Surround/Body

- a complimentary text to the Sound/Body

by
Petar Mrdjen

Institutions Name: Stockholm University of The Arts
Level of Degree: Master’s degree
hp 15 hp
Program: The master’s program in film sound design
Term and year: VT 2023
Supervisor: Ester Martin Bergsmark, Carolina Jinde
Examiner: Erik Gandini
# Table of Contents

intro 3
the bitter sweet spot 4
spatial distortion in narrative cinema 7
surround-scapes as resistance 11
  - empathic room 11
  - voyeuristic room 12
  - composite room 12
speaking over the sound 14
the not-yet-seen and sound-images 16
bibliography 18
Intro

I exhale. In the corner is a pile of cardboard boxes that I had just opened up. Stretched across the floor are many carefully measured lines of white duct tape. On the computer are countless browser tabs - from waves articles to small forums - telling me how to set up my new speakers. Equidistant triangles and other exact angles. All that is left to do is calibrate the volume so that each speaker reaches me equally. You see, I had decided to jump into surround sound. The thought of having a little personal cinema in my apartment to play with was not possible to pass on. I remember then sitting down in front of my digital audio workstation, with speakers surrounding me, and thinking…. Now what?

I had no idea what to do with it, what sound was supposed to go where, whether it should move or be still. I experimented with many surround things, never reaching a definite answer. I would go to the cinema to listen to other films - for research - but most of the time the sound never filled the room or left the screen, it is the spatial element of which I was mostly keen. What a waste.

In the pursuit of discovering what surround means to me, the intimacy it can bring to a room; why mainstream films use it shyly, or don’t use it at all.

I discovered a rabbit hole.
The bitter Sweet-Spot/

I got curious about moving one acoustic space to another. On the short film Järnidän (Dahlström, 2023) the opportunity presented itself. Most of the scenes took place during a rehearsal for a theatrical play, and it was shot in a genuine theatre. Since the location was aurally authentic to the space within the film I find it pertinent to capture it. I had scouted the location ahead of time, recording myself walking and talking at different positions in the room. I realised that it sounded incredibly strange. There were resonances that bounced around frantically in the space and positions where sounds collected awkwardly, it was not fit for recording clean dialogue. Either I would have to consider dubbing the film, or embrace the space.

With that in mind I went to the theatre a few days early and set-up microphones in a surround rig of sorts. One microphone at every corner of the room, at the positions that uniquely identified the acoustic character of the space, as well as a “center” mic at the edge of the stage. This setup was timecode synced to the production sound rig, but not monitored. The sound technician made sure to press record on both devices but the resulting surround recording remained a surprise until post-production.

The recording was pretty "dirty" with noises and hums of the room, the film team whispering and moving about, as well the pretty strange acoustics of the space itself. While editing this on near-field speakers my sound colleague commented on the sonic aesthetics of this experiment, noting it sounds "wrong for a film", because the feeling of "being there" was heightened to a degree that felt documentary instead of fiction. The theatre was so dominant that it created a distance to the characters on screen. However when it was transferred to the surround speakers of a cinema you could feel the whole environment in its liveliness. Since the theatre and cinema were quite close in size it became a near 1:1 transplant. Depending on where you sat you would experience a very different and unique character of the space and the characters within. In the left-rear-surround speaker low frequencies would pool up and footsteps would hit the listener like waves, this one was placed on one of those awkward corners. Under the front-right speaker you could hear whisperings of the director sitting by the video village nearby, commenting on the images, “oh wow, that looks so good”, “yes! We have it!”. Behind her there was a server system humming away. Truly just like “being there”.

Customarily the sound engineer works with the space to eliminate such noises and reflections. Turning off any sound making machines that you can, dampening the reverberations with sound blankets and similar objects; perhaps putting up baffles between the actors and crew, or moving the crew out of sound-sight. The purpose of this is to secure the cleanest possible recording of the actors dialogue. I ignored those customs and asked my sound team to do the same, by rigging up the whole theatre with microphones I captured all sounds - even the potentially unwanted.

Since the film was shot with two cameras in a documentary run-and-gun approach, and then edited with a disjointed dogma aesthetic, the visual experience felt fragmented and frantic. In contrast a 1:1 transplant of rooms aided in an extremely stable spatial experience. One that lived its own existence within the cinema space, that the audience personally resided within - and was aurally unique to their position.

"There are 'good' seats and 'bad' seats in the cinema." Alan Williams, 1980

In 2020 my partner got diagnosed with a brain tumour, at the time of its discovery the tumour had already decayed his peripheral vision, taking his ability to perceive the left field of vision. Now when
we go to the cinema we sit on the right side of the hall so that he can sit at an angle and see more of the screen. A sonically compromised spot underneath the right side speaker. While cinemas are designed acoustically to minimise the amount of seats with a compromised aural experience, a challenge that grows considerably more difficult with the addition of more speakers. For instance, in a mono soundtrack there is only one speaker, centered behind the screen, whilst you can sit at different angles and distances to it, the sound-image will not skew considerably. However Mark Kerins notes in his book *Beyond Dolby (stereo)* that “as the channel count goes up in a theatre, the aural experience becomes heavily dependent on the point-of-audition.” (Kerins, 2011, p280).

The sound-maker crafts the soundtrack in (ideally) acoustically treated rooms, sitting equidistant from all speakers; which have been calibrated to perform an equal amount of sound pressure from each direction, this is the “sweet spot”. The sound-maker works from a bodily experience of intuition and reaction, they make decisions and experiences from that single-point-audition - being as they cannot occupy two spatial positions at once. When their finished work is exhibited, how many audience members get to experience their art in the same way as the sound-maker did? Meta-spatiality (points of audition within a surround-scape) means that it is impossible for everyone to have the same aural experience.

Listening to films outside of the sweet-spot I wonder if you can even call it an equal experience. Besides the very few (perhaps at most 4 seats) in the very middle that share the soundmakers point of audition, most will experience skewed sound-images. With people at the boundaries receiving an abstraction, Watching BlackkKlansman (Lee, 2018) in the far right-rear seat was a psychedelic, surreal experience. The surround channels were used vividly for loop-groups, in chant and in protest, and they overpowered the speech from the screen. I had to piece together the main monologue from the reactions, the only thing I could hear. As if I was sitting among an unruly audience, it was sweet.

It got me thinking about people in the periphery, so far outside of the privileged sweet-spot that they are barely inside the “circle”. I can’t bend the rules of physics and give everyone access to the sweet-spot, nor do I think going back to a less spatially critical sound technology (read: mono) is the way to go. Instead I wonder about a different approach to surround-scaping, one where the whole hall is active, where everyone experiences ‘their own sweet spot’. This idea is not uncommon in sound-installations. In 2018 I got to experience Digital Unrealities (Prague, Sara Pinheiro, etc.). A sound installation that worked with speaker setups in a room to eliminate any one sweet spot. People explored the space until they found their own spot, and just listened. The whole room in togetherness, yet uniquely personally, experienced the sound.

Perhaps it is not appropriate to ask a filmgoing audience to actively swap seats, or move around in a dark room. Nonetheless it has been my goal to approach surround with a very active spatial soundscape. One that the audience could participate in, ideally dense and rich to the point that the listener would find their own sounds to focus on, and follow. In the short film Sonntag (Cabaco, 2019) rather than curating a specific single-point perspective I would distort the space with layers upon layers with seemingly contradictory “angles” and textures. I took the power back from the image and gave it to the listener. Where I sit and you sit in the cinema is a different experience, not just ‘good’ or ‘bad’.

It should be said that for most of my sound-making career I’ve worked in stereo. Accessing a studio on low/no budget films is out of the question, and collecting the equipment and speakers necessary to work in surround is expensive. When I finally bought 5 identical speakers and a sound card capable of streaming that many channels I was lost in how to utilise them. Sonntag (Cabaco, 2019) was a learning experience for me where I spent a lot of my artistic time googling how things are meant to be done, and stumbled upon a Waves article (2017) with clear guidelines and principles. Yet like a curious child I was questioning these at every opportunity.
Why should all dialogue be panned to the center speaker?

Why should I not pan sounds drastically in a 5.1 setup?

What are you supposed to place in the rear channels?

By challenging the sweet-spot I was questioning why a film must "sound like a film". At the time it seemed logical to hear the dialogue of an actor standing behind the camera from the back. When I tried to follow my instinct the result was odd, certainly not bad, also not what I expected. Partially because there still is a deeply rooted expectation of what sounds film-like.
Spatial Distortion in Narrative Cinema/

I have often felt confused by the use of surround in mainstream cinema. To me the placement of sounds are illogical and the spatial image distorted. Often it is conveys detachment, an out-of-body experience.

Consider the opening scene of Pulp Fiction (Tarantino, 1994), where two characters sit at a table facing each other. It starts with a "scene description" in the form of a two-shot. Visually a diner is established, we also see two characters, their spatial relationship to each other and to the environment that they are in. The surround sound sets a similar stage, with the dialogue front and center and the environment represented in a vague "stage left" and "stage right". In this sound equivalent of a "two shot" the dialogue is spatially detached from the characters, since they visually occupy the left and right portions of the frame, yet their voices project to us from in between them from the center channel.

When the editor moves the sequence along to a close up, then a mirrored close up of the other character (shot-reverse-shot) the audience jumps from shoulder to shoulder visually, yet the sound remain still. The surround-sound-image of vehicles passing by the window doesn't change, the other guests eating along in the background doesn't change, and the source music playing at this diner doesn't change. The picture simulates the experience of the audience's attention moving from one face to another. Sonically we never left the "establishing shot" and the sound experience likens the screen to a theatrical stage.

Michel Chion (audio-vision) calls this surround practice the "super-field". Where the sound is designed to be consistent and unchanging so that the camera can jump around freely without feeling disjointed. He has a theory that it works due to magnetisation. That is to say, when we watch someone move on the screen we associate the sound they create as coming from their visual position. The audience mentally "pans" the sound to the image. Even as they move off-screen their voice seem to follow them, in spite of our senses. One reason for this is that we have a more accurate sense of space and position visually than aurally. Our ability to pin-point sounds outside of our vision is approximate at best. It is also aided by the tremendous mental distraction of watching the light-play (read: image). Without the screen the sound (according to Chion) manifests the physicality of its mediator (the speaker).

Throughout the history of cinema there has been many variations on multichannel installations. Some experiments placed additional speakers in the (now empty) orchestra pit of the cinema, where music could play from a soundtrack instead of live musicians. Disney's Fantasound made a big impact at the time but was difficult and expensive to install. The majority of films remained monophonic for a long time.

By the 70's Dolby stereo was the technology that was the most widely adopted. It worked with the pre-existing optical sound strip of film by encoding a multichannel recording and mixing it down. In the playback stage the encoded track would be decoded and played back in surround. The availability of that multichannel technology meant that there was a surge of experimental surround-scapes in film. For a casual audience sound became an equal draw to experience the films perhaps for the first time since synchronised sound; perhaps for the first time as a sound experience and not a technological mediator of the voice.

It wasn't a painfree mixing process. Dolby stereo had many flaws, sounds could bleed into the wrong channels, and a heavy use of the surround speakers would skew the sound field. Yet it became the
standard because it was also the cheapest surround technology to adopt. It was backwards compatible with pre-existing setups (as in it would play in mono if the right equipment wasn't in use). To prevent technological mishaps Dolby would take on an active part in the production of Dolby Stereo soundtracks. Limiting the sound-maker from being too "out there" and playful with spatialisation (Kerins, 2011). Hiding the limitation of their product by controlling the art.

Sound-makers were creatively pushing the technology to its limit, and Dolby pushed back. Aesthetic principles of surround-scaping were born out of that tension. The center channel would bear the most responsibility for the auraity of films, Dialogue, on-screen sound effects, acousmêtre voices all reside in the center. Dolby Stereo didn't deal too well with transients in the rear channels, they would produce unreliable results in the encoding process. Which means that slow sounds with a weak attack like winds would yield the best results; atmosphere and other sounds without sharp transient. Principles that inform surround-sound to this day. If you inquired a sound-maker today about why even off-screen dialog is placed there you would most likely get the answer of "legibility" or "consistency". Some films follow this principle to absurd degrees. One early shot of the film Smile (2022) portrays a room in a wide shot with a door to the far right, barely fitting into the frame. One character walks through the door frame and says their line, which comes from the center channel. They then close the door and we hear the door slam on the Right channel. Both sound sources occupy the exact same spatial position in the diegesis, yet they are reproduced in completely different parts of the cinema.

Today surround technology is digital, it has a greater dynamic range, channels are discrete, and digital audio workstations have many tools to easily manipulate sounds in the surround-space. Yet these kinds of spatial paradoxes (distortions) remain in all-but-few modern films, and they are usually premeditated on (Dolby Stereo) principles regarding surround-aesthetics that we have come to expect and associate with a hollywood-style production (Højer, 2014).

For instance, non-diegetic music rarely occupies the center channel, instead it is pushed away from it to leave enough space there for other sounds and (mostly) dialogue. Atmospheric and environmental sounds are usually placed in the surround field to mask the acoustics of the cinema with a tailored interior or exterior feel. This is too a remnant from the time of pre-digital surround technology. The rear channel used to be of lower quality than the front channels, the frequency reproduction was lesser, so was dynamic range, and they were known to produce a lot of self noise. Those limitations did suit environmental ambiances quite well.

It is worth noting another reason why modern surround-scapes can feel restrained and center-heavy. With the spread of home-theatre systems films need to either be mixed twice for both theatrical and home formats, or be designed with down-mixing taken into account. The latter required less time, so it became the cheaper option. This means a severe choke-hold on the rear channels, because if the surround-scape is crucial to the film's narrative what happens to it when it's played back on a TV?

The consequence is that filmmakers can't use surround as a dramaturgical tool, or rather their films can't narratively rely on it. From immersive sound to surround, to stereo, to mono. In each step something is lost, but if you were to design a mono soundtrack with surround as some bonus "ear candy" then you would not lose the story in the down-mixing process. So what do you really gain to work in surround at all?

Today we have real multichannel sound reproduction as a standard in all cinemas, and with the advent of sound bars and mobile immersive sound (i.e. Dolby Atmos) this is fast becoming a standard in domestic life as well. Yet the status quo remains that surround is to be used conservatively, and usually when that notion is challenged it's met with the rhetoric that "sound shouldn't be noticed"; that
the illusion film creates, of flow and verisimilitude, would crumble like some fragile sand castle; if the audience became aware of the room that they are in.

My approach to surround-scaping - which is what I shall call the act of working with sound in multichannel media - is birthed from a blissful resistance to “cinematic” surround principles, and a naive playfulness. What I noticed early in my experiments is that the resolution of spatiality is denser by the screen. Panning a sound between the left-center-right speakers creates an impression of fluid and precise motion. If this sound continues further to the side and all the way to the back, the illusion falters. The space between the Left and Left Surround speaker is so vast that a panning motion appears like the sound disappearing from one angle and appearing from another, rather than a sound in motion. I had been thinking about working with 5.1 as an extension of working with stereo, that the illusion of 360 degree motion was its strength. I changed my tactic.

If I approached this setup as 5 individual mono channels the sound became fuller and more interesting to me. I think of each speaker as a unique voice, and that there can be a relationship between them. Considering minor variations in signal between two channels create a stereo image, it could also be a configuration of 10 potential stereo channels. Instead of panning I can create relationships between two neighbouring speakers, but even diagonally across the room. In music there’s a technique called double-tracking, where you record two takes of the same part, like a vocal line or rhythm guitar; then you pan them to opposite speakers and get a thicker sound that’s a little more fluid.

Working on Sonntag (Cabaco, 2019), the biggest challenge was collecting the amount of sounds required to fill each speaker, and deciding what sounds should go where. It had been a concern early in the process when I first read the script. Once I was on location recording the film I realised that there is a lot of sonic activity all around.

One scene took place in a backyard that was enclosed with hedges, you could not see beyond it. If you listen there’s airplanes, the dull hum of a nearby highway, a railroad crossing, thunderous freight trains, a small waterfall, a neighbour mowing their lawn, the busy town whose sound skipped across the lake, wind in the leaves, barley shakes, and a territorial bird singing - a sound engineers nightmare.

Except this time I needed sounds, and the environment was credible to the film’s narrative. So I quickly jotted down little maps in my notebook, and started localising the many sound sources. Like sonic floor plans. Then I recorded them separately so I could have material to edit with. There was not a single limitation in my mind to what sounds should be a part of the film. If I heard it, I recorded it.

All this material contributed to an enveloping sound experience where the soundscape is rich and intense, and real; simply sounds that are part of our daily normality. It became material I was seeking out instead of cutting out. The real world was offering so many ideas and sounds to me for free and I’m not rude enough to decline.

Then when it got to mixing I got an idea that the soundscape could represent the point-of-audition of the main character, a surround-scape that ignores the camera, in order for us to listen to the world through them. I interviewed the actor, asking them what was going through their mind for every scene. In my own way I became an actor too, putting myself in their position and surround-scaping intuitively.

What sounds draw my attention?

Does the environment/situation feel vast or small?

What is going through my mind right now?
Why aren’t all films made this way?

The space-making opportunity that surround-scaping provided me challenged my perception and listening. How do we listen to a three dimensional space? How does our body filter it? How can I use this technology to create experiences, and experiential stories?
Surround-Scapes as Resistance/

To look forward I had to look back on my work. I have always been lucky in the sense that I have usually worked with directors that believe in my artistic voice. All my surround-scapes have been experiments in a way because I was given space to follow my intuition.

When I work with the space, I work with the unseen, outside the physical frame of the screen. Surround-scapes give me the agency to take space (literally) and give new meaning to the image. I can contextualise how we perceive it, by expanding the world beyond it to different degrees; or perhaps disturb it with sounds that draw our attention beyond it.

While it is still in relation (or reaction) to the image, the contextualising power is akin to a bass player in a band; who with one note sets the root note of a chord, partially due to (and in lieu of) the low-frequency-weighted hierarchy of functional harmony. I can be that bass-player because sound’s power is entrusted to me, and usually it's important to go big. If the director can understand what the sound does to the image, they can start to understand what a responsibility it is to have, I encourage any sound engineer to dare to be that person; and to that person I share this chapter with you.

Within my work I have identified a few reappearing surround-aesthetics, three surround-scape strategies that I will refer to as rooms. There have been previous attempts to name approaches to surround aesthetics, but they have usually focused on spatial staticness or kinds of motion, a categorisation of technique (Jay Beck). My rooms are philosophical and political stances, ways of being with the image rather than ignoring it.

EMPATHIC ROOM

Is to follow the Characters point of audition, common to radio-plays. You disregard the position of the camera, if there is a sound object physically to the left of the character we would hear the sound coming from the left speaker(s). Aronofsky’s Mother! (2017) maintains this strategy for a majority of the film, giving us the impression of listening vicariously through the characters' perception.

The music producer Steve Albini is known for panning his microphones from the musicians perspective, easy to spot in his drum recording technique; we get to hear them from the drummer's seat. They appear "mirrored" to us because of the usual approach, which is to create a stage-like perspective with the panning.

In Sonntag (Cabaco, 2019), I establish this rule in the first scene by mirroring the position of the camera. We see the main character from ahead when water starts flooding the room from behind him. True to the characters position the sound of water begins behind the audience. Brutally shattering the camera’s perspective to highlight the experience of our character.
VOYEURISTIC ROOM

This room is usually the default for video games. The distribution of sounds in a multichannel soundtrack usually happens algorithmically depending on the position and perspective of the game camera. Kojima's Metal Gear Solid V (2015) is a third person action game where the surround-scape is anchored to the in-game camera's frame. Most of the time you see your character's full body and hear them fully. However when you crawl the camera frames out your character's legs. The foley sounds of their boots end up playing from the rear speakers. Splaying out the character's body across the whole surround-scape. Creating a spatial vacuum between the screen and the rear surround field, where the player sits.

In Roma (Cuarón, 2018), the voyeuristic room was explored with the careful placement of sound objects in relation to the camera. The position of sound sources are panned relative to the frame, with each shift in camera angles the sound follows, the soundscape rotates appropriately with each gentle pan. If a character is talking behind the camera, their voice is placed behind the audience. At first glance the surround-scape in Roma seems shackled to the picture. However it has the sneaky effect of tearing down the illusion of cinematic artifice.

Because abrupt changes in sound positions draws a tremendous amount of attention to itself it materialises the camera as an artefact. What I mean to say is that unlike the super-field that creates a stable soundscape for the camera to jump around in freely, the voyeururistic room heightens each camera angle (and each cut) to the point where it cannot be invisible. In a sense breaking the illusion of smooth audio to break the illusion of film itself, it brings down the images with it.

It is an “objectifying” experience.

As a counterpoint The Matrix (Wachowski, 1999) uses the voyeuristic room as a tool of stability. The famous Lobby shoot-out scene (where Neo and Trinity walks into the lobby of their enemy’s lair, loaded with weapons) used the surround-scape to orient the audience through every angle. Positions between the different characters and directions of bullets are placed according to the camera, it allows us to understand what's happening and to make sense of the visuals. Interestingly, the scene ends with its establishing shot, and starts with a series of close ups of boots, bags, necks, etc. The active surround-scape here allows the camera to be fragmented and impressionistic without feeling jarring.

COMPOSITE ROOM

One concern that has held back surround experimentation is the question of reproducibility. No two cinemas sound the same, the size and shape can differ, so can building materials and speaker systems. Not to mention the volume of bodies within the room differs, and the noise they may make. These variables means that the sweet-spot designed film will eventually have to compromise at some point. Yet if the listening experience - in a way - can’t be monitored, what compromises could there be? How do you even mix something you cannot fully hear?

The previous two rooms are examples of a “single-point audition”, whether it's the character’s ears or the camera's “ears” we are just experiencing one perspective, essentially a representation of the
human listening experience. The anti-thesis to such perspectives is the “multi-point audition” where we are presented with multiple listening perspectives at once.

Multi-point audition resists being defined as one perspective. It is anti sweet-spot. When I transplanted one room into another (Järnorf, 2023), it didn’t eliminate the new space, it became super-imposed. It also became impossible to experience in one go. Every seat brings a new perspective and this experience is the composite room.

Working with this my approach is to get as many different spatial impressions that I can; through myself (sitting in different seats) and friends who can experience different things at once. It is fruitless to aim for (too much) intention in spatial specificity. Instead it is an invitation to reconsider the dramaturgical possibilities of surround-scaping; as well what it means to design for a whole room.

What does the periphery sound like?
Speaking over the Sound/

“There are voices, and then everything else.” Michel Chion on vococentrism,
Audio-Vision

Within the soundtrack there are hierarchies, mainly the dichotomy of spoken text and everything else that is deemed less crucial. ‘Text’ refers to the written word mediated through the actor's voice. Since mixing is a practice of subtracting, usually something has to go. Because films have the ability to over-represent themselves - with sound, image, and text - there is a ranked list of priorities. This list is not a written law, rather one that is understood among sound engineers around the world. The legibility of dialogue is ranked on top (do we understand what’s being said?), following that is non-diegetic music cues, then sound. When too many layers are being forced to project at the same time the sound mixer has to apply many different tricks to maintain this sonic hierarchy.

With Interstellar (2014) Christopher Nolan set out to prioritize music and sound effects far higher than most dialogue. Going against the unspoken hierarchy, At release the sound mix was met with backlash, audiences were complaining that you couldn’t understand the actors, and assumed that this was a bad thing. Due to the film being screened with subtitles outside of the US-UK market, the level of upset was much smaller. By turning down the text an emotional narrative is brought forward. When the protagonist explores different planets he is met with oppressive atmospheres that positions him - aurally - as minuscule in the galactic scope, we feel his isolation.

Though Nolan does betray himself visually at a few points. There is a scene after lift-off where one scientist explains the mechanism of a worm-hole by folding a paper in half and punching a hole through it with a pen; a metaphorical action that is crystal clear. Yet the scientist describes this metaphor, and breaks it down to the audience in an expository dialogue, perhaps in case someone couldn’t see it? Nevertheless, because the mixing places this monologue amidst the sonic texture you would be hard pressed to understand the words coming from his mouth.

However the scene was framed in an unremarkably classic way; with a mid-shot framing both his actions and his lips, which sets an expectation of understanding what the subject is saying. Had he framed out the scientist’s head the focus would not be on what he’s saying but rather what he’s doing.

Michel Chion names the word-led dramaturgy “Theatrical Speech” (audio-vision) where the cinematography, editing, and sound effects play to the text like punctuation. In reaction to and in support of the spoken word sounds become commas, question marks, weighting to the dialogue. It doesn’t have to be this way, there are plenty of films whose narrative is moved along by visual techniques, sonic techniques, or performance; and plenty with no discernible narrative at all. However you will be hard pressed to find them on the mainstream market.

One exception is All Is Lost (Chandor, 2013) that made the rounds over its incredibly sparse use of dialogue, with only a handful of spoken lines. At the time of release it generated some buzz, but it did not make its budget back within its local market (United States). Instead it relied on the international market to break even.

Incidentally, I find it fascinating that a film without text was regarded as a novelty. Before synchronised sound, text was relegated to the few screens in between visual action. Moreover, silent films were never truly silent. Without our listening focusing on decoding language there was space for the aural imagination to take hold. I still remember the clattering and puffing, clanking and sparking, sucking sounds of machineries in Metropolis (Lang, 1927).
With the adoption of synchronised sound film became “talkies”, to the detriment of all other audio-visual aspects. Language became an exclusionary factor, only to be appreciated fully by fluent speakers of said language. Even with translation there were subtleties of diction and subtext of grammar that got lost in the process. Trans-european film production was the norm as language or accent (as the national & socioeconomic identifier) did not matter as much in the silent era.

When Hitchcock transitioned the filming of blackmail (1929) as a silent film to a ‘talkie’, he had already casted the Czech actor Anny Ondra for the lead role. Her Czech accent, which had posed no problems during its silent inception, became an issue now that her character had to speak. To solve this Hitchcock had the English actor Joan Barry stand out of frame and talk over Anny as she was miming her lines.

“The life force of music is materialized on the brink of its own total disappearance.”
Andrei Tarkovsky, Sculpting In Time

Consider dense mixes (like action sequences) where there can be a lot of music and dialogue at the same time, often you can audibly hear the music and sound effects ‘duck’ under the dialogue in volume. A mixing decision for intelligibility, so that we can understand the text being spoken.

When mixing the short film Killing R (Lopez, 2023) I ended up in a similar place where a lot of my sound was covered up by musical cues that sometimes were in competition with me over who can say the same thing “the most”. Other times we were residing in different stories and genres. To get through the mix there were moments that either the sound or the music had to take over completely, but what ended up connecting these elements were additional sound design that met the music halfway.

At the film’s climax there’s a dance with non-diegetic music in the spotlight. Whenever there was a test-screening this scene didn’t connect with audiences, you lost connection to the main character and the music made the moment pleasant, “too easy”. With foley and sound effects a sonic connection was established between audience and characters, and they became expressions of the character’s anger and frustration. Sound effects like thunder and creaking trains got to role-play a pillow being thrown and a painting smashed respectively. At times completely dissonant to the music.

On the other hand the music eventually became less rigidly fixed to its cues but became sound material that could be used freely in the film. Treating the music like sound effects allowed for a sonic dramaturgy which was markedly more emotional and less conceptual; empathetic sound.

The effect that the ‘right’ music lends a sequence is undeniable. Yet it resists the film, music wants to be shared and listened to on repeat. Jonny Greenwood’s score for We Need To Talk About Kevin (2011) was never meant to be listened to separately. His compositional approach lies in creating musical building blocks that then can be edited to (and for) a film. Still it grew out of its film and became a separate release, its own entity.

As a sound-maker I have scored a lot of films with incidental sounds. My process of creating a soundtrack is very similar to producing a song. Rather than working with guitars and pianos (insert instrument of choice here) I create rhythmic beds with steps and textures, point effects that disappear as quickly as they appear. There are tonalities everywhere, I find them often in the hum of an air conditioner or vacuum cleaner. Winds sweep like glissandos and thunder rolls like a timpani. I create harmonies and motions in time with these sounds. This composition never escapes the film. Would. The soundtrack is finished in post-production. Then it is petrified.
The not-yet-seen and sound-images/

When I see Derek Jarman’s Blue (1993) I am struck by the fluidity of the image. At first glance it seems like a solid, unchanging blue frame, however as the film begins I see the frame as a clear blue sky, from the perspective of lying down in the grass looking up. Later in the film I get the impression that I am staring down into a still and clear body of water. Even later I am suddenly staring into a cold blue wall at a hospital. What drives the narrative in his film is sound and voice, always recontextualising that static blue frame into a series of different impressions. While the visuals never provide new information, the tension between sound and image ebbs and flows like a concerto. The work would, in my opinion, be lesser were it not for that blue frame.

In my practice there are always images, ones that I get from my subconscious and in my dreams. When I layer sounds or surround-scape I am in fact creating image impressions for myself, and for other listeners. With sound alone I don’t have the power to instil a specific stable image into every listener. That tasks fall to the listener who has the agency to consciously or sub-consciously manifest them. This lack of specificity is a unique quality to sound. Unlike text that shares a similar interpretative imaging experience sound is also a physical acoustical experience, you can feel your whole body resonate with it; at once visceral and vague.

Sound thrives when categorised, but does not adhere to its category in the strictest sense. For instance a sound can appear simultaneously diegetic and non-diegetic, like in the “grace” section of Atlantis (1991); fish applauding the aria of a manta ray, bringing the human experience provokingly close to nautical life. Like two atoms sharing electrons, it transcends past its category and becomes a new narrative matter.

I made a short film, Orgy For One (2022) where I started the project by just recording “foley”. Exploring different materials through my body. That was edited in surround using my body’s reaction, cerebrally I was looking for patterns and interplay in the sound, and some sort of musical arc. Corporeally I was focusing on whatever gave me a physical reaction like goose bumps or shivers. When the sound-track was finished I treated it like a script, letting it evoke images that I shot over three days. In short sessions exploring one visual idea per session.

In the mornings between shooting sessions I edited the visual-images to the sound where I found an interaction, during the editing process I would listen to the sound-track and get a new idea that I would shoot later at the following session. The concrete linking of visuals to audio has to do with finding sync points so the sound and image magnetise, but that sync point is fluid and changes for me every time I see it. Interestingly the first cut turned out too long which often happens for short films, but to keep the spirit of the sound-first approach I re-edited the sound-track and then conformed the visuals to the sound.

When sound-making comes at the very end of the filmmaking process it is already too late. By that time there is a structure and dramaturgical aesthetic established, and the creative force of sound is shackled; often it has to be tamed so production sound doesn’t jump out at you (ie. uneven volumes or noise), or is relegated to smoothing over errors of visual continuity.

The visual-images are concretely material and undeniably there, they can be frozen in time and easily analysed by it’s content and composition. It can be printed and shared, drawn on and explored to no end. Sound-images are evocative but fleeting, they can’t be frozen but need to be re-played; and also re-experienced many times. They can spark images and ideas of causality, or bring forth connotative
memories when filtered through the listener. Because the sound-image is ever changing it is also a continuous source of inspiration.

I want to end on a quote from Brandon LaBelle, from his book Sonic Agency (2020, pg.20),

“I search for it, and yet, it is already gone; even it recorded, I must play this sound again and again in order to understand its shape and density, its frequencies as well as psychoacoustical impact. As such, it may slip through the fingers to elude description. It never stands up, rather, it evades and is therefore hard to capture fully. Weakness, though, is put forward as a position of strength; a feature whose qualities enable us to slow down and attune to vulnerable figures and the precariousness defining the human condition. Sound teaches us how to be weak, and how to use weakness as a position of strength.”
Bibliography

BOOKS/


TEXTS/


VIDEO GAMES/

FILMS/


We Need to Talk About Kevin (2011) Directed by Lynne Ramsay [film]. Artificial Eye.


