An Immersive Approach to 3D-Spatialized Music Composition*

Tools and Pilot Survey

D. Ledoux
Université de Montréal
Montréal, Québec
david.ledoux@umontreal.ca

R. Normandeau
Univeristé de Montréal
Montréal, Québec
robert.normandeau@umontreal.ca

Figure 1: A panoramic view of the 32.2 loudspeaker dome environment used for the pilot survey.

ABSTRACT

Open-sourced 3D sound spatialisation software tools, developed by the Groupe de Recherche en Immersion Spatiale (GRIS) at Université de Montréal, were used as an integrated part of two music compositions, in an immersive, object-based audio approach. A preliminary listening experience has been conducted on two separate groups of students, in a 32.2 loudspeakers dome, as a pilot for a case study that aims to get a better sense of the immersive affect of complex spatialized compositions through the listener's reception behaviors. Data collected from their comments on these two different 3D-spatialized musics have been analysed to extract converging expressions of immersive qualities.

CCS CONCEPTS

• **Applied computing** → **Sound and music computing**; • **General and reference** → *Surveys and overviews*; • **Software and its engineering** → Open source model;

KEYWORDS

Immersive sound, spatial audio, music composition, music analysis, object-based audio, audio software, listening behaviours

ACM Reference Format:
D. Ledoux and R. Normandeau. 2018. An Immersive Approach to 3D-Spatialized Music Composition: Tools and Pilot Survey. In *Audio Mostly 2018: Sound in Immersion and Emotion (AM'18), September 12–14, 2018, Wrexham, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3243274.3243300

1 INTRODUCTION

The classical interpretation of "immersion" refers to a *new world* or *universe* of action and perception rendered possible by the means of a technical device: to be immersed in a space means that we can localize objects and events from within that space [2]. Therefore, to be immersed in sound can be understood from the listener's point of view, within an unusual, mediated listening configuration, as opposed to a usual, real-life listening situation. But, as Bruno Trentini states, to be immersed can also be understood as being able

*"3D" hereinafter refers to the production process of spatialized music – the three dimensions being spatial parameters: depth, azimuth and height –, not its perception.

AM'18, September 12–14, 2018, Wrexham, United Kingdom

ACM ISBN 978-1-4503-6609-0/18/09…$15.00
https://doi.org/10.1145/3243274.3243300

to render a proposed universe accessible to ourselves [10]. We could then argue that immersive audio devices may not directly produce a sense of immersion for the listener, but only makes it possible by augmenting the content's immersive potential. Although 3D-audio spatialisation tools can easily allow composers to use space and trajectories as musical parameters — just like pitch, rhythm and intensity — how are their employed spatialization strategies perceived and received by the listener? Acoustics and psychoacoustics knowledge have played a crucial part in the development of spatialisation softwares, mostly by relying on physics and human sound perception. But what can we say about how complex musical content, evolving in a multidimensional space, affects a listener's sensation of being immersed? Furthermore, what is the extent of that immersion, its variations, its qualities? Even though 3D-spatial audio technologies are an integrated part of the creative strategies employed to immerse a listener, looking only at the technological device isn't sufficient to answer all these questions. We must also consider and interrogate the listener's point of view to evaluate and qualify the sense of immersion made possible by these tools and techniques. This paper reports on the preliminary stage of a case study on (immersive) listening behaviours of 3D-spatialized music, which aims to reach a conceptualization of the immersive (music) listening phenomenon. The pilot survey presented herein was conducted as part of a creative process and do not hold any scientific pretentions; this paper's goal is to share converging themes of expressions, extracted from listeners' verbalization data, and discuss the transposition of some general concepts of immersion to the specific phenomenon of sound immersion. It is believed that these surveys can offer relevant insights for other composers and sound artists who make use of 3D-spatialisation technologies in an immersive approach.

2 A SURVEY ON RECEPTION BEHAVIOURS OF 3D-SPATIALIZED MUSIC

2.1 Methodology

The whole research's qualitative method of analysis of empirical data – to which this pilot survey is thefi rst stage – is based on Pierre Paillé's adaptation of the "grounded theory", called *théorisation ancrée*. In short, it presents "an iterative approach of progressive theorization of a phenomenon" where "its evolution is neither expected nor related to the number of times a word or proposition appear in the data"[7]. In that sense, it consists more as an act of conceptualization by establishing an ongoing questionning process between the data collected and its analysis. While François Delalande also used a similar approach to construct his listening behaviour theory [4], this pilot survey is thefi rst iteration of a successive research process that aims to inform us on the listener's sense of immersion in the context of 3D-spatialized music.

2.1.1 Participants. Two separate groups of students (excluding the authors) listened to different acousmatic pieces of music: thefi rst group (G1), comprised of 12 participants, listened to *Ville Aux Cent Clochers* (2018), an unedited work of 11m30 composed by David Ledoux; while the second group (G2), comprised of 8 participants, listened to *La part des anges* (2011, 12), composed by Robert Normandeau [6] and lasting 14m30. Each group's listening experiment took place at different times and no participant took part in both of them.

2.1.2 Materials and musics. Participants of both listening experiments were sitting on chairs in a concert-like situation within a calibrated 32.2 channels dome (seefi gure 1), measuring 23 ft. in diameter and 8.3 ft. high from the lower ring of loudspeakers. Both musics were composed using the SpatGRIS/ServerGRIS spatialization tools (described in Section 3). Being made by different composers, the musical pieces present on the other hand very different sound sources and aesthetical approaches. Ledoux's work makes extensive use of recognizable sound sources, mostly originating from urban-related environements, in order to reproduce a sonic journey through a musically-augmented, urban-like environment. Contrariwise, Normandeau's work presents a more abstract and ambient-like music, in which "the evaporating sound material comes from a variety of capella songs of very old musical traditions" [6]. Although diverging in compositional strategies and aesthetics, both pieces were nevertheless intended to give its audience a sense of immersion.

2.1.3 Procedure. After a single listening of the music, participants were asked to comment on their experience in writing. For this pilot, a single open-ended question, borrowed from Antonio Alcazar [1], was trialled: "what have you heard and how did you listened to it?". As Alcazar states on his approach: "the qualitative methodology is that which offers the most useful tools since it conceives the reality as multiple andfl exible", adding that "its goal is the capture and the reconstruction of signified through an inductive process" [1]. Written comments were then analyzed by grouping expressions under recurring themes, eventually leading to key concepts in the next stages of this research.

2.1.4 Limitations. It must be pointed out that the number of participants is not sufficient to conclude any statistical evidence; our goal here is to extract converging (immersive) listening tendencies. Moreover, sitting inside a dome of loudspeakers may influence the listeners' expectations towards a sense of immersion. Here, our interest lies mostly in the immersive affect of the spatialized music content, beyond the employed immersive technological device. That is why both pieces of music were made using the same spatialization technology and presented in the same loudspeaker environment. Nevertheless, to minimize any novelty effect, all participants had previously experienced music in a multichannel dome listening environment at least once. Notwithstanding, G2 was generally more experienced[1] than G1 in such listening environment.

3 OBJECT-BASED AUDIO SPATIALIZATION TOOLS: SPATGRIS AND SERVERGRIS

This section briefly describes the 3D-spatialization tools employed in both compositions used for this survey.

3.1 Overview

Initially called ZirkOSC, the SpatGRIS plug-in (Mac AU/VST and Windows VST format) started being developed in 2012, by the

[1] G1 participants qualified themselves as ranging from novice to intermediary, while all G2 participants qualified their experience level as intermediary, to the least.

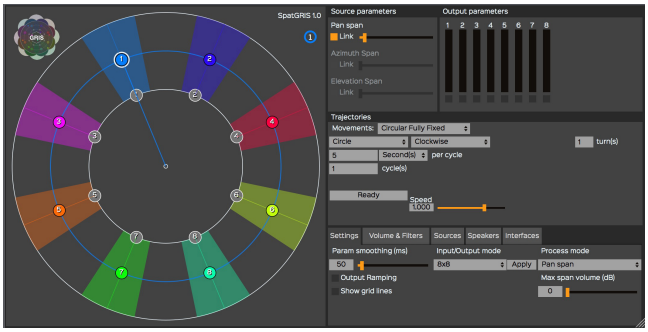

Figure 2: The SpatGRIS graphic user interface.

Groupe de Recherche en Immersion Spatiale (GRIS)[2] at the Université de Montréal, as a complementary Open Sound Control (OSC) interface controller for the Zirkonium spatialization software, designed by The Institute for Music and Acoustics at the ZKM | Center for Art in Media (Karlsruhe, Germany). The Zirkonium is a standalone application that uses Vector Base Amplitude Panning (VBAP) [8] to render virtual sound sources within a configurable dome-shaped loudspeaker setup. The ServerGRIS software has been developed since 2017 as a replacement to the Zirkonium, although its purpose remains to offer an accessible (free and open-sourced) way for composers to use space as a musical parameter with minimal interference with their creative workflow.

3.1.1 SpatGRIS. The SpatGRIS's interface (seefi gure 2) allows to control the position of sound sources over the surface of a dome-like virtual space by sending OSC data to an external audio spatializer called ServerGRIS (see section 3.1.2). Having the SpatGRIS inserted on a DAW's track allows the user to "record" the spatial coordinates and movements of each audio channel as automated data along the sequencer's timeline. Each channel output can then be considered as an audio object [11], whose distribution can be spanned vertically and/or horizontally, all over the dome's surface, without having to leave the DAW's window.

3.1.2 ServerGRIS. Each output channel from the DAW goes into the ServerGRIS (seefi gure 3) – via JackRouter (installed with the software) – in parallel with the OSC data from their corresponding SpatGRIS (seefi gure 4). The selected VBAP algorithm then positions the audio objects, based on their corresponding spatial coordinates, within the selected virtual speaker layout. Users can either import existing virtual speaker setups or create a new one. This object-based audio approach has the advantage of dissociating sound source's localisation from the loudspeakers position and allows pre-spatialized music to be portable from one dome to another, regardless of the amount of loudspeakers and their relative position.

3.2 Summary and future developments

The combined usage of SpatGRIS and ServerGRIS offers an accessible and integrated solution for composers interested in spatial music with minimal interference with their working habits. The development of these tools is based on a collaboration with composers in order to adapt them to their workflow. A notable future addition is the implementation of a spatialization algorithm for asymmetrical loudspeakers configurations, similar to DBAP [5] with added "layers" of elevation, useful for sound installations and exhibits.

[2]SpatGRIS and ServerGRIS tools are available for download through the GRIS's website: http://gris.musique.umontreal.ca

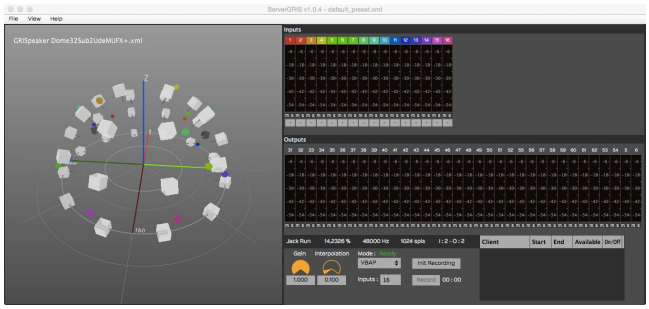

Figure 3: The ServerGRIS graphic user interface showing a virtual representation of the 32.2 loudspeakers dome setup used for the listening experiment.

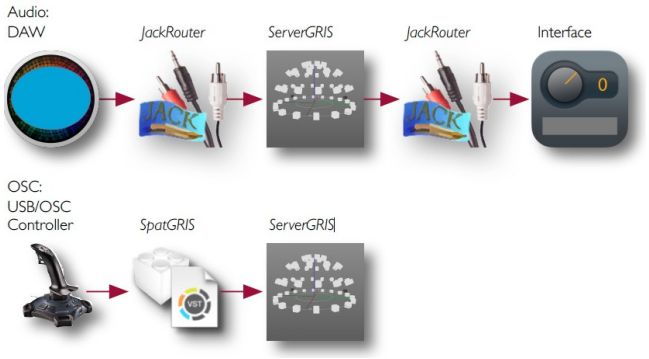

Figure 4: The system's architecture combines audio signals and OSC data, working in parallel.

4 RESULTS

4.1 Overview

The aesthetical description made by participants differed greatly between the two pieces. Made by different composers, having their own compositional strategies, this was to be expected. On the other hand, some listeners from both groups referred in some way to a sensation of immersion or noted a certain degree of spatial exploitation and movements.

4.2 Group 1 - *Ville Aux Cent Clochers*

Aesthetically speaking, G1 participants[3] commented on *Ville Aux Cent Clochers* mostly by referring to realism, or recognizable sounds and environments, in which they felt immersed, interlaced with more abstract, musical sounds. They used expressions like "in the

[3]In sections 4.2 and 4.3, group participants will be identified as "P", followed by a number (e.g. G1P1 = participant 1 from group 1)

middle of"(G1P1), "entered another world"(G1P1, G1P11), "inclusive character that plunges ourselves into the piece's universe"(G1P2), "we're attending to the festivities"(G1P3), "we're part of"(G1P5), "we are the centre of the events"(G1P7), "I lived"(G1P9), etc. Furthermore, other comments, like "the sound was going through me, stabbed me and hugged me", made by G1P10, indicate what Delalande calls an "empathic" reception behaviour: "the listener who adopts this attitude is mindful of the [physiological] sensations produced by the sound in him", "like if they were themselves subjected to those [sound] movements" [4]. A common thread in these comments indicate a sense of presence: "a state of consciousness, the (psychological) sense of being in the virtual environment" [9] or what Delalande simply calls "immersive listening" [4]. For these participants, it seems evident that at some point their sense of immersion went beyond the device's physical environment – the perceptual immersion – and passed the threshold of the mental environment, towards a sense of presence.

4.3 Group 2 - *La part des anges*

Comments from G2 on *La part des anges* mostly referred to musical parameters, describing sonic weavings; evolving textures and timbres; perceived elapsed time, distances, space size, spatial movements and trajectories; sound duration, motion and redundancy; punctuations, etc. Most of the listeners described the perceived form, sound materials and musical functions, while only two of them used metaphoric expressions like "undergroud celestial place" and "angels' voices" (G2P1); "the origin of the world" and "transported into an interstellar world" (G2P6). Except for G2P1, G2's common listening behaviour was clearly "taxonomic", which adresses "the listener's desire to get an overall synoptic view of the piece" by segregating "large morphological units" and by memorizing, qualifying and establishing relationships between them, possibly employing descriptive metaphors to do so [4]. Nevertheless, all of them considered spatialization a noticable aspect of the music.

5 DISCUSSION

Aside from the pieces' aesthetical differences, an explanation to why G2's general listening behaviour differed from G1's may be due to the participants' pre-existing knowledge of sound typo-morphology. As Alcazar noted in his experiment's conclusions, electroacoustic music specialists greatly tend towards a taxonomic listening behaviour because they've been trained in the technical, formal and analytical aspects of this type of music [11]. Alternatively, both groups perceived some spatial movements in multiple directions and with distance variations. Assuming that all participants minimally felt, at some point, perceptually "immersed" in sound, such sensation would be similar to Slater and Wilbur's definition of immersion [9], resumed by Gordon Calleja as "the hardware's effect on the user's senses"[3]. But Calleja also advises against a monolithic approach to immersion, stressing that a distinction must be made between "immersion" and "presence" [3]. Our observations of recurring expressions themes and listening approaches specific to each group tend to show that such a distinction is necessary if we want to better understand the phenomenon of sound immersion when listening to 3D-spatialized music.

6 CONCLUSIONS

Delalande believes that immersive listening generates "a certain abandon, music perceived as a surrounding environment, like a sensory bath" and where "the whole music would be perceived as external, in contact with the [listener's] body by the skin, by the senses" [4]. But he made no observation of such listening behaviour in his experiments, as no 3D-spatialized music or device were involved. Furthermore, our results seems to indicate at least two levels of immersive listening: thefi rst being external, rendered possible by the technological device, the second being internal, constructed by the subject's intentionality in response to the perceived audio content (the object). Precisely assessing and qualifying a listener's sense of immersion requires a more comprehensive survey. Investigating the listener's point of view by employing a funnel questioning approach might provide valuable data for audio content creators to further understand affective qualities of sound immersion. The next step of this research will have individual listeners describe their experience of both spatialized musics presented above,first in writing, which will then feed probe questions asked during a follow-up interview. Our goal is to validate the distinction between "perceptual immersion" and a "sense of presence" while also extracting more detailed qualitative data in an attempt to conceptualize immersive listening behaviours, which may serve as the basis for new analytical methods, adapted to 3D-spatialized music.

ACKNOWLEDGMENTS

This work is supported by the Social Sciences and Humanities Research of Canada (SSHRC), under grant number 435-2016-1080, by the Fond de Recherche du Québec - Société et Culture (FRQSC), under grant number 2017-RC-198003 and by the Canada Foundation for Innovation, under grant number 33558.

REFERENCES

[1] Antonio Alcazar. 2015. L'écoute attentive comme base pour l'analyse des musiques électroacoustiques. French. Trans. from the Spanish by François Delalande. *Musimédiane*, 8, (June 2015). http://www.musimediane.com/numero8/ALCAZAR/_fr/index.html.

[2] Malika Auvray, Charles Lenay, J Kevin O'Regan, and Julien Lefèvre. 2005. Suppléance perceptive, immersion et informations proprioceptives. *Arob@se*, 1, (Dec. 2005), 94–113.

[3] Gordon Calleja. 2011. *In-Game: From Immersion to Incorporation*. The MIT Press, Cambridge, Massachusetts. 224 pp. ISBN: 978-0-262-01546-2. https://is.muni.cz/el/1421/podzim2016/IM082/Gordon_Calleja-In-Game__From_Immersion_to_Incorporation_-MIT_Press__2011_.pdf.

[4] François Delalande. 2013. *Analyser la musique: pourquoi, comment?* •, (Ed.) INA. 244 pp. ISBN: 978-2-86938-201-5.

[5] Trond Lossius, Pascal Baltazar, and Théo de la Hogue. 2009. Dbap - distance-based amplitude panning. *Proceedings of the International Computer Music Conference (ICMC)*. http://hdl.handle.net/2027/spo.bbp2372.2009.111.

[6] Robert Normandeau. 2014. La part des anges (2011-12). https://www.electrocd.com/en/oeuvre/32391/Robert_Normandeau/La_part_des_anges.

[7] Pierre Paillé. 1994. L'analyse par théorisation ancrée. *Cahiers de recherche sociologique*, 23, 147–181.

[8] Ville Pulkki. 1997. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*. • 45, 6.

[9] Mel Slater and Sylvia Wilbur. 1997. A framework for immersive virtual environmentsfi ve: speculations on the role of presence in virtual environments. *Presence: Teleoperators and virtual environments*, 6, 6, (Dec. 1997), 603–616.

[10] Bruno Trentini. 2014. *Lesfi gures de l'immersion*. Bernard Guelton, (Ed.) Presses Universitaires de Rennes, Rennes. Chap. Pour une immersion non transparente, 25–59. ISBN: 978-2-7535-3401-8.

[11] Nicolas Tsingos. 2018. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*. Agnieska Roginska and Paul Geluso, (Eds.) Routledge, New York and Londres. Chap. Object-Based Audio, 244–275. ISBN: 978-1-138-90000-4.