Experimental Sound Design for Wave Field Synthesis: Employing Low-Quality Voice Recordings in Spatial Audio For Art Installation

Mariana Bahia

Sound Studies (Master of Arts)
Universität der Künste, Berlin, Germany.
kontakt@marianabahia.com

Abstract

This paper explores the intersection between high-end technology for immersive media and conceptual sound installations using low-quality audios. It discusses the concept of fidelity in sound reproduction and traces how this concept relates to the subject of identity and self through the recorded voice. I analyse art projects that use low-quality audio to address issues of self and how these projects can help presenting an alternative understanding of experimental sound design for spatial audio. Finally, I explore the issue through my own artistic research called This Alienness and Me, where I investigate the feeling of "not belonging" through personal voice recordings made with a cellphone. In my project the audios were juxtaposed using Wave Field Synthesis to create an immersive environment. Key-concepts: Spatial Audio, Experimental Sound Design, Low-Quality Audio, Wave Field Synthesis, Fidelity, Voice, Identity.

1. Introduction

Rendering virtual spatial sound data through Wave Field Synthesis (WFS)[1] requires high-end technology: audio processor, excellent speakers, adequate software, optical cables, and so on. Consequently, one could associate the use of WFS with high quality and uncompressed audio. However, scientific studies already show that compressed audio at higher and lower bitrates can be use combined

---

[1] For and introduction on Wave Field Synthesis see: Brandenburg, K., Brix, S., Sporer, T. (2004). Wave Field Synthesis: From research to applications. In 1*2th European Signal Processing Conference*. Ilmenau. Germany.

with WFS for different purposes. MPEG-4 BIFS (BInary Format for Scene description) encoder tools, for instance, can be used together with other tools for audio production to allow a transition from channel-oriented to object-oriented coding. (Plogsties, Baum & Grill, Conveying Spatial Sound Using MPEG-4, 2003) In the project CARROUSO (which stands for creating, assessing and rendering in real time of high quality audio-visual environments in MPEG-4 context) the codec was used with WFS to allow the transmission of a sound field to a remote space. (Brix, Sporer & Plogsties, CARROUSO – An European Approach to 3D Audio, 2001)

Concerning speech, different studies suggest that WFS can be applied into teleconferencing for improvement in sound quality. The positioning of sound sources into a space can favour the accurate localisation and identification of different speakers in teleconferencing with remote participants. Furthermore, this characteristic can considerably improve speech's intelligibility under "cocktail party" conditions. (Boone & de Bruijn, Improving Speech Intelligibility in Teleconferencing by using Wave Field Synthesis, 2003)

Although those examples above are mostly directed to commercial applications, WFS system's flexibility can also provide new paths towards creative freedom. Compressed audio files can be used not just as a matter of practicality, but also as a conceptual choice in creating an artistic work. On top of that, degraded audio quality and/or speech's intelligibility can be purposely employed as an aesthetic choice, providing very interesting results in the context of spatial audio and, more specifically, WFS.

In this paper I will address theories that deal with listening practices intrinsically linked to technology. I will focus primarily on Jonathan Sterne's book The Audible Past, more specifically on his perspective on the matter of fidelity in sound reproduction. (The Social Genesis of Sound Fidelity, 2003) I will also consider Lydia French's Chican@[2] Literature of Differential Listening, (2014) in which she proposes the acknowledgement of different listening practices and methodologies to trace our auditory past. Those theories can help bringing into discussion the different relationships between people and technology, and how these relationships can affect our own notions of self. Furthermore, they can help to understand the connection between fidelity and different cultural backgrounds. In this sense, my objective is to reflect upon how the technical choices can be intrinsically related to the conceptualisation of an art work.

Further, I will analyse two art works that explore the juxtaposition and repetition of voice recordings to yeld a recontextualisation of self, identity and space. The works are: Objects for Aggressive Sonic Occupation, by Pedro Oliveira and Luiza Prado (2014) and Balancing Act, by Amble Skuse.[3] While those works mainly explore the relationship between female discourse and noise, they do it by dealing with sound quality in a conceptual level. Through the reproduction and repetition of juxtaposed sound, the technical quality of the audio is slowly compromised and the speech becomes unintelligible within time. Discourse is then resignified and space is gradually taken over in different ways.

Finally, I will address my work This Alienness and Me, (This Alienness and Me, An Artistic Research on the Feeling of Not Belonging, 2017) where I used my personal voice recordings made with a cellphone allied to WFS to create an immersive installation. During the mixing of my audio, I played with the "cocktail party" effect by juxtaposing 24 voice recordings into the WFS. By playing with volume and movement automations, I promoted constant shifts between intelligibility and unintelligibility of speech within the compositional time. In this sense, I used my cellphone recorded

---

[2] The use of the @ in the end of the word 'chican@' is to avoid the need to write chicana/o, or chicana and chicano (feminine and masculine declanations of the word in Spanish). (French, L., "Chican@ Literature of Differential Listening, 2014) 'Because Chican@ studies encompasses the creative spirit, work, and history of Chicana women, Chicano men, and Chican@ gender queers, the @ more appropriately renders that inclusivity than the dividing line of the slash.' (French, L., Lydia French, 2013)

[3] Skuse, A. (2017). *Balancing Act*. Performance. Sonic Cyberfeminisms. Lincoln. U.K.

audio and the possibilities to play with intelligibility to encourage the visitors' interactivity in the installation and stimulate their constant change of perspective while inside.

2. The Matter of Fidelity in Sound Reproduction

Regarding hi-quality in sound reproduction, Jonathan Sterne argues that fidelity is more related to social processes, than to high-end technology. (The Social Genesis of Sound Fidelity, 2003) Analysing two early 20th century Victor Records advertisements, he argues that the ads were providing the readers with 'philosophical claims', by daring them to distinguish the difference between listening to opera singers live and their voices recorded. (2003, p. 215) One of the ads showed a picture of an opera singer on one side, and a picture of a gramophone on the other side. Between the two pictures, the question "which is which?" appears followed by the text: "You think you can tell the difference between hearing grand-opera artists sing and hearing their beautiful voices on the Victor. But can you?". (2003, p. 217)

> *"Everywhere we turn in the search for true fidelity, the desire to capture the world and reproduce it 'as it really is' yields a theory of correspondence between representation and that which is represented. [This implies] a potential or real loss of being in the original sound."* (2003, p. 218)

Sterne, on the other hand, maintains that reproduction (or even the possibility of it) already precedes the original. That means, if there is no (possibility of) reproduction, then there is no original. He continues:

> *"Sound fidelity is a story that we tell ourselves to staple separate pieces of sonic reality together. The efficacy of sound reproduction as a technology or as a cultural practice is not in its keeping faith with a world wholly external to itself. On the contrary, sound reproduction ... always implied social relations among people, machines, practices, and sounds. The very concept of sound fidelity is a result of this conceptual and practical labour."* (2003, p. 219)

Lydia French (Chican@ Literature of Differential Listening, 2014) argues that Sterne seems to "express a certain amount of faith in [the] realism" of sound reproduction technology depicted in visual and literary media. By doing so, he makes an unilateral "correlation between the ads teaching and actual listening practices." (2014, p. 8) In French's view, this is a contradiction with Sterne's own perspective about sound fidelity as something intrinsically connected to social aspects, and not a one-to-one correlation between original and copy. In her words:

> *"Sterne traces that realism in the manuals and advertisements that demonstrated how to use emergent sonic reproduction technologies. Though his treatment of media, including the visual and literary media of advertisements and manuals, is always positioned within the understanding that experience itself is always already 'intensely social' ironically, Sterne's nuanced treatment of listening for fidelity itself privileges the representations of reality and experience found within the ads and manuals." (The Social Genesis of Sound Fidelity, 2003) (Chican@ Literature of Differential Listening, 2014, p. 8)*

For French, this problem is more evident when Sterne analyses the presence of women on those ads. (2014, p.8) While Sterne (The Social Genesis of Sound Fidelity, 2003) argues that the ads were accurately depicting the participation of women in mass culture, French calls attention to the fact that those ads were showing only white women. "By [Sterne's] logic, the relative absence of [non-white]

women or men … would suggest that they were not fully participant in the growing mass media networks." (Chican@ Literature of Differential Listening, 2014, p.8)

Additionally, French (2014) proposes an alternative understanding of what can be constituted as history and, in consequence, what can be traced as real. She declares that considering other listening practices, such as differential listening, is essential to the comprehension of our auditory culture. She calls differential listening:

> *"Creative and motivating listening practices that account for sound as both interior and exterior, a product of subjective consciousness as well as of material media and technologies … Both the subjective interior form of listening … as well as the audile techniques learned from various sound reproduction technologies are socially imbricated. Neither 'the individual' nor the media are mutually exclusive categories in this understanding, but more importantly, neither are independent of their culturally- and historically specific situations."* (2014, p. 2)

Sterne's view on the matter of sound reproduction as a social process, (The Social Genesis of Sound Fidelity, 2003) combined with French's perspective (Chican@ Literature of Differential Listening, 2014) can provide an interesting conceptual base to understand how fidelity can be considered, while approaching WFS through a conceptual perspective. WFS can also be taken as an instrument, or a medium, in which relationships between people and technology can be in constant change. In this sense, WFS can be repurposed from being just a high-end technological tool into an aesthetic choice for developing an artistic work, in which the relationships between the individuals and their cultural and technological backgrounds become more important.

3. Recontextualization of Self Through Sound Quality Degradation

Critical designers Luiza Prado and Pedro Oliveira (Objects For Aggressive Sonic Occupation, 2014) created a sound object meant for women to speak up when ignored by men during a conversation. According to them, the object is a tool for self-affirmation against misogyny. It consists of a flower tiara with embedded microphone and speakers. According to Oliveira and Prado, when a man talks over a woman, she can take the object out of her head, repeat what she said on the microphone and "the taser repeats it over and over again, increasing pitch and volume and reclaiming back the sonic space that was taken from [her]" (2014). Within the overlapped and aggressive repetitions of the woman's voice, the self is recontextualised and acoustic space is taken back. Speech's intelligibility is gradually destroyed and transformed into loud and hi-pitched noises that overtake her surroundings. Within the repetitive juxtaposition of her voice, the woman overcomes her acoustical invisibility and resignifies her sonic space. More than dealing with a physical space, the taser can be understood as a tool to reclaim a place in a conversation dominated by men, as well as a social place for her voice and herself.

Conversely, Amble Skuse, in her work Balancing Act[4], performs by superposing her own voice with an electronic generated voice, while presenting 1464 names of female composers. The artist argues that female composers are usually taken as non-existent throughout composition history. Skuse explores this notion through concepts of noise and invisibility. In her words: "Noise, our voices become noise, our names become noise, inaudible, unheard, filtered out."[5] By juxtaposing all those names through sound reproduction and repetition, she reconfigures the notions of what is noise into reclaiming acoustic

---

[4] ibid.

[5] ibid. This quote was extracted from Skuse's abstract at the Sonic Cyberfeminsms printed program.

space, and giving those composers a place on stage. More than that, she is also raising awareness of their actual existence in composition history.

In both cases, significance through the spoken word is lost within time, and is remodelled by the juxtaposition of the mechanically reproduced voice. In both works, acoustic space is reclaimed. In Oliveira and Prado's, the woman recontextualises and reassures her own existence through the new versions of her speech. Using the mobile tiara made by Oliveira and Prado, a woman overcomes the imposed silence and reclaim auditory space, while recontextualising her speech and reaffirming her identity. In Skuse's performance, female composers' names form a mass of unintelligible speech within their superposition. This mass of sound takes over not only the stage, but the whole concert hall space. In both cases acoustic space is also recontextualised: in Oliveira and Prado's work, it is gradually redeemed through aggressive repetitions, while in Skuse's it is reclaimed within the juxtaposed addition of new names.

4. This Alienness and Me: An Artistic Research on the Feeling of Not Belonging

4.1. Inspiration and theoretical process

My work called This Alienness and Me (2017) was an artistic research that begun in January, 2016 and ended in February, 2017. This project was developed during the Masters in Sound Studies at the University of the Arts (Universität der Künste - UDK) in Berlin and originated, besides my Masters' dissertation, an installation that was exhibited at UDK's WFS Studio through February, 2017.

The inspiration for this work was a feeling. After moving from Rio de Janeiro, Brazil, to Berlin, Germany, I started to feel alone and disconnected, like I was watching life passing by through a window: I was there, but I was outside. While facing my feeling of disconnection, I found comfort in technology. I started to exchange voice messages with a close friend that was also living abroad. Talking on the cellphone in my mother-tongue while wandering through the streets of Berlin helped me feel I was no longer an outsider. I was somehow functioning in that society. The streets were mine, as I walked through them. I was re-appropriating that place and contributing to its soundscape, by bringing my private discourse into the public space.

During this process, I noticed that I enjoyed playing back my recordings. I could listen to what I just said through a new perspective, and this helped me reorganise my thoughts. I started to experiment with that, but without sending my recorded messages to anyone. In other words, I decided to talk to myself through the phone, and listen back to what I just said. This turned out to function as a fun (and reflective) practice, that I developed spontaneously to deal with my feeling.

Later, I decided to call that feeling "this alienness" and take it as the subject matter of my research, my own sense of self, while feeling I didn't belong either to a place, to a situation, to a group of people, or all of those at the same time. In other words, I researched my own sense of self while I felt "this alienness". I decided to explore this practice of "talking to myself on the phone", which I later called "mobile voice recordings" as part of my methodology. With the cellphone, I would record myself describing my "feeling of alienness" while in different environments and situations (e.g. inside a building I have never been before, walking down the street, or even in my room on my bed). For each situation, which I called "case of alienness", the main question was: "How does this feeling emerge to me?" I investigated six cases of alienness and for each case I recorded four voice descriptions using my cellphone. Later on I would explore those recordings to create my immersive environment using Wave Field Synthesis.

4.2. The Installation

In the installation, visitors would find 96 speakers disposed in a circle. Inside this circle, one chair was positioned in the middle of the room while seven other chairs were positioned peripherally, facing the chair in the middle. On the central chair, the visitors would find glasses with a tag that said "Try Me". The glasses had mirrors that pointed to the floor. The room was dark and colourful lights were spread out around the floor following the speakers' circle. Everyone was welcome to either sit on the chair in the middle and put on the glasses, walk around the room, or sit on one of the peripheral chairs. Each choice brought a particular perspective and new visual and auditorial approaches to the environment. If the visitors chose to sit on the chair and put on the glasses, they could see themselves sited, while the lights would make interesting reflections back to their eyes. If they decided to walk around and/or sit on one the peripheral chairs, then according to their own perception, awareness and movement inside the room, they would be able to have their personalised "mix" of the sounds coming from different sources.

Interaction inside the environment was developed by personal choices inside the installation. Each person could choose how and what to hear, either by focusing more on sound "colours", or by trying to follow what was being said in each recording. The number of visitors allowed inside was not limited and they could come in and out as they wanted.

4.3. Mixing This Alienness and Me in WFS

For the WFS, I composed a 29-minute piece that was played in loop. The piece's first moment was composed of six pairs containing the two first recordings for each case of alienness. Those twelve recordings were played simultaneously, each pair from a specific place in the room. Most of those recordings contained both speech and background sounds, which were recorded in different environments and situations. With WFS, I could organise each case of alienness spatially. For this part of the composition, each pair of recordings was positioned in a fixed place.

At first, my focus during the process of mixing was on quality of speech. Middle and higher frequencies were raised, while lower frequencies were either attenuated or cut, depending on each track. I aimed at giving the impression that the visitors should be free to move around the room and listen to my personal recordings, while at the same time, purposely guiding the visitors into what I wanted them to hear or not. To achieve that, I played manually with volume automations, by raising the volume of moments where I wanted the spoken words to be understood, while lowering the volume of other moments. This automation was applied in all the twelve tracks. In this sense, one could perceive and follow a line of thought through the significance of my spoken words emerging from different points - or sound sources - in the room. This effect had better results from the central perspective. To choose the chair in the middle, therefore, would mean to grasp words or phrases from different spatial sources and to be able to put the pieces together and understand the "message".

Conversely, it was important to have moments in the work where the juxtaposition of my recordings would be auditorialy overwhelming and exhausting from all the positions inside the room. In this sense, I purposely affected the intelligibility of my speech, by leaving some of the background of the different recordings along the piece's timeline. At those moments the background noises would deliberately overlap the speech and blur the understanding of what was being said. In this sense, the visitors could stay in the centre of the room, in order to have a broaden impression of what was being told, by combining the spoken elements coming from different directions. On the other hand, they could also choose to follow only one discourse placed in a specific localisation point in the room. This, however,

6

would be a more exhaustive task. Even if they made an effort to understand just one point clearly, this point could suddenly become unintelligible, as another sound could take over the space on the macro level and blur that specific discourse they were trying to follow.

On a second and third moments of the piece, a third set of voice recordings for each case of alienness was used. On those, I read written descriptions of my feeling of alienness, while recording at home from inside my room, also with the cellphone. Some of those recordings were still fixed at some point inside the studio room and overlapped each other. The other recordings weren't fixed at one spatial point. I changed their localisation manually so they would move circularly around the room. I felt that the circular manual movement provided a better way of transmitting the message, without turning the visitors' attention too much towards technology and the equipment.

At this moment, significance was most of the time very clear and the visitors would have an auditorial stimuli mostly from a macro level only. This means that now the tracks being played were not so independent from each other, as they were on the first part of the composition. On the other hand, two or more speeches could still sometimes overlap each other within the circular movements. This could work as constant incentive for visitors to change their auditorial perception. During their experience, visitors could, therefore, shift their auditorial perception to focus on the spoken words, and follow one or more lines of thought. On the other hand, they could also choose to listen to the environment through a macro perspective, by focusing more on the sound of the speeches and the new sounds that emerged from their juxtaposition.

5. Conclusion

Working with the WFS provided me a way to create an environment that was not analog to any other place, but with qualities of an imaginary place (or more), that was in constant change according to the movement, awareness and perception of each visitor. The shift between intelligibility and unintelligibility of the discourses played in the room was not a matter of simply mixing choices. More than that, through the mixing I could provide visitors with a space where they would have some freedom to walk around, get in or outside, change their modes of listening, interact with others.

I also aimed at creating an installation where each person could be aware that their choices inside the room would matter. The place would be in constant transformation through each person's connection with it. In this sense, the installation could become a place where I could also talk about the other. The visitors, therefore, would also be responsible for the images of themselves, that would emerge to them while they were inside that space.

I used the WFS as a medium to work on my mobile voice recordings, place them inside the studio room and play with words coming from different directions. The system functioned as a tool, that helped me approach my subject matter. During my mix, I could also place myself in the middle of the room and perceive how my sense of self emerged to me with the sound of my voice juxtaposed and expanded - into many voices. In the installation, the noises of my alienness were resignified, through the juxtaposition of my mobile voice recordings and their recontextualisation into the soundscape of the environment I created. I hope I could provide an interesting perspective concerning the use of Wave Field Synthesis conceptually and encourage other works in this direction.

Acknowledgments

References

Bahia, M. (2017). *This Alienness and Me, An Artistic Research on the Feeling of Not Belonging*. Master thesis. Sound Studies. University of the Arts. Berlin. Germany.

Boone, M., de Bruijn, W. (2003). Improving Speech Intelligibility in Teleconferencing by using Wave Field Synthesis. In *114th AES Convention*. Preprint 5800.

Brix, S., Sporer, T., Plogsties, J. (2001). CARROUSO – An European Approach to 3D Audio. In *110th AES Convention*. Preprint 2801.

French, L. (2014). Chican@ Literature of Differential Listening. In *Interference. no. 4* [Online]. Retrieved from http://www.interferencejournal.com/articles/sound-methods/chican-literature-of-differential-listening (accessed January 4, 2017).

French, L. (2013). Lydia French. In *Central Thoughts* [Online]. Retrieved from https://centralthoughtsjournal.-wordpress.com/fall-2013/prose/lydia-french/ (accessed February 28, 2017).

Oliveira, P., Prado, L. (2014). Objects For Aggressive Sonic Occupation. *A PAREDE* ツ *hello[at]*apare.de [Online]. Retrieved from http://a-pare.de/2014/objects-for-aggressive-sonic-occupation (accessed February 28, 2017).

Plogsties, J., Baum, O., Grill, B. (2003). Conveying Spatial Sound Using MPEG-4. In *AES 24th International Conference: Multichannel Audio.* The New Reality. Banff. Canada.

Sterne, J. (2003). The Social Genesis of Sound Fidelity. In *The Audible Past: Cultural Origins of Sound Reproduction (pp. 215-286).* London: Duke University Press.