or exploring microtonal tunings, which is possible with selected apps (Burt 2016). Playable waveforms are available in apps such as ‘Samplr’ by Marcos Alonso and ‘iDensity’ while microtonal apps include ‘ThumbJam’ by Sonosaurus LLC and ‘Shoom’ by Yuri Turov. I use these apps frequently in my live performances (KOshowKO 2014) and, similar to Johnston (2015), I find their designs enhance the scope for improvisation. The way these apps afford improvisation, is linked to their GUI and aspects of playability as well as expressiveness. When I made a move from performing live with traditional digital technologies, such as Ableton Live and associated MIDI controllers to performing...
with mobile devices, I felt liberated to start improvising a lot more. It has been intensified as I became drawn to the vast array of unusual GUIs and able to access multiple app interfaces relatively quickly within a performance of a single composition and on a single device. This, in turn, led me to becoming less bound to predefined song forms. As I will discuss further, improvisation is often incorporated in the context of live sequencing as well.

A touchpad, which Johnston refers to as a post-keyboard interface (2015), is another feature that is frequently integrated in music apps, and which enhances sound manipulation options available to a performer. The use of a touchpad for effects processing is part of the sound design and production process where live manipulation of effect parameters becomes a critical part of live performance. Touchpad enhanced music apps are performance tools that allow a tactile approach, which helps with streamlining of certain operations, such as manipulation of various parameters of an instrument in real time. A touchpad can be also used as a note triggering interface, an example of which is the ‘Shoom’ app, which is an expressive XY pad synthesizer, which includes three identical sound engines and is capable of playing any pitch in the audible range and does not limit a user to a particular scale.

In addition to innovative GUI designs, the iOS platform affords performing with software emulations of classic hardware synths. This is an extension of the shift from hardware to software that we have seen in software plugins on desktop computers. Emulations of existing hardware often gain in expressiveness or new features after being ported to iOS. Examples include ‘iSEM Synthesizer’ by Arturia, a model of the 1974 Oberheim SEM synthesizer and ‘iVCS3’ by apeSoft, which is an emulator of EMS VCS3 synthesizer from 1969.

The novelty of performing with these apps is threefold. Firstly, tablets and phones offer access to instruments within a lighter, smaller and therefore more portable device than hardware synthesizers or desktop computers. Secondly, the user interaction is directly connected to instruments’ GUI through hand gestures rather than through additional devices such as MIDI controllers and computer mice. This interaction allows for a tactile approach, more akin to using the original hardware. Thirdly, emulations of older synthesizers often include features not available on the hardware. For example, ‘iVCS3’ has an added sampler allowing capturing of audio via a microphone or from other apps (Nagle: 2016).

**Acquisition of knowledge**

The affordance of the acquisition of knowledge is related to four other affordances (see Figure 1) and can be manifested in several ways. Musical knowledge can be acquired through the use and examination of custom keyboard layouts, which is a feature of multiple iOS synthesizers, MIDI control-
lers and music composing tools. These apps enable musical performance within a predefined musical scale or mode and therefore can be used to improve one’s understanding of various aspects of music theory, including intervals, scales and chord progressions. The use of such apps “makes improvisation easier due to the absence of unwanted notes” (Johnston: 2015, 24).

Limiting options can help less technically proficient musicians to progress but it can also benefit ones that are more experienced because, for example, limiting available notes to a predetermined scale can make them think differently about melodies (Claudio: 2017). The scale definition feature affects the production workflow as it simplifies playability across multiple instruments within one composition. This in turn, allows the performer to focus on other nuances such as improvisation, articulation, melody building or live sequencing, since multiple iOS sequencers also include the ability to work within predefined musical scales. This ability means that acquiring the knowledge of music theory is equally possible on many sequencers as well as instruments.

Musical knowledge can be acquired in other ways than only scrutinising predefined scales and limiting notes available to the performer. Apps such as ‘Navichord • MIDI controller’ by Denis Kutuzov (Figure 6) demystify the relationships between notes of which chords are made and allow music producers to write and reorganise chord progressions. This is a dynamic process afforded by the touchscreen GUI where visible connections between notes change depending on the scale type and degree selection.

The relation of knowledge acquisition to mobility is primarily linked to the ability to facilitate learning in various, not necessarily musical environments. Learning can be conducted with the use of performance and composition enabling tools like ‘Navichord • MIDI controller’ but also with an array of interactive apps that teach music theory as well as a range of aural perception skills. Examples include ‘Music Theory and Practice by Musicopoulos’ and ‘Ear Training Course by Musicopoulos’. Both were created by SpartanApps and do not rely on Internet connectivity which further enhances mobility.
An additional element that facilitates learning of various facets of iOS music making are the globally dispersed communities of practice. The iPad Musician Facebook group as well as the Audiobus forum are two of the largest online communities of iOS music makers. Members discuss music production tools and obtain help from others, including the developers of various apps. There is a symbiotic relationship of app developers and their users visible in the above-mentioned communities. Since app developers frequently answer questions about their products and respond to requests for features from the users, the discussions can be seen as a form of market research as well as an opportunity to advertise new products or features.

Another area of my academic research is Remote Music Collaboration Software (RMCS) and iOS apps can also facilitate work in this sphere. Collaborative music production practice provides a fertile ground for knowledge acquisition and learning from people with different or more advanced skills than ours (Koszolko: 2015). There is a growing number of collaborative apps exploring different approaches to communal music making. Examples of self-contained collaborative apps taking form of relatively simple DAWs include ‘Songtree - Music Maker’ by Songtree S.r.l. as well as ‘Soundtrap’ by Soundtrap AB. Allihooopa (closed down in 2019), was a networking platform for music makers that facilitated a different approach by allowing remote collaboration on music created with various third-party apps. Examples of apps that allowed to share audio to Allihooopa include DAWs such as ‘KORG Gadget’ as well as instrument apps such as ‘Model 15’.

Sequencing

iOS-based sequencing can be performative, meaning a real-time activity in the context of live performance. In addition, it can be also executed as a part of the song arranging process, which is more likely to happen in studio settings. There are iOS DAWs such as ‘Cubasis’ by Steinberg Media Technologies GmbH or ‘Auria Pro’ by WaveMachine Labs, Inc., which to a large degree behave like desktop DAW software. Sequencing with iOS DAWs is a process resembling work with desktop tools more than any other area of the iOS music production.

In addition, some apps feature built in sequencing, which reduces reliance on DAWs. These apps can be standalone sequencers relying on externally supplied MIDI inputs or self-contained music making environments capable of live sampling, signal processing and sequencing. The previously mentioned app ‘Yellofier’ serves as an example of an in-app sequencing that can function in the context of improvised live performance. The app has been co-designed by Boris Blank, a member of the electronic group Yello, whom use the app in their live shows (BR TV 2017).
Apart from a rather traditional, user-initiated step-sequencing, iOS apps allow for self-generation of MIDI sequences. The use of generative sequencing is a process where “a composition algorithm serves as a generative engine for music creation” (Roads: 2015, 339). Sequencing approaches involving randomisation and self-generations are equally useful in the studio and on stage. An app with particularly well developed generative sequencing is ‘Rozeta Sequencer Suite’ by Bram Bos.

The use of machine-determined compositional sequences can lead to questions about the virtuosity of the musician (Bowen: 2013) and since several iOS apps can facilitate randomisation and self-generation of patterns, the debates concerning virtuosity are relevant in the context of iOS apps and more broadly speaking, producing music with mobile technologies that afford such approaches. As observed by Hugill “machine musicianship grapples with essentially the same issues as human musicianship” and elements of human musicianship include “technical mastery, critical judgement, aural skills, musical literacy, ensemble abilities and creativity” (2012, 174). And while virtuosity can be questioned when musicians using generative sequencing systems apply little or no critical judgement or have limited musical literacy, it is also worth acknowledging that “generative strategies are conceptually attractive” (Roads: 2015, 339) and that their attraction lies in the fact “that they let composers control sonic processes on a level that would be impossible without algorithmic assistance” (2015, 341).

Mobility

In her research on the mobile user experience, which preceded the launch of the Apple app store and availability of music apps, Ballard attempted to define parameters of mobile devices (2007). As she noted, this definition was not without its challenges as, for example, laptop computers can also be categorised as mobile. Defining the unique aspects of mobile electronic technologies, Ballard’s the ‘Carry Principle’ includes the following characteristics: wireless communication, multifunction device, battery powered, small, personal, and always on (2007, 231). This richness of characteristics, inclusive of multi-functionality, poses a challenge to end users since they need to redefine assumptions of what constitutes musical instruments as well as their perception of the music practice (Jones 2015).

Mobility, inherent to all iOS devices, affords music creation with apps in a variety of locations, and on the go, without the need for any additional hardware, with the possible exception of headphones. Jakob Haq is a mobile music producer from Stockholm who frequently reviews and experiments with various apps on his YouTube channel. As stated in the interview that I conducted with Haq in Stockholm, iOS apps allow him to make music during otherwise idle times, such as long commutes to work (Haq 2018). Fur-
Furthermore, Ben Kamen, who makes music apps under the name of Olympia
Noise asserts:

I really like approaching the iPad as an instrument, and tend to use apps more
that fit well into performance and live music making. It can also be a great,
portable way to generate ideas away from all the cables and chaos of a studio.
(Cited in Wilson 2016)

Members of the iPad Musician Facebook group frequently publish photos
labelled ‘current studio view’ displaying a wide range of locations where
they make music. Work in those locations typically relies on the battery
power of touchscreen devices and is often carried outdoors. Showcased set-
ups are frequently very simple, involving just one device with headphones,
although occasionally, larger setups are also present (see Figures 4 and 7).

As visible in Figure 7, mobility is not only afforded because of the small
size of individual iOS devices but by the fact that they are battery powered.
Wireless communication is also an asset when it comes to technologies like
Ableton Link, allowing for wireless synchronisation of apps on separate
devices. Figure 7 showcases four iPads and two Samsung Galaxy tablets
with additional hardware used in a remote location and utilizing battery
power for all gear. This photo is also demonstrating how mobility is inter-
twined with the aspect of performing live with iOS devices.

A significant question in the context of mobile music production is how
much can be achieved on iOS devices without the confines of the physical
studio? The latter often incorporates monitors, microphones, instruments,
interfaces, MIDI controllers and acoustically treated spaces, which all infl-
fluence the final production in various ways. As I indicated earlier, I use a fully
iOS-based mobile setup primarily in my live stage performances. P
roducing
complete musical compositions with nothing else but an iPad is less likely
for me in the studio, although I frequently use iOS apps as auxiliary sound
sources captured and later edited with a desktop DAW. The reason for this
approach is twofold. My desktop setup is a result of over 15 years of acquir-
ing and learning about a large collect-
on of software and associated work-
flows. This means that I have more
options available to me during the music
production process and that I can
achieve desired out-
comes faster while

Figure 7: Willard Van De Bogart performing in Clark
Valley in the Anza Borrego Desert in Southern California:
2016. (Photo credit: Michael "Lemon" DeGeorge)
mixing and producing music with a desktop DAW. My ethnographic research confirms that iOS musicians are divided on the usefulness of ‘in the box’ music production on iOS and while some of them are adding additional hardware or use desktop DAWs, many others are content with nothing else but touchscreen devices.

Conclusions

As I demonstrated, affordances of iOS apps are often intertwined and while I attempted to draw certain demarcation lines, the discussion on a specific affordance often transitions into another one. Creative immersion in the ecosystem of iOS apps opens up access to a multitude of workflows and new ways of interaction, playability and knowledge acquisition. Those innovations can empower producers to redefine their methods of music creation in the contemporary digital space and challenge their predefined notions of music production.

Music apps allow for the attractive combination between immediacy, control and unique sound design possibilities. They function in a relatively young and comparatively restrictive environment of the Apple iOS, which reduces certain options, for example, flexible connectivity and data exchanges with external hardware. Discussions among many producers within the iPad Musician group on Facebook indicate that while some users embrace an exclusively iOS-based music production, some others find available connectivity and setup options too limiting and as a result, use the iPad as a sketchpad for musical ideas, which are later developed with the help of desktop DAWs. Nevertheless, the iOS app store offers a very large array of apps featuring innovative designs that often challenge old paradigms of how music production tools are designed, how users can interact with the GUI and in what environmental settings music can be created.

Exploration of various workflows afforded by iOS apps demonstrates that in comparison to desktop music production methods centred on established DAWs, iOS tools stimulate flexible working methods where producers can explore new approaches to composition as well as multiple ways of interconnecting and manipulating the sound generating and sound processing apps. In this context, we witness, “the formation of intimate relationships between a musician and their tools” where “passive commodity consumption is resisted, and innovative/potentially lightweight workflows emerge” (Ferguson and Brown: 2016). Workflows inherent to the iOS environment allow for an all-in-one approach where recording via built in or external microphones, processing, post-production and live performance are possible with the use of dedicated apps. Moreover, these workflows allow enhancements of music production with video clips and artwork creation with the use of specialised apps.
Multiple apps available on iOS present innovative sound design approaches that open up uncharted sonic experiences for music producers, yet the connectivity between multiple apps requires the understanding that not all methodologies that are possible with desktop DAWs can be implemented on iOS. Currently there are several protocols facilitating connectivity—Ableton Link, Audiobus, Audio Units and Inter-App Audio—however, not all apps support the same protocols, which may increase the perception of reduced stability or reliability of the platform. Dobrev’s frustration with “the lack of a seamless, universal, and consistent standard for beaming audio, MIDI, and controller data between the iPad’s fantastic apps” (2017) can be seen as a craving for a stable production environment. Yet, as discussed by Lefford “[s]tability is problematic in domains like art and music production that value innovation” (2015). This represents a dilemma that producers utilising iOS tools often face: they are presented with a set of new approaches to creating and manipulating sound, but their utilisation often demands the abandonment of the ways of work known from desktop DAW environments. My own experience further confirms the dependency on desktop DAW workflows, however, only in the studio and not on stage. Differences between iOS and desktop methodologies highlight that the former are an expression of the exploration of uncharted territories of GUI and software design which allow producers to celebrate “uncertainty through improvisation and algorithmic processes” (Ferguson and Brown: 2016).

Various touchscreen music production tools that I incorporate in my creative practice and of which a few examples are featured in this paper, offer a distinctive qualitative character as well as new ways of playability and musical expression. Moreover, such technologies facilitate experimentation, acquisition of musical knowledge and enable tactile interaction. Despite its limitations, iOS music production apps offer significant advantages over various other existing technologies. These advantages include mobility, app affordability, new ways of designing the graphical user interface and an innovative approach to music production and performance.

References
BR TV. (2017) ‘Zu Gast im Studio, Yello’. In: BR Mediathek [Online]. Available at:  


Dobrev, L. (2017) ‘Why the iPad is an amazing – and awful – tool for making music’. In: Gearnews [Online]. Available at:  


https://youtu.be/xSFH01tPdD0 (Accessed: April 2018)  

Hein, B. (2017) ‘One of Kendrick Lamar’s new songs was created entirely on iPhone’. In: Cult Of Mac [Online]. Available at:  


KOhshowKO. (2014) ‘SAMPLR and Novation Launchkey apps live jam (KOhshowKO - In the Night)’. In: Philosophy Of Sound, YouTube [Online]. Available at:  


ble at: http://www.kinephanos.ca/2016/cloud-connectivity-and-contemporary-electronic-dance-
music-production/ (Accessed: July 2017)


studio/ (Accessed: February 2018)


Discography:


Paul Novotny: Creating an Immersive Fold-Out – Look Ahead

Abstract

A comprehensive autoethnographic case study, detailing the practice, engineering science and research used to create a stereo to 5.1-fold-out of the Canadian piano and bass jazz duo recording, Look Ahead. This paper asserts that a stereo to 5.1-fold-out, rather than fold-down is a preferable method because it affords each version a creatively distinct virtual acoustic environment by unique preparation. Fold-out results in two separate and unique masters that share a strong common foundation of ensemble sound. The conclusion reveals universal insights gleaned from the practice and research, henceforth assisting creators to adapt a stereo sound-field into a multichannel immersive sound-scape.

Introduction

Some years ago, I recall hearing my first playback of a jazz trio in surround-sound. I sat in the ideal listening position, looking ahead to the left, center and right speakers (L/C/R). The first thing I heard was a solo piano introduction, but surprisingly it came from the rear left and right speakers (LS/RS). As I turned around to face the music, the bass and drums entered from the L/C/R speakers which were now behind me. I was confused. My focus was not on the music because I was seeking a listening position that made the trio sound like a cohesive ensemble. In the end, I could not. That experience heightened my curiosity and I wondered what my music might sound like in 5.1, what process I would use to make it, and if others were asking these questions too.

This paper is an autoethnographic case study of the practice, engineering science and research involved in making the jazz duo recording, serendipitously titled, Look Ahead, recorded at 24 bit/96kHz for playback in stereo and 5.1. It reveals the thinking, process and reasons for a stereo to 5.1-fold-
out, rather than a fold-down and it describes why fold-out is preferred by audio engineers for multichannel music production.¹

I introduce terminology that differentiates stereo and surround-sound audio contexts with the hyphenated terms, sound-field and sound-scape. It seemed efficient to define the studio-made L/R stereo front image of a recording as a sound-field, while sound-scape refers to a studio-made 360-degree multichannel surround-sound audio image.

**Sound-field versus sound-scape**

Sound field (no-hyphen), created by L/R speakers in a room is a widely-used term but basically its job is to represent the nature of the sound we hear in that acoustic environment.² Components of the sound field include sound pressure level, anechoic or dry sound sources and their related perspectives of distance, horizontal position and room size. These traits are perceived from early reflections, reverb diffusion and spectrum texture in the recording but heard more simply as the performance staging and acoustic environment in recorded music. When music is constructed in a studio control room, the reflections of a virtual sound-space become very important because they indicate that the performance is occurring in either a natural environment or an exaggerated hyper-real environment. Alan Moore (Moore: 1992) described an abstract model called the soundbox – a ‘virtual textural space’ capable of replicating proximal and distal perception of depth spatiality as intimate, private, social and public.

The 2.0 sound field capably conveys these four perspectives through stereo L/R speakers, but like the human “field of view” and visual screen technology, the constructed L/R “sound field” is limited to the horizontal peripheral edges. Moore also uses the terms “vertical location” to express musical register (high and low notes) and “horizontal position” for L/R panning. William Moylan (Moylan: 2002, p.14-15) describes elevation thoroughly and says, “to date, the vertical plane has received little attention in audio because of playback format difficulties.” However, current multichannel surround-sound for games, 360-degree sound for virtual reality (VR) and artificial reality (AR) are becoming more mainstream and so expansion of Moore’s soundbox theory may now be appropriate. Use of the term “vertical

---

¹ Wolpert, Jeff, Adjunct Professor of Music & Media, wrote, “in the engineering community, a stereo to 5.1-fold-out is an accepted practice for music production”. University of Toronto, 11-19-17

² The British Society of Audiology provides information on international standards of sound fields used in audiometry. “ISO 8253-2 (1998) describes three types of sound field which are defined by the allowable variation of sound pressure level, produced by the output of a loudspeaker, in a small space surrounding a reference point. The reference point is roughly at the midpoint of the head of a hypothetical listener.” It also goes on to define a free sound field as anechoic, a diffuse sound field where walls and ceiling have significant effect and a quasi-free sound field where the walls and ceiling have a moderate effect.
location” to express musical register has become unclear since elevation can now be well represented in multichannel playback formats such as Dolby Atmos, Auro 3D and DTS:X.3

The purpose of Look Ahead in 5.1 is to expand the limitations of the recorded virtual front L/R sound-field (hyphenated) into a virtual multichannel immersive sound-scape, thereby improving the perception of all “soundbox” traits. Like a TV monitor with more pixels, the SMPTE/ITU–5.1 (L, R, C, LFE, LS, RS) multichannel format provides improved positional definition, acoustic depth of field and spatial context because there are more speakers, thereby reducing the intrusion of diffused reflections caused by playback from speakers in an indoor space.

Since every sound-source is naturally mono, a 360-degree immersive sound-scape augments the expert listening experience with better spatial detail.4 It can provide a cohesive virtual re-construction of an acoustic environment for studio-made music. A core difference between sound-field and sound-scape is that the sound-scape’s virtual reflections and reverb can travel toward the listener and then pass behind him/her, hitting what would sound like a rear wall, acoustically locating the listener in the middle of a 360-degree virtual sound-space.

A challenge in construction of the 5.1 sound-scape is creating a believable transition from the L/C/R front to the LS/RS rear. Two helpful techniques are delay and a gentle low-pass filtering of LS/RS signals. But when those techniques are folded-down (down-mixed from sound-scape to sound-field), unexpected problems in the perception of intimate versus public space can occur.

A static fold-down (down mix) collapses the LS/LR signals into the front L/R, creating unpredictable and undesirable spatial reflections, but a fold-out

---

3 I checked both Moore’s original 1992 book (located at the Toronto reference library) and the second edition available on Kindle (re-issued 2018), both refer to ‘vertical height location’ as ‘musical register’. The most current technological development suggests this concept needs updating. At the SIRT, AES, SMPTE conference (Toronto, Pinewood Studios-02/21/18) senior manager of broadcast operations at Bell Media, CTV, Francis Nunan declared “The audio channel is dead, object audio is where we are headed.” He stated that the reason for this is the need for a ‘universal delivery payload’ rather than the current system of ‘versioned broadcast media.’ This advancement will influence music playback as well as broadcast media because playback mediums (clubs, theatres, headphones etc.) are not standardized and rarely sound the same as control-room DAW mix playback. Gaming engines such as ‘Unity’ and ‘Wwise’ are already responsible for rendering spatialized audio in VR and AR, this now includes altitudinal height or elevation. Spatial rendering of audio is moving toward computer graphics processors, meaning that construction of Moore’s “soundbox” traits and “textural strands” is becoming automated and informed by meta-data. As Moore states, “Technology and texture is an ongoing evolution. Technology limits what can be conceived.” Object audio and immersive multi-speaker playback is evolving rapidly and expanding the limits, offering exciting possibilities for sound-scape creation and social reception.

4 In his book, Introduction to the Sociology of Music, Theodore Adorno (1976: 3-14) describes an expert listener as “profoundly competent,” “exceedingly rare” and “able to comprehend the multiple interrelationships present in music during the act of listening.”
avoids these issues because both the 2.0 sound-field 5.1 sound-scene versions are uniquely prepared.

In the appendix, example one supports this observation by providing a side-by-side analysis test—5.1 fold-down to stereo compared to the prepared 2.0 mix—of the bass solo in My Favorite Things, illustrating this undesirable occurrence. The LS/RS reflections were very appropriate in the sound-scene version, but not in the folded down sound-field.

The side-by-side fold-down of My Favorite Things also provided an opportunity to compare Loudness Units Full Scale (LUFS) and peaks. The result showed two masters that looked virtually identical according to the numbers, but sounded and felt very different. The fold-down test can be heard well using stereo headphones.

Research methodologies and report style

This paper utilizes a mixed research method. The recording was a heuristic arts-based improvisatory exercise that was guided by intuition and previous experience. The writing style of this case study is autoethnographic (Ellis, Adams, and Bochner: 2011, p. 273-290) and it attempts to organize the details of practice and research in the logical sequence of pre-production, production, post-production, mix and mastering. The primary questions were:

1. What, when, why and how did pre-existing multichannel formats originate?
2. What social reception was imagined for multichannel audio and what principles might have guided its creators to resolve their work into the eventual pre-existing playback contexts.
3. What universal insights support all immersive audio formats?
4. What principles guide the construction of acoustic environment and ensemble cohesion?
5. According to the history of multichannel sound, the first use of surround-sound as a storytelling device was pioneered by Walt Disney in the movie Fantasia. Multichannel sound employed psychological tension and release to heighten audience expectation and emotion using multichannel audio. This compelled me to look for a psychological and musicological explanation of ‘tension and release’ as it applies to the organization of sound.

5 A-B Side-by-side analysis of 5.1-fold-down versus the actual stereo master. This example displays unpleasing reflections because of the fold-down. (See appendix, Example #1)
These questions suggested four specific research directions: 1) Multi-channel-sound history, 2) David Byrne’s theory of “creation in reverse,” (Byrne: 2012) 3) R. Murray Schafer’s theories on acoustic environment, (Shafer: 1977) 4) David Huron’s ITPRA theory, which describes the five stages of expectation, and reduces them to two experiences of pre-outcome and post-outcome, commonly known as tension and release. (Huron: 2006, pp. 33-34, iBooks) With reflection I gradually realized that development of Walt Disney’s Fantasound was motivated by Disney’s innate understanding of tension and release, or more specifically, David Huron’s ITPRA theory. That notion led me to view all creators of sound-field and sound-scape as “choreographers of emotional expectation and experience”, who instinctively employ ITPRA principles. My conclusion presents seven foundational insights that can guide decisions when creating a sound-scape. It also advocates for the use of a fold-out process to afford the project two uniquely prepared masters, that share and benefit from a common ensemble sound.

Research begins with a chronological overview of multichannel sound

The earliest example of two-channel sound (2.0) over wire comes from Clément Ader in 1878. He placed twelve telephone transmitters in left-right positioning at the stage of the Paris Opera, then ran wires through sewers for two miles to the Electrical Exhibition for reception of the sound. From 1881 to 1930 this stereo listening experience which Ader called “Binauriclar Audition” was commercialized as Théatrephone in France and Electrophone in the UK. Soon, in 1933 Alan Blumlein invented stereo binaural-sound and matrixing. Right from the beginning, two audio channels started engaging audiences and today, multichannel audio such as Dolby Atmos continues to heighten audience expectations and emotions in cinemas and night clubs, such as the Ministry of Sound in London. The pre-existing and future contexts of immersive audio continue to develop, and may soon be ubiquitous due to VR and AR. Table 1 provides a chronological overview showing when these and other key audio advancements came to be.

Table 1: The history of multichannel sound.

<table>
<thead>
<tr>
<th>Date</th>
<th>Multichannel technological development</th>
</tr>
</thead>
<tbody>
<tr>
<td>1881-1930</td>
<td>Théatrephone and Electrophone are commercialized (Scientific American: 1881, Hertz: 1981 pp. 368-372)</td>
</tr>
<tr>
<td>1931-33</td>
<td>Alan Blumlein invents stereo binaural-sound and matrixing (Blumlein:1958) (Gerzon: 1992)</td>
</tr>
<tr>
<td>1934</td>
<td>Bell Labs experiments with 3 channel sound and “audio perspectives” (Pickering, Baender: 1953)</td>
</tr>
</tbody>
</table>

1940-41 Fantasia, Walt Disney introduces 6 track recording deployed in 5 channel Fantasound, (L, C, R, RS, LS) seeking an immersive audio audience experience. Multitrack/panning & overdubbing were invented in this era (Klapholz:1991) (Torick E. 1998)

1952-53 Cinerama (7 channel) and Cinemascope (4 channel) (Dientsfry: 2016)

1970-75 Sansui QS-Quadraphonic-Quintaphonic | Dolby matrix technology (Dientsfry: 2016) Ambisonics was developed. It is a full-sphere agnostic surround-sound technique


1982 Dolby Surround (Julstrom: 1987)

1983 THX Ltd. a company founded by George Lucas that developed a high-quality assurance system and certified accurate reproduction through surround-sound playback systems

1985 Dolby AC-1, Matrixing (Julstrom: 1987)

1987 Society of Motion Picture & Television Engineers (SMPTE) accepts the 5,1 numeric-convention system (Dientsfry: 2016)

1981 AC-2

1991 AC-3

1992 Dolby SR*D

1993-94 93-Digital Theatre Sound (DTS) Sony Dynamic Digital Sound (SDDS) 94-Holophone Microphone system for immersive audio was invented

1995 First AC-3 home theatre decoder

1998 First AC-3 5.1 sports broadcast

1999 Dolby Digital Surround EX-6.1 *(Starwars)* | Dolby E

2000 Dolby Pro-logic II

2004 Dolby Digital Plus (DD+)

2012 Dolby Atmos (3D audio adding height). Approx. 300 theatre installations 2013. SRS Labs “MDA” is a competing “object” based audio system to Dolby Atmos.


8K Super Hi-Vision currently testing by NHK-Japan, (Sigmoto, Takehero, Nakayama, Yasushige, Komori, Tomoyasu, Chinen, Toru, Hatanaka, Mitsuyuki: 2017)


---

6 Ambisonics: http://www.ambisonic.net/

7 Dolby AC-1 was the first digital coding technology, AC-2 had better audio quality and AC-3 included 5.1 channels at a bitrate of 320 kbp/s. AAC is part of the MPEG 4 codec and utilises lossy compression. Lossy compression removes data to make the file size smaller and it compromises definition.

8 In 1994, Canadian inventor Michael Godfrey achieved “the most realistic listening experience” with his Holophone microphone. It uses eight microphones placed in a dummy head and is regarded as the finest surround-sound microphone available.

9 Dolby Pro Logic I decode technology started in 1987. It is frequency limited, matrixed 4.0 surround-sound derived from a stereo mix. The surround-channel is slightly delayed keeping attention focused forward. Pro logic II provides 5 full frequency channels, Ilx provides 6.1 & 7.1 capability and IlZ provides height information to support Dolby Atmos. Dolby Digital AC-3 supports 5 full-bandwidth channels at 640 Kbit/s but DD+ supports up to 15 full-bandwidth channels at 6,144 Mbit/s.


11 SRS Labs developed the Sound Retrieval System technology. In 2008, approximately 36 million SRS equipped flat-screen TV’s were shipped. In 2012 SRS was acquired by DTS Inc.

210
Byrne’s “pre-existing context”
Because the history of multichannel film sound identifies what technological audio formats exist and how they originated, thoughtful creators can now consider them to be pre-existing music playback contexts. David Byrne writes, “I had a slow-dawning insight about creation. That insight is that pre-existing context largely determines what is written, painted, sculpted, sung or performed.” He calls the insight “creation in reverse.” (Byrne: 2012, p.18) In my opinion, sound-field and sound-scape are separate, virtual, pre-existing contexts, capable of enlivening audience engagement with their unique sound. However, to be most effective, they each require specific creative preparation, not adaptation using a static fold-down process.

Murray Schafer: the sound we hear
When describing natural sound, Composer R. Murray Schafer contends that “…outdoor sounds are different than indoor sounds,” and he uses the term “soundscape” to discuss the acoustic environment as the sound we hear. (Schafer: 1977) Byrne also observes that an outdoor stage and indoor concert hall are opposite acoustical environments that can uniquely inform creative decisions. He asserts that the creative path musicians unconsciously take fits into these pre-given listening contexts. Such insights by Byrne and Schafer helped me to understand that the sound-field should be uniquely adapted into sound-scape in order to be heard as a convincing virtual acoustic environ-

Image 1: Performance oriented set up, the bassist is close to the piano keyboard.
ment. I will discuss David Huron’s ITPRA theory in the upcoming section titled Reflective Thoughts.

The Practice: constructing Look Ahead, aesthetic essence and pre-production

Establishing the aesthetic essence of a recording is the foundation of its ensemble sound. The performance aesthetic we established was based on an intimate, improvisational playing style employed by Oscar Peterson, i.e., the bass was positioned close to the piano keyboard. In addition, shorter reflections in the headphone sound-space complimented the spatial aesthetic of our relaxed home studio stage.\(^ {12}\) (See Image 1)

What’s more, the sonic aesthetic was enriched using a diverse combination of esoteric microphones, pre-amps and Universal Audio (U.A.) plugins (tab. 2).

Table 2: Esoteric microphones, pre-amps and plugins provided our sonic aesthetic.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Microphone</th>
<th>Pre-amp</th>
<th>Convertor</th>
<th>Plug-in</th>
<th>Preset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bass</td>
<td>AEA R88</td>
<td>AEA-TRP</td>
<td>Apollo 16</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Bass</td>
<td>AKG-C-391</td>
<td>Manley Vox Box</td>
<td>Apollo 16</td>
<td>U.A. Studer A800 printed</td>
<td>GP9 30ips Noise off</td>
</tr>
<tr>
<td>Bass direct (D.I.)</td>
<td>Underwood bridge pickup</td>
<td>Summit Audio TD100 to a U.A. 6176</td>
<td>Apollo 16</td>
<td>U.A. Studer A800 Manley Massive Passive printed</td>
<td>GP9 30ips Noise off 2K notched out</td>
</tr>
<tr>
<td>Piano “in”</td>
<td>Calrec-SoundField</td>
<td>SPS 422B</td>
<td>Apollo 16</td>
<td>U.A. Studer A800 printed</td>
<td>GP9 30ips Noise off</td>
</tr>
<tr>
<td>Piano “out”</td>
<td>2 U87’s</td>
<td>Langevin Dual Vocal Combo</td>
<td>Apollo 16</td>
<td>U.A. Studer A800 printed</td>
<td>GP9 30ips Noise off</td>
</tr>
</tbody>
</table>

Repertoire and performance

Pianist Robi Botos provided the lead sheets for Praise and Budapest while I contributed lead sheets for A Gentle December Day and Porters Hymn. These original songs shared common musical traits—they were slow in tempo, contemplative and quiet, with dynamics ranging from ppp to mf; features that we believed would enhance the way our sound-source touched the virtual sound-space in stereo and 5.1.

\(^ {12}\) Conversely, Keith Jarret situates the bassist at the far end of the piano.

212
Sound-field leads to Sound-scape-recording the bass and piano

The Look Ahead sound-field begins at the center position of bass, recorded with a carefully centered stereo ribbon microphone, a mono hyper-cardioid condenser and a direct bridge pickup signal (D.I.). The tall stereo ribbon mic was placed in an acoustic shield to reduce rear reflections and focus the intimate stereo L/R spatial sound-field of the bass (image 2).

The piano mic technique utilized two outside Neumann U87 microphones, providing a focused center image that blended with an inside-placed stereo Calrec-SoundField (C-SF) microphone. We preferred the U87s for headphone monitoring because the C-SF was too intimate, but still necessary. (See discussion of piano panning and processing is in the mix techniques section.)

Reverbs

An artificial sound-space was created using multiple instances of stereo and 5.1 reverbs, mixed and panned to overcome dead-spots between the front right to right side (R/S) and front left to left side (L/S) channels. Since there were no drums, the sound-space became an active participant in the ensemble, responding with dynamic spectral excitement. The 5.1 mix used the stereo ‘dry’ ensemble as its sound-source to ignite the stereo and 5.1 reverb chambers, creating a responsive and immersive virtual sound-scape that could be savored in the center listening position.\(^{13}\)

\(^{13}\) Stereo playback in a room is always 3 dimensional due to room reflections, it is never planar. The listening location within a room can be referred to as a Euclidean space, defined by Cartesian coordinates.
Post production and mix template

We tracked into Pro Tools with Universal Audio (U.A.) processing. Studer A800 tape saturation provided warm saturation and this workflow decision afforded us more computer processor power in the mix stage. A film-style mix template that accommodated discrete stereo and 5.1 mixing simultaneously, was ideal for the fold-out. It provided easy sub-grouping and bussing of individual tracks to 5.1 sub-masters and the LFE. It must be understood that a fold-out from a fixed stereo master-print, known as faux 5.1, is not ideal because individual instrument tracks cannot be accessed. The mix template provided instrument sub-masters that allowed us to use gentle compression at several stages, gradually increasing levels while preserving a natural and warm character to the mix. Efficient solo and mute functionality allowed for speedy edit auditioning and easy analysis of blend relationships when joining individual reverbs into one coherent sound-space (image 3).

Routing

The stereo piano and bass were routed to their own instrument sub-masters, which then led to a dry-band stereo sub-master. Three stereo reverbs (U.A. EMT-140/250, Waves R-360 5.1-folded-down to stereo) were bussed to their

---

The Ability to compress, equalise, send or effect tracks or groups and any stage of routing is ideal. Only the bass D.I. (direct-signal) was routed to the LFE to make the fundamental deeper. Low frequency roll-off was applied at the sub-master stage using the Waves 360 sound manger. Roll off began at 80 HZ at 24 db per octave finishing at about 240HZ. The LS, RS were rolled off from at 200 HZ to 80 HZ. This kept the bass localised to the front image.
Paul Novotny: Creating an Immersive Fold-Out – Look Ahead

own stereo reverb sub-master, reducing the stereo mix to two faders, dry band and reverb — the sound-source and sound-space. These were then routed to a full-mix sub-master and lastly to the full-mix-print track. For construction of the 5.1 sound-space, three Waves 5.1 R-360 reverbs were used, each with different reflection and reverb traits, and an additional EMT 250, with a 60-millisecond pre-delay and 2.2 seconds of reverb length was blended into the LS/RS. This accomplished the perception of reverb travel from front to rear in the sound-space. It formed the rear spatial-counterpoint to the original stereo sound-space which remained in the L/R front and added some commonality between the sound-field and sound-scape versions. Together, all these reverbs created a thick and complex sound-space. Global reverb EQ involved rolling off high frequencies and emphasizing the 200 to 800 Hz range of the audio spectrum by +2.5 db. A very small amount of the outside piano microphones and the front bass hyper-cardioid mic were blended into the center channel for sound-field support, creating a pleasing and somewhat immersive L/C/R sound. The inclusion of a discrete center channel served to eliminate the phantom center of a L/R sound-field, providing more accurate positioning in the horizontal panning (image 3).

Mix philosophy

To create a virtual sound-field or sound-scape, a mix engineer must combine analytical listening, personal sonic aesthetics and audio engineering skills. Composer R. Murray Shafer’s thoughts about soundscapes propounded in his book The Tuning of The World. (Schafer: 1977) are both illuminating and revelatory:

“What the soundscape analyst must do first is to discover the significant features of the soundscape, those sounds which are important either because of their individuality, their numerousness or their domination.”

“Outdoor sounds are different than indoor sounds. Even the same sound is modified when it changes spaces. The human voice is always raised outdoors.”

“When there is no sound, hearing is most alert.”

These thoughts can help mixers to understand acoustic environment by providing insight into the extended psychological and cultural meaning of their decisions. For example, a jazz bass drum does not sound like an electronic dance music bass drum. Surround-sound introduces aesthetic and technical considerations that are more complex than stereo. I believe that awareness of Shafer’s soundscape philosophy can assist a mixer to evolve a sound-field into a convincing sound-scape which would be heard as a convincing “virtual acoustic environment.”
Mix techniques: Piano

The piano width was recorded intimately, using an inside-placed Calrec-Soundfield (C-SF) microphone. B-channel signal (WXYZ) was not used but the mic was placed about nine inches above the strings and directly over the Yamaha C-7 logo on the soundboard. The cardioid pickup pattern was set to about 1:30 on the rotary dial of the C-SF SPS422B pre-amp and the width was set to ultra-wide. The two outer Neumann U87s formed more of a mono image. It’s essential for the orientation of the L/R piano image to be heard as if one were seated at the instrument: low notes on the left, high notes on the right. The U87s were predominant in the piano blend while the C-SF image was panned slightly inward, reserving the outer edges of the L/R sound-field for the virtual sound-space.

Bass

Creating a stereo version first on either a 2 channel or a 2.1 bass managed speaker system provides an excellent foundation for fold-out to 5.1. Experience has shown me that fine-alignment of an LFE subwoofer is not an exact science, but if a stereo mix plays with confidence in bass-managed 2.1, then sub-woofer alignment in 5.1 should be close. Unique to the sound-scape, some D.I. was routed directly to the LFE to tighten-up the bass sound. The C-391 provided the focus and the AEA ribbon added spatial texture.

Reference

For a comparative mixing reference, I chose two songs, Daydream and The Oracle, from Kenny Barron and Dave Holland’s duo recording The Art of Conversation because I liked their ensemble sound. Izotope’s Insight plug-in provided analysis of their levels, tone balance and sound-field. Since these examples were mastered at approximately -16 LUFS, the average stereo mix level of Look Ahead was reduced by about four decibels (db) to provide extra dynamic range. This four db was carefully regained during mastering.

Automation

Instead of fader automation, clip-gain was the primary edit tool for blending. This technique sent consistent levels to compressor thresholds, keeping the mix dynamically open and even.

---

15 Barron, Kenny, Holland, Dave, The Art of Conversation, Izotope analysis: Daydream, -19.3 LUFS Left Peak -1.0, Right Peak -0.1 | The Oracle -15.9 LUFS—Left Peak-0.3, Right Peak -0.2
Reverb sends were routed as pre-fader, so the ratio of dry-to-wet could be easily auditioned and automated to change the balance between wet and dry. The sound of stereo reverb panned across four or five speakers is not pleasing because it is unidimensional. The Waves R-360 5.1 reverb solved this by joining the front-stereo and rear-stereo reverbs together with diverse multi-channel reverb signals, eliminating the “reverb dead-zone” on each side between the left/right front and LS/RS rear speakers. This presented a multidimensional sound-space.

Roll off

The stereo version of Look Ahead was mixed first and then the 5.1 version was created, maintaining a universal balance of sound-source and sound-space. The sound-space employed an LFE (low frequency effects) LPF (low pass filter) roll-off of 24 db per octave started at about 80 HZ and the LS/RS channels began their HPF (high pass filter) roll off at about 200 Hz, ending at 80 HZ. This crossover was done with the Waves 360 manager, and helped to keep the staging focused on the front L/C/R audio image.

Unique mix opportunities in 5.1

The 5.1 sound-scape presented several other unique opportunities such as custom re-balancing of the sound-space with the sound-source at opportunistic moments. For example, on Now’s the Time, at approximately forty-one seconds, a hand-on-piano string muting technique with careful pedaling was used to produce a percussive and sustaining string effect. This pleasing sound benefited from an increased level of sound-space that created an immersive, mild, slow chorusing effect. Example two (see appendix) focuses on this section and peels back the layers of sound-source, sound-space, sound-field and sound-scape in a side-by-side comparison.

The 5.1 sound-space was exaggerated on the endings of Praise and A Nightingale Sang in Berkeley Square — the chamber level rises by about twenty percent, creating a haunting goodbye as the dry-band fades into the distant and immersive reverb chamber.

Arco bass solo moves from a public to an intimate space

Also unique to the sound-scape, the introduction on Porters Hymn includes an arco bass solo, set far back on the left side of the virtual stage. When the pizzicato melody enters, the perspective returns to the front center, illustrating Moore’s “soundbox model” (Moore: 1992) by first presenting the bass in a distant public space and then quickly moving it to an intimate position.
Mastering

Mastering brought the LUFS level up to about -16.0 with peaks limited to -0.7. U.A. mastering plugins such as the Manley Massive Passive, Ampex ATR-102, Millenia NSEQ-2, and the Precision Limiter improved the tonal focus, adding saturation and peak limiting while extending the high and low end spectral range.

Reflective thoughts lead to ITPRA theory

For fifteen years, I have folded-down 5.1 to stereo in broadcast post-production, but this is the first time that I folded-out. I now see that it allows unique decisions to be made that can enliven emotional engagement in the sound-scape, while maintaining the ensemble from the pre-existing sound-field. However, I think that it’s still important for surround-sound music creators to observe how film/broadcast and media-arts will stage their sound in future multichannel formats, notably how NHK’s upcoming convention—8K Super Hi-Vision with 22.2 audio—will provide home-theatre audiences with the ability to select language, adjust dialogue volume, and hear audio-objects in immersive 3D.\(^\text{16}\) New technology brings challenge and opportunity by affording multichannel music creators more options for playback, at the same time making it important to respect the disciplined efforts of those who follow the pedagogy of composition, orchestration and arranging to pursue cohesive music ensemble. Conversely, 22.2 multichannel home theatre audio may also inspire new possibilities for compositions and audio productions that aim for hyper-realistic sound-scapes.\(^\text{17}\)

Starting with Fantasia in 1940, film audio history shows an evolution of technology and sound that tried to further stimulate human emotional expectation and experience with immersive sound-scape. Walt Disney wanted to choreograph the musical-sound of a bumblebee twirling around his audience to heighten their emotional experience. Maestro Leopold Stokowski suggested that Disney engineers speak with Bell Labs about their research in multichannel audio, leading to Fantasound. Soon after that, in the ’50s, emotional experience induced by music was explored in Leonard Meyer’s seminal book, *Emotion and Meaning in Music*, asserting that music communicates emotion by “choreographing expectation” (Meyer: 1956).

In 2006, David Huron’s book, *Sweet Anticipation*, identified the ITPRA theory, a psychological theory of expectation (Huron: 2006). ITPRA aims to better understand musical tension and release by identifying five internal

\(^{16}\) 8K, Super Hi Vision, involves a 22.2ch. sound microphone, portable recording and editing device, 3D audio mixing system, mixer for live feeds, 3D reverberator, 22.2ch. headphones, processor etc.

\(^{17}\) Lopez, Francisco, Electro Acoustic artist whose music style is based on immersive sound-scape, inspired by R. Murray Schafer’s book, *The Tuning of The World*, https://open.spotify.com/artist/2NIo9CS1lxz2pArKee4hes
emotion-response systems: 1) *Imagination*, 2) *Tension*, 3) *Prediction*, 4) *Reaction* and 5) *Appraisal*. ITPRA can be grouped into two categories of *pre-outcome tension* and *post-outcome release*. The first three *pre-outcome* states (ITP) describe how anticipation builds toward a stressful level of tension, and the last two *post-outcome* states (RA) describe the release of that tension. Decisions about meter, syncopation, tonality, spatiality and cadence leverage the ITPRA theory of expectation to elicit an arc of emotional responses from the audience. Expanding upon Meyers observations, I now view creators of *sound-field* and *sound-scape* as “choreographers of emotion and expectation.” Their individual sense of ITPRA guides the organization of sound, which can be expressed anywhere between the virtual contexts of a small room, to a natural outdoor setting or, even further outward, to an interstellar space battle portrayed with cinematic exuberance.\(^{18}\) I suspect that the distinctive traits of immersive surround-sound could even help canny record producers of concert music to find their particular listening audience, without allowing the technology to deconstruct the ensemble or change its focus on the performance and composition.

**Performance touches the sound-space**

*Look Ahead* arrives at its performative destinations via the virtual acoustic environment of an indoor concert chamber with staging that supports intimate music. Schafer’s statements: “What the soundscape analyst must do first is to discover the significant features of the soundscape” and “When there is no sound, hearing is most alert” proved to be foundational in creation of that virtual acoustic environment. The dynamic of this recording resides predominantly between *pianissimo* and *mezzo-forte* and rarely goes beyond *forte*, this trait invites the condition of alert hearing that Schafer speaks of. A live-performance dynamic is foundational to the mix because it excites the *sound-space*. A multi-speaker listening environment can support quiet dynamics very well due to the diffused or “spread-out” channel energy that naturally occurs from a multi-speaker installation. In *Look Ahead*, surround-sound never draws attention to itself by panning an instrument to an unexpected place. This encourages the listener to suspend their believe in the virtual *sound-scape*. Trusting that the listener would become solidly immersed in this virtual sound, I increased *sound-space* level by about 2-db for *Praise* and *Budapest*, because the slow tempo—duration between the beats—allowed the chamber to ring clearly without “smearing.” These specific adaptations also illustrate Byrne’s theory of creation in reverse as they

\(^{18}\) Bhatia, Amin, *The Interstellar Suite*, Produced by Frank Marrone, Triplet Records, 2015, TR-10018
did not occur to me while working in the *sound-field*, but the *sound-scape* called out for them.

**Supporting examples and data analysis**

**Example #1:** A side-by-side comparison of a reverse engineered fold-down from the 5.1 master and the prepared stereo 2.0 master. It illustrates how the static convergence of LS/RS results in an inconsistent *sound-space* that portrays undesirable and varying levels of intimacy. The bass solo of *My Favourite Things* was selected to show these undesirable results. (see appendix)

**Example #2:** Reveals the audio-channel layers from stereo *sound-source* and *sound-space* to *sound-field* and *sound-scape* in a side-by-side playback of the piano hand-mute section of *Now’s the Time*. (see appendix)

**Data analysis**

While both prepared masters of *My Favourite Things* sound entirely different, they share common cohesive ensemble and loudness data. Image 4 provides a comparison of the LUFS and peaks between the fold-down to 2.0 and the prepared 2.0. The numbers are remarkably close, and this is representative of the entire album. However, the masters sound very different, proving that one must always listen and not just look at meters.

![Image 4: My Favourite Things, side-by-side analysis of levels. Left: Prepared stereo 2.0 & Right: 5.1-Fold-down to stereo 2.0.](image-url)
The prepared stereo analysis is displayed on the left instance of Izotope Insight and the fold-down and its down-mixer settings are on the right (image 4).

**Conclusions from research and practice**

The practice and its four selections of research: 1) Surround-sound history; 2) David Byrne’s theory of “creation in reverse”; 3) R. Murray’s Schafer’s theories of soundscape analysis; and 4) David Huron’s general theory of human expectation—ITPRA, distill into seven foundational guiding insights that may assist the creator when adapting sound-field to a sound-scape.

1. Since 1881, multichannel audio has been heightening audience experience and it continues to do so, suggesting that immersive audio should not be overlooked by music creators. It’s possible that music sound-scape could become a more significant music distribution format because it seeks an expert listener who desires a premium product.

2. Creative decisions about how to utilize multichannel contexts are informed by the available technology which is evolving constantly.

3. These pre-existing contexts determine what the reception of immersive experience can be. They range from a re-creation of a natural acoustic environment to a hyper-real sound-scape.

4. Acoustic environments—soundscape—influence and inform the creation of virtual sound-fields and sound-scapes.

5. Composers and creators of immersive sound can be considered “choreographers of emotion and expectation.”

6. Even though data analysis of loudness and peaks can show very little variance between a fold-down to 2.0 stereo and the prepared 2.0 stereo version, the virtual acoustic environments of sound-field and sound-scape can sound very different, resulting unique listening experiences and emotional outcomes.

7. Adapting a stereo sound-field 2.0 mix from its mix template into a sound-scape is a preferred method for two reasons: 1) It results in two individually prepared masters that share cohesive ensemble in their unique virtual sound-spaces; and 2) a fold-down can result in unpredictable changes to depth perspective when listening for the sounds of intimacy, personal, quasi-public, and public spatiality in the down-mixed sound-field.
**Closing thoughts**

My first surround-sound playback experience taught me not to confuse the listener by pulling an ensemble apart. A creator needs to understand the aesthetic essence of their music and then integrate it with the pre-existing technological formats accordingly. Once the playback formats are chosen, accepted engineering practice indicates that the ideal way to organize the sound is with a ‘film-style’ mixing template because it provides accessibility to individual instrument tracks that can be routed (bussed) to stereo and multichannel sub-masters.

The essential take away is that a fold-out from a sound-field can provide a cohesive foundation that confidently leads to a soundscape. The result is two diverse masters that share a common and familiar ensemble sound. For the listener of both, these commonalities and differences between sound-field and soundscape may even help to keep the music fresh regardless of how many times it’s listened to.

I hope that this case study inspires practitioners of the recording arts to experiment with immersive music production while maintaining ensemble sound. Hopefully, students and professionals will consider these insights when they adapt their own music from stereo sound-field to immersive soundscape.

**Acknowledgments**

Thanks to Canada Council for the Arts, Robi Botos, Amin Bhatia, Frank Morrone, David Greene, Jeff Wolpert, Tom Reynolds, Harry Lewis, Michael Francis Nunan, Hayward Parrott, Chad Ischick, Doug Wilde, Matt Zimbell, Peter Cook, Steve Bellamy, Bruce Cassidy, Michael Colgrass, Barry Elmes, Al Henderson, Rob van der Bliek, Jay Rhan, Michael Coghlan, Dr. Norman Ludwin, Keith Loach, George Novotny, Adrian Novotny and Dave Dysart for the many thoughtful discussions regarding practice, study and analysis of music composition, recording, engineering, mixing, film, broadcast and audio post. Special thanks to Universal Audio, HHB Canada, Yamaha, Humber College and York University for their continued support. I acknowledge the support of the Canada Council for the Arts, which last year invested $153 million to bring the arts to Canadians throughout the country.

**References**

**Bibliography**


www.jstor.org/stable/10.2979/filmhistory.28.1.06

Discography
Novotny, Paul, Botos, Robi, Look Ahead, Triplet Records, 10019, 2016

Websites
http://techtland.time.com/2013/06/06/dolby-atmos/-Accessed 11 13 17
http://www.nhk.or.jp/8k/index_e.html-Accessed 11-8-17
http://www.telephonecollecting.org/Bobs%20phones/Pages/Essays/Ader/Ader.htm 04-25-18
Terminology

Sound-field: A stereophonic recording played back presents an *acoustical environment* with the sound coming from front L/R speakers.

Sound-scape: An immersive multichannel recording played back presents an *acoustical environment* with the sound coming from front L/C/R speakers and rear side, LS/RS with an LFE (low frequency effects) channel.

Soundscape: R. Murray Schafer’s term for *acoustical environment*. David Byrne uses the term *acoustic context*, and I interpret that he means *acoustical environment*.


Sound-source: The dry audio source containing no virtual reflection, echo or reverb.

LUFS: Loudness Units Full Scale.

Appendix

Analysis example #1: Side-by-side fold-down analysis–5.1 to 2.0 compared to the stereo master, illustrating why fold-out is necessary, Bass solo from My Favourite Things, https://www.youtube.com/watch?v=fGZu2w1gZIQ

Analysis example #2: Unpacking the Sound-scape 2.0 Sound-field to 5.1 Sound-scape, Now’s the Time, https://www.youtube.com/watch?v=93tagVFvUvQ&t=15s

Look Ahead in 24/96 lossless Flac, 5.1 sound-scape and stereo 2.0 sound-field high-resaudio.com/en/album/view/dpdwkx/paul-novotny-robi-botos-look-ahead

Look Ahead in 2.0 stereo, 24/44.1 Mastered for iTunes, MFIhttps://itunes.apple.com/ca/album/look-ahead/1113854143
Daniel Pratt & Andrew Bourbon: Life In Between
Phase: Understanding and Manipulating Microphone
Relationships With Visualisation Tools

Abstract

Phase relationships are a complex and mystifying phenomenon for early stage recording engineers and university students. In this paper, we take the analytical capabilities of time shifting plugin Auto-Align and use it to develop new methods of understanding phase interaction. We utilise the visualisations and time shifting features to assist in recording a multi-miked drum kit and in the post-production soundstage for a thirty-two piece big band. We explore new methodologies for phase interaction and microphone manipulation by running these two experiments and documenting the process using a combination of text, audio, and video.

Introduction

In this paper, we examine the phase interaction of multi-microphone recording and mixing with the intent to develop an in depth understanding of relationships ‘in between phase’ to produce better recordings and mixes. In our curriculum design in both The Queensland University of Technology (QUT) and The University of West London (UWL) we discuss phase in technical descriptions relating to the acoustic and electronic summation of multiple sources. Phase relationships are also explained in practical recording workshops as well as theoretical lectures. One of the conceptual challenges that our students face is hearing phase variance and implementing the appropriate action to remedy what they hear. For educators, explaining phase becomes problematic when that variance represents a shift that cannot be solved via a simple binary polarity reversal.

We explore the use of metering and phase manipulation in the recording and mixing of audio. On the recording side, we investigate this phenomena through the creation of an educational drum recording video. The video examines the capture of phase interaction information, the analysis, and correction of resulting issues. The data analysis of the phase relationships informs
the physical movement of microphones. We demonstrate real time phase interaction measurement using the innovative metering in the Sound Radix Auto-Align plugin. This information aids us in demystifying phase interaction between microphones and enables us to develop new methods for both microphone placement and the teaching of multi-microphone recording.

We then manipulate the complex phase relationships in a Big Band recording featuring thirty two microphones across twenty sources using a range of mono and stereo techniques. We use metering and analysis techniques to inform our alignment and manipulation of this pre-recorded work. In both of these case studies the sonic impact of the phase analysis and manipulation are presented as video and audio examples. The data will be used as a pedagogical tool for the demystification of phase in the teaching environment. Throughout this process, we are guided by the following questions, how can we use a tool dedicated to alignment to improve the understanding of phenomena like phase relationships? Can educators use this tool as a teaching device to accelerate the education of young engineers in the understanding of complex phase issues that take years of practice to fully understand?

**Simplifying Phase Interaction**

Phase interplay is a phenomenon whereby frequencies of multiple waveforms combine with one another and either cancel or amplify the sound source. Explanations of phase interaction often arrive with complex equations and scientific terminology, this initial overview is a simplification of these explanations with practical uses applied to microphones and speakers. To better recognise how a phase relationship occurs we must first understand how microphones and speakers work. “When we record sound, the diaphragms in our microphones essentially replicate the action of our eardrums, vibrating in accordance with those [sound pressure] waves” (Keller, 2011). Consider that both microphones and speakers are transducers that perform mirror images of the same process. For example microphones transform acoustical energy to electrical energy and speakers transform electrical energy into acoustical energy (Sigismondi, et al, 2014). As a microphone diaphragm moves backwards and forwards it transforms acoustic energy into positive and negative electrical current over a period of time. A speaker conversely takes this positive and negative electrical energy and transforms it into forwards and backwards motion which creates the acoustic energy that we hear as sound. When we use more than one microphone, we introduce the possibility of interference. Corbett (2015) explains these two typologies of interference as constructive and destructive. Constructive interference occurs when two microphone diaphragms are moving in the same direction; destructive interference happens when the microphone diaphragms are in con-
trary motion (ibid). When a sound is captured through multiple microphones and converted into electrical energy, constructive interference encourages free speaker movement. Destructive interference restricts the movement of the speaker which cannot move forwards and backwards at the same time. This destructive interference is often referred to as phase cancellation. Cancellation becomes more problematic as you add more microphones to any single source such as a drum kit or guitar amplifier. This is why “The principle idea is to get all the microphones working together constructively” (Weiss, 2014, para, 4). Clinch (2011) states that "inexperienced engineers often complain about thin sounds when mixing multiple microphones together even though each individual microphone may sound great" (para, 2). Understanding the relationships between microphones is crucial because "in most recording sessions, we’re dealing with multiple instruments and multiple microphones" (Keller, 2011, para, 13). These cancellations are neither good nor bad but a choice. Senior (2008) argues for moving microphones not only to remove destructive phase issues but to use minor phase cancellation as a creative choice.

The Problem With a Binary Understanding of phase

One of the problems associated with explanations of phase interaction is the binary nature of the equipment, which offers only mirror images of phase relationships. A polarity button on a console provides the ability to reverse the phase relationship with an in or out push button. Coppinger (2012) refers to this as a polarity reversal. An experienced engineer is skilled at listening for the sonic change that occurs between multiple microphones when they invert the polarity of one microphone (Paterson, 2007). However, due to its on or off nature, the polarity button establishes a conversation that only offers a binary understanding of phase relationships. For example, a mirroring phase relationship of one hundred and eighty degrees does not account for a floor tom on a drum kit, which is often ninety degrees out of phase with the
overhead microphones. In this case, a phase reversal will bring a floor tom to a two hundred and seventy degree relationship, which is the same level of audible phase cancellation. Corbett (2015) argues for an understanding of in-between phase relationships, which recognise phase as any relationship within three hundred and sixty degrees of a wave cycle. If we think of phase as merely binary of in or out, then we do not understand the nature of the relationship and we miss the full potential for the creative manipulation of microphone relationships. As a result, a phase relationship is neither constructive nor destructive; it is a tonal colour that encompasses three hundred and sixty degrees of frequency dependant manipulation. Microphone movement is the common methodology for manipulating tonal colour with more creative intent than just a binary response to phase relationships. However, a multi-microphone setup will inevitably have varying degrees of phase cancellation that Paterson (2007) suggests fill the role of creative tonal colours. Paterson (2007) states that “there is no known way of alleviating this, and indeed it has become an accepted part of the sound engineer’s art to accept this and indeed harness it to creative effect” (para, 4).

Three Arguments for Pedagogical Approaches

There is an agreement in the literature that phase cancellation leads to an undesirable sound when using multiple microphones (Sigisomondi et al., 2014, Paterson, 2007, Senior, 2008, Corbett, 2015, Savage, 2011). However, at QUT for example, there are only twenty-four two-hour tutorials per week allocated for teaching students how to hear and understand all practical recording and mixing concepts for the entire year. The first year is concerned with establishing a baseline of rudimentary theoretical knowledge, the second and third year move into more advanced audio engineering. Hearing microphone relationships are only a small part of the full curriculum designed to develop well-rounded, production-capable music students. As a result we have limited time to embed "a relevant object of auditory knowledge [emerging] through interplay between a domain of targeted listening and a set of discursive practices played out in the context of specific sound-engineering activities" (Porcello, 2004, p.734). Condensing years tacit phase relationship understanding must address three main issues.

1. The explanations for detecting phase need to be less vague and binary in their delivery.
2. The assumption that early stage recording engineers can hear phase problems is flawed.
3. If the suggested correction is to move a microphone we must find a method of indicating which direction or how far.
Firstly, there are many variables that create destructive phase relationships such as microphone placement, room reflection, speaker placement, and equalisation. Without access to visualisations, it is difficult to explain what phase sounds like to an untrained student. Explanations from the literature include “typically a thin-sounding signal with little or no bass sound” (Keller, 2011), or “as a hollow sound in which certain frequencies, or tones, appear to be missing” (Lashua, Thompson, 2016, p.82). or “a hollow, filtered tone quality” (Bartlett, 2017, p.15). Explanations like these make sense to experienced recording engineers but are not useful to an engineer that doesn’t have the experience to understand what descriptors like hollow or thin mean in the context of a multi-microphone drum setup.

Secondly, the assumption that a student can hear phase relationship problems in a stereo setup is flawed. The understanding of relationships between phase and frequency in an audio recording environment is complex and time consuming (Paterson, 2007). Thirdly, moving a microphone is not a helpful suggestion to an inexperienced engineer, especially when there is no indication of which direction or how far. Weiss (2014) indicates that engineers are "simply going to have to move the microphone to different proximities and listen for what sounds best" (para, 8). Sigismondi et al. (2014) suggest that engineers "place the microphone at various distances and positions until you find a spot where you hear from the studio monitors the desired tonal balance and the desired amount of room acoustics" (p.5). Paterson (2007) suggests that microphone placement is “the art of the sound engineer, who will make minuscule adjustments to the positioning of microphones in a session, evaluating the monitor mix to choose final placements”. These suggestions are helpful, but ultimately take years of practice to hear and correct. Adding to this, different shaped rooms introduce different phase relationships so microphone placements do not translate between spaces. This potentially leads inexperienced engineers on a guessing game based on an assumption that they can hear the differences between various microphone movements. This confusion is compounded by the overwhelming amount of options when we move a microphone as stated by Senior.

Tweaking the distance between [microphones] subtly shifts the frequencies at which the comb-filtering occurs. Inverting the polarity of one of the mics yields another whole set of timbres, switching the frequencies at which the sine-wave components in the two mic signals cancel and reinforce, so the potential for tonal adjustment via multi-miking is enormous (Senior, 2008, Para, 21).

On top of this potentially overwhelming issue, there is also a personnel concern. Microphone movements for drum recording require a producer to listen, a drummer to hit drums, and an engineer to move microphones.
Finally, relying on microphone movement alone does not consider that there are other ways to manipulate phase relationships like equalisation. As stated by Savage:

“applying EQ will alter the phase relationship of the sound that is being processed. This is because there is a certain amount of time required for the EQ to process the frequencies that it is acting on, and so those frequencies get shifted in their time relationships to other frequencies that make up the sound. This time shift creates changes in the phase relationship” (Savage, 2011 p.50).

*Live Sound Design Approaches to Phase Relationships: Analyse First then Listen*

One subject area outside the studio environment where an understanding of phase is essential is live sound system design. In live sound design, system alignment takes place through transfer function measurement, and is never trusted to the human ear alone. Though they are aligned using technical measuring tools, some aesthetic choices around the performance of the system are the result of either matching a target response curve or the choice of the experienced engineer. One of the key aims is to create a uniformity of performance throughout the space, with a minimised variance in frequency response and intensity through strategic placement and alignment of speakers. Destructive phase relationships cause ripples in the frequency response of the speakers so a system that is improperly aligned is challenging to manipulate by ear.

In the simplest of systems a single source is able to provide even coverage, however in practice the characteristics of a space normally make it impossible to achieve spatial uniformity without the addition of extra sources such as speakers for reinforcement of the system. Contemporary approaches to live sound design see line array speaker systems employed to increase the efficiency of sound distribution, utilising multiple aligned speakers to cover difficult to reach areas. Individual speakers in the main speaker array, and any auxiliary speaker arrays are designed to focus on isolated target zones. The focus of live sound system design is to avoid multiple sources hitting the same target zone. However, it is inevitable that there will be points in the space where there is interaction between sources leading to acoustic summing of signals much like overlapping regions in an audio crossover. It is these interaction zones where designers focus their attention in order to optimise the phase coherence of the summed response of two systems interacting at the same target zone. This means engineers must time and phase align speakers in the physical space which, as stated earlier, is achieved primarily through technical measurement and refined by ear. It is also important to clarify that the signals coming from the sources will be almost identical subject to speaker voicing, unlike sound received at
microphones at different positions in a space as found in the recording examples.

A summation zone is any area where two independent speakers outputting the same signal are acoustically summed in physical space. This is the exact reverse of multiple microphone signals summing down into speakers. To create a phase coherent summation zone there are two key factors to take into account: relative level and phase (McCarthy, 2016). The impact of two sounds being acoustically summed is a resultant change in overall level, with potential summed output levels of between +6dB and -60dB, depending on the relative level and phase of the two sources. First consider the impact of relative level, with 2 sources arriving at a single receiver, but each at a different amplitude. If the level offset between the two sources at the point of reception is greater than 10dB, then the maximum ripple (change in frequency response) is limited to +3dB. A ripple of +3dB which is considered an acceptable result in PA system design for large venues. Designers strive to minimize the number of summation zones where multiple sources are within 10dB in level offset to maintain the best possible frequency domain behaviour.

The second point of consideration in exploring a summation zone is that of relative phase between sources. In a line source all the elements in the array are propagating with equal phase and amplitude, though in practice amplitude is manipulated to provide even coverage on complex audience planes where path length from source to the plane varies significantly. When a second source combines with the primary array, the relative phase has a significant impact on the summed level at the receiver. If the two sources are within 120 degrees of phase rotation, the maximum ripple at the receive is again +3dB, with a potential ripple of +30dB should the sources arrive with more than 120 degrees of phase difference. System designers are therefore looking to minimise the potential ripple by creating systems that interact within 120 degrees of phase rotation, and have a minimum offset in level of 10dB. It is important to note that, regardless of level, offset with less than 120 degrees of phase rotation, the impact on frequency response is less than +3dB, but with a change in that response throughout the phase rotation range. This means that any variance in phase up to 120 degrees is challenging to hear.

In the recording environment, this study of live acoustic summation provides the student with an appreciation of the impact of relative amplitude and phase that does not occur in traditional studio-based music technology education. When placing a room mic and an overhead in the same room it is likely that the signals arriving into the recording device and then reproduced in the control room will be within 10dB in relative level. This level difference offers the potential for broad frequency response ripple. By moving the microphone, the level difference will not be significantly manipulated, leaving little change in frequency response due to amplitude.
For example, a phase relationship change will only be significant when moving a room microphone to within 120 degrees of an overhead microphone. This movement will result in the combined microphones presenting a more balanced frequency response. It is important to note that unlike our theoretical approach to PA system design, the signal transduced by the microphones operates differently. The summation theory discussed here relies on the same signal being produced by both sources and being received at the point of summation. In the studio, indirect reflected sound is used creatively to give an impression of size and space and means that the single source becomes a multitude of sources as it is re-radiated through the boundaries in the recording room. Regardless of this difference, the phase response is manipulated through microphone movement, resulting in different frequency responses and therefore different musical impact subject to placement. Should there be negative tonal and spatial impact from using multiple microphones on a single source the engineer has two choices to improve or change the resultant response, the first being to manipulate the relative level of the microphones and the second to manipulate the relative phase. One of the challenges we face with pedagogical practice is assisting students in distinguishing the sonic impact of relative phase. With significant frequency response ripple and destructive changes in the summed level of the combined sounds, hearing the destructive impact of interactions beyond 120 degrees is unambiguous to the listener. However, Learning to hear the interactions in the first 120 degrees of rotation takes time and practice. We propose that visual aids assist students with the development of these skills by identifying phase relationships that they are not yet understanding and allowing students to practice listening to more subtle phase shift in recorded text.

**Live Sound Design Approaches to Phase Relationships: Analyse First then Listen**

One subject area outside the studio environment where an understanding of phase is essential is live sound system design. In live sound design, system alignment takes place through transfer function measurement, and is never trusted to the human ear alone. Though they are aligned using technical measuring tools, some aesthetic choices around the performance of the system are the result of either matching a target response curve or the choice of the experienced engineer. One of the key aims is to create a uniformity of performance throughout the space, with a minimised variance in frequency response and intensity through strategic placement and alignment of speakers. Destructive phase relationships cause ripples in the frequency response of the speakers so a system that is improperly aligned is challenging to manipulate by ear.
In the simplest of systems a single source is able to provide even coverage, however in practice the characteristics of a space normally make it impossible to achieve spatial uniformity without the addition of extra sources such as speakers for reinforcement of the system. Contemporary approaches to live sound design see line array speaker systems employed to increase the efficiency of sound distribution, utilising multiple aligned speakers to cover difficult to reach areas. Individual speakers in the main speaker array, and any auxiliary speaker arrays are designed to focus on isolated target zones. The focus of live sound system design is to avoid multiple sources hitting the same target zone. However, it is inevitable that there will be points in the space where there is interaction between sources leading to acoustic summing of signals much like overlapping regions in an audio crossover. It is these interaction zones where designers focus their attention in order to optimise the phase coherence of the summed response of two systems interacting at the same target zone. This means engineers must time and phase align speakers in the physical space which, as stated earlier, is achieved primarily through technical measurement and refined by ear. It is also important to clarify that the signals coming from the sources will be almost identical subject to speaker voicing, unlike sound received at microphones at different positions in a space as found in the recording examples.

A summation zone is any area where two independent speakers outputting the same signal are acoustically summed in physical space. This is the exact reverse of multiple microphone signals summing down into speakers. To create a phase coherent summation zone there are two key factors to take into account: relative level and phase (McCarthy, 2016). The impact of two sounds being acoustically summed is a resultant change in overall level, with potential summed output levels of between +6dB and -60dB, depending on the relative level and phase of the two sources. First consider the impact of relative level, with 2 sources arriving at a single receiver, but each at a different amplitude. If the level offset between the two sources at the point of reception is greater than 10dB, then the maximum ripple (change in frequency response) is limited to +3dB. A ripple of +3dB which is considered an acceptable result in PA system design for large venues. Designers strive to minimize the number of summation zones where multiple sources are within 10dB in level offset to maintain the best possible frequency domain behaviour.

The second point of consideration in exploring a summation zone is that of relative phase between sources. In a line source all the elements in the array are propagating with equal phase and amplitude, though in practice amplitude is manipulated to provide even coverage on complex audience planes where path length from source to the plane varies significantly. When a second source combines with the primary array, the relative phase has a significant impact on the summed level at the receiver. If the two sources are
within 120 degrees of phase rotation, the maximum ripple at the receive is again +3dB, with a potential ripple of +30dB should the sources arrive with more than 120 degrees of phase difference. System designers are therefore looking to minimise the potential ripple by creating systems that interact within 120 degrees of phase rotation, and have a minimum offset in level of 10dB. It is important to note that, regardless of level, offset with less than 120 degrees of phase rotation, the impact on frequency response is less than +3dB, but with a change in that response throughout the phase rotation range. This means that any variance in phase up to 120 degrees is challenging to hear.

In the recording environment, this study of live acoustic summation provides the student with an appreciation of the impact of relative amplitude and phase that does not occur in traditional studio-based music technology education. When placing a room mic and an overhead in the same room it is likely that the signals arriving into the recording device and then reproduced in the control room will be within 10dB in relative level. This level difference offers the potential for broad frequency response ripple. By moving the microphone, the level difference will not be significantly manipulated, leaving little change in frequency response due to amplitude. For example, a phase relationship change will only be significant when moving a room microphone to within 120 degrees of an overhead microphone. This movement will result in the combined microphones presenting a more balanced frequency response. It is important to note that unlike our theoretical approach to PA system design, the signal transduced by the microphones operates differently. The summation theory discussed here relies on the same signal being produced by both sources and being received at the point of summation. In the studio, indirect reflected sound is used creatively to give an impression of size and space and means that the single source becomes a multitude of sources as it is re-radiated through the boundaries in the recording room. Regardless of this difference, the phase response is manipulated through microphone movement, resulting in different frequency responses and therefore different musical impact subject to placement. Should there be negative tonal and spatial impact from using multiple microphones on a single source the engineer has two choices to improve or change the resultant response, the first being to manipulate the relative level of the microphones and the second to manipulate the relative phase. One of the challenges we face with pedagogical practice is assisting students in distinguishing the sonic impact of relative phase. With significant frequency response ripple and destructive changes in the summed level of the combined sounds, hearing the destructive impact of interactions beyond 120 degrees is unambiguous to the listener. However, Learning to hear the interactions in the first 120 degrees of rotation takes time and practice. We propose that visual aids assist students with the development of these skills by identifying phase relationships that they are not yet understanding and
allowing students to practice listening to more subtle phase shift in recorded text.

**Understanding Auto-Align**

In our pedagogical practice at QUT and UWL we regularly address student concerns based around phase. Our students often worry that they can’t hear phase interactions that are obvious to the lecturers who have more experience in hearing these relationships. In order to develop a better understanding of phase for these students to improve their recordings we propose that we utilize tools that perform visual analysis to assist in their understanding of phase relationships in both the recording and mixing environment. Auto-Align is an automatic time aligning plugin developed for use in digital audio workstations. Its primary function is to detect and time align multiple waveforms to reduce phenomena like destructive phase cancellation and comb filtering while improving the dynamic intensity of multi-microphone recordings. Auto-Align is commonly used to correct microphone placement issues on drum overheads, multi-miked guitar amps, or to align a bass guitar recording that consists of a direct signal and a miked bass amp (SoundRadix, 2017).

Time aligning waveforms is a conventional technique that mix engineers use to correct problematic microphone placements. Savage (2011) recommends that mix engineers experiment with minuscule shifts in audio files known as nudging to create improved phase relationships between multiple microphones capturing the same source. Keller (2011) states that "You’d be amazed what a difference just moving a track by one or two milliseconds can make"(para, 20). Auto-Align is one of a few new audio plugins that take the guesswork out of shifting audio to improve phase relationships. It achieves this by employing a detection algorithm that listens to the audio between multiple microphone recordings and then offers the engineer several selections for phase coherent wave positioning. This type of detection allows mix engineers to select from a reduced choice of phase-coherent positions, changing the task of nudging audio and listening into an efficient automated process.

Auto-Align also has some additional features for the analysis and calculation of phase relationships and microphone placement. These analysis tools consist of a circular phase analyser, a delay relationship display, and a distance evaluation that offers estimated microphone distances in both centimetres and inches. It is this expansive suite of measurement tools that enables recording engineers to understand their microphone placement and phase relationships with an accuracy that redefines phase and microphone analysis. This new knowledge of phase relationships leaves us with a choice of continuing with established methods or using a new understanding to challenge established norms surrounding the recording process. A new
measurement device calls traditional methods of recording into question and requires investigation to discover if these established practices can be redeveloped using newly refined methods (Bacon, 2012).

**New Visualisations Paint a Detailed Picture**

In the digital plugin market, there are several options for visualising phase interaction between microphones. Izotope Insight provides a detailed analysis of spaciality, phase interactions, and loudness information. There are also some stereo equalisation plugins that come with small phase scopes. By and large, they all use a similar two-dimensional approach to phase metering as the indicator moves to +1 the signal is more phase coherent.

Our investigation of Auto-Align reveals a paradigm shift in the visualisation of phase relationships, using a three hundred and sixty degree scope that establishes a detailed analysis of the relationships between microphones. Pointing north indicates the most constructive relationship, but the scope allows for detailed analysis of every single phase rotation. This level of choice means that an engineer can decide exactly how in or out of phase they want their microphones with high phase coherence offering clarity and punch and less phase coherence offering more depth and space. This new approach to phase metering also utilises colour information to indicate which frequencies ranges are causing constructive or destructive relationships. The more detailed approach to metering allows us to understand and manipulate microphone relationships to a far higher degree than previous analytical tools.

In addition to the phase meter Auto-Align provides distance calculator that gives information on distances between microphones. Using a combination of both the phase and distance readings it is possible to virtually move microphones in the DAW to test phase relationships before venturing into the studio. Theoretically, this means we can know which direction
and how far to move the microphone, removing the usual guesswork associated with the practice.

Finally, Auto-Align uses a detection algorithm that suggests multiple in phase measurements for microphone placements. These measurement points are displayed on a delay meter which gives users a choice of phase coherent microphone placements if they are inclined to experiment with different options and different levels of phase coherence to create colouration (Paterson, 2007). Having different choices offers engineers the ability to make informed choices on how much constructive or destructive interaction they want. The choice is an important factor depending on whether an engineer wants a sound that is “diffuse or blended, instead of sharply focused” (Bartlett, 2017, p.114). This ability to make informed choices between out of phase and in phase microphones allows the recording or mix engineer to design their desired spaciality around multi-microphone recordings by focussing on which instruments are more or less in phase.

Such a detailed suite of analytical tools creates several options to enhance the creativity of the recording process without negatively affecting the phase relationships of the microphones. For example, an engineer can choose the best sounding spot for room microphones and not have to measure them and sacrifice the sound of one microphone to ensure that the measured relationship between a spaced pair of microphones is correct.

**Methodology**

Our methodological approach combines the use of nominalistic data generated by two experiments combined with participant observation conducted during the investigations. We use participant observation to bring our in-the-world experience to the research and present a humanistic dimension to the more nominalistic data that we generate (Atkinson & Hammersly, 1994). We triangulate this nominalistic and observational data with our tacit experience.
taken from our audio teaching practice. We aim to triangulate this multi-method approach to create a more in-depth analysis of our test results (Flick, 2018). We present two systems for using visual tools, in this case Auto-Align, that involve more consideration than just loading the plugin and using auto detection algorithms. These two tests generate both audio and visual data for later analysis to triangulate with observations that the two participants collected while conducting the experiments. Firstly, we use the live sound system design approach of McCarthy (2016) to understand phase relationships in the post-production mixing of a big band. Secondly, we use the same procedure, using Auto-Align as a measurement tool to explore and manipulate phase issues while tracking a drum kit before the sound is committed to tape.

A better understanding of phase concepts aids educators to reinforce the importance of microphone placement. However, a written medium lacks the practical engagement that we see and hear when physically recording and mixing music. For a lasting pedagogical tool, we use a combination of visual aural and written mediums that offer students and educators a more in-depth understanding through repeated viewings. In the case of the drum recording, an edited video serves to demystify and highlight both microphone movements and relationships in a practical environment. In the big band mix, audio examples aid to highlight the changes in amplitude, spatiality, and punch that are problematic semiotic descriptors with vague meanings. In particular, the video offers the opportunity for outside engagement through impact with industry partners and educators around the world. Thus creating a more in-depth understanding and confidence that visual, aural, and written demonstrations give early stage recording engineers.

**Drum Recording**

For the first measurement experiment using Auto-Align, we chose a multi-miked drum kit using a combination of close and distant microphones to present a multitude of destructive phase relationships. We recorded the drums at QUT recording studios, Kelvin Grove in Australia to test our theories in a controlled and professional environment. According to Keller (2011) "It’s hardly surprising that the more microphones used in a recording, the more potential for phase problems. In modern music recording, that usually points to the drum kit. (para, 17). We adopt a similar process to Weiss (2014), which begins with establishing a 'recorderman' overhead setup.
The recorderman overhead microphones are set up with one microphone approximately 120 cm directly above the snare drum and the other placed over the right shoulder of the drummer measured equidistant from the snare and kick drum. We then close mic all the snare, rack tom, and floor tom on the kit with microphones approximately 8 cm above each drum. We double mic the kick drum with a dynamic microphone just inside the back hole of the drum and a large diaphragm condenser outside the kick drum approximately 15 cm. We also use two room microphones to capture the space so that we have multiple close and far microphones to measure. These rooms are set approximately 220 cm away from the drum kit and measured so that they are equidistant from the kick drum. The recording was conducted at QUT Skyline Studios Using the Neve Custom 73 Console. Tab.1 is a list of all the microphones used in the experiment.

Once the drums were set up Dan recorded a simple drum line using a metronome. This recording is the 'before correction' example of a multi-microphone drum setup. After the recording, We used Auto-Align to measure each microphones phase relationship to the primary overhead above the snare drum. Using the readings from Auto-Align, Dan re-recorded the drum groove for comparison with the earlier 'before correction' example. Each example used the same gain structure, so the only audible difference occurs from microphone movements. For a more detailed explanation of the experiment and to hear a comparison of both recordings, please watch the embedded video1.

---

1 Video Example - Drum Recording: The following video gives an overview of the drum recording experiment. It practically demonstrates the use of visual alignment analysis to inform microphone placement. The video contains examples of drum recordings before and after microphone movement: https://www.dropbox.com/s/16ofal0qkkbni4p/Drum%20Phase%20Inst%20Vid.mp4?dl=0
Discussion

It is important to note that neither drum recording is perfectly in phase but the second recording was adjusted for the maximum constructive phase interaction possible. On closer inspection, this experiment yielded some surprising results that challenge the orthodoxy of drum recording. From a logistical perspective, Dan was able to take a three-person job of moving listening and playing and reduce it to a one person method with a surprising efficiency. We propose that this method of phase measurement is of particular help to smaller studios or self-produced songwriters without the resources to hire assistants. It also benefits producers that prefer to record in the room with a band as phase relationships are nearly impossible to hear when you are standing near a drum kit and attempting to evaluate microphones. Finally, this method benefits students of audio engineering by offering a comprehensive method for measuring and choosing phase relationships in multi-mic recordings. It also affords students the confidence to experiment with different microphone placements which opens up a new realm of creative manipulation in the recording environment. Having excellent recorded phase relationships also reduces the post-production work for any location recording where ideal listening environments do not exist such as mobile broadcast units.

Of particular surprise was the movement of the room microphones. It is common practice to measure room microphones so that they are equidistant from the source you want to emphasise. Up until this point, measuring from the kick drum has been the method of choice for our teaching practice. The aim is to create best phase relationship for room microphones and use those microphones to feature the kick drum in the room. In this case, Dan wanted the kick drum to feature in the room microphone recording, so he initially ensured that the room mics were equidistant from the kick drum. However, after measurement with a distance analytical tool, the right room microphone moved forward thirty centimetres, and the rear microphone moved back thirty centimetres. We posit that Auto-Align is also accounting for the shape of the room and measuring microphones considering the direct microphone relationship, the surrounding space as well as reflective wave information. This surprising development is of value to anyone who is recording in an unfamiliar drum room as phase measurement offers the opportunity to understand the relationships within any given space. This means that engineers don’t need to spend time accurately measuring rooms or guessing at microphone placements.

From a pedagogical perspective, the visualisation of phase relationships between the microphones on the plugin offers teachers a precise method for explaining what occurs to a recording when you move individual microphones within a network. This deeper understanding opens up possibilities for students to experiment with microphone placement and
manipulate the tonal colour of their recordings without the possibility of accidentally creating destructive relationships. This assurance in microphone relationships shifts the teaching emphasis from mastering microphone placement to a more confident, discovery-based experimental approach to recording. A visual representation also affords the opportunity for students to familiarise themselves with out of phase placement. Using analysis tools, students can determine the level of phase colouration they intend to achieve. This information gives students the confidence of knowing the precise phase relationship that they have as well as the ability to creatively manipulate recordings.

**Big Band Mixing**

Due to the thirty two microphones on the recording session, a recording of a big band provides a compelling opportunity to explore elaborate re-alignment in post-production. The methodology undertaken draws inspiration from live sound system design, with sectional mics aligned using an approach similar to that employed in McCarthy's (2016) ABC approach.

The initial processing involved alignment of bleed across all the microphones, with the drums providing the fundamental source for alignment. The alignment technique is the same as the one documented in the studio drum recording video presented earlier in this paper. Once the drum alignment was completed, Auto-Align was used to provide multiple in-phase suggestions as earlier demonstrated in the drum aligning video. These phase points are then auditioned and selected by ear. In this sense, we use the plugin to reduce the time shift selections to a manageable set of in-phase points. In live sound system design, the phase alignment in a single system takes place at the crossover point between speaker elements in a system.

This crossover point commonly occurs between a subwoofer and the principal part of a line array which is already phase aligned a part of the speaker design. The crossover point is the region in which there is maximum interaction between elements each generating the same frequency, and as such system designers strive to ensure that the majority of the target zone for the array are receiving both the subwoofers and tops ‘in phase’ (fig. 7). In the big band recording example, a similar approach is taken to alignment,
focussing on the target frequency range of the element being aligned to ensure optimised sonic performance in the alignment process.

In the case of the drum recording example, moving the room microphone 6 feet from the original position to ensure the desired response will not have a severe consequence in performance timing. As the microphone is in real physical space, it will provide significant tonal improvement when combining the room mic and close mics. These changes will occur as the relative intensity of the room mic increases and the balance of direct and reflected sound alters. In the case of the big band recording, however, virtually moving one microphone through multiple delay choices in post-production that forms part of a sax section will have negative consequences in the cohesion of the musical timing of each sections performance. To maintain the best possible timing in the performance across instrumental sections the choice of delay value is made based on two considerations. The first being the improvement in sonic performance, and the second is that the decision must minimise the virtual distance moved for each microphone as much as possible. Using a simple automatic alignment selection can see a delay suggested on a saxophone mic that effectively stages the saxophone behind the trumpet. This type of virtual movement is inappropriate for aligning big band microphones as it ignores the sound staging that audiences expect from big band recordings.

Once all the microphones have been aligned to the drum bleed, there is a change in the presentation of the drum sound in the recording. In this case, the impact of the room reflection is reduced resulting in a drier, closer sounding drum recording. The snare drum in particular comes across as being considerably less hollow with more sound of the body of the drum propagating into the recorded space. An example of the drums aligned and unaligned with all thirty two microphones open is provided in the example folder.²

In live sound design, the ABC approach sees the top section of a line array focussing on the furthest audio plane, this is labelled plane A. The next audience plane is targeted by the middle section of the array, this is labelled plane B. With the frequency response and amplitude at that plane a function of the combination of the top two sections of the array, rather than just the middle section. The lowest audience plane, C, is targeted by the lowest section of the array, but again is influenced by the previously focussed array sections. It is imperative that the individual sections are targeti

² Audio Examples - Big Band: The following is an example of the big band before and after alignment. The drum only alignment featuring all 32 microphones is also provided for reference. Note that the only change in the audio files is the phase aligning of instruments. The recordings and volumes of the instruments are the same in both examples: https://www.dropbox.com/sh/4i9bd9enbx5hoqi/AABQeGjpmkEZcQyndD0Qr-ca7dl=0

242
allocated audience zones, but also that the array as a whole is working as a single unit to provide polar pattern control over the desired frequency range. The lowest controlled frequency is a function of the length of the array, meaning that the response at the listening plane is influenced by more than the section of array targeting that plane. Approaching the big band recording requires a similar methodology, with the need to target not just the bleed from the drums into the sectional mics (sectional mics become plane A) but also the local bleed from other players in the section. The saxophone section provides an example of this approach, with the Baritone saxophone providing the primary reference. The ABC approach is used to align the saxophone section as if it were a line array, bringing each element into the best possible alignment at the target frequency range. The tenor saxophone is added to the baritone, and other close measured in-phase points are auditioned, focussing on the tonal relationship between the two instruments at the recommended points. The points that are auditioned are not from an alignment measurement between the two saxophones, but still from the drum alignment (plane A), preserving the phase relationship with the bleed while optimising the relationship between sectional instruments. The process continues, adding each sectional mic (plane B and C) in turn and tuning in to the existing section using alignment options provided by the relationship with the drums (plane A).

Discussion
The end result of this alignment process across all thirty two channels of the recording is a significant change in delivery. Before alignment there is a sense of a big band being in a space, with spatial cues arriving at the listener along with the direct sound, creating a sense of distance and a lack of intimacy and urgency. After alignment the staging of the band has changed, with a move from a band situated in a room to a sound that has clearer definition and a greater sense of immediacy, particularly in the brass stabs at the end of the short example provided. The sense of delivery and detail creates a greater sense of energy delivery, with the band now appearing to be located closer to the front of the soundstage rather than further into the room, with the splash of energy in the reverb now feeling like a result of the transient energy of the band, rather than a space that the transient is filtered through. Through a time and phase alignment process inspired by live sound system design approaches the balance of instruments can be manipulated without significant changes in frequency response, allowing considerable post production manipulation of sounds that sees significant changes in overall response without alignment.

From a pedagogical standpoint, the realigning of big band microphones using a considered methodology offers students the chance to understand how better microphone placement can affect the delivery of a recorded
performance. The opportunity to visualise and understand such relationships without having to move microphones allows teachers to demonstrate extensive phase relationships in a less time-pressured environment. In this post-production example, the visual and automatic options provide the chance for students to hear multiple different placement suggestions from the algorithm. However, it also affords the opportunity for students and teachers to engage in critical discourse as to why certain positions work better than others. Finally, this experiment offers an insight into the choices that engineers need to make when visualising a complete recording as well as an opportunity to evaluate microphone placements that need correcting in future recordings.

Conclusions

In both of these phase experiments, we used comprehensive visualisations to inform our choices in pre- and post-production. In the experiments and our pedagogical practice, the value of visual tools to inform recording practice cannot be overstated. In our tests, the ability to utilise a precise visual analysis of microphone relationships removes the guesswork from creative microphone positioning. In our teaching environments in QUT and UWL, phase visualisation tools now play a significant role in presenting an explicit picture of microphone phase relationships to students. However, it is important to remember that there is no correct answer with phase choices. The notion of perfect phase relationships for a multi-microphone recording is an impossibility. Using a multiple microphone setup involves creative decisions between different microphone types and distances. The result of this distance mismatch creates a sense of space around the drum kit due to the natural phase cancellation. As stated by Paterson (2007), these unavoidable phase differences offer the chance to decide on the phase colouration. This colouration is a phenomenon that the listener is familiar with due to the ubiquitous nature of multi microphonic drum kit approaches in past recordings. This experiment improves the understanding of phase relationships to make deliberate choices as to the aesthetics of multi-microphone recordings. We intended to utilise these systems as teaching tools so that early stage recording engineers and students can understand and manipulate microphone choices with a deeper understanding of the consequences of their actions. Experienced recording engineers learn to hear and minimise the phase cancellation so that the microphone relationships allow the speaker to represent the recorded text with minimal nullification due to destructive relationships. In other words, the speaker can move with more freedom because it is not being asked to move backward and forwards at the same time due to ill-considered microphone placement. This listening method is a skill that takes years to acquire. The experienced engineer draws on years of tacit
knowledge to quickly understand and correct phase issues. We assert that it is possible for the early career engineer using visual tools to produce their desired phase and frequency relationships in multi-miked drum recordings.

Further Study

Our use of optical phase analysis in post-production reveals a more considered approach to correcting phase issues. It is important to note that the tool we utilised is a post-production tool and is not designed for analysing microphone placement in recording scenarios. However, it was the only phase analysis tool that offered distance analysis as well as comprehensive frequency interaction, time delay, and phase analysis. As a result, it is the use of Auto-Align as an analytical tool that shows real promise for creative manipulation of phase in physical space. This experiment in phase analytics has opened up opportunities for experimentation using different microphone setups and relationships. As such, we intend to produce a series of videos that continue to explore the use of post-production analytical tools but with unconventional microphone placement in spaces that are less forgiving than QUT Skyline Studios. As stated in the literature review equalisation can drastically alter phase relationships between microphones. We intend to explore some aggressive equalisation and processing techniques with recordings to see if we radically shape drums using creative processing then reposition microphones for better phase coherence. We also intend to explore phase measurement in a variety of settings to test its validity for different applications such as guitar recording, orchestra recording, and various other acoustic sources. We feel that real-time phase measurement in the tracking environment offers the chance for us to develop a methodology of multi-microphone tracking that understands and creatively manipulates phase relationships to push analog recording into new frontiers of creativity.

References


Daniel Pratt, Shane Hoose & Wellington Gordon:

Transnational Flow in Cloud-based Music Production:
Organisational Communication and Collaboration Between
Australia and America

Abstract

In this paper, we take three researchers from different parts of the globe and experiment with long-distance songwriting through a sociological and organizational lens. During this process, we combine production expertise and facilities into a networked recording studio environment with expanded toolsets centred on songwriting production. We investigate a new flow concept that we name transnational flow (TNF). This flow state exists as a form of non-synchronous group flow somewhere between Csikszentmihalyi’s (1995) individualist flow theory and Sawyer’s (2007) group flow theory. The paper investigates the difficulties of managing multiple flow states in a complex organizational network between Australia and North America. We use Karl Weick’s Sensemaking (1995) and Jazz metaphor (1998) theories to form an organizational framework that enables the creative flow process to occur over a complex network with a time difference of seven hours.

Introduction

Several digital workstations have implemented collaborative elements to their platforms with varying degrees of success. Questions of stability, consistency, and complexity surround projects when artists collaborate across different counties to produce new music. Rather than focus on the technical aspects of collaboration this paper investigates the communication strategies surrounding a transnational workflow. We take three experienced songwriting producers from three different locations across Australia and North America and explore the interpersonal and technical aspects of group communication and flow. We posit that there is a networked state of non-synchronous flow that occurs in this framework without concrete borders or time restrictions.
In this project, we treated the writing and recording of a song as an organizational exercise informed by distributed creative technologies, communication theory, and group flow theory. All members of the group participated in every aspect of the decision-making including the writing, recording, and mixing of a new song. This full integration enabled us to investigate group interplay from the inception to the conclusion of a project. Due to the logistical nature of a project between North America and Australia, there was a binary interaction of individual creative flow states and socially negotiated logistical decision-making. The creation of the song occurred as a networked flow of ideas, and the organization process to facilitate these flow states occurred in weekly sensemaking meetings. We tested whether it was possible to maintain a state of networked group flow as a songwriting production team without existing in the same space and time zone. Our research investigated how a socially constructed organizational approach reduced factors like distance and time in the creation of new musical content.

**TNF in the literature**

Transnational flow (TNF) is a networked system of flow states occurring in multiple regions of the globe. It utilizes technology to connect separate flow states and builds them into a group flow interaction. It draws on Csikszentmihalyi (2000) and Sawyer’s (2007) sociological work on the subject of flow and group flow. TNF is a non-synchronous hybrid that draws on elements group flow and individual flow, it occurs as a network of semi-isolated flow states; this means that TNF lacks the direct physical interaction that usually happens in a creative collaboration such as jazz improvisation where group members respond to cues in real time. Unlike Sawyer’s (2007) group flow theory, TNF is reliant on the analytical sensemaking models and Jazz Metaphor approach of Weick (1995) to facilitate the creation process. TNF differs from Sawyer’s (2007) group interaction theory used to explain Jazz improvisation and group interaction in theatre because the networked group is involved in a form of delayed networked flow states. However, TNF is not an individualist form of flow as discussed by Csikszentmihalyi (2000). An example of a more traditional form of non-synchronous flow occurs in recording studio settings where group members arrive at different times to perform parts in a staggered manner. However, in this TNF example, the group members are never in the same location and have no mediating factor such as a shared studio facility, engineer, or producer.

Weick’s (1995, 1998) flexible organizational approach allows us to transverse the increased level of complexity derived from working with mixed technologies in multiple regions of the world. This structure establishes a binary of creative flow states (Csikszentmihalyi, 2000, Sawyer, 2007) and retrospective analytical states (Weick, 1995, Sutcliffe &
Obstfield, 2015, Gioia, 1988, Rutledge, 2012, Allan, 2009). As a result, this form of group creative flow follows a model of sensemaking meetings, recording in a non-synchronous flow state, analyzing the process and repeating until the project ends. Thus creating an interactive binary process of analysis and creativity.

Distributed Creativity

TNF is a form of distributed creativity that needs an analytical framework coupled with a creative organizational philosophy and a networked central nexus. This investigation represents a non-individualistic creative process that we refer to as distributed creativity (Sawyer & DeZutter, 2009). Sawyer and DeZutter (2009) define the group creative process as “[o]ne that generates a creative product, but one in which no single participant’s contribution determines the result” (p.81). Campelo and Howlett (2013) adopt the distributed creativity model to “refer to situations where collaborating groups of individuals collectively generate a shared creative product”. Brown (in Salavuo 2006) discuss the cognitive diversity of online music communities, referring to the concept of distributed expertise as “individual members possess knowledge of a particular subject, which exceeds the knowledge of the whole community” (2006: p. 255). In music production, a single recorded artifact is produced as the result of coordinated output by skilled experts and machine labor and through collaborative exchanges (Lefford, 2015). Tasks are not accomplished by separate individuals, but rather through the interactions of those individuals (DeZutter, 2009). To sustain this collaboration, experts come together in a shared virtual production, which is delineated by agreement among collaborators. Usually, the context for collaboration consists of a combination of conceptual and physical configurations (Lefford, 2015).
Flow

Flow is a cognitive feedback state that exists as people experience peak mental activity. It is a subjective state where the human brain processes and adjusts in a seamless motion (Csikszentmihalyi, 1990). According to Sawyer (2007), flow is a difficult state to attain and relies on particular circumstances. To reach a flow state a person must have the skills to match the task, clarity of goals, immediate feedback on achieving their targets, and complete concentration on the task (Sawyer, 2007). The knock on effects of a flow state are increased confidence, clarity of goals, the transformation of time constraints and a loss of self-consciousness (Walker, 2010). Csikszentmihalyi (2000) states that flow contributes to the understanding of "enjoyment, here and now - not as compensation for past desires, not as preparation for future needs, but as an ongoing process which provides rewarding experiences in the present" (p.9). In the recording field, this concept of positive psychology is backed up by Howlett (2009) in his discussion of recording session flow and the added benefits that arose from a positively flowing recording session. Flow occurs in creative projects like music recording because of the intrinsic reward that naturally takes place in creative, driven activity. This, in turn, attracts autotelic personalities that thrive on the intrinsic rewards that imaginative invention provides (Csikszentmihalyi, 2000).

Keith Sawyer takes the notion of flow and de-individualized the process. He argues that collectives with an accumulated body of tacit knowledge have the potential to engage in a group flow experience that exceeds the boundaries of mono flow states (Sawyer, 2007). In this sense, groups engage in a synchronization of flow states but with a wider palette of interconnected knowledge. In a collaborative setting, such as jazz improvisation, a diverse base of tacit knowledge increases the potential for creative outcomes. This decentralized form of creative decision-making results in the recognition of, as well as deference to, the "emergent flow of the group" (Morrow, 2012). However, for productive and insightful group flow to occur, collectives need to meet certain conditions. As previously mentioned, TNF borrows from group flow but lacks some of the direct response to cues due to the delayed nature of flow states separated by less compatible time zones. As an emerging theory, TNF falls between individual and interactive group flow states. To set conditions for TNF, we borrow from both Sawyer (2007) and Csikszentmihalyi (1990) but add intersectional organizational elements from Weick (1996). This triangulation of theories allows us to create the conditions to connect individual flow states that interact over vast distances within a technologically sophisticated framework.
Conditions for TNF

1. TNF group members need to have a minimum level of tacit knowledge and skillsets specific to the project.

2. TNF Group members need to have a common shared knowledge base and can have further specialization in certain areas.

3. TNF has no concrete boundaries but combinations of spaces that create abstract conceptualizations of the working creative environment.

4. TNF requires the separation of analytical states from creative states to facilitate the individual flow moments for networking.

5. TNF needs an overarching pragmatic and improvisational philosophy that facilitates and accounts for the unpredictable nature of networked recording sessions occurring over long distances.

Sensemaking

A network of decentralized flow states relies on regular sensemaking discussions to untangle the complexity associated with the creative invention in a decentralized structure. Sensemaking is best described by Weick, Sutcliffe, and Obstfield (2005) as "the ongoing retrospective development of plausible images that rationalize what people are doing" (p.409). It is a socially driven group discovery mechanism whereby members examine and discuss complexity to increase comprehension and come to a unified understanding (Rutledge, 2012). Sensemaking differs from the usual cognitive looping that producers and music makers encounter when operating as individuals. Sensemaking is more of a “social constructionist process in which people respond to the unusual, the confusing, the unexpected and change by using cues from the environment and their mental maps to construct meaning and literally make sense of events” (Allen, 2011).

Karl Weick (1995) grounds sensemaking using seven properties with the intent of placing "boundaries around the phenomenon of sensemaking” (p. 18). These seven properties are the construction and questioning of identity, retrospective analysis, adherence to sensible environments, social inquiry, ongoing examination, the extraction and mapping of cues, and the drive towards plausibility over accuracy (Allan, 2009). In Weick's (1995) own words “This sequence is crude because it omits feedback loops, simultaneous processing, and the fact that over time, some steps may drop out” (p. 18). When applying sensemaking in a practical setting like TNF, it is necessary to divide these properties into either implicit or explicit categories. This separation means we reduce these crude markers into a more practical model that is applicable during a creative meeting. To develop our context-specific model of sensemaking we draw on the adaptations of Rutledge (2012).
To simplify the process into a practical model, we assert that certain properties of sensemaking occur implicitly to our experience of TNF. Firstly, meetings for a TNF project are implicitly social by their nature. Secondly, the framework, tools, and roles occur in a context specific environment. Thirdly, the communication, file sharing, and recording technology coupled with the historical conventions of music production served to direct our environmental boundaries. Our roles and environmental limitations serve to inform our grounded identity constructions. It is when we question these constructed identities that our sense of flow is interrupted and the practical implementation of sensemaking begins (Weick 1995). Understanding the implicit and explicit sensemaking properties that apply we are able to construct a practical model based on Rutledge’s (2009) model of sensemaking.

To untangle a complex problem we begin to "bracket" (Rutledge, 2009) to isolate the interruption that removes us from the the flow of experience. Once outside that flow, we use retrospective analysis to determine and map the nature of the disruption (Weick, 1995). During the mapping process, we try and discard phrases to understand the nature of the interruption. As soon as the mapping occurs, we then suggest a series of plausible solutions. Plausibility is the fundamental criterion of sensemaking" (Weick, Sutcliffe, & Obstfield, 2005) because it removes the debilitating need for precise answers and moves the social group towards action (Rutledge, 2009). The final stage of this social cognitive exchange is taking action and pragmatically analyzing the results of the action. This process is ongoing and it is repeated at every meeting so that the creative vision of the recording session is constantly revised, refined, and updated.

**The Jazz Metaphor**

![Figure 2. The practical model for sensemaking process based on Rutledge (2009)](image)

Improvisation as a mindset for organizational analysis is part of Karl Weick’s (1998) later organizational communication work which bears similarities to Sawyer’s (2007) jazz-laden group flow metaphors. Using an improvisational Jazz based metaphor to support organization is hardly a new concept in Organizational theory (See: Senge, 1994; Hatch, 1999; Hum-
phreys, Brown & Hatch, 2003; Moorman and Miner, 1998a, 1998b; Barrett, 2000, Mantere, Sillince, & Hämäläinen, 2007). Weick (1998) presents organization as the antithesis to creative sociology such as flow (Csikszentmihalyi, 1990). He infers that an improvisational mindset for any organization adds flexibility to the structure and accounts for interpretation, embellishment, and variation. Each of these steps is an incremental departure from the original intent of the plan. In TNF we use sensemaking as a methodology for dealing with complexity in the recording process. However, the improvisational mindset enables us to begin our research without a complete plan. "Thus improvisation deals with the unforeseen, it works without a prior stipulation, it works with the unexpected" (Weick, 1998, p.544). Taking this mindset into consideration early in our recording, we experimented with our initial recording setup.

**Methodology**

To understand the relational, interactive, and process-oriented nature of TNF we employ a multi-qualitative approach to investigating it as phenomena. We thematically organize our data analysis into three examples of networked flow states, sensemaking, and improvisational mindset. We triangulate between recorded meeting data, observational reflections, and treat the song creation as an observable organizational process. The transnational distribution of the project members creates an enforced separation that offers a chance for us to record our sensemaking sessions and observe our organizational method as a separate analytical state while relying on observational data to understand our creative states.

**Participant Observation**

Participant observation serves as our primary methodology to engage on an interactive level with our emerging data. Atkins and Hammersly (1994) point out that participant observation is “observation carried out when the researcher is playing an established participant role in the scene studied.” This qualitative approach provides us with the ability to develop a deeper understanding of process and relational elements with which to situate our research. Acting from within the system we contextualize and interpret unfolding data of significance to expand our awareness of events (Schwartz & Schwartz, 1995). As participating actors in the process we generate data that can meaningfully add to our collective understanding of human experience (Guest et.al, 2013). Interview data provides opinions on a project in hindsight but doesn’t provide insight on the mechanics of process that happen from the perspective of the participant. One of the difficulties that neutral
observation faces is the accusation of participating in a “form of tourism rather than data collection” (Guest, et.al 2013).

In the case of any TNF exercise, we are incapable of allowing ethnographic tourism. Each member must understand the concepts with a depth of insight only possible with an intimate knowledge of recording process. TNF requires a particular type of implied knowledge that sets a minimum standard for participation in a project. Morrow (2007) suggests that for a recording group to experience a cohesive form of flow "the members have to share tacit knowledge and demonstrate comparable skill levels". Our approach was to assemble a group of researchers that have a diverse enough skill set to enhance our potential for group flow. Each researcher needed a baseline of tacit knowledge appropriate to key recording concepts such as gain staging, equalization, microphone technique, and musical ability. There is an invasive aspect to neutral observation and “studio access is generally open only to those involved in the production” (Thompson & Lashua, 2014 p.749). This means that observers “are considered interlopers, at worst they are obstacles to the recording process” (ibid). The non-participating observer may alter the course of events, even when the observer is temporarily absent” (Schwartz & Schwartz, 1995 p.346).

Interaction analysis

Due to our use of Skype video conferencing we have access a rich source of non-verbal and interactional information to expand the scope of our transcriptions. Interaction analysis provides a methodology we use to analyze distributed creativity—a method of analyzing verbal gestures, body language, and conversation during collaboration. (Sawyer, 2007) Using interaction analysis, researchers explore group creative processes and gain new insights into how creative products emerge collaboratively from groups. Multiple viewings of recorded video reveal the richness of both verbal and nonverbal actions, allowing researchers for the first time to study the rich multimodal nature of human interaction (Sawyer/DeZutter, 2009). Transcription methods vary in detail depending on the researcher’s interest; in some cases, only talk is transcribed; in other cases, nonverbal details such as eye gaze and body position are important and recorded along with speech (ibid).

The Technology to Sustain a Creative Network

The network for this project consisted of recording facilities in three different locations around the globe, Kentucky, Maine USA, and Brisbane Australia. In each location, participants had access to different analog and digital recording equipment. The distribution of tacit knowledge and studio facilities in this project presented a compartmentalised approach to recording
where each member performed specific roles. In a musical sense, participants were specialists, but not solo artists. Drums were tracked in Kentucky, bass was tracked in Maine, and the guitar and vocal tracks were recorded in Australia. This larger abstract recording studio also provided the ability to tap into a global network of performers, opening up project studios to a wider scope of musical guests who contributed to a growing pool of tacit knowledge. Additional musicians recorded overdubs in Kentucky such as theremin, saxophone, keyboard, and tape effects.

We used a variety of analog and digital equipment during the recording. Cockos REAPER served as the digital audio workstation (DAW) for this project because it was an easily accessible platform. Using a unified DAW minimized conversion time and confusion when exchanging sessions. The use of analog processing including time based processing like reverbs and delays reduced the need for a unified plugin library. Furthermore, specialised plugins were “printed” into the session to avoid compatibility issues. This combination of unique processing tools allowed us to define our sound.

We saved the Reaper session on the Dropbox file sharing platform. The file sharing platform enabled our session to reside on a server in the cloud, rather than on local machines. File management through Dropbox provided a central nexus for our transnationally distributed creative team. This nexus allowed participants to interact without all having to be present at the same time. Each contributor could upload, download, comment and make changes to a session independently from the other collaborators.

We held regular weekly meetings throughout the songwriting and production process. Using Skype as our communication platform during these meetings enabled us to conveniently make immediate decisions.

Figure 3. Small networked facilities combine to create one large abstract studio.
regarding the musical ideas, direction, and production techniques. These sensemaking sessions facilitated the group flow and assisted the progression of the project. Working in weekly stages allowed us to revise the previous achievements and plan for the following week of creative invention. The sensemaking meetings allowed us to manage any technical or accidental complications proactively. The video platform of Skype enabled us to gauge each other’s reaction to ideas reflexively. Having the ability to view and socially respond to nonverbal cues added another level of feedback that is not possible through written or verbal communication.

File naming systems and organization were an especially important aspect of this project. Working collaboratively in the cloud can lead to some potential file conflicts. The numbering system was broken into whole numbers depending on which part of the song we worked on. For example the drum and bass recordings over the first weeks were numbered 1. As we gradually updated the drum and bass parts we included decimal increments to the whole number so that we didn’t accidentally create file conflicts. For example 1.1, 1.2, 1.3. If we worked on any significant parts we would also add a descriptor to the numbers so that we could access the session quickly during our weekly meetings.

Results and Discussion

Our study yielded significant examples of TNF conditions in action. The following three examples are thematically organized into examples of Flow and Group Flow in a distributed environment, Sensemaking, and Improvisational approach.

Example 1 - Reflections on the Beginning of the Project

In this example we extracted the reflective entries of the participants to examine how this process began and where the experience sits in relation to the concepts of Flow and Group Flow and the emergence of the creative network.

Daniel Pratt Reflection: This song idea began on a coincidental plane ride with myself and Shane on the way back from ARP 2016 in Denmark. By pure chance, we happened to be sitting next to each other on the flight from Aalborg to Copenhagen. At the time Shane and I had not spoken much, but we immediately started talking about writing music together. It was a vague fun plan centered around connecting two studios and having fun writing a song. After I returned to Australia, we both remained in contact and chatted via Skype about music. One day Shane just sent through a drum track he'd been working on, and it immediately triggered my creative instincts. As a producer, my workflow revolves around bouncing off ideas and adding my
take to things. This drum track inspired me to spend a few hours playing with complimentary guitar lines, and, before I knew it, we had effortlessly arranged a song that had an upbeat, energetic vibe. Wells, joined the group after responding to a Facebook comment and suddenly the three of us were bouncing ideas around and writing a song. This process grew organically, and I became excited to wake up in my different time zone to hear what had happened while the other guys were awake and tracking away. It’s a motivating feeling to get up in the morning and hear these great ideas coming from the other side of the world. As the project grew in complexity we decided to start a weekly meeting so that we could plan the song to avoid a haphazard approach and focus our creative energies. (Dan Pratt, Reflection - September 2017)

Shane Hoose Reflection: This collaboration began last December—shortly after the ARP2016 conference. I usually practice with backing tracks in a variety of styles at home. Although most of the stuff that I play falls specifically under the "rock" category, I feel that it is important to be able to play the other styles well. Entering into the process, I honestly had no idea where this thing was headed, or if it would even proceed or would be successful. Since I was essentially playing by myself and only with a metronome, I was very careful not to overplay anything from the beginning. Otherwise, the drum part would be too cluttered, and there would not be adequate space for anything else. I figured that it would be best for me to keep things somewhat basic as well because I wasn’t necessarily sure what styles everyone else was comfortable with. I aimed for a solid backing track that would serve as a bed for anything else that was added. I remember one of the earliest discussions before we actually did anything was an agreement to use Reaper as the DAW. In other words, we agreed on some specific parameters for the project. If I remember right, all of our communications were by email until Wells joined. It was early February when Wells came into the discussion for the first time. That was when the Skype meetings began. That was an important point in the process. The unique thing about this collaboration for me is that I knew absolutely nothing about anyone’s prior musical or production experience or style entering into this. From this standpoint, it seemed like totally shooting in the dark. (Shane Hoose, Reflection - May 2010)

Wellington Gordon Reflection: When I first listened to Dan and Shane’s track, I was excited. Sonically it bared resemblances of many of my favorite style of music. Shane’s drums had a classic Motown vibe, Dan’s guitar also suggested a Motown feel but also it’s skank like rhythm that reminded me of reggae. Dan’s choice of reverb on the guitar and drums also created a sense of nostalgia for me. The combination of the drums and guitar placed the sonic signature close to artists like Amy Winehouse, the Daptone’s catalog or Lilly Allen. Dan’s voice and accent also gave me a sense of familiarity and reminded me of some alternative British groups. (Wellington Gordon, Reflection - April 2017)

These three reflections provide insight into the reflexive group flow that occurred in the project. They refer to the organic process that built a technological framework around the group's creative invention. Dan, Wells,
And Shane began weekly meetings as soon as they decided that the project had reached a level of complexity that required organization. Each participant noted that the previous recording induced an excitement and focus to their writing process. For example, Wells explained his nostalgia and excitement when hearing the flow states of Dan and Shane for the first time. He explains how this triggered his past knowledge of musical acts that informed his production practice and how they relate to the feel of the song. As a flow process, each member triggered the other into a collective creative state associated with Sawyer's (2007) group flow. Dan noted that Shane’s drum work inspired him into an extended flow state. It was interesting to note that the difference in time zones and the distributed form of creativity gave the participants the ability to both influence each other’s creative invention as well as to exist in a constant personal state of flow. In drawing conclusions from this analysis we try to avoid leaning on the emotional responses of the participants but we do notice that these flow states occurred as a direct reaction to previous flow states and not of their own accord. This makes it difficult to untangle emotional response to stimuli to the triggering of a flow state. As a result, we claim that these states fall somewhere between Csikszentmihalyi (2000) and Sawyer's (2007) flow and group flow theory and draw from the characteristics of both individualistic and group flow theories.

Example 2 - Tape Loops to the Rescue : Sensemaking to Restore flow

In this example we take three transcripts from the recorded meeting videos to examine a sensemaking process that occurred during a major change to the bridge section of the song. The bridge section was detracting from the individual flow of the group members and caused the creation of the song to halt. These conversations provide insight into how the group made sense of the interruption and how Shane’s contribution provided a flow state that restarted the non-synchronous flow of the group. It is important to note that the group cannot use Sawyer’s (2007) group flow to rewrite the bridge section in real time by reacting to cues and presenting ideas in a shared space. As a result, the group relies on sensemaking to discuss solutions and restore the flow of the project.

Meeting Video #6

**Dan:** The brass stuff is taking away from the way the guitars bass and drums interacting, and I feel like it's not working. Then there is the autumn leaves section, I’m nervous that we will sound like irrelevant white academics trying to play funk. I was thinking that we might be moving away from the funk and moving towards a more Portishead vibe. Remove the bridge and horns and add maybe street percussion.
(Dan makes the effort to look into the camera - so that it appears that he is trying to make eye contact - a form of sincerity when offering constructive criticism and new direction to the song)

Shane: What do you mean street percussion?

Dan: Roto toms, people clapping and talking, like a street party.

Shane: I kind of like the idea.

(Shane nodes and smiles)

Dan: It would be fun.

(Dan becomes more animated with enthusiasm when continuing the conversation)

Shane: I could get some timbale and other percussion.

Dan: Let’s change the horns and look towards the keyboardist and see what he can do.

Shane: The keyboardist has rhodes, organs, moog synths.

(both visually show excitement for the new plan)

Meeting Video #9

Shane: Alright. I added some more craziness in . . . (Dan smiles while listening in response and says ohh yeah) Just a little bit. I actually came up with one or two other ideas too for this thing. That middle section where it was originally the circle of fifths type of progression where I think that we are collectively thinking that it is going to do something really different there.

(Shane is more animated than usual when explaining his ideas about the bridge section - showing excitement)

Wells: I am open to seeing where it goes for sure. I wasn’t heartbroken that we went different directions or anything.

Shane: What would it sound like too if maybe we added some guitar feedback in there too maybe

Dan: I’m in. I’ll pop that in. I’m about to spend the day rebuilding the studio.

Shane: I’m not saying that it is something that we have to use, but it is something that came to mind.
Dan: Crowd noise too.

Shane: I did get us some crowd noise too this past week.

Dan: I’m listening to the crowd noise right now.

(Dan Smiles while listening to the crowd noise)

Shane: For some reason I stepped outside on Friday night and they were setting off fireworks, and I thought it was a good idea.

Dan: I might throw in some random yelling too. I am working with some high school students on Friday. I might try to get them to do some yelling. Like “yay” and “hey”, stuff like that. The crowd noise sounds great.

(Both Shane and Wells smile and nod in agreement)

Shane: I also threw a tape loop in there to see what it would sound like.

Dan: You called it “tape loops”?

Shane: Yeah.

[Dan listens to the tape loops]

Dan: Oh. Dude, that is cool.

(Dan Smiles in approval)

Meeting Video #14

Shane: I think that the moment that we decided to veer away from the Motown sound and head in a different direction that led us down the road that we ventured down was a good decision in hindsight.

(Shane looks satisfied as he reiterates how the group’s choice to move in the current direction has been a positive experience)

Dan: Yes. It was a really good decision to abandon in than to try to push it. There are things that we do as a collective where Motown is a very good basis for starting, but then it can go somewhere else. And I think that your tape tricks are a thing, and it is a cool thing.

(Shane smiles and nods in agreement) (Dan has enthusiasm in his voice and a satisfied look on his face when mentioning Shane’s approach to using analog tape)
Shane: I thought of that too as I was messing with that stuff, but we all have different tools that we use. But, the tape machine is one of those tools where not everyone has one now, and the people that do have them are not necessarily use them to do unusual things like obliterating things and taping them back together in random order.

Dan: Yes. The tape machine as a creative tool rather than a mastering device.

These three transcripts provide an insight into related sensemaking processes. In the first transcription we observe the identification and bracketing of the area that needs improvement. In the second transcription the group reacts to the plausible solution to a problem that halted the creative flow of the group. In the third transcript we identify the action point that occurred as a result sensemaking process. In the first video Dan is communicating an irritation of the senses (Pierce, 1778), causing him to doubt his process and investigate the source of his irritation. Weick (1995) states that the questioning of identity construction occurs when our senses are irritated, causing members to step out of the flow and analyse the interruption that is impacting the creative process. This interruption to the flow causes a member of the group to begin the process of bracketing (Rutledge, 2009), Dan brackets his concerns around a chordal movement in the bridge of the song. He describes this area as the Autumn Leaves section.

The bridge follows the chord progression of Autumn Leaves and the harmonic movement is lifted from a Jazz standard which serves to date the section of the music. This informal naming of the bridge informs the members on the positioning of the disputed area, while at the same time identifies the harmonic problem with the specified segment. The team identifies the context-specific obstacle and adopts mapping phrases to assess the boundaries of the problem before they begin the process of plausible solutions. During the mapping process, Dan uses phrases like irrelevant and trying to play funk to help identify the issue to the group. He also states that the recording started with a Motown flavor but has since moved into a more modern sphere and the bridge is holding the music back.

After this bracketing and mapping, Dan presents a plausible replacement for the section. The proposition is to remove the Autumn Leaves chordal

![Figure 4. The first two stages of a practical sensemaking model.](image-url)
approach and replace it with a rhythmically focused street party soundscape. This new approach is a reaction to the perceived problems caused by the more harmonically focused writing methodology. Dan and Shane discuss the details of plausible solutions, Shane offers different forms of instrumentation that he can contribute. During this conversation, a more detailed picture of the new approach to the bridge emerges. In this sense, Shane’s new recording of tape loops stimulates the other members of the group and restores the flow of the project. In this reflexive discussion, Dan and Shane are developing ideas that place boundaries around the developing creative objective for the bridge. Setting contextual limits around creativity gives the group the ability to constrain creativity into a more sensible environmental boundary appropriate to the stylistic approach of the recording.

When Dan suggests there is an issue, both Shane and Wells asks for clarification. After this, the parameters were defined, bracketed, and mapped until the group agrees on the boundaries for the problem. Once the mapping occurred the trio begin to plan plausible solutions that move towards action. During this process, there is an action-driven motive emerging through the group dynamic. This outlook displayed by Wells exemplifies the improvisational mindset to the organization of the recording as opposed to constraining themselves to the already established process of ideation. The information generated by bracketing and mapping then informs the presentation of plausible solutions.

![Figure 5. Full sensemaking model, taking action on the bridge.](image)

**Example 3 - Different Worlds Collide**

Weick (1998) presents a binary of organizational structures in his improvisational approach that is counter to the musical framework. He asserts that the constrictive nature of organizations serves to repress creative invention rather than nurture it (ibid). His Improvisational mindset maintains an attitude of detaching from rigid structures and adopting an receptivity to organizational leaps. In this example we investigate the improvisational mindset producing unexpected results in TNF.

**Meeting Video #3**

Dan: When you did a bassline on a conflicted track on Dropbox. Rather than looking at that as a problem, we started playing around with the two conflicted copies and I started copying and pasting things over and all of a sudden
we had a new vocal line and a bassline that never would have happened any other way. Meeting Video #3

Wellington Gordon Reflection: At the point that I started to collaborate with Shane and Dan- I was still a bit disoriented with the recording process and file management system that they had put into place. Unknowingly, I opened the wrong session, which looked like the appropriate one. I recorded a bass part based on the arrangement of the drums, guitars and scratch vocals. When we reviewed the work I had done, we soon realized that I had opened the wrong session. Dan, being an optimistic person, tried to make due with the situation. My perspective was that the song’s form and arrangement was still being determined/flushed out- so there was a sense of improvisation and flexibility given to all of our ideas. In some ways I felt that Dan was still familiarizing himself with Reaper’s capabilities, pushing Reaper’s editing features and this was an opportunity to see how he could manipulate and work in this platform. The transformed and resulting arrangement was created by Dan’s vision while Shane and I offered positive feedback in support of his musical instincts.

In this example, the group encounters a phenomenon that frequently occurs in the writing and recording of music. A happy accident. However, the unique conditions of TNF lead towards consequences that are unlikely to occur in a traditional recording environment. While recording their parts in the verse, Wells (bass) and Dan (vocals) accidentally recorded on different sessions. The consequence was that they both contributed musical lines without influencing each other. Fortunately, due to the agreed labeling system and the added security that Dropbox saves conflicted files, it was simple to merge the two recordings into a new session and listen to them interact. Dan notes in the transcript that this could not have happened in any other way. This single interaction became integral to the atmosphere of the verse and continued to reshape the whole approach of the song.

This outcome highlights the fifth TNF condition of a pragmatic and improvisational philosophy. Members of the group kept an open-minded approach to the benefits that accrue when organizational leaps transpire. Weick (1998) frames these organizational leaps on a continuum of "interpretation, embellishment, and variation" (Weick, 1998, p.544) using terminology lifted from Jazz improvisation. This is an example of variation which is the most transformative of the three organizational leaps. The music transformed in style and approach due to the accident of opening the wrong file and recording two flow states independently without any form of interaction. In this example, the change worked because Dan and Wells had no affinity to either idea. To account for the shift in direction Dan re-recorded the vocal line to remove minor clashes and the organization of the future instrumentation was modified to account for the new musical direction. The group embraced the nature of an improvisational mindset, which expanded their compositional options to consider mistakes as possible
advantages. This ethos kept the flow of the writing process fluid and responsive to creative leaps.

Conclusions

This project demonstrated varying degrees of TNF as defined by our literature support, methodology, and research design. As a result, we created a new song, it is about a used car salesman attempting to fix his soul to sell it to The Devil.¹

The results of this study determined that collaboration and networked creative flow states can occur between Australia and North America and that it is possible to maintain a state of networked group flow without being in the same locality. It also revealed that creative use of abstract facility networks expanded the scope and roles within the context of our constructed environment. The group engaged in both individual and non-synchronous group flow states as well as sensemaking and organizational improvisation in the music production process between Australia and North America. TNF relied on regular sensemaking discussions to untangle the complexity associated with the creative invention in a decentralized structure. In this pilot study we employed sensemaking sessions as well as an improvisational approach to organize the writing and assist in the creative flow of the group.

Our first attempt at transnational flow was a successful pilot, however, more research is required to expand the narrow scope of the project. We used this pilot study to develop models to enhance our organizational and creative flow but we need to test this on a more far reaching project. Our next aim is to record an entire album and increase the managing complexity of the project in order to test our strategies. We intend to push the boundaries of our transnational group flow to see if it is capable of adapting to greater complexity. We also need to facilitate an environment with more time pressure. There were no behavioral issues that required management which would be an interesting outcome for the research but each member of the project contributed a positive outlook. It is impossible in this sense to create conflict for the sake of writing an exciting paper. The obvious drawback is that manufacturing disagreement is disingenuous and there is no way to achieve this without coloring the research. Additionally, Australia and North America share cultural commonalities which reduce the complexity of the exercise. More cultural and gender diversity must occur to develop a more inclusive sensemaking model.

¹ Unfortunately, The Devil isn’t buying please click the following link to listen to the song: https://www.dropbox.com/s/s6a8n1ehkc0db/Soul%20Man%20Don%27t%20Know.wav?dl=0
References


265


Méi-Ra St-Laurent: "It’s kind of in the middle": The ‘Mid-Fi’ Aesthetic: Toward a New Designation of Black Metal Aesthetic of Recording. The Case of the Québec Black Metal Scene of recording

Abstract

The designation of the black metal sound as “lo-fi” by academics is problematic, since it doesn’t reflect the contemporary black metal production. Fieldwork in the Québec scene show that the members designate their sound as being “in the middle” – or mid-fi as I advocate it. In this paper, I first define “hi-fi”, “lo-fi”, “authenticity” and “mid-fi” in the context of black metal. Then, I present different black metal sound aesthetic using audio extracts, frequency analysis, spectrogram images and ethnographic data. This analysis allow better qualify black metal aesthetic production today using the Québec scene as an example.

Introduction

Between April 2015 and December 2017, I conducted nineteen ethnographic interviews with twenty-five members of the black metal scene in the province of Québec (Canada) as a part of my doctoral thesis in musicology1. From the first interviews, I realized that the members of the scene described their sound aesthetics as neither lo-fi nor hi-fi. This aesthetic designation of the “in-between”, which I came to designate as “mid-fi” – a term I will define in this paper – seemed to reveal two paradigm shifts in the aesthetic production of black metal.

First of all, this change in the aesthetic designation challenges the very way in which the style has been understood so far by academics. Indeed, this subgenre of extreme metal, developed in the late 1980s, is linked to the lo-fi

---

1 My doctoral thesis is about the phonographic and identity-based narrative of the black metal scene in Québec.
aesthetic of the first Norwegian black metal bands of the 1990s. It has also
been described by scholars as having a minimal, raw, ugly, amateur and
noisy sound production (Reyes: 2013; Thompson: 2012; Hainaut: 2012;
Hagen: 2011). However, it has to be said that many groups have moved
away from these standards over the last twenty years and that the mere “lo-
fi” designation is no longer sufficient to describe today’s black metal sound.
Secondly, this dissociation from the lo-fi aesthetic is present in the music of
many contemporary black metal bands (namely Cult of Fire, Mgła and Batu-
shka) and also suggests that the type of aesthetic production chosen does not
seem to have the same importance that it previously had in determining the
authenticity of the music, a fundamental value in black metal, since its in-
ception. But before elaborating, it is necessary to ask the following ques-
tions: how has the aesthetic production of black metal evolved since the
1990s? What does “mid-fi” refer to? And can we really use this term to de-
scribe the sound aesthetics of Québec’s black metal bands?
To be able to better categorize the aesthetic used by Québec’s scene, I
will first define the notions of fidelity, hi-fi, lo-fi and authenticity, while
considering the discourse of the Québec scene’s members about these el-
ements. I will also offer a provisional definition of the term mid-fi. Secondly,
I will take a closer look inside the black metal sound production from its
beginnings until today by focusing on both lo-fi and hi-fi productions, in-
cluding those from Québec. To do so, I will use sound excerpts, frequency
analysis and spectrograms of the targeted recordings. In the last section of
the paper, I will explore how the technological changes affected the way the
production standards of black metal are now understood by members of the
Québec scene.

Defining the basic vocabulary
Before machines could reproduce sound, the term “fidelity” had a very fluid
meaning (Sterne: 2003, p. 216). As Jonathan Sterne indicates, its first use
dates back to the end of the 19th century and:

[it] indicates both a faith in media and a belief in media that can hold faith, a
belief that media and sounds themselves could hold faithfully to the agree-
ment that two sounds are the same sound. (Sterne: 2003, p. 222)

2 However, more recently, Podoshen at al. argued that: “The production of black metal music
can range from very slick – with attention paid to highlighting the intricate musicianship of
band members to muddier, dirge-like production, with churning riffs and unearthly vocals, to
the lo-fidelity, harsh and tinny production in guitars, drums and vocals seen from underground
black metal bands” (2017, p.4). Not coming from a musicological or sound production field of
study, this nuanced posture advocated by authors is quite new in the academic discourse and
echoed to a certain extent the results I have obtained in my own fieldwork.
In other words, in a context where people were still in the process of understanding the relationship between sounds made by humans and those reproduced by machines, a faithful sound meant that it could be understood as “worthy of faith” (p. 274). Rather than focussing on a particular sound aesthetic, sound fidelity was more concerned with reproducing, to the best extent possible, a performance. Alternatively, it was also concerned with distinguishing the sonic features emanating from various machines and technologies (2003: p.274-275).

According to Yannick Lapointe, the term “high fidelity”, commonly known as “hi-fi”, appeared in the 1930s³, where the junction of these two words suggested the desire to be “highly faithful to something” (2012: p.18)⁴. Dario Rudy and Yves Citton also explain that a hi-fi recording is often conceived as equivalent to the live performance, where the recording is then understood to be “true” and “authentic”, capturing the intentions of the composer (Ruby and Citton: 2014, p.113). The notion of “liveness” also remains important for Aden Evens, who adds that “[high fidelity] is achieved by the accurate recreation of the same pattern of sound waves at the listener’s body as would have occurred at the body of a listener who was present at the recording” (Evens: 2005, p.10).

Moreover, the concept of low fidelity (“lo-fi”) originated at the beginning of the 19th century, where an interest in imperfect, rough, simple sounds can be understood as a reaction to Enlightenment aesthetics (privileging perfection and balance). During the 19th century until the beginning of the 20th century, we saw the development of primitivism, romanticism and realism, functioning as a negation of the conventions of musical composition of this period (Harper: 2014, p. 63). In the second part of the 20th century, the desire to reject social conventions with the help of music was also reflected in the folk revival, the rock of the counterculture and independent (indie) rock music. Later, in the 1980s, even though technical imperfections were often understood as a guarantee of more authenticity, an appreciation for sound imperfections was rare (2014: p. 378). Indeed, as Harper points out:

[punk] favoured ‘rawness’ and a disrespect for technocratic convention, but performance imperfections and phonographic imperfections ran the risk of compromising the all-important intensity of the music. (Harper: 2014, p. 377-378)

³ As Lapointe pointed out, the term “hi-fi” (as well as the term “lo-fi”) can be linked with many significations. It can refer to the sound system, recording devices, and can also be understood as an adjective (e.g. hi-fi/lo-fi equipment, hi-fi/lo-fi phonogram or hi-fi/music) or even an aesthetic conception (Lapointe: 2012, p.19-26). In this paper, the latter signification will be privileged.

⁴ Citation translated from French by the author of this paper.
It was truly in the 1990s that lo-fi was recognized as a category, particularly with indie rock and grunge, and it was linked with the disillusionment of society felt by Generation X (2014, p. 378).

Then, for Harper

Lo-fi as it is normally understood is less a genre or mode of music-making than a confluence (rencontre), in the reception of certain recordings, of various aesthetic currents that run through the late-twentieth and early twenty-first century – known under further headings such as primitivism, realism, postmodernism and archaism – and their manifestation in the changing relation of portable magnetic-tape-recording technology to the wider landscape of music-technological (and particularly industrial) production. (Harper: 2014, p.5)

This specific sound aesthetic is often linked to a set of unwanted noises that he calls “lo-fi effects”, such as harmonic distortion, tape hiss, the noise of unplugged cables during the performance (2014, p.23), performance noises (e.g. mouth clicks), or the environmental noises (e.g. unwanted voices in the background) (2014, p.23-27).

Kromhout also focussed on the way the lo-fi production was understood by the media and artists. He explains:

[…] for Pitchfork, the term lo-fi also refers to a loose genre or conglomerate of genres: artists who deliberately reject the standards of hi-fidelity (polished, crispy, clear, ‘accurate’ sound) and share a recording aesthetic focused less on the most accurate recording. Opposed to the ideal of absolute transparency of hi-fi, lo-fi aims at non transparency; it is noisy, ‘inaccurate’ and consciously establishing itself as a recording. (2009, p.1)

This type of sound imperfection can be achieved using obsolete or broken equipment or by not using the proper recording device. However, Kromhout additionally pointed out that the term “lo-fi” remains more or less difficult to define, because there are

[…] many different kinds of lo-fi recordings and many different stages of lo-fidelity, from recordings completely drowned in noise to predominantly hi-fi recordings incorporating lo-fi elements. (2009, p.3)

5 According to Wadhams, harmonic distortion can be described as: “the unwanted addition (by an acoustic environment or electronic device) of harmonics of a pure tone when that tone is propagated in the environment or input to the device” (1988, p.65). Harper adds that: “it typically occurs when a signal is amplified beyond the dynamic range of a device – ‘overdriven’ or ‘overmodulated’ – and thus clipped” (2014, p.20; original emphasis).

6 Original emphasis.
To this, Harper adds that when referring to popular music, the sound quality of a recording cannot be fundamentally good or bad, nor lo-fi or hi-fi. Rather, it is understood as being good or bad in relation to the technological environment from which it arises (Harper: 2014, p.11-12).

However, as Université Laval’s sound engineer Serges Samson pointed out during an interview, this classification of aesthetic productions may also differ depending on the style or genre. For example, the lo-fi production of blue grass remains difficult to compare with the lo-fi production of black metal, meaning that we must establish what is hi-fi and what is lo-fi for a specific style of music (Interview, October 30 2017). This also means that these aesthetic categories remain highly subjective, making it even more important to consider how they are understood by members of the various scenes.

Another important ingredient, which is present in both hi-fi and lo-fi aesthetics, is the notion of authenticity. More specifically, Allan Moore specifies that authenticity is not present in music as such but is created during the act of listening (2013, p. 265). Then, he adds, it must be seen as being “ascribed to” rather than “inscribed in” a recording or a performance (2002, p. 220). It must also be understood as “a matter of interpretation which is made and fought for from within a cultural and, thus, historicised position” (Moore: 2002, p. 210). Moreover, he explains that there are several types of authenticity. Authenticity can be used to reflect two elements, being the “purity of practice” or the “honesty in experience” (2002, p. 213). On the one hand, authenticity is related to purity when fans and musicians alike refer to the origin of something and see it as being genuine and more honest than the use of a contemporary musical language, which would rather participate in the disintegration of the musical style in question (Moore: 2013, p. 263). On the other hand, authenticity is linked with honesty, when it comes to a song understood as being unmediated, or composed “on the spot” (Moore, 2002, p. 213). For example, Norwegian black metal bands of the 1990s tend to associate themselves with Venom, an older band of the NWOBHM, which dwells on punkish aesthetic, a much simpler musical aesthetic than the death metal aesthetic (I will come back to this element).

Finally, authenticity is not always related to the sound. Indeed, it is also possible to speak about social authenticity, as Moore explains when a defined community legitimizes this value as a fundamental characteristic (Moore: 2002, p. 215). For example, in the context of black metal, Danish researchers Kuppens and van der Pol explain that this value is usually understood in the context of the black metal community as:

---

7 The New Wave of British Heavy Metal (NWOBHM) is a group of heavy metal bands emanating from Great Britain in the 1980s (e.g. Iron Maiden, Def Leppard, Motorhead and Venom).
[…] an absolute and inherent quality, [which] is never uncontested, but rather subject of a continual political struggle between parties who seek to ‘naturalize’ their particular definition of authenticity. (Kuppens et van der Pol: 2014, p.152)

This value can be used by the community to evaluate the discourse or to describe the persona or music of a band (2014, p. 152). In that case, even if “authenticity” must be understood as a social construction, it remains nonetheless a value to consider when qualifying the aesthetic of black metal.

Members of the black metal scene in Québec whom I interviewed, described hi-fi and lo-fi in black metal using mostly the same vocabulary and concepts as the definition given above. For example, Fiel (drummer of Forteresse, the main band of the Québec scene) explained regarding lo-fi that: “The sound will be more boomy […] and it’s going to reflect the sound that we can get with a tape recorder” (Ethnographic interview, November 2, 2017). They also pointed out the fact that neither one nor the other aesthetic production seems to characterize their sound. As Moribond (a founding member of Forteresse) added:

The lo-fi aesthetic still has its share of followers, but I find that there is a lot of modern projects that are venturing into the realm of higher production values (such is often the case within the "orthodox" black metal current, for example). In a way, this has brought to light the possibility of having a higher quality production without sacrificing the essence of black metal. (Ethnographic interview, November 2, 2017)

For his part, Thorleif (leader of the band Délétère), describes the sound of the scene as being a “lo-fi of quality”. While being perfectly aware of the pleonasm introduced with this expression, he keeps using it because he does not know how to designate this aesthetic choice (Ethnographic interview,

8 In Québec, the black metal scene was established in the mid-1990s. However, the lack of financial and promotional resources limited the development of the scene. It was only really in 2006 that a specific black metal scene was formed, when the Montreal label Sepulchral Productions signed three emerging black metal groups, namely Forteresse, Gris and Sombrès Forêts. The members I have interviewed are emanating from a specific portion of the scene, known under the moniker of “métal noir québécois”, which came into use also around 2006. The groups belonging to this specific community are also using similar elements in their music, that is, French lyrics, references to the history and culture emanating from Québec, and incorporation of folk music in black metal songs. The community also advocates a sovereignist position regarding the future of the province and the main bands associated with this movement are Forteresse, Brume d’Automne and Chasse-Galerie (St-Laurent 2019. “Le métal noir québécois: Analyse du récit phonographique et identitaire d’une scène marginale”. To be published).

9 Pseudonyms have been used to designate the members of the scene who were interviewed in order to respect their anonymity.

10 The orthodox black metal current refers to black metal bands (e.g. Ofermod, Malign, Deathspell Omega and Watain) that are employing lyrical themes linked with satanism and occultism, while usually advocating an aesthetic sound production of higher quality.
December 16, 2016). Fans and musicians of the Québec scene, such as Stan (founder of the band Tsar Stangra), describe the aesthetic production of the Québec bands as being “kind of in the middle” and being “neither too dirty nor too clean” (Ethnographic interview, October 3, 2016). Finally, Moribond came to a similar conclusion when he told me:

Québec black metal bands are not going to be searching – the most important groups at least – for a super lo-fi sound, nor a super hi-fi sound. The most important for us is the texture and the ambiance that we are conveying on each recording. (Ethnographic interview, November 2, 2017)

It is after many discussions with members of the scene, academics and sound engineers, that I have come to use the term “mid-fi” to designate this aesthetic of the in-between characterizing the Québec scene. “Mid-fi” is not used in the academic literature surrounding metal music and sound production. However, this term is often employed in various audiophile forums to refer to recording equipment of good quality that is not overpriced. Therefore, I define “mid-fi” as a recording aesthetic of good fidelity, while not being completely free of sound imperfections, and that can satisfy most listeners. This preliminary definition also reflects the way most of the scene members in Québec conceive of their sound, as shown earlier.

In the following section, I explore the categorization of hi-fi and lo-fi productions in the context of black metal and examine to what extent it is possible to use the term mid-fi to qualify the Québec scene. To do this, I use frequency analysis and spectrogram images representing portions of the recordings considered to be significant. The frequency analysis will illustrate the frequency range at specific moments in selected excerpts and the spectrograms will be used to demonstrate the level of definition throughout the duration of chosen excerpts. These tools allow me to place each excerpt on a left-right axis referring to lo-fi and hi-fi aesthetic production.

The black metal context: exploring the lo-fi, hi-fi and mid-fi aesthetic sound production

As mentioned above, academics have defined the black metal style as belonging to lo-fi aesthetic production. When doing so, they often refer to the first recordings of the Norwegian black metal bands of the early 1990s, including the Darkthrone’s Transilvanian Hunger’s album. This album, recorded in 1994 at the Necrohell studio (Norway) using a four-track recorder, is often considered to be the lo-fi aesthetic reference for black metal fans. It has also been associated with the “necrosound” aesthetic, since the musicians wanted their sound to contrast as strongly as possible to the hi-fi aes-
thetic production used by technical death metal bands of that moment (Reyes: 2013, p.255; Aites and Ewell: 2009, 13:45). As explained by Reyes:

[black metal] pursued not just any kind of amateur, low-fidelity work. Rather, it reified only particular sounds, creating an ugly, raw, grim aesthetic specific to the dominant sensibilities of extreme metal at the time. (2013, p. 247)

By using this aesthetic, black metal bands also wanted to reproduce the sonority of the proto-black metal bands of the 1980s, which also used low-budget production (e.g. Venom, Bathory, Celtic Frost, etc.). These bands were also associated with a low-budget type of recording, since their albums were produced with minimal equipment and within a short time (Reyes: 2013, p.250).

Therefore, the first excerpt chosen for this analysis is “Transilvanian Hunger” from Darkthrone. When listening to this song\(^{11}\), we notice that the sound is thin and that the spam of frequencies reproduced seems very small. The sound of the instruments, and particularly the drum, are characterized by the lack of sharpness of the attacks. The voice is positioned relatively at the back of the mix and is very distorted. According to Samson, this is probably due to the poor quality of the audio recording and the obsolete state of the recorder (Interview, November 7, 2017). The next image (figure 1) represents a screen capture of the frequency analysis of the targeted song, using 400 band frequencies with Pyramix software. With each of these images I present only a specific portion of the excerpt, which appears at the top left corner (here the screen capture was taken at 31:45 seconds from the beginning). The horizontal axis represents the frequencies covered between 20 to 20000 Hz, and the vertical axis depicts the decibel level reached. I focus mainly on the data presented on the horizontal axis. On this first image of the frequency analysis, we can see that the frequencies range between 150 and 5000 Hz (if we do not consider the normal slopes of the frequencies at both extremes). What is also interesting to mention is the presence of three larger spikes after 5 Kh. According to Samson, they can be related to sound arte-

\[\text{Figure 1: Frequency analysis image of an excerpt of “Transilvanian Hunger” from Darkthrone}\]

\(^{11}\) It is possible to listen to the song “Transilvanian Hunger” from Darkthrone through Youtube. The address is the following: https://www.youtube.com/watch?v=4iaj2w7Bp58
facts\textsuperscript{12}, which are usually linked with a sonic error due to the restriction in the audio capture of a recording device (Interview, November 7, 2017).

To have a better idea of how this lo-fi aesthetic can be translated visually with each excerpt, I also used spectrogram images, realized with Sonic visualizer software. On these images (e.g. figure 2), the horizontal axis refers to the time progression in seconds, while the vertical axis indicates the frequencies (from 43 Hz to more than 20 kHz). The different colors used allow viewers to identify the energy level of certain frequencies. Thus, the yellow bands on the bottom refer to the frequencies most present in this excerpt. The irregular orange curves that can be seen in the lower right corner represent the vocal frequencies.

This first spectrogram image also informs us on the level of definition\textsuperscript{13}. Indeed, the vertical lines – namely the transients –, which usually refer to guitars or cymbal attacks (as shown in the red oval) are not prominent. On the contrary, everything seems to be very smooth, which characterizes the absence of definition. Moreover, the faint lines that can be seen after 10 kHz are the artefacts that were discussed earlier. This spectrogram image is also typical of a black metal lo-fi aesthetic. Indeed, as Hainaut explained, black metal sound is highly defined by the settings of the guitars. Therefore, the high gain will tend to increase the upper partials of the sound, and consequently, the mid and lower frequencies will be less prominent. This will produce a highly homogeneous harmonic spectrum, where the energy will not be focused on specific formants but will be spread more uniformly across the spectrum – as it is shown in this first spectrogram image (2012, p.99).

If for many academics and sound producers, this production has been labeled as the example of all things you must not do when producing an album, for black metal fans, this is understood as an aesthetic statement of great importance. Thus, in the context of black metal, this type of aesthetic is often heard as being authentic, since it is intended for an audience of insiders, who understand its codes. Many black metal bands

\textsuperscript{12} Samson also points out that the presence of such artefacts is not necessarily the result of the recording process. Depending on if the song was digitalized afterwards, these artefacts could have appeared at the mastering stage (Interview, November 7, 2017).

\textsuperscript{13} “Definition” can be defined as the level of sharpness of an image.
today, such as Akitsa from Québec (e.g. Grands tyrans, 2015), continue to use this kind of low-budget/lo-fi aesthetic. Thus, given this information, I will position this excerpt on the extreme left side of my axis, as being an example of extreme lo-fi aesthetic in the black metal style (figure 3).

During the beginning of the 1990s, other groups of the Norwegian scene (e.g. Burzum, Emperor, Mayhem) recorded several albums that were also understood as being lo-fi by the black metal fans. However, they were slightly better produced than the Darkthrone excerpt I just presented. More specifically, these bands recorded their albums at the Grieghallen studio (Bergen, Norway), where they worked with sound engineer Eirik “Pytten” Hundvin.

The second excerpt, “I am the Black Wizards” comes from Emperor’s first album, In the Nightside Eclipse, recorded in 1994 with Pytten at Grieghallen Studio. The comparison of this excerpt with those of three other bands also produced by Pytten (namely Mayhem, Immortal and Burzum), reveals that the excerpt by Emperor seems to reflect the production aesthetics specific to these groups.14 Listening this excerpt15, we can hear that the production has a little bit more fidelity than the last one, since the lower frequencies are more prominent. We can also quite easily hear that all the instruments and their definition is greater (e.g. we can hear each cymbal attacks distinctively). However, the global sound remains muddy, which could be due to the omnipresence of reverb added on all the instruments.

The frequency analysis (figure 4) shows us that the frequency range is already much wider than the Darkthrone excerpt, extending from 80 to 8500 Hz, and where the curve is more regular, having a slight hollow between 300 Hz to

---

14 It’s important to note that the subsequent albums of these bands have shown a higher degree of fidelity, which can be due to several factors, such as the preferred aesthetic choices or even the greater experience of the musicians. For this analysis, I’m referring to the albums produced in the beginning of the 1990s, which better characterized the lo-fi sound we can hear during this period in Norway.
15 It is possible the listen to the song “I am the Black Wizards” (Emperor) using this Youtube link: https://www.youtube.com/watch?v=YgQRR19goFg
1 kHz in order to leave frequency room for the keyboard (which is featured later in the song).

The spectrogram (figure 5), meanwhile, confirms the greater definition mentioned before, which is immediately visible thanks to the many transients we can clearly see, and which represent cymbal attacks. We can still see a strengthening of the frequencies between 1000 and 3000 Hz, where the definition seems less present, due to the lack of sharpness of the transients.

Even if this production has a greater level of fidelity, the lack of data prevents us from classifying it definitively on the axis. In the meantime, I propose to understand this recording as an example of a well produced lo-fi and place it momentarily between the lo-fi and mid-fi poles (figure 6).

I now explore the other end of the spectrum. From the mid-1990s, black metal began to be known outside of the underground scene and many other groups broadened its aesthetic production standards. Therefore, in the early 2000s, several black metal bands developed a sound production of a higher quality, as is exemplified by the Norwegian group Dimmu Borgir. Indeed, from the beginning of the 2000s, the group softened its black metal sound, even including a complete symphonic orchestra (e.g. *Death Cult Armageddon*, 2003). Many other bands also followed a similar tendency, such as Anorexia Nervosa (e.g. *New Obscurantis Order*, 2001) or Cradle of Filth (*Thornography*, 2006). The excerpt I chose for analysis, “Born Treacherous”, comes from the most recent album of Dimmu Borgir, *Abracadabra* (2010)\(^{16}\). When listening to the song, we can feel a more uniform distribution of all

\(^{16}\) It is possible to listen to the song “Born Treacherous” (Dimmu Borgir) using this Youtube link: https://www.youtube.com/watch?v=6hJmg4OyA1w
frequencies and especially a greater presence of the lower ones (e.g. the bass guitar is clearly heard compared to the previous examples).

Looking at the image of the frequency analysis (figure 7), we can identify a very large frequency range, from 80 Hz to 6 kHz, but with only a significant reduction at 10 kHz, which is more extended than with the two previous recordings. The peaks we can see in the left portion of the image refer to a strong attack on the bass guitar in the lower frequency range.

The spectrogram image (figure 8) clearly demonstrates distinct transients and a reinforcement of the mid-high frequencies up to 6 kHz. This image even allows us to distinguish the transients at the places where the frequencies are the most concentrated (between 2000 and 3000 Hz), which illustrates here a recording of high fidelity. The greater repartition of the yellow color (not showing separated bands as much as in the other spectrogram images), also illustrates the better distribution of all frequencies in the spectrum. Finally, the spaced transients on the left side of the image represent the bass guitar attacks mentioned earlier. I thus position this excerpt on the extreme right side of the axis, as being a hi-fi production (figure 9).

Even if this hi-fi aesthetic ensured a larger popularity for Dimmu Borgir outside of the black metal scene, it had a rather devastating effect in the underground scene. Indeed, fans began calling this group “sell-outs”, which refers to:

the process by which artists or songs sell beyond their initial market which, in turn, loses its sense of possession, exclusive ownership and familiar belonging … ‘selling out’ means selling to outsiders. (Thornton: 1996, p.124; original emphasis)

Some groups, such as Moribond explained, managed to maintain their underground status despite their aesthetics of better fidelity. This is the case with bands Dissection (Germany), Bexehen (Finland) and Watain (Sweden). Regarding the latter, most of its albums were produced by Necromorbus Studio, located in Stockholm. The aesthetic production of Watain is described by Moribond as being “not entirely clean but offering a higher level of dynamic” (Ethnographic interview, November 2, 2017). In this regard, Mgła (e.g. Exercise of...
“It’s kind of in the middle”: The ‘Mid-Fi’ Aesthetic: Toward a New Designation of Black Metal Aesthetic of Recording. The Case of the Québec Black Metal Scene of recording

Futility, 2015) seems to also be associated with this kind of hi-fi aesthetic, while remaining well regarded in the black metal scene internationally.

However, Watain is also known to offer extremely transgressive performances during their shows17, remaining “authentic” to the eyes of the fans and musicians but also preventing the group from gaining popularity outside the underground, despite the better quality of its recordings.

The selected excerpt from Watain is the song “De Profundis” from the album The Wild Hunt (2013)18. Listening to the excerpt, we can hear that the reverb is mostly present on the drums and voice. All instruments are audible, and their positioning seems well-balanced, even if the bass frequencies are less prominent here than in the previous excerpts. Looking at the frequency analysis (figure 11), we can see it covers 35 Hz to 10 kHz, which means that the frequency range is greater than in the previous example.

As in the last spectrogram image, frequencies are also well distributed, and we can distinguish a precise degree of definition thanks to the many small transients we can see all over the spectrum (figure 12).

Thus, even though the spectrogram of Dimmu Borgir showed a slightly better definition, I also place the Watain’s excerpt close to the right end of the axis (figure 13).

Having explored both ends of the axis, it is now possible to examine the bands of the Québec black metal scene to understand why their aesthetic is best described as “mid-fi”. I will focus here on two excerpts: “Le lai de la vermine” from Délétère (Per Aspera ad Pestilantium, 2017)19 and “Là où

---

17 They are known to throw pork blood on the crowd – which they are covered in as well – and to play next to rotting carcasses of animals, causing nauseous smells.
18 It is possible to listen to the song “De Profundis” (Watain) following this Youtube link: https://www.youtube.com/watch?v=7CT3eTBk85I
19 The song “Le lai de la vermine” (Délétère) is available through this Youtube link (at 01:37): https://www.youtube.com/watch?v=HVzqVZe48R1
nous allons”, from Fortresses’ last album (*Thèmes pour la rébellion, 2016*). These excerpts illustrate two types of production used in the scene. Indeed, bands such as Délétère, Csejthe or Forteresse (with its first albums) are auto-produced, and others, such as Monarque, Chasse-Galerie and more recently, Forteresse, use the help of professional studios. I also chose these excerpts because these two bands are labelled by fans as having an aesthetic production that is “in the middle”. When listening to Délétères’ song, we first notice that all the frequencies seemed to be taken in one block, since the recording doesn’t have a large variation in frequencies. Indeed, the mid and lower frequencies seem less clearly transmitted compared to the higher frequencies of the cymbals, which can be heard really clearly. However, the voice and instruments seem well defined, since we can clearly distinguish each word or instrumental parts.

Looking at the frequency analysis image, we can see that the main frequencies of this recording are situated between 100 Hz and 5000 Hz, which is slightly similar to the Emperor’s excerpt (figure 14).

The spectrogram (figure 15), on the other hand, shows a high level of energy between 2000 to 4000 Hz, reflecting the sound block mentioned before. It also allows us to see clearly each attack of the cymbals in the higher frequencies, thanks to the many defined transients. However, it also demonstrates that the lower frequencies are less defined, since transients are not discernable in that portion of the image.

The last excerpt, “Là où nous allons” from Forteresse, was produced by Necromorbus studio in Stockholm. In this regard, Fiel and Moribond told me that by making the choice to be professionally produced for the first time, they wanted to give a more powerful sound to their music, without sacrificing the aggressiveness of their black metal sound. They also explained that they wanted to achieve a “wall of sound” effect (not to be confused with Phil Spector’s “wall of sound”) (Ethnographic interview, November 2, 2017). Listening to the excerpt, we can hear quite distinctively this “wall of sound” in the way the guitars are positioned on each side, playing the same melody and using

It is possible to access to the song “Là où nous allons” (Forteresse) at the following Youtube link: [https://www.youtube.com/watch?v=ACVOpYnmSg0](https://www.youtube.com/watch?v=ACVOpYnmSg0)
the same harsh sound. The lower frequencies are also much more present in this excerpt than in the previous one and all instruments remain clearly audible. However, the voice seems to be positioned a little bit more at the back of the mix than in the last example.

The frequency analysis image of this excerpt (figure 16) presents a frequency range extending from 80 Hz to 7 Khz, that is, it is slightly narrower than the Watain excerpt. However, the range of frequencies remains larger than Délétère’s excerpt.

The spectrogram (figure 17), for its part, shows a higher concentration of energy in the mid-high region of the spectrum (the different yellow bands), exemplifying the “wall of sound” effect. However, by doing so, it also removes a little bit of definition to the recording. Indeed, thick yellow bands can be seen up to about 5000 Hz, masking the transients of this region. Besides that, we still can see a considerable number of transients in the higher area of the image.

This short analysis allows us to have a more nuanced idea of the place of each excerpts on the lo-fi/hi-fi axis (figure 18). We can see, not surprisingly, that Darkthrone’s and Dimmu Borgir’s excerpts present the widest differences, one clearly belonging to the lo-fi and the other to the hi-fi aesthetic production. Emperor’s and Délétère’s excerpts, with their slightly better production, are located closer to the middle of the axis, where Délétère seems to present an aesthetic of a slightly better fidelity than Emperor, who remained closer to the left side. Finally, Watain’s and Foreress’s excerpts are extremely similar, probably since they were produced at the same studio. However, the difference can be found in the lack of definition in the mid-high region of Foreress’s excerpt, which lessens the fidelity. Therefore, Watain can be placed close to the right side of the axis, while I place Foreress at the junction point of the mid-fi and hi-fi categorization.
The technological advances and the changes in the hi-fi/lo-fi aesthetic conceptions

Changes in recording technique during the last twenty years have influenced members of the scene in their aesthetic conception of black metal sound. Indeed, according to Auster, a sound technician and musician who has been involved in the scene – first as a musician, and later as a sound engineer – for more than 20 years, the biggest difference between the first black metal productions and the latest ones is linked with the improvement of technology, the presence of a “standard lo-fi sound” and the greater expertise of the sound engineers. The lo-fi black metal aesthetic of the 1990s was also linked with the absence of norms and formats. More specifically, he recalls:

The production trends were probably less formatted at the time. There wasn't a standard on how black metal should sound. Well, there was some trends but the movement was starting, so engineers and bands were going in lots of different directions. Some albums sounded muffled and others really bright or thin. Effects were all over the place and the sound levels varied a lot too. [...] So, there was more color or sonic texture in the recordings from one album to another and from one band to another. (Ethnographic interview, December 12, 2017)

Today, as he explains, there is a greater neutrality in the sound production because of the higher quality of the equipment. Another important element to consider, according to Auster, is that today’s technicians are more qualified when it comes to producing black metal artists, which also influences the quality in the aesthetic productions. By that comment, he means that producers at first didn’t know how to elaborate this specific sound and atmosphere, rendering a sound that was too clean for what bands expected. However, as he points out, the aesthetic choices have not changed dramatically. Indeed, as he mentions:

[...] I don't think most the modern black metal bands changed their sonical requirements. Lots of them are still aiming for this lo-fi sound. But with today's recording equipment, the sonical results changed. I think the most significant change was the transition from analog to digital. Now, there is less color, noise and imperfections coming from the machines, so it sounds cleaner. As everything sounds flatter and closer to reality, this means that if you want a more textured sound, you must craft it consciously as a choice. (Ethnographic interview, December 12, 2017)

Many other members also agree with the importance of technological development in the aesthetic production of black metal, such as Thorleif. According to the latter:
By downgrading and decreasing the quality of your production, but in a thoughtful way, you make a choice. It gives a color to your music and it is used to convey an emotion. (Ethnographic interview, December 16, 2016)

Moreover, when I told Auster about the way the scene members were considering their music (as being “in the middle”), he explained:

I understand what they mean by that. It’s about finding a balance between a lo-fi production and a good production, where you hear everything. I have to agree with this. When instruments tones are too colored the audibility is affected. In extreme cases, you don’t even distinguish which instrument you’re hearing in different frequency ranges. [...] Some recordings were raw to a point of becoming a puddle of indistinct sound mud. It takes multiple listening to actually grasp the music in this sonic chaos. While in a good production, everything is in its place… But when everything is too much in a defined spot without any overlap, it can sound fake. You’ll miss the life of the instrument and you can’t believe it’s a live band anymore. I think sounding fake is what most black metal bands are trying to avoid. (Ethnographic interview, December 12, 2017)

What emerges from this comment is the insight that, according to the standards of black metal, a good, hi-fi production is heard as being fake; however, producers no longer aim for a lo-fi sound either. In addition, these new technological possibilities have also had a negative effect – meaning, an effect that is understood as negative by members of Québécois black metal bands. Antoine (drummer for the bands Monarque and Délétère) and Sébastien (leader from Sorcier des Glaces and sound engineer) explain, technological advances gave the opportunity for everybody to reproduce the lo-fi sound in their bedroom and transformed the specificity and the originality of lo-fi into a trend (Ethnographic interviews, October 31, 2016; December 9, 2016). This establishment of a tendency goes fundamentally against the value that scene members share, that is, the willingness to stay underground by avoiding any kind of aesthetic trends. Therefore, the mid-fi sound seems to have allowed them the possibility of exploring other aesthetics without scarring their authenticity and originality.

**Conclusion**

To conclude, I return to the two paradigm shifts I mentioned in the introduction. First, I claimed that black metal music cannot be understood anymore as being associated with only lo-fi aesthetics. In this article I demonstrated (with many examples) that the aesthetic production of this style is much more diverse than what academics have accounted for, to date. Thus, I propose to use the term “mid-fi”, which reflects the way the members of the black metal scene in Québécois understand their sound. Thereafter, I argued
that it is possible to qualify the aesthetic production of Québec scene as being “mid-fi”. At this point of my study, I am making only an initial claim, since many more analyses must be made to distinguish the presence (or the absence) of a tendency in the Québec black metal scene. However, according to the chosen extracts we can associate the mid-fi aesthetic with a sound presenting the harsh and boomy characteristics of lo-fi, but with certain hi-fi features, specifically the definition of the attacks and the possibility to clearly hear each of the instruments used.

In order to pursue and improve this analysis, it is also important to acknowledge its limitations. Indeed, even if the tools used can provide more details regarding the frequencies being represented, they cannot fully replace simply hearing of the examples (eg. a spectrogram might present a lot of transients even if the recording sound fairly lo-fi). Also, in order to better select future examples, it will be important as well to consider the fact that bands improved through the years (they learned more about sound production and they gained better control of their instruments), which greatly influenced the quality of their productions. For example, the first albums of Forteresse (e.g. Métal noir québécois, 2006; Les hivers de notre époque, 2008; Crépuscule d’octobre, 2011) can be considered as belonging to lo-fi aesthetic, as can the first album of Déletère (Les heures de la peste, 2015).

The second paradigm shift is linked with the way the new standards used by many contemporary black metal bands suggest that the determination of authenticity is no longer linked with lo-fi aesthetic production. Indeed, I have shown that a lo-fi aesthetic has been used extensively by many bands, which transformed the exclusiveness of this sound into a tendency. This abundant use seems to have modified the way authenticity is perceived by scene members, meaning that it does no longer refers to the simple polarization of lo-fi = authentic/hi-fi= inauthentic. However, this opposition is not solely applicable to black metal but is as a matter of fact present in all the history of popular music. Indeed, as Moore suggests:

[…] the distinction between ‘authentic’ and its opposite (‘entertainment’ at some times, ‘commercial’ at others) underpins the history of popular music for the time of Elvis Presley onwards, and that such a history proceeds as a pendulum, swinging from one extreme to the other, frequently with much disagreement among fans and critics as to which term to apply to which music – again such attributions are to be fought for (Moore: 2002, p. 211).

In this context, if the need for black metal bands to build a very distinct aesthetic from death metal was at a prerogative in the beginning of the 1990s, it does not seem to have as much importance today, where we find more and more crossovers between styles. Indeed, the aesthetic codes that were previously reserved for black metal are now used by death metal and
vice versa (eg. we can find several death/black groups, such as Belphegor or Behemoth). Moreover, my analysis of Forteresse’s excerpt also showed a higher fidelity than what was expected based on interviews. Indeed, the fact that members of the Québec scene are referring to Forteresse – and to other bands of the Québec scene that have an even better production (e.g. Monarque, Chasse-Galerie) – as having a “mid-fi” aesthetic, implies that scene members are developing more nuanced perception of what it means to be “authentic”. This could be explained by the fact that even if the lo-fi aesthetic is not as related as before to a higher level of authenticity, a good production is still perceived as inauthentic, up to a certain point. Indeed, describing these recordings as belonging to a hi-fi aesthetic could lessen the way in which the group is esteemed by underground fans of black metal; therefore, qualifying them as “in the middle” protects the legitimacy and the authenticity of the scene.

Finally, although this paper is oriented towards a specific scene, I believe that black metal in Québec is exemplifying a more general tendency in the contemporary aesthetic production of black metal around the globe. More analysis in that direction will be done in future studies.

Acknowledgments

I would like to express my gratitude to the following persons for helping me to complete this paper: Serges Samson, Gérald Côté, Jacqueline Fortier and Heather MacLachlan.

References

Bibliography


21 However, as Moore pointed out: (…) this commercial/authentic polarity is illusory, since all mass-mediated music is subject to commercial imperatives, but what matters to listeners is whether such subjection appears to be accepted, resisted, or negotiated with, by those to whom they are listening (Moore, 2002, p. 218).


**Interviews**

Antoine. (2016) Ethnographic interview realized on October 31th.


Sébastien. (2016) Ethnographic interview realized on December 9th.


Thorleif. (2016) Ethnographic interview realized on December 16th.

**Discography**

Anorexia Nervosa. Osmose productions, 2001


Paul Thompson & Phil Harding: A ‘Service’ Model of Creativity in Commercial Pop Music at P&E Studios in the 1990s

Abstract

Producers in pop and dance music genres have a significantly different role to music producers in other music genres such as rock (Frith in Frith & Zagorski-Thomas: 2012). A prominent difference is that pop music producers are often part of a production team that involves direct collaboration and participation with songwriters, programmers, musicians, artists, record company A&R executives and managers. Pop music songwriting and production teams are therefore more frequently part of a larger creative collective (Hennion: 1990) in creating a musical product. This historical study introduces the record producer as ‘team leader’ and the creative production workflow at P&E Music Studios located within the Strongroom Studio complex in London during the 1990s. It investigates the ways in which the production team worked within the creative system of pop-music making and presents the pop music ‘Service Model’, which illustrates the various stages of the commercial pop songwriting and production process at P&E during the 1990s.

Introduction

Simon Frith stated ‘the aesthetics of the popular [music] continues to be at best neglected and at worst dismissed’ (Frith: 1996) and 20 years on, the study of how popular music is made is now only beginning to be addressed in some areas of scholarly literature. Pop music in particular, with its labels of ‘manufactured’ or ‘teen’, has been overlooked as a valid area of popular music research because it is so often viewed in the popular imagination as an: ‘inauthentic exploitation of the masses’ (Gracyk: 1996, p.175-176). However, as Theodore Gracyk points out, even musical genres that are considered to be more closely related to art-making, such as rock music, take place: ‘in a system of music making and distribution that has been commercial since its inception’ (Ibid: p.179). The system of music-making in pop music typically involves a production team and, although the artist takes
centre stage, a team of music producers, songwriters, programmers, musicians, record company A&R executives and managers often work behind-the-scenes in collaboration with the artist.

Andrew Blake maintains the term ‘record producer’ is the: ‘greyest of grey areas’ (2009: p. 36) because of the ambiguity surrounding the role. The record producer has been characterized as ‘nexus’ (Howlett, 2009) because the record producer: ‘acts as a means of connection between the artist, the technology and the commercial interest’ (Howlett: 2009, p.1). Richard James Burgess (2013) has categorized the role of the music producer in terms of their functional typology identifying six broad categories: ‘Artist-producer, Auteur-producer, Facilitative-producer, Collaborative-producer, Enablative-producer, Consultative-producer’ (Burgess, 2013: p.9). Within these functional typologies, record producers have to deploy a vast array of skills using their background, experience and knowledge. Record producers:

‘have been (and are) individual entrepreneurs, freelance operators, record label owners and record label employees. They have been people managers, whether Svengalis, artist and repertoire developers, or gifted amateur psychologists able to guide temperamental artists through a recording session. They have been events managers...They have been music managers: session fixers, composers, arrangers, synthesiser and drum machine programmers, and conductors’ (Blake, 2009: 36)

In contemporary commercial pop music, the central producers at Chieron studios, Denniz PoP and Max Martin, have been characterized as ‘song machines’ who work within ‘hit factories’ (Seabrok, 2015); equating them to the Tin Pan Alley writers in NewYork’s Brill Building, Motown and the UK’s own Stock, Aitken and Waterman (SAW) ‘Hit Factory’ of the 1980s. Although the record producer is acknowledged within the majority of these studies as requiring leadership qualities, none have identified or explored the record producer’s role as ‘team leader’. The team leader producer can be likened to the ‘entrepreneurial producer’ (Howlett: 2009) which describes the person who:

‘initiated, facilitated and inspired the production of a large amount of music and without whose talents the corpus of popular music would be the poorer. Prime examples would include Chris Blackwell (Island Records), Richard Branson (Virgin Records), Clive Calder (Zomba/Jive) and Dave Robinson of Stiff Records (Howlett: 2009, p. 24).

The following study therefore contributes to the further characterisation of the record producer as ‘team leader’ by exploring the creative production workflow, and the involvement of team leader Tom Watkins, at P&E Music Studios during the 1990s. Beginning first by outlining the design of this study, the creative system is introduced and each of the elements are re-contextualized so that they apply to the context of pop music production.
Finally, the agents with the process are discussed and the pop music ‘Service Model’ is depicted, illustrating the stages of the commercial pop songwriting and production process at P&E during the 1990s and the involvement of the team leader producer.

Methodology

This historical study draws upon a series of semi-structured interviews with those involved with the production or mediation of pop records during the 1990s at P&E Music Studios located within the Strongroom Studio complex in London. Participants included Ian Curnow, who was a programmer and producer at P&E Music; Tom Watkins artist and producer manager for East 17; pop music journalist Matthew Lindsey and pop music songwriter John McLaughlin. Interviews were conducted from May 2014 to December 2015 and focused upon the practices and processes within the P&E Music Studios from 1992 to 1999. The interviews were recorded using a Dictaphone and later transcribed and analyzed for common themes, ideas and observations. As one of the authors of this paper, Phil Harding began working at P&E Music Studios as co-owner and producer from 1992 and was involved in recording and mixing all of the artists that worked at P&E Music Studios through to 1999. These artists included East 17, Boyzone, Deuce, OTT and 911. Harding’s autoethnographic reflective data such as personal diaries, alongside press articles, sound recordings, and information collected for the book *PWL From The Factory Floor* (Harding, 2010) have all been used as data for this study and to support some of the themes and observations highlighted in the interviews because: ‘it is only “with the falling of the dusk”—after the day’s action is done—that reflection and analysis can take place. In the heat of the action philosophy is far from the participants consciousness’ (Howlett: 2009, p.3). For ease of presentation, the majority of responses from interviewees have been paraphrased and integrated into the main body of the text.

Creativity and Pop Music Production

Although commercially driven, pop music production still takes place within an identifiable cultural tradition and involves using the language of this tradition to create something new. The ‘systems model of creativity’ (Csikszentmihalyi: 1988, 1996 & 1999) suggests that creativity is the result of a complex and dynamic system in action. This system has three essential and interconnected elements: (1) a set of symbolic rules, practices and guidelines called a ‘domain’, (2) an ‘individual’ who brings something unique into that domain and (3) a ‘field’ of specialists or experts who recognize and
substantiate that novelty (Csikszentmihalyi: 1996, p.6). In order to create something new, an individual must first acquire knowledge and understanding of previous creative works in that area and learn the content and rules of the domain. Creative individuals must also understand the parameters that control the selection of creative work by the field (Csikszentmihalyi: 1996, p.47) and for an idea or product to be creative it must be valuable to a particular social group (the field), have an element of originality and implemented into the cultural matrix or symbol system (the domain) (Csikszentmihalyi & Wolfe: 2000, p.81).

In order to make a new pop record, the production team needs to draw from the domain to choose a unique selection of elements from this symbol system and then present it to the field for verification (Csikszentmihalyi: 1996). This can be seen in action as a new pop record is commercially released. The field, which is made up of the press, audiences, other artists, engineers and pop producers all comment on the record’s uniqueness or creativity. The field decides whether the record is a creative record or not (economically) either by buying it or, historically, by incorporating it into the domain. The systems model of creativity has also been illustrated on smaller scales, for example on a group scale during the making of the pop record, by first contextualizing the domain and field so that they apply to the

Figure 1. ‘Revised Systems Model of Creativity Incorporating Creative Practice’ (Kerrigan, 2013, p. 114.)
specific context (Kerrigan: 2013; Thompson: 2016) and then observing the interaction of the elements of the system (fig.1). The creative system is therefore scalable and: ‘applies equally well at the individual level and also at the group, organizational, institutional or sociocultural level’ (McIntyre: 2013, p.91).

The Domain of Pop Music Production

So in order to be creative, creative practitioners must first internalise the domain of pop music and pop music production, as this is essential knowledge. The domain of pop music is expansive so for ease of analysis, the cultural matrix and symbol system of pop has been divided into four broad areas: musical, technical, cultural and commercial.

The musical area of the domain in pop is centered on the contemporary western song, its structure and form (i.e. verse chorus, middle eight etc.), its lyrics and lyrical themes, its instrumentation and arrangement. Pop music producers must acquire knowledge of all of these essential parts to the domain. The Technical part of the domain includes the vast array of recording formats, recording and music technologies. This essential knowledge area also includes a working knowledge of computers, samplers, microphones, mixing consoles, room acoustics and monitoring equipment and the ability to use all of these.

The socio-cultural area of the domain has developed in response to the environment of the recording studio in pop music production. This relates to the social and cultural etiquette of collaborating with artists, how to communicate with them, their managers and their record company A&R representatives about musical and technical ideas. Finally, the Commercial area of the domain includes remuneration systems of pop production, royalties and the different types of business deals that exist within the music industry. Pop music producers need to develop an understanding of these in order to continue to operate financially within the field of pop production.

3.2 The Field of Pop Music Production

Mihaly Csikszentmihalyi states the field: ‘includes all those who can affect the structure of the domain’ (Csikszentmihalyi: 1988, p. 330) and in the case of pop music production this includes artists, engineers, songwriters, programmers, record producers, artist management, record label A&R representatives, the pop music press, TV, internet, radio, audiences and social media commentators. These individuals, groups and institutions form the field of pop music production and they not only comprehend and use the domain, but through their process of evaluation of creativity, help to decide what is included in the domain of pop music.
3.3. The Personnel of Pop Music Production

The crucial member of the personnel for a Service Model of Creativity in Commercial Pop Music is the team leader, who needs to possess a high level of cultural, symbolic and economic capital (Bourdieu: 1984) in order to ‘get things done’. At PWL Studios in the 1980s the team leader was Pete Waterman (Thompson & Harding: 2017). In the case of P&E Music Studios, based at The Strongroom in the 1990s, the team leader was Tom Watkins. Working ‘in service’ to Tom Watkins during the 1990s at P&E Music were Ian Curnow (producer, songwriter, musician and programmer), Phil Harding (producer, songwriter, recording and mix engineer), Julian Gallagher, Dan Dodd (studio assistants at P&E Studios), Richard ‘Biff’ Stannard and Rob ‘Fingers’ Kean who were both co-producers and partners to Tom Watkins at Massive Management. In addition to those key studio personnel were session vocalists Tee Green and Andy Caine who regularly worked in the studio. There was also a fully-staffed office for Tom Watkins’s Massive Management Company that began at his home in Maida Vale, London and by 1994 had moved to a commercial office space in Shoreditch, East London. These offices were within walking distance of The Strongroom Studios. The Massive Management office from early 1995 retained a staff of up to twelve people taking care of the day-to-day management of their artists East 17, Deuce, 2wo Third3 and J-Pac as well as the P&E Production team. The lyricist / ‘top-liner’ in P&E Productions would change, depending on the project – importantly though, these top-liners were never Harding or Curnow; the top-liner would always be someone else. For East 17 it was band member, songwriter and rapper Tony Mortimer, for Deuce it was Tom Watkins and Rob Kean and for 2wo Third3 it was Richard Stannard or Rob Kean.

Each role within the production process has varying degrees of creative agency that relates to the power relationships that operate within the specific context and pop music production more generally. For this study we are focusing on the importance and uniqueness of the team-leader, which in P&E’s case was Tom Watkins. When Tom Watkins began collaborating with P&E he was renowned for his previous artist management success with Pet Shop Boys and Bros, both of whom achieved sustained levels of commercial success in the 1980s. Tom had guided their careers from the beginning and, in the early 1990s, was still highly regarded by some UK record company executives as a successful entrepreneur that could find, form, develop and mould commercial pop acts. In other words, he had accumulated notable cultural, economic and symbolic capital.
Cultural, Economic and Symbolic Capital

Shuker (2008) describes ‘music as cultural capital’ by stating that ‘music consumption is not simply a matter of personal preference. It is, in part, socially constructed, serving as a form of symbolic or cultural capital’ (Shuker: 2008, p.181). Cultural, economic and symbolic capital within a pop music production team is deployed in order to manage relations of power and establish a sustainable mode of production that allows creative activity to occur and (hopefully) achieve commercial success. Clearly defined roles and complementary skillsets amongst the agents within the pop production team are therefore necessary for a successful pop music production. The following section highlights the conditions that brought the P&E team together and their cultural, economic and symbolic capital (Bourdieu, 1984).

Pop Music Managers & Tom Watkins’s Capital

Team leader Tom Watkins controlled the P&E team in terms of cultural, symbolic and economic capital. Watkins began his career in the music industry as the manager of UK bands Giggles (signed to EMI Records) and Grand Hotel (signed to CBS Records) in the late 1970s. Watkins then turned his hand to forming a design studio, named XL, which quickly became a favourite of Trevor Horn’s ZTT Records; designing everything from the ZTT studios and offices through to leading the campaign for Frankie Goes to Hollywood in the early 1980s. By 1983 Watkins had returned to artist management, discovering and signing Pet Shop Boys and then Bros.

Watkins’s involvement was both unusual and distinctive for this period of pop music because he managed his pop acts as well as his producers Phil Harding and Ian Curnow. Other managers such as Louis Walsh (Boyzone and Westlife) and Nigel-Martin Smith (Take That) only managed their artists. Typically, pop music producers and songwriters in the 1990s would have their own independent representation.

After the commercial success of the Pet Shop Boys Watkins found himself in a role of ‘service management’, in which he performed his management duties at the request of the act rather than dictating to them. After a series of successful singles and albums managing Pet Shop Boys he turned his attention to manufacturing Bros, co-writing the songs on the their debut album under a pseudonym and achieving notable commercial success. However, Watkins relinquished songwriting duties for Bros’s second album, and overall control; a decision that he believed was a key factor in their commercial demise. Watkins ensured that the next time he manufactured a pop act, he would take full control of all aspects of the process, as he explains:

‘I would like mentioned in my obituary that I am a control freak and I believe that every time I take full control it happens [success]. The minute I relin-
quis any control, I think it’s very dangerous. You just push me once over the edge and that’s it because I was holding and watching every single aspect. It only worked when I was in my professional environment working with Gilberts [accountants], working with Paul Rodwell [lawyer], working with Neil Ferris [promoter] and there was a perfect understanding’ (Watkins: 2014, personal interview).

Watkins forbade any marketing and promotion, whether it was external or within the record label, to be carried out without his permission. Watkins had control of the entire process, from the videos to the promotion campaigns and explained:

‘That’s why I formed my own film company, I formed my own graphics department, I employed my own graphic stylist and everyone else because I simply wouldn’t trust people [record company executives]. I have often been quoted as saying it [control] is a ‘necessary evil’ right, because I knew how important it was once you had those quality goods. You do it for money and all the rest of it but to me it was a secondary thing [money], getting it right was far more important (Watkins: 2014, personal interview).

Watkins’s surrounded himself with comments a team of business people he could trust and delegated jobs to personnel that had the right skills, allowing him to concentrate on what he did best, steering the pop production process; from the songs, to the production through to design, styling and promotion. Watkins was therefore able to ensure every link in the production chain was sufficient to enable commercial success (Harding & Curnow: 2010). Creativity comes about because of a system in action (McIntyre: 2012) and the level of control that Watkins describes above was important to him to achieve creative and commercial success throughout the 1990s. In the early days of Watkins’s management of Bros in the 1980s there was a similar amount of control but this was relinquished when Bros demanded control of the songwriting for their second album. Bros’s songs on the second album were not well received, principally because they lacked sufficient knowledge of the domain and an applied understanding of the criteria for selection operating within the field of pop. This caused Watkins’s attitude to change from East 17 onwards and he decided to take full control of everything, having learned his lessons from Bros. Watkins’s relationship with Pet Shop Boys was entirely different however compared his previous acts as Watkins was viewed as:

‘the big bad man that went to the record label, thumped his fists on the desk and got the money for Pet Shop Boys [for recording and promotion]. If Pet Shop Boys were informed by Tom it was more as a reaction against what Tom was trying to put across. For instance, for their ‘Please’ album cover, Tom turned up to the studio and showed them a mock-up of this huge foldout sleeve with maximum content. They balked at it and ended up with a plain white sleeve with a picture of themselves the size of a postage stamp on it.
They would say it was a very creative period though because Tom was someone they could constantly react against artistically. Tom wanted dancing girls in one of Pet Shop Boys videos and Neil was mortified that he was expected to stand there with dancing girls behind him – so it didn’t happen’ (Lindsay: 2014, personal interview).

Because he wasn’t in complete control of the ways in which the band were marketed and promoted Watkins eventually parted ways with Pet Shop Boys, which fuelled his desire by the 1990s to be in full control of the artists he managed. Team leader working relationships require commitment from both sides, especially from the producer’s point of view. It was clear early on that Watkins’s relationship with P&E created conflicts of interest, not least because Watkins was managing East 17 (and his other Massive Management artists), as well as P&E. When making a decision, Watkins had to choose between one or the other and he typically took the side of the artist and was always willing to challenge P&E.

Few descriptions of what a music manager’s role entails actually encapsulate the requirements of a manager in the manufactured Pop and BoyBand genre of the 1990s. Generally, management in the 1990s for manufactured artists such as East 17, Boyzone, Take That and so on, all started by brainstorming ideas of the type of pop band they wanted to formulate and then manage. East 17 are an exception to this as songwriter and rapper Tony Mortimer approached Tom Watkins with his song demos and Watkins offered advice and feedback with a recommendation to come back with some ‘like-minded’ friends. The other artists (especially BoyBands) typically attended auditions and were chosen, styled and mentored by the managers. The Boyzone and Westlife manager Louis Walsh (2007) calls it: ‘Audition Hell; it is hard being judged and it’s hard being rejected, but it’s part of this business and it’s certainly part of the audition process’ (Walsh: 2007, p.25).

Watkins predicted the early success of the East 17 singles and debut album and had persuaded the band’s songwriter, Tony Mortimer to sign to his publishing company, Porky Publishing. These developments were crucial in Watkins’s accumulation of economic capital. Watkins used his successful cultural and symbolic capital by persuading recording companies that his new artists, together with the P&E production team could provide hit records for them. After the 1992/93 successes of the East 17 singles such as ‘House Of Love’ and the debut album ‘Walthamstow’, Watkins then signed Deuce to London Records and 2wo Third3 to Epic Records. Tom Watkins was therefore both an instigator and provider of the work for P&E Music and placed himself in the novel position of both knowing what his clients required and how to translate these requirements to the individuals within his team. This generative and interpretive role is unique in pop production and places Tom Watkins in the role of team leader at P&E because his cultural,
symbolic and economic capital contributed to the existence of the P&E production team and studio.

Watkins’s cultural capital was therefore deployed within the evaluation process and his knowledge of the domain and the mechanisms and criteria for selection formed part of this evaluation. This is because: ‘the influence of the market – what will sell – is important in shaping the content and form of the musical product’ (Robinson et al: 1991, p.238).

1990s Pop Music Producers & The Harding / Curnow Capital

Phil Harding and Ian Curnow formed as a remix, songwriting and production team at Pete Waterman’s PWL Studios in the mid 1980s. Harding was already an established recording and mix engineer from the Marquee Studios and spent the first year of the Stock Aitken Waterman incarnation engineering their early mixes and recordings for acts such as Dead Or Alive at The Marquee. Having joined the SAW team in their move to PWL Studios in early 1985, Harding became chief engineer and oversaw the technology installations at the studio, the hiring of more recording engineers and the expansion of a second studio in 1986. During that studio expansion, a Fairlight programming room was built below the PWL Bunker studio and Ian Curnow was hired to manage the room, supply programmed keyboards and samples to the SAW team and deal with the increasing amount of external remix work that was being offered to Harding. Pete Waterman suggested that Harding and Curnow’s skills were complimentary to each other and that he would market them as a remix and production team on behalf of PWL. Remix and production hits soon followed for the Harding / Curnow partnership for acts such as Jermaine Stewart / Rick Astley / Blue Mercedes / Jesus Jones and Pet shop Boys. It was during a Pet Shop Boys collaboration that the Harding & Curnow talent as a remix and production team came to the attention of Tom Watkins, the Pet Shop Boys’ manager. These commercial pop music successes for Harding & Curnow earned them significant cultural, symbolic and economic capital and allowed P&E to leave a declining PWL in 1992 and create their own production facility within the Strongroom Studio complex in Shoreditch, East London. Their pop music domain knowledge had been somewhat limited by the way in which PWL was structured, with Pete Waterman and Tilly Rutherford insisting on conducting all business meetings and negotiations. In comparison to some contemporary pop production teams such as Denniz PoP and Max Martin at Cheiron Studios in Sweden, SAW’s commercial success has been viewed as short-lived (Seabrook, 2015). Consequently, there was an eagerness for Harding & Curnow to learn more and develop a better understanding of the domain and field of pop music production. Their first assignment after leaving SAW was to mix East 17’s first single ‘House of Love’ and Ian Curnow explains:
‘We were like the rabbits in the headlights and we just went for it and I think that’s what separates the good from the bad, not wishing to sound pompous. At that moment in time we were on it, we were a hot remix and production team willing to take risks and work with new technology. At those moments, if you dig deep and do something spectacular, which we did with ‘House Of Love’ [the first East 17 single]. We were also fresh from leaving PWL, keen to prove ourselves and this was a really exciting project for us to get our teeth into. It fired us up – Tom came in with tons of energy and enthusiasm about the project. Give us [P&E] the goalposts and we’ll go for it’ (Curnow: 2014, personal interview).

As Curnow mentions above, Watkins provided the brief for P&E to work towards and later provided a cassette tape of other East 17/Tony Mortimer song demos to listen to:

‘we went off to the New Music Seminar in New York in June 1992 and we hired a car and drove into the upstate New York countryside to listen to them [demos] properly and we looked at each other and said ‘what the fuck are we going to do with these?’ From that demo tape we chose 2 or 3 songs that we felt we could produce a good result from and that was the next stage of our work with East 17 – productions rather than just remixes towards their first album ‘Walthamstow’. There were no bridges and no song development on most of them (Curnow: 2014, personal interview).

The P&E Music Micro-domain

As the pop music market showed its enthusiasm for P&E produced records, those involved in the process and principally Tom Watkins, grew in their own self-confidence. This resulted in the emergence of a P&E house-style (or distinctive sound), in which tried and tested production methods and sounds were re-used, Zagorski-Thomas (2014) calls these ‘sonic cartoons or schematic mental representations’. This increased the symbolic capital of P&E and Tom Watkins as hit-makers. In this way, the P&E production team created their own unique sound through a process of internal re-evaluation. By early 1995 the P&E team had achieved the UK Christmas #1 (1994) with ‘Stay Another day’ by East 17 and a whole album of co-written and produced material of the pop act ‘Deuce’, ready to be released and launched by London Records during 1995. The creative system can therefore be seen operating inside the recording studio on a group level in which they drew from a limited area of the domain to rearrange it in a new way and then internally evaluate their contributions with Tom Watkins as the lead authority within each process. On a group level then, elements of the creative system can be scaled and the domain becomes a microdomain, which contains related elements of the pop song. The social organization inside the recording
studio is scaled to represent a microfield, which accepts or rejects new creative ideas. The scaled creative system is shown in figure 2.

Watkins began the creative process as the strongest agent with the most amount of agency, gathered the necessary agents with their own musical and technical knowledge of the domain and steered them to deliver a creative product that would be deemed commercially successful by the field of pop music. When that product did succeed the recognition would return to Watkins, which further enhanced his cultural and symbolic capital. This was expected and accepted by the two other main agents in the production process – Harding & Curnow – whose contributions were always appreciated, credited and well represented.

The P&E Music Service Model in Action

The macro process at P&E during the 1990s can be depicted as a flow diagram (figure 3) that models the involvement of Tom Watkins (team leader) at various stages during the production. It is important to highlight the element of time throughout the service model because the duration of a single pop project is often defined by the power and capital of the team leader.
Generally the agency of the team leader is not absolute as they won’t be involved at particular stages of the songwriting and production process (see figure 3). This helps the team leader to maintain a more objective perspective on the musical product as it develops. It’s also important to note that although the stages of the service model are linear, within each of these stages the creative process was non-linear as Sawyer states: ‘creativity did not descend like a bolt of lightning that lit up the world in a single brilliant flash. It came in tiny steps, bits of insight, and incremental changes’ (Sawyer: 2013, p.2). The P&E Music creative process demonstrates the steps that Sawyer suggests and most notable are the points of Tom Watkins’s engagement with the process.

Very little began at P&E without Tom Watkins’s agreement; Watkins would choose the song to record or would have some involvement in writing a new song, which would often be a substantial involvement as both team-leader and the lyricist / top-liner. Harding & Curnow and their team would compose a minimal backing track that included drums, bass, keyboards and a rough arrangement of the song. Vocals, including lead and backing vocals, would then be recorded and arranged. The production team then completed the rest of the record in sympathy with the vocal; principally because this is considered to be the most important element of a pop recording (Harding: 2017). Watkins was not involved in the stages of vocal recording and music programming as he was content to leave these parts of the process to the rest of the production team. He did however return during the mixing stage and, after listening to a mix, Watkins offered his opinion on elements that required revision, which sometimes involved song arrangements and song structures. On occasion this involved revisiting previous stages to amend or add particular parts. In rare instances, it involved beginning the entire process from the beginning in order to react to new musical tastes, trends or market requirements. For example, East 17’s ‘Someone To Love’ (1996) on their third album was recorded twice because both Watkins and songwriter Tony Mortimer rejected the first version at the mix stage. Their request was for Harding & Curnow to return to the original song demo by Mortimer and to re-produce the whole track again to this format and arrangement. This was an extreme example of how far a production team is willing to travel to be ‘in service’ to the client but in this case the client was the team-leader and the songwriter. Watkins had the symbolic and cultural capital to persuade the P&E team to take such drastic measures; his motivational abilities certainly had everybody absorbed in the conscious activity of working together as a team to produce hit records. Watkins engaged the PE Music team fully in what Keith Sawyer (2003) calls ‘Group Creativity’. The Service Model and its stages are depicted in fig. 3 below:
Conclusions

In Pop music it is normally the artist who receives the most attention because their contribution is the most prominent on the final record. However, during the production process inside the recording studio, the contribution of the artist is often a small one as a team of songwriters, programmers, lyricists, top liners, musicians and producers work ‘behind the scenes’ in creating a pop record. Each member of the team has their own respective domain knowledge and varying degrees of agency. Driving the process is the team leader who has the most amount agency and at P&E in the 1990s, Tom Watkins had accumulated enough cultural, symbolic and economic capital to undertake the role of team leader.

The simplified service model of pop production at P&E illustrates the role of the team leader and their involvement within the decision-making process and their involvement at both the beginning and the end of the pop production process. Because of Tom Watkins’s accumulated economic, symbolic and cultural capital the pop production process in the recording studio began and ended with him. As team leader, Watkins may choose the song to record or have some involvement in writing a new song, using his expertise as a
lyricist and top-liner to compose the melody and the lyrics for the song. The backing track would then be composed by the production team of Harding & Curnow, which involved programming drums, bass, keyboards and sketching out a rough arrangement of the song. Vocals, including lead and backing vocals, would then be added and the production team would complete the rest of the record in sympathy with the vocal. Within each of these production stages, the creative system can viewed in operation on a group level as the production team draws from different areas of the domain of pop music and then verifies each of these creative ideas or actions in relation to Watkins’s and the broader field’s mechanisms and criteria for selection. At the stages where Watkins is involved, Watkins evaluates Harding and Curnow’s creative contributions with consideration for the mechanisms and criteria for selection of the field of pop music and, in this way, a microsystem of creativity can be seen in operation during the pop production process.

References


**Discography**


Abstract

This paper explores the various types of sampling and digital manipulation used within “33 ‘God,’” the fourth track on 22, A Million. Within this track, Vernon utilizes unique modes of musical and vocal expression, featuring samples as symbols of otherness—other times, other places, other emotions—presented through the timbre of others’ voices in order to convey his own memories and mental states. Vernon’s authorial voice within the track is both complicated and strengthened by this presence of other “characters” in his narrative. The samples will be examined through my term “sonic totem,” which is offered to capture the relationship of sound objects and memory within a narrative song. The fact that Justin Vernon’s is not the only voice heard on the track will be analyzed as signifying, as described in the theory of Henry Louis Gates, Jr. (1988), and related to theories by Simon Frith (1996) and Lori Burns (2010).

Introduction

The impetus for this project stems from an innate fascination with Justin Vernon’s 2016 album, 22, A Million. While the lyrical syntax and formal structure of these songs are not much changed from Bon Iver’s earlier output, Justin Vernon’s expression and attitude appears significantly changed. Darkness and anger in some tracks (especially “10 Deathbreast”) paired with manic joy or elation in others (especially “45”) presents a depth of expression yet to be seen from Justin Vernon/Bon Iver. Vernon’s voice covers greater range as well—which is additionally enhanced by intense digital manipulation. As the front man, songwriter, producer, and lead singer of Bon Iver, Justin Vernon presents himself as the author of Bon Iver’s songs, and his live performances, interviews, and press releases further the notion that his songs are outcroppings of his personal experience.

Although Vernon is the leader of Bon Iver, his “crew” deeply influences the final musical result: in liner notes he acknowledges many people and producers who shaped 22, A Million.
Like other Bon Iver projects, Vernon’s unique voice and melodic phrasing are highlighted within 22, A Million, and showcase his ability to move from his gritty falsetto through to a baritone range. The emotionality of his voice is typically considered powerful and moving, but to some critics and listeners, the digital manipulations present in 22, A Million render his natural voice (which is already a bit strange) into uncanny, unnatural, or over-emotional tendencies.

The scope of this inquiry deals with Justin Vernon’s use of digital manipulation in “33 ‘God’”, track four of 22, A Million. Within this track, and throughout 22, A Million, Vernon utilizes modes of musical and vocal expression in new ways than in earlier projects.

First, although Vernon typically employs cryptic or abstract titles, track names for 22, A Million are even more abstract as collages of letters, symbols, and numbers. Second, nearly every track radically expands or eschews typical instrumentation and musical texture associated with indie or alternative rock; instead, most tracks employ and prioritize digital samples as the main timbral elements. Vernon began to experiment with this type of digital manipulation on 2009’s Blood Bank EP, especially with the track “Woods.”

As a third new expression element, Vernon seems to favor timbres that are unpredictable or glitchy on this album—sounds that cut out randomly or stutter and repeat themselves. The five tracks that do not include digital samples, “715 Creeks,” “29 #Stratford Apts,” “21 Moon Water,” and “8 Circle,” feature production techniques that manipulate and mutate “natural” or acoustic instruments and voices.

The fourth and final new expression element of 22, A Million is Vernon’s use of other’s voices within his own composition by sampling, as exhibited on “22 (Over Soon),”, “10 Deathbreast,” “33 ‘GOD’”, “666,” and “00000 Million.”

To manipulate samples within 22, A Million, Vernon and his team used an OP-1 synthesizer+sampler device, created by Stockholm-based company Teenage Engineering, to create and manipulate samples. Vernon also develop into what it is. Bon Iver, the band, includes many musicians who have alternate bands or solo project themselves and Vernon acknowledges their input on his musical creation.

3 To view the track names and their stylization, I suggest looking at the album’s liner notes. These notes are available in any publication format (CD, LP, Digital booklet). Here is a recreation of the track names and their stylization. “22 (OVER S∞∞N)”; “10 d E A T h b R E a s T”; “715 - C∑∑KS”; “33 ‘God’”; “29 #Stratford APTS”; “666”; “21 M◊◊ WATER”; “8 (circle)”; “____45____”; “00000 Million”.


4 A variety of techniques were used within the album to achieve glitchy effects. For example, “29 #Stratford Apts” features straightforward lead vocals (from Vernon and band mate, S. Carey), acoustic guitar, harmonizing vocals, and piano. However in production, Vernon and his team recorded the vocals onto tape, then crumpled up and damaged the tape in order to create warbled vocals that cut in and out on the final track.

304
Elizabeth Navarra Varnado: The sampling aesthetic of Bon Iver’s “33 ‘God’”

oped a system of hardware and software with producer Chris Messina that harmonizes live sounds coming from another player. This device, called the “Messina,” was created as an extension of the Prismizer effect (created by Francis Starlight, first made popular by Chance the Rapper’s *Coloring Book*). The Messina is set-up like other harmonizers but can be performed live with low-latency: A microphone inputs into an audio interface, which is hooked up to a laptop running the harmony engine software. A midi controller is mapped to parameters within the software so that Vernon can manipulate the harmonies during live performance. The “harmonized” audio is then converted back to analog. A combination of additional off-the-shelf analog gear and plug-in settings create the special recipe that sets the Messina apart from other iterations of this type of machine.\(^5\)

This live performance playability was the most important feature to Vernon. In an interview, Messina described the tool: “Normally, you record something first and then add harmonies later. But Justin wanted to not only harmonize in real time, but also to be able to do it with another person and another instrument. The result is one thing sounding like a lot of things. It creates this huge, choral sound.”\(^6\) On *22, A Million*, the Messina is most often used to harmonize and add glitch effects to vocals and saxophones, especially within “715 - CRΣΣKS”, “666 i”, “21 MØN WATER”, “8 (circle)”, and “____45____”.

**Sonic totems: Defined**

Most of the production techniques used on *22, A Million* may be revealed through a focused examination of the fourth track, “33 “God.”” This is also the best example of Vernon’s use of samples within the expression of a personal narrative. Just as digital effects veil Vernon’s “real” vocal timbre, his narrative voice is veiled and projected onto the samples he uses. These samples exhibit a concept I term “sonic totems:” sound objects that represent a connection between timbre and memories of a specific space or place in time (see diagram below).\(^7\)

---

\(^5\) Messina has obviously not shared all components used to create the “Messina” sound, but watching live videos of Vernon perform “715- Creeks” will show the multi-faceted capabilities of this harmonizer.

\(^6\) In an interview with Emilia Petrarca, “The Engineer Behind Bon Iver’s *22, A Million* Clears Up Any Confusion About Its High-Tech Sound,” Messina discusses how the software and hardware of the Messina work together. An electronic music friend helped me decipher what software/hardware Messina refers to.

\(^7\) I created this term to describe specifically the connection between timbre within samples and memory. “Sonic totems” is distilled from the concept “AudioAutobiography.” This term was created in a musicology seminar on “music and memory” at the University of Kentucky (Spring 2017). An AudioAutobiography is a group of sounds (musical or not) that represent a memory of the author’s, and communicate this experience to the listener. “Sonic totem” further describes the connection between a very specific sound and memories of place and time.
A sonic totem represents temporal “otherness” each time it occurs. This otherness is the most important part of a sample: even if the listener cannot identify the origin of the sample, the imbedded timbre of that sample triggers some attempt at recollection and a sense that the sample is taken from somewhere else. In earlier publications, Erik Askerøi has used his term “sonic markers” to discuss this connection between timbre and time in music (2013, 2016). My term extends this by focusing on the object-ness of samples, and how these objects are tied to memories. Sonic totems are objects that are used or displayed in order to retrieve or describe a specific point in space and time, and can be equated to old photos, souvenirs, or other objects used in an artistic collage. Referencing Lori Burns’s schema for interpretation of a narrative voice in song, the meaning of a sonic totem can have different shades as it moves from the real author (in this case, Justin Vernon), through the recording’s narrator and implied reader/listener, to the real listener (Burns 2010, p.161).

Using the diagram below as a guide (tab.1) each time a sonic totem is put into play, a layer of space or time memory can be added on to its sonic totem. For example, “It’s Your Thing” is a track that could represent the early 1970s “golden age” of funk. This early 1970s drum sound is considered a “classic” hip-hop beat, and when this groove is used as a repeating layer within a hip-hop beat, it represents the late 1970s or ‘80s, a “golden age” of hip-hop. While these first two are examples of cultural or historical memories connected to music, “It’s Your Thing” could also be a personal sonic totem, if it represented to the author their own “golden age” (summer as a twenty-year old). Additionally, if a listener holds the sonic totems as representative of something else, say, the death of a loved one, the sonic totem can take on ever-more faceted meaning.

Tabel 1. Sonic Totem in sound, space and time.

<table>
<thead>
<tr>
<th>Sonic Totem</th>
<th>Space</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>lo-fi, syncopated, drum groove. “Classic” hip-hop sound</td>
<td>“It’s Your Thing” by Cold Grits</td>
<td>early 1970s, the “golden age” of funk</td>
</tr>
<tr>
<td></td>
<td>as a repeating layer within a hip-hop beat</td>
<td>late 1980s, a “golden age” of hip-hop</td>
</tr>
<tr>
<td></td>
<td>*played at a party</td>
<td>*summer at 20 years old</td>
</tr>
</tbody>
</table>

This term also references meaningful objects used in filmmaking to denote passage of time or change in a character’s mindset, location, age, etc. A good example is the “totem” used in Inception to determine what plane of existence characters find themselves.

8 In her analysis, “Vocal Authority and Listener Engagement,” Burns extends Seymour Chatman’s narrative communication model and applies it to narratives in popular music.
By analyzing samples as sonic totems, “33 ‘God’” becomes a collage of sounds that incorporate memory of the general musical past and memories from Vernon’s personal past. With this track, Vernon has moved beyond the perceived authenticity of his own quiet voice, lost in the woods of sincerity and folk music, and explores the ability of the electronic “other” to evoke a powerful, intertextual narrative.9

Identifying and tracing purposes of samples

To aid with the analysis of “33 “God,”” I created a spreadsheet that can be read like a score. It denotes what samples are used where, and how they are placed into Vernon’s autobiographical narrative. This spreadsheet can be found in Appendix A, along with directions for its use. Within the spreadsheet and throughout this paper, terms defined by Mark Katz and Serge Lacasse will be employed that have become widely used in the discourse on digital sampling. These terms are useful in identifying and tracing the purpose of a sample within the narrative of “33 ‘God’”. Autosonic quotation refers to musical borrowings that quote a musical segment by digital (or analog) sampling, but the sample can be manipulated. Performative quotation, a term used by Mark Katz, further defines the distinction of autosonic quotation as “quotation that recreates all the details of timbre and timing that evoke and identify a unique sound event” (Katz 2004, p.141). This means that the sample is not aggressively modified, so it sounds as if it has been “dropped” from the other recording. Autosonic samples serve a functional purpose within a track: either looped as part of a foundational accompaniment beat, or featured within the main “lead vocal” space of the track as commentary on the meaning of the track. Autosonic samples serve to enhance the meaning of a track by giving it credibility to those in-the-know and/or by “dropping in” historical timbres that give historicity to the track. There is a third type of sample present in “33 ‘God’” that isn’t defined by Katz or Lacasse. I term this type “manipulated acoustics”: defined as samples of acoustic instruments (for example: guitar, piano, human voice, saxophone) that have been digitally or manually altered through recording techniques or equipment—in this case, the OP-1 or the Messina—and turned into mediated musical units. For the sake of continuity, I will use the term “unmediated” to define recorded sound within the track that sounds “true to

---

9 The scope of this paper does not allow for a full discussion of why Vernon/Bon Iver is considered authentic by listeners. However, Vernon presents himself essentially as a singer-songwriter, and much scholarship has discussed how this type of performer appears authentic to their audience. (Frith 1996, Moore 2002, Negus 2011)
life,” although there will still be EQ and tweaking that’s done to any “live” sound while in the recording process.

Table 2. The samples in “33 ‘God’” are drawn from the following sources.

<table>
<thead>
<tr>
<th>Song name</th>
<th>Artist</th>
<th>Album, Year</th>
<th>Samples used</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Morning”</td>
<td>Jim Ed Brown</td>
<td>Released as a single, 1970. RCA Victor.</td>
<td>Introductory music with strings, Jim Ed Brown singing: “When we leave this room, it’s gone” and “Here in this room, this narrow room, where life began when we were young last night.”</td>
<td></td>
</tr>
<tr>
<td>“It’s Your Thing”</td>
<td>The Isley Brothers, covered by Cold Grits</td>
<td>Released as a single, 1969. ATCO Records.</td>
<td>Introductory drum beat.</td>
<td>This sample is not credited in the liner notes, but is found on Whosampled.com.</td>
</tr>
<tr>
<td>“All Rendered Truth”</td>
<td>Lonnie Holley</td>
<td><em>Just Before Music,</em> 2012. Dust-to-Digital.</td>
<td>Lonnie Holley singing “All my goodness to show”</td>
<td></td>
</tr>
<tr>
<td>WCFW Radio</td>
<td>Radio Chippewa</td>
<td>N/A</td>
<td>Percussive vocal syllables, in a female vocal register.</td>
<td>Vernon identified this sample in an interview, Sept. 2016. The station broadcasts near where Vernon lives.</td>
</tr>
<tr>
<td>Psalm 22 quote</td>
<td>N/A</td>
<td>N/A</td>
<td>A voice reading Psalm 22, in a low male register.</td>
<td>Credited in liner notes to “A reading by Granger Community Church in Granger, In.”</td>
</tr>
</tbody>
</table>

Influence of hip-hop tradition

Vernon’s choice of “It’s Your Thing” has significance in the story of “33 “God,”” as a classic hip-hop sample. Vernon uses “It’s Your Thing” in a traditional hip-hop style, as a repeating layer in an accompaniment beat, for just one portion of “33 “God,”” This track has been used in the beat for several recent rap tracks and functions as a classic or “old-school” hip-hop sam-
Table 3 shows another example of its use as a sonic totem in “33 ‘God’”

**Table 3. Sonic Totem in sound, space and time, example from “33 ‘God’”**.

<table>
<thead>
<tr>
<th>Sonic Totem</th>
<th>Space</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>lo-fi, syncopated,</td>
<td>“It’s Your Thing” by Cold Grits</td>
<td>early 1970s, the “golden age” of funk</td>
</tr>
<tr>
<td>drum groove</td>
<td>as a repeating layer within a rap</td>
<td>2010s to signify the “authenticity” and lineage of the producer</td>
</tr>
<tr>
<td></td>
<td>33 “God” by Bon Iver</td>
<td>2016</td>
</tr>
<tr>
<td></td>
<td>• signifies lineage of the producer</td>
<td>• signifies influence of Kanye West</td>
</tr>
<tr>
<td></td>
<td>• signifies “sexiness” in the song</td>
<td>• signifies “sexiness” in the song</td>
</tr>
</tbody>
</table>

The funky groove of “It’s Your Thing” is interpreted as signifying sexiness because of the way it underscores lyrics that hint at a sexual relationship but are kept ambiguous.

Staying at the Ace Hotel, if the calm would allow

Then I would just be floating to you now

it would make me pass to let it pass on

I’m climbing the dash

That skin

This will be discussed further within the song analysis. For hip-hop producers, the draw of “It’s Your Thing” and other classic breaks from funk and soul records is the recognizable timbre of these breaks. Even under digital manipulation, such as pitch or tempo shift, the grit and grain of pre-recorded material holds true and catches the ear of a listener as something “other,” something that has come from somewhere – and “somewhen” else. Musicians and producers seek to use sounds that ground their compositions within

---


11 The format of these lyrics is replicated from the liner notes of the album. This preserves the phrasing of the lyrics within the song. This is also replicated in the spreadsheet of the song, located in Appendix A.
the history of music that has come before them, just as scholars quote each other to show how new research ties in with existing and historical scholarship. Hip-hop producers show historicity by literally dropping sounds from the past into their works, but producers and musicians in other genres often quote licks, drum beats, and fragments of melodic lines from earlier works to ground them in the past. For hip-hop producers and fans, “historical timbre” is one important factor that signals “good” or “authentic” hip-hop. Musical quotation is at the center of the genre, and the authenticity of hip-hop tracks is wrapped up in whether the sound of a chosen sample is judged to be authentic or not (Schloss 2004, p. 67). Raps often contain verbal references to people from the author’s past and the general history of hip-hop, and samples usually back-up these references. Hip-hop scholar Joe Schloss identifies this referential tendency as signifying, a theory first outlined by Henry Louis Gates Jr., as a basis of analysis for African-American literature. In his 1988 text, The Signifying Monkey, Gates points out how “black writers read, repeated, imitated, and revised each other’s texts to a remarkable extent” (Gates 1988, p. xxii). Gates recognized that “repetition and revision are fundamental to black artistic forms,” including music (Gates: 1988, p. xxiv). He states: “I decided to analyze the nature and function of Signifyin(g) precisely because it is repetition and revision, or repetition with a signal difference” (Gates 1988, p. xxiv). Gates’s “signal difference” identifies the space and time categories on my diagram of the sonic totem. When a new layer of space or time is added into the “otherness” of a sonic totem, this is what Gates calls a signal difference—the signified elements (space and time) have changed, while the signifier (sound object) stays the same.

In his theory, Gates states that signifying should not be confined only to so-called African-American literature, because “all texts Signify upon other texts, in motivated and unmotivated ways” (Gates 1989, p. xxvi). By considering the lyrics and musical elements of rap as text, the displays of repetition, imitation, and revision among producers and artists of hip-hop and rap music are clearly evident, and make these genres prime examples of Gates’s theories.

According to popular music scholar Mark Katz, repetition of this text “can be used to boast, insult, praise, or moralize, [and] generally play on the many possible meanings and interpretations of a given statement” (Katz

12 Askerøi also offers the theory of ‘retronormativity’, defined as “the mechanism of repositioning the ‘past’ in the ‘present’ and implying, in turn, a nostalgia for vintage technological artefacts and their aesthetic impacts on musical sound” (Askerøi 2016, p. 381). However, although samples from “33 3God” are from the past, they are not used in a way that preserves historical timbre, except perhaps for the drum groove of “It’s Your Thing.”

13 Schloss quotes producer Jake One as evidence of this judgment within the hip-hop community.

14 Signifying(g) is written with the parenthetical g to separate it from Peircian signifiers and to denote its origin in African-American vernacular culture.
Elizabeth Navarre Varnado: The sampling aesthetic of Bon Iver’s “33 ’God’”

Katz cites Public Enemy’s “Fight the Power” as an example of “double-voicedness” as it is so-called in the original theory by Gates (Gates 1988, 51). In the opening minutes of the song, a vocal sample from James Brown’s “Funky President” is used: James Brown’s voice and rapper Flava-Flav’s are heard at the same time singing: “People, people.” “Fight the Power” has overtly political lyrics, and the raps of Chuck D and Flava-Flav reference samples used within the accompaniment: “1989, the number another summer (get down)/ Sound of the funky drummer” (Referencing the break from “Funky Drummer” in the accompaniment). Samples within “Fight the Power” express opinions regarding the Vietnam war, white American history overshadowing African American history, and pay homage to African-American musicians important within hip-hop tradition. Within the first minute, samples of James Brown, Afrika Bambaataa, George Clinton, The Jacksons, and others are mixed into the beat of the track (Katz: 2004, 156). In most cases, only a hip-hop aficionado would recognize these samples, but the listener’s perspective is only one facet of signifyin(g). Public Enemy prioritized these digital samples as compositional building blocks; they are used as sonic totems that uphold and support the meaning of the lyrics. Further, it’s not just the words of previous rappers that are featured in the samples, but the grooves and vocal interpolations (specifically those of Afrika Bambaata) are also used to create a sonic landscape where their passionate raps can reside. Public Enemy used their lyrics to vocalize their philosophy, and used sonic totems to signify the compositions or physical voices of older musicians. To sample the voice of an honored role model is a powerful way to “back up” or bring legitimacy to one’s narrative. Rappers and producers consistently use sampling to enhance narrative lyrics and provide an aural history of their own rap lineage.

Because multiple voices are portrayed in a song that is both narrative and historical Katz’s analysis of “Fight the Power” provides a model with which to analyze Vernon’s “33 ‘God’” (although the latter lacks political themes). Both songs portray a strong narrative with sung/rapped lyrics that are supported by the musical elements in underlying samples. Gates’s theory of signifyin(g) is additionally important to the analysis of “33 ‘God’” because the samples used by Vernon are repetitions, imitations, and revisions of their original recordings.

15 An older example is “The King” by Grandmaster Flash and the Furious Five, released by Elektra in 1998. The track features samples of Dr. Martin Luther King Jr.’s “Dream” speech. A second good example is Kendrick Lamar’s “DUCKWORTH” released in 2017. Lamar’s lyrics tell a story of epic-like proportions, citing the background and lineage of his own producer Top Dawg. The samples and other production elements (such as a rap in rewind at the end of the track) reinforce and enhance the thesis of the rap: “Pay attention, that one decision changed both of they lives.”
The authorial voice within “33 ‘God’”

In this and in most of his songs, Vernon’s abstract/word associative lyrics outline a first-person narrative, and generally consider the past tense. In “33 ‘God’,” Vernon’s lyric “I’d be happy as hell if you stayed for tea” within the opening stanza announces that this is a personal narrative. Vernon’s songwriting generally eschews verse-chorus structure, and he tends to write in three-part structures or stanzas, basing his musical material on the trajectory of the narrative lyrics. If an idea or theme in the lyrics returns, a musical phrase or melody will also return. These typical elements are found within “33 ‘God,’” where Vernon’s lyrics reference spirituality, loneliness, love, and perhaps an affair or love unrequited/unconsummated. A reflective or nostalgic tone is presented.

The impression that “33 ‘God’” is, in fact, autobiographical comes from Vernon’s reputation for intimate, “self-expressive and self-revealing” lyrics (Frith 1996: p. 183). Vernon upholds a secretive personal life, avoiding press photos, fan pictures, and interviews, but he does give occasional clues that his songs are related to or descriptive of events in his life. As lyricist, Vernon’s is the ultimate authorial voice (Frith 1996: p. 184). Frith’s theory about the voice, questions how the singer’s voice relates to other voices, or other points of view, present within a song (Frith 1996: p.183). In the case of “33 ‘God’”, Vernon further complicates the meaning of authorial voice by adding other voices into the mix of vocal perspectives within the narrative. What does this implicate for the meaning of these other voices, these sonic totems, presented by Vernon?

These sonic totems presented by Vernon are commentary on his own lyrics, allowing his authorial voice to remain intact. Vernon splices and repeats the samples to create sonic totems that portray a memory of sound-place-time, and “double-voicedness.” To pin-point this double-voicedness, focus on the lead vocal space of “33 “God.”” The lead vocals of a track are a space where the lyrics reside, a space where the meaning of the song can be analyzed. Justin Vernon’s lead vocals are heard as the loudest element of the

---

16 Before releasing 22, A Million, Vernon gave one press conference to discuss its material. He stated that this was the only interview he would give regarding the album. Within the hour-long press conference, Vernon discussed why he chose several of the samples, how the Messina was created, etc. Specifically referencing “33 ‘God,’” he said: “This song is about a very messy night in London, and I met this guy Paolo Nutini, and I was like, ‘Who is this guy?’ A few days later after this very important happening, I was sitting in the hotel room in Spain when Volcano Choir was playing over at Primavera, and I just heard that, ‘Find god and religion,’ I kind of heard it amongst the lyrics of the song I was already working on for that, and I was like, ‘Great! More samples. Yeah!’” A video of this press conference is available on Bon Iver’s website.
mix of “33 ‘God’”, but the samples of Jim Ed Brown, Paulo Nutini, and Lonnie Holley are also heard as lead vocals.17

**Song analysis: Sonic totems as structural elements**

The track opens with a familiar idiom regarding a one-night stand or affair, from Jim Ed Brown’s tune “Morning”: “When we leave this room, it’s gone”. A sample of a string intro from “Morning” is featured here, with its pitch and tempo modified. This modification produces a romantic, sliding, and somewhat off-tuned timbre in the strings. The WCFW sample provides female murmurings that further suggest “When we leave this room, it’s gone,” is in reference to some romantic entanglement. Finally, the intro concludes, and Vernon’s voice comes in to sing a question: “Is the company stalling? We had what we wanted: your eyes.” Jim Ed Brown’s: “When we leave this room, it’s gone” is repeated three times within the first minute of the song, reinforcing this subject matter in the narrative and serving as a structural boundary for the opening stanza of the song. Here, the accompaniment groove is provided mostly by polyrhythm between the syncopated piano line and the WCFW Radio sample. At 1:08, the final repeat of “When we leave this room, it’s gone” leads into Paolo Nutini’s voice and words: “I find God, and religion, too.” At 1:20, the background accompaniment falls away to a sparse vamp, that hangs for a moment before launching into the backbeat provided by “It’s Your Thing”, a new fat and funky bass line, and deeply-echoing drums. To begin the second stanza, Vernon sings, “staying at the Ace Hotel, if the calm would allow, I would just be floating to you now.” This is the dynamic high point of the song thus far, built upon the foundation of the classic funky backbeat and peppered with Justin Vernon’s high-pitched interjections that are echoed by synth samples: “oohs” similar to those sung by James Brown, Michael Jackson, and Prince (around minute 1:33). While the lyrics are fairly ambiguous, these musical features impart a sexy edge to Vernon’s performance and match it to the subject matter of the earlier Jim Ed Brown sample. Consider again Paulo Nutini’s vocal timbre at “I find God, and religion too.” This timbre is not matched or echoed by Justin Vernon except at the emotional climax of the track, beginning at 2:12. At this moment, Vernon’s voice enters with the most abstract poetry yet to be sung: “Well we walked up on the bolt in the street After you tied me into the -way of the apartment of his bede…”18 At 2:29, Vernon’s perspective

---

17 The song analysis is greatly aided by the spreadsheet! The spreadsheet denotes in bold text when the lead vocal is Justin Vernon’s, and in neutral text when the lead vocal space is filled with autosonic samples of Jim Ed Brown, Paulo Nutini, or Lonnie Holley.

18 Bede is not a word that can be found in the Oxford English Dictionary, and a reference to St. Bede seems ill-fitting. One reader suggested it may be a misspelling or mispronunciation of bed, but this has not been discussed by Vernon.
shifts into less enigmatic, first-person statements. His voice drops into a lower register to sing “I didn’t need you that night, not gonna need you anytime. Just gonna take it as it goes, I could go forward in the night/well I better fold my clothes.” These lyrics do not point to any particular romantic or spiritual crisis if they are taken out of the context of the other samples in the track, however, when they are presented in correspondence with the sampled lyrics, their meaning is projected as some declaration of closure or surrender. These lyrics at 2:29 also contrast with the other abstract and ambiguous lyrics that Justin Vernon sings in the rest of the track.

Following this at 2:38, the groove built by the “It’s Your Thing” backbeat and funky bass returns and the listener is left with the sense that whatever may have been building up in this story—never quite happened. Typical conventions used to signify resolution at the climax of a rock song—such as a crescendo to a drum break and a return of a chorus, or a key change, or some “wicked” instrumental solo that ascends higher and higher and higher until a screeching high note is wailed out—are not used. All that Vernon gives is a return to the earlier Cold Grits backbeat groove…then, Lonnie Holley’s voice enters into the lead vocal space: “All my goodness to show.” Finally, a lonesome quote of Psalm 22 is recited, twice: “Why are you so far from saving me? Why are you so far from saving me?”

Song analysis: Sonic totems as alternate voices

Vernon’s samples and manipulated acoustics act as sonic totems that create an association between timbre and memory; the lyrics of “33 ‘God’” are given interpretive power only within the context of the musical imagery and additional lyrics provided by the samples. While Justin Vernon’s voice and lyrics present a character that is ambiguous, abstract, and impressionistic (although still intimate and emotional), Jim Ed Brown’s voice presents the regretful, nostalgic lover; Paolo Nutini’s remembers an emotional and spiritual experience; Lonnie Holley’s voice pleads, and finally, the low rumble of Psalm 22 is a spiritual cry of anguish or defeat.

Throughout the track, Jim Ed Brown’s lyrics are explicitly about an affair, where as Vernon’s lyrics are never this straightforward. Vernon’s lyrics need the funky Cold Grits sample and Brown’s lyrics to direct the listener’s understanding of his abstract words. Brown’s “When we leave this room, it’s gone” frames the first section of this narrative musically and ideologically, but there is no place where Vernon’s lyrics match the straightforward, non-metaphorical lyrics of Jim Ed Brown’s country-pop tune. Additionally, the spiritual and emotional elements brought to the track by Paulo Nutini’s lyrics (“I find God, and religion, too”) are never echoed with any spiritual lyrics sung by Justin Vernon. Only the Psalms 22 sample, which enters at the end of the track (3:10), reiterates this spiritual element. Even then, “Why are you
so far from saving me?” is only an allusion to a spiritual statement, since it is taken from a religious text that is assumed to be addressing God. Vernon presents other voices as sonic totems that represent his experience, rather than singing some of these direct statements himself. However, Vernon includes the sampled text within his own lyrics in liner notes, strongly signifying their place within the narrative. The text of the sonic totems is integral commentary within the narrative, not an afterthought or just a background vocal. To distance his personal voice from the listener, during the most direct first-person narrative statements, Messina harmonies are continually added, and the voice expands into the background of the mix.

Sonic totems of “Morning” and “All Rendered Truth” present the most candid descriptions within the narrative, but Vernon digitally manipulates them and alienates these voices (Brøvig-Hanssen and Danielsen, 2016). The sonic totem of “Iron Sky” is presented by Vernon as a performative quotation, as if Paolo Nutini were singing directly and “humanly” on this track. This type of performative quotation (Gates’s “double-voicedness”) is described by Vernon to be an artistic necessity. There are some musical lines that he claims he “couldn’t have sang (sic)” and that the original artist must be quoted. Vernon elevates these performative samples to insinuate that those melodies and lyrics belong to those artists and have the most meaning coming from their voices. These samples are important sonic totems of Vernon’s experience, placed into this musical narrative to be shared with listeners.

**Conclusion**

To conclude, I present “33 ‘God’” as an ambitious example of narrative song writing, created through digital sampling processes drawn from hip-hop aesthetics. Vernon weaves his own memories within the memories of others with “sonic totems” or digital samples. Sonic totems become vessels for memory in the moment where Vernon’s memory might intersect with the listener’s experience, as both author and listener creates meaning from the sonic totems. When a listener recognizes a timbre, and this is possible even if they can’t identify the song, an emotional connection is generated that causes introspection and memory-searching for the source of that sound.

Going forward, I believe that the sonic totem concept can be a useful tool to describe the “self expressive and self revealing” power of sampling within

---

19 Research into this area is being completed by Emil Kraugerud, first presented at the 2017 Art of Record Production conference, under the title “The role of stereo center placement in constructions of intimacy.”

20 For more of this theory of the ‘uncanny’ see Grafton Tanner, Babbling Corpse: Vaporwave and the Commodification of Ghosts (2016). Or the concept of “hauntology” as described in Simon Reynolds, Retromania.
narrative song composition, an addition of voices that does not detract from the composer’s authorial voice (Frith 1996). My term encompasses the historicity sound-place-time memories tied to certain timbres, and how these timbres and musical units can be exploited to create narratives with deeper relationship between text and timbre. Beyond the power of clever or beautiful lyrical content, beyond the power of a well-placed guitar lick or vocal swell, the current development of sampling techniques has the power to elevate the singer-songwriter compositional method that has been around since early folk singers. Sampled voices and timbres are already lauded as adding expressive power and commentary to narratives in rap music, and “33 ‘God’” shows that they add that strength to indie/alternative/experimental music as well.

References

Bibliography


  http://pitchfork.com/reviews/albums/22420-22-a-million/

Web Resources
Published on Youtube, August 29, 2016.  https://youtu.be/6C5sB6AqJkM
Bon Iver official website, 22, A Million press conference video, September 2, 2016, 1:32:55,
https://boniver.org/pressconference
http://www.whosampled.com/Bon-Iver/33-GOD/
  Various search engines used within this site.

Discography

Appendix A:
The following spreadsheet was created as an analysis tool for Bon Iver’s “33 “God.”” It should be read from left to right, then down. Timestamps in the furthest left column denote the duration of the track. The “Lead Vocal” column shows the lyrics of the song. The format is recreated from the liner notes of 22, A Million, including capitalization, quotations and line breaks. The “Accompaniment” column gives my best guess at samples or instruments used within the accompaniment. “Sample Source” gives the song title of the sample. Any highlighted cell is a sample, and each color denotes a different kind of sample, as defined in the body of the paper:

- green: autosonic, vocal sample
- orange: autosonic, accompaniment
- purple: performative
- light blue: manipulated acoustic.

Any unmediated sounds, vocal or accompaniment, are not highlighted.
In the introduction and at 2:50 and 3:05, there are samples that are audibly present, but I cannot identify them, denoted by question marks. Whether these are manipulated acoustics created by Vernon or tiny samples from another recording is difficult to tell, but by 3:05 these unidentifiable samples form a climax of wild and diverse timbres that matches the highest emotional point of the lyrics as well.
Lead Vocal	Accompaniment	Sample Source
00:00 Intro:	piano	&	synth	lines,	sample	fragments
00:14 String sample - full version "Morning"
00:25 WCWF Radio sample looped throughout until 1:12 "Morning"
00:31 Is the company stalling (Vernon, unmediated voice)
We had what we wanted:	your	eyes
00:40 With no word from the former
I'd be happy as hell, if you stayed for tea.
00:51 I know so well, that this is all there is"
00:57 A child ignored
These will just be places to me now
01:08 When we leave this room, it's gone" "Morning"
01:10 String sample - full version "Morning"
01:12 I find God and Religions, too..."
01:27 Staying at the Ace Hotel, if the calm would allow
Then I would just be floating to you now
I'm climbing the dash
That skin
01:44 String sample - fragmented, in background "Morning"
01:58 Synth lines, drum hits. "Morning"
02:02 Drumbeat:	Cold	Grits	intro	looped
throughout	"It's	Your	Thing"
02:24 Call & response:	unintelligble	vocal	sample
Said I woulda	walked	across	any thousand lands
02:29 Do not really if you can't
I didn't need you that night,
was gonna take it as it goes
I could go forward in the night
I will better fold my clothes
02:39 Drumbeat:	Cold	Grits	intro	looped
to	3:08 "It's	Your	Thing"
02:38 Unintelligble vocal sample 2x "DsharpG"
02:44 "All my goodness" "All Rendered Truth"
03:05 "All my goodness to show" "All Rendered Truth"
03:10 Why are you so far from saving me* 2x "DsharpG"
03:11 "33 "God"", Bon Iver. 22, a Million Jaguar, 2016.
Abstract

Multi-channel sonic experience is derived from a myriad of technological processes, shaped by market forces, configured by creative decision makers and translated through audience taste preferences. From the failed launch of quadrophonic sound in the 1970s, through the currently limited, yet sustained niche market for 5.1 music releases, a select number of mix engineers and producers established paradigms for defining expanded sound stages. Whereas stereophonic mix practices in popular music became ever more codified during the 1970s, the relative paucity of multi-channel releases has preserved the individual sonic fingerprint of mixers working in surround sound. Moreover, market forces have constricted their work to musical genres that appeal to the audiophile community that supports the format. This study examines the work of Elliot Scheiner, Bob Clearmountain, Giles Martin, and Steven Wilson to not only analyze the sonic signatures of their mixes, but to address how their conceptions of the soundstage become associated with specific genres, and serve to establish micro-genres of their own. I conclude by arguing that auteurs such as Steven Wilson have amassed an audience for their mixes, with a catalog that crosses genre boundaries, establishing a mode of listening that in itself represents an emergent genre – surround rock.

Surround Sound Auteurs and the Fragmenting of Genre

The history of stereophonic sound... is a history of discontinuity... it is a multichannel history consisting of numerous intersecting flows of sound, music, scientific inquiry, financial investment and invention, and listening contexts and practices that lead only occasionally toward any kind of unitary movement, cohesion or success. (Théberge, Devine, and Everett 2015)

So authors Paul Théberge, Kyle Devine, and Tom Everett frame the development of stereo as both format and cultural marker. The road to surround sound is equally convoluted, with a series of loudly heralded, yet soundly rejected introductions of expanded listening formats such as Quad in the 1970s, and the competing formats of SACDs and DVD-As at the turn of the
millennium, and quite possibly incorporating the latest buzz – Dolby Atmos. Surround sound is a losing bet that won’t quite go away. Despite multiple market failures, a select group of audio professionals and creative musicians still pine for a world that envelopes the listener, and though the amount of music being offered to the public in such formats is miniscule in comparison to all recorded music, there is a small, but growing segment of recorded music audiences that have embraced surround sound not only as a listening preference, but as a cause, a mission. This article examines some of the figures who have shaped this listening experience, who have crafted the templates, and who have authored not only specific album mixes, but the very definition of what surround sound sounds like. I will focus on two points – what designates someone as a surround sound auteur, and what is it that is actually being authored.

Eliott Scheiner became an early proponent of surround, and served as mixing engineer on significant number of 5.1 releases at the turn of the millennium. In most cases, these were canonic “classic rock” catalog re-issues from the 1970s, reflecting the industry model of introducing new formats via established titles that had successfully introduced the CD over a decade earlier. Scheiner had serious credits as a recording/mixing engineer from Van Morrison’s Moondance, to Steely Dan’s Gaucho, and his stature in the industry guaranteed that his work in the new format would be state-of-the-art, yet reliably familiar. The dictate was to maintain the basic stereo soundfield of the original mixes, just extend it somewhat into an added dimension. But Scheiner had other ideas. Recalling his initial exposure to surround sound, Scheiner said, “When I heard the whole 5.1 concept for the first time, it blew my mind… The thought of being able to create a new environment for music, a whole fresh, new approach, really energized me.” (Walsh 2001)

However, Scheiner initially opted for a partially fresh, not particularly new approach to crafting surround mixes. The standard 5.1 array of L-C-R, with L and R rear channels supplemented by a low frequency subwoofer emerged from movie theater audio experiments in the 1970s. And reflecting Théberge, Devine, and Everett’s summation of stereo’s complex network of progenitors, the center channel – crucial for film dialog – perplexed mixers who were expert at creating a phantom center by balancing L/R signals for stereo. Scheiner has stated that his early mixes grappled with center channel questions, though he soon developed an approach that generally ignored the center channel, and thus resembling the quadrophonic mixes briefly issued in the mid-70s.

Yet, a comparison of quad mixes and Scheiner’s 5.1 mixes highlights some significantly different approaches to soundfield placement. Just as early stereo releases highlighted discrete channel separation, many early quad mixes featured a dramatic separation of channel information. Scheiner initially countered this with a more immersive approach that paralleled the
move towards a fairly codified stereo soundfield during the early 1970s, with
a great deal of “phantom” center, both L/R and F/R.

But as Scheiner gained confidence, his mixes moved towards something
that was certainly a new approach, with information appearing discretely in
any of four channels, and occasionally moving across the entire soundscape
— left to right; front to back.

When a lot of film mixers do scoring, they set up the room with ambient mics
and all you hear coming out of the rear are the ambient mics. But I'm a big
believer in using the rear speakers for more than that—we've got the speakers,
let's use them. I think if the upcoming buyer of a DVD is going to spend
some money, they want to hear more than just some reverbs out of the rear.
They're paying a higher price for the piece, and I think they want to be blown
away. I think they want to experience something new. So I tend to surround
the listener with music. I take chances. I'll put strange things in the rear.
(Scheiner, in Jackson 1999)

But though Scheiner was enthusiastic, most record companies were reluc-
tant to commit significant investments in untested formats and experiences.
Scheiner was a safe bet, and companies frequently hired him to oversee re-
mixes without consulting the original production team or artists themselves.

Record companies are going to the original guys who mixed some classic
records and saying, 'We want to redo this, but we can't afford you’… The in-
tegrity of the original is at stake… The record might sound okay because it's
in 5.1. But it might not have any bearing or resemblance to what the original
record was. I think you have to maintain some kind of integrity. (Scheiner,
ibid.)

One interesting episode reveals the tensions between insider/outsider au-
thorship of 5.1 remixes. Following the body of work Scheiner had crafted in
a few short years, he was tasked with mixing Queen's *Night at the Opera* in
surround. As a label-driven initiative, Scheiner’s mix was rush-released
without the approval of the surviving band members, though evidently with
the participation of Roy Thomas Baker, the original producer. Guitarist Bri-
an May quickly and publicly voiced his displeasure, and demanded that
Scheiner create a different 5.1 remix of the album that followed May’s spec-
ifications. Though Scheiner acknowledged the right of performers to be sati-
sfied with work issued in their own name, he also argued that his mix was
more exciting, more original, and implied that an artist need not always have
the final word.

Brian also had me move certain things. In 'Bohemian,' when it cuts to the
opera section, I originally put the piano in the rear for those quarter notes: da
da da da, da da da da. Everything went to black except there. But Brian said,
'Can we bring it out to the front a little?' So I did - and it just doesn't have the
same impact for me. I thought my version was good because it drew your at-
tention, it did something different, it wasn't ordinary. But that's my opinion. Brian's the artist. His name's on that record, not mine. (Scheiner, in Richardson, 2002)

But market forces stymied his momentum. No matter how loudly he championed surround sound, even going so far as to lend his name to a 5.1, high-end car audio system briefly included as a standard feature in Acura models, a confused introduction of competing formats for surround resulted in a lack of market interest, and the initial enthusiasm for surround as the “next big thing” quickly waned. Mixes that had been prepped for release were permanently shelved, and Scheiner began to focus on performances that might still utilize surround for the home video market, rather than audio-only content, leaving the surround market to other industry peers and upstarts.

Countering Scheiner, Bob Clearmountain, one of the first superstar mix engineers ignored the potential remix market with a few exceptions – notably remixes of albums for which he was the original stereo mix engineer. His argument was that most albums of the previous decades had been conceived as stereo, and thus surround remixes were akin to colorizing black and white films. However, regarding his remix of Roxy Music’s Avalon, Clearmountain expressed the notion that the many textural elements in the recording tested the limits of the stereo soundfield, and that had the 5.1 format been in existence at the time, he would have crafted a 5.1 mix as the ideal solution, grateful that the option to craft the new mix allowed the music to be heard as intended.

While (we) were working on the original stereo mixes in 1982, I can recall imagining the sound image as being more than just ‘stereo.’ There were so many wonderful things going on, I wished I’d had more places to put them than just two speakers. I wanted to be totally immersed in the album’s soundscapes, and tried as best I could to create as much depth as possible, nothing less than such a brilliant production deserved. That is why I was thrilled … to be involved with the 5.1 Surround mix of Avalon … It is how I’d always imagined this album should be presented – the surround experience actually drawing you inside the music… (We) did our best to keep the vibe and content of the original mixes intact while literally adding a new dimension to the listening experience…. If there was ever an album that cried out for a surround mix, I believe Avalon would be it. (Clearmountain, 2003)

Thus, the notion of authorship is seen as an extension of the original mix, rather than a translation of another author’s work. Like Scheiner, Clearmountain became a convert. But further highlighting the role of market and industry, Clearmountain, of his own accord, took the initiative of creating surround mixes simultaneously with the stereo mixes for every project he worked on since the early 2000s, only to see them all languish in storage (Clearmountain, 2015).
Pedigree matters, even if the connection to the original recordings is somewhat indirect. Nick Davis, engineer on later period Genesis was brought in to oversee stereo and 5.1 remixes of the entire catalog, including over a dozen albums that Davis had no original credits. Similarly, James Guthrie, the engineer for Pink Floyd’s *The Wall*, and Roger Waters’s solo career recordings, was commissioned to craft 5.1 mixes of *Dark Side of the Moon* and *Wish You Were Here*, though he was not a part of the original production team.

Like Clearmountain’s work on *Avalon*, Jerry Harrison’s 5.1 mixes of the Talking Heads catalog were generally well-received in large part because of his status as a central participant in the making of the original recordings. More experimental placement, in some cases fairly different approaches from the stereo soundfield and element mix levels were more easily accepted than if a similar mix had been crafted by an outside engineer. And like Clearmountain, Harrison touted 5.1 as the ideal format for certain titles such as *Remain in Light* and *Speaking in Tongues* that featured complex layers of instruments and voices. At the same time, he expressed some anxiety that surround sound might not offer much more of a heightened experience for other, simpler productions in the band’s catalog (Harrison, 2005).

For a brief moment, even contemporary recordings were issued in surround sound simultaneously with their stereo counterparts. As resident surround auteur, Elliot Scheiner was enlisted to oversee many of these mixes. Free from the canonical ghosts of stereo past, Scheiner employed far more drastic and experimental approaches to his mixes for Beck’s *Sea Change*, and perhaps most notoriously for Flaming Lips’ *Yoshimi Battles the Pink Robots* where the soundfield was treated as something in active motion rather than as fixed position, with a great deal of panning movement throughout the soundfield. Such dramatic approaches to the mix placement met with controversy from listeners either championing the heightened effect, or lamenting the dizzying distraction of such an active soundfield. But none of these contemporary releases were able to establish the format in the marketplace, and subsequent albums were only issued in stereo. Like Quad in the 1970s, the enthusiasm that artists and producers felt having their work represented in a larger soundfield was not met by their audiences who were content to experience the music in stereo, often under headphones, and increasingly in cars and later portable devices, none of which were suited to surround listening.

One significant experiment posed the most potential for establishing a larger market for surround sound – The Beatles’ *Love*. The project involved not only crafting mixes in stereo and surround, but treating the sacred musical elements preserved on a series of multitrack tapes in sometimes radically altered form, following the then-current “mash-up” approach, wherein elements of different recordings were layered on top of one another (Danger Mouse’s *The Grey Album*, a melding of The Beatles’ “White Album” and
Jay-Z’s *The Black Album* being perhaps the most well-known example). A radical approach to tampering with iconic recordings might have been the source of much controversy, but as the project was helmed by original producer George Martin, the question of authorship wasn’t called into play. In truth, the new mixes were created by Martin’s son Giles, a team approach that allowed an authority figure to guide and approve, while a mix engineer not present for the original work could approach it from a new perspective (Wills, 2008). The pedigree of Giles Martin’s indirect connection to the Beatles catalog is crucial to the amount of discretion he was given over some of the most tightly controlled audio artifacts on earth. Such radical reworkings would be unimaginable for any other potential engineer (and the relatively tame approaches to surround on both the *Yellow Submarine* and *Anthology* re-issue projects of the 1990s bear this out). The case of Giles Martin might be thought of as “author-by-proxy.”

Hearing pieces of several Beatles recordings juxtaposed against one another allowed fans to experience deeply familiar music freshly anew, and Beatleologists could approach the results as forensic evidence, a lifting of the veil that allowed elements to be auditioned outside of the context of the previously released composites. And while it took over a decade, Giles Martin was later tasked with reimagining the soundfield for the 2017 re-issue of *Sgt. Pepper’s Lonely Hearts Club Band*, in both stereo and 5.1. The stereo mix is a fascinating mash-up of 1960s psychedelic musical arrangements re-cast within a 21st century codified stereo soundfield. Drums in pseudo-stereo, with a significantly increased amount of low end frequency information, etc. The result sounds like second and third generation Beatles-inspired pop – musically imitative, but functioning within more contemporary approaches to stereo. Highlighting the difference between influence and pastiche, consider the recordings of post-Beatles homage that replicate the oddly discrete placement of drums in one channel, guitars in the other. Such a mix recalls a soundfield experience identified as “Beatles,” distinctly different from music that might be described as “Beatlesque.”

Giles Martin’s stereo mix retained much of the power of the original mono mixes, while tipping a hat to later period conceptions of the stereo soundfield. His 5.1 mix of *Sgt. Pepper* eschewed the more radical placement of elements in his *Love* re-conception, resulting in an experience that was more immersive than expansive. The re-issue met with critical favor, and the lavishly packaged, expensive box set sold in significant numbers that it placed in many top ten sales charts in its initial weeks of release. It is too soon to tell if the release actually expanded the audience for surround, though sales were healthy enough that rumors abound that the *White Album* and *Abbey Road* may see similar re-issue treatments. The very existence of a 5.1 mix indicates that there is a perceived audience for surround sound, an audience that has encouraged and sustained a healthy re-issue program of 60s and 70s classic progressive rock.
Steven Wilson, founder of second generation prog band, Porcupine Tree was introduced to surround sound when observing Elliot Scheiner craft a 5.1 mix of that band’s 2002 release, *In Absentia*. The possibilities of surround sound as a new standard listening format greatly appealed to Wilson, who proceeded to craft 5.1 mixes of his subsequent Porcupine Tree projects. Prog musicians weaned on the catalogs of early 70s progressive rock not only maintained a tradition of musical complexity and instrumental virtuosity, but extended the fascination with studio craft, and carefully detailed mixes. It is difficult to determine the extent to which the influence of 70s progressive rock was compositional or performative, and that of the approach to recording and mixing.

As economic forces drove older bands to yet again re-package their catalogs, it was determined that new value needed to be added to the updated product. Beyond remastering decades old mixes, a brand new audio experience might appeal to jaded listeners – new mixes in high-resolution stereo, as well as 5.1 surround would afford the opportunity for old fans to hear the music anew. But who could be entrusted with reshaping soundscapes into something fresh and new, without conflicting with deeply embedded memories of the music? As a second generation prog musician with growing acclaim and stature, with a number of production credits that exhibited all the hallmarks of 70s-era studio craft, Wilson was positioned at the nexus of past and future.

Musicians such as Robert Fripp who had endured protracted legal actions in order to gain control over their master recordings were unlikely to hand over authorship to industry anointed figures such as Elliot Scheiner. But a fellow musician might approach the remixes from a fresh angle, while being deferential to the older generation reluctant to hand over their hard-won legacies. Impressed with Wilson’s work as both musician and producer, Fripp asked him to create new stereo mixes of King Crimson’s catalog. As an enthusiast for surround sound, Wilson suggested that 5.1 mixes might also be of interest to Crimson fans, as well as the audiophile market willing to embrace any music that came at them from all sides.

The initial release of the remixed *In the Court of the Crimson King* brought Wilson accolades from fans and reviewers, and not only was he asked to oversee the entire Crimson catalog re-issue, but was soon approached by the stakeholders in several other progressive-era bands – Yes, Jethro Tull, and Gentle Giant. Before he knew it, Wilson had a second career as the default auteur of 70s progressive rock for 21st century ears. With new audiences discovering him through his production work, Wilson’s solo career as a musician also benefitted from an expanding audience, with audiophiles embracing his own catalog as if it were part of the pantheon of canonical progressive rock.

But as a second generation progressive music fan, Wilson also grew up with 80s pop music, and his sense of an exciting mix owes as much to the
more fragmented, programmed, and post-automation music as it does to the trappings of progressive rock. So while he continued to oversee re-mixed releases from 70s rock acts, he also put his stamp on stereo and surround remixes of albums such as Tears for Fears’ *Songs From the Big Chair*, and the ongoing reissue program of the XTC catalog, as well as applying his perspective to contemporary prog/metal acts such as Opeth.

With a credit list of surround mixes that rivals Elliot Schiener, Steven Wilson has become a *brand* to a growing number of audiophile-oriented consumers, willing to take a chance on music that falls outside their normal musical tastes, simply to experience another iteration of a Steven Wilson soundscape. On audiophile chat threads, Wilson is frequently invoked as the ultimate authority on surround mixes, with occasional asides along the likes of, “I’m not really much of a (fill in the blank) fan, but if Steven Wilson has done the surround, I want to hear it.” Wilson is quite aware of this, and has capitalized on this emerging market by approaching his own music from a surround sound perspective, noting,

> There are, actually a lot of people out there now… a growing audiophile community who love things that sound great. … And there’s surround sound; there are people out there that care about that. They are a minority, but they are a substantial minority. And they’re growing. (Wilson, 2017)

If Jerry Harrison and Bob Clearmountain claim that certain productions cry out for surround sound treatment, the reverse corollary might also be true – other productions do not warrant expanded soundfields. This distinction implies that production style differences in recordings are as crucial as musical ones in determining genre distinctions. And production approaches that are consistent across a variety of artists and a broad spectrum of musical styles might align such work under a commonly shared heading. Even if tempo, lyrical themes, instrumental and vocal styles and timbres are widely disparate, there is an attention to sonic and mix detail that unites Yes, Tears for Fears, and Opeth fans. It is for this reason that I posit that Steven Wilson represents the emergence of a new sub-genre – “surround rock.”

Whereas audiophiles of the 1950s and 60s were often painted as obsessed with frequency response and conspicuous sound system consumption, current audiophiles spend as much time in chat groups praising the musical attributes of the recordings they embrace as they do citing impressive sonic replication. I argue that listeners feel that a carefully crafted mix honors their dedication to both audio *and* music, and they can be as excited by programmed percussion loops as they are by guitar arpeggios, provided the soundscape that results reflects an attention to detail in both songcraft and soundcraft. Just as genre distinctions place music in different categories and cultural spaces, surround mixes assign sounds to discrete, localized positions. Yet this fragmenting of sound can also create the sensation of an im-
mersive whole. If rock fans seek out music that provides dynamic energy, while jazz fans look for spontaneous invention, or classical fans a world of ordered precision, surround sound aficionados desire sonic immersion, whether it is Opeth, Roxy Music, or Yes. It is a type of listening that unites these recordings, and sets them apart from other musical experiences – a form of genre unto itself.

For many surround sound listeners, it is a history of mixing approaches that unites otherwise distinctly different musical genres. A fondness for spatially placed delays across a soundfield may be formed from an acquaintance with Pink Floyd, Lee “Scratch” Perry, or New Order, and generate a positive response when encountered in recordings in any musical genre. David Brackett, working from Mikhail Bakhtin, underscores the importance of historical practices in defining genre distinctions.

Musical texts, in the process of citing the conventions of genre, are ‘shaped and developed in continuous and constant interaction with’ the musical texts of others working in similar genres. Each musical text ‘is filled with echoes and reverberations of other’ musical texts ‘to which it is related by the communality of the’ musical genre. Every musical text ‘must be regarded primarily as a response to preceding’ musical texts of a given genre. (Brackett, 2016: 15)

Brackett goes on to posit that,

the meaning-producing relationship extends beyond what occurs between a text and a listener, and depends on a feedback loop in which ideas and assumptions about genre circulate among music producers (musicians and music-industry workers), audience members, and critics. (ibid: 16)

Steven Wilson’s statement about a growing subculture of surround sound and audiophile listeners illustrates an awareness of audience that reflects the type of “feedback loop” that Brackett describes. Listening modalities formed over a period of time, that include even a rough analysis of audio production practices result in audiences able to discuss the placement of sonic elements in a soundfield, degrees of digital limiting and compression, etc. as well as responses to musical performance and composition. Indeed, for many of these listeners, there is no distinction made between musical and sonic text.

Musical/sonic relationships are present in all forms of recorded music, but the frequent comments made by audiophile listeners that make this relationship explicit, as well as the self-identified niche audience, cult or sub-culture designation of being an “audiophile,” correlates listening practice and production technique preferences to individual identity. Indeed, it is the limited range of this subculture that highlights hi-resolution/surround sound as a genre distinct from other forms of audio recording, and further blurs the distinction between “musical” and “audio” texts for these listeners. A similar
argument could be made for fans of monaural recordings – whether as fans of Enrico Caruso, Charley Patton, Phil Spector, or Jack White. It is the preference for specific forms of audio playback that sets these audiences apart from the stereophonic normative “mainstream,” a subculture that emerges in opposition to a norm, a community framed by genre, a genre based upon the listening experience of particular audio replication techniques, established and identified by sonic auteurs such as Scheiner, Martin, Wilson and others.

References

Bibliography
https://vimeo.com/9860895

Filmography
All Together Now: A Documentary Film, Wills, A. (dir.), Apple/Parlophone DVD 5099921706790, 2008.
Ask Me Nicely (the making of “to the bone”), Steven Wilson, to the bone, Caroline 016BR, 2017.
Jason Woolley: The cultural politics of using technology to support the aesthetic in jazz record production

Abstract

This paper discusses the attitudes some contemporary jazz musicians have toward the use of technology and the thresholds of studio ‘intervention’ they are willing to cross in order to achieve their preferred studio recording aesthetic. The discussion draws upon personal experiences of mainly free jazz production, and also the returns of a pilot survey of jazz musicians, who were polled on the subject of the use of technology in the production of jazz recordings. Grounded theory was utilised as a methodology to code the survey responses into two main categories of ‘idealistic’ and ‘pragmatic’ responses. This coding enabled a discussion of the themes which were evident in the responses. This limited study found that clear consensus on the extent to which studio techniques should be used in the production of jazz recordings was not apparent in the data. There were instances where some studio techniques were considered acceptable, whilst others which were as similarly technically invasive as the acceptable, were not acceptable. It also appeared that it was generally acceptable to edit pre-composed elements of the musical performances, but not acceptable to edit improvisations. This perhaps reinforces the notion that for the jazz performer the ephemeral ‘moment’, which many studio productions aim to represent, occurs when they are improvising. Finally, it appears for these participants, whether a studio intervention such as an edit is audible in the final recording is irrelevant. The fact that a studio intervention was required or occurred due to perceived deficiencies in the recorded performance, crosses the boundary of acceptability in terms of their own interpretation of ‘authenticity’ in jazz recordings.

Introduction (background and context)

Both jazz and classical music are played by highly trained performers, who have spent a lifetime perfecting their ability to read music, and in the case of jazz, to improvise a certain way. Thus these forms of music are generally recorded live, without overdubs, and with minimal, if any, post-performance enhancement. (Moorefield 2005: xiv)
We are aware from a number of writings that for many jazz musicians there exists, what we might refer to, as tensions between the jazz community’s ideological notion of ‘authenticity’\(^1\), and how practitioners respond and order their practice upon this, including how they engage with processes of recording and producing a recording of jazz music. This tension manifests itself in the potential for studio techniques to be used to present a ‘dishonest’ representation of a performance. Consider John Hammond, jazz producer, in 1953, quoted by Khan on the use of editing via tape splicing. ‘It made it possible to make more dishonest records than ever before, through splicing’ (Kahn: 2000, p.110). As Khan points out for ‘traditionalists’ like Hammond, ‘[I]f it altered the sanctity of the improvised jazz moment, it was simply wrong’. (Khan: 2000, p.110) Although certainly not exclusively, those jazz musicians who aim to achieve a ‘live’ in the studio aesthetic, whilst striving for great sounding records, appear to also want to ensure that the production process is transparent, and in being transparent, there is no sense that, when the recording is disseminated, what the listener is hearing is something which has been manipulated beyond the artist’s ability to play. Jago (2012) has made discussion of the challenges of reconciling the use of the studio with jazz authenticity when presenting the example of Tristano, and although the topics discussed share comparison, the context is somewhat different to this study in that the focus in this research is on the ‘live in the studio’ recording aesthetic rather than the use of extended studio techniques as an artistic tool. This study attempts to understand the attitudes jazz musicians have toward the application of studio techniques such as pitch correction, rhythmical edits, overdubs, and/or other methods for ‘improving’ or creating a more convenient version of a recorded ‘live’ performance in pre and post-recording contexts. I will occasionally refer to the application of these sorts of techniques for the latter strategic aim as being studio ‘interventions’. Although some of the techniques might be considered as extended studio techniques, they can also be considered interventions when the purpose is to create a more convenient sonic aesthetic.

Across a broad range of music which falls under the category of jazz there exist varying levels of engagement with studio based sound manipulation techniques. We might consider that in the case of those practitioners who seek a ‘live’ aesthetic from the studio, studio manipulation is supposedly minimal, and that this contrasts to instances where the use of ‘extended’ studio techniques in a post-recording context are integral to an artistic vision. It is interesting that at both ends of this spectrum, ‘to the point of parody’ according to Hegarty (2007, p.10), artists often like to present a kind of account of the production process which some would argue attempts to frame the work within the boundaries of the acceptable and authentic in order to

---

\(^1\) I will define ‘authenticity’ in the context of this study later in the paper.
‘gain acceptance with an audience’ (Zagorski-Thomas: 2010, p.261). Some might also argue that liner notes are a marketing technique and/or merely an addition to the product that is there to contribute to the listener’s experience, by providing a context and perhaps story backdrop to the production.

In the context of this study, the use of studio techniques as correctional tools are unlikely to be explained and presented on album liners, as for many jazz practitioners and fans this would demonstrate studio strategies that stretch the boundaries beyond the notion of an authentic representation. This study is concerned with consulting jazz practitioner’s views on where studio interventions are appropriate and if there is a threshold point where the amount or type of intervention travels beyond the acceptable.

‘Authenticity’ for ‘traditionalists’/ ‘purists’ and beyond

As Grazian and others have written, authenticity is a construct rather than an ‘objective quality’ (Grazian in Bennett and Peterson 2004). Although more than one notion of authenticity is apparent in the practices of the respondents discussed in this writing, the main focus of this paper is concerned with the status of jazz recordings as documents or ‘inscriptions’ of the moment (Corbett in Gabbard 1995). Therefore, in this context a recording of jazz might be ‘authentic’ when it is considered a document or inscription of the moment that was the original recorded performance. In theory such a recording can therefore not contain anything other than the musicians performing together at the same time, or the soloist performing alone in what was in the moment. This type of recording of jazz is considered authentic because it purports to represent what happened in the moment during the performance that was recorded and was ‘unaltered’ by further overdubs, musical edits or processing. Milner has discussed how recording engineers such as Laico sort to create for the record listener a sense of being in the ‘presence’ of the musicians by ‘…mak[ing] the records as “transparent” as possible’ (Milner: 2007, p149). Of course when a deeper consideration of the recording process is made, we understand that a recording is never a neutral or a ‘transparent’ representation of the moment (Page 2009).

As authenticity is a construct, a shared belief between groups of individuals, there are those within the data sample, and beyond, that might argue a different notion of what is authentic in the recording and production of jazz. In this paper, when referring to the term authenticity, it is the former definition of the term which is of most interest to the discussion of this study.
**Liveness and the fiction of recording**

Auslander (1999) has asserted that an important part of jazz music, for the genre’s players and fans, is in its ‘liveness’ within the ‘spur of the moment’ (Taylor quoted in Jago: 2013, p.no page number). This is in contrast to Rock music, where often the aim of live performance is to recreate the idealised studio recording (to affirm the authenticity of the music). Whilst recreating the sound of a live recording for jazz musicians is important, authenticity for ‘purists’ is affirmed from their ability to recreate moments like (but not the same as) all the moments heard on their fixed recorded version. Therefore we might consider that in many instances, the challenge for some recording jazz performers (and their producers) is to avoid unwanted mediation in the production process which distorts the moment beyond what was actually played in that moment and this might be achieved perhaps by limiting studio interventions within the production process of their performances. Many will agree that the avoidance of mediation in a studio recording context is impossible, as the studio processes always impart something, and as Crooks puts it, ‘...the sonics of jazz recording are not true to the original performance–no recording is’ (Crooks: 2012, p. no page number). We might also consider Anderson’s assertion that ‘[A]ny recording is a ‘fiction’, a falsity, even in its most pure form’ (Anderson quoted in Page: 2009, p. no page number), and furthermore Williams’ discussion on the status of sound recording as being a ‘reading’ of an event.

The practice of sound recording implies by definition a reading, a deciphering, an attending to a sonic event. This is the meaning of the difference between the three-dimensional physical "sound" and its recording as a one-dimensional analogically encoded event (Williams: 1980, p. 12).

**What do the jazz musicians themselves actually think and consider?**

We know something about how some jazz practitioners and fans view the recording process and how they consider that it can be considered to have impacted negatively on the ideological authenticity of their recorded music, from anecdotes and interpretations of other clues such as album liner notes and scholarly writing. Jago (2012) has documented Atlantic’s inclusion of a disclaimer on Lennie Tristano’s ‘The New Tristano’ outlining that extended studio techniques had not been used. Jago goes on to explain that this was in response to a critical and fan backlash to the overt use by Tristano of extended studio techniques on his earlier record ‘Tristano’ and here Jago appears to be suggesting that the liner notes were a commercial decision by the record label in response to a backlash to Tristano’s earlier work. Tristano’s use of extended studio techniques was an artistic consideration, and not a studio intervention as I have outlined here. Jago has documented that although many might consider the use of extended studio techniques to be beyond the
threshold of jazz authenticity, Tristano himself did not. Jago’s rational for the backlash to Tristano’s work can be suitably applied to explain why studio interventions might not be acceptable for some practitioners and fans.

The record is held to be an artifact of a live event, and even if it were not recorded in a “live” setting complete with audience, but rather “live in-studio,” the general assumption has been that the musicians were subject to the same constraints – the same improvisational demands – as they would have been in a club or concert setting (Jago 2013).

Intervening to change some element of a performance on a ‘live’ Jazz recording takes the constraints on the performers beyond the club or concert setting.

Within the study of Popular Music, Zagorski-Thomas (2010) has discussed the example of the liner notes of Queen’s early albums, which may or may not have been (according to Zagorski-Thomas) for the purpose of anchoring Queen’s early records within the acceptable paradigm of Rock authenticity, as it stood at the time. Zagorski-Thomas outlines what appears to be a contradiction in that whilst Queen exploited the use of extended studio techniques to achieve their artistic vision of over-dubbed guitar parts, Brian May and Queen felt the need for an album liner note disclaiming the use of synths in the production of the music, perhaps to reassure their fan base of their authenticity as a Rock band. This example has some parallels to the Tristano example given by Jago, where the artistic aim results in the creation something that cannot be recreated live without some form of extension to their respective ensembles, be that in the form of additional ensemble members or the addition of electronic augmentation such as backing tracks. It also suggests something of a threshold to the application of studio techniques, in that at the time of the Queen albums discussed, it was acceptable to overdub to a point where recreation beyond the live capabilities of existing band members was impossible but not acceptable to use a synth, as ‘[S]ynthesisers, [were] once seen not as musical instruments but as machines that had no place in rock...’ (Auslander 2008, p.84)

I am interested in instances where the studio techniques are being used to enhance a ‘live’ performance. That is to say, no additional musical parts are added other than those individual and singular contributions made, solo or ensemble based, during the recording and therefore if overdubbing occurs, it does not involve anything beyond overdubbing more than each ensemble members initial part. This study is interested in the use of the studio as a tool in order to achieve the artists ‘ideal’ live recording in the studio, and where it is that the threshold for studio processes and interventions is crossed, rendering the finished recording as unauthentic for the artists themselves. Previous studies on this topic have often concentrated on the reception and perception of the listener and critic, this study aims to look toward an understanding of
the view of the artist/practitioner. So for the artist, for example, is it acceptable to ‘correct’ the less appealing aspects of a recorded ‘live’ performance using studio techniques? And in instances where studio techniques are applied to ‘correct’ or enhance a recorded moment post recording, what is the threshold upon which a recording travels from ‘authentic’ to ‘unauthentic’ in the perception of the performer(s).

My own experience of engineering jazz recordings revealed tensions between jazz musicians wanting to respond to the ideological demands of authenticity, and the desire to produce a sonically satisfying recording of their moments. As one respondent to the survey said, the challenge when producing a jazz record with an intentional ‘live’ aspect is to present an account of the music, through recording, which reconciles with the ‘...album format.’ On two recent free jazz productions for The Markov Chain (2014), and Spirit Farm (2015), I found that the tensions between ‘authenticity’ and the impact recording/post recording process has on the perception of this did exist. However, there was no clear consensus from any of the 8 performers involved in these separate projects on where the threshold before the impact of studio intervention renders the recorded representation beyond the authentic was or is. To give an example, performers in the one session requested limits to the number of microphones to be used in order to limit post-recording rebalancing, whereas performers in the other session did not make any similar demands. Understanding the rationale of the jazz performer in a studio context I feel would contribute to practice. With this goal in mind, I felt it would be useful to poll jazz performers beyond the confines of my practice and experience, through online survey polling of jazz musicians who partake in recording sessions.

Method: The challenges of survey design

An anonymous online questionnaire survey of jazz practitioners was considered a possible way of supporting and expanding upon my personal practice led insights. Whilst the survey has quantitative and qualitative aspects to its design, due to the limited number of responses, the discussion of this paper focusses on the qualitative aspects of the data collected. Given the challenges with survey ‘noise’ such as ‘socially desirable bias’ outlined by Neumann (2014), it was felt that participant anonymity would serve to gather more accurate responses from the jazz performers. The latter is considered a two-fold design consideration. In the first consideration the anonymous framework would help support and encourage respondents to respond faithfully and honestly without concern for presenting socially un-desirable responses. The 2nd consideration of the same concern guided the choice of online survey as opposed to other approaches to data collection such as structured interviews. It was felt that within this pilot study, asking participants to com-
communicate their responses personally to the researcher may have resulted in data noise in the form of socially desirable bias. Furthermore, the ‘online’ approach with a web link to the survey rather than email or other means to identify participants, gave the respondents further assurance that there was no option for the researcher to trace electronic identity or other information. However, none of the latter two considerations are considered assurances that participants did respond with complete honesty.

Method: Jazz style and sub-genre

Given the complexities and debate on defining jazz practice, the survey design had to consider the question of jazz sub-genre from two main perspectives. Firstly, there appears to be no absolute collective agreement within the jazz community on terminology when describing practice. The survey navigated questions on terminology by asking practitioners to describe their practice in their preferred terms which, as expected, resulted in a number of different terms being used by participants such as ‘free’, ‘free-jazz’, ‘plain jazz’, ‘improvised music’, ‘straight-ahead’, ‘large ensemble’, ‘big band’, ‘small ensemble’, ‘electronic’, ‘traditional, and ‘modern’. The 2nd consideration was the possibility that practitioners were likely to practice more than one sub-genre or elements of a sub-genre. Therefore the survey design allowed participants to specify more than one sub-genre of jazz practice if they wished. The advantages were that data collection would not exclude any type of jazz practitioner and would not be hampered by entering into a debate on terminology which is not considered the aim of the study.

A pre-considered disadvantage to this methodology is that where a participant indicates more than one style or sub-genre in their practice, it might be more challenging to identify which of their practices they are referring to in response to a survey question. Therefore a ‘partially open ended’ questioning framework was utilised in the survey design (Neumann 2014). A multiple choice element formed part of the survey design and each participant was given the opportunity to choose more than one answer from a number of presented responses including ‘none of the above’. Participants were also encouraged to comment and qualify their answer to a multiple choice question by referring to their practice in the free text comments section available for each question.

Method: Survey design - Language

The survey design consciously avoided language such as ‘authentic’, mostly because I wanted to avoid the survey responses, especially in the free text areas, veering toward a discussion on the topic of authenticity in jazz (although some respondents did have a view to offer regardless of my intentions). Also, it was felt words like authenticity might well influence
responses toward a socially desirable bias resulting in additional noise within the data.

Alongside ‘socially desirable bias’, other considerations of where ‘noise’ may occur included the possibility that participants might not understand some of the language and terminology required to outline the studio recording context of the questions. Therefore, the studio process questions were as ‘plainly’ worded as possible in order to assist in the gathering of views from those practitioners within the sample who were not as experienced and familiar with studio production process terminology.

Participants and recruitment

Jazz practitioners who engage with studio recording and production as performers were invited to engage with the survey via an internet link to the survey which was created using the BOS online survey tool (https://www.onlinesurveys.ac.uk/). My institution currently requires researchers gathering data by survey to use the latter survey tool. The survey design was subject to ethical approval including providing participants with a Participant Instruction Sheet prior to their engagement with the survey itself. Invites were sent out over a six-week period via email from myself and also via the networks of jazz musicians and researchers I have contact with, and participants generally would have been from networks within the UK, Portugal and Sweden. There were no other restrictions to the eligibility to take part in the survey other than the above and the data gather showed that participants spanned an age range of adults up to the age of 60 or above.

Data Limitations

There are of course limitations to what can be concluded from the data collected. The partially open ended framework of the survey design and also the opportunity for multiple responses from participants which occasionally allowed participants to give more than one answer if they desired, means the quantitative data is not reliable enough to draw any conclusions. If the data collection could have been restricted to, for example’ practitioners who perform and record free jazz only, and the survey responses had been more numerous, a much more traditional scrutiny of the quantitative data would have been possible. In my analysis, the quantitative elements are occasionally referred to in the discussion of the qualitative information supplied by participants in the free text comments box of each question.

This is a small pilot study which was open for participation for a relatively short period and avoided techniques for gaining broader engagement such as unsolicited contact. I asked members of my own small network, whom I mostly knew through recording activities, to engage with the survey, and to also distribute the survey link to their own network of contacts. Thirty-three
Results and Analysis

For the analysis of the data, I used Glaser and Strauss’ Grounded Theory approach as detailed by Urquhart (2013) as a method for the analysis of qualitative data and this involved data-driven coding or ‘open coding’ of the survey responses (Gibbs 2012). Although I am ‘...an observer of the social world and a part of that same world...’ I wanted to begin with as near to an open approach as possible (Gibbs: 2012, p9). This seemed appropriate as before the coding began, anecdotal evidence seemed to indicate that the survey data would potentially contain a broad range of viewpoints which might not necessarily mirror the viewpoints of those jazz musicians I had produced recordings for in the past. Furthermore, a similar approach has also been successfully used by Brand et al (2012) on a smaller number of twenty-four responses in their qualitative study of jazz performers and audience members recruited from a series of eight gigs at The Vortex Jazz club in London.

When reading through the responses collected I noted some themes within the responses to a number of the survey questions. In my analysis I coded these themes on whether they were either ‘idealistic’ or ‘pragmatic’. By idealistic, I mean where the decision making process would appear to be motivated by a rationale that is abstracted from the sonic aesthetic of the outcome, and relies more on an ideological belief of what is acceptable practice in the production of jazz recordings. A decision or choice that is motivated by a belief which places the utmost importance on the process irrespective of whether this might negatively impact on the sonic aesthetic of the outcome. As a polarity to idealistic, pragmatic coding refers to where the response indicates that a clear strategic goal of achieving the best sonic aesthetic, regardless of the processes involved, is of the most importance to the respondent. Once the codings were completed, I identified some theoretical insights for discussion on how these jazz musicians interact with the recording studio and associate processes, and how this interaction is reconciled with the demands of their community of fans and fellow practitioners.

Pragmatic

As discussed, this is defined as where the response in some way indicates that the quality of the sonic aesthetic supersedes any concerns about the pro-

---

2 This approach was not appropriate for all the questions and responses within the survey framework. Some questions were designed to enable a contextual analysis e.g. age of respondent, and sub-genre of jazz practice.
roduction process introducing elements that are not part of the original moment. For instance, this could be where the music and/or conceptual elements are the most important aspect of the outcome, rather than adherence to a guiding set rules which govern what is for instance politically or culturally acceptable as studio practice in a jazz production context. Where there is no ideological or other concern that the studio process is impacting on the authenticity of the recorded performance as a document of music in the moment, and the only concern is that there is a convenient musical and sonic outcome, and/or the conceptual starting point is maintained. Here are some example responses that fall into this coding:

The process is driven by the work, at times it is relevant to capture the moment of spontaneity unchanged at others the integrity may be better served assembling in post to represent the intention. Process is flexible and unique driven by the conceptual goal of an individual project. [in response to survey question 5]

Getting the idea out how you want it to be is the most important, so all the above are acceptable. [in response to survey question 5]

*Idealistic*

As explained above, this is defined as where the response in some way appears to indicate that a belief, whether social or cultural, dictates what is acceptable in terms of recording workflow when producing the recording. For instance, this might be where the purpose of the production process is to present the ‘liveness’ of the moment in a truthful, honest, and unaltered way. The responses categorised in this coding bracket, demonstrate an approach which places the utmost value on a belief that jazz performance is a ‘live’ activity, and the recordings of jazz performances must represent a faithful representation of that moment – away from the ‘adulteration’ of post recording interventions. Here are some example responses from the survey data:

…the more improvisation there is, the less post-production there should be. [in response to survey question 5]

They’re all acceptable practices but with my own work, I’m not fully comfortable with excessive edits/overdubs within improvised contexts. [in response to survey question 5]

---

3 Although ‘Idealistic’ refers essentially to an ideological motive, to avoid confusion and a debate outside of this paper, I have avoided the use of the term ideology as a coding category. For example, ‘pragmatic’ practice may be driven by commercialisation which could also be argued as an ideological motive.
Limited overdubbing etc., regardless of listening stance of audience, as long as the overdubs are improvised (never the same twice). [in response to survey question 5]

Also assigned to the idealistic category were those instances where an importance appears to be placed on the collective viewpoints of the activists of the session. This appears to suggest instances where some type of ensemble ‘democracy’ is at play, where the collective decide what is best for the music, rather than a dictation by any individual activist. A democratic idealistic approach which has the potential at least, to supersede the musical and/or conceptual outcome as a factor of importance. Here are some example responses from the survey data:

The answer would emerge through discussion with everyone involved. We'd reach a consensus and go with that. [in response to survey question 9]

All depends on the wishes of the other performers. [in response to survey question 9]

Discussion

Theme 1 – What is acceptable as an intervention and where is it acceptable?

Jazz is, however, a genre where the cultural norms are rather different in terms of the settings and conventions of performance. It places improvisation at its core, and so its dominant discourse puts the performer, in the specific here and now, right into the driving seat (Crooks, 2012).

When considering processes such as how much studio intervention is acceptable to the jazz artist in the production of a jazz recording, the data unsurprisingly demonstrates a polarity of views. There are those coded as ‘idealistic’ where the viewpoint collected indicated that factors outside of the importance of the sonic aesthetic appear to influence the participant’s practice and use of the studio. There are also those I have coded as ‘pragmatic’ where viewpoints indicate the participant’s goal is to achieve the best sonic aesthetic regardless of the processes involved and included comments such as ‘[T]he studio is a tool to be used in any way possible in order to produce the best result’. For those respondents that gave viewpoints coded in this study as ‘idealistic’, there was some indication of where the threshold for studio intervention procedures would be deemed to contravene the boundaries of an authentic recording as defined for this study. In some of the free text responses to the survey questions, some participants indicated that stu-
dio interventions were unacceptable on areas of the music which were originally improvised by the musician(s). Comments included the following:

I wouldn't want any editing done on a free improv[sic] recording…

I'm not fully comfortable with excessive edits/overdubs within improvised contexts.

Personally I don't like to do any overdubbing at all, as my aim is to present accurate representations of the music made in the moment. This is because improvisation by the players makes up most of the music being recorded, meaning that there are no 'mistakes' that need to be corrected with overdubbing.

Limited overdubbing etc., regardless of listening stance of audience, as long as the overdubs are improvised (never the same twice).

With improvisations I would be more restrictive towards post-processing.

This is in contrast to pragmatic comments made on the composed or fixed musical elements of a recording.

...but on a piece that was fully composed, editing and splicing takes where necessary [is ok].

If it’s plain jazz with mainly acoustic instruments then you should just do the necessary work in the post-recording process.

I would tend toward unlimited post-recording action but only in the sense of faithfully presenting the original concept.

However, if I found myself having to record heavily-notated big band music, for example, where the accurate performance of the notation is priority, I would use overdubs to correct mistakes if necessary.

These free text comments appear to indicate that, for this small survey, and for some respondents within the data collected, the threshold of studio intervention is not found in the nature and amount of processing, the threshold appears to be at what type of musical element is subject to the intervention. It appears that for some respondents, the threshold is whether the section which is subject to an intervention was originally an improvised passage of music. For some participants it is within the boundaries of their authenticity to intervene where there appears to be a less than ideal take of a performance of composed music, but to intervene on an unsatisfactory improvised element, other than to re-record an alternative different version of the whole musical piece, sits outside of the boundaries of authenticity. We might conclude that for these participants, the ephemeral interaction of the moment
occurs when they are improvising. Anything that is pre-composed and recorded exactly as scored prior to the recording session could be considered as being pre-created and outside of the moment. For some respondents, by applying studio editing techniques to unsatisfactory improvised elements, we are distorting the moment and creating an alternative reality of what happened in that moment, and perhaps by doing so, creating a record of another and different moment of interaction.

Theme 2 – Jazz performer’s perceptions of where and how the studio intervenes and the sliding scale of acceptability

From the responses to this survey, as a group, the respondents demonstrate a broad range of viewpoints on what is acceptable as studio practice. Individual responses also demonstrate what we might consider either as a non-acceptance that production techniques are never impartial, or, that there is a sliding scale of what is, and what is not acceptable as studio practice. This is evident when individual responses appear to demonstrate idealistic and also pragmatic responses to different questions in the same overall return. Consider the comments from the same participant shown in example 1 below:

**Example 1:**

Personally I don’t like to do any overdubbing at all, as my aim is to present accurate representations of the music made in the moment. [in response to question 5]

Sometimes when a group has recorded multiple extended improvisations, I think it’s acceptable to lift out sections to make tracks for an album… I see this excerpts approach as a way of reconciling free improvisation with the album format… [in response to question 12]

In example 1, the respondent aims to record music in the moment, but considers the reordering of the moment in post-production acceptable practice. Corbett would consider that the recording strategy referred to in example 1 is compositional practice rather than free improvisation.

…recording involves the post-facto selection, editing, organization, sequencing, titling and packaging (all compositional, not improvisational considerations) of music that has been made by means of improvisation but is now repeatable and fixed (Corbett in Gabbard 1995).

Example 2 (shown below) also demonstrates idealistic and pragmatic responses in the same individual’s return. Whilst there appears to be a commitment to capturing the moment, it appears to be acceptable to edit the moment if it is not a convenient moment.
Example 2:

I prefer to record live as part of an ensemble. Jazz is an "in the moment" music which largely depends on the rapport between the musicians in time/space. [in response to question 8]

I have released recordings with "mistakes" but also have inserted corrections if the mistake means the music suffers to a degree that I find unacceptable. [in response to question 9]

Whilst in the context of musical performance, the dynamic, pitch and timing are all part of the performed moment, the responses in Example 3 (shown below) demonstrate that the participant is comfortable with the post-recording rebalancing of the dynamics of the recorded performance, but not as comfortable altering pitch and timing of the recorded performance. I would argue, that in terms of the moment, altering dynamics, pitch or timing are all equally invasive to the presentation of the original moment.

Example 3:

To change/enhance the levels and dynamic balance of the music where this is lacking in the original performance. [in response to question 10]

I would only alter pitch or timing in exceptional circumstances. [in response to question 11]

As outlined above, at this stage it is not clear whether participants accept or have fully considered the impact recording and the processes involved have upon of their music. It is my intention that this matter will be explored further in my plans for the future development of this research.

Conclusion

Within the data collected, respondents demonstrated approaches and attitudes to the studio recording and production of jazz that I was able to define and code as ‘idealistic’ or ‘pragmatic’. The idealistic approaches often relate to the production recordings of improvised music which may be where, for these respondents, the jazz moment resides. The responses coded as pragmatic often demonstrated an ‘anything goes attitude’ to the use of the studio ‘... as a compositional tool’ (Eno in Cox and Warner: 2004, p.115). Some responses demonstrated a contradiction in approaches to studio practice, with individuals advocating clear idealistic studio strategies in some areas of the survey, but demonstrating that they were also willing to undertake pragmatic action in order to enhance their original recording in other areas of the survey. For example, it
could be argued that whilst no overdubbing might help to create a somewhat ‘honest’ document of a jazz moment, the use post-recording editing creates something which is further away from the original moment. In the case of these musicians, there appears to be no agreement on where the limit of studio intervention is, and there appears to be something akin to a ‘sliding scale of acceptability in terms of jazz production practice (Grazian in Bennett and Peterson 2004).

A final thought on this pilot study is that it is interesting to consider that a studio intervention, as a corrective tool, when applied effectively, should be inaudible to the listener, therefore rendering it ‘invisible’. The only people that will be aware of its existence are the production team and the artist(s). However, as this data collection reveals, for some participants the fact that an intervention was needed seems unacceptable regardless of whether anyone beyond the production team is aware or not of any corrective intervention in the final version.

Future research

Further research aims to gather additional data through semi-structured interview as this will hopefully provide the opportunity to discuss with jazz practitioners the points raised in this pilot study in more detail.

Acknowledgements

Thanks to Adam Fairhall, Jose Dias, Anton Hunter and Pete Dale for comments and support. Thanks also to PCAR at MMU for supporting the development of this research. Thanks to those who completed the survey.

References

Bibliography


**Discography**

The Spirit Farm, _The Spirit Farm_. CD SLAM. CD299, 2015.


Miles Davis, _Bitches Brew_. LP Columbia Records 1970.

**Appendices**

**Survey Questions**

1. Please confirm which category best describes you as a Jazz practitioner
   a. I am a Jazz performer aged between 16-29 years old.
   b. I am a Jazz performer aged between 30-39 years old.
   c. I am a Jazz performer aged between 40-49 years old.
   d. I am a Jazz performer aged between 50-59 years old.
   e. I am a Jazz performer aged between 60 years old or above.
   f. Other

   If you selected Other, please specify:

2. Please describe the type(s) of Jazz music-making you take part in when recording. I understand there may be more than one type you take part in, so please include them all. Some Jazz musicians align themselves to practice terms such as ‘Free’ or ‘Straight-ahead’ or ensemble terms such as Trio or Big Band. Whilst these terms are fine, if other terms and language are more appropriate to you as a practitioner, please feel free to use them. This question asks you to type in the comment box below.

344
3. What are your priorities when choosing to organise or perform in a Jazz recording session?
   a. Budget
   b. The availability of a good sound engineer
   c. The availability of good equipment
   d. Other (please specify and comment in the box below)

4. In your opinion, when a Jazz recording is produced for dissemination, please tick which are the most important of the following statements. (Choose more than one if you wish)
   a. The production process and strategy should strive to faithfully represent the performances captured during the recording session without any post recording editing or manipulation.
   b. The production process and strategy should strive to present the performances captured during the recording session in the best way possible, and in order to achieve this aim, post-recording editing and processing should not be overly used.
   c. The production process and strategy should strive to present the performances captured during the recording session in the best way possible even if this involves unlimited post-recording editing and processing at the mixing stage.
   d. Other (please specify and comment in the box below)

5. In your opinion, which of the following are acceptable as part of Jazz music production. (Tick more than one if appropriate).
   a. Unlimited overdubbing, editing and processing of the recorded performances as part of the production process regardless of whether this is noticeable on the final recording.
   b. Unlimited overdubbing, editing and processing of the recorded performances as part of the production process as long as this goes unnoticed by any one listening to the final recording.
   c. Limited overdubbing, editing and processing of the recorded performances as part of the production process regardless of whether this is noticeable to anyone listening to the final recording.
   d. Limited overdubbing, editing and processing of the recorded performances as part of the production process as long as this goes unnoticed by anyone listening to the final recording.
   e. No overdubbing, editing and processing as part of the production process.
   f. Other (please specify and comment in the box below).

6. In your opinion, which of the following are most important to Jazz music fans? (Tick more than one if appropriate).
   a. The recording is a faithful reproduction of the performances captured during the recording session.
   b. There were no edits or overdubs utilised in the production process.
   c. None of the above. The only thing that matters is the music sounds great, the production process is irrelevant.
   d. Other (please specify and comment in the box below).

7. If you were to take part as a performer in a Jazz recording session, which microphone techniques would you prefer?
   a. A single microphone placed close to my instrument.
   b. Multiple microphones placed close and/or in proximity to my instrument
   c. A single mono or stereo microphone placed in the room to capture the whole ensemble
   d. Other preference, no preference or unsure (please comment in the box below).

8. When recording with a Jazz ensemble which of the following applies to you?
   a. I prefer to record live as part of an ensemble.
   b. I prefer to overdub my parts by layering my parts on top of previously recorded parts.
   c. Other (please comment in the box below)

9. If a player or players made noticeable mistakes during a Jazz ensemble recording session, which of the following actions would you be happy to take?
10. Considering the recording and production of Jazz, in your opinion which is the most important mixing strategy?

a. To maintain the levels and dynamic balance of the music as it was performed.
b. To change/enhance the levels and dynamic balance of the music where this is lacking in the original performance.
c. Allowing the levels and dynamic balance of the music to be altered even if this leads to development of a new musical outcome which was not the planned intention of the performer(s).
d. Other (please comment in the box below).

e. Other (please comment in the box below).

11. In your opinion, which of the following are acceptable when producing a mix of a Jazz recording?

a. Rebalancing of individual instrument levels between the instruments of an ensemble.
b. Reorganising the positions of ensemble instruments in relation to each other so that they vary from the actual positions they held in relation to each other during the recording session. c. (E.g. Piano was on the left of the room/stage during recording, but is placed in the right stereo speaker in the final mixed recording).
d. Adding artificial reverb.
e. Introducing additional effects such as Chorus, Distortion or Flanging to the recording.
f. Rebalancing the level and dynamic articulation of an individual instrument.
g. Editing timing mistakes made by a performer.
h. Applying tonal enhancement to any part of the recording using equalisation techniques.
i. Editing and correcting recorded pitch mistakes.
j. Adding audio samples which were not part of the original performance.
k. None of these are acceptable. (if you wish to comment, please do so below).

12. When you or someone else produce a Jazz recording you have performed on, which of the following structural edits would be acceptable?

a. To creatively edit an instrument out of sections of the recording.
b. To creatively edit an instrument out of the entire recording.
c. To remove or reorder whole sections of the overall recording.
d. To edit in order to loop/repeat a section of the recording which was not repeated previously by the ensemble.
e. None.
f. Other (please comment in the box below).
Appendix: Abstracts - Parallel sessions The 12th Art of Record Production Conference Mono: Stereo: Multi (ARP2017) Royal College of Music, Stockholm 1-3 December 2017

Session 1 A

Christos Moralis: Bridging the gap between the studio production and the live performance in popular electronic music: The ‘Performable Recordings’ model

Abstract: As Nicolas Collins and Julio d’Escriván Rincón mention (2007), ‘It is perhaps a general human habit to view the technological and the organic as opposites. It is certainly the case that the phrase ‘live electronic music’ strikes many a music fan as oxymoronic’. Since electronic music is all about the consistency in its sonic attributes, performing it live suggests spontaneity, variance, and incongruence between the studio and the live produced musical descriptors. This paper will explain what the ‘Performable Recordings’ model is and how the gap between the studio production and live performance could be bridged. This production and performance model is based on the research on ‘liveness’ and on the computer made music expanding the research to include the rhythmic, dynamic, pitch, timbral implications that are characteristic to live performances. Furthermore, this model aims to balance the aesthetics and the consistency of the sonic attributes between the studio production and the live human performance. More specifically, is ‘a type of music production that enables the artist to perform live using, in real-time, the same mixing and post production process that applied during its creation’. According to Adam Parkinson (2006), ‘Liveness and mediatization can co-occur…Mediatization, may in fact, amplify perceptions of liveness. From this viewpoint, audiences call something ‘live’. Since, electronic music is all about the consistency in its own inherent traits and even though is all about the human that steps in and alternates the musical descriptors with his/her live performance, the ability to maintain consistency in the sonic attributes, while preserving also the human emotional expression, can bridge the gap between the studio production and the live human performance.'
Emil Kraugerud: The role of stereo centre placement in constructions of intimacy

Abstract: Since the normalisation of a “diagonal mix setting” (see Dockwray & Moore 2010), there has been an unwritten rule in production of popular music that lead vocals should be placed in the stereo centre of a mix. There are exceptions to this rule, but in contemporary mainstream pop it has become a standard, along with a tendency to push the vocals as far as possible to the fore of the sound box. This is effective in terms of drawing the listeners’ attention to the music, via an increased sense of intimacy with the singer. As part of a PhD project aiming to investigate intimate space in recorded popular music, and the ways in which such spaces contribute to the communication of musical identities, this paper will investigate the background for central placement in the construction of intimacy. In the context of the so-called “loudness war”, mainstream pop music is dependent on a sense of immediacy between singer and listener to be able to cut through the “noise filter”. Accordingly, the paper will discuss the central placement of lead vocals both as a result of the aesthetic preferences of the producer for constructing a sense of intimacy, and as a practicality for “reaching through” to the listener. This will be based in a combination music analyses and readings of literature addressing various aspects of virtual space (see, for example, Dockwray & Moore 2010; Doyle 2005; Moylan 2015; Zagorski-Thomas 2014).

Toby Seay: Towards signifying an engineering ethos through Heavy Metal and Country Music

Abstract: Much has been written about the production of Metal and Country Music. Since these two genres rarely intersect, one might assume that production methods for each are on different ends of the recording spectrum. While Metal production seeks “definition and intelligibility” (Mynett, 2012), Country music production has been described as “clean and crisp” (Porcello, 2002). A “vital parameter” of Metal is its “sonic weight” or “heaviness” (Mynett 2017). But is this concept of heaviness much different than a Nashville producer’s goal of sounding “big as a house” (Buckingham)? While the tools and techniques used by recording engineers in both genres are not unique (Williams 2015), the intended goals may be perceived as quite different. However, might these goals be more universal? This paper explores the notion that the production of Metal and Country Music (and by extension, most recorded popular music) is guided by a common engineering ethos. In doing so, the author will compare written literature, conduct interviews of engineers in both fields, and focus on the use of engineering techniques such as sample supplementation/replacement, distortion, and direct-to-reverberant ratios within a mix. Audio examples will be used to illustrate an engineering ethos through key elements of production.
Session 1 B

**Brendan Anthony:** The producer’s vision: A study into the multifaceted cognitive design of the popular music recording aesthetic

**Abstract:** Research into popular music record production and its associated creative practice has highlighted that a song’s production is often influenced by a multitude of stimuli and these can be musically, sonically and socio-culturally diverse. Technology’s influence on musical aesthetics is also at the forefront of scholarly investigations because the democratization of recording technology suggests that the musical spaces producers operate in have changed. Artistic direction however, is still the producer’s responsibility and the current landscape for record production is filled with a multitude of creative practice options that shape the recording aesthetic. These can include live or overdubbed performances and electronic programming versus acoustic instrumentation and when combined with technological choices these decisions ultimately frame the creative stages of pre-production, recording, and mixing.

So how does the producer ensure a production process that engages appropriate influences, and subsequently manifests a suitable musical result? This paper theorizes that the producer’s vision is the constant underpinning of the production rationale and therefore this subsequently designs the recording process and affects musical and sonic aesthetics. It is here that the producer uses multi-modal perception to target genre related outcomes of musicality and the sonic palate, and nurture the capturing of appropriate performances.

However the paper argues that this cognitive vision is an individualised trait that is inspired by a ‘field of knowledge’ from which producers innovate. This paper reports on a qualitative investigation into the producer’s vision via a survey of five producers whose experience range from national success in Australia to international acclaim. The paper demonstrates how the data analysis unpacks the discourse surrounding the producer’s vision and is supported by research from the fields of creativity, musicology and popular music production.

**Phil Harding & Paul Thompson:** Collective Creativity in Commercial Pop Music Production: A Service Model

**Abstract:** In his introduction to The Art of Record Production: An Introductory Reader for a New Academic Field (Frith & Zagorski-Thomas, 2012), Simon Frith proposed that producers in pop and dance music genres have a significantly different role to music producers in other music genres such as rock. A prominent difference is that pop music producers are often part of a production team that involves direct collaboration and participation with songwriters, programmers, musicians, artists, management and record company representatives. Pop music songwriting and production teams are therefore more frequently part of a larger creative collective (Hennion, 1990) in creating a musical product. The following paper describes the creative production workflow system at Pete Waterman Ltd. (PWL) Studios during the
1980s and investigates the way in which Phil Harding and Ian Curnow (P&E) worked with manager and entrepreneur, Tom Watkins in the 1990s. Drawing upon a series of interviews and data gathered during an extended ethnographic and auto ethnographic study, this paper presents the pop music ‘service’ model, which underlines collectivist rather than individualist thinking and illustrates how evaluation is present (and co-current) at the ideation stage in the generation of creative ideas (Sawyer, 2003) at various stages of the commercial pop songwriting and production process.

**Tuomas Auvinen: Differences and Similarities in the Creative Agency of Producers of Pop, Rock and Classical Music**

**Abstract:** In my presentation, I will explore differences and similarities in the creative agency of the producer in the production process of urban pop music produced in a home studio, rock music produced in a conventional studio facility and classical concert hall music produced in a concert hall setting. Starting from the premise of record production being a collaborative effort, I approach agency as the capacity to make and effect decisions within a structure or even to alter it to some extent, and creativity as contributing to the domain of existing works through exercising aesthetic decision-making. Based on these understandings of agency and creativity, I will examine how different cultures in different production settings and different studios conceived as cultural spaces affect the construction of the producer’s agency within creative communities in the production process. Furthermore, I will discuss how differences in understandings of the ontology of the music contribute to the level of creativity, i.e. the contribution to the domain of existing works, that a producer agent can possess. I base my presentation on extensive ethnographic fieldwork of three case studies on production processes, which took place in the course of 2015-2017. The presentation will summarize and discuss some of the central findings of my forthcoming PhD dissertation. This presentation is intended to be in the short presentation format.

**Session 1 C**

**Pete Gofton: From Shoreditch To Sainsburys; An Ethnography of the Vinyl Music Industry**

**Abstract:**

Over the past decade or so a growing market has emerged for vinyl records, with specialist record labels and industrial-commercial practices catering for it. I wish to undertake an ethnographic study of this industry, incorporating an examination of the format itself, the record labels that sell it, the marketplace they operate within, and their relationship to the consumers. Through use of physical product, situation within culture, society, and a utilisation of some of the approaches of ethnography, an interesting paradigm has
emerged, rife with contradictions: post urban, yet also grounded in locality and identity. Artisanal and oppositional, yet married to post-digital industry. Borne of the politics of both cultural and physical gentrification, yet an expression of individual politics and identity. By examining the links between the object, its shifting identity in culture, and the increasing corporatisation of that culture, we may come to a greater understanding not only of the music industry, but of the way in which, post digital, culture is increasingly used as an economic engine.

Marco Antonio Juan de Dios Cuartas: The incursion of stereo in Spanish popular music: the English influence in the definition of a local sound throughout the professional exchanges between Madrid, Barcelona and London.

Abstract: The figure of the engineer and the musical producer arises in Spain from the ‘imitation’ of the Anglo-Saxon model. The English professional experience of Alain Milhaud – a French-born producer settled in Madrid who must be considered the first musical producer associated with the rock genre in Spain – with the recording of the song Black is Black in 1964 by the Spanish rock band Los Bravos at the Decca studios in London with the engineer Bill Price as well as other professionals such as the audio engineer Pepe Loeches – who moved in 1969 to London to work at Pye Records studios where he developed his professional activity until 1975 –, helped to import into Spain some of the usual working methodologies in English recording studios. Loeches’ return to Spain implies, in some way, the importation of some of the Anglo-Saxon working methodologies into the production of our country, exerting an important influence in other professionals and recording studios of Madrid during the decade of the 70s: Kirios, Eurosonic and Musigram. On the other hand, the recording studios of the major companies were technologically nourished with those devices that had been updated in their studies of origin. To the EMI studio in Barcelona, it came material from Abbey Road, allowing to make stereo recordings with consoles and tape recorders previously used in productions of The Beatles. In spite of this, the introduction of the stereo in Spanish popular music productions does not seem to be delayed as compared to other markets like England or the United States. Audio Engineering magazine, in its issue from March 1961, noted in an article about Hispavox studios in Madrid its experiment with multiple microphones and the development of the first stereo recording techniques. The incursion of stereo production techniques into Spanish popular music and the influence of the professional exchanges that took place during the 60s and 70s among London, Madrid and Barcelona within an increasingly globalized musical industry, represent the main objective of this research.
**Tore Størvold: Sonic Cartography: Record Production and Geographic Space in Ólafur Arnalds’ Island Songs**

**Abstract:** *Island Songs* (2016) by the Icelandic composer and multi-instrumentalist Ólafur Arnalds, is an audiovisual work spread across platforms: a musical travelogue consisting of recordings, music videos, an interactive website, texts and photos. The creation of *Island Songs* involved Arnalds travelling to seven different locations around Iceland, one per week, where he would compose and record one song at each location. This process was continuously documented on social media, with dedicated #IslandSongs feeds on Twitter, Facebook and Instagram. The production, mediation and reception of *Island Songs* thus explores issues of locality, home and belonging through globalized media technologies. This paper mobilizes the media theory of Jody Berland to shed light on *Island Songs* as a cultural technology of space. Record production is viewed here as a technology involved in the shaping of geographic space: how it becomes inscribed, negotiated and narrated. The paper contextualizes *Island Songs* in contemporary Icelandic society, where the relationship of Iceland to the global has been a central tension following the financial crash of 2008. A distinctly spatial narrative of Iceland as nation state has been observable since its independence, with cultural tropes of distance, remoteness and isolation being key. The paper employs audiovisual analysis in a cultural reading of *Island Songs*, aiming to introduce some theoretical issues regarding record production and geographic space.

**Paul Novotny: Stereo to 5.1–Creating an Immersive fold-out**

**Abstract:** Look Ahead is a jazz piano and bass, duo recording of performance-music, tracked, mixed and mastered at 24 bit/96khz for stereo and 5.1 playback. Esoteric microphones and pre-amps contributed tonal diversity and contrary to standard practice, the stereo mix was folded-out to 5.1, rather than folded-down to stereo. It was pre-determined that a “sympathetic openness” in the playing and sound was a desirable aesthetic, thus the "performance oriented" physical setup was a blending of the traditional Oscar Peterson and modern Keith Jarrett piano/bass set ups. These choices set forth a coherent foundation toward an intimate, immersive and dynamic performance recording. The stereo sound-field begins at the phantom center position of bass and the 5,1 mix builds outward, maintaining a natural coherence between both versions. The upright bass was recorded with a carefully centered stereo ribbon microphone, a mono hyper-cardioid condenser and a “DI”— the piano utilized two outside mics (U87’s), providing a cohesive center image that is blended into an inside-placed “ultra-wide-stereo” Calrec-Soundfield mic, limited to approximately 90% of pan-width, reserving the outer L/R edges for reflections and reverb. Multiple reverbs were mixed and panned to avoid a dead-spot between the R-RS and the L-LS. Since there are no drums this “chamber” became a featured participant of the ensemble, providing unexpected and contrasting responses to percussive attacks. The conclusion asserts that a stereo sound stage built on traditional performance and recording values provide a connecting foundational coherence when folding-out. A ste-
Jo Lord: 3D Audio for Music

Abstract: The research project investigates and explores the development, practical application and aesthetic suitability of 3D mix technique for record production, the current developments and changes in consumer listening behaviour and the demand for an aesthetic, affordable and convenient 3D audio reproduction. The practical element focuses on re-composing stereo pieces for 3D over headphone consumption using perceptually-motivated production and an Auro 13.1 speaker array. The study’s first objective analyses the translation differences between the 13.1 speaker reproduction and the binaurally encoded 3D versions, negotiating the practical and aesthetic adaptations required for mixing 3D for headphones using a multichannel speaker array and binaural encoding. An important and fundamental objective posing questions such as; ‘How should we mix 3D music for current user listening trends?’ “How will variation of encoder, headphone and listener affect the perceived musicality and 3D translation?” Research as practice takes an ecological focus in the development of 3D production technique. The study’s second objective works alongside the first, investigating auditory perceptual phenomena, peripheral and kinetic staging, timbre, conceptual blending and sonic cartoons. Drawing upon questions such as; ‘How could we utilise 3D audio to benefit composition and music production?’ ‘How could we utilise our understanding of human perception to better 3D music production?’ and ‘How could we arrange a sound stage for 3D music?’ The second objective’s listening tests analyse the importance and effectiveness of the 3D aesthetic against original stereo mixes, examining localisation, re-composition, musical application and creative effect. The translation of array to headphones and the creative 3D aesthetic are subjectively assessed in both 13.1 and 3D binaural playback during a series of randomised listening tests using a consistent sample of expert and non-expert consumer volunteers. The AV presentation playfully examples interesting 13.1 audio playback demonstrating 3D production, 3D sound staging and sonic cartoons developed in research.

Session 2B

Elizabeth Varnado: Sonic Totems: The sampling aesthetic of Bon Iver’s ‘33 God’

Abstract: In September 2016, Bon Iver released their third full-length album, 22 A Million. Full of digital samples and gritty, manipulated textures, the album is brazen and bold if compared to Bon Iver past. Known for sweeping emotionality and Justin Vernon’s melancholic, free-association lyrics, Bon
Iver took a four year hiatus after releasing Bon Iver, Bon Iver in 2011. In that time frame, Justin Vernon worked with Kanye West, hip-hop’s self-proclaimed god, and released a few tracks with moody electronic artist, James Blake. With only these few and far-between performances as clues, rumours flew about whatever Vernon was sweating through to release next. In this paper, I will explore the various types of sampling used within “33 God,” the fourth track on 22, A Million. I will discuss how Justin Vernon’s use of sampling aligns with the tradition of hip-hop, specifically in his prioritization of timbral authenticity, and his tendency toward signifying, as described in the studies of hip-hop by Joseph Schloss (2004) and Mark Katz (2004), extended from theories by Henry Louis Gates, Jr. (1988). I will explore how Justin Vernon weaves digital samples (using an OP-1 Synthesizer and a device especially built for this project) into his own vocal presence on the track to create an audio-autobiography, where Vernon’s perspective, assumed authentic because of his singer-songwriter reputation and delivery, is combined with the perspective and vocal delivery of those he samples: Jim Ed Brown, Paolo Nutini, and Lonnie Holley. I will investigate Vernon’s use of these samples as “sonic totems” and symbols of “otherness”—other times, other places, other emotions—which are layered and collaged in order to convey his own memories and mental states, through the voice and timbre of others. I would like to present this paper as a long-form presentation of completed research. I would require AV equipment with laptop hook up to present diagrams, charts, and to play musical selections.

José Manuel Cubides-Gutierrez: The Portable Studio: The City as a Recording Studio

Abstract: The way we ‘consume’ music has changed over the last decades due to the development of new and more affordable technology. The producer and engineer have become consumers creating a whole new marketing target for audio technology developers. Additionally, the final recording is listened to on inexpensive equipment such as low-cost earphones and laptop speakers despite the expensive and delicate process the music has gone through. Also, the fact that nowadays everybody can record and distribute music easily from the comfort of their own homes has blurred the lines between professional and amateur music producers generating an oversaturated market of audio equipment affordable for a wide range of the population. Given that this argument has been examined a lot from the democratisation perspective, my doctoral research has been exploring the creative possibilities of the utilisation of inexpensive audio equipment and the ground in between raw found sound and highly processed samples where all their original character is removed and they sound indistinguishable from the instrument they are emulating. For instance: making percussive, harmonic and melodic sounds out of car doors, train announcements, and healthcare equipment sounds amongst others. The aim is to demonstrate the range of creative possibilities for music composers and producers using only a portable recorder and a DAW for post-production. This includes experiments such as creating binaural recordings using in-ear microphones and mixing them into multi-
track productions. The research has been based on ideas by Paul Théberge (1997), Brøvig-Hanssen and Danielsen (2016), Bull (2000), Miller (2008), Gandy (2014), Baker (2012) O’Rourke (2013), and Zagorski-Thomas (2017) amongst others. Finally, this presentation will finish with a brief reflection and analytical comparison between different methodologies employed for making popular music recordings out of inexpensive audio equipment.

Session 2C

**Kirsten Hermes: Enhancing creativity through research: developing a novel spectral clarity measurement tool for use the mix process**

**Abstract:** Mixing music is a complicated process and hence, automatic mixing or metering tools are beneficial. In a prior research study (Hermes et al., 2017), the spectral clarity of sounds — one particularly important parameter of music mixes — was investigated. Two predictors of spectral clarity are the harmonic centroid (a weighted centre mass of energy of a sound spectrum, Hermes et al., 2016) and spectral inconsistencies related to sharp peaks roughly in the middle of the frequency spectrum (Hermes et al., 2017). Naturally-occurring sounds can become more clear when low-Q equalisation is applied to boost the less-audible higher frequency regions (raising the harmonic centroid). However, if equalisation exaggerates or introduces timbrally unpleasant spectral inconsistencies, then these can mask or distract from other sonic components and lead to a clarity reduction. In order to work towards the development of automatic mix quality measurement tools, two simple computational tools were devised to detect changes in these parameters (Hermes et al., 2017). In the current study, the impact of these tools on the creative process is tested and starting points for further research are established. As an electronic artist, the author has been writing, producing, performing, mixing and mastering original songs for approximately ten years. Since the entire creative process is undertaken by just one person, external, objective feedback on the mix process can be useful. Therefore, the clarity metric is used as guidance for a vocal mix in a house music production. Feedback is also gathered from a group of additional audio professionals. It is established that the tools can be beneficial not only for automatic mixing; but also in the manual mix process. Improvements of the clarity model are discussed.

**Xuefeng Zhou, Qinhai Li & Xian Cai: Timbre Perception in Affordances and Aesthetics**

**Abstract:** This paper addresses timbre perception in three issues as the following: (1)Timbre perception is possible to musicians. According to Vuust et al. (2012), the characteristics of the style of music played by musicians influence their perceptual skills and the brain processing of sound features em-
bedded in music. And results from Garner et al. (2015) showed that training benefits are dissociable for the brain events that reflect distinct sensorimotor processing stages. Beside perceptual skills and training benefits, selective attention to a specific object in auditory memory does benefit human performance not by simply reducing memory load, but by actively engaging complementary neural resources to sharpen the precision of the task-relevant object in memory (Lim et al. 2015). These results deduce a possibility of timbre perception although experimental materials of the second and third research are not ecological music. (2) Using a sentence of an ecological music, this study employed the Audacity and five pianists’ samples recorded near the performer and in auditorium. Single sounds, a global state and timbre contour of ten samples were evaluated. (3) Then the standards of timbre were discussed according to views of different cultures, e.g. the west and east. Further study of perception experiments need to be carried out. We hope to interpret how are these socio-cultural and technological changes related to the sounds of recorded music.

Session 3 A

Brecht De Man, Kirk McNally & Joshua Reiss: Behind the Mix: Exploring the influence of a music producer’s background upon their creative practice

Abstract: This paper looks at the question of how a music producer’s background and local music culture affects the aesthetic and technical decisions made when realizing a music mix. Given the same multitrack recording, two engineers can deliver wildly different products, or indeed, very similar ones. Using a dataset of mixes created by engineers from significantly different geographical locations we investigate where approaches are shared and where they differ. This question has been dealt with on the macro scale by Zagorski-Thomas, where he questions whether UK vs USA rock music of the 70’s has a discernible sonic signature. Zooming in, the same question has been posed by Cummings, Horning and Seay, who examine the local culture of a city and a studio(s), respectively, in order to gauge the effect upon the musical output. In each of these cases the entire creative process is encapsulated in the analysis of the end musical work, from recording to mix, to release. This work has simultaneously identified the intriguing nature of the question, but also the difficulty in answering it. The study here limits the music production process to mixing. A shared set of pre-recorded multitrack recordings were mixed by a number of engineers with different backgrounds and from different countries. Analysis of both the final musical output and individual mix elements (stems) allows for further insight into how this common material is approached by these different music producers. Statistical analysis determines the extent to which differences in approach are likely to be influenced by the mixer’s background. The work extends previous work in this area to include objective comparison of audio features and processing parameters between
mixes of similar songs, including loudness balance and reverberant ratio. Furthermore, mix evaluation by listeners from the same groups of experts reveals to what extent preferences, focus, and perception differ as a function of personal background.

**Andrew Bourbon & Michail Exarchos: From Southern (T)Rap to Club Atmos: A Live Performance of Samples, Beats and Modular Synthesis Exploring the Inter-stylistic Evolution of ‘Synth-Hop’ in Surround**

*Abstract:* Although Hip-Hop is primarily considered a sample-based musical form, its historical journey has been frequently ‘contaminated’ by the use of synthesisers. From the appearance of analogue mono-synths in early Electro and G-funk, to the use of synthesised drum sounds (Roland TR-808) in Southern Bass music and Boom Bap, a plethora of rap subgenres have challenged their primary dependance on sampling in order to navigate the legal landscape and signify geographical or stylistic divergence. Contemporary Hip-Hop has expressed a new-found synthetic obsession as a result of these forces, continuing a process of stylistic morphing that began with synth-heavy Southern (US) subgenres such as Crunk and Trap. But under the apparent sampling and synthesis binaries lies a larger inter-stylistic mechanism, mirroring the interaction between Hip-Hop and other musical genres. On the one hand, Hip-Hop’s sampling philosophy is expressed in its inclusive modus operandi looking outside of itself—Funk, Rock, P-funk, Disco, Pop and Electronica—to redefine its instrumental bed. On the other hand, fewer are the cases when Hip-Hop has had the opposite effect on Electronic Dance Music (EDM). Yet with the subgenre of Trap, the common synthetic denominator between it and Electronica has energised a reciprocal and international dynamic, with Trap gradually morphing to instrumental electronic music, reaching European dancefloors, and crossing over from EDM’s underground back to the rap mainstream. At the same time, European dancefloors have been experimenting with multi-channel speaker formats (Ministry of Sound), while sampling drum machines, such as—Hip-Hop’s primary tool—the Akai MPC, have evolved to represent workflows favoured by both rap producers and EDM DJs combining live performance capabilities with sampling-synthesis integration (MPC X). The Live Performance leverages these new technologies in order to express the inter-stylistic opportunities afforded by the historic evolution of rap and electronic subgenres, their convergence, and the aesthetic implications of technological affordances identified in the creative tools deployed and dissemination formats explored.

**Michail Exarchos: Mono Samples | Stereo ‘Joints’: Exposing Hip-Hop’s Modus Operandi in Surround**

*Abstract:* Sample-based Hip-Hop production re-contextualises the sonic past, juxtaposes multiple sonic ‘pasts’ with the present, and proposes a multi-dimensional phonographic construct as a result. While the harmonic an
rhythmic qualities of the rap construct attempt to ‘glue’ its inherent multi-dimensionality and exponential juxtaposition, the sonic staging of contemporary Hip-Hop often exposes and exaggerates the effect. The phenomenon is the result of a complex web of contributing factors, ranging from developing aesthetic values, to the effect of sample legislation, to pragmatics borne out of technological necessity. Rap’s aesthetic and pragmatic mix preoccupation with the centrality of the beat, bass and voice—combined with the pursuit of optimum loudness via middle-and-side mastering techniques—position the sampled past in literal and metaphorical distance, exaggerating the illusion of ‘depth’, and directing the listener to the immediacy of the present. As contemporary music production moves from a post-modern duality to a meta-modern multidimensionality, does Hip-Hop exemplify this notion in its simultaneous collapse and celebration of temporal and spatial distances? How can these stereo production traits—audible in the work of contemporary rap producers such as Kanye West—be further explored in multi-channel formats? In the author’s own professional output, samples of Greek folk legends and reggae orchestras have been juxtaposed with electronics, beats and rap, leading to the first 5.1 rap mixes commissioned in the Greek market (and the first MTV Best Greek Act award in 2008). As part of a larger research project exploring the effect of vintage production techniques on contemporary Hip-Hop, the investigation deploys a reverse-engineering methodology, (re)constructing mono samples referential to past eras and positioning them in the hip-hop mix. Using (auto)ethnography and reflexivity, the applied exploration expands the staging effect from stereo to surround, examining the implications of sonic signatures derived from traditional formats (tape, vinyl) on the stylisation of a genre, and navigating future opportunities for innovative staging within a multi-channel context.

Session 3 B

Ragnhild Brovig-Hanssen: Dynamic Range Compression’s Influence on Perceived Timing

Abstract: To what extent does the signal processing operation Dynamic Range Compression (DRC) influence our auditory perception of temporal placement at the micro level in music? Compression is used on most popular music recordings and its function is usually explained as narrowing or compressing an audio signal’s dynamic range, often with the purpose of making the music sound louder, or making sonic features within the music more consistent in dynamic range. In EDM music, compression is also often used to reduce the amplitude of one sound source at the moment when another sound source reaches the threshold—an effect usually referred to as side-chaining. While compression is usually described in terms of dynamics, this paper examines the extent to which compression (master compression, individual compression and side-chaining) also influences our perception of the compressed signals’ temporal placements. This examination relies on Danielsen’s
hypothesis that perceived temporal location of a sound is affected by the sound’s sonic features, including temporal envelope (Danielsen 2015). A compressor is a complex amplitude modifier that influences the compressed sound signal’s envelope. As such, it is hypothesized that the use of compression manipulates temporal relationships at a micro level. This hypothesis will be tested through the use of two methodological approaches and draw on recent studies on the relation between sound and timing (Danielsen 2006, 2010, 2012, 2015; Goebel and Parncutt 2002; Hartmann 1995; Hove et al 2007; Lakatos 2000; Tekmann 2002; and Wright 2008). The first approach is interpretive musical analyses of selected EDM music, in which I will use a combination of auditory and graphical analyses, supported by calculations of relevant physical measurements using the MIR Toolbox for MATLAB. The second approach is qualitative semi-structured interviews with EDM producers and engineers, in which they reflect upon the compressor’s influence on timing at the micro level.

**Hans T. Zeiner-Henriksen: Sidechain compression and pulse perception in music production**

**Abstract:** The use of DAWs in music production has during the last decade(s) given the musicians/producers an incredible control over the many parameters of sound. Various plug-in effects and synthesizer parameters can be programmed to make detailed modulations that enrich the production. Based on theories within embodied cognition and my earlier work on bass drum sounds with descending pitch movements on the downbeat (Zeiner-Henriksen 2010) I now turn to off-beat sounds that in a similar manner may establish reference points (or movement affordances) that are vital for how we may perceive the pulse. Among others, Troye Sivan’s track «Fools» from 2015 will be used to exemplify the phenomenon.

**Anne Danielsen: The perception, aesthetics and cultural values of ‘glitched’ grooves**

**Abstract:** Rhythmic feels produced by inserting glitches into a played groove post production or warping samples in different ways are now widespread in a wide range of popular music styles, from mainstream pop to experimental hip-hop. In this paper, I present an analysis of the song ‘1000 Deaths’ from the album Black Messiah (2014) by D’Angelo. I will, first, map the micro-rhythmic relationship of the groove, and, secondly, relate its microrhythmic design to examples of similar past and present practices. The aim is to explore the wide array of aesthetic and cultural meanings that these production practices have taken on, bridging the gap between technical descriptions of sound/sound production and interpretations of its cultural and contextual meanings.
Session 3 C

Joe Bennett: “How someone controlled you”: The Digital Audio Workstation and the Internet as influences upon songwriting creativity

Abstract: This paper investigates the interrelationship between studio-based songwriting processes and artistic product, focusing on two digital tools that became available to songwriters towards the end of the 20th century: the Digital Audio Workstation (DAW), and broadband Internet connectivity. Building on recent research into computer-enabled music creativity (Marrington, 2011, 2017; Mooney, 2010), the paper asks whether DAW interfaces, combined with various online tools, has provided ‘cognitive extensions’ (Magnusson, 2009) for songwriters. If these digital tools have had a material influence on songwriters’ creative activities and workflows, how might this be evidenced in the musical characteristics of contemporary popular songs? Groups of successful hits are compared between the analogue and digital eras, with the aim of identifying a correlation between musical characteristics (chord loops, phrase lengths etc) and creative processes and workflows. Several professional songwriter case studies are provided, including a ‘digital immigrant’ (Prensky 2001) who began to write songs before digital tools were available, and a ‘digital native’ (ibid) has always used DAWs and an Internet connection in his songwriting. The participants were asked to describe their creative processes in detail, and to reflect on how these tools may have influenced their decision-making and artistic direction. From these and from previous studies, the author attempts to identify behaviours and affordances engendered by digitally-enabled songwriters, and to make inferences regarding these tools’ influences on song characteristics.

Daniel Pratt, Shane Hoose & Wellington Gordon: Transnational Group Flow: Writing and Recording Music in Three Different Locations

Abstract: Writing and recording music in multiple geographical locations with a decentralised group structure presents a series of challenges. These include technical integration, creative flow, organisational cohesion that allows for the unexpected, and maintaining a unity of musical vision. This paper frames the difficulties of creative collaboration across geographical distance as organisational challenges to be overcome. Using the group flow theoretical models of Sawyer, and the improvisational sensemaking focus of Karl Weick’s work, we investigated the creative challenges of distance and audio recording. Working with three music production academics in songwriting and recording sessions, we used Dropbox and Reaper as open platforms for an exchange of ideas. Each practitioner brought their own tacit knowledge to the project. The expanded use of recording spaces in three different geographical communities of practice presented opportunities for musical experimentation and organisational interpretation. These creative challenges do not exist in the traditional single studio environment, and as such we investigated
whether the challenges of distance and creative flow need reframing as creative organisational opportunities. We implemented and evaluated organisational theory designed to address the challenges inherent in this recording set up.

**David Myhr: What goes on in my mind? Observing melodies in the making**

**Abstract:** This paper posits a set of epistemological methods to enable researchers to investigate melody-writing in solo songwriters. Using ethnographic and autoethnographic methods combined with annotated video, I provide examples, from my ongoing research project, of real-time songwriting activities and address the opportunities and challenges of introspection as an investigative method. Sloboda (1985) proposed four possible methods of inquiry by which researchers might investigate composers' creative processes, one of which is “live” observation of composers at work. Collins (2007) uses a verbal protocol methodology that requires the composer to narrate the evolving composition in real time, which not only interrupts the creative process itself, but risks falling victim to the observation effect. Bennett (2014) suggests that observing co-writing processes has an immanent advantage over solo writing through the fact that the necessity to communicate ideas makes it an evidence-generating activity. The methods proposed combine digital video documentation, immediate post-hoc introspection and qualitative data analysis. The songwriting session is recorded in full, followed by immediate post-session review and captioning of identified ‘key moments’. This enables the composer to reflect immediately upon the session, adding extra insight such as explaining moments of silence/reflection that would be otherwise lost in the observation.

**Session 4 A**

**Alex Case: Oops, Do It Again - Gated Reverb From the 80s to Today**

**Abstract:** Among the more absurd sonic concoctions to come out of the recording studio, gated reverb offers a unique aesthetic possible only through loudspeaker-mediated sound. Born in the 80s, it relied upon creative, even counterintuitive application of some of the newest signal processing technologies of the time. The genesis of gated reverb was part discovery, and part invention. Its further development was motivated by rebellion, and confusion. Peter Gabriel did it first, with “Intruder” (1980). Phil Collins made it famous, with “In the Air Tonight” (1980). But David Bowie likely inspired it all with tracks like “Sound and Vision” (1977). This paper tours the development of gated reverb, with audio illustrations showing when, how, and why. What began as a radical reshaping of timbre has evolved into a more subtle form. Gated reverb remains relevant in contemporary music production, not just for
80s pastiche, but as a tool for overcoming masking through the strategic leveraging of its unique psychoacoustic properties.

**Tor Halmrast:** *SAM PHILLIPS’ SLAP BACK ECHO; LUCKILY IN MONO*

Abstract: “Slap back echo” was created by Sam Phillips for Elvis Presley’s early Memphis recordings. Using cepstrum and autocorrelation, we find that the tape delay used in Sun Studios was 134-137 ms, which is so long that the echo is perceived as a single, distinct echo in the time domain, and not the comb filter coloration of timbre in the frequency domain defined as Box-Klangfarbe. Such coloration would be perceived if a distinct, separate, reflection gave a comb filter with a distance between the teeth (CBTB: Comb-Between-Teeth-Bandwidth) comparable to the critical bandwidth along the basilar membrane in the cochlea. When Elvis changed to RCA Victor’s studio in Nashville, “RCA was anxious to recreate the “slapback” echo…To add them to Elvis’ vocals Chet [Atkins] and engineer Bob Farris created a pseudo "echo chamber" by setting up a speaker at one end of a long hallway and a microphone at the other end and recording the echo live”. Analysis of these recordings gives that the echo is somewhat shorter (114 ms and 82 ms), and much more diffuse, so “slap echo” was not actually recreated. The main findings is that even though the delay time of the Sun Studio “slap tape echo” is long, the echo is still perceived as rather “close”, because the echo is in mono. Panned in stereo, the feeling of being inside a small room would disappear. In addition, we analysed also a shorter delay, as for a possible reflection from the floor of the studio back to the singer’s microphone. These results are more unclear, but we found that such shorter delay would have given Box-Klangfarbe, but if this actually was a floor reflection, the measured deviation of the delay time must mean that the singer moved his head during the recordings (a highly reasonable assumption for Elvis!)

**Kai Arne Hansen:** *Interpreting Sound Recordings in a Time of Media Convergence: Aesthetics, Technologies, and the Migratory Behavior of Audiences*

Abstract: While recent technological developments have led to a range of new possibilities for the recording, production, and distribution of sound recordings, equally significant changes have ensued with regard to audiences’ usages and experiences of music. These changes concern not only how we access and listen to sound recordings, but also how we make sense of them. In light of what Henry Jenkins (2006) has described as the migratory behavior of media audiences, this paper considers the multi-modality of our present-day music experiences. By attending to the primacy of the artist persona in a contemporary pop music context, I call attention to how sound recordings are interpreted vis-à-vis other pop commodities and discourses surrounding the artist. I suggest that, as the representational strategies that promote and aestheticize the artist persona across multiple platforms become increas-
ingly pervasive and sophisticated, listeners become accustomed to enriching their musical experiences by seeking out additional content and information through various media. By merging recent theories of intermediality and transmediality with a critical musicological approach to interpretation, I attempt to demonstrate how symbols and signs dispersed across multiple media platforms are aggregated in the experiences of listeners and fans. To this end, I focus on the recent output of one commercial pop artist to take up how recorded sound operates alongside other media content to imbue our musical experiences with various meanings.

Session 4 B

Steven Gamble: From recorded sound to musical soundworld in popular music listening

Abstract: Recorded popular music affords virtual, spatial, narrative environments for its listeners. Much scholarly inquiry has focused upon either technological decision-making in music production or the creation of meaning in music perception. These studies give much insight into technology and expression on one hand, and aesthetics and interpretation on the other. These dichotomies are at times necessary, but may conceal as much as they reveal. At what point does ‘hard right’ on a pan pot become ‘creeping up on my right’ in a listener’s headphones? An emerging body of work instead draws upon an ecological approach to perception and embodied cognition in order to directly investigate the relationship between the two. Perhaps the best known example is Moore’s (2001) soundbox: what leaves recordists’ hands as the mix/master becomes the soundbox in the listener’s imagination, and from there we can begin to investigate the meaning any given track affords for that listener. Another useful concept is Zagorski-Thomas’s (2010) development of staging in recorded music. Without rigorous empirical evidence, however, a great deal of care must be taken in claims about listener perception. Using a theory of affordance, I discuss the kind of cautious claims we can make about the potential for meaning to arise in the listening process, referring to a single progressive metal track, Karnivool’s (2009) ‘Goliath’. This theory of affordance, based in Gibson’s (1977) ecological theory and substantiated by Lakoff and Johnson’s (1980) experientialism, may be a helpful way to investigate the emergence of meaning between sonic data and listener experience. To do so, it is important to clarify this position from its present use in musicological work: it does not merely mediate between the objective and subjective, but fundamentally questions this distinction to begin with.
Eirik Askeroi: Sonic Markers in Popular Music

Abstract: Throughout recording history, the use of different technologies and instruments has left lasting impressions on recordings, suggesting an almost intrinsic relationship between sound and period. With an increasing momentum from the early 2000s and onwards, producers, engineers and mixers, especially within the realm of popular music, have taken compositional advantage of this relationship, constructing what I label sonic markers: Musical codes that have been historically grounded through a specific context, and that, through their appropriation, serve a range of narrative purposes in recorded music. In the proposed paper I will explore this concept as a careful navigation between the two following positions: 1. Sonic markers as narrative strategies—how sonic markers can be used as a means of constructing musical identities in pop production. 2. Sonic markers as constructions—how sonic markers themselves are constructed through different narrative strategies. Presented against a backdrop of various musical examples, I will argue that sonic markers constitutes an inroad for interpreting the expressive dimensions of musical sound in recorded music for practitioners as well as academics. Exploring the potential effects of musical codes in recorded music requires close readings of musical texts against a contextual backdrop of journalistic and scholarly writings (articles, interviews, blogs, books), biographies, documentaries, semi-biographical movies (biopics), music videos and social media. The overall aim of this paper, then, is to demonstrate how sonic markers operate within different discursive formations and supply “sonic evidence” for identity formations in popular music.

Toivo Burlin: The making of virtual space in the King Crimson Box sets Larks tongues in as-pic: Complete recordings, Starless, Road to Red, On (and off) the Road and Thrak Box: Live and Studio Recordings 1994–1997

Abstract: When the founding fathers of progressive rock, British (-American) group King Crimson, produced their groundbreaking LP In the Court of the Crimson king in 1969, they established an eclectic yet recognizable style, with a strong musical and production virtuosity. From 2010 and onwards all of King Crimson's back catalogue was systematically re-released, some of their albums was also re-packaged in deluxe box sets: In the court of the crimson king, Larks tongues in aspic: Complete recordings, Starless, The Road to Red, On (and off) the Road and Thrak Box: Live and Studio Recordings 1994–1997. They all covering both studio recordings and live recordings that was produced in and around the period of the original recording sessions,
in a various set of mixes, alternate takes and sound codec’s. In this presentation the focus is on sketching out an analysis of the mixing strategies on the surround mixes in DTS-HD Master Audio and LPCM 5.1 Surround of the studio recordings and live recordings, as it can be heard, using models for analysis of recordings, the Representation Model and the Room Model.

Paul Theberge: ReCon: Recording Consoles, Reconsidered

Abstract: In the rock press and in recent films and television programs, a number of accounts of audio production have marked the recording console as an object of popular mythology, establishing a nostalgic constellation of associations between analog mixing consoles, recording studios, artists, hit records and, more generally, the history and aesthetics of popular music. Ironically, this fetishizing of analog recording consoles follows a lengthy period in which digital consoles have increasingly gained market share over their analog counterparts. The recording console is a technology whose primary function is to simultaneously enable and reduce multiplicity: it combines, modifies and shapes multiple inputs and configures them for a small number of discrete outputs (and consumer formats): mono, stereo and multi. The ways in which consoles have been designed to accomplish these tasks, however, is variable. This paper will survey a select number of historical shifts in console design in order to understand the specific technical affordances offered by them and their aesthetic significance. Of initial interest are the changes in console design that coincide with the historical development of multitrack recording practices. The console was, arguably, a singular device within the multitrack studio – a device whose affordances enabled multitracking as a technical process, mediated communications between musicians and producers, and supported the sound engineer’s claim to artistry. With the increasing complexity of contemporary studio recording, live touring, and cinema post-production practices, however, analog consoles have given way to multilayered digital control surfaces, where only a fraction of the mixing and processing functions are available to the sound mixer at any given time, resulting in new levels of abstraction and the affordances of full automation, digital storage and instant recall. These changes in design require a re-evaluation of the console: as a form of remediated technology, a set of shifting aesthetic practices, a network of human/ non-human interactions, and as emblematic of cultural values in a digital age.
Session 4 C

Kirk McNally, Toby Seay & Paul Thompson: Another Take: Teaching Music Production Using Multitrack Recordings

Abstract: Historically, the apprenticeship model of training in the recording studio allowed student engineers to learn from the masters of their craft. However, recent studies have shown that the recording studio sector has suffered a significant decline within the broader musical economies (Leyshon, 2009), which has resulted in the fragmentation of the knowledge capital that was traditionally found in larger recording facilities. University music production programmes have assumed the role of instruction, changing the nature of knowledge transfer from entirely informal and in-situ, to more formal experiential and theoretical (Thompson & McIntyre, 2013). Educational institutions offering music production programmes are therefore challenged to provide appropriate resources to support learning that would have otherwise been learnt on-the-job. However, currently available literature and AV resources often overly focus on technical knowhow and typically fail to connect technical decisions to the potential aesthetic consequences upon the musical output (Askerøi, Viervoll, 2016). In other words, there is a “virtual absence of pedagogical resources” (Zagorski-Thomas, 2016) that help to explore and connect technical, aesthetic and musical relationships. This paper explores the ways in which three institutions, Drexel University in Philadelphia, USA the University of Victoria in British Columbia, Canada and Leeds Beckett University in Leeds, UK are working to address some of these questions. We argue that multitrack materials afford a rich historical, practical and aesthetic resource for use in music production programs that help to connect technical decisions to aesthetic consequences within a musical context. Importantly, we illustrate the ways in which music production education can tap into the well of historical knowledge by using multitrack materials and, how moving the master into the classroom, can provide a greater level of access to their knowledge and ways of working.

Sören Johansson & Nyssim Lefford: The student’s view of the producer’s role: Analytical methods, interpreting content and practical project management

Abstract: In the sound and music production Bachelors program at Dalarna University, we have conducted a case study to assess how students understand the producer’s multifaceted role (following on,
for example, Burgess, 2013) and if and how they utilise analytical methods in their producing practice. Our curriculum emphasizes critical thinking, preparation for working life, and communication and collaboration. These aims are especially apparent in the producing coursework—because, practically speaking, producing involves analyzing what is heard and communicating about interpretations. Regardless, students are reluctant to embrace analysis as a practical tool. To address this, we offer a (pop) vocal production workshop that integrates scientific research, theoretical musicology and practical concepts. It delivers techniques for production analysis and a framework for connecting vocal affect to the vocal persona revealed as the lyrics are performed. Perception research provides concepts about affect. Other techniques are drawn from production and musicological literature, from practitioners and theorists both. (Wadhams, 2001; Cone, 1974; Tagg, 2013; Hennion, 1983) We investigated if students use these ideas in later work to plan, produce or communicate with session participants. Three students participated while working on their senior-year production projects. We started with semi-structured interviews to ascertain how they understood the producer’s responsibilities and their knowledge of production techniques. Later, they responded to a questionnaire about their specific projects. Results show they understood the producer’s multifaceted role and the need for some analysis of the material, in preparation and production. They struggled with concrete examples of methods. They unanimously expressed hesitance to discuss lyric interpretation or interfere with vocal performances, while simultaneously acknowledging the importance of lyrics in pop production. They demonstrated management skills, but their most analytical reflecting pertained to release formats and distribution channels. This study has raised new questions about what the curriculum could or should address.

Daniel Pratt, Brecht De Man & Joshua D. Reiss: Developing mix evaluation skills in higher education: A technology-aided approach to self-directed learning

Abstract: Mixing a multitrack recording is a complex and critical task involving many creative decisions, where any two engineers can produce wildly different results from the same source material. To understand more about these phenomena, De Man developed an online tool for evaluating multiple mixes of the same song. This tool has been used to better understand different mix approaches taken by sound engineers from a range of geographical locations and education levels, and mix evaluation has become a useful pedagogical tool for educators interested in technology-aided course design. Students respond to
the comparative nature of the mix evaluation tool and understand their mixing practice on a deeper level because of this comparative analysis. Our thesis is that the evaluation tool acts as a passive learning device to develop students critical listening and mixing skills. A group of third year students from QUT in Australia participated in a series of mixing workshops using the mix evaluation tool as a pedagogical device. We educated the students in critical listening and mixing techniques and challenged them to produce a mix in one week using a restricted set of tools. During the mix sessions, staff observed the students’ engagement with the evaluation tool as a method of developing their mixing practice. The students were then given the option to revise their mix before submitting it for summative assessment. We observed whether students are passively motivated to improve their mixes after participating in the evaluation. Finally, we interviewed a selection of students to better evaluate their impressions of the mix evaluation tool. Through a triangulation of observation, data collection from the evaluation tool, and interviews with staff and students we investigate the mix evaluation tool as learning device beyond its primary function as a data collection device.

Session 5 A

Magda Mayas: Piano mapping/ Memory piece: a work for amplified piano and multi channel playback

Abstract: I am doing research as a PhD student at the University of Gothenburg and I am proposing a performance presentation in a multi-speaker set up. I am exploring and experimenting with a grand piano, objects, microphones, speakers and space in the realm of improvised music. The relation between these elements and their relation to me is constantly shifting, and with it my defined role as a performer or listener becomes blurry. Finding a more engaged relationship with space, sound and the listener and transmitting my listening experience to others is the main concern of my research practice. I amplify the piano in a surround speaker set up with the piano in the middle of the space and the audience around it. Through spatial mapping I am expanding the piano to the size of the entire room, creating an immersive feeling for the audience of being inside the piano. I am controlling the routing and distribution of the microphones through the speakers through a simple max patch, changing between different presets during the performance, emulating the piano’s geography in the space and constantly shifting it. “Memory piece” is a documentation and recording of processes of change while playing inside this “piano map”
of a surround set up. Segments of these recordings are created into a separate piece, which serves as a sparse installation element in a new surround set up that I improvise with in a live performance. Overlapping multiple realities of environments and piano maps, I am creating a space within a space and in doing so, turning space into an instrument as well, adding manifold sonic and psychological layers. I encourage the audience into different perceptual modes by mediating their listening to the blurring of different spaces and to the disembodied sounds.

Aaron Liu-Rosenbaum: Immersion, Wanderlust and the Ambient Experience: A Narrated Multi-Channel Soundscape Performance

Abstract: As sound technologies proliferate and metamorphose, the "art of record production" is recast into interesting new forms with new aesthetic possibilities. One such form is soundscape composition, whose goals can range from sonically preserving real places, to creating artificial ones. The success of a soundscape composition depends, in part, on its ability to transport the listener to a new soundworld, and multi-channel audio recording and diffusion play a key role, as they offer a means of auditory immersion and control heretofore unavailable. This quest for immersion also transforms the spaces where music is consumed, such as in the Philips Pavilion at Expo ’58 in Brussels, whose hyperbolic architecture served to emphasize the spatial effects of the multi-channel music heard through its several hundred integrated speakers—or more recently, the multi-channel speaker array installations making their way into concert halls, cinemas, and even university research centres. For this presentation, I will briefly discuss some technical and aesthetic considerations in the production of immersive, multi-channel soundscapes, and will follow it with a performance of one of my soundscape compositions. The piece will be composed to exploit the sound system in the venue.

Session 5 B

Larry Whelan: The Roland TB-303 and timbre: Klangfarbenmelodie for the rave generation

Abstract: The Roland TB-303 Bass Line synthesizer is one of the most important devices in the history of electronic music and it’s widely acknowledged as an example of technology that’s significantly influenced the development of musical genres in which it was used.

369
Most accounts of the TB-303 to date have tended to focus on its technology and history, from its initial release by Roland intended as an accompaniment device through to its adoption, creative repurposing and widespread use by electronic music producers from the mid-1980s onwards. The purpose of this paper is to focus on the musical qualities of the TR-303 and it will be shown that its importance stemmed from offering possibilities of timbral manipulation in a field where timbre was becoming predominant over pitch. I will examine the TB-303 in relation to Arnold Schoenberg’s concept of Klangfarbenmelodie, who envisaged: “if it is possible to create patterns out of tone colours that are differentiated according to pitch, patterns we call ‘melodies’ … then it must also be possible to make such progressions … out of that which we call simply ‘tone colour’, progressions whose relations with one another work with a kind of logic entirely equivalent to that logic which satisfies us in the melody of pitches.”[1] The TB-303 was ideally suited to explore progressions of both pitch and timbre, and to explore transitions between the two - and this was the typical pattern of use in acid house and other forms of electronica. The musical context was of course quite different from what Schoenberg had envisaged. And in an examination of the psycho-acoustics of timbral manipulation, it will be shown that there cannot be a close equivalence between progressions of pitch and timbre. Timbre is “an informationally unconscious phenomenon, hazy in definition, difficult to articulate”[2] - and this was exploited by artists who used the 303 for extreme effects which musician Marshall Jefferson complained was less about creating musical moods than “disrupting thought patterns”.[3]

Andrew Bourbon: Hit Hard Or Go Home- An Exploration of Distortion On The perceived Impact of Sound in The Context Of A Mix

Abstract: Distortion is a powerful tool used by mix engineers to enhance the presentation of audio through enhancement of an existing sound. Distortion can take on many forms, with common distortion tools either emulating an existing hardware tool or providing new interfaces for the enhancement of audio, adding new material to manipulate the energy delivery of a chosen sound. This enhancement has the potential to not only draw attention to the listener of particular components of a sound, but also changes the fundamental delivery energy of a sound. Energy manipulation also takes place through other processing techniques including compression, which in the manipulation of the envelope of morphology of the sound may also contribute distortion subject to the wavelength of the sound being compressed, the nature of the gain reduction cell or the nature of the gain makeup amplifier. Dynamic processing will not be considered as part of this
paper, instead focusing on the impact and application of distortion on the delivery of sound. In his work Zagorski-Thomas (2014) explores concepts of sonic cartoons, with his continuing work exploring the relationships between actions and energy on the perceived delivery of a sound. A range of experiments, measurements and listening examples will be presented, exploring the potential for manipulating the perceived action on a sound using distortion as the primary process. Examples from drum kit processing through to bowed stings and vocals will be processed, with a focus on exploring specific distortion footprints and application techniques. The impact of the application of distortion directly onto the sound, as a parallel to a single sound in mono and stereo, and upon multiple sounds in parallel will be measured and presented with reference to the manipulation of the perceived action of the sound.

Session 5 C

Johnny Wingstedt & Thomas Florén: Music Producers, Knowledge and Multimodal Communication in the Recording Studio

Abstract: This short paper presents and discusses an ongoing pilot project studying communication and collaboration between producers and singers in the recording studio. From a knowledge perspective, it is assumed that 'the formal, the informal and the tacit are embedded in and inseparable from the ongoing interaction of the participants' (Heath, Luff & Knoblauch, 2004). It is therefore necessary to, in addition to the verbal/vocal, also examine multimodal and embodied means of social interaction – such as gesture, facial expression and gaze. In this, the specific conditions established by the mediated setting of studio communication via 'talkback' systems and sound-proofed glass windows are considered. Furthermore, the communicative use of artifacts, tools and technology needs to be studied, in order to get a deeper understanding of how creative and collaborative work is organized in the studio. This includes how the musical sound (live or recorded) and related technologies are used for interactional purposes besides mere artistic expression. The study of the activities in the studio is, on a micro level, being performed using multi-camera documentation and multimodal interaction analysis. On a macro level, in-depth interviews and stimulated recall sessions, focusing on the role of the producer, are used for getting the participants’ perspective on how they view their own role regarding e.g. situational, institutional, strategic and knowledge-related matters. Preliminary results touch on issues concerning how power relations, agency, status and (often tacit)
knowledge are negotiated. The joint workflow of collaborative and creative processes, as well as communicative resources, choices, and strategies are illuminated. The continued plan for this project is to include a larger variety of proficiency and professional levels, and musical genres. Given the increase in studies on contemporary media societies, such as consumer habits, media cultures and technologies, more research is correspondingly called for on issues concerning collaborative media production processes.

**Phillip McIntyre: Adopting a Multimodal Research Approach to Mastering for Record Production: An Examination of the Systems Model of Creativity through Practitioner-Based Enquiry (PBE)**

**Abstract:** Bob Katz has asserted that mastering “is the last creative step in the audio production process, the bridge between mixing and replication” (Katz 2002, p. 11). With this declaration as a guide this paper presents the third in a series of papers on mastering for record production. The first two studies took a traditional research approach by undertaking a qualitative ethnographic study on mastering (McIntyre 2008) followed by a quantitatively based experiment on mastering (Paton & McIntyre 2009). This specific paper adopts a less traditional and more innovative methodological approach, that of practice based enquiry (PBE) (Murray & Lawrence 2000). In setting out this research this series of papers adopts a multi-modal approach to academic investigation which eschews the mono-cultural research practice of adopting a singular research position. In doing so the contribution of this paper to the series is to reveal the creative and technical practices of a record producer seeking to further his own mastering skills and at the same time documenting and analysing that creative process. This PBE approach to mastering for record production moves past traditional approaches to instigate innovative creative practice and allows for a deeper understanding of this important technical process.

**Session 6 A**

**Zack Moir: IDDM (Multi-Channel Electro-Acoustic Performance For Live Saxophone and Tape)**

**Abstract:** IDDM is a composition which explores the condition of Type-1 diabetes and the effect it has on the daily life of those who suffer from it. Drawing on a number of streams of health data from the composer's personal diabetes management, pitch and temporal aspects of the music map to and sonically represent the near constant fluctuation of a number of aspects of
this condition. Sounds sampled from diabetes paraphernalia (blood glucose meter, insulin pump, needles, etc.) are used to create a meaningful and relevant sonic environment in which the live improviser performs, responding to and trying navigate the constantly shifting harmonic and timbral landscape created by the interaction of the sonified health data. The multi-channel presentation of this work allows for the improviser and the audience to be entirely immersed in the sonic environment, thus fully representing the experience of dealing with the barrage of rapidly changing, and often confusing health data. A blog post pertaining to the composition of this work can be found here: http://thinkingaboutmusic.com/?p=1998 A video of the premier performance (performed in quadraphonic sound, but stereo in the video) can be found here: https://www.youtube.com/watch?v=VvBQGGvZwkQ

Mark Durham: Multi Channel Sound Design: Instruments for 360 Degree Composition

Abstract: The continuing development and uptake of multi-channel audio is creating new potential for sound creators, with many current delivery formats supporting sound authored in three dimensions, for example Ambisonics in VR/AR, Auro 3D and Dolby Atmos for cinema and the home. There is also a growing interest in music composed specifically for highly spatial audio systems such as 4D Sound System, Envelop and Dolby Atmos for nightclubs. Often the focus of this technology is delivery and presentation - rather than content creation. This is reflected in the tools and process of authoring spatial audio, where typically the most common approach is to upmix: create in mono or stereo, then spatialise the content further down the line at the mixing stage. This research looks to address multi-channel content creation at its genesis, by exploring how digital instrument architectures can be designed for spatialisation, rather than being adapted to it. As part of the research a prototype device has been developed that the user can perform within a three-dimensional sound space. This terminology deliberately questions the role of performance in traditionally non-instrumental context such as mixing and spatialisation. Our approach looks towards new technology to bridge the gap between programming and performing with digital instruments in three-dimensional space, utilising motion control, gestural mapping and machine learning to interface instrument and performer with a spatial audio system. As part of a larger practice based research project, this research aims to explore how software and hardware tools shape the processes and sound palette available to sound designers, and to what extent customised software can contribute to a unique sonic signature.
**Matt Barnard & James Bagshaw: Frames to Spheres: The Abstract Spatial Paradigm of Ambisonics in a Recording and Production Context**

**Abstract:** Recorded popular music established its presentational conventions in the stereo age. This stereo-centricity of music production and reproduction has made transitions into new spatial paradigms largely problematic and often under-utilised. The paradigm of the ‘frame’ that stereo affords has become defacto, mirroring stage-centric presentation modes of music performance and informing most stages and aesthetics of production. Alongside practical themes, this paper explores a three-dimensional approach to music presentation, using the vehicle of ambisonics. The spherical domain appears, in line with our natural mode of audition, and the principle of ‘channel = speaker’ dissolves along with much developed spatial practice. How does this sphere function and how is it utilised? Alongside overarching ideas such as these, the speaker-agnostic characteristic of ambisonics introduces the challenge of abstract spatial definitives of works: the b-format domain of ambisonics is only realised during a decode, where the sphere is translated to the real world. As a case-study, the recording of drums is explored. The paper will address some fundamental issues of recording technique as mediated and influenced by the prospect of being played back on a multichannel set-up. With the two-channel configuration abandoned (and therefore the correlation to left and right ears), how could the drum kit be reimagined in this 3D space? How could its elements be recorded to reflect such a spherical spatialisation? And what changes to the physical configuration of the drum kit could be made to enhance such effects? And, crucially, how might such a new drum kit layout lead to new ways of playing, new grooves, new relationships between those familiar sounds? The paper will report on practical experiments and will present finished recordings.

**Session 6 B**

**Michail Exarchos & Glenn Skinner: Bass | The Wider Frontier: Low-end Stereo Placement for Headphone Listening**

**Abstract:** The placement of low end spectra in popular mixes has ranged from dogmatic centrality to extreme staging, driven by aesthetic, pragmatic and technical concerns, such as the directionality in perception, and the development of a range of playback and dissemina-
tion formats. From Phil Spector’s mono obsession, to limitations enforced by radio, TV and club systems, low end content has occupied the mix’s ‘centre’ for a vast majority of phonographic releases. On the other hand, jazz and pop records have experimented with extreme staging of ensembles, portraying real or hyper-real placements regardless of the resulting low-end speaker imbalance. However, current developments in production workflows (DAWs’ preoccupation with stereo channels), the stylistic evolution of contemporary genres, and increasingly personalised modes of consumption (personal computers, smart phones and headphones) have resulted in alternative approaches to low-end creation and production, sculpting ‘bass width’ in innovative, if less polarised, ways. But how do producers in contemporary pop, hip-hop and electronic genres expand and ‘animate’ low-end content to entertain their audience and take advantage of current listening modes? What production strategies are employed to secure mixes translating effectively to a range of formats and playback technologies? And is it possible to innovate in the bass spectrum without sacrificing sensible mix architecture? The authors deploy creative practice utilising their extensive combined professional catalogue to examine the creative possibilities—and limitations—for panoramic placement of low end in the mix. Techniques range from widening pre-recorded mono bass sources via processing, to extending sampled and synthetic low-end material for the purposes of specific stylisation and contemporary modes of consumption. The aim is to investigate the aesthetic and pragmatic causality behind the phenomenon, but also to explore innovative methodologies for producing stereo bass in mixes that cater for future listening modes and formats.

Douglas Heath: Monism vs. pluralism in record production: a comparison of contemporary heavy metal verses dub-reggae music

Abstract: This study explores the concepts and cultural practices of monism in heavy metal and pluralism in dub reggae within a recording production context. Utilising two unique contrasting case studies, the paper explores the current reality of popular music production in practice, and applies a practice-led participant-observation approach grounded in auto-ethnography addressing the ‘missing musician’ concept which is a perspective absent in the literature of popular music studies (Tagg, 2011; Frabbri, 2013) and metal music studies (Pillipov, 2012). The study contrasts the two genres from both the musicians’ preference and genre specific production techniques. Contemporary heavy metal music production has defining features of the genre’s high commercial standard of production as heaviness and sonic weight combined with definition and intelligibility of the instrumentation in-
volved (Mynett, 2012). Individualist and isolated recording techniques are diligently employed consistently in this genre to create this high level of production. This requires an intrinsic monistic approach from the musician’s performance through isolation of recording processes. Literature around ‘live-ness’ (Auslander, 1998; Croft, 2007; Emmerson, 1994; Grossberg, 1993; Sanden, 2009) focuses primarily on issues concerning the translation of studio performances to the stage. Pioneers of the dub reggae genre King Tubby and Lee Scratch Perry approached the act of live dub mixing in much the same way as an instrumentalist might approach performance on a conventional instrument. ‘Performance recordivity’ indicates the dub genre has always had a long history within studio music in a live context with a sense of live-ness and increased audience engagement (Knowles, Hewitt, 2012). Group inclusion and pluralism within reggae culture and music lends to ideals of group performance in interactivity. These concepts of live-ness and community commonly associated with the dub genre are investigated as the influences and the preferred technique of live multi-tracking of the band in the same room with minimal overdubs.

**David Brackett:** Listening to Electric Miles: Collaboration and Creativity in the Jazz Recording Studio

**Abstract:** This presentation considers a case study from the archive of modern recorded jazz—Miles Davis’s Jack Johnson (1971)—as a means to address the collaborative aesthetic, technical, and social dimensions of record production. The release of The Complete Jack Johnson Sessions (2003) contains a number of alternate takes and “inserts” that were cut up and spliced together to create extended tracks on several albums by Davis released from 1971 to 1974. Building upon Paul Tingen’s foundational work on Davis’ “electric” recordings, and drawing on recent work in popular music studies, ethnomusicology, and studies in the art of record production, I ask how this studio collaboration affects our notions of authorship and creativity in jazz. What is the relationship between the countless decisions in the studio and the sound that results? I will study studio practices and interactions in an integrated fashion, bringing together technological, practical, social, and creative/artistic components through a detailed consideration of the construction of a specific track from the album, “Yesternow.” The study of this little-discussed album by a canonical artist allows us to show how a specific approach to recording, as well as changes in recording technology, can be correlated with sonic and formal differences: the complete takes and infrequent splices from Davis’ early work up to In a Silent Way (1968) giving way to the ad-
Additive forms of Jack Johnson compiled from open-ended sections defined by funk- and rock-based ostinati.

Session 6 C

Ingrid M. Tolstad: Entangled musical spheres: Conceptualizing cross-cultural interactions in pop music production

Abstract: Notions of center and periphery are well-known within the musicological discourse (see for instance Hesmondhalgh & Born 2000). However, little attention has been paid to their presence in unfolding studio production processes. How are notions of center and periphery, of mainstream and ‘other’, articulated and negotiated in pop musical studio production? How can we conceptualize such notions, and the ongoing interaction between them, in terms that extend beyond their immediate dichotomy? Drawing on a year of studio fieldwork, this paper explores a series of Swedish and Russian pop musical interactions in which ‘the Russian’ was continuously constructed as an inferior and peripheral ‘musical other’, characterized as ‘lagging behind’, having ‘bad taste’, and being unprofessional. In comparison, ‘the Swedish’ or ‘Western’ was considered to represent an expertise capable of producing ‘up-to-date’, ‘cool’, high quality pop music in a professional manner. Strongly related to the extent of functionality within their copyright remuneration systems, Swedes and Russians interacted based on very different understandings of how music is financed, produced, organized, distributed, consumed and profited from. These discrepancies played out not only as disagreements and discussions about what kinds of resources should be put into what kinds of music making activities, but was also fundamentally intertwined with evaluations and priorities made concerning sonic and musical features. Distinctions made between center and periphery thus became entangled with notions and perceptions of aesthetic quality, professionalism and modernity. Drawing on traditional anthropological notions of economic spheres (Barth 1981, Bohannan & Bohannan 1968), these two Swedish and Russian pop musical entanglements are here conceptualized as a form of musical spheres (Tolstad 2016), representing systematically divergent paths of circulation, perceptions of aesthetic value, and practices in the making of pop music. The paper also shows how the dynamic character of this concept makes it transferable to other musical contexts.
**Dana Devlieger: "Pullin' Out of Here to Win:" The Changing Narrative of "Thunder Road"**

Abstract: This paper examines three versions of Bruce Springsteen’s “Thunder Road,” all recorded in 1975: an early live performance in February, the studio recording released in August, and a later live performance in October. Building on Zbikowski’s work on conceptual blending in German lied, I argue that conceptual blending can influence the listener’s understanding of narrative because musical choices can clarify lyric ambiguities through blends. I also assert that blends found in popular music are clearer than those of classical music because the primary text (the recording) is the same every time it is played. However, “Thunder Road” highlights an interesting issue present in popular music: the artist can continue to revise his arrangement of a song in live performances, drastically altering the song’s narrative. Springsteen’s alterations to the song’s music and lyrics during 1975 create three distinct protagonists, each negotiating their own relationship to the desire to escape from small town life. In the earliest recording, the protagonist comes off as young and restless, fueled more by boredom than desperation. The studio recording tells the story of a desperate young man seeking to escape from unpleasant circumstances. Finally, the later recording depicts a much older man who wants to escape but knows that he cannot. This paper investigates the different features of each recording that contribute to these altered narratives.

**Lachlan Goold & Phil Graham: The uncertain future of the Large-Format recording studio**

Abstract: This paper is an investigation of new approaches to music production that have emerged over the past thirty years and greatly accelerated over the past decade. The changes have primarily been facilitated by the proliferation of low cost digital music production tools. A secondary trigger has been the subsequent decline in recording budgets concomitant with the decline of revenues for recorded music (Burgess, 2008, p. 1). The confluence of these issues has meant large-format recording studio processes are giving way to lower cost “desktop”, “bedroom”, or “DIY” music production processes (Bennett, 2012, p. 8; Théberge, 2012, pp. 89-90). Little attention has been focused on evaluating and understanding the outcomes of these new technologies from perceptual, actual, and aesthetic perspectives specifically as they pertain to the meaning of the term “recording studio” as a particular kind of space. This paper seeks to address this gap in understanding through evaluations of a comparative set of recordings.
completed in disparate scenarios. It compares processes of recording in large-format recording studios with those associated with situations dictated by a set of limitations common to smaller budget "DIY" projects, and with those involved with a hybrid approach combining both methods. As part of a PhD project, research methods include participant interviews, participant observation, peer review of the material produced, and analysis of attitudes towards creative labour under these different circumstances. This paper will focus on the peer review of the material produced by seven highly regarded producers which indicates that while different recording spaces present different advantages and disadvantages, in terms of recording quality, the large-format recording studio appears to present no major technical or aesthetic advantage over DIY environments in terms of recording outcomes. Given the proliferation of DIY approaches to recording, this topic has important ramifications across the music and recording industries and brings into question the viability of the large-format recording space.

Session 7 A

Yngvar Kjus: Producing consumption: Facilitating musical immersion across online and offline media

Abstract: This paper examines the facilitation of music experience and immersion across live and recorded music formats. These two forms of music delivery (and their intersection) are often evaluated from an economic macro-perspective. Live music has been regarded as advertising for records, and upheavals from digital distribution in recent years have therefore been seen as a reversal of this effect. The relationship between the two domains has barely been approached in terms of the efforts of concert organisers and music distributors to enable communication between artists and their audiences. This paper reveals how digital media can be used to bridge artists’ concerts and the growing recorded archives of new online music services, thereby facilitating novel music experiences. It identifies evolving practices through which live-music organisers tap into online music archives to contextualise performances, and online music services bring their archives to life through interactions with concerts and festivals. Case studies involve the Norwegian streaming service WiMP/Tidal and the Øya festival and include several interviews with key personnel as well as music consumption statistics. The paper studies the efforts and techniques of these intermediaries towards (re)gaining the trust of artists and audiences. In doing so, it explores the analytical potential of
the psychological terminology of mentalization, which is, in short, the capacity to imagine and respond to what others are thinking and feeling. These processes affect the ways in which new technology is used to integrate and enhance social and psychological experiences of music. The paper assesses the power of live and recorded music intermediaries over how people receive, experience and immerse into music, considering, among other things, the new intersections between the curatorial work conducted by humans and the guidance offered by machines and algorithms of digital music archives.

**Jason Woolley: The cultural politics of using technology to support the aesthetic in Jazz record production**

**Abstract:** In the pursuit of the aesthetic, producers and artists are able to utilise technology to create and process recordings, and in doing so, shape the presentation of the performances contained within the recordings. We might consider the application of studio techniques in order to produce the recordings, as ‘studio interventions’. Ethical acceptance on the extent the technology of studio intervention can be applied can vary between artist, listeners and genres. Taylor argues that the goal of Rock production is to achieve the best aesthetic by utilising the ‘…manipulation of electronic sound…in order to get the best final result’ (Taylor in Jago 2013). Consider also the countless main stream pop releases where it is acceptable and even stylistic for the vocal line to be audibly supported by pitch correction processing. A genre where complex interactions and tensions exist between the affordances technology might offer the aesthetic, and which studio interventions are culturally acceptable to the artist and listener, is Jazz. This is perhaps because at the stylistic and cultural core of Jazz, both artist and listener seek the authentic ‘ephemeral interaction’ of the music, and anything other than this might be considered ‘dishonest’ (Jago 2013). Further to this is the case of Free Jazz, where for some, the use of technology to shape the recording of an improvised performance means the ‘freeness’ succumbs to the dominant ideology of the production context (Corbett in Gabbard 1995). This paper discusses the attitudes contemporary Jazz musicians have toward the use of technology and the thresholds of studio intervention they are willing to cross in order to achieve their aesthetic. The discussion draws upon personal experiences of mainly Free Jazz production, and also the returns of surveyed Jazz musicians, who were polled on the subject of the use of technology in the production of ‘Straight-ahead’ Jazz and Free Jazz records.
Martin Koszolko: The Tactile Evolution — Mobile Electronic Music Production and Digital Affordances of Apps

Abstract: A growing trend among the community of music producers is a return to the tactile experience of music creation tools and a focus on hardware that reduces the reliance on desktop systems. In this paper, grounded in my music production practice, I investigate the various affordances of mobile music technologies in the field of electronic music production. I examine a selection of iOS music apps as well as hardware and highlight affordances concerning production, performance, improvisation, portability and music theory. I also consider how mobile producers relate to the concept of studio space in the context of portable technologies. The producers discussed here undertake multiple roles reflected in Burgess’s understanding of a producer as an auteur, a self-produced artist leading the creative process (2013). The different feel and tangibility of mobile technologies impact upon the way producers choose to engage with particular formats. This paper explores an area of the contemporary digital landscape that is significantly changing the creative practice of music producers around the world. My analysis shows that many iOS music production apps as well as mobile music making hardware offer new ways of musical expression. Such tools enable a tactile approach to music performance and facilitate experimentation, and offer significant advantages over other existing technologies. Drawing on notions of portability and performativity, I consider how contemporary technologies have generated renewed approaches to electronic music production and performance. I conclude by suggesting that portable technologies allow new ways of interaction, music composition and playability, which empower producers to redefine their methods of music creation.

Session 7 B

David Ward Francis: A Creep Climbing Up the Walls: The Dynamic Timbre of Rebellion, Angst, and Escapism in Radiohead

Abstract: Albin Zak, in The Poetics of Rock, argues that amplified attention and control to timbre developed a poetic timbral language with timbral tropes between specific techniques and meanings. One such trooped timbral mannerism is rock’s connection of distortion to rebellion. In recent years, improvements in recording technology enabled effective contrasts between distortion and clean timbres. Radiohead dynamizes this trope and moves sonic events through coordinated continuums of brightness and aggression. Reducing timbral qualities
and proposing potential meanings remain problematic. Harris Berger, Cirro Scotto, and Robert Walser provide definitions of distortion and examine the relationships between distortion and ideologies in several sub-genres, yet, restrict the definition to production techniques. Based on David Huron’s approach to timbre, this paper advocates reducing timbral dimensions to three perceptually relevant dimensions of spectral centroid, balance, and aggression, plotting sounds and sound aggregates through a non-sequential space—Relative Timbral Space. This space combines technological and acoustic sonic origins into one space—allowing for acoustic distortion generation—and provides mappable space for meaning. Eric Clarke’s ecological listening analysis (refined by Brad Osborn and Mark Slater) enables metaphorical interpretation of resulting timbral relationships. Radiohead’s 1993 single, “Creep,” “Feral” (2011), and “Climbing Up the Walls” (1997) each make use of distortion, but modify the trope to affect meaning. RT-space is used to compare timbral relationships, finding subtle differences between each. Using ecological listening, timbral differences are found to contain cogent metaphorical statements. Small enclaves of contrasting bright timbres invoke angst. If that enclave increases in volume against the contrasting timbres, escapism is invoked. “Climbing” is found to form the most cogent metaphorical statement, encapsulating Radiohead’s transition to a new sound identity found in Kid A (2000).

**Alex Stevenson: Digital aesthetics in contemporary popular music performance**

**Abstract:** The introduction of digital sampling technology enabled electronic music producers to manipulate existing musical performances to achieve complex grooves that would have been significantly challenging, if not impossible for musicians to perform live. In hip hop, micro-rhythmic gestures created genre-specific grooves which, as discussed by D’Errico (2011), formed a significant component in the creation of sonic signatures for many influential hip hop producers. In drum’n’bass, samples of funk breakbeats were subjected to increases in both pitch and tempo, which along with the fragmentation and rearrangement of components of the sample, created complex rhythmic performances with timbral variations and characteristics far removed from those captured in the original recording of the drummer’s performance. Furthermore, in many experimental electronic music styles, as part of what Cascone refers to as a “‘post digital” aesthetic’ (2000, p. 13), digital artefacts and failures became defining musical characteristics. The glitches of digital playback failure and what Brovig-Hannsen & Danielsen refer to as ‘digital silence’ (2016) provide ex-
amples of digital characteristics that have become common place in contemporary popular music production. As a result of this, many contemporary popular musicians face the challenge of incorporating these musical characteristics performance practice. These musicians have therefor had to develop innovative performance techniques which allow them to (re)create these musical characteristics which conforms to a digital aesthetic, often without the use of digital technology. Drawing on the work of authors such as Keil (1995), Danielsen (2006; 2013) and Oliver (2015) utilising rhythmical and timbral analysis of musical works, this paper explores the impact of digital aesthetics on contemporary popular musicians and explores the ways in which the emerging techniques and practices can be better understood within popular music discourse.

Méi-Ra St-Laurent: « It’s kind of in the middle » The “mid-fi” aesthetic: toward a new designation of black metal aesthetic of recording. The case of the Quebec’s black metal scene

Abstract:
Since its advent in Norway in the late 1980s, black metal music is referred to a "lo-fi" sound aesthetic. For black metal fans and musicians, this designation not only allowed to define the type of sonority used, but also to measure the level of "authenticity" of a group. Indeed, a black metal band which used a lo-fi sound esthetic was considered to have a sincerer artistic approach than a black metal band using a hi-fi aesthetic. However, this designation has become problematic. Indeed, as early as the mid-1990s, many black metal bands moved gradually away from this lo-fi sound, but without embracing the hi-fi sound aesthetic, which would be considered as a betrayal of black metal’s principles (Reyes 2013). Moreover, several groups of contemporary black metal do not necessarily seek to imitate the sound aesthetic of the first black metal bands and will instead seek to reproduce a sound "in the middle", between lo-fi and hi-fi. However, academics working in the black metal milieu continue to designate the black metal sound as “lo-fi”, without necessarily taking into account the gradual change of sound aesthetics characterizing these groups. In this paper, I will be interested in this phenomenon by first defining what is a lo-fi aesthetic in the context of black metal, based on the definition of Kromhout (2009). Then, I will be interested in the evolution of this aesthetic by using different sound examples. Finally, I will explain concretely how contemporary black metal bands are considering their own sound aesthetics based on ethnographic data I have collected during different interviews I realized in 2016 with the black metal scene in Quebec. This will finally lead me to propose a new terminology allowing to better
qualify the sonority of many contemporary black metal bands, that is the “mid-fi” aesthetic.

Session 7 C

Jason Fick: Why Study Music Production at a University?: The Benefits of a Multidisciplinary Approach to Enhance Student Learning and Career Preparation

Abstract: Learning the craft of music production is traditionally grounded in practical, hands-on experiences, and a college degree is not often necessary to obtain an entry-level position in the industry. Despite this fact there has been a steady increase in music production-related degree programs at higher education institutions worldwide in recent decades. If this is an area that does not require a degree to be successful, why should anyone consider pursuing it as a course of study at a university? What information can be learned in higher education that cannot through on-the-job experience? This paper will discuss the benefits of a multidisciplinary approach to music production pedagogy that hinges on the intersection of science and liberal arts curricula. With a limited amount of specialized, entry-level positions in the industry, the music production professional of today must function as an audio expert capable of a vast range of skills. Enrolling in a program with an integrated, multidisciplinary curriculum can empower students with versatile skills that encourage connections between their immediate specialty and other areas, including the liberal arts and sciences. My teaching approaches balance hands-on, real-world scenarios with reflective writing and discussion, research and experimentation, and multi-disciplinary collaboration. This provides students opportunities to build critical thinking, analytical, communication, group work, and problem solving skills, all of which are basic competencies that current research proves lacking among new hires. Furthermore, university students have the opportunity to broaden their methods and application beyond their discipline through the general core curriculum. With these concerns in mind, I will discuss building a program at Oregon State University that aims to develop well-rounded individuals exposed to a variety of creative, technical, and multidisciplinary approaches to the field of music production.
Carlos Caballero: The forms of recording and sound aesthetic of tropical and urban music in Colombia. A parallel between music production of the 60’s and 70’s and the contemporary sound

Abstract: Medellin was the epicenter of Colombia’s music and record production between the middle of the 50’s and the end of the 90’s. Such production boom came with the so-called “paisa” sound, characterized by youthful and irreverent groups that mixed rock and traditional rhythms of the Atlantic coast, specifically between the 60’s and 70’s. In this period, a musical and social sound revolution was developed revolving around not only music but a universal sense of transformation. That era was characterized by common elements around activities related to record production, which marked a particular sound aesthetic given not only by the technological scope of the recording equipment of the time, but by the social characteristics that allowed a Third World country to position a sound that would transcend its borders and would carry to all the world its artistic manifestations by means of its most recognized groups of tropical music. At present, the sound aesthetic of music generated from this epicenter has similar ingredients that influence its results, both from the technological and social point of view, but certainly with very different aesthetic meanings. However, the cultural phenomenon is very similar, based on irreverent young people making music that is internationally accepted as an innovative commercial element (and to which popular artists of commercial success have been added), but that for traditional artists, folklorists, musicians and cultural investigators generates more animosity than admiration. Taking this into account, we can ask: Is it possible to make a cultural parallel between the phenomenon of Medellin’s tropical urban music in the 60’s and 70’s with the current phenomenon of urban music? How much did technology influence the final aesthetic result? What perception did the society of such era have of these artists, and what perception does it have now? Is it possible to establish similarities?

Phillip McIntyre & Paul Thompson: Tradition and Innovation as Complementary Pairs in the Creation of a Body of Recorded Work: The Beatles Journey from Mono through to Stereo

Abstract: Systems are characterised by emergence, scalability, interdependence, networks, nonlinearity and complementarity (McIntyre et.al. 2016, p. 15). It is the last of these, complementarity, that this paper examines. While polar oppositions typify most human communication they depend on and complement each other for existence. As Negus and Pickering indicate in their book Communication, Creativity
and Cultural Value (2004), it is “a common misconception to regard innovation and traditions as diametrically opposed to each other” (2004, p.91). They suggest that “rather than seeing them as deeply divided, we want to consider tradition and innovative forms and practices as informing and supporting each other. It is only by thinking about their interrelationship that we can understand processes of creativity and cultural change” (ibid). To put this another way, “‘new’ is meaningful only in reference to the ‘old’. Original thought does not exist in a vacuum. It must operate on a set of already existing objects, rules, representations or notions” (Csikszentmihalyi 1999, p.315). In the light of these ideas this paper will examine the historical development of the creative practices of the Beatles’ recorded oeuvre. The Beatles growth across their brief recorded history moved from what is now seen as an anachronistic medium, mono, to what was then seen as an innovative format, stereo. They absorbed the lessons of using a relatively traditional process to establish a novel and valued creative output.

Session 8 A

Amy Blier-Carruthers: The Influence of Technology on Performance - Classical Perspectives

Abstract: What effect does recording have on the performances that are captured, and the performers that are being recorded? It is an oft-repeated axiom that studio recordings are not the same thing as live performances, and this has been discussed from a largely theoretical standpoint, but comparatively little research has been undertaken into what actually happens behind the often closed doors of a classical recording studio. In order to begin to understand the influence of recording technology on performance, it would help to first consider some questions: What changes do classical musicians have to make, consciously and unconsciously, to produce a performance suitable for recording? Is the technology driving the aesthetic choices, or is it at the service of those aesthetic concerns? How do the accepted working practices affect the artistic output? How can musicians and production teams use the tools of the recording studio to their creative advantage? This paper will explore these questions, and offer a brief selection of case-studies, which cover examples which fit under the headings of ‘Mono’, ‘Stereo’ and ‘Multi’. These range from an analysis of how details of performance vary between the live concert and the studio recording (illustrated by rarely-heard sound examples from the Sir Charles Mackerras Collection of recordings), through a recent exper-
iment in which modern musicians have experienced the effects of recording onto wax using early-20th-century acoustic recording techniques, to current experiments in the ‘hyper-production’ of classical repertoire. It will conclude with examples of some recent and forthcoming attempts at using more progressive technological tools and working practices in classical music recording, and a consideration of new developing strategies for the training of conservatoire students to prepare them for the effects of recording in order to enable them to work creatively and collaboratively in the recording studio.

**Hans Lindetorp: Multiraction - An interactive musical experience in a multi dimensional audio environment**

**Abstract:** This demonstration presents the result of a research project aiming to discover new possibilities for interactive live music productions using cutting edge web audio technology. In the presentation the audience is invited to interact using their smartphones through which different layers and aspects of the music can be controlled. The music is pre-composed and produced in a non-linear fashion to be responsive to the inputs from the audience. Lots of audio clips are finally synchronized in real time and routed through a multi-channel speaker system. The project has been running for three years as an iterative process of software development and student productions where artistic visions has been an on-going driving factor for further software development. The result is a new web based musical framework, new strategies for productions and new courses at Royal College of Music where traditional music producers are introduced to the challenges of making interactive music. The technical foundation for the framework is Web Audio API, nodejs and socket.io and the project relates very well to the latest’s research from CoSiMa (a research project at Ircam, Paris). The study has led to important insights regarding still unsolved problems for music integration in interactive environments. One important aspect is the relation between the artistic freedom for composers on one hand and the technological limitations on the other. Another aspect is the dynamic and flexible time in a live musical performance and the need for rigidly quantized audio files in a loop-based, interactive environment. In the demonstration, the audience will be presented by solutions to some of the challenges and immersed by music where new questions arise: Who is the music producer? What is an audience? Is there an artist?
**Simon Zagorski-Thomas:** Looney Tunes: Sonic Cartoons and Semantic Audio

**Abstract:** The idea of sonic cartoons (Zagorski-Thomas 2014) is built on a theoretical model drawn from the ecological approach to perception (Gibson 1979) and embodied cognition (Feldman 2006, Lakoff & Johnson 2003). The presentation will begin with a brief explanation of how the neuroscience of mirror systems and this theoretical model can lead us to an understanding of recorded sounds in terms of activity performed by one or more human and/or non-human agents in an environment. A non-human agent is an object which has either been physically designed and built or found and re-purposed by a human agent and, thus, embodies their agency in some way. In addition, in much the same way that our understanding of abstract visual art is based on ‘what it is not’, multi-track and electronic music can be interpreted as something impossible and yet understandable. Based on an on-going series of experiments, this project seeks to identify a range of sonic invariant properties relating to types of activity, the levels of energy expenditure involved, types of human agent, types of non-human agents and types of environment. While these relate to a broader musicological research project, they also provide a framework for exploring processing that is based on semantic audio. Thus, for example, if a range of invariant properties can be systematically identified that relate to: 1) the size of a drum; 2) the way it is tuned 3) the physical properties of the skin, the shell and the beater; 4) the energy and shape of the sounding gesture; 5) the size, shape and material construction of the space 6) the orientation and proximity of the listener to the sounding object and the various ways these things interact with each other, this opens up the possibility of designing semantic audio-based plug-ins that approach shaping and creating sound from a variety of new perspectives.

**Session 8 B**

**Alan Williams:** Setting the Stage: Surround Sound Auteurs and the Fragmenting of Genre

**Abstract:** Multi-channel sonic experience is derived from a myriad of technological processes, shaped by market forces, configured by creative decision makers and translated through audience taste preferences. From the failed launch of quadrophonic sound in the 1970s, through the currently limited yet sustained niche market for 5.1 music releases, a select number of mix engineers and producers established
paradigms for defining expanded sound stages. Whereas stereophonic mix practices in popular music became ever more codified during the 1970s, the relative paucity of multi-channel releases has preserved the individual sonic fingerprint of mixers working in surround sound. Moreover, market forces have constricted their work to musical genres that appeal to the audiophile community that supports the format. Building upon the work of Moylan, Théberge, Sterne and others, this study examines the work of Elliot Scheiner, Bob Clearmountain, Dave Fridmann, Steven Wilson and Giles Martin to not only analyze the sonic signatures of their mixes, but to address how their conceptions of the soundstage become associated with specific genres, and serve to establish micro-genres of their own.

Andrew Bourbon & Daniel Pratt: Existing In Between Phase

Abstract: This paper examines the phase interaction of multi-microphone recording and mixing. Its intent is to develop an in depth understanding of relationships 'in between phase' in order to produce better recorded text. For any student studying recording and mixing, phase is a subject that is often discussed. It is used in technical descriptions relating to the acoustic and electronic summation of multiple sources and is explained in practical recording workshops. One of the conceptual challenges that students face is hearing phase variance and implementing the appropriate action to remedy what they hear. For educators, explaining phase becomes problematic when that variance represents a shift that cannot be solved via a simple polarity reversal. We explore the use of metering and phase manipulation in the recording and mixing of audio. On the recording side, we investigate through the creation of a drum recording documentary. The documentary examines the capture of phase interaction information, the analysis, and correction of resulting issues. The data analysis of the phase relationships will inform both analog processing choices and the physical movement of microphones. We demonstrate real time phase interaction measurement using the innovative metering in the Sound Radix Auto Align plugin. This information helps us to demystify phase interaction and enables us to develop new methods for microphone placement. We then manipulate the complex phase relationships in a Big Band recording featuring thirty two microphones across twenty sources using a range of mono and stereo techniques. We use metering and analysis techniques to inform our alignment and manipulation of this pre recorded work. In both of these case studies the sonic impact of the phase analysis and manipulation will be presented. The data will be used as a pedagogical tool for the demystification of phase in the teaching environment.
**Shara Rambarran: ‘The Ghost Inside’: Exploring the aesthetic retro sounds and vintage production in the works of the producer, Danger Mouse**

**Abstract:** Brian Burton, better known as Danger Mouse, notably recognized as a hip hop music producer, has become a demanding figure in the music industry and has worked with the likes of Damon Albarn/Gorillaz, A$AP Rocky and Beck. Ever since his controversial mash up, The Grey Album (2004), Burton has occasionally sidelined his musical roots as a hip hop producer to focus on his trademark sound. While genre-blending is still evident in his productions, there are elements used in his works that has shaped Burton’s individuality as an auteur and producer. He has achieved this by applying retro sounds in his music through the use of vintage instrumentation and production. This paper examines the inclusion of retro sounds in music resulting from the use of vintage instrumentation and production. Based on Schaeffer’s concept of ‘reduced listening’ (1966), the paper argues that the sounds are not heard for ornamental pleasure, but for cathartic experience displayed by musical signs from the 1960s/70s that are covertly transformed into haunting sonics. The analysis of these retro signifiers are approached by: understanding its socio-cultural context with reference to Derrida’s hauntology (1993); and, aesthetically exploring the affect and effect of the sounds that are produced. Arguably, it is the musicians who contribute in recreating the sounds by playing the instruments in the studio. It is also however, the placing of the vintage instruments in its recording space, use of technology, and overall production that captures and finalizes the retro signifiers in sounds, and therefore allows Burton to revisit this arrangement in other projects. The paper explores this by referring to Grandaddy, Michael Kiwanuka and Red Hot Chili Peppers as case studies, and argues that the recordings demonstrate Burton’s fixation in bringing past sounds to the present resulting in a retro-futuristic and multi-audio experience for the listener.

**Session 8 C**

**Carl Flattery: Memory and Place in Songwriting and Production: The work of Hannah Peel**

**Abstract:** Hannah Peel is a singer-songwriter, multi-instrumentalist, composer and arranger who’s work continually reflects the themes of memory and place through the lyrics, arrangements and production.
Her most recent solo album ‘Awake But Always Dreaming’ (Peel, 2016) was written in response to her Grandmother’s dementia. After seeing the effects of music on her Grandmother, how connections are made through sharing musical experiences, she was inspired to write the album. She linked up with a research project into dementia and used the scientific research to inform the main themes of the album combined with her own personal response to her relationship with her grandmother. The idea of memory and place can be observed throughout Hannah’s career. In the two albums made as part of the collective The Magnetic North, conveying memory and place is the central ambition, being autobiographical and linking into ideas of Psychogeography. The first album ‘Orkney: Symphony of the Magnetic North’ (The Magnetic North, 2012) focused on the legends surrounding Orkney, which is the birthplace of lead singer Erland Cooper. The second album, ‘Prospect of Skelmersdale’ (The Magnetic North, 2016), would prove to be a more unlikely theme, looking at the new town of Skelmersdale, where the band’s guitarist Simon Tong had spent his teenage years. Within these albums, memory and place is communicated through the lyrics, arrangements and additional sound effects and archive voice recordings. In Hannah’s ‘Rebox’ projects, memory, though less intentional, maybe more subliminal, is a key trait of the composition and arrangements. Her use of a music box to reinterpret her favourite pop songs conjures up feelings of childhood. Her project, ‘Mary Casio’, is a collaboration with a colliery brass brand with Hannah performing 70’s analogue synths and telling the story of 86 year old Mary Casio, based in the mining town of Barnsley.

David Thyrén: Searching for Sophia in Music Production Education – Dag Volle as an example of indifference, individual skills and special knowledge in Swedish record production

Abstract: In the project Searching for Sophia in Music Production Education, a team of researchers from the Royal College of Music, study various formal and informal learning processes connected to education in music production in higher education. This paper is part of that project. Previous research shows a variety of competences that are used and needed among musicians, music producers and others active in the art of music production. Our interest has a background in the worldwide export of music from the Nordic countries. Several of the Nordic international successful songwriters, musicians and music producers have more of an informal background than a scholastic formal education. In this paper, I focus on the Swedish record producer Dag Volle [Denniz Pop] (1963-1998). Volle enjoyed international success in the 1990’s, recording and producing e.g. Ace of Base, Backstreet
Boys and Michael Jackson. Volle was also the mentor of Max Martin. Volle was not a trained musician, he couldn’t read music and had no formal knowledge of music theory. He started his career as a DJ in the 1980’s and ended up as one the world’s most successful record producers, before his untimely death in 1998. The results indicate that Volle was a creative person, almost childish in a way. He was driven by his love of music. Volle felt slightly inferior for not being able to read music but didn’t let it affect his working ethos. He worked very long hours in the studio, often under pressure, but managed to have fun all the same. Volle had great intuition and didn’t hesitate to make fast decisions that initially seemed risky but later paid off in an unlikely way.

Simon Sjöstedt, Felix Brag, Sophie Verdonk, Erik Petersson & Ludvig Klint: An analysis of creativity aspects in songwriting books

Abstract: This study shows how different aspects of creativity are expressed in current literature about songwriting. The study includes twelve recently published handbooks that were selected after a search on online bookshops with songwriting and synonymous terms as keywords. Knowledge-critical analysis was used to study who the authors of these books are, who the recipients of the books are expected to be and how concepts of creativity are mentioned and used in the texts. Although the analysis shows that there are major differences between the analysed books, all authors, who all have extensive professional experience from the music industry, uses words like “gut feeling” or similar expressions to describe what primarily is needed to develop creativity and success within the field of songwriting. Craftsmanship, the importance of workflow and persistence combined with hard work that will pay off in the form of acquired knowledge and the ability to be creative when inspiration strikes, are also perceptions that clearly appear in the analysed literature. There is a lot of variation in the content and the layout of the different analysed books. This can mainly be explained by the fact that the handbooks are written for different audiences but also that the authors have different industry backgrounds. One possible shortcoming in the analysed literature is that the contemporary integration of music production and songwriting, as described in other current literature on record production, is not addressed in the extent that would be reasonable to reflect the music production industry of today. This study will be followed up by three new sub studies. Firstly an analysis of current hits where theories from the analysed songwriting books are used, secondly a production study based on the same theory and thirdly, a follow-up analysis of similar digital books and other online teaching materials.
**Hans Gardemar & Jan-Olof Gullö: Artists, musicians and music producers: Same but different?**

**Abstract:** The purpose of this project is to analyse core characteristics of key players in the music industry: artists, musicians, music producers and others. The background of the project is the growing interest among students to study popular music and music production in Scandinavian higher education. Their interest can partly be explained by the successful music exports from the Nordic countries. A clear problem for the universities is that many of the students who are interested in various kinds of music related education often much more would prefer an artist career than to become a trained musician or music teacher. At the same time, very little of the training offered, in higher education, is focused on developing talents into full-fledged artists. Therefore, there is a gap between what is offered and what many students desire. Another problem is that many of the teachers in higher music education often have very limited experience as artists, but often have a good and extensive experience as musicians. Therefore, students in higher music education perhaps develop musician skills to a far greater degree than artist skills. A similar problem can be found in music producer education, where many of those who teach may have extensive experience as sound engineers rather than as producers. In order to create knowledge about this problems and to develop the higher education in music, this project collects multifaceted data through interviews with key players in the music industry: artists, musicians, music producers and others as well as by and analyses of their activities.
Photo D. From keynote Linda Portnoff’s presentation. (Photo credit: Eleonor Gislason-Ferrari)