Immersive Continuity:

Long Takes, 3-D Sound, and the Impression of Reality in the Cinema of Alfonso Cuarón

René Fernando Idrovo Zambrano

PhD

University of York

Theatre, Film, Television and Interactive Media

April 2021

ABSTRACT

This PhD thesis provides a sound-driven analysis of Alfonso Cuarón's approach to filmmaking, arguing that his film style, which the author calls immersive continuity, stands out as the most effective method for transporting the spectator into the film's narrative world. Grounded on the spatiotemporal continuity of the long take, Cuarón was able to exploit the three-dimensional capabilities of Dolby's most advanced sonic platform, Dolby Atmos; thereby, making the most of the cinematic apparatus, his immersive film style facilitates the occurrence of highly affective 'out of body' audio-visual experiences, the closest that for now we can get to the mythical total cinema.

Moreover, apart from analysing Cuarón's audio-visual aesthetics and their potential spectatorial effects, this thesis investigates the creative practice behind the sound design of *Roma*, whose soundtrack was conceived as a Dolby Atmos mix from the very beginning. Based on information provided by some of the key practitioners involved on the creation of the soundtrack, the author determines the creative strategies and workflows that were put in practice for the creation of what is perhaps the most realistic audio work in cinema history, undoubtedly the quintessential example of an object-based Atmos mix.

Furthermore, based on an extensive observation of present-day cinema, the author argues that there is an ongoing stylistic turn towards a broader application of 3-D sound, which is evidenced not only in *Gravity* and *Roma*, but in a number of recent mainstream productions, among which stand out highly regarded directors such as Alejandro González Iñárritu, Bong Joon-ho, Darren Aronofsky, among others.

Finally, aiming to offer insights on how to create films with 3-D sound, this thesis concludes with a practical proposal in the form of guidelines and suggestions that could lead filmmakers to plan and design their films with three-dimensional sound in mind. A 3-D storyboarding methodology is presented, which is put in practice to visualise the distribution of images and sounds in the space of exhibition.

LIST OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First of all, I would like to thank God for offering me the challenge and the opportunity of doing this PhD in the prestigious University of York. To my family special thanks for their love and support during these demanding times. I also would like to thank all my supervisors: Mr Simon van der Borgh, Dr Mariana López and Dr Sandra Pauletto, whose support and feedback have been very useful for strengthening my arguments and ideas. Special thanks are also due to the members of my thesis supervisory panel, and to all those scholars who directly or indirectly contributed academically during this PhD journey. Thanks to all my interviewees for kindly offering me their time to talk about film sound, specially to Sergio Díaz, whose knowledge was key for answering this thesis's queries. Thanks to all those friends that have crossed paths with myself during the last few years, eternal blessing to all of you. Thanks also to Universidad UTE and its rector, Dr Ricardo Hidalgo Ottolenghi, for approving an unpaid leave so I could do my doctoral research abroad.

Finally, it is worth acknowledging that this project would have not been possible without the financial support of the National Government of the Republic of Ecuador through the National Secretary of Higher Education, Science and Technology (SENESCYT).

DECLARATION

I declare that this PhD thesis is my own work and that all source material has been referenced. I also declare that parts of this research have formed part of conference presentations, symposiums, video essays and journal articles. Although one of these publications has a second author, this thesis only uses those sections that are my own. A full list is included below:

- Idrovo, R. and Pauletto, S. (2019). Immersive Point-of-Audition: Alfonso Cuarón's Three-Dimensional Sound Design Approach. *Music, Sound, and the Moving Image*, 13(1), 31-58.
- Idrovo, R. 'Immersive Point-of-Audition: Alfonso Cuarón's Three-Dimensional Sound Design Approach' at the *Sound/Image19 Conference* on 9-10 November 2019 hosted by the University of Greenwich [Conference Presentation].
- Idrovo, R. 'Alfonso Cuarón and the Immersive Cinematic Apparatus in the Times of Progressivism' at the *TFTI Postgraduate Virtual Symposium 2020: Global Concerns in Storytelling* on 18 June 2020 hosted by the University of York [Conference Presentation & Video Essay].
- Idrovo, R. 'The Immersive Continuity of Roma: Netflix and the Cinematic 3-D Soundtrack' at the *Coming Soon to a Screen Near You Symposium* on 23 April 2021 hosted by the University of York [Conference Presentation].
- Idrovo, R. (2021). The Immersive Continuity of Roma: Towards the Consolidation of an Alternative Audio-Visual Style. *Music, Sound, and the Moving Image* [Approved article, pending publication].
- Idrovo, R. (N.D.). The Gravitational Pull of Gravity. *New Review of Film and Television Studies* [Pre-approved article pending revision for publication].

INTRODUCTION

Background Overview and Thesis Scope

Technology and cinema have always been interconnected. The developments of the former have historically produced changes in the later, adjusting conventions and audience expectations in a progressive manner. Given that different periods follow different conventions, the expectations of today's audiences are not the same as the ones of earlier times (Bordwell, Thompson and Smith, 2017, p. 453). Similarly, the possibilities offered by today's technologies are not the same as the ones offered in the silent era. Before the 1930s, it was difficult to make a film with sound[1], and before the 1960s, most films had to be made in black and white as making a film in colour was not yet affordable (ibid). Such technical restrictions have historically influenced film aesthetics, styles and conventions, which tend to be continuously modified as new developments appear. Indeed, cinema changes, and as Altman argues, "the action of sound is one of the prime reasons for that change" (1992b, p. 37). Nowadays, *immersive sound* is one of the most important developments that have been incorporated to the cinematic apparatus, presenting challenges and opportunities for moulding once again the ways we make and experience a film.

But ever since its most early developments, film sound has had the ultimate desire of enhancing the realism[2] of the cinematic experience. Its goal has always been to improve the illusion of presence in the space of the diegesis; that is, its aim has been to achieve what André Bazin referred to as "the myth of total cinema*"* (1967, p. 17-22). Bazin—whose writings were first published in the mid-forties in the French magazine '*Cahiers du Cinema*', and translated into English in the posthumous publication 'What is Cinema?' (1967-1971)—argues that those who invented the cinema imagined it "as a total and complete representation of reality; they saw in a trice the reconstruction of a perfect illusion of the outside world in sound, color and

[1] Alan Crosland's Don Juan (1926) was the first film to include a synchronized musical score and sound effects; while Crosland's The Jazz Singer (1927) was the first film to include a synchronized dialogue in several sequences.

[2] In the context of this thesis, the term 'realism' is not referring to an aesthetic that aims to portrait reality in a factual or documentary fashion, realism is understood here as an aesthetic that intends to produce a sense of spectatorial presence in the depicted space by means of recreating the sensorial ways in which we experience reality—namely the cinematic representation of human vision and hearing.

relief" (1967, p. 20). Consequently, for Bazin, "[i]t would be absurd to resist every new technical development aiming to add to the realism of cinema" (ibid, p. 21), and ultimately, it would be absurd to impede the achievement of such a mythical total cinema. Evidently, today's immersive sound technology not only contributes to such a realistic cinema, but it aims "to generate a sense that one has left the real world and is 'present' in the virtual environment" (Mera, 2016, pp. 93).

This brief introduction presents a historical overview of cinematic sound. We shall depart from the advent of *monophonic audio* in the late twenties, which shall be followed by a quick overview of a number of sonic inventions that occurred throughout the years. Ultimately, this section shall end up introducing Dolby's latest sonic platform, Dolby Atmos.

The Dawn of Sound in Cinema: Intelligibility vs. Fidelity

During the early years of sound in cinema, commonly referred to as the transition era (from 1926 to 1934), there was a big debate, a confrontation, an intense battle between *intelligibility* and *fidelity* as models of sound representation (Altman, 1992a; Lastra, 1992; 2000; Wurtzler, 1992). With the advent of the Vitaphone, an analogue sound-on-disc system utilised for Alan Crosland's *The Jazz Singer* (1927) and many other early sound films, the establishment of new representational conventions was at play. Aiming to conserve perceptual fidelity, some sound researchers and practitioners defended the idea of matching sound to image scale and perspective, a sound approach that Lastra refers to as the "phonographic" model (ibid, p. 138). On the other hand, some others supported intelligibility and aimed to capture direct and non-reverberant sound with homologous dialogue levels, an approach that Lastra calls the "telephonic" model (2000, p. 139).

The phonographic paradigm of sound representation aimed to privilege the originality of the sound event over all other considerations. The idea was to represent all aspects of the sound in question as a faithful duplication (Lastra, 2000). Consequently, Lastra argues, the phonographic model aimed to place the auditor as much as possible within the *pro-filmic space*, that is to say, the original scenario in which the filming took place. As such, the defendants of such an idealistic sound production practice aimed to offer "a simulation of real human hearing" (ibid, 160); "which entailed 'looking' and 'listening' to the profilmic

performance from the same position" (ibid). In contrast, the telephonic model of sound representation—deployed in films such as John Francis Dillon's *The Girl of the Golden West* (1930)—assumed that there was a sort of hierarchy of importance among all the different sounds. Thus, the representation of space could be fabricated, with the intention of foregrounding desired sonic elements, such as dialogue, at the expense of putting aside those sound that have a lower hierarchy (ibid, p. 182).

Altman asserts that "[t]he single most important question occupying Hollywood sound technicians during the late twenties and early thirties was this: what relationship should obtain between sound scale and image scale?" (1992a, p. 46). The ideal was to assure sound localisation by means of a faithful and realistic relationship between both (Altman, 1992a); and although such a systematic audio-visual correlation was probably never practiced, scale matching and sound perspective were considered as 'obvious' goals in much of the technical literature of that era (Lastra, 2000). As Wurtzler comments, "this debate on Hollywood sound practice can offer a privileged historical instance in which can be read the process of establishing realistic representational conventions" (1992, p. 97).

Joseph Maxfield, by that time a chief engineer at Western Electric, is acknowledged as one of the most important defenders of the so-called phonographic model in the early sound era (Altman, 1992a; Lastra, 2000). As Lastra puts it, "what was true for Maxfield was true for the rest of the field" (2000, p. 184). On a paper published within the context of the early years of film sound, Maxfield (1930, p. 85) argued that "[t]he problem to be solved is that of obtaining a sound record which correlates with the picture in such a manner that a member of the audience is given the illusion of being an actual spectator in the scene". Lastra refers to this as the "invisible auditor" or "embodied-audience-member" approach to realistic sound representation (2000, p. 160), a concept taken from a previous paper written by Maxfield and Harrison in 1926, which mentions:

> Phonographic reproduction may be termed perfect when the components of the reproduced sound reaching the ears of the actual listener have the same relative intensity and phase relation as the sound reaching the ears of an imaginary listener to the original performance would have had. (Maxfield and Harrison, 1926, p. 327)

Maxfield (1930) observed that the variations between *direct sound* and *reverberation* created the illusion of movement, which could be used to represent sounds that follow the image across the screen. By that time, the so-called invisible auditor approach to sound

representation was considered as the ultimate ideal in the context of film sound. In the praxis, however, it made it difficult to cut from close-ups to long shots without introducing extreme acoustical changes on reverberation and consequently on dialogue intelligibility (Lastra, 2000). Such sudden variations in sound textures, Lastra argues, did not work appropriately in the context of classical representation and narration, which instituted spatial hierarchies determined by the narrative, with the dialogue track above everything else. Unfortunately, as Doane notes, respecting sound perspective—matching "close-up sound" with close-up picture, and "long-shot sound" with long-shot picture—compromised dialogue intelligibility (1985, p. 59), and thus was not able to take hold.

In despite of such difficulties, the desire for matching sound and image scale continued during the thirties. In an article published in 1930, John Cass, another well-known sound professional of that period, criticized the common practice of using more than one microphone, a technique whose goal was to record intelligible dialogue lines. For Cass and others, the illusion of reality was lost when using homogenous dialogue levels, between a close-up and a long shot, for instance. Doane notes that during the early years of sound in film, it was common to use several microphones to record sound from different positions around the set, destroying this way the spatial depth conveyed by the image and the illusion of perspective and of certain spectator position (1985, p. 59). As Wurtzler puts it, the positioning of such an invisible cinema subject was "threatened by a sound film practice that privileges intelligibility of the voice over the matching of sound and image perspective" (1992, p. 98). Cass observed:

> When a number of microphones are used, the resultant blend of sound may not be said to represent any given *point of audition*, but is the sound which would be heard by a man with five or six very long ears, said ears extending in various directions […] When this scene is projected, the eye will jump from a distant position to an intermediate position, and from there to close-up positions on important business. The sound will run throughout as though heard from the indefinite position described above. Since it is customary among humans to attempt to maintain constant the distance between the eye and the ear, these organs should move together from one point to another in order to maintain our much mentioned illusion. My observation has been that this lack of coordination of eye and ear is the most flagrant fault in sound recording at the present time. (1930, p. 325, emphasis added)

Such an idealistic attempt for spatial fidelity in sound representation during the early years of film sound created tension between sound professionals and other practitioners such as cinematographers and editors (Lastra, 2000). Cass claimed that it was needed a "better understanding of photography by sound engineers, and better understanding of sound by

cameramen" (1930, p. 326). Meanwhile, Frank Lawrence, by that time president of the editor's union, manifested:

> Every blooming sound expert in the entire world is at present convinced that the only way to make satisfying sound pictures is to sacrifice every other feature of value in Filmland to the proper recording of sound. . . . [The] more experienced divisions of motion picture production [should not] permit such rot. . . . Sound experts will have to get in step with the motion picture fraternity (Lawrence, 1929 as cited in Lastra, 2000, p. 167).

It is important to point out that sound professionals during the dawn of the so-called talking pictures, were mainly engineers and technicians borrowed from other industries, such as telephone companies and radio broadcasting studios. Altman thus argues that the early years of cinematic sound had the mark of a version of reality brought from other modes of representation, first the silent cinema and then the radio (1992a, p. 55). Theatre was, in fact, their closest reference for discerning the way auditory perspective should be heard in the cinema, as stressed by Franklyn Hunt (1938, p. 351), another contemporary sound professional and researcher within the transition era:

> An auditor at the theater knows by his sense of hearing as well as by sight when an actor moves about the stage, and he can tell in what parts of an orchestra the various instruments are playing […] In a sound picture theater, on the contrary, the sound always comes from one fixed source, namely, the loud speakers behind the screen. This makes the reproduction resemble sound heard when an auditor listens directly to speech or music with only one ear. Under these conditions the perception of direction is seriously impaired.

Indeed, although these technicians were experts in their previous field, it took time for them to realise that their new field had different norms, and that the ones they knew from their previous fields were not appropriate anymore (Lastra, 2000, p. 168). In their transition from one industry to another, sound professionals had to face new standards, the norms dictated by the cinema industry. Ultimately, due to the exigencies of narrative cinema, sound technicians were not actually able to put the invisible auditor approach in practice.

Yet, some of them continued to develop techniques that could respect the ideal of fidelity without putting at risk dialogue intelligibility. For instance, in their quest for realism, Maxfield and other sound practitioners started to manipulate reverberation on the set, and learned that with controlled reverberation they could place sounds at any distance from the camera (Lastra, 2000). Carl Dreher, another well-known sound engineer of the transition era, suggested that "[t]he principal characteristics of good recording in the motion picture field may

be classed under two heads: (1) intelligibility of dialog; (2) naturalness, or acoustic fidelity to the original rendition" (1931, p. 756). Dreher claimed that by that time intelligibility was a big concern in terms of sound recording mainly because of the technological equipment that was available. Yet, "[t]he elimination of film ground noise which is now in progress", Dreher noted, "promises to be a powerful force in promoting natural recording. With ground noise eliminated or materially reduced, it is possible to decrease the levels considerably without losing intelligibility" (ibid, p. 758). Regarding the relation between these two fundamental pillars, intelligibility and naturalness, Dreher argued:

> Some critics believe that close-up quality should be maintained even with a long-shot picture. More discerning observers are inclined to the view that the sound should in general follow the action, so that the loudness will decrease with increasing distance in the picture, but that, even at its minimum, the sound must be loud enough to be clearly understood if the action of the play requires it. (There are, of course, instances when a mere murmur of voices or certain intelligible lines standing out above unintelligible material are all that is required.) (ibid, p. 757)

In spite of all their efforts, dialogue intelligibility gradually became accepted as the norm, which aimed to prioritize narrative understanding over everything else (Lastra, 2000). Sound "technicians no longer sought to match sound scale to image scale through 'correct' microphone placement, but instead sought to produce a continuous sound track of nearly level volume and unbroken close-up characteristics" (Altman, 1992a, p. 54). Joseph Maxfield himself began to criticize sound recordings with too much reverberation, arguing that such an approach overemphasises incidental noise (Lastra, 2000). According to Harold B. Franklin, by then head of Fox West Coast Theatres, the greatest advantage of sound cinema was the "ability to present every word so clearly and distinctly that no one need strain to hear what is being said, at least when recording and reproducing is properly conducted" (As cited in Altman, 1992a, p. 58). As a result, Lastra argues:

> [I]t became the norm not to match visual and acoustic 'scale,' not to locate the microphone with the camera, not to respect the acoustics of the space of production, and not to offer a perceptually based 'coherent point of audition' with which the spectator could identify. Instead, technicians developed a flexible set of norms that sought to enhance intelligibility through close miking and soundproofing. (2000, p. 188)

Along with the new norms came new technological innovations such as directional microphones, more mobile microphone booms, as well as more research in the field of film sound (Lastra, 2000). Thus, the thirties saw the adoption of the mobile mike, which suspended from a boom was able to move silently above the actors, capturing in great detail all the

dialogue lines (Altman, 1992a, p. 53). With these new devices and accessories, the microphone could be kept within a relatively homogeneous distance of the speaker. Ultimately, a new model of sound representation that balanced foreground and background sounds was forged. The ideal of sonic realism was still alive, but had to evolve from the obsession of absolute fidelity to a model that recognised that strict fidelity to the actual space is not as important as producing spectator effects by manufacturing a desired diegesis (Lastra, 2000, pp. 193-194).

At last, sound engineers had to recognise that a classical film is constructed by the combination of fragments rather than autonomous units. As Lastra (2000) explains, it would be difficult to manufacture a continuous soundtrack without maintaining the sort of sonic hierarchy demanded by dialogue intelligibility, and without controlling the changes of sonic perspective when different shots are edited together. Therefore, sound technicians were encouraged to focus on recording only the dialogue on set, using a microphone located close to the actors, as background sounds could be added later (ibid, 2000, p. 175). The confrontation between intelligibility and fidelity was thus replaced by what Lastra refers to as the "foreground/background system" (2000, p. 202), an approach that was more flexible and consistent with the demands of Hollywood's narration, as it allowed recording close and intelligible foreground sounds, while leaving background elements to be added later in the process that is now known as postproduction.

This new approach had according to Dreher (1931) two main purposes: to enhance the quality of the original recording, and to introduce sound elements with greater effectiveness than if they were captured during the original recording. In regards to the latter, George Lewin, another well-known sound technician of those years, asserted that "any background noises which may be necessary are easily put in later by dubbing" (1931, p. 42). Maxfield himself, the most important defender of absolute fidelity, acknowledged that 'a certain faking' was needed in order for the recording to be effective (Lastra, 2000).

Hence, the new system of foreground and background sounds opened the possibility of keeping both continuity and some sort of naturalness in the soundtrack during the early thirties, it benefited dialogue intelligibility without sacrificing the background sounds, as the latter could be added during what they referred to as re-recording. As Belton notes, technological developments led to changes in recording, editing, and mixing practices, which were the result of the improved possibilities for controlling the soundtrack and thus "duplicate the sound not of the pro-filmic event but of that event's photographic image" (1985, p. 70). The soundtrack

passed from being ‘recorded’ to being ‘built’, with mixing practices that did not aim to conserve the integrity of any pre-existent reality anymore, but to build its own reality based on a previously recorded image (ibid). The concept of sonic realism had to pass from the idealistic invisible auditor approach to a system that constructs the soundtrack out of many different layers and hierarchies. The once ‘obvios’ goal of having a coherent point-of-audition for each individual shot was left behind in the service of dialogue intelligibility. Ultimately, the ideal of achieving spatial realism through matching sound and image scale and perspective, although still alive, had to gradually evolve during the transition era to fit within the framework of Hollywood’s continuity and narrational demands.

Spatial Sound Technology Innovations and Aesthetics

Since the early years of film sound, assuring sound localization by means of a realistic spatial correlation between image and sound has been one of the big concerns for sound practitioners. According to Altman (1992a, p. 47), there have been historically three main approaches to this end, all of them related to technological factors:

1. manipulation of the place of exhibition, largely through speaker placement and switching mechanisms (1927-31);
2. manipulation during production, especially of microphone choice and placement, along with control of sound levels during editing (1929-present);
3. development of multichannel technology, eventually including stereo-phonic localization capability (1930-present).

Although Altman published his work in 1992, his argument continues to prevail today about even the latest and most advanced approaches to cinematic spatial sound.

Similarly, Doane (1980) notes that sound technology has historically been innovated in response to the need of more realistic reproduction. In 1940 for instance, Walt Disney’s *Fantasia* debuted with an experimental multichannel sound system, *Fantasound*. Developed by RCA and Walt Disney Studios, Fantasound consisted of three horizontal speakers behind the screen—fed by three recorded channels—and a surround channel that fed ninety-six small speakers located around and behind the audience (Kerins, 2011). However, in despite of its magnificence, Fantasound did not achieve the success that its creators expected, and only six theatres in the United States were equipped with this technology as it demanded exhibitors not

only to add a big number of speakers, but also an additional projector (ibid). The amount of equipment required, the time and costs that were necessary to make the installation, and the advent of wartime, all played against the proactivity of this very first multichannel sound system. Hence, Fantasound disappeared after being deployed in only one film and a few theatres (Kerins, 2011, p. 24).

Later on, in the mid-fifties, the traditional cinema viewing experience was transformed as part of the widescreen revolution with systems such as Cinerama, Cinemascope, and Todd-AO, which brought changes to sound as well. As Belton observes, cinema sound passed from being reproduced through one behind-the-screen speaker, to as many as five speakers located behind the screen, plus dozens of surround speakers spread all around the auditorium (1992, p. 154). Audience spectatorship went from distraction—quite common in the twenties because of the theatres' architecture itself—to 'participation', a catchword widely used in the fifties to describe high levels of audience absorption during the cinema viewing experience (Belton, 1992), which could be compared with the idea of *immersion*[3] in today's cinema.

Following André Bazin's theories of cinema, Belton argues that film technological advances have been commonly associated with the production of what he refers to as "*greater realism*" (1992, p. 159). During the 'widescreen era', *Cinerama*, the first exhibitor to offer huge widescreen projection, advertised to be a medium "that creates all the illusion of reality . . . [that] you see things the way you do in real life […] [and that] you hear with the same starling realism" (As cited in Belton, 1992, p. 159). Along with an impressive fifty-one-foot widescreen, Cinerama began to practice what was referred to as *'travelling' sound*, allowing sounds to move from back to front and from left to right (and vice versa) across the auditorium, which contributed not only to an apparent realism but also to the spectacle of the cinematic experience (Belton, 1992, p. 161). The footsteps and even the voice of a characters walking from right to left across the whole length of the screen would emanate first through the right loudspeaker, then though the centre and finally through the left, "the source of the sound matching, point to point, the character's various positions on the screen" (Belton, 1985, p. 68).

[3] In the context of this thesis, the concept of immersion is understood as the replacement of "as many real-world sensations as possible with the sensations of a virtual environment" (Mera, 2016, p. 92). Acoustically speaking, this means the recreation of the way we hear in real life, where sounds reach our ears from all directions in accordance to the position of the sonic sources in space. Immersion can thus enhance the sensation of spatial presence in the diegetic world and thereby increase the levels of engagement with the film's narrative.

Examples of such an approach are found, for instance, in Cooper's *This is Cinerama* (1952), Walsh's *Battle Cry* (1955), Kubrick's *Spartacus* (1960), among others.

However, the apparent realism of the so-called 'travelling' sound was perceived by others as rather artificial. Not only sound effects, but also music and even dialogue moved from speaker to speaker, to the extent that "[f]or the first time in film history, offscreen dialogue was literally offscreen, emanating from surround speakers on either side of the auditorium"[4] (Belton, 1992, p. 163). Although it was a big novelty, 'travelling' dialogue was not accepted by the vast majority of studios that thought, as claimed by Loren Ryder—a Paramount's sound engineer of the 1950s—that stereo playback could be effective for sound effects and music, but found the movement of dialogue to be very annoying (Belton, 1992). It was quite difficult to replace the well-established monophonic convention that had been the standard for over twenty-five years, within which dialogue had to come out from the behind-the-screen central speakers (Belton, 1992, p. 164). As a result, 'travelling' dialogue gradually vanished in favour of central dialogue, which was complemented by stereo music and sound effects (Rumsey, 2018, p. 183).

It is important to notice, nonetheless, that it was not only the monophonic tradition that kept 'travelling' dialogue from being a successful novelty, but also all the technological limitations of the fifties. Yet, technological advances have not stopped. After the failure of multi-channel experiments, such as the quadraphonic-based system *Quintaphonics*[5], monophonic reproduction continued to be the standard during the sixties and most of the seventies (Kerins, 2011). But in 1976, a new sound system debuted and gradually became the new standard of theatrical sound, *Dolby Stereo*. Developed by Dolby Laboratories, this new sound system not only provided more frequency and dynamic range, but also encoded four channels of audio down to two, which made it cost effective as the two channels could be printed—replacing the monophonic optical track—into the 35mm film, a feature that previous multichannel systems had not accomplished (Beck, 2008; Kerins, 2011).

Certainly, the characteristics of such an innovative sound system brought a lot of benefits; yet, Dolby Stereo also presented a lot of limitations that had to be circumvented by filmmakers in order to achieve the desired effects (Kerins, 2011). Kerins (2011) observes, for

[4] See, for instance, Koster's The Robe (1953).

[5] A quadraphonic-based system that expanded the 4 channels of quadraphonic sound to five (left, centre, right, left surround and right surround)

instance, that many films associated with Dolby Stereo—such as George Lucas' *Star Wars* (1977), Steven Spielberg's *Close Encounters of the Third Kind* (1977), Richard Donner's *Superman* (1978), and Francis Ford Coppola's *Apocalypse Now* (1979)—used, in fact, audio setups that were not part of the original specifications of the Dolby system. *Apocalypse Now*, the film in which Walter Murch[6] gave origin to the term *sound designer*, was branded as a Dolby Stereo release, but it actually used a multichannel configuration that was closer to Quintaphonics (Kerins, 2011, p. 32).

Indeed, filmmakers were not satisfied with the limiting possibilities offered by 35mm Dolby Stereo. For instance, Walter Murch, *Apocalypse Now*'s sound designer and editor, preferred the 4-channel quadraphonic sound over Dolby Stereo and claimed that the former allowed to locate sounds through 360 degrees, something that was not possible with the standard configuration of the 35mm Dolby optical (Kerins, 2011). Even Dolby Laboratories recognised that the 70mm six-track discrete magnetic format offered better quality than their own 35mm optical sound (Beck, 2008, p. 70), the former being in fact what Coppola and Murch decided to use for *Apocalypse Now*. In this film, Murch panned helicopters throughout the surrounds, but only a small number of spectators in a few showcase theatres were able to hear this, as the film's 35mm Dolby Stereo prints did not include a surround channel (Kerins, 2011, p. 79). Beck observes:

> In spite of his complex use of 70mm split surrounds to envelope the audience in the diegetic acoustic space of *Apocalypse Now*, it is commonly forgotten that Murch discarded the surround tracks entirely when preparing the Dolby Stereo 35mm mix because of the impossibility of re-creating his sound design and complicated effects in the matrixed format. (2008, p. 71)

Within these limitations, Dolby Stereo standardized the practices of film sound and reinforced the screen-centric approach of monophonic reproduction. Its technological restrictions, Beck claims, "limited its stylistic use and discouraged sound practitioners from exploring cinema sound's full potential" (ibid, p. 68). Since both, filmmakers and audiences demanded a truly multichannel surround sound system, Dolby Stereo was perhaps, as Beck stresses, "a giant step *backwards* in the history of film sound evolution" (2003, p. 132, emphasis in original).

---

[6] Walter Murch is one of the most influential film editors and sound designers in modern cinema. He was the first sound practitioner to be credited as a sound designer for his work in Coppola's *Apocalypse Now*, which won the Academy Award for Best Sound in the 52nd Academy Awards.

However, in the quest for realism, technological advances continued. In 1992, the release of Tim Burton's *Batman Returns* (1992) marked the beginning of a new era, one that Kerins (2011) refers to as the *digital sound age*, it was indeed the beginning of what would become the new standard of theatrical sound reproduction, the era of *digital surround sound* (DSS). Since its conception, DSS expanded the narrative capabilities of cinema in a way that was not possible with monophonic or Dolby stereo sound (Kerins, 2006). Thanks to the possibilities brought by this new technology, Kerins (2006) claims, filmmakers were able to create realistic sound fields and spaces, such as those evidenced in the D-Day opening sequence of Steven Spielberg's *Saving Private Ryan* (1998). Based on the immersive potential of DSS, Kerins (2011) introduced the concept of the *ultrafield*, a term that evolves from Chion's *superfield* (1994 [1990]) to adjust to the characteristics of DSS, as according to Kerins, the superfield was based on Dolby Stereo's capabilities.

Chion defines the superfield as "the space created, in multitrack films, by ambient natural sounds, city noises, music, and all sorts of rustlings that surround the visual space and that can issue from loudspeakers outside the physical boundaries of the screen" (1994 [1990], p. 150). While the superfield offers continuity in the soundtrack, by means of a *continuous aural environment* that becomes the common denominator of all the picture cuts in a particular location, the ultrafield sacrifices that type of continuity and instead offers a *continuous space* in which sound sources maintain a specific relative position in the diegesis. Kerins elaborates:

> The superfield relies on a *constant* sonic space to ground freely moving and cutting images; the ultrafield allows for a *shifting* aural environment that functions in *dialogue* with the visual one. As an example, *Saving Private Ryan*'s aural elements often drive cuts in the image track—yet one might just as easily say that the visual elements of the sequence drive changes in the soundtrack. (2011, p. 92, emphasis in original)

In the highly immersive D-Day sequence at the beginning of *Saving Private Ryan*, the sound perspective and the sounds relative position change in accordance with the position of the camera, which according to Kerins is what differentiates the DSS-based Ultrafield from the superfield, its Dolby Stereo-based predecessor. "In short, the ultrafield is the three-dimensional sonic environment of the diegetic world, continuously reoriented to match the camera's visual perspective" (Kerins, 2011, p. 92).

Decoding all the characteristics of what Kerins calls the "*digital surround style*", he describes contemporary film practice as a "*cinema of immersion*" and argues that its different aesthetic traits "work *to immerse the audience in the diegetic world of the film*", both visually

and aurally (ibid, p. 130, emphasis in original). The idea, Kerins argues, is for the audience to be literally placed in the diegetic space of the film, changing the conception of the cinematic experience from something "to be watched from the outside"—with events happening on the screen in front of them—to something that they are personally experiencing as they are located in the middle of the action. Kerins claims that the cinema of immersion characteristic of the so-called digital surround style creates a "continual 'present'", which keeps the spectator's attention on that precise moment. It places the audience within the world of the characters, and thus replaces the physical space of the theatre with a fabricated diegetic space. Ultimately, Kerins argues that the digital surround style presents a tight sound/image synchronisation, with changes in the image's point-of-view driving coherent changes on the soundtrack. "Through all this, the cinematic experience shifts from being *presented with* a movie (to simply see and hear) to being *enveloped by* one" (Kerins, 2011, pp. 130-131).

Such *diegetic immersion*, the core of the so-called digital surround style, has according to Kerins a direct connection with three technical characteristics brought by digital surround sound. First of all, the powerful rumbling of the dedicated low-frequency effects (LFE) channel allows filmmakers not only to make us hear the low-frequency sounds, but also feel them, which "creates the very physical sensation of being at an earth-shaking event, not just *watching* one" (2011, p. 134). Second, DSS's large dynamic range and signal-to-noise ratio facilitate a very accurate representation of how sounds would be heard in real life, from a whisper to an explosion, the loudness level helps to convey the idea of being present in the filmic space. Third, the number of discrete, full-range channels—five in a DSS's standard 5.1 configuration—allow to create a close aural representation of the fictional world that surrounds us, an enveloping diegetic space that can create the impression of being in the middle of the action (pp. 134-135).

Certainly, diegetic immersion was made possible by DSS capabilities, and as Kerins (2011) predicted, the latest sound technological advances seem to encourage its application. Nowadays—almost a century since the release of *The Jazz Singer* (1927), the first film with synchronised dialogue—immersive sound systems with height are the biggest trend in terms of cinematic spatial sound, not only for the cinema but also for the home entertainment industry. It is important to highlight that acoustically speaking, immersion refers "to sounds as coming from all directions around a listener, which normally is an inevitable consequence of natural human listening in an air medium" (Wenzel, Begault and Godfroy-Cooper, 2018, p. 5). Hence,

when effectively utilised, these new technologies "can give the listener an experience of *being there* through sound" (Roginska and Geluso, 2018, p. 1).

In this respect, Miguel Mera (2016) argues that film aesthetics are passing through an evolutionary process, from the superfield, via the ultrafield to something that he simply refers to as *three-dimensional (3-D) sound*. Mera (2016) argues that 3-D sound is an expansion and development of the characteristics of the ultrafield, partly afforded by the increased visual and spatial depth of the so-called 3-D cinema, requiring a more dynamic use of both music and sound. Mera suggests that a three-dimensional spatialisation of music and sound leads to the experience of *spatial presence*, "the psychological perception of 'being in' the virtual environment in which one is immersed; the impression that a mediated experience is 'real'" (2016, p. 93). However, the present thesis argues that in order to achieve a complete diegetic immersion, every sound element that is part of the three-dimensional diegesis must coherently move around the spectator, resembling the way we hear in real life. Yet, even with the most advanced technologies, the vast majority of films still keep dialogue and other important narrative sounds behind the space of the screen, which highlights the recurrent predominant role of the screen-centric convention and suggests that the so-called diegetic immersion has actually been applied under certain restrictions. Certainly, as Jeff Smith stresses "even the most adventurous applications of the digital surround style are confined to individual scenes. Most films use the digital surround style intermittently. Other films do not use it at all" (2013, p. 336). In fact, "it is currently the exception rather than the rule" (ibid). Kerins himself admits that

> while new technology has increased the cinema's capabilities for diegetic immersion, and more films are taking advantage of that capability, not every moment of a given film will follow this approach, and indeed not every movie will even deploy it at all. (2011, p. 142)

In any case, realistic approaches to sound spatialisation have slowly increased throughout the years. For instance, specific moments in Kathryn Bigelow's *Strange Days* (1995), David Fincher's *Panic Room* (2002), Alfonso Cuarón's *Children of Men* (2006), and Darren Aronofsky's *Black Swan* (2010) demonstrate how DSS can convey a coherent 360-degrees sonic space in which dialogue and sound effects are localised in accordance with the position of the sources in the diegesis. In the realm of animation, Lasseter's *Cars* (2006) illustrates how such a sonic strategy can be used more consistently throughout an entire movie. And in 2013, powered by the immersive capabilities of the most up-to-date technologies of

spatial sound mixing and reproduction, Alfonso Cuarón's *Gravity* (2013) became the first Hollywood live-action feature film to have fully, coherently and unrestrictedly applied a three-dimensional sound design approach from the beginning to the end, undoubtedly a milestone in the history of film sound.

Gravity: A Three-Dimensional Approach to Sound Design

The work of Alfonso Cuarón is difficult to define when considering his broad creative trajectory. Including his low-budget debut as a director, Cuarón has directed eight feature films: *Solo con tu pareja* (1991), *A Little Princess* (1995), *Great Expectations* (1998), *Y Tu Mamá También* (2001), *Harry Potter and the Prisoner of Azkaban* (2004), *Children of Men* (2006), *Gravity* (2013), and *Roma* (2018). Additionally, Cuarón was involved in directing a segment of the French feature film *Paris, je t'aime* (2006), a production that included a total of 22 directors. Furthermore, it is worth mentioning that including his work for television, short-films and school projects, his portfolio includes 17 credits as writer, 24 as producer, 17 as director, 11 as cinematographer and 10 as editor (IMDb, 2021). This extensive career honours the spirit of his alma mater, *Centro Universitario de Estudios Cinematográficos* (CUEC/UNAM), an institution that since its foundation in 1963 has aimed to create filmmakers that could professionally perform a variety of roles: e.g. editor, producer, cinematographer, screenwriter, and director (Coria, 2019). Along his journey, Cuarón has explored a diversity of genres and styles; yet, by observing his latest productions, it seems clear that this Mexican filmmaker has reached his maturity, and his signature style has been consolidated.

In *Gravity,* Cuarón tells the story of a rookie astronaut as she tries to make her way back to Earth and survive under extreme conditions in outer space. Dong Liang notes that "[t]he film is, in many ways, another milestone of cinema sound since the Dolby Digital trailers" (2016, p. 2). Similarly, to emphasise the importance of this film and its possible impact in contemporary cinema, film theorist Kristin Thompson quotes film critic Don Shay's: "Every once in a while, a film comes along that is a game-changer. This is one of those films" (as cited in Thompson 2013). Besides its splendid use of cinematic silence—which Bender (2014) takes

as an example of 'reported realism'[7]—what distinguishes the sound design of this film is the way sound effects, music and specially dialogue move in all directions within the three-dimensional filmic space. Liang notes:

> In consistently pushing the dialogue out of the screen (or rather, the front speakers), *Gravity* presents a significant challenge to the established codes of surround sound, and signifies a triumphant return to a crucial idea in the history of sound space (2016, p. 7).

Although advances in audio-visual technologies do not dictate the way filmmakers do their work, the possibilities created by new tools can certainly have an impact on filmmakers' aesthetic choices. In this respect, it is worth acknowledging that Cuarón had the opportunity to mix the sound of *Gravity* using Dolby's latest innovation, Dolby Atmos, a sound platform that according to Gianluca Sergi is "a prime candidate to become the 'industry standard'" (2013, p. 108). Cuarón comments:

> Every time throughout all my years, when I start mixing my films, I was always asking for possibilities in the mixing room that were not achievable. And now finally, with this system [Dolby Atmos] you know, it's this dream come true […] in which you really can explode the possibilities of depth and separation as never before. (2014, Interview by Dolby)

Instead of discrete channels, this *object-based* platform defines and uses audio objects (up to 128 sound elements) together with metadata that can individually specify their aural attributes and position in space. Additionally, its optimal setup has up to 64 speakers, including two subwoofers and a determined number of ceiling speakers, which allows much more immersion than the traditional multichannel arrays. Furthermore, each Atmos speaker is individually driven and offers full frequency range, which means that now the transit of sounds from the front to the surround (or height) speakers can be more smoothly executed. Another important improvement achieved with Dolby Atmos is its capacity to be down-mixed to any other format (stereo, 5.1, 6.1, 7.1, etc.), meaning that if a theatre does not have Atmos installed, the mix would be translated to the particular configuration that is available (Sergi, 2013, p. 111). All these new capabilities are making Dolby Atmos the new platform of choice for more and more filmmakers and exhibitors.

[7] Bender's research aims to understand 'the critical and public discourse which emerges in response to stylistic devices that may or may not be factually accurate to real life but which are taken to be realistic representations nonetheless' (2014: n.p.)

Technological innovations on their own, however, are not sufficient to create an immersive cinematic experience, but rather the ways in which that technology is aesthetically used what can produce such an effect (Mera, 2016, p. 102). In this respect, this thesis argues that even though two of the most innovative sound professionals—academy award-winners Glenn Freemantle and Skip Lievsay—led the sound department and hence were in charge of designing the aural world of *Gravity*, it is Cuarón himself who is the actual creator of the film's sound concept. As director, he embraced an audio-visual method that permitted the materialisation of a completely three-dimensional soundtrack. Grounded on an unconventional utilisation of *long takes*[8], Cuarón 'broke the rules' by allowing an aggressive spatialisation of sound effects, music and dialogue in all directions across the auditorium, connecting sound with the spatial position of each shot within the diegesis. Such a precise three-dimensional coordination between image and sound is clearly something that neither the sound designer/supervising sound editor nor the re-recording mixers could have done without a supporting directorial approach. In other words, Cuarón provided the framework that allowed the sound team to fabricate an extraordinary sonic immersion. This, of course, does not mean that all creative agency shall be attributed to Cuarón, yet it is clear that the film's approach to sound spatialisation is a key component of the director's own film style.

In fact, Cuarón's interest in such an audio-visual approach was already evident in some of his previous films. In the car scene of *Children of Men*, for instance, when Theo Faron (Clive Owen) and the group of rebels are attacked by a band of refugees, the camera rotates by 360-degrees and Cuarón maintains the continuity of space by also rotating the soundtrack. Throughout this scene, the audio-visual perspective of the circling camera is maintained, hence the spectator is placed in the middle of the action, not only visually, but also aurally as sound effects, diegetic music and dialogue move all around the 5.1 setup in accordance with the position of the sources in relation to the moving camera. Skip Lievsay, who has worked with Cuarón as sound designer and re-recording mixer, stresses:

> Alfonso likes the idea that the words are attached to people. If you've ever seen Children of Men, you see that nearly all the dialogue in that movie is panned [to match the actor's position on screen]. I like that idea. It removes a certain veil of film shenanigans where the old simple idea is that all dialogue comes from the center channel—which is a film construct. It doesn't happen that way in life, and it isn't really necessary to do that in the modern film formats. (2013, Interview by The Verge).

[8] 'Shots with durations well beyond the industry standard' (Udden, 2009, p. 26-27)

Years later, on his next film, Cuarón decided to replicate that technique as the core of *Gravity*'s entire soundtrack. In this film, Cuarón notes: "sound is constantly traveling, is very dynamic, is geographically very literal, meaning if one character is talking behind you, the sound is going to come from behind" (2013a, Interview by Sound Works Collection). Hence as Cuarón acknowledges, sound becomes fundamental for the storytelling in *Gravity* "because it is part of the immersive experience […] and you go almost being a third astronaut", who follows the journey of the other two (2014, Interview by Dolby). Ironically, Cuarón and his sound team broke today's conventional rules by materialising the so-called 'invisible auditor' approach sought by Maxfield and others during the early years of cinematic sound. In other words, Cuarón broke with convention by achieving in the cinema an unprecedented *impression of reality*, enhancing with 3-D sound the cinematic illusion of presence in the narrative world.

The Long Take as the 'Look' of Immersion

Although such a three-dimensional utilisation of sound seems beneficial for the immersive experience of films, its application is complex as it requires filmmakers to seriously consider sound during the whole creative process, something that is discussed by Manolas and Pauletto as a condition to enlarge the diegetic space of films (2009, p. 43). More specifically, in order to represent the 'binaural' point-of-audition of an invisible spectator that occupies the position of the camera, filmmakers need to plan, shoot and edit their film acknowledging that three-dimensional sound will accompany each visual perspective. As sound practitioner Jeremy Price comments:

> You have to be quite careful for it not to become distracting. *Gravity* was the perfect example because of the way it was done, it just worked really well, it lent itself to it […] I always want to do as much panning as possible, but it is amazing how difficult that can be just because of camera placement. (personal interview, 2019)

Certainly, as re-recording mixer Dan Johnson corroborates, "you are led by what is happening in the story, what is happening on the screen" (personal interview, 2019). In other words, embracing this sound design approach challenges screenwriters, directors, cinematographers, and editors to execute each of their individual crafts visualising how the relation of shot-to-shot will work when three-dimensional sound is added. Evidently, embracing this approach is difficult, and perhaps inadequate, if the movie is fast-cut, but the cinema of Alfonso Cuarón stands out for its unconventional utilisation of long takes.

Throughout their creative career, cinematographer Emmanuel 'Chivo' Lubezki has helped Cuarón to elaborate a visual aesthetic that is undoubtedly beneficial for the effective deployment of 3-D sound. Lubezki, a virtuoso of the long take, has been Cuarón's collaborator for more than 3 decades. His long takes are recurrent not only in *Gravity*, but also in *Y Tu Mamá También* (2001) and *Children of Men*; three films that showcase what Bruce Isaacs refers to as "the Cuarón/Lubezki signature collaboration"[9] (2016, p. 476). Referring to his relationship with Cuarón, Lubezki comments:

> The truth is that ever since I met him, Alfonso has always been one of my most important teachers. I worked with him in film school as his gaffer when he was the cinematographer, as his boom man when he was sound mixing, as his second AC when he was a first AC, and finally as his cinematographer when he became a director. I know him well. He is my teacher and also one of my favourite filmmakers. (B, 2013, p. 49)

On his part, Cuarón acknowledges Lubezki as his "co-filmmaker". "He is not just doing what most people think of as the cinematographer's job. On *Gravity* he was everywhere, collaborating every single step of the way" (ibid).

In his analysis of *Children of Men*, Udden notes that although Bazin is not mentioned in their discourse, Cuarón and his collaborators defend the utilisation of "attention-grabbing long takes" based on the Bazinian theories of cinematic realism (2009, pp. 26-27). For Bazin, those who invented the cinema imagined it "as a total and complete representation of reality" (1967, p. 20). In the classical Hollywood cinema, one of the pioneers of such a realistic approach was Orson Welles, whose debut in *Citizen Kane* (1941) is distinguished by the utilisation of long takes and deep-focus cinematography, two cinematic resources that are also evidenced in the Italian neo-realism. Directors like Welles, Rossellini, and de Sica, Bazin argues, "were determined to do away with montage and to transfer to the screen the continuum of reality" (Bazin, 1967, p. 37). Moreover, Bazin acknowledges the work of Jean Renoir as the precursor of Orson Welles. In Renoir' films, Bazin notes, "the search after composition in depth is, in effect, a partial replacement of montage by frequent panning shots and entrances. It is based on a respect for the continuity of dramatic space and, of course, of its duration" (1967, p. 34). In contemporary mainstream cinema, such descriptions could also be applied to the work of Alfonso Cuarón.

9 'Long-take, hand-held camerawork with inconspicuous movement captures the harmonious relationship of objects within a single spatio-temporal field'. (Isaacs 2016: 476).

It is important to point out, however, that Bazin treasured the long take for its capacity to capture continuously a material profilmic reality, namely a real event happening before the camera. On the other hand, what Cuarón seems to actually seek is not a faithful recording of reality *per se*; his aim is to manufacture the impression of reality through the exploitation of the latest cinematic technologies of image and sound. For Lubezki, the main thing about the long take

> is that it is immersive […] it feels more real, more intimate and more immediate. The fewer the cuts, the more you are with [the characters], it's as if you are feeling what they are going through in real time. This is something that Alfonso and I discovered on *Y Tu Mamá También* and *Children of Men* (ibid, p. 41)

Isaacs argues that *Y Tu Mamá También* signified "the formative development" of the Cuarón/Lubezki creative collaboration (2016, p. 476). He notes that the film's visual style is based on long takes and inconspicuous camera movement that captures the harmonious relationship of characters and objects within the same spatiotemporal field (ibid). Similarly, Benjamin B (2006, p. 62) highlights the fact that the documentary-like visual aesthetic of *Children of Men* was achieved by means of the utilisation of handheld camera-work with very little film lighting, as well as the avoidance of standard shot breakdowns, such as the typical *shot/reverse-shot* patterns[10]. Lubezki explains that their point of reference was taken from actual news footage and war documentaries (B, 2006, p. 72), suggesting that they followed a realistic aesthetic with no room for *coverage*[11] and "beautiful close-ups of the gun, or the trigger in slow motion" (ibid). Indeed, as Cuarón stresses, the language that he has constructed with Lubezki "is not one based on close-ups", they "include close-ups as part of a longer continuous shot. So this all becomes choreography" (interview by ICG Magazine).

Lubezki confesses that he and Cuarón have a strong disinclination for traditional coverage that intends to present shots of two actors by intercutting them using shot/reverse-shot patterns (B, 2006, p. 72). Such a strategic narrative resource, as Burch (1982) asserts, was developed during the silent film era (between 1906 and 1929) precisely to overcome the absence of sound; yet it has prevailed since classical cinema as the most conventional visual device. Bazin notes, for instance, that around 1938, the characteristic film editing procedure "was by shot-reverse-shot, that is to say, in a dialogue scene, the camera followed the order of

[10] Shot/reverse-shot patterns are a very common tactic within the 180-degree system. The idea is cutting back and forth from one end of the axis of action to the other (Bordwell, Thompson and Smith, 2016: pp. 233-234).
[11] Footage of the same scene captured from different camera angles (e.g. wide-shots, close-ups, medium-shots, etc.)

the text, alternating the character shown with each speech" (1967, p. 33). This editing tradition, Bazin argues, "was challenged by the shot in depth introduced by Orson Welles and William Wyler" (ibid); and so it is by the Cuarón/Lubezki signature style. In the case of *Children of Men*, for instance, Lubezki explains that they

> decided to have every shot be a shot in itself and avoid the A-B-A-B of coverage, even though we couldn't get away from doing it sometimes. The more I work this way, the more I realize that conventional coverage is what makes movies feel the same. You go to see a comedy, a drama, or a horror movie, and they all somehow feel the same. It's as if the cinematic language hasn't really evolved that much. Many films just cover the dialogue without really exploring the visual dimension (B, 2006, p. 62).

Moreover, Lubezki notes that they decided to use long takes "to try to get the audience to feel they are there" (B, 2006, p. 62). This asseveration confirms that Cuarón's ultimate goal is not precisely the portrayal of reality *per se*, what he actually intends is to facilitate the experience of immersion through an illusory impression of reality, namely the sensation of being present in the narrative world. Ultimately, in alliance with Lubezki, Cuarón has developed a 'Bazinian'audio-visual style, with long take-driven visuals that allows him to manipulate the soundtrack, directing the spectators' attention with audio cues instead of visual close-ups (Whittington, 2011, p. 5).

Aims and Research Questions

Grounded on contemporary theories of film and media, this thesis intends to evaluate the potential impact that three-dimensional sound design might have on the spectator, focusing on the fact that such an immersive strategy exploits the illusory capabilities of present-day's cinematic sound technology to the full, and thus enhances the impression of reality that the cinema can offer. Following such a theoretical justification, this thesis shall determine the extent to which 3-D sound has been deployed in current cinema, especially (but not only) in the films that present a Dolby Atmos mix. Furthermore, based on the aesthetic strategies deployed primarily on Cuarón's most recent films, this thesis aims to establish his immersive approach to filmmaking as a film style that considers 3-D sound as one of its core components. Ultimately, apart from analysing Cuarón's films from an aesthetic point-of-view, his creative practice shall be deconstructed in order to determine effective ways for creating films with 3-D sound.

These goals shall be achieved by answering the following research questions:

- What opportunities does 3D sound offer as different mode of address for filmmakers who wish to engage audiences in a more immersive fashion?
- What visual and sonic aesthetics are found in Cuarón's latest films, and have other filmmakers employed similar audio-visual traits?
- What creative strategies, technical operations and workflows have helped Cuarón and his collaborators to employ the full potential of immersive sound, and what additional screenwriting and preproduction strategies could facilitate the creation of films with three-dimensional sound design?

Notes on Methodology

In order to answer the research questions presented above, a mixture of methods relying on primary and secondary sources of information will be followed. The first research question shall be answered by exploring a variety of contemporary film and media theories, as well as the empirical findings that some of those theories have brought. Based on the work of other theorists, this thesis shall offer its own hypothesis in relation to the effects that 3-D sound may have on the spectator, with premises that could be proven by future researchers through empirical experimentation.

Aiming to answer the second research question, this thesis will firstly explore the technical aspects of today's spatial sound technology, focusing mainly on the study of object-based audio in the form of Dolby Atmos as the leading alternative for the creation and playback of immersive cinematic sound. The thesis will then analyse some of the films that are subject of this investigation, embracing the multichannel methodology proposed by Mark Kerins (2011, pp. 201-252) as a fundamental base, which will be complemented by examining visual aspects—such as average-shot-length (ASL), framing, composition, movement, and so forth—and their relation with the soundtrack. A survey of a large number of recent films shall be

carried out aiming to discover whether there is an aesthetic turn going on since the dawn of Dolby Atmos and its creative utilisation on the film *Gravity*.[12]

Regarding the third research question, this thesis shall explore the film *Roma*, not only in terms of its aesthetics, but also in terms of the practical procedures that were followed by its creators. The testimony of Cuarón and his collaborators will be accessed through behind the scenes content, interviews, articles from specialised journals, and other secondary sources. Furthermore, in order to access specific information, declarations gathered from personal interviews with some of Cuarón's collaborators shall be used, especially with the ones that were directly involved in the process of sound supervision and design. Finally, a set of guidelines encompassing creative operations and filmmaking workflows shall be proposed based on all the knowledge accumulated throughout the research process. This practice-based stage of the PhD will consist on the creation of an original screenplay with its accompanying storyboard, two documents that in the form of appendixes shall lead to a final reflective discussion.

Thesis Structure

Following this introduction, Chapter 1 provides a technical description of present-day cinematic sound technology, within which Dolby Atmos stands out as the most promising platform for the production and delivery of immersive 3-D sound. A survey of articles and other literature is carried out, which is supported by personal interviews with Dolby's personnel and film sound practitioners in relation to the challenges and opportunities that the Atmos system presents.

Chapter 2 then explores the concept of illusion in cinema, a theoretical overview that works as a fundamental spine throughout the whole thesis. Beginning with Baudry's *apparatus theory* (1974-5 [1070]), the chapter explains how during the 70s such a Marxist film theory claimed that the cinematographic apparatus was used as an instrument of ideology, something that according to Baudry was thanks to the impression of reality that the theatrical experience

[12] Conditions of viewing: All films were watched on a Dolby Atmos enabled home-cinema equipment using a 5.1.2 configuration with 'upfiring' speakers. In the case of *Gravity* and *Roma*, which are this thesis's main case studies, additional viewing sessions were performed in the Dolby Screening Room at London Soho.

offers, a premise that he approached through the glass of psychoanalysis. The chapter then explores cinematic illusionism through other rather rational points-of-view, which the author ultimately connects to his own hypothetical explanation of the potential effects of cinematic 3-D sound. Ultimately, Chapter 2 suggests that, in conjunction with the camera's monocular 'vision', the three-dimensional representation of natural human hearing, namely our immersive point-of-audition, helps to augment the cinema's capacity to create the impression of reality, which may facilitate the processes of transportation[13] and identification[14] in cinema, two constructs that Tal-Or and Cohen (2010) highlight for being two central paths for facilitating audience involvement in entertainment.

Using the theory and technological facts as the basic skeleton of analysis, Chapter 3 then explores the audio-visual aesthetics of *Gravity,* which leads to the conclusion that the long take driven visual aesthetic used by Cuarón in collaboration with cinematographer Emmanuel Lubezki is an essential aspect for the creation of films with 3-D sound. In Chapter 4, the author supports his argument through an in-depth exploration of Cuarón's *Roma* (2018), an intimate independent film that is taken in this thesis as a model of film creation, and as the consolidation of what is referred here to as the *immersive continuity style*. This thesis sustains that such an unconventional film style is visually characterised by the utilisation of long takes, deep-focus, wide-shots and slow camera movement, all of which are complemented by the immersive action of 3-D sound. Moreover, based on commentaries and personal interviews with key sound practitioners involved in the creation of the film's soundtrack, Chapter 4 identifies the creative strategies and workflows that were used for producing what this thesis acknowledges for being the most realistic audio work in cinema history, the sound design of *Roma*.

Chapter 5 then observes how sound three-dimensionality has become more recurrent in contemporary mass-audience cinema, meaning that the screen-centric convention of keeping dialogue in the centre speaker behind the screen is no longer as much a concern as it was before. A number of excerpts from a variety of recent films are analysed in terms of the use of sound

[13] In the context of this research, the concept of 'transportation' is understood "as an integrative melding of cognitive, emotional, and imagery engagement in a story" (Green and Sestir, 2017, p. 1), something that, as this thesis shall argue, may be facilitated through the utilisation of 3-D sound.

[14] In the context of this thesis, the concept of 'identification' is understood as "a mechanism through which audience members experience reception and interpretation of the text from the inside, as if the events were happening to them" (Cohen, 2001, p. 245).

and its connection to the image, which suggests that Cuarón's proposal has casted a sort of 'gravitational pull'.

Ultimately, inspired by Cuarón's audio-visual aesthetics and practical strategies, Chapter 6 proposes an original filmmaking method for including 3-D sound within the creative decisions that take place early in the screenwriting stage during development, and throughout the preproduction of the film. The main aim of this thesis is to establish immersive continuity as a film style with its own rules of film construction, that is, the goal is to propound a set of guidelines that shall facilitate the application of 3-D sound as one of the essential characteristics of the aforementioned film style. Considering the proposed framework of film construction, the immersive continuity style will be put in practice for the creation of an original screenplay and its accompanying storyboard, two documents that shall illustrate how this thesis's screenwriting and preproduction strategies can help on the creation of films with 3-D sound.

Finally, this thesis presents a final discussion in the form of conclusions and areas for future research.

CHAPTER ONE

Cinematic Sound Technology in the Era of Immersive Sound

Bazin argued that "with panchromatic stock in common use, with an understanding of the potentials of the microphone, and with the crane as standard studio equipment, one can really say that since 1930 all the technical requirements for the art of cinema have been available" (1967, p. 30). If such a claim was made based on the technology that was available during that time, today one could argue that the cinema has reached a point of technical perfection, certainly the closest that we have been to what Bazin (1967) denominated 'total cinema'. Film art, as was observed in this thesis's introduction, has constantly evolved in parallel to technological developments, which have gradually move forward towards the 'achievement' of such a total cinematic experience.

Today, the cinema is able to immerse us with sounds that emanate from around and above the audience, creating this way highly realistic—although imaginary—three-dimensional worlds. This chapter aims to present the state of the art in cinematic sound technology. A detailed description of both channel and object-based audio will be provided, followed by a summary of some of the most relevant immersive sound platforms available in the market. Ultimately, given that Dolby Atmos has consolidated its position as the leading solution for the production and delivery of immersive sound, the last section will explore all the technical characteristics of Dolby's latest sonic innovation, which will be complemented by the opinions of a number of sound professionals in regards to the potential and challenges that such a system presents.

Channel-Based vs Object-Based Audio

For many years, the intention of sound technology developers has been to offer an apparatus capable of recreating reality as accurately as possible, that is, their aim has been to give filmmakers the possibility of creating lifelike three-dimensional soundtracks. In present-day's cinema, a sound mix with such characteristics can be delivered as channel feeds (channel-

based audio) or as spatial scene descriptions (object-based audio). Channel-based audio systems deliver each channel of audio to a specific loudspeaker (normally a group of loudspeakers) in the cinema. Object-based audio, on the other hand, defines sound elements—audio objects—with individual metadata that indicates how they will be reproduced during playback (Füg, Marston and Norcross, 2016, pp. 1-2).

Channel-based technology is still the most common way of using audio for cinematic and broadcasting purposes. Traditional digital surround sound formats such as 5.1 and 7.1 are examples of two-dimensional channel-based audio layouts, the X-axis being represented by the front speakers, and the Y-axis by the surround speakers. Dolby Laboratories refer to *Dolby Digital 5.1* as "the established standard for home theater, broadcast, and cinema surround sound" (Dolby, 2020a). Dolby Digital 5.1 integrates five discrete full-range channels and a sixth channel that carries the low-frequency effects (LFE). The '5' full-range channels are connected to specific speakers/arrays distributed in the cinema in the form of left, centre, right, left surround, and right surround, whereas the '.1' channel goes in the form of a subwoofer (twice as loud as the other channels) located in front of the audience. Figure 1.1 illustrates a typical Dolby Digital 5.1 cinematic configuration.

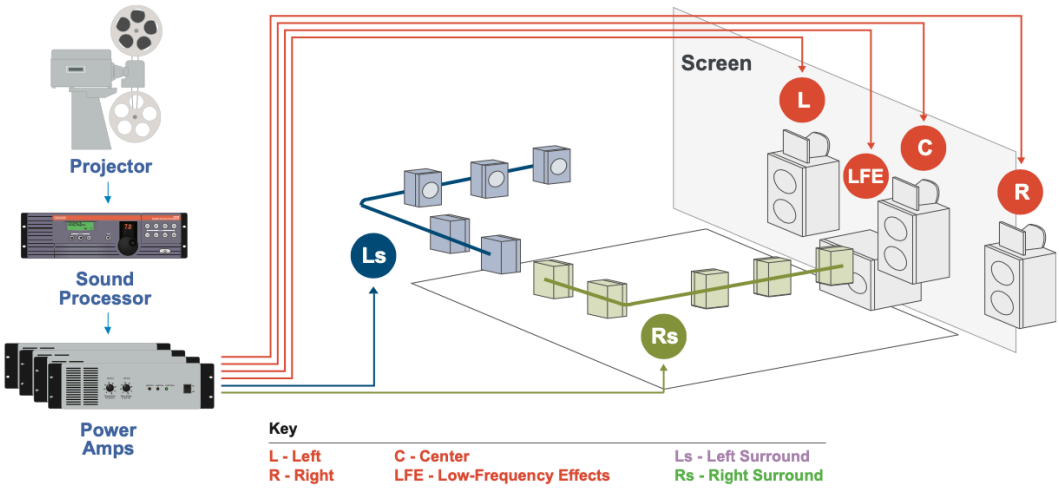


Figure 1.1 – Dolby Digital 5.1 Configuration

With independent channels located in front and around the audience in a 360-degrees manner, Dolby Digital 5.1 facilitated the creation of enveloping soundscapes, a capability that was as well adopted by companies such as DTS and Sony. Further developments were then gradually launched by Dolby and its competitors, establishing the 7.1 format (which divides both surround channels into two) as the upgraded alternative to the 5.1 layout. This, of course,

increased the possibilities of sonic immersion by adding more speakers in the horizontal plane. *Dolby Surround 7.1*, for instance, is a channel-based sound system that allows up to 7.1 discrete channels for a quite realistic 360-degrees sound spatialisation as illustrated in Figure 1.2.

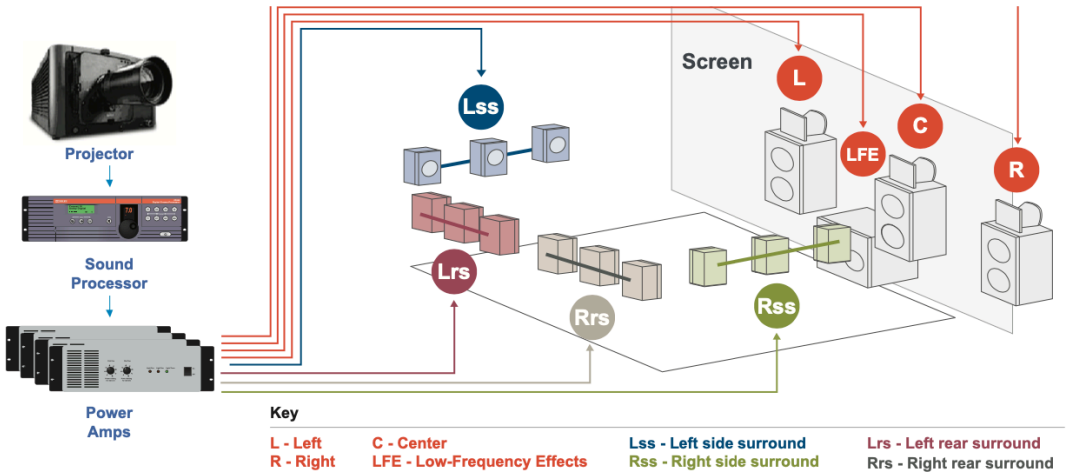


Figure 1.2 – Dolby Digital 7.1 Configuration

Furthermore, in the ongoing quest for improving the cinema's spatial realism, developers have recently introduced more advanced sound platforms that add more loudspeakers around and above the audience, moving from what has been commonly known as surround sound to what is currently referred to as *immersive sound*.

Auro Technologies in partnership with Barco, announced in 2011 a three-dimensional channel-based audio system by the name of *Auro 11.1*. Whereas surround sound systems like Dolby Digital 5.1 are two-dimensional, Auro 11.1 adds the third dimension (height), with a determined number of overhead speakers (Claypool, et al, p. 2). Auro 11.1, the company explains, is based on the Auro-3D concept that compounds three layers of speakers (Figure 1.3). The lower layer, which is designed from the traditional 5.1 format, is intended to reproduce sounds at ear level. The second layer, which consists on speakers located up the walls at about 30-degrees of elevation, is intended to reproduce specific elevated sounds and reflections. Finally, the top layer, with speakers above the audience, is there to reproduce special effects that would come from above, as well as what Auro Technologies calls the "Voice-of-God" (Auro Technologies and Barco, 2015, p. 6).

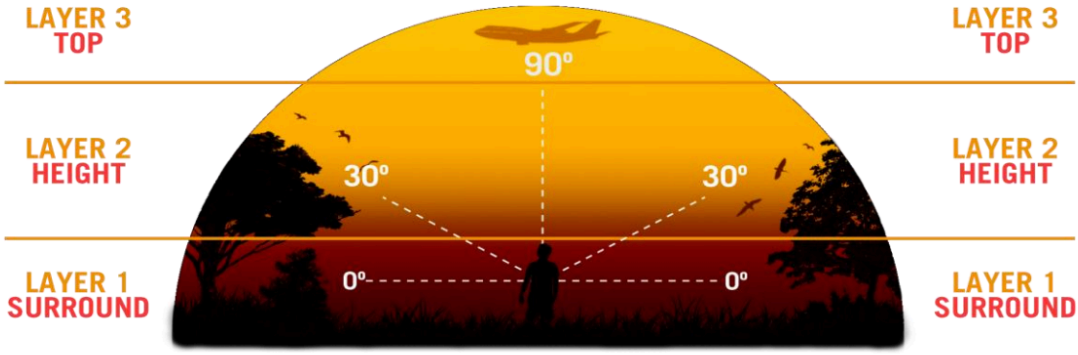


Figure 1.3 – Auro 3D Concept

In January 2012, Auro 11.1 debuted in theatres with Anthony Hemingway's *Red Tails* (2012), a Lucasfilm production in which the sound of airplanes passing around and above us intends to make us feel as being there. Since that year, more that 125 films have been released in this format, and this sound system has been installed or committed in over 550 cinemas around the world (Auro Technologies, 2021).

Although the 11.1 sound format is clearly more immersive than the traditional 5.1 configuration, it is also less practical as its improvements are based on increasing the number of channels, which is one of the limitations of channel-based audio (Lossious and Anderson, 2014, p. 1338). Hence, as a response to the need of a single and open format for immersive sound, three competing companies launched a new generation of immersive sound systems designed under an object-based approach. Dolby Laboratories developed its latest cinematic sound innovation, *Dolby Atmos*, DTS launched *DTS:X,* and Auro Technologies in its alliance with Barco introduced *AuroMax*. In relation to the adoption of object-based technology, Robert Karlsson, Technical Operations Manager at Dolby Europe Limited, points out:

> The feedback that we got was that [channel-based audio] wasn't good enough. So, this is why we had to think differently, and the only way to really make a change was to change the way we describe sound, and how we place sound, and by doing so we can then make it adaptive to any particular configuration you may have. Without that, we would have been stuck in just this channel-based world where people would be saying 13.1, 15.1 or whatever. So, it was critical to move to object-based audio at that point. (Personal interview, 2019)

Regarding the home entertainment sector, Rob France, Head of Home Content Engineering Partner Success and Support at Dolby Europe Limited, argues that the way people consumes entertainment today is different than the way they did when Dolby Digital 5.1 was

developed, which he explains is one of the reasons for moving from channel-based to object-based technology (Media Production Show Conference, 2019). France elaborates:

> When we developed 5.1, there were only a couple of audio systems in the home, there was TVs and there was stereos […] Whereas today headphones are key listening environments, speakers on mobile phones are key listening environments, soundbars were coming through and now, of course, we have smart speakers, and all this different environments that we have to listen to entertainment. (ibid)

Hence, considering the need for a more versatile way of delivering audio content, object-based sound platforms are channel agnostic. That is to say, their technology is not based on a determined number of channels but on audio objects that can adjust to any speaker layout. In channel-based audio, the sound mix had to be made for a specific speaker configuration, such as stereo, 5.1, 7.1, 11.1 etc. On the other hand, in object-based audio, the spatial position of sounds can be defined in relation to speaker-independent 3-D coordinates (Oldfield, Shirley and Spille, 2013). The process of playing back audio objects is based on information that determines the location of each object in relation to the position of the available loudspeakers (Zhang et al, 2017, p. 8-9). Moreover, object-based audio can be *down-mixed* to a variety of formats depending on the sound system that is available, covering the traditional 5.1 and 7.1 to more immersive configurations that include overhead loudspeakers (Oldfield, Shirley and Spille, 2013, p. 2719).

Rob France, explains that in the production side of the chain, there is nothing particularly new about object-based audio. An object, he says, is basically an audio track with some positional metadata (Media Production Show Conference, 2019). France explains:

> For years we have been working in digital audio workstations having a track with some positional data, so there is nothing fundamentally different about that. The only difference now is to actually allow that element to leave the production environment without being mixed into a final channel-based format. All the mixing is done, all the EQ and processing applied, but ultimately, we keep these elements separate. (ibid)

Aiming to standardise the use of these innovations, the Society of Motion Picture and Television Engineers (SMPTE) announced on September 2018 the publication of new SMPTE ST 2098 standards for immersive audio. Brian Vessa, founding chair of SMPTE's Technology Committee on Cinema Sound Systems and executive director of digital audio mastering at Sony Pictures Entertainment, acknowledges that "[i]mmersive audio makes a significant impression on audiences, and because of the real value it adds, we're seeing an increasing number of movies being mixed for the immersive environment" (SMPTE, 2018). Vessa

explains that "[b]y supporting delivery of a standardized immersive audio bitstream within a single interoperable digital cinema package, the new SMPTE immersive audio standards simplify distribution while ensuring that cinemas can confidently play out immersive audio on their choice of compliant immersive sound systems" (ibid). In other words, the standardisation of cinematic immersive sound under an object-based approach allows filmmakers to deliver one mix that will accommodate to the different sound platforms available. On the other end, the SMPTE standard encourages exhibitors to choose the best sound platform according to their needs, preferences and budget, being Dolby Atmos, DTS X and AuroMax the three strongest alternatives.

It is important to notice that although these three immersive sound systems make use of object-based technology, their mixes actually include two types of audio content: beds and objects (Auro Technologies and Barco, 2015, p. 14). The so-called beds are basically surround stems addressed as arrays to be reproduced on a group of loudspeakers (Rumsey, 2013, p. 343), they integrate channel-based content—such as music and atmospheres—that does not require to be treated under an object-based approach. Beds are typically more static and do not take specific positions in the mix, hence they can be rendered as channels or 'static objects' (Auro Technologies and Barco, 2015, p. 15). Hence, beds can be created out of traditional multichannel configurations like 5.1 or 7.1, or based on more immersive formats with overhead speakers like the 7.1.2 layout. For the sake of example, a 7.1.2 bed is a 10-speaker layout, which adds 2 overhead speakers to the common 7.1 setup as explained below:

> [7] Refers to the number of traditional surround speakers (left, center, right, left surround, right surround, back surround left, back surround right).
>
> [.1] Refers to the number of powered subwoofers.
>
> [.2] Refers to the number of overhead—or Dolby *Atmos* enabled speakers (Dolby, 2018c).

Objects, on the other hand, are all those elements that can be precisely localised and moved through the speakers across the auditorium, and hence are individually rendered with metadata that indicates their position for playback, their size, diffusion and other sonic characteristics (Rumsey, 2013, p. 342). Such a hybrid approach is a common factor that describes the way an AuroMax, a DTS:X, or a Dolby Atmos mix work. Each of these systems, however, has its own specificities, which will be described in the following sections.

AuroMax

AuroMax is the upgraded version of the Auro 11.1 immersive sound format developed by Auro Technologies in its collaboration with Barco. In order to align with the latest technological trends and standards on cinematic immersive sound, this platform integrates the flexibility of object-based audio to the three-layered concept of Auro-3D. AuroMax, its creators explain, is specifically designed to work under the above-mentioned hybrid approach that combines channels and objects simultaneously. Hence, the ideal AuroMax workflow consists of creating the channel-based beds of immersive atmospheres first, and then adding all the specific objects to the mix. Objects can then be localised on more specific positions around and above the listener, depending on the number of reproduction channels (speakers and amplifiers) available. AuroMax's speaker configurations are based on the traditional 5.1 digital surround sound layout, which helps to maintain the best possible listening experience and compatibility when content is reproduced on different 5.1-based formats (e.g. 5.1, 7.1, 11.1 and 13.1) (Auro Technologies and Barco, 2015).

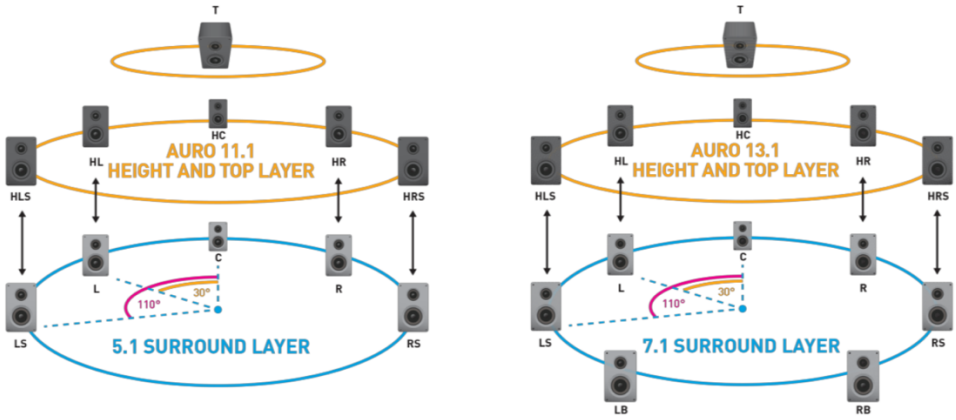


Figure 1.4 – Auro 11.1 configuration (left) and Auro 13.1 configuration (right)

In a native AuroMax mix, objects are thus combined with beds created in either of the Auro-3D formats, which incorporate 12 channels in the 11.1 configuration or 14 channels in the upgraded 13.1 configuration as illustrated in Figure 1.4, leading to 12 or 14 beds respectively. In the cinema, an AuroMax system adds more loudspeakers, but whereas Dolby Atmos addresses each speaker separately (as will be explained later in this chapter), AuroMax treats its loudspeakers as 'zones' that are essentially small groups of speakers distributed across the cinema's walls and ceiling. That is to say, a cinematic AuroMax sound system divides each

channel into smaller groups of individually addressed zones, which reduces the costs of having a separate amplifier for each loudspeaker in the theatre (Auro Technologies and Barco, 2015).

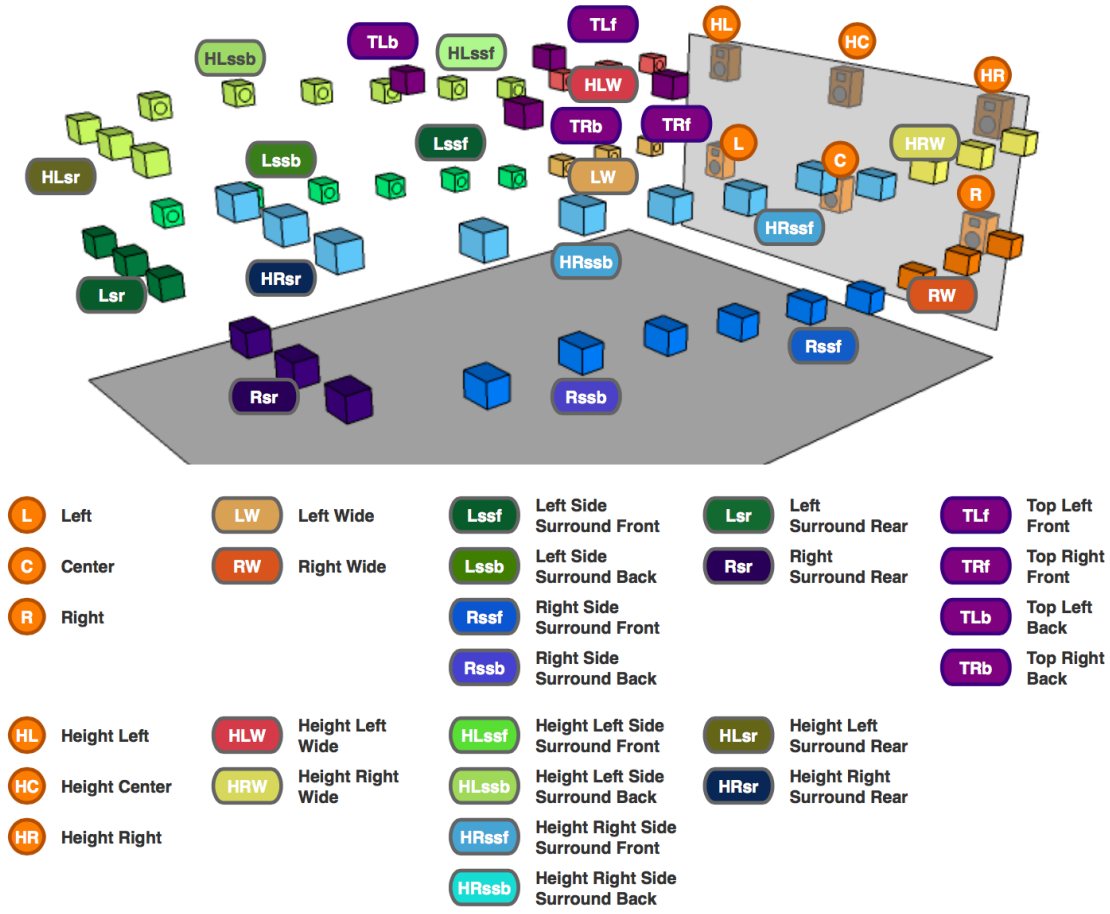


Figure 1.5 – 26.1 AuroMax Configuration

The ideal AuroMax theatrical configuration (26.1) integrates 26 zones or groups of loudspeakers, the surround and height layers are divided into two zones for each wall, and the top layer is divided into 4 zones (Figure 1.5). This configuration adds a set of speakers between the screen channels and the front-most surround speakers, filling the gap that exists between the screen and the surround channels in traditional theatrical configurations. The smallest AuroMax theatrical layout (20.1) is based on the original 11.1 layout and includes 20 different zones, which basically means reducing the top layer to only 2 zones, and not including the speakers between the screen and the front-most surrounds as shown in Figure 1.6 (Auro Technologies and Barco, 2015).

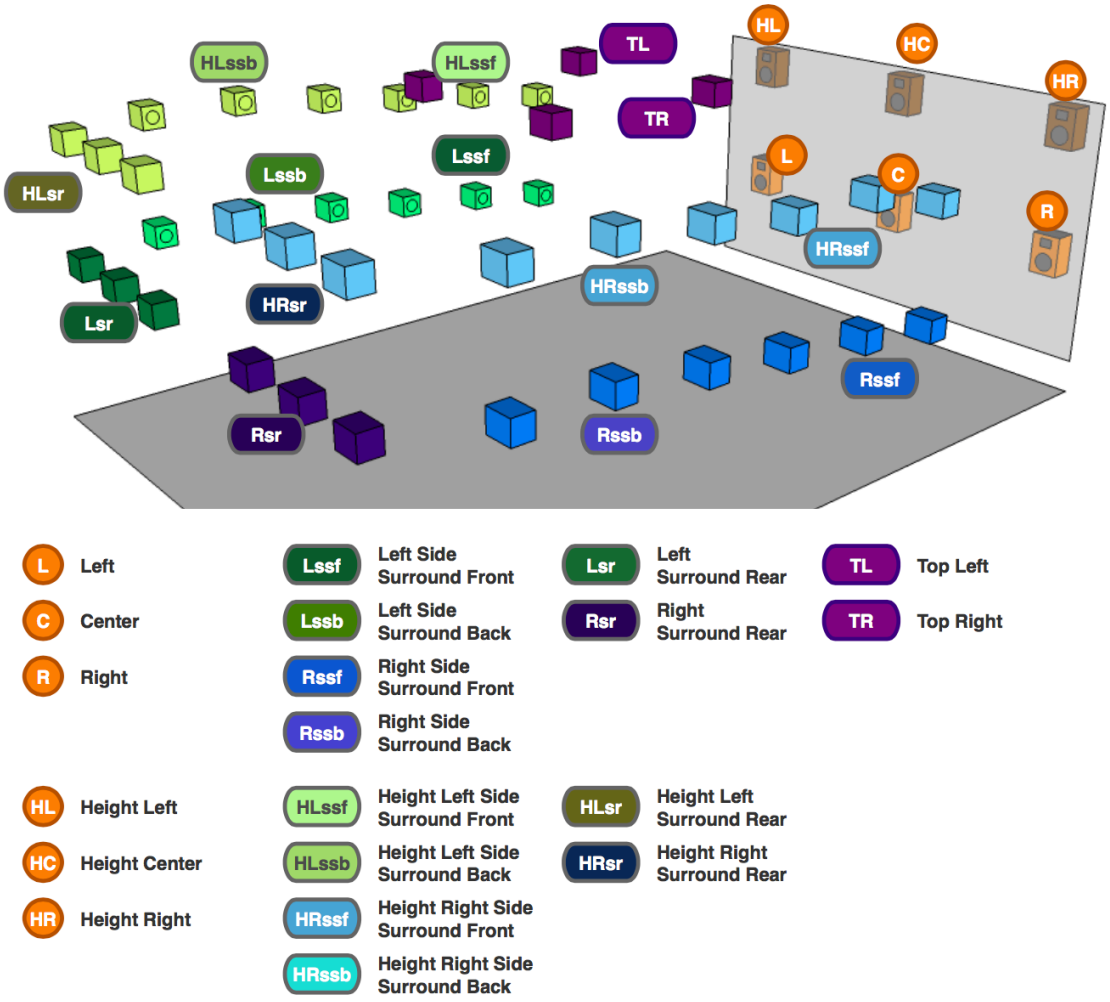


Figure 1.6 – 20.1 AuroMax Configuration

For any AuroMax configuration, their creators recommend using bass management—which basically routes the low frequency information from the full-range loudspeakers to subwoofers capable of reproducing the bass content and low frequency effects (LFE). In order to do this, they suggest installing a subwoofer at each side of the theatre, either on the sidewalls or on the back wall (Auro Technologies and Barco, 2015). This aims to provide an equal level and frequency response to all the loudspeakers in the auditorium, which smooths the transit of sound from screen to room and encourages more aggressive panning practices. Blockbusters such as Pierre Coffin and Kyle Balda's *Minions* (2015), Chris Columbus's *Pixels* (2015), and Baltasar Kormákur's *Everest* (2015) are some of the titles that have been released in AuroMax.

In terms of tools for content creation, AuroMax is designed to natively work with the Auro-3D Creative Tool Suite, an affordable software package that includes a set of plug-ins under the names of Auro-3D Authoring Tools and Auro-Matic Pro. Some of these plug-ins

include 3D panning, encoding, and downmixing functionalities. The AuroMax Digital Cinema Package (DCP)—which includes all the image/audio files and metadata for delivery—includes a 5.1/7.1 surround mix in addition to the AuroMax master file. (Auro Technologies and Barco, 2015).

DTS:X

Following the steps of Auro 11.1 and the demand for more immersive formats with height content, DTS launched in 2012 a new sound system by the name of *DTS Neo:X*. This channel-based system offered up to 11.1 discrete channels, adding front-height and wide speakers to the original 7.1 format (DTS, 2012). After Dolby's debut on the object-based world, however, DTS adapted their technology to this new way of delivering sound, and in 2015 the company launched an object-based platform by the name of DTS:X. Among the specific characteristics of this object-based system, it is worth noticing that DTS:X does not recommend a particular speaker layout. According to Dave Casey, DTS's senior director of product development, DTS:X works "across the whole gamut of speaker configurations" (As cited in Palenchar, 2015, p.8), meaning that the system is thought to be functional on any existing layout. Palenchar observes that DTS aim to position DTS:X as a flexible format, compatible with Dolby Atmos configurations and the capacity of coexisting in Audio/Video Receivers (AVRs) along with Atmos and Auro-3D (2015, p. 8).

Although there are a few cinemas with DTS:X theatrical sound installed, this sound platform is mainly focused on the home entertainment sector. With a variety of AVRs and soundbars with native compatibility, DTS:X has consolidated its position as Atmos' strongest competitor within the consumer market (Cohen, 2019). According to the company's official website, however, there are less than 100 Blu-ray titles featuring DTS:X technology for the home, among which stand out films such as Alex Garland's *Ex Machina* (2014), Guillermo del Toro's *Crimson Peak* (2015), and David Leitch's *Atomic Blonde* (2017). In fact, a big percentage of the DTS:X titles are re-mixes of previously released movies, such as *Independence Day* (1996), *E.T.: The Extra-Terrestrial* (1982), *Apollo 13* (1995), among others (DTS, 2021).

Regarding tools for content creators, DTS offers the DTS:X Creator Suite, a package of plug-ins that allows mixing and delivering of immersive audio content for digital cinema,

Ultra HD Blu-ray, Blu-ray and other premium formats (DTS, 2020b). Compatible with Avid Pro Tools—the industry standard digital audio workstation (DAW)—these tools include mixing, rendering, and monitoring capabilities, allowing a variety of multi-speaker deliverables, as well as the creation of immersive audio content for headphones (ibid). In relation to the later, DTS offers an individualised listening technology by the name of DTS Headphone:X, which intends to simulate 3-D sound on headphones by means of *binaural audio* technology.

In essence, the intention of binaural audio is to replicate the way that humans use both ears in order to hear any given sound field. Each ear receives sound waves on a particular way depending on the physical characteristics of the head, ears, nasal and oral cavities and torso. These differences are then interpreted by our brain, which allows us to localise sound. The way an ear receives a sound from a point in space is determined by what is known in acoustics as head-related transfer functions (HRTFs). In binaural recording and reproduction, HRTFs are measured in order to represent these differences in a mathematical way and thus provide spectral cues to determine sound localization. As Kenny explains, the technology behind DTS Headphone:X involves the manipulation of HRTFs (2013, p. 25). But given the fact that HRTFs tend to be quite broad as they are not specifically design for each individual, it is expected that the 3-D effect on headphones would vary from ear to ear.

Although DTS has concentrated their efforts on conquering the home entertainment space through headphones technology and native compatibility with AVRs, soundbars and others, its all-time rival has strongly invaded not only the home market, but also the theatrical space with a system that, as their creators proclaim, is currently the global leader in immersive audio, Dolby Atmos.

Dolby Atmos

Since its theatrical debut in Pixar's *Brave* (2012), Dolby Atmos has gradually achieved the top position in the field of cinematic immersive sound. Stuart Bowling, Senior Worldwide Technical Marketing Manager at Dolby Laboratories, mentions that Dolby's last sound innovation was developed as a response to visual technological advances. He explains:

> This has been a real evolutionary process […] we were approached by a major exhibitor who was looking to the future and saying, '[w]e are going to start changing the way in which we design and build our auditoriums. With all the changes in imaging technology, something has to happen with audio. What would that look like?' (Fuchs, 2012, as cited in Sergi, 2013, p. 109).

As Gianluca Sergi (2013) observes, Dolby Atmos is defined by its creators as a revolutionary immersive sound system that addresses all key areas of film sound: filmmakers have more tools, distribution is easier for studios, exhibitors get greater 'pull', and audiences enjoy more the cinematic experience. In terms of its technological design, one of the most important factors that Atmos introduced was precisely the utilisation of audio objects, which, as Dolby claims, can virtually be placed anywhere to create a realistic three-dimensional soundscape (Dolby, 2018a).

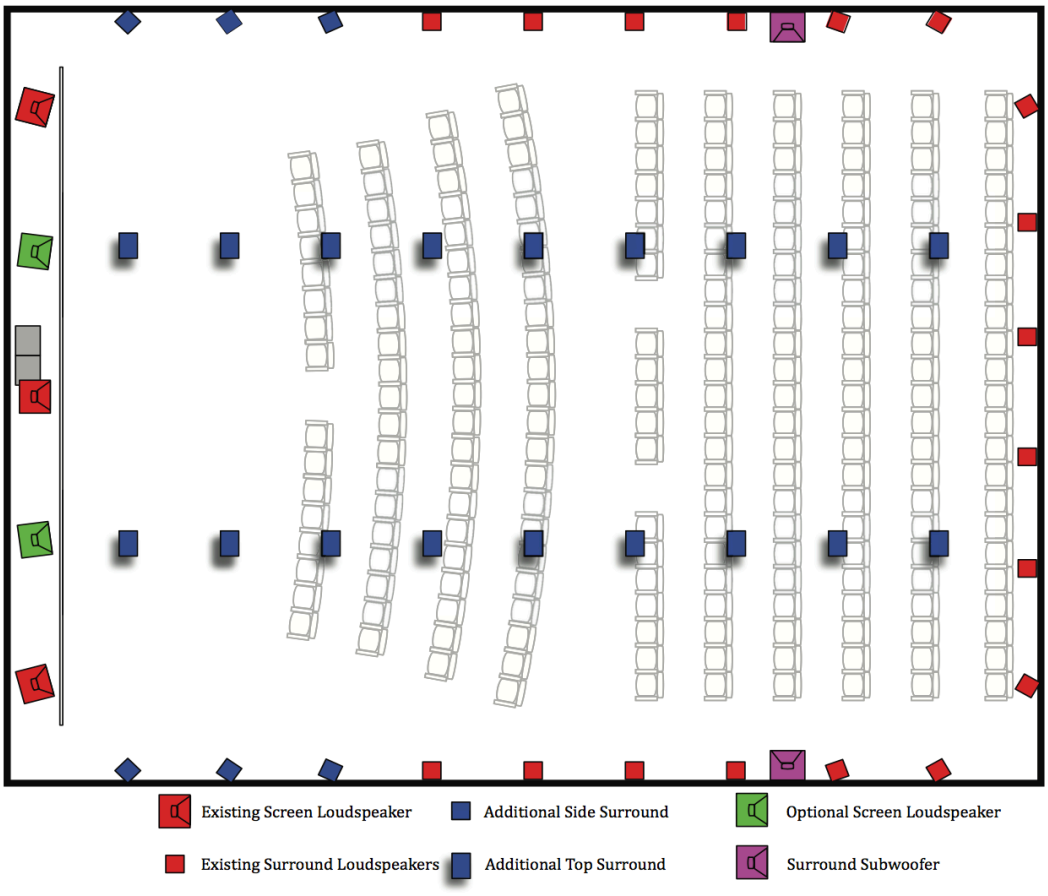


Figure 1.7 – Dolby Atmos Theatrical Configuration

Previous cinematic surround sound arrays forced the soundtrack to be reproduced within a limited number of channels, while in Atmos, in its full configuration, there are up to 64 individually-driven speakers available, each one of them offering full frequency response. Every speaker is powered independently and gets its own separate audio feed, meaning that in

contrast to the traditional channel-based configurations where several speakers are part of the same zone (e.g. left surround zone, right surround zone, etc.), each speaker in an Atmos configuration gets to be its own independent zone. These speakers are not only located horizontally, as Atmos adds a vertical dimension with two overhead speaker arrays. Moreover, Dolby Atmos adds more speakers around the auditorium (Figure 1.7), which provides better localisation capabilities, increased definition and improved audio-visual coherence.

The ability to precisely position sources anywhere in the surround zones also improves the audio-visual transition from screen to room. If a character on the screen looks inside the room toward a sound source, the mixer has the ability to precisely position the sound so that it matches the character's line of sight, and the effect will be consistent throughout the audience. Moreover, Dolby claims that Atmos is their first 'room-centric' sound system, meaning that the 'sweet spot'—the zone in the theatre in which the spectator gets the optimal listening experience—is no longer a concern because sounds get to any member of the audience with equal power, range and directionality (Sergi, 2013, 111).

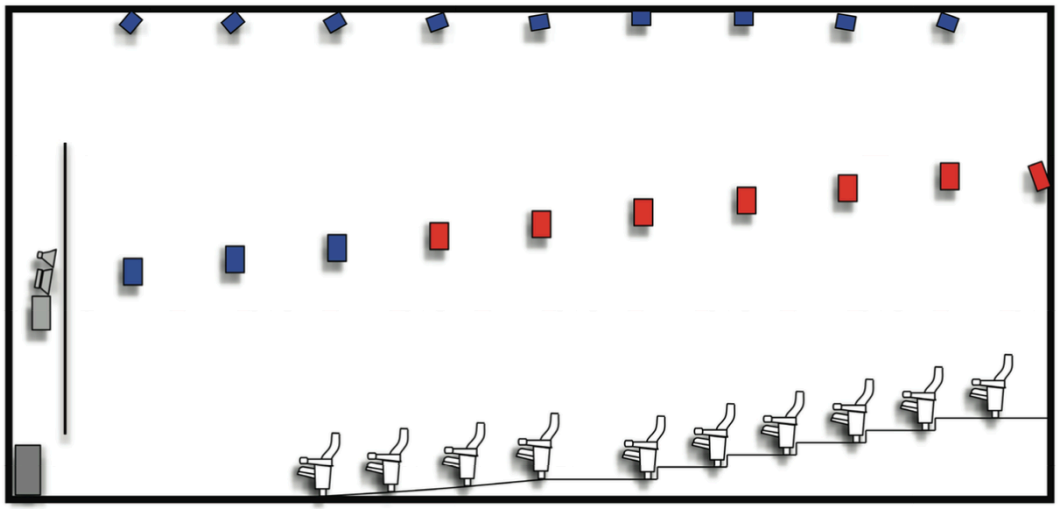


Figure 1.8 – Dolby Atmos wide side speakers along the whole length of the wall

Surround speakers in the traditional theatrical sound systems (e.g. Dolby Digital 5.1) offer a reduced amount of power and frequency response compared to the screen channels. This is clearly a limitation during mixing, as panning sounds from the front to the surround channels leads to notable differences in level and timbre matching. Dolby Atmos addresses this problem with the implementation of identical full frequency-range loudspeakers, improved room equalisation, and bass management. For better results, besides the subwoofer traditionally located behind the screen, Dolby suggests the implementation of two additional subwoofers

located on each side wall. Additionally, to smooth the transit of sounds from screen to room and vice versa, Dolby recommends installing side surround speakers along the whole length of the wall as illustrated in Figure 1.8 (Dolby, 2014).

In summary, Dolby claims that Atmos improves the quality of the cinematic experience based on three fundamental elements: sounds originating overhead, enhanced audio quality and timbre matching, and greater spatial control and resolution (Dolby, 2014).

As previously explained, object-based audio systems like Dolby Atmos make use of beds and objects depending on the characteristics of the sounds. Beds are more adequate for static audio content like ambient sounds and reverberation, whereas objects are more appropriate for dynamic audio elements. Rob France explains that beds "is just another name for channel-based audio, but Dolby Atmos is fundamentally beds plus objects, channel-based audio plus objects, and you can choose to use them whenever you wish" (Media Production Show, 2019). France elaborates:

> We often, in the conversations I have, talk to people who look at Dolby Atmos as just being object-based audio, just being that classical 'I want a helicopter flying overhead', which is a really great use of an object—you have a clean sound that you can position overhead and it moves. But there are a lot of sounds where still channel-based audio works much better for. Things like reverbs, if you want the space of a room or a hall, actually putting reverbs into channels still works a lot better. (ibid)

Karlsson explains that a Dolby Atmos master file has the capacity of packaging 118 audio objects plus a bed—e.g. 118 objects plus a 7.1.2 bed (Personal interview, 2019). The beds are mapped to the speaker or speaker array assigned to each channel, while the objects are rendered in real time and positioned based on the physical location of the loudspeakers available in the auditorium, up to a maximum of 64 loudspeaker outputs (Dolby, 2014). France explains that rendering is the process of mapping beds and objects into the particular speaker or headphone environment that is available. For example, if one wishes to render an Atmos mix into a listening environment that has a 5.1.4 speaker array, the mixer would have to set the renderer up to derive those 5.1.4 channels, which will position each audio source into that particular configuration depending on their positional metadata (France, Media Production Show, 2019). "It is a clever sort of bussing and panning structure ultimately" (ibid).

The key thing, France explains, is that Dolby Atmos uses different renderers that will work for each specific environment, so if physical speakers are available—a 7.1 theatrical layout, for instance—then it would be adequate to use a renderer that will spread the audio

elements in a typical channel-based manner; but if the intention is to deliver the mix onto headphones, then the renderer should be consistent with that individualised listening environment. Ultimately, the renderers "are trying to take the best content you have created and reproduce it as close as they can with ultimately the speakers or the headphones that are available at that time" (ibid). With-height Atmos content can thus be rendered to traditional multichannel formats, which eliminates the need of having different mixes for each deliverable. Dolby Atmos is able to produce automatic re-renders to traditional channel-based formats like stereo, 5.1 and 7.1, which derive from the original Atmos master file (Connor, 2017). France elaborates:

> Within the professional environment I don't just need my one final mix, I may need multiple different elements. I may need a 7.1 version or a 5.1 version. So, within the Dolby Atmos production tools, there is this concept of re-rendering where you can specify 'ok I want a 7.1, what elements of the mix should go into that 7.1 re-render?' […] And it becomes much simpler to map across the different deliverables and create those different elements, because you've got the co-building blocks, and you can just select what you need to come down of that downstream. (Media Production Show Conference, 2019)

Karlsson explains that Dolby Atmos offers mixers the possibility of having a live feedback of a 5.1 or a 7.1 channel-based downmix while they are still mixing in object-based Atmos (Personal interview, 2019). He argues that "by thinking of sound in the 3-D space the quality of the 7.1 and 5.1 mixes that come out of it is higher, because when thinking more about sound in the 3-D space the surrounds just naturally sit better where they want" (ibid). This is confirmed by sound practitioner Enrique Greiner, who notes that a 7.1 downmix from Atmos is much better than a native 7.1 mix. "The way to position things, and how they move, Atmos translates it to 7.1 in a much better way than if you had mixed it directly in 7.1" (personal interview, 2019, translated by author). Similarly, sound practitioner Javier Quesada asserts that the process of downmix from Atmos to 7.1/5.1 "is incredibly transparent". Quesada believes that

> Atmos has achieved this convertibility like no other format, starting from Atmos and going down to 7.1, 5.1, LCR, web. It is part of the virtues of Atmos, you do not have to worry about making various mixes. You do the Atmos mix well and you can be sure that it will translate well. (personal interview, 2019, translated by author)

In the DCP, a Dolby Atmos mix is hence stored along with other rendered deliverables (e.g. Dolby Digital 5.1 or 7.1), allowing the reproduction of different formats when necessary without the need of additional workflow steps for creating the 5.1 or 7.1 versions.

Consequently, if a particular theatre is not equipped with Dolby Atmos, for instance, compatibility is still assured as the DCP contains a rendered 5.1 (or 7.1) version of the mix. Furthermore, the Dolby Atmos master file can be used to create other deliverables (e.g. a stereo mix) (Dolby, 2014).

And besides its strong engagement with cinematic theatrical reproduction, Dolby Atmos has also innovated the home entertainment industry with Atmos-enabled products such as AVRs, soundbars, mobile phones, tablets, games, PCs, headphones and more. Karlsson explains:

> The Dolby Atmos experience would be dependent in terms of your playback format, in terms of the re-monitoring, on how you set up your system [...] Over headphones what we are aiming to do is provide the most immersive headphone experience that we can, just like we are trying to provide the most immersive soundbar experience in comparison to other soundbars in the market that don't have Dolby Atmos. So, Dolby Atmos in the cinema is one level of spatial resolution and detail, with speakers up to the number of 64 depending on the coverage, and then in the home it can be up to something like 32, but typically we are seeing 7.1.4 systems. And then in the home that is going down to 5.1.2 to this soundbar configurations and then to the headphone setup. So, each one of those products is slyly different in terms of replay, but what mixing in Dolby Atmos allows you to do is optimise the content best for that headphone, for that soundbar, for that 7.1.4 room to get the most immersive experience you can. So, the idea behind Dolby Atmos over headphones is providing an Atmos-based mix, you can then create the most effective sort of 3D immersive audio experience. (Personal interview, 2019)

Aiming to become the standard for home entertainment, Dolby Atmos is now available for digital media players such as Apple TV or Amazon Fire TV, and within video-on-demand services such as Netflix, Amazon Prime, Vudu, and iTunes (Dolby, 2018a). Ultimately, says France, "the aim of Dolby is about bringing immersive audio everywhere […] So we want immersive audio to be there on whatever platform the consumers listen on" (Personal interview, 2019). In June 2017, for instance, Netflix released Bong Joon-ho's *Okja* (2017), its first Atmos-enabled content; and nowadays the streaming giant offers a variety of original and third-party films and TV shows with Dolby Atmos soundtracks—e.g. *Guardians of the Galaxy 2* (2017), *Mowgli: Legend of the Jungle* (2018), *Roma* (2018), *Extraction* (2020) and others.

In this regard, Rob France acknowledges that in the consumer space, the vast majority of people that have a Dolby Atmos enabled system at home do not install speakers in the ceiling (Media Production Show, 2019). That market, he says, "is less than 1% of the potential costumers out there" (ibid). Therefore, Dolby had to develop other technologies that permit to

bring the immersive experience to the home. One of these solutions are the so-called 'up-firing'[15] speakers, which aim to simulate the existence of overhead elements by bouncing the sound waves off the ceiling as illustrated in Figure 1.9. This technology works through a combination of *psychoacoustic signal processing* and *speaker directivity and angling* (Dolby, 2016), the former being conducted by the modification of certain frequencies through filtering, while the latter being achieved by firing the acoustic energy upward (ibid). France asserts that "the vast majority of consumers listening to Dolby Atmos today get it through a system that is reflecting sound off the ceiling" (Media Production Show, 2019).

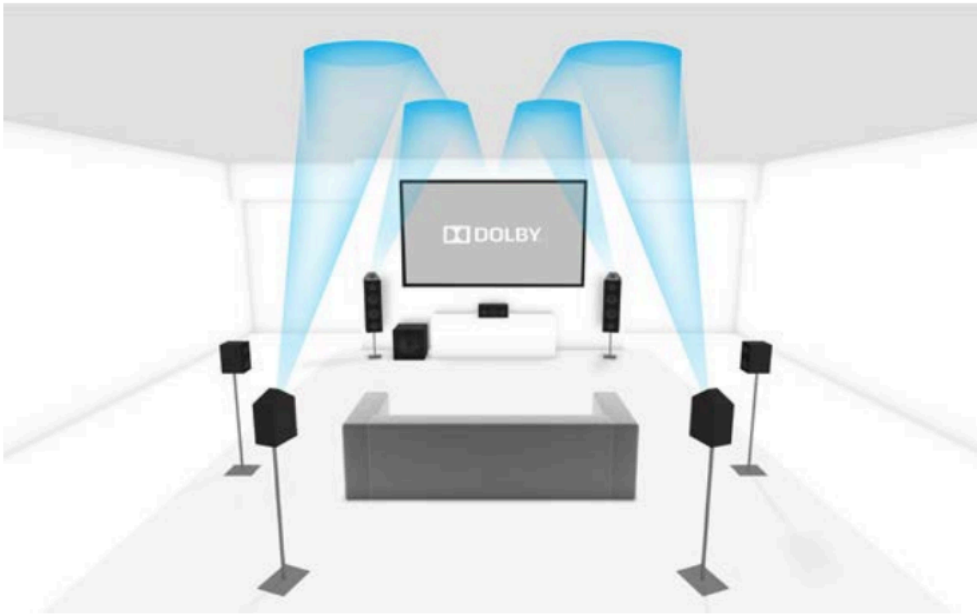


Figure 1.9 – Dolby Atmos 'Up-Firing' Speakers

An alternative way of bringing the Dolby Atmos experience to the consumer environment is related to the process of *virtualisation*. This technology is grounded on the application of height cue filters, HRTFs and cross-talk cancellation, which combined aim to create the illusion of sound emanating from above and around the listener even through a pair of headphones. In headphones, "it does not sound as good as a soundbar", France admits, "but our aim with all of these systems is always to make the sound substantially better than a stereo experience" (Media Production Show, 2019). Ultimately, although the cinema is the most appropriate space for experiencing the immersive capabilities of Dolby Atmos, nowadays one can playback an Atmos title using home theatre systems, soundbars, TVs, or simply by plugging a pair of headphones.

[15] Also referred to as Dolby Atmos enabled speakers.

Furthermore, Dolby Atmos' growth has forced the development of new immersive tools that allow an efficient connectivity for content creators. In this respect, the 12.x and later versions of Avid Pro Tools enhance Dolby Atmos integration (Avid, 2018). Pro Tools | Ultimate allows working with different Dolby Atmos speaker configurations for sound design and mixing, supporting stem formats like 7.1.2, as well as built-in rendering technology (Avid, 2018). Similarly, Dolby Atmos supports native panning integration with Nuendo. In addition, Dolby has released affordable tools for content creators such as the *Dolby* Atmos *Mastering Suite*, which allows delivering Atmos content for cinema, Blu-ray and OTT (over-the-top) media services such as Netflix; and the *Dolby Atmos Production Suite*, which allows to edit and premix three-dimensional Atmos sessions, even when using a laptop with Pro Tools and a pair of headphones for binaural monitoring (Dolby, 2018a). The Dolby Atmos Production Suite can also be used to deliver Atmos mixes for OTT productions, meaning that now sound designers and mixers can create and deliver an Atmos mix for platforms such as Netflix for a very affordable price.

All these capabilities have encouraged the adoption of Dolby Atmos, which is gradually becoming the new standard of cinematic sound mixing and reproduction. To date[16], over 1785 features have been released (or committed) in Dolby Atmos, there are more than 6000 Dolby Atmos screens installed (or committed) within more than 90 countries, over 65 multiplex offering Dolby Atmos theatrical exhibition, and over 230 Atmos-equipped mixing facilities (Dolby, 2021), which includes all US major studios and most leading postproduction-houses around the world (Sergi, 2013). In addition, Dolby asserts that 79% of costumers prefer an Atmos enabled theatre and a stunning 91% have reported being likely to return for a Dolby Atmos experience (Dolby, 2019). Matt Cuson , former Senior Director of Product Marketing, Cinema Products at Dolby Laboratories, highlights that "[t]he long-term goal […] is for (Atmos to be in) 'every movie and every theatre'" (Fuchs, 2012, as cited in Sergi, 2013, p. 108). Karlsson asserts that

> the rate of adoption of Dolby Atmos has been faster than Dolby Digital ever was in 1991 onwards. And that's part because of Dolby's experience in working with the content creators, like Alfonso Cuarón and others, the exhibitors, the cinemas themselves and what they need and how they can do it, and how we distribute these mixes everywhere […] The cinema is becoming, for Dolby Atmos, quite a mature

[16] As of June 10, 2020

> market, it is not a question now of is it going to last?, it is more a question of when are you going to put it in? (Personal interview, 2019).

For *Roma*'s (2018) sound designer and supervising sound editor, Sergio Diaz, Atmos is "the ideal format for all types of films"; he observes that Atmos "should become the universal platform" (personal interview, 2019, translated by author). With Atmos, Diaz asserts, "you really go into the heart of each moment, the only thing you do is enriching the scene more and strengthen the film through immersive sound" (ibid). Academy Award winner Skip Lievsay, who has worked as sound designer/supervisor and re-recording mixer for filmmakers like the Coen Brothers, Darren Aronofsky, Alfonso Cuarón, and others, comments:

> I think Dolby Atmos is a really great format because it has a kind of unlimited potential. You can do minimal Dolby Atmos movement and panning, and you enjoy increased fidelity in theatre. Many more high-quality channels than normal 7.1 or 5.1, and the potential to really move stuff around and get the soundtrack moving and the excitement that that creates I think is a really breakthrough for cinema sound. (personal interview, 2019)

Cuarón and the Exploitation of Dolby Atmos

Undoubtedly, "[a] powerful weapon such as *Atmos* can create a large disruption, hence a substantial window of opportunity for change" (Sergi, 2013, p. 118). Sergi (2013) suggests that Dolby Atmos is a good reason to start changing the way sound and image departments collaborate. In order to fully exploit Atmos' cinematic capabilities, this immersive sound platform challenges filmmakers to adopt a sound design approach that integrates with the image earlier in the production process. Sound practitioner Emma Butt comments:

> I think [Dolby Atmos] is giving us a chance to get directors to see there is more ways of using sound to help enhance their film, and I think that's only a good thing. The more they understand about it and the more they try things out and they experiment with it, the better because it gives Dolby Atmos the chance to grow and become bigger and become basically the new 5.1, become widely used by all of us. (Personal interview, 2019).

Indeed, as Karlsson asserts, "the more people understand about Dolby Atmos, the more they can then engage on how they want to do the production part of the project" (Personal interview, 2019). In this regard, Academy Award winner sound designer/supervisor and re-recording mixer Randy Thom comments:

> I think [Dolby Atmos] is potentially a great tool, I think that we need to figure out how to use it well […] [*Gravity*] is a great example of a first really solid first step in

> terms of using it well and there have been several other films that I think have used it well. But I think in order to take full advantage of it you need to write for it, to design a scene for it […] But unfortunately the way it is now, just like with the rest of sound, we do not have a plan like that, we do not start out with a plan and everything just kind of falls out the way it falls out and very often it does not make any sense of visually to try to impose that kind of soundscape. And so, your hands are tied sonically because the scene has not been designed for it […] We need to find ways to at least simulate that kind of sonic environment as early as we can in the process (Thom, 2015 Interview by SoundWorks Collection).

In previous channel-based formats made out of arrays of surround speakers, putting dialogue in the surrounds meant that it was actually going to sound coming out of a group of speakers. This, Karlsson comments, "never felt natural because when I'm talking to you it is a discrete sound from one acoustic source. Whereas, in the cinema there were […] seven or eight of these acoustic sources talking to you" (Personal interview, 2019). Nowadays, with Dolby Atmos and its object-based technology, "you can discreetly move a voice around the room and it still sounds like this point source effect" (ibid). Therefore, according to Karlsson, Dolby Atmos opens the possibility of breaking the screen-centric convention, and using dialogue in a better way (ibid). Yet, in relation to the panning of dialogues, sound practitioner Emma Butt argue that:

> directors don't understand [Atmos] yet, and that's a big problem. So, in traditional 5.1 mixes dialogue is always centred and it doesn't move. You can move it slightly if the characters move from left to right onscreen, but it is very rare that your dialogue will ever come to the side or back speakers, it is really unusual. And as soon as you do that to a director and you play it to them, if you move their dialogue anywhere else in the room, they don't understand it. They don't understand why they are not hearing it right in front of them. And I think until directors understand sound more, we are not going to see enough creativity happening with Dolby Atmos. (Personal interview, 2019).

Re-recording mixer Dan Johnson comments:

> Traditionally the dialogue in 5.1 is sat in the centre speaker. I guess because partly in 5.1 the surrounds are more of a diffuse field, so they were only a slight kind of upgrade on the Dolby Stereo technology […] The surround speakers were generally not as good as the screen speakers. They are smaller, they are hidden somewhere, they are behind something, they have single drivers, they haven't got base, they haven't got top treble. If you panned dialogue to the back it sounded worse in 5.1 and in Dolby Stereo, it didn't sound very good […] So I guess mixers were pretty couscous in 5.1 with what we panned in the surrounds […] But I guess with Dolby Atmos now is more full-range, it is that proper 360-degree pan, you've got bass management of the surrounds, and you've got much better quality speakers. There is not that kind of technical reason for not doing it. (Personal interview, 2019).

In relation to the potential impact that Dolby Atmos can have on filmmaking, Randy Thom comments:

> I think Dolby Atmos does have the potential to change the ways films are made. I wish more film sequences were designed by the director with Dolby Atmos in mind. I think that's the way that Atmos would really be able to make a change. As it is now, I think very few directors think in terms of Atmos when they are writing or shooting their films, it's something they delay thinking about until postproduction. (personal interview, 2019a)

Ceirtainly, Dolby Atmos and the latest advances on spatial sound mixing and reproduction can facilitate, and in fact, in certain cases, have facilitated the creation of truly immersive narrative worlds; yet, as Kerins claims, that big potential is only feasible when filmmakers allocate enough time and money to that end (2015, p. 147). In this regard, sound practitioner Enrique Greiner comments that if you want to use Atmos effectively you have to do it since the preparation that has to take place before the film is shot, and not only for the final mix.

> If you don't do that, it's a bit like what happened with 3-D movies. There were movies that were shot in 3-D because they shot them with two cameras, but then they found a computer algorithm that did that for you also with movies that were shot in 2-D, using only one camera. So, they did a sort of trick and returned them to you as a kind of weird 3-D. Doing Atmos after a movie has already been shot is like that. (Greiner, personal interview, 2019, translated by Author)

Dolby Consultant Robert Karlsson explains that some of the best Atmos mixes that they have got have been the ones that have taken Dolby Atmos as the core mixing structure, from which all other deliverables are derived. "They follow Dolby Atmos from the shoot, maybe with the long shots to mixing it from scratch, doing most sound design and more detail using sound objects" (personal interview, 2019). Karlsson, who was in charge of providing technical support for the pre-mastering of *Roma* (2018) and a variety of Dolby Atmos films, acknowledges the work of Alfonso Cuarón as being exemplar for its effective utilisation of object-based technology. He comments:

> [Cuarón] has always been a huge fan of Dolby Atmos. So *Roma* has a very large number of objects in the mix, *Gravity*, they are two examples [of effective Dolby Atmos mixes] […] There are some films perhaps which have used a lot of sound objects, but because they are bombarding you a high level, in high density, there is that possibility that they are not going to be able to resolve or hear the detail of them […], you may have been better off just using the bed. So, it is about knowing when to use them in relation to how we hear and localise sound. […] The reason why *Roma* works so well at times is because there is no music and so it very much feels like a very big soundscape piece

> […] So every little sound you hear, and the detail is amazing. There would be other films where if there is a driving musical score throughout the whole shot, then you probably won't be able to hear that detail. (Personal interview, 2019)

In a similar way, when asked about the films that have actually used the full capabilities of Dolby Atmos, Dolby Specialist Rob France replied:

> The one that always stands out to me is always *Gravity*, it was one of the first movies that was done in Dolby Atmos […] None of us know how space sounds like, so they had the total flexibility to come out with their idea of what space sounded like. But having the individual voices coming out from different places in the surrounds, starting to move around, and particularly when they start to get hit in the sort of boom arm […] that swings around in the air, it is a really compelling experience. (Media Production Show Conference, 2019)

Indeed, in his two most recent films, Cuarón was able to create an effective 3-D soundtrack because he planned his movies with three-dimensional sound in mind, considering the film's fictional setting as an important carrier of narrative information. In relation to Cuarón and his work, re-recording mixer Dan Johnson comments:

> He is not just a filmmaker who gets a script and thinks 'how am I going to shoot it? What is it going to look like?' He is like, 'how is it going to sound? How am I going to use my sounds?' What may work with sound that means that I can cut these lines from the script or whatever it might be. He is always thinking about it from the beginning […] In a lot of productions it fills like they get to almost right at the end and they go 'right, what are we going to do with the sounds?'

The truth of the matter, however, is that the vast majority of films that have been mixed on Dolby Atmos do not exploit its spatial capabilities to the full, and only a few filmmakers have truly experimented with a complete sound three-dimensionality. Kerins notes that a lot of filmmakers acknowledge the greater creative possibilities offered by the new immersive sound formats, and recalls Cuarón's appreciation for Dolby Atmos (2015, p. 115). Yet, Kerins points out that not all of them have shown interest on making significant changes to the ways they approach sound design in their films (ibid). Karlsson acknowledges that *Gravity* and *Roma* did a great job by exploiting the full capabilities of Dolby's last innovation, and in his opinion, there will probably be more experimentation to come with filmmakers trying to do new things. Yet, "we are still coming from a world where […] dialogue off the screen speakers is very rare […] People don't want to take risks sometimes" (ibid). For sound practitioner Javier Quesada, Dolby Atmos has a great potential, and "it's just a matter of directors and producers being encouraged to break conventions, and instead of wanting to please fans who are already used

to something, betting on taking audiences to new terrains" (Personal interview, 2019, translated by author).

Indeed, in order for Dolby Atmos to be completely exploited, those long-lasting conventions need to be broken. Ultimately, as Karlsson comments, "people are still figuring it out" (Personal interview, 2019). One of the challenges of this thesis is hence to discover the extent to which Dolby Atmos has aesthetically changed the way filmmakers use sound design, and beyond, the extent to which those changes are giving or could give rise to new directorial, cinematographic, and editorial styles.

Conclusion

Although channel-based technology is still the most common way to deliver audio in the cinema and the home entertainment space, it seems clear that object-based technology is increasingly becoming the new standard for mixing and playing back immersive audio content. The hybrid functionality of the new sound platforms asks for production and postproduction workflows that integrate beds (channels) and objects within the same mix, exploiting this way the capabilities of each model to the full. The recently published SMPTE ST 2098 standards for immersive audio are an indicator that the delivering of audio content for the future goes on that direction, no matter what platform is used, the same master file would be compatible with all the object-based playback platforms available.

Even though AuroMax and DTS:X are still in the market, Dolby Atmos has consolidated its position as the global leader in immersive audio, a position that seems to be difficult to overmatch. With more than 6000 screens installed or committed, Atmos has become the preferred alternative for delivering cinematic immersive sound. Moreover, its increasing compatibility with streaming platforms like Netflix, Amazon Prime, iTunes and others, Dolby has made a giant step within the home entertainment sector, obtaining the top position in a market once dominated by DTS.

Dolby Atmos is undoubtedly a great tool, but in order to exploit its immersive capabilities to the full, content creators should really understand its potential applications and uses. The standardisation of immersive sound in the form of Dolby Atmos asks for innovative audio-visual aesthetics. Dolby Atmos differentiates from previous channel-based formats in

many ways, but its ability to localise sounds in specific positions around and above the listener, and thus to enhance the possibilities of creating realistic three-dimensional narrative worlds, is perhaps the most transcendent one in terms of aesthetic choices. Accordingly, the scope of the observation that shall be carried out as part of this thesis shall particularly focus on the liberation of sounds (especially dialogue) from the long-lasting screen-centric tradition.

Given that Alfonso Cuarón has proven himself to be a pioneer on the three-dimensional utilisation of present-day's immersive audio tools, his film style will be used as a case study to help this thesis answer its research questions. To this end, an in-depth exploration of his audio-visual aesthetics and production methods shall be carried in the chapters to come. Moreover, aiming to discover the extent to which Dolby Atmos has inspired a broader application of three-dimensional sound design, a number of Dolby Atmos releases will be analysed in terms of the utilisation of the sonic space.

CHAPTER TWO

3-D Sound, Illusion and Transportation: A Theoretical Framework

The rise of immersive sound and the proliferation of Dolby Atmos and other immersive sonic platforms have established new creative possibilities within the cinema industry. The cinematic apparatus that once was studied and theorised based mainly in its visual elements (i.e. the camera, the projector, the screen), today also covers a series of sonic devices that have the power to transform the cinematic experience into an illusory three-dimensional event. Certainly, the cinema in the era of immersive sound presents new theoretical and practical challenges that shall be embraced in order to parallel academia and creativity with the technological innovations of present-day.

In order to introduce the key concept of illusion in cinema, this chapter will begin with a summary of Jean-Louis Baudry's *apparatus theory*, which in the 1970s opened the debate around the effects that films might produce on the spectator through the impression of reality characteristic of the medium. Moreover, apart from explaining cinematic illusionism through the frame of psychoanalysis—a key intellectual area that together with Marxism gave shape to the structuralist apparatus theory—this chapter will explore other rather rational explanations for the experience of illusion in film. Ultimately, given that such theories fail to deeply explore the effects of 3-D sound on the enhancement of the impression of reality, this chapter provides a theoretical introduction to what this thesis terms *immersive point-of-audition*, a sound design strategy that, as shall be argued, may facilitate the process of transportation into narrative worlds, a hypothesis that is approached through Green and Brock's (2000) *transportation theory*.

Illusion in Film: The Cinema's Impression of Reality

Following Bazin (1967), who in his essay 'The Myth of Total Cinema' argues that those who invented the cinema imagined it as an illusory medium, capable of creating the impression

of an inhabitable lifelike world, contemporary film theory offers new perspectives in relation to what could be called the *cinematic impression of reality*. One of the most significant contributions is the work of psychoanalytic film philosopher Jean-Louis Baudry, who in his essays: 'Ideological Effects of the Basic Cinematographic Apparatus' (1974-5 [1070]), and 'The Apparatus: Metapsychological Approaches to the Impression of Reality in Cinema' (1976 [1975]), presents a big part of what is referred to as *apparatus theory*. Baudry wrote his work during a period of political confrontations, which in France reached its peak during the events of May 1968, when a series of demonstrations and strikes against capitalism, consumerism, and American imperialism took place. Apparatus theory stands out, therefore, as one of the products of the Cold War and the renaissance of Marxism in the Western European arena, which integrated with psychoanalysis within French structuralism.

Baudry developed his theory based on Louis Althusser's approach to ideology, a concept that he understood as "the system of the ideas and representations which dominate the mind of a man or a social group" (Althusser, 1984 [1971], p. 120). According to Baudry, the cinema is an instrument of ideology because of its power to create the *impression of reality*, which he claims could be used to impose dominant ideas in the minds of the spectators. Baudry argues that while being exposed to a narrative in the dark and isolated space of a movie theatre, the *passive* spectator adopts the perspective of the camera as their own, identifying with it and with the characters portrayed by it (1974-5, p.45). Grounding his theory on psychoanalysis, Baudry suggests that when we watch a movie, we actually see ourselves in it and become part of that mediated reality (ibid).

Baudry describes the cinematographic apparatus as a set of operations—executed by a set of instruments—that take place throughout the filmmaking process. He argues that the camera occupies an intermediate position between the raw material ('objective reality') and the projection of the finished product; and distinguishes the processes of '*decoupage'* (shot breakdown before shooting) and '*montage'* (editing, or final assembly) as part of the operations that take place before and after the camera, as the site of the inscription, captures the raw material. Ultimately, Baudry notes that between the finished product and its consumption, the projector and the screen restore the light and the movement seized from 'objective reality' (1974-5, p.40).

Hence, Baudry argues that the camera occupies a central position in the cinematographic apparatus, as it allows the construction of an image that is analogous to the

perspective projections typical of the Italian Renaissance[17] (1974-5, p.41). Through monocular vision, the subject is placed in a determined position, a position that is established given that all the objects in the frame are seen and organised in relation to a fixed point (*perspectiva artificialis*) what Chion refers to as the *point-of-view* (1994 [1990], p. 89). Moreover, in consonance with Bazin's (1967) ideas, who had argued that the cinema adds to the realism of the photographic image the extra dimension of movement, Baudry highlights that the cinema projector puts together a series of fixed images captured by the movie camera, and hence restore the continuity of time and movement (1974-5, p.42).

This argument was also supported by Christian Metz (1974), another major contributor to apparatus theory, who drawing on the work of Roland Barthes highlights that unlike the still photography, the movement of the cinematic image is always perceived by the spectator as being present. Metz argues that "the movie spectator is absorbed, not by a 'has been there,' but by a sense of '[t]here it is'" (ibid, p. 6). As such, Metz notes that movement adds to the cinema a higher degree of reality, it gives volume and corporality to objects, and induces the spectator to see the images as being present. In between the lack of reality of the still photography and the 'too real' characteristics of the theatre, film is for Metz the optimal point for creating the impression of reality. In addition, the movie camera is not only capable of recording movement, but as Baudry (1974-5) notes, it can move and through its movement provide meaning.

Baudry then goes on to argue that the cinematic experience presents the conditions that are necessary for the reconstruction of Lacan's *mirror stage*[18]. Baudry explains that such a psychological phase, which according to Lacan (1966 [1949]) occurs between the ages of six and eighteen months, permits the identification of the self, the 'I', by means of seeing one's own body reflected on the mirror. Arguing that body immobility (stillness) and the predominance of the visual function are two conditions that happen during early childhood and upon cinematic spectatorship, Baudry uses the mirror as an analogy to claim that the 'reality' represented by the cinema is taken by the spectator as that of a 'self'.

But given that the image projected on the screen is not that of the spectator's own body, Baudry argues that cinema permits two levels of identification. The first level—what Metz

[17] The Italian Renaissance is a period in Italian history that took place in the 15th and 16th centuries. In Renaissance art, the illusion of three-dimensional space and realism was of central importance.

[18] See Lacan (1966 [1949]) 'The mirror Stage as Formative of the *I* Function as Revealed in Psychoanalytic Experience'.

(1982) refers to as secondary cinematic identification—is derived from the character who works as the centre of identification. The second level—what Metz (1982) refers to as primary cinematic identification—facilitates the occurrence of the first by taking the 'body' of the camera as their own. Ultimately, Baudry's theory suggests that through the relation between the camera and the subject, the basic apparatus of the cinematographic process transforms the spectator into a "*transcendental subject*" (1974-5, p.46), that is some sort of invisible spectator that experiences the events from an illusory central position within the filmic space. As such, Baudry sees the cinema as an ideological apparatus that works in service of the ruling class, and thus facilitates the maintenance of the dominant ideology.

In his second essay, Baudry (1976 [1975]) draws on Plato's allegory of the cave[19]. He argues that the spectators' immobility is a characteristic of the cinematographic apparatus, something that he claims is similar to the prisoners' immobility in the cave, which Plato describes as follows:

> Imagine the condition of men living in a sort of cavernous chamber underground, with an entrance open to the light and a long passage all down the cave. Here they have been from childhood, chained by the leg and also by the neck, so that they cannot move and can only see what is in front of them, because the chains will not let them turn their heads. At some distance higher up is the light of a fire burning behind them; and between the prisoners and the fire is a track, with a parapet built along it, like the screen at a puppet show, which hides the performers while they show their puppets over the top (Translated by Cornford, 1976, p.103)

Plato suggests that those prisoners would not have been able to see themselves or one another, but only the shadows projected by the fire-light on the wall in front of them; and that if they were able to speak to one another, they would assume that the voices came from the shadows that they saw. Plato continues his allegory supposing that the cave produces an echo that comes from the wall facing the prisoners, hence if there were people speaking as they cross behind them, the prisoners would only be able to suppose that those voices originate from the shadows moving in front of their eyes. Ultimately, Plato concludes that the prisoners in the cave would only be able to recognise as reality "the shadows of those artificial objects" (ibid). In Baudry's analysis, the prisoners in Plato's cave are victims of "an illusion", of a "hallucination", of a "dream". The prisoner is "the prey of an impression, of an *impression of reality*" (1976, p. 107 emphasis in original).

[19] See Plato, 'The Republic'

Drawing on Freud's psychoanalytic theories, Baudry then elaborates his hypothesis by comparing the impression of reality achieved by the cinema with the impression of reality that we experience in our dreams. He notes that in the darkness of the movie theatre, the spectator remains relatively passive and immobile watching a series of moving images that are reproduced in front of them. Under such conditions, Baudry argues that the cinematographic apparatus produces "a state of artificial regression", a state in which the subject may find satisfaction explaining this way their desire for cinema and the pleasure they find in it (1976, p. 119).

In his essay 'The Imaginary Signifier', Christian Metz (1982 [1975]) contributes to the strengthening of apparatus theory by connecting it to other psychoanalytical conceptions. Building upon Baudry's approach to the Lacanian concept of the mirror stage, Metz argues that the screen in the cinema is similar to a mirror, but not precisely the one in the mirror stage; for unlike the mirror projection, the spectator's own body will never be projected on the cinema screen. Hence according to Metz, the cinematic experience is like "a series of mirror-effects organised in a chain" (1982, p.51), a comparison that he explains in the basis of two concepts: *presence* and *absence*.

He explains that in theatre we find real objects (presence) that are invoking fictional ones (absence). A chair on the stage during a theatre play, for instance, is trying to represent a chair in a fictional world. In the cinema, on the other hand, we find fictional objects in the form of a projection (absence) that at the same time are invoking fictional objects that belong to a fictional world (absence). In that sense, Metz argues that what is at play in the cinema is not a presence that is invoking an absence (like the case of the child looking at their reflection in the mirror), it is actually an absence that is invoking an absence, like in a series of mirror-effects. Therefore, as Rushton (2002) explains, cinema is for Metz "*doubly* imaginary" (Rushton, 2002, p. 109, emphasis in original) because an imaginary object (the projections on the screen) produce an imaginary scene. As such, Metz argues that the cinema is "an imaginary apparatus referring to imaginary scenes and situations" (ibid, p. 109); and because those imaginary scenes are represented in an imaginary way, Metz claims that the representation is psychologically more convincing than other forms of art such as the theatre (ibid).

Departing from this idea, Metz then takes the Freudian concept of fetishism[20] to argue that in the cinema the spectator experiences a doubling up of belief: they consciously know that the images projected on the screen are not real, but at the same time they unconsciously believe that they are real. Hence, Metz thesis suggests that the spectator at the cinema has to "double up its belief" (ibid): we are willing to believe in the 'reality' of the events projected on the screen even though we know very well that they are imaginary. Metz argues, therefore, that the audience is not deceived by the cinematic illusion, they know very well that what is presented on the screen is fictional; and yet, in order to the spectacle to be effectively unfold, they must believe in the 'reality' of the film.

Although apparatus theory has had its influence in film studies, it has also received lots of criticism[21]. Noël Carroll (1988), for instance, rejects psychoanalysis as a tool for explaining cinema as a whole. For Carroll, cinema spectatorship is a rational phenomenon, and argues that Baudry and Metz should have proven that the cinematic apparatus is irrational before psychoanalysing it. Although Carroll does not deny that certain films may be irrationally constructed or that certain films may produce irrational responses from spectators, he argues against Baudry and Metz in the sense that their theories do not speak of specific films but of cinema as a whole. Therefore, he argues, Baudry and Metz fail to demonstrate that the spectator's interaction with the cinematic apparatus is necessarily irrational, something that Carroll claims may not be possible to demonstrate.

The work of Baudry and Metz, however, has provided a fundamental base for other theoretical approaches that have intended to explain cinematic illusionism from different angles. Richard Rushton (2002), for instance, argues that although it is obvious that a big part of the spectator experience falls into conscious thought, much of the information that is presented reaches the spectator in an unconscious way. "These unconscious residues", Rushton argues, "would be as fundamental to the filmic experience as those things to which we can consciously point" (ibid, p. 118). But leaving psychoanalysis aside, Richard Allen (1993; 1995) explains cinematic illusionism from a rather rational perspective. As Richmond (2016) puts it, whereas Carroll asks scholars to stop calling the cinema illusory, Allen intends to change what

[20] The theory of fetishism is based on the coexistence of two contrary beliefs: 'All human beings are endowed with a penis' (primal belief) and ' Some human beings are deprived of a penis (evidence of the senses)' (Metz, 1982, p. 68). The fetishist boy finds an object to work as a substitute for his mother's lack of a penis. See Freud, S. (1927). 'Fetishism'.

[21] See for instance Bordwell, D. and Carroll, N. (1996).

we mean when saying that the cinema's representations are illusions. Allen argues that the cinema is propitious for experiencing what he calls *projective illusion*, a form of sensory illusion that, without encouraging the spectator to believe in the reality of what they see, favours the perception of the image as a fully realised world. Allen claims that the cinema spectator can be aware of the medium and yet experience the image as an illusion; an affect somehow similar to what Noël Burch, paraphrasing Metz, refers to as a "*willing* suspension of disbelief" (1982, p. 18, emphasis in original), which Burch argues is the basis of the cinematic illusion that he dubs the *diegetic effect*, namely a sense of being present within the diegesis. Allen argues that when experiencing a projective illusion, one perceives the fictional world as a world with "all the perceptual presentness or immediacy of our own" (1993, p.40).

Following some of the ideas proposed by Baudry, Metz, and Burch, Allen asserts that the illusion of being present in the narrative world is achieved through 'identification' with the camera's point-of-view; and is nourished by the possibility of movement offered by the cinematic medium, which is complemented, he observes, by the action of *diegetic sound*. Allen claims that the cinema spectator expects to experience movement; thus, if such an expectation is disappointed, their disposition to experience the film as a projective illusion will be diminished (1995, p. 112). Moreover, he highlights that spectator expectations are central for the experience of such an illusion, and suggests that the advent of colour and sound introduced a more demanding threshold; including the expectation of a sound treatment that responds to the events occurring in the narrative, or as he calls it, a "profilmically motivated sound" (ibid).

Allen thus emphasises the importance of establishing the relationship between onscreen and offscreen space for facilitating the experience of illusion. He explains that given that the cinematic image has movement, objects and characters constantly enter and leave the screen. Although we cannot see what is happening offscreen, sound gives a sense of presence to the events that are occurring beyond the limits of the frame, which allows the audience to be aware of the offscreen space. Therefore, Allen argues that when audience awareness of offscreen space is combined with the perception of movement supported by a soundtrack that foregrounds both the presence of the image and the continuum between onscreen and offscreen space, "the *optimum preconditions* for the experience of projective illusion are established" (1995, p. 108, emphasis added).

But beyond such requirements and technology-driven expectations, Allen (1995) notes that what helps to maintain the experience of projective illusion is the film's narrative. Contrary

to Burch's (1982) thesis, which affirms that what he calls the diegetic effect can take place even in the absence of narrative, the story constitutes for Allen a fundamental aspect in the experience of a cinematic illusion. He argues that the narrative helps to sustain the illusion in time, and only with the extra dimension of time, the spectator may lose their awareness of the real world and hence witness the cinematic event as something that is really happening in the exact moment of projection. Allen hence suggests that the cinematic apparatus on its own does not have the power to maintain the aforementioned illusion; rather it is the ways in which the narrative is presented within the apparatus what can achieve such an effect.

Classical Hollywood narration, with its genre and character-centred conventions, is for Allen the best example of the way the film's narrative allows us to experience the story as if it were a continuous present, and hence to experience the film as a projective illusion. Allen observes that common techniques (such as cutting, shot/reverse-shot patterns, eyeline match, the 180-degree rule, reframing upon the characters, and sound that matches the image and highlights the dialogue) promote the experience of such an illusion. Although a film like *Gravity* (2013), with its minimal plot and experimental aesthetics, challenges some of these assertions, the presence of characters is certainly what allows us to pass from experiencing what Allen calls "spectator-centered projective illusion" (in which we take up the point-of-view of the camera as our own and inhabit the narrative world by emotionally responding to the film as if we were witnessing the events in person) to experience what he refers to as "character-centered projective illusion" (in which we inhabit the narrative world through a sort of empathic identification with the characters on the film).

Allen notes that the typical experience of cinema "is one in which the spectator moves between spectator-centered projective illusion, character-centered projective illusion, and medium awareness" (1993, p.45). Such a description is consistent with the one offered by Rushton, who suggests that what he calls 'out of body' experiences do not take place for the entire duration of the film. In fact, Rushton notes that those moments are rare, they are instants of "unconscious perception" in which one actually "*believes* in the reality of the screen world in which one is engrossed" (2002, p. 114, emphasis in original). Drawing on Metz, Rushton notes that those moments are precisely the ones in which the cinema is successful for giving rise to wondrous fantasy worlds, moments to which the classical narrative cinema aspires. Yet, he argues, films can also destroy our fantasies when the tools are not effectively used (ibid, p. 116). In other words, using Allen's terminology, the spectator can pass from experiencing a

projective illusion to medium awareness depending on the ways the tools of cinema (e.g. narrative, cinematography, sound design, and so forth) are presented.

Allen's approach to the cinematic impression of reality advances the illusory aspect of apparatus theory to a great extent. Besides explaining the spectator's illusory experience from a rational perspective and highlighting the importance of the narrative for the consolidation of a cinematic illusion, one of his more important additions is the consideration of diegetic sound as an essential factor for the fulfilment of what he refers to as projective illusion, something that is completely absent in Baudry's original proposal. Yet, Allen's thesis still lacks a more in-depth exploration of the ways in which the soundtrack contributes to the achievement of the aforementioned illusion, a gap that this thesis intends to fill up considering the intervention of immersive cinematic sound.

This image-centric bias is also encountered in the work of Scott Richmond (2016), who explains the illusory effects of the cinema in the basis of phenomenology, human perception and proprioceptive effects in response of what we see on the screen. Using a phenomenological description (that is one based in his own conscious experience of the phenomenon), Richmond asserts that while watching *Gravity* in the cinema (especially but not only in IMAX 3D) he experienced an intense *bodily illusion*. He stresses:

> As Cuarón's camera continually moves through space, as I sit in the theater I find myself saturated by a sensation of my own body moving through onscreen space, at once weightless and frictionless, terrifying and illusory […] I found my heart rising in my throat, my arms tensed, my hands clutching armrests. I found myself pressing my whole body back into my seat and my feet into the floor in an attempt to anchor myself. My response was intense, but it was not intentional (2016, p. 1-3).

Richmond argues that Kubrick's *2001: A Space Odyssey* (1968) and Cuarón's *Gravity* (2013) are two examples of "crucially *cinematic* films", as their "sensational, perceptual, and proprioceptive effects" are mostly achieved within the infrastructure of a movie theatre (2016, pp. 3-4).

Within his thesis, Richmond notes that proprioception, which he defines as "the set of perceptual processes whereby we orient ourselves in and coordinate ourselves with the world" (ibid, p. 6), is fundamental for the cinema's intention of materialising the illusion of a world that unfolds "*before us onscreen*" (ibid, emphasis added). Richmond explains that proprioception means self-perception within the world, thus he defines *proprioceptive aesthetics* as "an aesthetics of self-perception modulated by the cinema" (2016, p. 8). Along

the lines of Hoberman (2013), who manifested that watching *Gravity* on an IMAX screen is "to appreciate the power of illusion—what André Bazin described as the experience of 'total cinema'" (n.p.)—Richmond argues that the totality of the cinema lies in its capacity to immerse us in a fictional world. The total cinema is hence for Richmond a proprioceptive cinema (2016, p. 143).

Richmond stresses that *Gravity* and *2001* share such proprioceptive aesthetics, a concept that he argues lies at the heart of the cinema, connecting its aesthetics and technological qualities. He explains that by proprioceptive aesthetics he aims to indicate three things at once: 1) a series of possibilities that the cinema offers as a technological system; 2) a series of films that make use of such possibilities in a productive way; and 3) a broader field of aesthetic practices and media that take the "paradoxically double presence of the body"—the spectator's illusory presence[22] in the diegetic world and its actual presence in the real world—as the focus of aesthetic creation (ibid, p. 9). In other words, Richmond notes that not all films succeed to provoke such proprioceptive effects; yet, today's cinematic apparatus offers the challenge and the opportunity of manufacturing such an 'out of body' illusion, something that can occur as long as content creators deploy the appropriate set of aesthetics.

Richmond highlights *Gravity* as being "the most successful—although certainly not the only—example of Hollywood's recent obsession with proprioceptive aesthetics" (ibid, p. 124-125). In this film, proprioceptive aesthetics are evidenced in the continuing modulation of the spectator's perception of space—particularly a sense of "position, orientation, and attitude" in the diegetic space (ibid, p. 8)—which creates the sensational illusion of movement through the diegetic world that is manifested together with a set of bodily responses: a lump in our throat, a tightening in our chest, the involuntary engagement of the muscles in our core, pushing into the back of our seat, our hands gripping the armrests, and so forth (ibid). Drawing on the work of Christian Metz, Richmond thesis thus suggests that the cinematic impression of reality is

[22] The International Society for Presence Research defines presence as "a psychological state or subjective perception in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience. Except in the most extreme cases, the individual can indicate correctly that s/he is using the technology, but at some level and to some degree, her/his perceptions overlook that knowledge and objects, events, entities, and environments are perceived as if the technology was not involved in the experience" (International Society for Presence Research, 2000).

not a type of illusion that deceives the spectator to believe that what they experience is real. He elaborates:

> I do not confuse the illusion of bodily movement with actually moving […] I feel a quasi movement whose distance from real movement is measured by responses in my body […] Proprioceptively aware of my situation in the cinema, when I feel as *though* I am moving, I feel it as illusion, *merely* as *though.* Which is to say, I *also* feel my stillness in the cinema. (2016, p. 58)

Although Richmond admits that some films deploy the so-called proprioceptive aesthetics more than others, he highlights that his thesis is not about films; his work propounds a theory of the cinema that takes the modulation of perception as its fundamental pillar. Given today's proliferation of all kinds of screens—namely TV screens, laptops, smartphones, and so forth—Richmond notes that theorising cinema is important because the cinema still matters as a fundamental proprioceptive medium. In the cinema, more than anywhere else, a film can materialise its proprioceptive capabilities and hence create "the impression of a world unfolding before us" (2016, p. 126). That is, the cinema matters "as an architecture and a particular kind of screen" (Richmond, 2016, p. 23), it matters as a medium where we can more easily get transported into alternative worlds, not only through the utilisation of enormous curved screens, but also by means of powerful surround-soundscapes (Recuber, 2007, p. 317).

Cinematic 3-D Sound and the Immersive Point-of-Audition

In the era of immersive sound, any image-centric theoretical approach seems insufficient to explain they ways in which the cinema creates the illusion of being present in the narrative world. Indeed, as Richmond (2016) argues, the movie theatre is the ideal space for creating the impression a world that unfolds before us; but in order to explain such an illusion we cannot set aside the sonic mechanisms that complement the image projected on the screen, and thus underestimate the role that sound plays for manufacturing the impression of reality.

Kerins argues that the impression of reality is best achieved by those films that deploy a lifelike spatial synchronization of image and sound (2011, p. 278-279), something that is in fact highly uncommon. Cuarón's most recent films are important to analyse precisely because they stand out as rare examples where such a perfect audio-visual match is found. In the vast majority of films, however, the monocular cinematic vision is conventionally contradicted by

the standardised sound practice of keeping the voice and other diegetic sounds locked in the front-centre speaker, overlooking the actual position that the sources of sound occupy within the three-dimensional filmic space. Conventionally, the movement of the camera is only partially supported by the soundtrack, as even when immersive sound technologies such as Dolby Atmos are in use, dialogue—and other narrative sounds—is most of the time kept immobile in despite of the trajectory that the characters and/or the camera follow. Such a screen-centric sound design approach limits the illusory potential of the cinematic apparatus by disrupting the spectator's central position that is the essence of all the above-mentioned theories. Hence, how could one claim that the cinema creates the impression of reality, if the perspective from where we see does not match the one from where we hear? Or put in another way, wouldn't the impression of reality be enhanced if the perspective from where we hear matches the one from where we see?

Metz noted that still photography was "not sufficiently lifelike: It lacked the dimension of time; it could not render volume acceptably; it lacked the sense of motion, that synonym for life" (1974, p. 14). Similarly, it can be argued that the conventional screen-centric sound treatment derived from classical narrative cinema and its system of continuity is not sufficiently lifelike: it lacks the sense of motion that, as Metz highlights, is one of the essential characteristics of real life (ibid). Therefore, this thesis argues that a sound design approach that coherently moves objects in the three-dimensions has the power to transport the spectator into the narrative world through the enhancement of the impression reality, only partially created by the moving image, and thus, as Metz once put it, "to render the world of the imagination more real than it had ever been" (Metz, 1974, p. 15).

However, representing our aural perspective within the cinematic world does not happen by default. In cinema, visual perspective is represented by simply showing what the camera records. Colour photography allowed filmmakers to provide a pristine representation of the way we see, something that is even more realist when 3-D cameras are employed. However, the sound devices that are part of the cinematic apparatus do not work automatically in the basis of a single unidirectional perspective as the one recorded by the camera and later projected on the screen. Unlike the visual artefacts, their sonic counterparts allow for a multiplicity of perspectives to be heard at the same time. Hence, in order to offer a realistic sonic perspective that matches the one of the visuals, filmmakers need to manipulate the sound—through panning, volume levelling and other acoustic adjustments normally done in

postproduction—in such a way that the speakers at the cinema provide the spectator with a 'pair of ears' that work in accurate spatial synchrony with what they see on the screen. That is to say, whereas the image track always presents a singular point-of-view at each moment in time, the soundtrack demands further processing in order to provide a lifelike representation of human hearing, conveying this way a single, coherent and three-dimensional *point-of-audition*, namely the point within the diegesis from where we hear.

In her essay 'The Voice in the Cinema: The Articulation of Body and Space', Mary Ann Doane highlights the cinema's intention of offering the audience a lifelike experience (1980, p. 35). Doane notes that sound offers filmmakers the possibility of representing a "fuller (and organically unified body)" (ibid, p. 34); and argues that what she calls the "fantasmatic visual space" constructed by the camera is supplemented by sound spatialisation techniques designed to localise and give depth to the voice, "and thus lend to the characters the consistency of the real" (1980, p. 36). She stresses that the utilisation of room tone, reverberation and sound perspective manifests the cinematic desire to recreate the way the voice fits within a given space (ibid, p. 36-37), enhancing this way the cinema's impression of reality. For Doane, sound technology innovations such as the Dolby systems are aimed at hiding the work of the cinematic apparatus, and thus diminishing the distance perceived between objective reality and its representation (1980, p. 35). She notes that the desire for presence in the space of the story works as a standard for measuring quality in the sound industry (ibid), which nowadays is clearly evidenced in Dolby Atmos' advertising slogan, which claims that this platform "*transports* you into the story with moving audio that flows all around you with *breathtaking realism*" (Dolby, 2018, emphasis added).

Doane notes that the aural representation of a certain position in the space of the diegesis, enhanced by the application of depth and sound spatialisation techniques, is what constructs a singular and stable point-of-audition that together with the image creates the 'hallucination' of a sensory world (1980, pp. 45-46). Aiming to define this concept, Michel Chion argues that the point-of-audition is the aural equivalent of the point-of-view (1994 [1990], p. 89). He refers to the latter's ability of being spatial or subjective, a characteristic that also adapts to the former (ibid, p. 90). Chion argues that the point-of-view is the spatial position from where the spectator sees a given shot, which can sometimes be the perspective of a character (ibid). It is important to emphasise that within this definition Chion is not referring to the well-known POV shot, but rather to any perspective used for a given shot, that is to say

the monocular perspective of the camera, which only on specific moments may represent the perspective of a character.

In a similar manner, Chion reflects on the concept of point-of-audition by enunciating two questions: "[f]rom where do I hear, from what point in the space represented on the screen or on the soundtrack?"; and "[w]hich character, at a given moment of the story, is (apparently) hearing what I hear?" (ibid). Thus, according to Chion, the point-of-audition can take two designations that are not always linked: a *spatial designation* if it is representing the point within the diegesis from where we hear, or it can take a *subjective designation* if it is representing a character's subjective hearing (ibid, pp. 90-91). Yet, in both cases, Chion is referring to a point within the filmic territory, to which we are transported through the power of sound.

Approaching this concept as a creative strategy, Rick Altman refers to *point-of-audition sound* as a technique that "is identified by its volume, reverb level, and other characteristics as representing sound as it would be heard from a point within the diegesis, normally by a specific character or characters" (1992a, p. 60). Consequently, Altman suggests that point-of-audition sound places the auditor on a specific spatial position inside the filmic space. "As such, point-of-audition sound always has the effect of luring the listener into the diegesis not at the point of enunciation of the sound, but at the point of its audition" (ibid). Altman argues that the point-of-audition is commonly the body of a character, and therefore he observes that this technique asks us to identify with the character that hears for us. "Instead of giving us the freedom to move about the film's space at will, this technique locates us in a very specific place—the body of the character who hears for us" (ibid).

An example of this can be found in O'Meara's (2012) analysis of Hartley's *The Unbelievable Truth* (1989) and Baumbach's *Greenberg* (2010). O'Meara argues that sound design in the form of point-of-audition, helps to construct audience identification, especially with those characters that hear sounds that others cannot hear (2012, pp.122-128). Similarly, Whittington's analysis of the sound design of Cuarón's *Children of Men* (2006) highlights how specific sound effects heard from the character's perspective—such as Theo's tinnitus ringing and ambience sounds—lure the audience to identify with the character by representing his point-of-audition (2011, pp. 4–12). Moreover, with examples taken from films like *Saving Private Ryan* (1998), *There Will Be Blood* (2008), and *Children of Men* (2006), Grajeda

approaches point-of-audition sound as one of the primary ways that filmmakers use to generate identification with a particular character in the text (2016, pp. 127-135).

In a similar manner, following Altman's subjective conceptions of point-of-audition sound (1992a, p.60), Martine Huvenne argues that the sound design of *Gravity* is always representing the relation between the protagonist and her environment. "Each time she turns, the sound turns with her", Huvenne claims, "[a]nd we as the audience also turn with her through the sound" (2017, p.47). Although such an interpretation is appropriate for certain moments of the film, Huvenne tends to assume that the main character is always the central axis of the sonic space, and does not investigate the creative possibilities of that 'invisible astronaut' mentioned in Cuarón's own discourse, which he describes as follows:

> So then it falls to us to find a way to deliver that objective view, but then transform it into a more subjective experience. So you are caught up in the trajectory of her story, seeing things as she does, through her point of view, but then continuing around to see her and the Earth and the stars behind her. The audience is then seeing things as their own subjective view, as they follow the character along in her journey. (Cuarón, 2013b, Interview by *ICG Magazine*)

Therefore, Huvenne's analysis is only concentrating on the subjective designation of point-of-audition, which reads well for particular instances of the story, but does not explain the sonic approach of the whole film.

Taking into account both spatial and subjective designations of Chion's original definition, this thesis observes that the sound design of *Gravity* is based on always maintaining some sort of immersive point-of-audition that works in synchrony with the point-of-view represented on the screen. In other words, the soundtrack of the film helps to locate the spectator in the centre of the filmic space by means of offering a three-dimensional representation of the aural world that surrounds our visual perspective. This way, sound fulfils the cinema's impression of reality and lures us to become a sort of invisible auditor, that all-powerful astronaut that is present in the narrative world and is able to detect vibrations, radio transmissions and the position of the original sources in space. Moreover, the film's sound concept seduces the spectator to identify with Dr Ryan Stone (Sandra Bullock), the protagonist of the film, by making us hear—in specific moments of the story—from her own focalised body, that is to say, by sharing her own aural perspective or inner world as if in certain instances we possessed her particular subjective reality.

Drawing on Doane's, Chion's and Altman's seminal ideas, and considering immersive sound technology as an important component of today's cinematic apparatus, this thesis presents the concept of *immersive point-of-audition*, understood as the three-dimensional aural representation of the point within the diegesis from where we hear. As such, this sound design strategy helps to locate the spectator on a specific point within the space of the narrative, and thus lures us to be *transported* into story world by means of making us hear from that specific point in the three-dimensional filmic space. Moreover, it is argued that given that this technique may also be used to represent the subjective aural perspective of an embodied character, it may also facilitate the process of audience *identification* with specific characters encountered within the film. Ultimately, the construction of such an immersive point-of-audition is meant to augment the cinema's capacity to create the impression of reality by means of offering a three-dimensional representation of natural human hearing, which in conjunction with the camera's monocular vision may facilitate the processes of transportation and identification in cinema, two constructs that Tal-Or and Cohen (2010) highlight for being two central paths for facilitating audience involvement in entertainment.

Transportation Theory and Audience Effects

Certainly, when watching a film, especially in a movie theatre, we may be able to feel as being there with the characters, or as Green and Brock (2000) put it, we may get transported into the narrative world. Green and Brock define *transportation* "as a convergent process, where all mental systems and capacities become focused on events occurring in the narrative" (ibid, p. 701). This theory is grounded on Gerrig's (1993, pp. 10-11) work and his metaphor of being transported, which suggests six elements as part of the experience of transportation:

1. Someone ('the traveler') is transported
2. by some means of transportation
3. as a result of performing certain actions.
4. The traveler goes some distance from his or her world of origin
5. which makes some aspects of the world of origin inaccessible.
6. The traveler returns to the world of origin, somewhat changed by the journey.

Green and Brock (2000) take Gerrig's metaphorical description as a basis for conceptualizing what they refer to as *transportation into narrative words,* which Green and Sestir define "as a as an integrative melding of cognitive, emotional, and imagery engagement in a story" (2017, p. 1). Green and Brock (2000) suggest that a transported individual will experience three main consequences: 1) "parts of the world of origin become inaccessible" (e.g. not noticing other people entering the room); 2) "experience strong emotions and motivations" (e.g. thinking of what could have happened to change an unhappy ending); and 3) "people return from being transported somewhat changed by the experience" (e.g. changes on individual's beliefs and attitudes) (Green and Brock, 2000, p. 702).

Based on empirical evidence, transportation theory suggests that transported individuals are completely absorbed by the story and thus are more likely to change their attitudes and real-world beliefs to match the ones implied in the narrative (Green and Brock, 2000; Green, 2004). Transportation is therefore not only a powerful momentary experience, but also a key factor for the influence that media can have on those who consume it (Green and Sestir, 2017). As Bilandzic and Busselle put it, a transporting narrative "may instigate reflection about the appropriateness of moral rules, norms and values" (2011, p. 35). According to Green and Clark (2013), this mode of narrative persuasion occurs through multiple means. First of all, transported individuals are likely to reduce counter-arguing of explicit and implicit messages. That is to say, transportation is directly connected to what is commonly referred to as *suspension of disbelief*, meaning that the media user supresses certain information and stop counter-arguing the content of the story (Böcking and Wirth, 2005).

Moreover, according to Green, Brook and Kaufman (2004), the process of transportation may be a prerequisite for experiencing *identification* (2004, p. 318), a mental process that Cohen defines as "a mechanism through which audience members experience reception and interpretation of the text from the inside, as if the events were happening to them" (2001, p. 245). Cohen argues that identification with characters involves a temporary adoption of someone else's point-of-view (ibid, p. 248), meaning that the absorbed individual experiences the world not as being a spectator but rather as being a character, adopting their identity and perspective (ibid, p. 251). Ultimately, grounded on empirical research[23], Green and Clark (2013) highlight that transportation may lead individuals to adopt the beliefs of the

[23] See Sestir and Green, 2010

characters with whom they identify, or even to shift their self-concept and become more similar to such characters (Green and Clark, 2013).

Furthermore, Green and Clark (2013) argue that an increasing perception of realism or the resemblance of a direct experience tends to have a greater impact on beliefs and attitude changes. Although transportation theory focuses on the power of the narrative to transport the reader (listener/viewer) into the world of the story even without any special effect whatsoever, Green and Clark (2013) highlight that the illusion of reality that can be achieved in a medium like the cinema can facilitate the occurrence of transportation and its effects on the spectator. This claim resonates with Bilandzic and Busselle's (2011) experimental findings, which suggest that a film's perceived realism can facilitate the occurrence of narrative experiences such as transportation and identification. Furthermore, Green, Brock, and Kaufman observe that in relation to interactive technology, the process of transportation is the result of a mediated simulation of presence in the virtual space (2004, p. 122).

Lombard et al. (2017) observe that presence experiences are common, and occur not only in virtual, augmented, and mixed reality, but also in more accessible media such as film. In the realm of cinema, such an experience is commonly known as immersion, which as Mera notes, "is achieved by replacing as many real-world sensations as possible with the sensations of a virtual environment" (2016, p. 92). That is to say, the experience of immersion in cinema can be interpreted as a synonym of what is commonly known as spatial presence, "in which mediated environments are perceived as nonmediated" (Lombard et al., 2017, p. 1). In today's cinema (within which she includes the category of cinematic virtual reality), Cortés-Selva explains that the experience of immersion can be divided into physical and cognitive (2016, p. 176). In physical immersion, she notes, the user is virtually 'present'—embodying an avatar—in a three-dimensional space that can be virtually created or captured by means of 360-degrees cameras (ibid). Cognitive immersion, on the other hand, 'transports' the audience member in the sense that it creates the illusion of presence—as in reality the spectator remains sited in the theatre—in a diegetic world, which is achieved as a result of a series of narrative, stylistic and technological strategies (ibid).

Among such strategies, Cortes-Selva includes the audio-visual techniques of depth of field, camera movement, and surround sound, which combined can enhance the experience of immersion (ibid, p. 177). Hence, while physical immersion is only possible through virtual reality and other interactive technologies, cognitive immersion can in fact take place within the

space of a movie theatre as long as the impression of reality is accurately manufactured with stimulus that simulate our real-world perceptions. In cinema, the process of transportation is thus directly linked to the process of immersion, an experience that can be enhanced through 3-D sound, a cinematic strategy that serves as one of the 'audio-visual cues' that, together with the knowledge that we already possess of our world's spatial characteristics, help to insert the spectator within the film's 'setting', defined by Lefebvre as the place where the story events occur (2006, p. 21).

It is important to emphasise, however, that transportation levels may increase or decrease depending on a series of variables, such as: narrative quality (higher quality narratives are more transporting), individual differences (some individuals are more easily transported than others, a condition that is commonly known as transportability), or situational factors (e.g. individuals' goals, environmental factors such as screen size, number of discreate speakers available, distraction, and so forth) (Green and Clark, 2013). As such, all the possible effects of transportation would only take place under optimal conditions (Green and Brock, 2000). As an example, a smoking scene in a highly transporting movie should have a stronger effect on spectators' propensity to smoke than a smoking scene in a less transporting movie (Green and Clark, 2013); and ultimately, any narrative message is likely to have a greater impact when conveyed through a highly transporting film.

It is also important to emphasise that transportation is a desired state, as we enjoy escaping from reality and getting into different worlds far from everyday life, even (given the safeness of fiction) into fictional worlds that are frightening or cheerless (Green, Brook and Kaufman, 2004, p. 315). Green, Brook and Kaufman (2004) highlight the premise that enjoyment is the reason for which people spend a lot of time with media (ibid). As Green and Sestir note, the enormous amounts of money spent each year on movies "attest to individuals' desire to leave reality behind and experience other worlds" (2017, p. 1). For Bilandzic and Busselle enjoyment is "an essential motivator for audiences to use media and maybe even the origin of all other motives" (2011, p. 29). In fact, Bilandzic and Busselle (2011) empirically found that narrative experiences such as transportation and identification are connected to enjoyment in film. Exploiting the cinematic impression of reality stands out therefore as a fundamental strategy for facilitating transportation and identification, and through this produce enjoyment.

Finally, it is worth mentioning that the level of perceived realism that can be achieved in the cinema is undoubtedly higher than the one achievable on a TV screen. Yet, the difference between one another can be insignificant if filmmakers conform to screen-centric approaches and hence do not exploit the lifelike three-dimensionality that only the cinematic soundtrack can provide. Therefore, in order to enhance the theatrical experience over its home-based counterpart, and hence to increase the possibilities of transportation and its effects on the spectator, it would be of great importance to review and analyse the ways in which filmmakers like Alfonso Cuarón have succeeded on the exploitation of immersive 3-D sound, a highly powerful component of present-day's cinematic apparatus.

Conclusion

The impression of reality that Bazin saw as the cinema's ultimate goal, has continued to be a fundamental concept throughout contemporary film theory. Baudry's apparatus theory, although highly criticised and controversial, stands out as one of the first attempts to study the connection between cinema and spectatorial effects. Today, transportation theory shall be highlighted as one of the most recent endeavours for explaining such a relationship, and although its foundations do not refer specifically to film, its premises offer enough space to study the cinematic impression of reality and its persuasive potential.

The consolidation of immersive sound as an important component of present-day cinema offers new opportunities for the reassessment of these and other film and media theories. As argued above, the representation of a lifelike immersive point-of-audition may help to transport the spectator into an illusory central position within the filmic space; and as transportation theory suggests, individuals may return somewhat changed after being transported into the narrative world. It is argued as well that transported individuals may more easily identify with sympathetic protagonists, a feeling that may be intensified through the three-dimensional representation of the character's point-of-audition.

Research suggests that perceived realism has the power to facilitate transportation, meaning that the persuasive effects of such a narrative experience may be boosted when 3-D sound is effectively deployed. Given that transportation is a desired estate for the cinema

spectator, exploring ways to facilitate its occurrence foregrounds as an important contribution to knowledge and understanding in the field.

In the following chapters, the theories and concepts here presented will serve as the fundamental skeleton for the analysis of a set of films that the author considers to be exemplar on the utilisation of 3-D sound, a sonic strategy that enhances the cinema's impression of reality and through this props up transportation and its effects on the spectator. 3-D sound is in fact a key aesthetic component of *immersive continuity*, the film style that gives name to the present thesis. Let us explore its aesthetics in the following chapter.

CHAPTER THREE

Gravity: The Audio-Visual Aesthetics of Transportation

Gravity (2013) is a science fiction thriller that follows the journey of Dr Ryan Stone (Sandra Bullock), a medical engineer on her first mission in outer space. While working on a set of hardware upgrades on the Hubble, a sudden space debris takes place, which destroys the shuttle and cuts communication with Mission Control. Stone and veteran astronaut Matt Kowalski (George Clooney) are the only ones that survive the disaster, but the latter dies during their quest to reach the International Space Station (ISS). As the sole survivor of the mission, Stone passes through a series of vicissitudes before achieving the impossible, she returns safely to Earth aboard a space capsule.

In this thesis's introduction, this film was acknowledged for its three-dimensional utilisation of Dolby Atmos, an immersive sound platform whose characteristics were described in Chapter 1. In Chapter 2, the film was taken as a reference for conceptualising what this thesis denominates immersive point-of-audition, a sound design strategy that may facilitate the process of transportation by providing an immersive representation of natural human hearing. Certainly, the film's three-dimensional treatment of sound is worth investigating, and so is its unorthodox visual style. Thanks to a visual aesthetic anchored to the long take, the sound team of *Gravity* was able to establish a specific perspective from where we, as members of the audience, hear the sounds that belong to the cinematic universe, a perspective that always matches the one conveyed via the camera's point-of-view. Ultimately, going against the conventions of contemporary mainstream cinema, Cuarón employs the full potential of immersive sound by maintaining at all times a coherent, realistic and three-dimensional representation of a given aural perspective, that is to say, beyond conveying a point-of-view, Cuarón endows the camera with its own point-of-audition. Based on the theories, concepts and technical information provided in the previous chapters, this chapter will explore the audio-visual aesthetics of *Gravity,* a film that features many of the stylistic characteristics that are the core of this investigation.

The Sound of *Gravity*

In the first part of his essay 'Sound, Space, Gravity: A Kaleidoscopic Hearing', Dong Liang (2016) highlights the treatment of the voice as perhaps the most important factor that differentiates *Gravity* from other films. The anchoring of voices to bodies, he argues, "is accomplished by a consistent mapping of the spatial location of the voice: a voice is associated with a body, because it is perceived as coming from the precise location that a body is found on the screen" (ibid, p. 5). This treatment of the cinematic voice, Liang notes, was resurrected by Cuarón from a crucial idea that originated during the early years of sound, which as pointed out in this thesis's introduction, intended a perfect spatial match between image and sound with the ultimate goal of making the spectator experience the movie from within the filmic space. Among other factors, the technical limitations of those years obstructed this sonic strategy to take place. Today, such a limitation does not longer exist, and Cuarón has finally deployed a three-dimensional sound design approach that has the power of transporting us into the narrative world.

- **3-D Sound and Transportation into *Gravity*'s Narrative World**

Using a discourse that resonates with the theories of illusion in cinema discussed in the previous chapter, Cuarón acknowledges that the intention in the case of *Gravity* was to make the spectator experience the film as if they are sitting in the theatre but they feel as if they are "floating in space" (Cuarón in IMax, 2013). His goal was hence to transport us into the story by means of some sort of virtual simulation of a lifelike three-dimensional world. Let us analyse for instance the first conversation between the two key characters, a scene that can serve as an illustrative example that explains the sonic approach of the whole film.

Figure 3.1 – Frame taken from *Gravity*. Stone's voice and movements are a bit panned to the front-left while Kowalski's to the front-right.

At minute 00:06:14 the camera shows a two-shot of the protagonists: Dr Ryan Stone is seen on the left and Kowalski is seen on the right (Figure 3.1). Accordingly, Stone's voice and movements are slightly panned to the front-left while Kowalski's to the front-right, both heard quite clearly from a close perspective. Additionally, her exaggerated breathing not only makes it evident that she is not feeling comfortable in the vacuum of outer space, but also suggests our closeness in terms of depth and perspective. In between their conversation, the distant and more filtered voice of Mission Specialist Shariff (Phaldut Sharma) is heard behind us, suggesting that we are in the middle of the three-dimensional filmic space. In addition to the voices of these three characters, we can also hear the voice of a man in Houston—Mission Control—and the voice of a woman who operates 'Explorer' from inside the spaceship. The man's voice emanates from the front-right in accordance with the position of Earth, while the woman's voice is heard on the rear in accordance with the position of the communications cabin in the spaceship.

Figure 3.2 - Frame taken from *Gravity*. We can hear Stone's voice on our left while she is offscreen.

Resembling Chion's concept of "*on-the-air*" *sound* (1994 [1990], pp. 76-77), all the voices have a special low-cut filter that simulates a radio transmission; however, Stone's and Kowalski's are less filtered and hence more clearly heard, which can be interpreted as both being closer to our perspective, suggesting the existence of some sort of *spatial signature*,[24] which under 'normal' acoustical circumstances would include reverberation and therefore differences in the ratio of reflected and direct sound, even on a story that takes place in outer

[24] Altman uses the term spatial signature to suggest that a sound heard from different positions would result in different representations of the same event, or to put it in Altman's own words 'a different narrative of the same event' (Altman 1992b: 24). For example, a sound event heard from a close perspective will be very clear and intelligible, while the same event heard from a distance will probably be more reverberant and unintelligible.

space. That is to say, the amount of filtering replaces the amount of reverberation and creates the illusion of depth and sound perspective. Hence, besides signifying the characters' "social distances from one another" (Gabriel and Sonnenschein, 2016, p. 91), this technique helps to establish our point-of-audition.

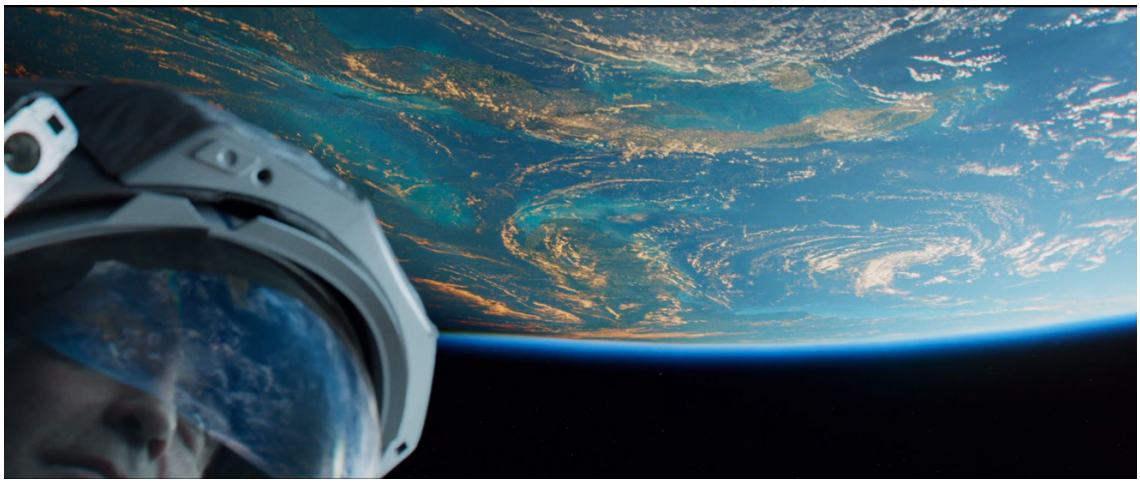


Figure 3.3 - Frame taken for *Gravity*. A close-up of Kowalski looking at the marvellous blue planet.

Figure 3.4 - Frame taken from *Gravity*. We see Kowalski reappearing through the right side of the screen.

Thereafter, Stone moves and keeps talking offscreen (Figure 3.2), we can hear her voice on our left, but it gradually moves to the right side of the screen as the camera follows her and pans in a 180-degrees manner. For the second part of their conversation, we see Stone on the right, Kowalski on the left, and Shariff on the far background, and thus their voices occupy their new correspondent position in the aural realm. After seeing Shariff fooling around from a distance, the camera reframes and shows a close-up of Kowalski looking at the marvellous blue planet (Figure 3.3). As the camera slowly pans in rhythm with the enveloping non-diegetic music, we get to see a splendid panoramic view of Earth, which leaves the characters offscreen.

Kowalski's voice is heard on our left, "so what do you like about being up here?", Stone exhales and replies, "the silence, I could get used to it", which we hear behind us. After a 360-degress camera rotation, we see Kowalski reappearing through the right side of the screen (Figure 3.4), "terrific", he says, and his voice and chuckles are clearly heard emanating from the front.

Conventionally in cinema, as we have discussed, all voices, including the offscreen ones whose source we cannot see, are locked in the same single position: the centre of the screen emanating from the centre speaker behind the screen. Yet, as Doane notes, "[t]here is always something uncanny about a voice which emanates from a source outside the frame" (1980, p. 40). The screen-centric "voice-off", as she names it, is thus unnatural in despite of being conventional. It is a voice that emanates from an "empty" screen; a voice without a body whose efficiency is grounded on the knowledge that the offscreen character will be visible again through reframing, reuniting this way the voice and its source (ibid, p. 41).

Chion suggests the term "*acousmêtre*" to refer to an offscreen voice that has not yet been linked to a body (or more precisely a mouth) seen onscreen. Unlike the treatment of the voice in the cinema, he argues, the acousmêtre in the theatre is physically located outside the stage, and therefore is heard from an off-stage space (1999 [1982], p. 22). The treatment of the offscreen voices in *Gravity* could be said to be closer to the theatre acousmêtre. In the sequence described above, for instance, the voices of the characters move in accordance with their position in the diegesis, emanating from the front speakers when we see them on the screen, or from the surround speakers as the characters move offscreen across the *acousmatic* field, that is the offscreen space that surrounds the visual field within the three-dimensional world of the movie.

Ultimately, Cuarón's sound design approach aims to bring us closer to the way we experience reality. As Chion observes, human vision is partial and directional, meaning that the screen at the cinema does a great job representing the way we see. On the other hand, hearing is omnidirectional; "[w]e cannot see what is behind us, but we can hear all around" (Chion, 1999, p. 17). The treatment of the voice in *Gravity* breaks the long-lasting screen-centric tradition implemented in cinema since the era of monophonic sound. The film stands out as a contemporary example of the ways in which the utilisation of voices whose sources we cannot see "deepens the diegesis" and reinforces the existence of "a space in the fictional world which the camera does not register" (Doane, 1980, p. 40). The construction of a single, stable and immersive point-of-audition reinforces the relationship that exists between sound

and image, not only by means of voice synchronization (lip synch), but also through the unification of the character's body and voice as elements of the same three-dimensional space. Ultimately, Cuarón's sound design approach not only 'brings to life' a space that we cannot see, but helps to transport us into the narrative world by increasing the cinematic perception of realism.

Although *Gravity*'s highly realistic sound design approach minimises the utilisation of sound effects in the scenes that occur in outer space (a place where sound waves cannot travel, and therefore sounds could only be received in the form of vibrations or radio signals), there are also a few sequences that take place within interior atmospheres, which confirm that 3-D sound can also be effectively deployed in fully-audible scenarios. In such environments, the imaginary presence of the spectator in the filmic space is conveyed by a coherent arrangement of voices and sound effects within the 3-D space, which in such cases carry their own tone, reverberation and other acoustic characteristics.

Figure 3.5 - Frame taken from *Gravity*. Within the ISS's capsule, we hear every single sound as Dr Stone takes her spacesuit off.

For instance, after finally entering into the International Space Station (ISS) (00:37:34), Stone regulates the air pressure of the capsule and thus we gradually begin to hear all the sounds emitted within the diegesis. Immersed in an enveloping atmosphere, we hear every single sound as she takes her spacesuit off. Sonically, we are able to perceive all the different components of the spacesuit as they move across the room, clashing with each other and with the surfaces within the zero-gravity space capsule (Figure 3.5). After taking a few seconds to breathe, Stone goes to the communications cabin and we see her 'flying' through the corridors. As the camera follows her trajectory throughout several sections of the space station, the

representation of its immersive point-of-audition helps the spectator to experience the bodily illusion of movement, something that is achieved by means of providing a dynamic arrangement of the different room tones through which we pass. At a certain point, for instance, both image and sound focus on a few fire flames that have sparked as a result of a short circuit. As the camera shows a close up of the damaged device, we hear the sparks very close in front of us; however, as it moves away and keeps on following Stone, we hear them at a lower volume emanating from the rear-left, where they gradually disappear.

Once Stone gets to the radio unit, she puts a pair of headphones on and begins to speak aiming to communicate with Kowalski, but there is no response. We see her in the centre of the screen, and her monologue, together with the static of the radio system, is heard in the centre speaker, increasing in volume as the camera approaches and frames a close up of her face. After the camera pans to the right and focuses on a panoramic view of Earth shown through a hatch, we see her face reflected on the glass (Figure 3.6), but we hear her voice, with some sort of echo created by its reflection on the circular window, panned to the left side of the auditorium. As she confirms being "the sole survivor of SDS157", the camera moves back and hence we hear her voice moving from left to centre.

Figure 3.6 – Frame taken from *Gravity*. We see her face reflected on the glass, while her voice is panned to the left side of the auditorium.

This sequence shows how 3-D sound helps to transport us into an interior location by providing a coherent, realistic and immersive representation of hearing within that space. Here, the dynamic atmospheres inside the International Space Station contribute to experience what using Richmond's (2016) terminology could be called proprioceptive perception of movement, a sort of self-perception within the space of the movie. Ultimately, throughout the whole film, specific sound elements are precisely localised and coherently change position in accordance

with our visual perspective, facilitating this way the process of transportation through a vivid impression of reality.

- **3-D Sound and Identification with the Protagonist**

As discussed in Chapter 2, transportation theory suggests that the narrative experience of transportation might be a prerequisite for experiencing identification with fictional characters (Green, Brook and Kaufman, 2004). That is to say, this theory indicates that transported individuals are more likely to identify with the individuals found within the narrative. In cinema, transportation may lead to identify with the characters that we see the most, namely the film's protagonists. Yet, as Cohen argues, cinematic identification can also take place in function of "the primary point of view" (2001, p. 257), a premise that this thesis complements by suggesting that identification with a given protagonist can be amped up through the immersive representation of that character's point-of-audition.

Figure 3.7 - Frame taken from *Gravity*. POV shot of Dr Stone from inside the helmet. We see exactly what she sees, and as a counterpart, we hear exactly what she hears.

An example of this can be found at minute 00:19:36, when a cut to a POV shot is employed to show exactly what Stone sees, and as a counterpart, her subjective point-of-audition is three-dimensionally represented to make the audience hear through her individual perspective, from inside the helmet (Figure 3.7). This is a clear representation of the subjective audio-visual experience of the character; we are not only seeing through her eyes, but beyond, we are hearing through her ears. From inside the helmet, we see and hear Kowalski as he tries to communicate with Houston. His filtered voice comes from all around us, but it is softer on the centre speaker, that is to say, it is a representation of how his voice is heard inside Stone's

helmet. Hence, this time his voice is not heard as a traveling radio voice as in previous objective sequences, but as an audio signal that Stone receives through a sound system inside the helmet.

At the same time, Stone's own voice (and breathing) is heard more closely and louder on the centre speaker, it has a certain reverberation that sounds equally in all directions, which provides a realistic representation of the way she hears her own voice within the small space of a helmet. Additionally, we hear her heartbeat being accentuated through the LFE channel, and the vibrations that are produced every time the rope that keeps the two astronauts linked gets stretched and sharply pulls her body towards Kowalski. Furthermore, the frequency sweep of the jamming radio signal is heard moving all around, merging with the three-dimensional music score that also travels across the auditorium as a resource that is employed throughout the film to supersede the absence of sound effects. Throughout this sequence, 3-D sound facilitates a greater level of identification with a character that we embody not only by seeing through her own eyes, but additionally, by hearing three-dimensionally through her ears.

Typically, sound design is thought to promote identification with protagonists by giving us access to their somehow altered way of hearing (e.g. hearing loss or tinnitus), as well as their thoughts, dreams or mental formations (O'Meara, 2012). In such cases, we as an invisible auditor keep the objective point-of-view provided by the camera, but in a way share the subjective hearing or mental activity that the character experiences. That is to say, even though in this type of situations we are not embodying the corpus of the character nor their position in the diegesis, we still come across with a highly subjective event as the information that we hear comes from the individual experience of a specific character.

Such a situation occurs in *Gravity*, for instance, when Stone gets knocked out for a few seconds while trying to put out a fire in the ISS. Similar to the tinnitus ringing experienced by Theo Faron (Clive Owen) in Cuarón's *Children of Men*, the disturbed subjectivity of Dr Stone is represented sonically by filtering all sounds and adding a high frequency pitch, a resource that was also very effectively used in Spielberg's *Saving Private Ryan* (1998) to represent Captain Miller's (Tom Hanks) state of mind during the astonishing D-Day opening scene. In such cases, representing the three-dimensional sonic space can help, yet is not essential, because what actually promotes identification is the filtering of sound waves and not the spatial position of the aural elements.

There are other situations, however, in which 3-D sound can facilitate identification by representing the inner world of a character, as long as such a world maintains all the spatial characteristics of the one we know. In *Gravity*, this type of strategy was deployed in the scene in which Stone decides to kill herself by lowering the oxygen levels of the space capsule (01:01:53). She closes her eyes, and after a few seconds Kowalski knocks on the window. A crescendo of non-diegetic sonic layers expands as he turns the door handle, a disturbing noise that is suddenly terminated when he opens the door, which highlights the absence of sound in outer space. Under complete silence, he enters into the capsule, closes the door and turns up the oxygen levels, and gradually we begin to hear again. Their voices, the emergency alarm and other diegetic sounds are heard in accordance with their position on the three-dimensional diegesis; however, a reverberant version of Kowalski's favourite country song is heard in all directions, which subtly suggests that he is actually part of Stone's inner world. For the most part of the sequence, we see a two shot of them as they talk (Figure 3.8), but on the last part of Kowalski's deep and emotional speech, the camera focuses on Stone, leaving Kowalski's offscreen voice emanating from the right (Figure 3.9). As Stone begins to realise that what she is experiencing is not real, the oxygen alarm gradually becomes present again and we see that in fact no one is sitting beside her, which one can interpret as a sort of hallucination that was presented to us by sharing Stone's mental disruption.

Figure 3.8 – Frame taken from *Gravity*. We see a two shot of the characters as they talk.

Figure 3.9 - Frame taken from *Gravity*. The camera focuses on Stone, leaving Kowalski's voice emanating from the right.

This example shows how 3-D sound can also facilitate identification with fictional characters by inviting us to experience the way they hear within their subjective inner world; that is, within some sort of internal diegesis that exists within the diegesis of the film. Ultimately, the process of identification with fictional characters can occur through multiple means. Transportation itself can facilitate identification to occur, but the latter can be supported by giving the spectator access to the way a specific character hears, which can be realistically represented through 3-D sound.

- **3-D Sound and the Non-Diegetic Terrain**

As it has been shown, the concept of immersive point-of-audition and its theoretical connotations help to explain how the sounds that are part of the diegesis are arranged in function of a single audio-visual perspective. However, it is also important to analyse they ways in which non-diegetic elements (e.g. the musical score) may be incorporated when embracing this sound design approach. Music is an essential component of mainstream cinema, including the soundtrack of *Gravity*. Nonetheless, within the context of this film, it is in the equilibrium with a 'silent' spatial reality where non-diegetic music finds its beauty, a sound treatment that fits into Kulezic-Wilson's (2008) ideas of a less musical sound design. For Kulezic-Wilson:

> Music in film is trying too hard these days, and even those who are not film music specialists are starting to notice it. This can be felt not only in the inevitability of scoring every action/transition/emotional/suspense scene, but also in the stubborn determination to fire all of the musical weapons all the time. So no wonder that an increasing number of composers are trying to resist this trend and convince directors to

> use less music in their films, knowing instinctively that this will not only benefit the film itself but also their own efforts to create effective scores. (2008, p. 127)

Using the work of directors like the Coen Brothers, Gus Van Sant and Darren Aronofsky to illustrate her argument, Kulezic-Wilson (2008) suggests that sound design has the potential for becoming the new score; and the work of Alfonso Cuarón can also be used to support such a claim. Although in *Gravity* the musical score is abundant and quite necessary as it replaces the absence of sound effects[25], Steven Price's composition still offers enough space for the diegetic 'silence' to manifest itself in several moments throughout the film.

In fact, diegetic sound is all what we hear for the most part of the twelve and a half minutes opening sequence analysed earlier in this chapter. It is only after Houston orders them to abort the mission that the score appears, yet it does so in quite an unconventional way. Following the disorienting concept of the film, its music also breaks the screen-centric regime of Hollywood tradition by allowing the score to freely move in all directions across the auditorium. As *Gravity*'s re-recording mixer, Skip Lievsay explains, incidental music is normally fixed and does not move unless there is movement happening in the orchestra itself (personal interview, 2019). In *Gravity*, on the other hand, music moves all the time in the three dimensions, emancipating from the fixed surround sound representation typical of mainstream films, which is for Mera one of the fundamental characteristics of what he calls 3-D sound (2016, p. 103). Along the lines of Kulezic-Wilson, Mera observes that *Gravity* demonstrates how the borders between music and sound design are collapsing, something that for him works in service of creating the illusion of presence in the diegetic space (ibid). Price argues that what they were trying to do was to "reconceptualize what action music could be. All of the composition derives from the movement of the screen, the characters, the emotion" (Price in *Gravity*: Special Features, 2015). Lievsay comments:

> [Cuarón and Prince] very wisely made the choice to have music associated with the objects, so when things were roaring by, we took things in the score and we panned them to match the action in the frame, so the score becomes the sound effects path. It occupies the space that would normally be for sound effects. (personal interview, 2019)

Moreover, the intensity of the film's musical score fades in or out in synchrony with the events that are happening in the story. For instance, after Stone finds herself lost in the vacuum of outer space, the moving camera approaches the protagonist and enters into the

[25] In Gravity's special edition Blu-ray, Cuarón invites the audience to experience a version of the film without music.

individual atmosphere of her helmet. At this point, the transition between the two spaces, the vacuum and the helmet, is accentuated by fading out several musical layers, enriching in this way the realism of the scene. A similar technique is employed throughout the film to intensify the effect of diegetic silence, such as the moments when the characters enter into the space capsules and the music is removed just as the hatch gets closed. In other words, the film's musical score is non-diegetic, yet it adopts the spatial characteristics of the diegetic world and this way enhances the sensation of presence in the narrative world.

In summary, although the concept of immersive point-of-audition primarily provides a framework to explain the treatment of diegetic sound elements, its three-dimensional characteristics can expand to the non-diegetic terrain. This is evidenced in *Gravity*'s incidental music, which supports and enhances the spatial attributes of the film's soundscapes not only by giving them the chance to realistically manifest themselves in several moments throughout the movie, but also by merging with them and hence giving rise to a unified sound design score.

The Look of *Gravity*

In his seminal article 'Intensified Continuity: Visual Style in Contemporary American Film*'*, David Bordwell (2002) notes that fast editing is one of the characteristics of the 'look' of what he terms *intensified continuity*, the dominant editing style of today's mass-audience films. Bordwell stresses that editing has become faster through the years, with average shot lengths (ASL) that have passed from hovering around eight to eleven seconds in the period between 1930 and 1960, to three to six seconds in the years of 1999 and 2000. This observation has been reconfirmed in more recent publications, such as the eleventh edition of Bordwell, Thompson, and Smith's 'Film Art: An Introduction', in which they note that "today a 2-hour film might contain 3000 shots or more" (2016, p. 246). Using Barry Salt's data available on the website "*Cinemetrix*" (which contains ASL measurements for nearly 11,000 films), the author of this thesis has calculated that the ASL on American films since the year 2000 until 2012 has a media of around four and a half seconds[26]. Although all these numbers may very well describe the editing style of the vast majority of contemporary movies, they do not apply

[26] The measurement was done using the data available on Cinemetrix on 31 July 2020. It included a total of 1,323 films categorized as American.

to the visual aesthetics that seem to support a highly spatialised deployment of 3-D sound, and the 'look' of *Gravity* only serves to ratify this argument.

In collaboration with Lubezki, Cuarón has developed an unconventional visual aesthetic that subverts the norms of Hollywood's contemporary editing system. Without considering the titles introduction and the credits sequence, the author has calculated that the ASL of *Gravity* reaches 26 seconds among a total of 194 shots, confirming a visual style that contradicts the typical quick cutting strategy of contemporary Hollywood (Bordwell, 2002, pp. 16-17; Salt, 2009, pp. 376-378; Kerins, 2011, pp. 102-103). Certainly, as Isaacs notes, the long take means liberation from the spatiotemporal regime of tradition (2016, p. 476). The opening sequence, for instance, presents an impressive long take that lasts almost twelve and a half minutes, and shots of more than one minute in length are constantly found throughout the film. As such, the average length of each shot in *Gravity* is much longer than the 'norm'; and even though some of these long takes were artificially assembled from multiple shots in postproduction, in the end they "give the *appearance* of taking place in 'real' time and 'continuous space'" (ibid, emphasis in original); they achieve that 'documentary look' that Udden (2009) relates with the seminal ideas of André Bazin[27].

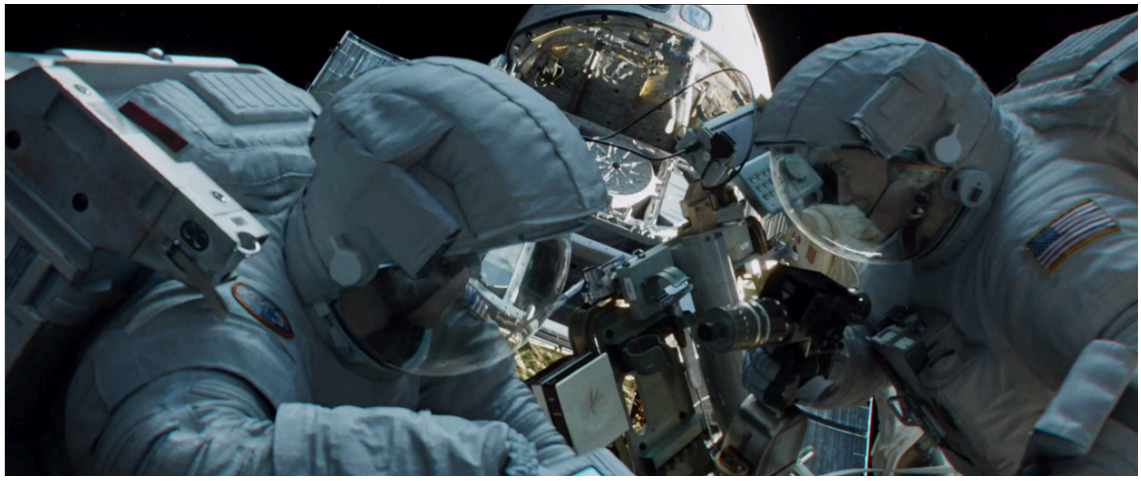


Figure 3.10 - Frame taken from *Gravity*. An example of a shot with deep-focus cinematography.

Another important consideration regarding the visual style employed in *Gravity* is the director's intention to present the story from an objective point-of-view, conveyed through the utilisation wide-angle lenses that capture the action from a distance while keeping all visual

27 See Bazin, A. (1967) What is cinema? Vol. 1: 'The evolution of the language of cinema'; 'The virtues and limitations of montage' and Bazin, A. (1971) What is cinema? Vol. 2: 'An aesthetic of reality: Cinematic realism and the Italian School of the liberation'

layers in focus thanks to their large depth-of-field. Such wide-shot driven aesthetics allow them to create highly realistic compositions in depth. Deep-focus is in fact a key element in *Gravity*, a film full of realistic panoramic views with front-ground, mid-ground and background elements distributed in focus along the whole visual field. See, for instance, Figure 3.10, which shows a two-shot of the main characters, while, at the same time, we are able to see Mission Specialist Shariff being in sharp focus in the background. Regarding such visual aesthetics, Cuarón comments:

> When you have a long take from a wide angle, you're prioritizing neither the character nor the surroundings. Context and character have the same weight. In fact, maybe the context is more important. The character merely flows through that context. What has interested me ever since *Y Tu Mamá También*, is the relationship between subject and context. With a close-up, you're emphasizing the subject, the actors, the characters. I love close-ups, but I don't condone using them to make the narrative easier. In the language of TV, and most commercial films, whenever a character speaks, you have a close-up. And then you cut to another close-up of the reply. You can watch those movies with your eyes closed. There's no language in them (Cuarón in *Road to Roma*, 2020, original translation).

Furthermore, Richmond notes that Cuarón's visual aesthetics rely on keeping the camera in constant movement, orienting the spectator's attention by arranging objects, events, and characters "in an unerringly mobile frame" (2016, p. 2). Cuarón discovered that in order to make the audience experience the sensation of weightlessness in space, he needed to create the illusion of movement within that space (ibid). In this film, such an effect was visually achieved by means of what is known as *virtual cinematography*, a modern cinematic technology that, as Nick Jones notes, has the ability to surpass the physical restrictions of the traditional camera apparatus (2013, p. 254).

Lubezki explains that their intention was to keep their shots 'elastic' through the utilisation of fluid long takes; "for example, to have a shot start very wide, then become very close, and then go back to a very wide shot" (B, 2013, p. 36-37). To achieve such an effect, "we needed the freedom of a virtual camera", says visual-effects supervisor Tim Webber, "so we created a virtual world and then worked out how to get human performances into that world" (ibid, p. 38). In an interview for *American Cinematographer*, Lubezki revealed that the actor's faces behind their helmets are the only real element in the exterior scenes; all the other elements are CGI, which according to Lubezki is "part of what cinematography is becoming" (B, 2013, p. 36-37). Cuarón acknowledges that their intention was to "surrender to the environment of space", and the only way to achieve that was through the utilisation of such

technologies (ibid, p. 38). Referring to the impression of reality that they were able to fabricate through virtual cinematography and CGI, Cuarón stresses that "not much would have changed in terms of the visuals" if the film was actually shot in outer space (ibid).

Certainly, whereas their long take aesthetic offers the opportunity for a consistent sound spatialisation, traditional editing obstructs the fulfilment of the cinematic illusion. In order to enhance the cinematic impression of reality, 3-D sound provides a lifelike representation of the camera's (or the character's) immersive point-of-audition, matching the framed visual perspective that is projected on the screen with its corresponding three-diemsional aural perspective. Thus, if a speaking character were seen on screen, their voice would accordingly emanate from the front speakers. However, if what follows is a cut to a different perspective and hence the same speaking character were behind the camera (in the offscreen space), their voice would have to emanate from the rear speakers, which could consequently produce a rather distracting effect.

In *Gravity*, a sort of "sonic teleporting" (Liang, 2016) is evidenced, for instance, in the scene that presents a conversation between the two protagonists as they make their way to the International Space Station (00:24:30). Cutting between shots becomes more recurrent through this particular scene, thus the characters' voices constantly switch from the front to the rear speakers, which in a way affects the fluidity of the soundtrack. Certainly, as suggested by Liang, the film's three-dimensional audio-visual match finds significant resistance when cuts are employed (2016, p. 10). Nonetheless, as Kerins (2011, pp. 173-174) suggests, in as much as audiences learn and adapt to new filmmaking styles, the key is to teach them on how to watch and listen to each particular film. As Richmond puts it, the opening sequence of *Gravity* works as the 'initial learning level' commonly found in video games, "teaching the player the fundamental mechanics of the game" (2016, p. 124). Hence, the so-called sonic teleporting may be more easily assimilated in *Gravity* as Cuarón 'trains' us from the very beginning on how to understand the movie's audio-visual contract. In this regard, re-recording mixer Dan Johnson comments:

> If it feels that it has been done deliberately for the whole thing, then you buy into it. It is like any creative thing, you have to set it up, you set the way it is going to work. […] But it is not a sort of thing that you can just add on at the end (personal interview, 2019).

In summary, it is evident that 3-D sound benefits from a visual style that rethinks the way planning, shooting, and editing practices are conventionally executed. Such a style is

aesthetically characterised by an extensive utilisation of long takes, a resource that Cuarón normally complements with the employment of deep-focus cinematography, wide-shots, and slow camera movement. These four stylistic traits illustrate in general terms the visual style of *Gravity,* a film in which the Bazinian aesthetics of cinematic realism have found a contemporary ambassador. The cinema of Alfonso Cuarón deviates from the conventional intensified continuity and presents an alternative visual style, which grounded on the long take, opens the possibilities for expanding the filmic space through a truly immersive and three-dimensional sound design. Ultimately, moving away from Bordwell's intensified continuity, such long take-driven aesthetics, combined with the immersive action of 3-D sound, shapes what from now on will be referred in this thesis to as *immersive continuity*, a film style that is drawing us near to the mythical experience of total cinema.

Conclusion

This chapter has provided some illustrative examples that explain how 3-D sound may transport the spectator into the narrative world, and the ways in which it may facilitate a greater identification with characters. Although it is undeniable that other factors also contribute to the experience of transportation (e.g. the power of the narrative, the actors' performance, or the exhibition environment), this thesis suggests that a realistic deployment of 3-D sound may be an important addition for the cinematic illusion, which, as argued in Chapter 2, can facilitate transportation and its effects on the spectator.

As pointed out by Manolas and Pauletto, a full and coherent deployment of 3-D sound requires the consideration of sound throughout the entire filmmaking process, which can affect directing, cinematography, and editing practices (2009, p. 43). It is understandable therefore that filmmakers may be resistant to challenges to their traditional ways of making films, as this is only one approach among a variety of creative possibilities. It is remarkable, nonetheless, to witness new aesthetics forming, and how examples of new audio-visual approaches are starting to be embraced by highly regarded practitioners.

The long take visual aesthetic found in *Gravity* stands out as a cinematographic mechanism that undoubtedly facilitates a realistic utilisation of 3-D sound. Together with techniques such as deep-focus, wide-shots and slow camera movement, the long take conforms

what this thesis has termed immersive continuity, a film style that is sonically characterised by its three-dimensional utilisation of sound design.

As object-based audio in the form of Dolby Atmos becomes the new standard for the production and delivery of cinematic sound, it seems likely that more and more filmmakers will consider new ways to marry the sound with the picture and new ways to consider sound during the whole workflow. Even though the application of spatialised sound is possible with more traditional formats (e.g. 5.1 or 7.1), the three-dimensional capabilities of Dolby Atmos might be an important factor for a broader adoption of the aforementioned audio-visual aesthetics.

As Sergi suggests "[a] powerful weapon such as Atmos can create a large disruption, hence a substantial window of opportunity for change" (2013, p. 118). It will be of great use to new filmmakers to understand how Cuarón and others have started to address these challenges. In the following chapter, *Roma*, Cuarón's most recent film, will be analysed in terms of the aesthetics and practical strategies that facilitated the creation of what this thesis considers to be a quintessential Dolby Atmos soundtrack.

CHAPTER FOUR

All Roads lead to *Roma*: The Consolidation of the Immersive Continuity Style

Since *Y Tu Mamá También,* Cuarón has demonstrated his interest for the long take, a cinematic resource that is now one of his most common characteristics. Cuarón acknowledges that his fascination for the long take began in film school, and explains that it is "the unity of time and space" and "the relationship between foreground and background, character and environment" what drawn him to this technique (Netflix, 2018c). In relation to the visual style of *Children of Men*, for instance, Cuarón argues that he used "fluid long takes to take advantage of the element of real time" (Udden, 2009, p. 34). Similarly, Cuarón acknowledges that the *Hubble 3D* (2010) documentary was his biggest reference on creating *Gravity* (IMax, 2013), a film in which realistic long takes with deep-focus play a fundamental role. Moreover, *Roma*, his most recent production, may be compared with the classic black and white films of the Italian neo-realism, a 1940s film movement in which the long take aimed to portray the continuous reality of the poor, in their own location and with non-professional actors. It is thus curious that the title of the movie brings to mind Roberto Rossellini's masterpiece, *Roma Città Apperta* (1945). But leaving aside all the comparisons and influences that Cuarón's latest film might have, *Roma* signifies the consolidation of this filmmaker's cinematic style; and as he admits, it is the first film that truly embodies the kind of cinema that he aspires to make (Cuarón in *Road to Roma*, 2020, original translation), a cinema that beyond the search for authenticity, aims to transport us into the narrative world through a highly realistic audio-visual illusion.

As argued previously in this thesis, the Bazinian cinema of Alfonso Cuarón transcends the visual realm, and it is his three-dimensional sound design approach that distinguishes him from the majority of contemporary directors. Although the absence of colour contradicts the myth of total cinema that Bazin (1967) once idealised, the soundtrack of *Roma* 'colours' the image with sounds that vividly move all around the diegesis. In reality, as Chion argues, human vision is partial and directional like the screen in the cinema; hearing, however, is omnidirectional (1999 [1982], p. 17). "We cannot see what is behind us, but we can hear all around" (ibid). Cuarón's audio-visual style seems to fully embrace this basic principle,

resulting in a highly realistic approach to sound design. In *Roma*, 3-D sound helps to represent vividly a past reality, a life-like world to which the spectator is transported by means of, among other factors, the sensation of presence in a space created out of the director's own personal memories. Cuarón comments:

> I wanted then to place the camera in a physical environment, for the camera to be surrounded by the sounds of that space. When we are watching the film, we're not only seeing what is onscreen, but the existence of that place that isn't on the screen. (in O'Falt, 2019)

Cuarón, *Roma*'s writer, producer, director, cinematographer and editor, admits that after *Gravity* he had the opportunity to do bigger movies, but he decided to go back to his country to make a more intimate film using the tools and the knowledge that he had gathered throughout his filmmaking career (Netflix, 2018c). Ultimately, following Cuarón's creative evolution, one may poetically say that all roads led to *Roma*, a film in which the immersive experience of total cinema has practically left the myth.

In this chapter, *Roma* will be analysed aiming to establish the immersive continuity style in terms of its aesthetics and the practical filmmaking strategies that permitted its effective application. Based on interviews and testimonies of key practitioners involved in the film's creative process, this chapter shall attempt to give insights about how to create films with three-dimensional sound design. Finally, the last section will discuss the pros and cons of using Netflix for distributing a film with a quintessential cinematic soundtrack.

The Immersive Continuity of *Roma*

Although Cuarón's interest in realism had been previously exhibited in films like *Y Tu Mamá También*, *Children of Men*, and *Gravity*, *Roma* is undoubtedly where his search for Bazinian authenticity transcends all boundaries. Based on his personal memories, the film tells the simple but truthful story of Cleodegaria "Cleo" Gutiérrez (Yalitza Aparicio), an indigenous young woman that works as live-in maid for a middle-class family of white Mexicans. The main characters, played mostly by non-professional actors, depict each member of Cuarón's own family, and Cleo plays the role of Liboria Rodríguez, 'Libo', the family's maid to whom Cuarón dedicated the film. Although Cleo is certainly the protagonist of the film, *Roma* is really a story about Mexico City, their people, their identity and social problems.

As her employers, Sofía (Mariana de Tavira) and Antonio (Fernando Grediaga) pass through a separation and eventual divorce, Cleo gets pregnant, but her boyfriend, Fermín (Jorge Antonio Guerrero), refuses to acknowledge that the baby is his and disappears. Cleo receives acceptance and help from her employers while waiting for her baby to come. On a day of student demonstrations, Cleo goes to a furniture store with Teresa (Verónica García), Sofia's mother, and by accident sees Fermín again, who as a member of the paramilitary group 'Los Halcones', enters the store looking for a young man that gets killed by one of the armed men. As a product of the stressful moment, Cleo's water breaks. She is taken to the hospital, but unfortunately her baby is stillborn. After a few days, Cleo goes on a holiday to the beach with Sofia and her children, where Cleo saves two of them of being carried of by a strong current in the sea. The love for each other gets stronger, and live goes on.

- **The 'Look' of *Roma***

Adding to the aforementioned aspects of authenticity, the realism of *Roma* is supported by its 'look'. In Chapter 3, this thesis coined the term immersive continuity to refer to a film style that, grounded on the long take, deviates from the quick-cutting conventional norms of narrative continuity. Following a Bazinian cinematic aesthetic, the long take stands out as the core of *Roma*'s visual style. The film includes 303 shots, with the average shot length (ASL) being 26 seconds, coincidentally the same ASL found in *Gravity*[28]. Comparing those numbers with the 3 to 6 seconds average found in contemporary Hollywood (Bordwell, 2002, p. 17), *Roma* clearly stands out as another unconventional film. Long takes convey time and space without the human manipulation of montage, a resource that is minimally used in *Roma*. The opening sequence (00:00:47), for instance, presents the story without a single cut for 4 minutes and 34 seconds, the childbirth scene (01:42:51) is presented in one continuous 4 minute 11 seconds take, and the beach scene near the end of the movie (01:58:07) comprises a single 5 minute 29 seconds shot.

Although the majority of scenes use montage to present the story from different perspectives, Cuarón's long take aesthetic is in evidence throughout the whole movie with a total of 113 shots that last more than 20 seconds each. Broadly speaking, cutting is used in

[28] See Chapter 3.

Roma only to facilitate the transition from one location to another—from the patio to the living room, from the living room to the kitchen, from the kitchen to the outdoor street, and so on—however, the action that is happening in each location is frequently presented within continuous shots. Cuarón's disinclination to use traditional montage is illustrated for instance during the first ten minutes of the film (from 00:00:47 until 00:10:14). Here, montage is strictly used as a 'teleporting' mechanism, that is, cuts are used to move the viewer from space to space, but not to separate each space into smaller fragments.

Furthermore, without relying on close-ups, Cuarón constantly presents the action through wide-angle lenses and depth-of-field, a cinematographic resource that Bazin described as "a dialectical step forward in the history of film language" (1967, p. 35). Such a veneration for the spatiotemporal continuum of reality is supported by slow camera movements that follow the subjects as they move and interact within the diegesis. Ultimately, with long takes, wide-shots, deep-focus and a subtle camera movement, Cuarón honours time and space and allows the viewer to freely observe all the layers of the *mise-en-scène*. All these visual resources can be illustrated, for instance, by analysing some of the shots that are part of the introduction.

At the beginning of the film, Cleo is cleaning the patio floor (00:00:47), of which the camera shows only the running water during a three-minutes titles sequence, that among others credits the work of José Antonio García (Direct Sound), Sergio Díaz (Sound Design), Skip Lievsay (Sound Design and Sound Mixing) and Craig Henighan (Sound Mixing), suggesting the degree of importance that sound will have throughout the movie. Then, the camera tilts up and we see a wide-shot of Cleo working near the garage door (Figure 4.1). As she walks towards us, the camera pans left showing a small patio occupied by a few plant pots, a couple of buckets, a washing machine, bicycles, a dog (Borras), a few birds, and several other objects, all of them clearly seen as meaningful elements of the scenario. The camera stops its rotational movement when it reaches the door of the exterior toilet—the maid's toilet—in which Cleo enters. For a few seconds the camera stays still showing only the movement of the birds playing inside their cages. Once Cleo exist the toilet, the camera pans right framing the protagonist, who enters the house through the kitchen's door.

Figure 4.1 - Frame taken from *Roma*. A wide-shot of Cleo working near the garage door.

Figure 4.2 - Frame taken from *Roma*. The camera slowly pans from left to right.

After the opening long take, a first cut is used to transport us to the ground floor of the house (00:05:21). For 16 seconds, the camera slowly pans from left to right showing two dining rooms and the area below the stairs through which Cleo ascends (Figure 4.2). A second cut is then used to show us the upper floor, the space in which the bedrooms are distributed (00:05:37). Following a similar pace and movement, the camera pans slowly from left to right as Cleo carries on doing her daily activities. As the camera reaches the other end of the floor, we see Cleo making the bed inside a bedroom. The camera finishes its trajectory on the right side and bounces towards the left after Adela (Nancy García García), Cleo's colleague, shouts reminding her that it is almost 1pm. Cleo quickly picks up all the dirty clothes and runs.

Figure 4.3 - Frame taken from *Roma*. Deep-focus is used to show Cleo putting clothes into a washing machine in the patio.

Figure 4.4 - Frame taken from *Roma*. Through a wide-shot we see Cleo, the neighbours' made and the honey vendor that passes by chanting his selling call.

At this point, a third cut is employed to take us back to the ground floor (00:06:35). The camera keeps on panning towards the left as Cleo goes down the stairs and runs towards the kitchen. As the camera reaches the dining room at the left side of the house, deep-focus is used to show Cleo putting clothes into a washing machine in the patio, which we observe from inside the house (Figure 4.3). She takes off her kitchen apron and runs. Montage is used once again to teleport us to a new spatial perspective, this time we see the front door from outside the house (00:06:58). Cleo gets out and rapidly puts a sweater on as she runs to pick up Pepe (Marco Graf), the youngest child in the family. The rotational movement of the camera follows her trajectory and offers a first glance of the neighbourhood. As she vanishes in the distance, we see the neighbour's maid sweeping the exterior of the house next-door, and a honey vendor chanting his traditional selling call, all of which we see through a wide-shot (Figure 4.4). Cutting is used for the fifth time to relocate us outside 'Kinder Condesa' (00:07:16), showing

Cleo and Pepe exiting through the main door of the boy's kindergarten. Although they are the two characters that we follow, we are equally able to see the rest of the people that are framed within the same wide-shot.

Throughout the introductory sequence described above, there are four fundamental resources that can be distinguished as part of the overall visual aesthetic of the film. First, with a substantial reduction in the number of cuts, the long take is present as the main visual strategy that Cuarón uses to represent reality and its spatiotemporal continuity. This visual device is complemented by a second fundamental resource, depth-of-field, which is extensively employed to provide the same amount of detail for all the visual elements located within three layers: foreground, mid-ground and background; hence, as Bazin argues, the spectator "is called upon to exercise at least a minimum of personal choice" (1967, p. 36). Third, wide-shots are extensively used to show not only the main subjects but also the space within which they move and interact with others. And fourth, slow camera movements are constantly used to navigate with the subjects as they move within the diegesis. Regarding the latter, it is worth noticing that besides using camera panning, movement is also executed in the form of tracking shots, which are seen at several points throughout the movie as a tool to continuously present the space as the characters transit within it.

For instance, when Cleo and Adela run towards 'La Casa del Pavo' restaurant (00:21:53), a tracking camera follows them with a wide-shot that fluidly presents the space that they inhabit (Figure 4.5). Similarly, a tracking shot is used during several scenes in the hacienda (00:55:54; 01:03:29; 01:06:00); when Cleo and Señora Teresa (Verónica García) go to the cinema with the children (01:10:51); when Cleo, Señora Teresa and Ignacio (Andy Cortés) walk towards the furniture store (01:32:55), and in several other moments throughout the film. Cuarón comments:

> The choice to pan so much was a result of the format. But I knew that the film had to remain objective, and there is nothing more objective than a tracking shot from afar. […] I wanted to keep that distance to stay objective, and let the moments add up so as to convey emotions, and also evoke empathy, which was a side effect. The film doesn't follow one character's subjective point of view. I didn't want it all to be seen from Cleo's perspective. Cleo's only another character in this wider universe. The story is about this universe. The characters only traverse it (Cuarón in *Road to Roma*, 2020, original translation).

Figure 4.5 - Frame taken from *Roma*. A tracking camera follows Cleo and Adela. A wide camera angle is used to fluidly present the space that they inhabit.

In summary, following the stylistic tactics previously employed in *Gravity*, the long takes in *Roma* are normally complemented by the utilisation of deep-focus cinematography, wide-shots, and camera movement (either in the form of panning or tracking shots) four aesthetic characteristics that visually shape what has been referred here to as the immersive continuity style.

- **Immersive Continuity and the Three-Dimensionality of Sound**

Following the sonic approach introduced in *Children of Men* and fully deployed in *Gravity*, all the sound elements in *Roma* move in accordance with the position of the sources in the filmic space. For instance, if a character is behind the camera, their voice is coherently heard emanating from behind, or if the camera turns around, the whole sonic environment turns around with it. This extremely unconventional use of immersive sound reinforces the existence of a point-of-audition from where we, as members of the audience, experience the events that are happening, not only in front of us, but also around us. As such, the soundtrack of *Roma* enlarges the filmic space and endows the camera with a complete representation of natural human hearing. Hence, sound not only provides all the details that the framed black and white screen is not able to convey, but ultimately, it helps to transport the audience member into a vivid reconstruction of 1970s Mexico City. Cuarón stresses:

> The film is as much sound as it is the visuals, it's about the atmosphere around. Each place has a complete different sound because Mexico City is like that, and that's something I wanted to honour. Some of those sounds have prevailed, and after all these years you keep on hearing them. When I'm watching the film I wanted to feel that I'm back into that reality of Mexico in 1970. (Cuarón in *The Sound of Roma*, 2020)

Certainly, the soundtrack of the film acts as a sound field reconstruction that allows the spectator to experience how Mexico City sounded in the 1970s. Moving beyond the limits of the screen, sound turns the film into some sort of 'audio-visual installation', immersing us not only within the story of a family, but also within the history of a nation, with its identity, culture and social problems. Such a sociohistorical connotation is showcased, for instance, in one of the most intense sequences of the movie (from 01:31:50 to 01:36:43). On 10 June of 1971, a tragic event in the history of modern Mexico happened, a violent repression against a group of student demonstrators known as 'El Halconazo' or 'The Corpus Christi Massacre'. The brutality of this event is carefully portrayed in this sequence, showing how a group of paramilitaries murdered several protesters in cold blood.

The sequence begins with the camera located inside the car as Ignacio, Cleo and Señora Teresa go to the furniture store. The shouts of a massive group of demonstrators is heard from a distance, while the footsteps and voices of specific characters are heard around us as we go along the street. The direction of the voices, noises, shouts and student chants change as the camera moves and with each visual cut (there are only 11 cuts throughout the 5-minutes sequence), moving from the front to the back or sides in accordance with the position of the sources of sound in the three-dimensional diegesis, creating this way a vivid aural environment that surrounds us as the main characters converse.

This treatment of the soundtrack continues throughout the whole sequence (and throughout the whole film). When the tracking camera follows the main characters as they walk among the crowd, for instance, we hear the voices of specific groups of individuals moving from right to left, such as the voice of the police officer that appears on screen as he gives instructions to his subordinates. Then, at the beginning of the furniture store scene, we firstly hear the voice of Señora Teresa coming out of the rear left corner, where she is physically located while the camera shows the store's window, a position that changes when the camera cuts to a cabinet full of table clocks. As Señora Teresa negotiates with a seller for a better price, the shouts of the demonstrators are heard behind us from a distance, a noise that increases in volume as the peaceful demonstration becomes violent. The camera slowly rotates towards the right accompanied by its corresponding sonic rotation, and as it frames a close-up of Cleo and her employer, the voice of a frightened student asking for help is heard coming from the rear right corner, the zone in the diegesis from where he enters just before being killed by a group of armed men.

Throughout these and other sequences, the inseparable relation that exists between characters and environment is represented in great detail, presenting a series of individual narratives that are happening simultaneously within a space that tells a story on its own. The idea, says *Roma*'s supervising sound editor and sound designer Sergio Díaz, was "that the audience felt within each geography of the film" (personal interview, 2019, translated by author). Skip Lievsay, re-recording mixer and sound designer of the film, who had previously worked with Cuarón in *Y Tu Mamá También* and *Gravity*, stresses:

> Basically, the screen and the picture become a doorstep. And when the audience starts to get involved in the sound, they move to the edge of the door. And then, if you really are carful with the way you do the panning […] at a certain point the audience literally steps in through the door. And now they are in a place with the actors. (personal interview, 2019)

Lievsay explains, however, that this sonic treatment does not always work "because most films have more movement in the image than the audience can bear" (ibid). Hence, if a sound designer tries to follow that movement (or more precisely those visual cuts), Lievsay admits that the experience can become quite distracting (ibid).

Díaz argues that the way the film was conceived and shot allowed them to work in that manner (personal interview, 2019, translated by author). Lievsay acknowledges that they were able to fully exploit sound three-dimensionality because Cuarón "had set this up with his photography, with long takes and slow moves, and not a lot of cutting" (O'Falt, 2019). In other words, each shot is carefully conceived considering its position and orientation within the diegesis, which in terms of continuity implies thinking of the relations that will exist between shots when three-dimensional sound is added. Adam Gough, picture editor of the film mentions that he always knew that if there was ever a cut Cuarón would always want sound to change accordingly. Gough comments:

> And how he explained that to me is the camera is also the microphone. If we cut to a different angle, the sound was panned accordingly, and everything would sound the same so everything was tilt and shift with the camera on the edit. So, if we cut from inside for outside you hear the obvious change in atmosphere (in *The Sound of Roma*, 2020).

That is to say, Cuarón does not conceive the camera simply as an optical and unidirectional recording device, but as an invisible entity that sees the world in one direction, but is capable of hearing in an omnidirectional way. In *Roma*, says Cuarón, "the camera was this ghost of the present, traveling to the past, and just hovering, observing everything,

honouring time and space" (Netflix, 2018c). Hence, through a coherent audio-visual match, Cuarón succeeded in creating what Lastra calls the "invisible auditor" (2000, p. 159), an entity that we embody as we get transported into the cinematic universe. Ultimately, with his most recent film, *Roma*, Cuarón has consolidated immersive continuity as a film style that, grounded on the long take, permits a complete exploitation of three-dimensional audio tools.

The Sound of *Roma*: Unveiling the Creative Process

As the counterpart of his characteristic deep-focus cinematography, Cuarón has revealed that his intention with *Roma* was to sonically represent all the different visual layers "that you see from foreground, mid-ground and background" (Cuarón in Netflix, 2018a), that is, Cuarón wanted to complement the visuals with some sort of sonic depth-of-field. His idea was to emulate what we experience in our daily life, where "our visual perception is just a fraction of the spatial perception that we have—because part of our spatial perception is filled in by the sound" (Cuarón in Hunt, 2020). But in order to achieve such levels of realism, Cuarón and his collaborators followed a creative process that considered sound thoughtfully from beginning to end.

Obviously, Díaz, Lievsay and Henighan, in collaboration with the whole sound department, were the ones that sculptured the soundtrack of the movie. Yet, it is worth acknowledging that Cuarón was who created the film with sound in mind. "The richness of the audio track", Lievsay admits, "comes from his dedication to the fact and to realism" (Lievsay in *The Sound of Roma*, 2020). In the screenplay, Cuarón describes the scenes with sentences like: "Cleo washes clothes, occasionally humming to a song coming out of a yellow radio"; "The wind carries the hum of different radios and dogs barking"; or "The entrance to the movie theater is a bazar of STREET VENDORS hawking their wares, reciting lists over and over like a litany, like mantras in a chant" (Cuarón, 2016). As an example, Díaz comments that the credits sequence at the end of the film incorporates many of those sounds that are the essence of the city. "Vendors on the street, cars on the street, avenues behind us, families talking, kids playing, planes over our sky, and water running […] It's like a spiritual connection with the world, with the universe, and life continues, that's it" (GoldDerby, 2018). The script for this section reads as follows:

> The patio, in shadows now, floats in the afternoon quiet. Borras sleeps and the parakeets are quiet. Only the hum of the city in the distance.
>
> Cleo comes out of the kitchen carrying her load of dirty laundry and crosses the tiny patio to go up the metal staircase that leads to the roof.
>
> Her steps reverberate throughout the bony structure in a metallic moan that echoes through the tiny patio, waking the caged birds.
>
> Cleo reaches the step in front of her room and keeps walking upwards. A sweet potato vendor lets out his sad howl in the distance.
>
> Step by step, Cleo ascends. Yet further up, beyond the roof, the sky is pure. (Cuarón, 2016)

As illustrated in this short passage, *Roma* is undoubtedly a film that was scripted with sound in mind, a condition that, as sound designer Randy Thom (1999) explains, is essential for taking full advantage of sound. In *Roma*, representing a realistic aural diegesis was a fundamental directorial decision that the sound department had to creatively materialise. Díaz comments:

> The biggest challenge was to enter Alfonso's memory and bring that to the screen. Secondly, it was making a really immersive soundtrack, so the audience felt within each geography of the film. And having those elements separately was essential, because from the beginning the film was conceived to be mixed on Dolby Atmos. So, we had to have all the elements separately to be able to manipulate them in the mixing room as the director wanted. I mean, if we are in a scene where we have many sound elements playing at the same time, we had to have this 360-degree distribution in the room so that all of them coexisted in such a way that you as an audience member felt part of that geography, you were not isolated from what is happening in the image. So getting all that was a real challenge and it took us a lot of time to collect, explore, edit and select those elements. (personal interview, 2019, translated by author)

The dynamic 3-D environments of the film are hence the result of the combination of individually captured audio objects, an approach that fully exploits the object-based functionality of Dolby's latest sound platform. The final sound mix was thus entirely constructed out of individual sound elements that Díaz and his sound team collected throughout the entire process (Díaz, personal interview, 2019). As a filmmaker who gives sound the importance that it deserves, Cuarón permanently directed the sound team aiming to reconstruct the soundscape of his memories. Yet, he hired Díaz to be there during the whole process, who as supervising sound editor was the person responsible for the final soundtrack. Considering that apart from postproduction, Díaz worked full-time during the 18-months of shooting, plus 5 to 6 months of his own research and sound collection during preproduction (Díaz, personal

interview, 2019), it is evident that the creators of *Roma* followed a workflow within which the processes of sound preproduction, production and postproduction were effectively harnessed and coordinated.

It is important to mention, nonetheless, that Cuarón did not give the screenplay to Díaz during preproduction; in fact, the only thing Díaz knew was that the story "was going to happen in 1970-1971 in the *Roma* neighbourhood" (Díaz, personal interview, 2019, translated by author). So he took the task of going to different cities in Mexico for collecting sounds that would serve as a basic sound library for the film (ibid), which, as sound editor Javier Quesada comments, "became an ever larger library of sounds that did not belong to 2017 or 2018, but had to sound like 1971" (personal interview, 2019, translated by author). Sound editor Enrique Greiner describes that such a sound library was very well organised, each car, for instance, had its own model, its year, and a picture attached to each sound (personal interview, 2019, translated by author).

Sound editor Javier Quesada asserts that given that they knew right from the beginning that the final mix was going to be done using Dolby Atmos, anything that was "put in the session had to be something that could be moved up/down, panned […] Everything was treated separately" (personal interview, 2019, translated by author). Such a recording strategy was also applied to the characters' dialogue. In fact, as re-recording mixer Dan Johnson comments, mixing in that manner would not have been possible if for example there was more than one character recorded simultaneously on the same boom microphone (personal interview, 2019). Sound editor Enrique Greiner comments:

> In Atmos, what you have to do is precisely to record each thing separately, because you need them clean. It is like a sort of 'photoshop', you have to cut and paste each little thing to make a collage, but it cannot have absolutely anything in the background. Like a green screen where you have to cut everything exactly, otherwise you would not be able to handle it; the background noise that would exist if you record it wrong is going to reveal that it is a recording that you are moving. So, you have to really use it with elements separately as it is in real life, and behind that put some kind of background. (personal interview, 2019, translated by author)

During shooting, therefore, Díaz and his sound team were in charge of capturing each character individually, plus other location sounds that could be useful for the final mix, commonly known as *wild tracks*. For the 'Halconazo' sequence previously analysed, for instance, Díaz recalls placing microphones at several positions around the location to record the atmosphere from all the different perspectives that they could use during the mix, managing

this way to record more than 2000 people (Díaz in *The Sound of Roma*, 2020, original translation). Díaz elaborates:

> Alfonso was upstairs in the furniture store, and he gave me instructions from there about what to record. At that point, I was in the middle with a huge microphone. And way at the back, by the subway, we had another team. Then there was another team behind me, two more hidden in the cars. And on the roof of the furniture store, there were other teams recording the sound from different angles. [...] Obviously, José [García, Roma's sound mixer] was in charge of what was going on inside the furniture store. Each actor had their own microphone. And later on, we made these things what we call 'wild tracks' […] It was a real advantage to have all this material from different perspectives and angles, to later be able to create a combination during the editing process. (Díaz in *The Sound of Roma*, 2020, original translation)

Díaz recalls setting up an audio workstation in the production offices. After each day of shooting, he would receive the selected shots in order to clean up the dialogue and incorporate other sound elements that he had already recorded during preproduction or as part of the audio recordings of each day. Díaz would then arrange a meeting with Cuarón and present his work. "He [Cuarón] walked into the office", Díaz recalls, "we both sat in front of a computer, both wearing headphones, and basically I showed him the shots, the scenes that he liked, and started doing development exercises during the shooting process" (Díaz, personal interview, 2019, translated by author).

Figure 4.6 - Frame taken from *Roma*. Cleo is shown among a large number of people, each of them with their own personal story.

Moreover, apart from recording all the possible audio elements on location, Díaz was in charge of gathering other sounds that would complement the ones recorded during shooting or in the preproduction stage. In the case of the street vendors, for instance, Díaz explains that they searched for the people that still remembered how to do those very traditional chants (Cine

Premiere, 2019). A good example is the scene outside the Cine Metropolitan (00:40:11), in which Cleo realises that Fermín (Jorge Antonio Guerrero) had left after receiving the news of her pregnancy. The city does not stop as she shyly exits the theatre hoping to find her lover among the crowd. Although she is surrounded by hundreds of voices, Cleo is aware of her loneliness. The chaotic atmosphere of the location immerses us within the scene and transports us to that specific moment in time and space, in which Cleo is only one individual among all those people, each of them with their own personal story (Figure 4.6).

In this scene, Díaz notices, "there is the balloons vendor, the one who sells the skeletons, the one who sells meringues, […] each couple has a specific dialogue, they carry their own story […] The best way to isolate each sound is recording each character that you see separately" (Cine Premiere, 2019). Hence, Lievsay recalls, they had a recording session with "350 actors coming to record voices", whose lines were all written by Cuarón (Netflix, 2018a). Díaz elaborates:

> So, we called a lot of people into the dubbing studio, and they had to play out the scene so that during the editing process, we could assign each one of them to different groups. Because Alfonso didn't want to just have murmuring in the background like 'walla.' He hated that word. 'Walla' didn't exist. We needed real conversations, like you have in real life. That session lasted more than three weeks. It required its own logistics. We gave everyone a number, and we had a five-microphone configuration. We had five booths, and we would put four to five people in each one of them. (Díaz in *The Sound of Roma*, 2020, original translation)

A similar strategy was employed for the beginning of the 'Halconazo; sequence (01:32:00). Díaz explains that along this particular scene there are people talking all around the street, and "each one of them is having their own story" (Cine Premiere, 2019). Aiming to manipulate each person independently, they recorded their dialogue separately in the studio so that in the mixing room they could be able to localise each specific sound in the particular position of its source (Diaz, personal interview, 2019).

This object-based approach was used for all the other sounds that would shape the aural environment of the city. In terms of the diegetic music, for instance, sound editor Enrique Greiner recalls:

> There was the rule, or the wish that when they were walking through the streets all the houses should have the window open. And in each window, there should be a different radio playing, for example; or that every passing car should have a radio on. Then I had

> to be checking lists and lists of songs that were hits in 1971, avoiding any repetition. (personal interview, 2019, translated by author)

Although sounds were gathered during the whole creative process, most of these sound-related activities took place in parallel with the process of picture editing, during which Díaz and his sound team worked on individual scene clips, whose soundtrack was gradually designed and edited according to the indications provided by the director in collaboration with his co-editor (Díaz, personal interview, 2019, translated by author). "[Cuarón] would say 'scene 6, 7, 8, 9 are already approved, this is the shot, send it to Sergio", Díaz recalls (ibid). He and his team worked through a server that he had installed, from which they collaborated and fed. There were sound editors from all the departments, "dialog editors, ADR editors, effects editors, Foley editors, ambience editors, and from all disciplines. And as each scene clip evolved, I returned it to the editorial department through Adam Gough, who was Alfonso's co-editor." (Díaz, personal interview, 2019, translated by author). Once Cuarón and Gough received the material, they would always give feedback and suggestions. Díaz comments:

> I was sending different stems, I mean sound elements already mixed, for example a dialogue track, an ambient track, a Foley track, a music track and everything that had been done, so that they could then upload to their editing platform (Avid). And from that moment Alfonso already gave us many notes, what he did not like, what we should change and what we should improve. So, we went through that process for many months during editing. And then, towards the closing of the picture editing process, Skip Lievsay joined us in order to have an overview of everything that had already been built, and they had a screening with Alfonso in a proper movie theatre to know whether we are covered, do we need anything else? how much do we have left? are we making progress? what do you suggest? And get into that conversation. (Díaz, personal interview, 2019, translated by author)

During the final mix, all those sounds were then layered in the form of objects that would be located in the correspondent position of each specific source, constructing this way our immersive point-of-audition within the three-dimensional diegesis of the film. "It was like a relay race", says Díaz, "one character enters, another exits, another comes, and you are hearing the whole cacophony in 360-degrees, within which you are truly immersed with what is happening from Cleo's perspective" (Cine Premiere, 2019). Ultimately, throughout all the way, the creative process aimed to take advantage of the immersive capabilities of Dolby's object-based sonic innovation; and, according to Cuarón, a film like *Roma* lent itself to it. "When we decided to engage with Atmos for *Gravity*' says Cuarón, 'it was a fantastic experience but I kept they saying the real use for Atmos is going to be intimate films" (Cuarón in The Sound of Roma, 2020).

It is not surprising, therefore, that Dolby Laboratories and its personnel constantly highlight the work of Alfonso Cuarón as one of their more important creative achievements.[29] Dolby consultant Robert Karlsson, for instance, who was in charge of providing technical support during the mixing of *Roma* and a variety of Dolby Atmos films, acknowledges Cuarón's latest production as one of the few motion pictures to have fully used Atmos' capability of localising sound objects in an effective way (personal interview, 2019). In this regard, Díaz argues that Dolby Atmos "works much better for projects that are 'organic', meaning those that are not gimmicky, because it puts you in a very real context within the sound universe, and puts you in context of what is happening in the film" (personal interview, 2019, translated by author).

It is important to mention, however, that in order to achieve the level of sonic detail evidenced in *Roma*, Cuarón and his collaborators invested a great amount of resources and hours of work for planning, recording, editing and mixing the soundtrack. Díaz admits that generating all the sounds needed for each specific scene was

> a real challenge because it meant multiplying what you normally do in a project, where the mix is 5.1 or 7.1 at most, but here that was only the base; from there you have infinity of objects that have to be spread independently within the whole room. So, we had to be covered in all the geographies of the film, so you can imagine the number of sound elements and the sound library that we created expressly for *Roma*, it is huge! (personal interview, 2019, translated by author)

Certainly, as sound practitioner Emma Butt comments, in order to do a complex Atmos mix like the one of *Roma*, "the sound team needs more time to cut in more atmospheres and more time to think about. If I'm standing on the street, what am I going to hear from above me, around me, below me? What do I need to cut in to make that happen?" (personal interview, 2019). Regarding the final mix of *Roma*, Lievsay admits that it is common to get only a couple of weeks to do everything, whereas for *Roma* they needed 10 weeks of full-time work (personal interview, 2019). So, as he puts it, "the question is do you have the patience? Is it important enough to invest energy in that idea?" (ibid). Perhaps the answer is to find a balance between perfection and the resources that are needed to achieve it. As Lievsay recalls, they completed the final mix of *Gravity* in two weeks plus an additional week for the Dolby Atmos mix, "so

[29] For example, Robert Karlsson, Rob France, Glenn Kiser acknowledged that during their personal interviews with the author of this thesis.

as busy as that was, with a really solid idea […] you can do something like *Gravity* in two weeks" (ibid).

Certainly, as Randy Thom comments, Dolby Atmos "is potentially a great tool", but "we need to figure out how to use it well […] In order to take full advantage of it you need to write for it, to design a scene for it" (Thom in SoundWorks Collection, 2015). It seems clear that with *Gravity* as a first successful experiment, now consolidated in *Roma*, Cuarón has showcased an audio-visual method that exploits the full potential of such an immersive sound system. Based on a script written with sound in mind, and a visual aesthetic grounded on the long take, the creative process that Cuarón and his sound team have followed is exemplary and hence worth to be studied. His filmmaking approach, not only provides the framework for achieving a complete exploitation of present-day cinematic audio tools, but beyond, Cuarón's methodology stands out as an alternative for the creation of fully three-dimensional, realistic and transporting cinematic pieces.

Netflix and the Cinematic Experience at Home

As Cuarón acknowledges in the Netflix featurette '*Roma*: The Sounds of Real Life', the film is made as much of sounds as it is of visuals (Netflix, 2018a). Emphasising on the utilisation of Dolby Atmos to create a realistic sonic universe, this 3-minutes advertising piece promotes the soundtrack as a fundamental element for the immersive experience of the film. Yet, how many people were able to experience *Roma* within the enveloping infrastructure of an Atmos-approved theatre? The truth of the matter is that *Roma* qualified for the Oscars and became one of the two most nominated films for the 91st Academy Awards (with 10 nominations) after a limited cinema release. This provoked a negative reaction from a group of Hollywood filmmakers, including Steven Spielberg, who commented: "I hope all of us really continue to believe that the greatest contributions we can make as filmmakers is to give audiences the motion picture theatrical experience" (as cited in Allen, 2019, n.p.).

In any case, it is worth acknowledging that although *Roma* got a minor theatrical exposure compared to other Academy Awards nominees, this film has drastically changed the way Netflix commonly releases its original content. Cuarón's agreement with the video on demand giant included an exceptional three-weeks theatrical opening before its release on their

streaming platform (McClintock, 2018), something that not all 'Netflix Originals' can have. For instance, the Coen Brother's *The Ballad of Buster Scruggs* (2018), another Oscar nominee, got only a one-weekend theatrical opening in four cities (ibid), and the vast majority of Netflix productions do not even have the privilege of a theatrical release (Sims, 2018). Scott Stuber, head of Netflix's film group commented:

> We love this film and our goal is to bring this cinematic experience to audiences everywhere, and our theatrical plan, combined with the reach of our service, has made *Roma* an event film and cultural moment […] *Roma* has played in festivals all over the world and we're expanding our theatrical engagements, not only in major markets, but also in places like Honduras, South Africa, and Iceland. (Netflix, 2018b)

The reason for such an exceptional release is that, as Cuarón acknowledges, "[t]he complete experience of *Roma* is unquestionably in a movie theater" (as cited in Sippell, 2018). As *Roma*'s sound editor Javier Quesada comments "a film that is made for Atmos and is seen on a television, perhaps with the same television speakers, will not translate in the same way […] It's not the same to experience the film in the cinema than to experience it at home" (personal interview, 2019, translated by author). Perhaps other Netflix Originals do not really require a theatrical experience, but a film like *Roma* demands it. It is hence discouraging that such a complete auditory experience, although exceptional, was relatively limited. While most commercial feature films are commonly screened in over 1000 cinemas in North America alone (Sippell, 2018), *Roma* was released theatrically in only 600 cinemas worldwide, 100 theatres in the United States and just over 500 theatres over 40 countries (Netflix, 2018b), of which not all included a Dolby Atmos infrastructure.

In defence of their distribution strategy, Cuarón argued that "[s]eeing *Roma* on the big screen is just as important as ensuring people all over the world have the chance to experience it in their homes" (Thompson, 2018). Cuarón highlights that the movie theatre undoubtedly offers the best possible experience as the film was created using state-of-the-art technology (65mm digital photography and Dolby Atmos sound); yet, he claims, *Roma* "was designed to be equally meaningful when experienced in the intimacy of one's home" (ibid). But if that is so, why did they occupy so many working hours for the creation of a three-dimensional Atmos mix? Perhaps the 65mm black and white photography would translate well enough on a modern TV screen, but the immersive Dolby Atmos soundtrack requires more than the simple unidirectional speakers that the common Netflix subscriber is more likely to have at home.

Undeniably, *Roma* is a cinematic film, it is a movie that can only be itself when experienced in a Dolby Atmos approved cinema.

In order to be able to hear a relatively transparent version of the film's cinematic soundtrack in the intimacy of one's home, a complete Atmos-approved home theatre system would be required. Or perhaps, given the down-mixing capabilities of Dolby's latest sound platform (see Chapter 1), a 7.1 surround home theatre could provide a 'close-enough' 360-degrees experience. Evidently, those devices are currently the exception and not the rule for the average Netflix subscriber.

It is important to highlight, however, that nowadays Netflix plays a fundamental role in the film distribution industry. In the second quarter of 2018, the year in which *Roma* was released, Netflix reported 130.1 million streaming subscribers worldwide, among which 57.38 million were in the United States (Consumer Technology Association, 2018, p. 114). In that year, Netflix reported more U.S. subscribers than cable TV (ibid), becoming the most important video-on-demand service in the country with the largest number of signed up households (Statista, 2021a). This tendency is also seen in other markets. In the United Kingdom, for instance, Netflix has reached over 15 million subscribers (Statista, 2021b), whereas, as a comparison, Sky's satellite TV was only able to reach 10 million in more than two decades (Sweney, 2018). Worldwide, the number of paid Netflix subscribers has grown from 21.5 million in Q3 2011 (Statista, 2021a) to over 203.67 million in Q4 2020 (Statista, 2021c), including a record of 36.6 million accounts added since the end of 2019, specially during the Covid-19 pandemic (Richter, 2021). On the other side of the spectrum, movie theatre attendance has decreased in the last two decades, with 2017 showing the lowest number of tickets sold in the domestic movie theatrical market from 2000 until 2019 (The Numbers, 2021). To all this, it is important to add the effects that the Covid-19 pandemic has had on the cinema industry, which obligated movie theatres to close their doors while abiding by new norms of social distancing. While 2019 reported over 1.2k million tickets sold, 2020 reported only 224 million (The Numbers, 2021).

Therefore, even though it is true that the movie theatre offers the best possible cinematic experience, something that is undeniable for an immersive film like *Roma*, it is also true that the possibility of watching a film on Netflix expands the potential number of viewers. After *Roma*'s three-weeks theatrical opening, more than 203 million households have had the

opportunity to view the film all over the world[30], and a percentage of that population may have actually experienced it using some sort of surround sound system. According to the Consumer Technology Association, the surround sound system household penetration in the United States by 2016 reached 41% (2016, p. 25). Moreover, object-based audio has pushed the development of soundbars, which in the USA are becoming "the most popular audio solution for home theatres, with factory-level sales rising 16.1 percent in 2017" (Consumer Technology Association, 2018, p. 21). Additionally, it is worth noting the recent introduction of immersive sound technology for headphones, which offers three-dimensional virtualisation capabilities that in theory can provide an 'all-around' listening experience.

In its alliance with Dolby Atmos, Netflix has produced immersive films like *Okja* (2017), *Extraction* (2020), and *Roma.* The streaming giant has clearly shown its strategic commitment with the improved home-cinema experience, and the offering of Atmos-enabled films, together with the post-pandemic lifestyle, may be an important pull for audiences to upgrade their movie-watching infrastructure at home.

Conclusion

With *Roma*, Alfonso Cuarón has once again shown his fascination for the immersive cinematic experience. Following the realistic approach of his science-fiction thriller *Gravity*, *Roma* signifies the consolidation of an audio-visual style that considers the three-dimensionality of sound as one of its fundamental pillars. The director's aesthetic choices aim to honour the continuum of time and space, something that is aesthetically evidenced on his extensive utilisation of long takes, deep-focus cinematography, wide-shots, and camera movement; all of which are realistically embellished by his three-dimensional sound design approach.

With all these characteristics, Cuarón's audio-visual style, which this thesis refers to as immersive continuity, stands out as an alternative filmmaking approach that deviates from the conventions of mainstream cinema, offering a feasible opportunity for exploiting the full potential of immersive sound. As the standardisation of object-based audio in the form of

[30] This calculation was made by the author considering the number of Netflix subscribers at Q4 2020 based on the information found in Statista (2021c)

Dolby Atmos continues to grow, the adoption of an audio-visual method that facilitates embracing sound three-dimensionality becomes more relevant. Certainly, intensified continuity is still the dominant visual style of contemporary mass-audience films, and it is likely to remain like this due to tradition and the long-lasting hegemony of the visuals over the aural; nonetheless, immersive continuity may become a strong alternative as Dolby Atmos and other immersive sonic platforms gain more momentum, challenging filmmakers to find new ways to take advantage of such powerful technologies.

The possibilities that Dolby's latest sonic innovation has to offer invite filmmakers to consider sound earlier in the creative process. In *Roma*, this was evidenced in a script written with sound in mind, in a set of workflows that connected the sound and image departments as two parallel sources of creation, and in a series of well-connected sound design strategies that always followed an object-based approach. Although Cuarón must be acknowledged as the person who visualised the story, both visually and sonically, having a supervising sound editor in charge of sound throughout the entire process was essential for achieving the final results.

In despite of all the benefits that distributing a film on Netflix can have, it is clear that a film with the characteristics of *Roma* deserves more than a simple home cinema experience. The film was made thinking on the effects that only the immersive darkness of the cinema can produce; 3-D sound was a key element throughout the film's creative process, and all the time and money spent in the production of its soundtrack is a proof of its importance. *Roma*'s three-dimensional sound design is in fact what really differentiates this film from the enormous number of movies that are released each year. Cuarón's immersive continuity is by excellence a theatrical strategy, hence *Roma* should have been intended for a major theatrical exhibition and not only as another TV movie.

In any case, the consolidation of Netflix and other streaming platforms as important agencies in the market of film consumption is a factor that cannot be underestimated. The growing adoption of sound technology for the home and the latest innovations in individualised listening devices indicate that an improved movie-watching experience is now possible in the comfort of one's home. Yet, the immersive capabilities of a movie theatre are today irreplaceable by the typical home-theatre equipment. In order to combat against the increasing number of streaming platforms, films that are intended for theatrical exhibition should consider embracing a more immersive approach, one that can offer audiences an 'out of body' experience. Current cinematic sound technology has the potential to enhance our sensation of

presence in the space of the movie, meaning that content creators have now the tools at their disposal, and it is a matter of creative choices what may or may not exploit the potential of the immersive cinematic apparatus.

As Bazin once argued, "[e]very new development added to the cinema must, paradoxically, take it nearer and nearer to its origins" (1967, p. 21), and in the hands of filmmakers like Alfonso Cuarón, they are taking us closer to the mythical experience of total cinema. In the following chapter, a number of films (most of them mixed using Dolby Atmos) will be observed in order to evaluate how the main traits of the immersive continuity style are being used by other filmmakers, and in what ways these may differ from Cuarón's initial approach. Ultimately, this academic contribution aspires to bring to light other emerging approaches, and beyond, it aims to support the development of a new immersive cinema of the future.

CHAPTER FIVE

The 'Gravitational Pull' of *Gravity*

After experimenting with unconventional sound spatialisation techniques in *Children of Men*, Alfonso Cuarón consolidated in his two most recent films a sound design approach that offers a three-dimensional cinematic experience. At the 86th Academy Awards, *Gravity* won seven Oscars (including Best Director, Best Cinematography, Best Film Editing, Best Sound Editing, Best Sound Mixing, and Best Original Score); while *Roma* obtained three awards (Best Director, Best Cinematography and Best International Feature Film) out of ten nominations at the 91st Oscars Ceremony. All these accolades confirm Hollywood's appreciation for Cuarón's audio-visual proposal. But could such an unconventional approach become the future of cinema? In Mera's analysis, movement and directionality in music and sound are likely to be the next step in the evolution of film sound, both in terms of technology and aesthetics (2016, p. 108). This does not mean that all movies will use immersive sound in such an extensive way. In fact, it is more likely that such an unusual treatment will continue to be the exception. Yet, Cuarón's 'gravitational pull' has proven to be strong, and his audio-visual traits are now found in a diversity of films and genres.

With the aesthetic innovations presented in his most recent productions, Alfonso Cuarón has proved himself to be a pioneer in the exploitation of immersive sound technology. Yet, he is not alone. Subsequent mainstream productions have also experimented with highly spatialised approaches to sound design, suggesting that the long-lasting screen centric convention is starting to change. Perhaps the best examples are to be found in *Birdman* (2014) and *The Revenant* (2015), both films directed by Mexican filmmaker Alejandro González Iñárritu in collaboration with Cuarón's long-time collaborator, cinematographer Emmanuel Lubezki. In these films, the film style that this thesis has termed immersive continuity is thoroughly deployed, with fluid long takes and 3-D sound encountered in almost every sequence.

Although three-dimensional sound design is still highly uncommon, it has gradually become more recurrent, not only in the mainstream theatrical realm, but also in streaming-based content. Proof of this are Dolby Atmos releases such as Darren Aronofsky's *Mother!*

(2017); Edgar Wright's *Baby Driver* (2017); Bong Joon-ho's *Okja* (2017), David Leitch's *Atomic Blonde* (2017); Steve McQueen's *Widows* (2018); Quentin Tarantino's *Once Upon a Time in Hollywood* (2019); Bong Joon-ho's *Parasite* (2019); Leigh Whannell's *The Invisible Man* (2020); Sam Hargrave's *Extraction* (2020); and many more, all of which contain moments where sound elements are distributed in accordance with the position of the sources in the 3-D space. In addition to this, it is important to mention that such an immersive treatment of sound has also been found in moments of Atmos-enabled TV series such as *Games of Thrones*, *Black Summer*, *The Haunting of Hill House, Dark*, among others.

In this chapter a number of films will be examined aiming to determine the extent to which Cuarón has cast the 'gravitational pull' that Mera (2016) predicted. Let us begin with a brief case study of Alejandro González Iñárritu, a highly regarded filmmaker who has lately stood out in terms of his complete embracement of immersive continuity, a film style that is drawing us near to the mythical experience of the Bazinian total cinema.

Following Cuarón's Steps: Alejandro González Iñárritu and Immersive Continuity

Iñárritu is perhaps the filmmaker that has more clearly adopted the immersive film style pioneered by Cuarón. As a matter of fact, Iñárritu acknowledges Cuarón not only as a close friend, but as a personal adviser (Chagollan, 2019). Throughout the whole creative process, they influence each other, receiving advise "from the conception through the script and then through the editing" (Iñárritu in Chagollan, 2019). Although until *Biutiful* (2010), Iñárritu was not considered a 'long take director', his two most recent (and successful) productions stand out for an extensive deployment of long continuous shots. But in contrast to directors like Sam Mendes, who has also begun to use long takes in films like *Spectre* (2015) and *1917* (2019), Iñárritu makes use of this visual strategy with the important addition of 3-D sound. In other words, Iñárritu has not only entered the list of 'long takers', but has also joined Cuarón as an ambassador of the immersive continuity style.

Although such an audio-visual style is still passing through an experimental phase, its application has been well received by the industry and its institutions. In addition to all the prizes that Cuarón received for his two most recent films, Iñárritu's *Birdman* and *The Revenant*

received nominations for Best Sound Editing, Best Sound Mixing and Best Picture at the Academy Awards of 2015 and 2016 respectively. The two movies won in the categories of Best Directing and Best Cinematography, and *Birdman* was selected as the best picture of the year. Emmanuel 'Chivo' Lubezki, director of photography of *Gravity* and Iñárritu's latest films, became the first cinematographer to win three consecutive academy awards; and in part, it was thanks to his work on these three movies that he earned the reputation of being a "master of the long take" in the digital age (Fazio, 2016). With such a large number of accolades, the Academy has ratified its appraisement for immersive continuity, a film style that despite being experimental has conquered the summit of Hollywood's cinema.

The visual experiment showcased in *Birdman* is in fact the most extreme example of what this thesis considers to be the 'look' that facilitates the deployment of three-dimensional sound design. The film tells the story of Riggan Thomson (Michael Keaton), who after being a famous Hollywood superhero celebrity in the 90s, tries to satisfy his obsession for recognition by writing, directing and starring in a Broadway play. Apart from the dream sequences shown at the beginning and before the hospital epilogue, this movie plays as one continuous, real-time shot. *The Revenant* is far less orthodox. In this epic Western, Iñárritu follows the journey of Hugh Glass (Leonardo DiCaprio), a frontiersman who is attacked by a bear and seeks revenge after being abandoned by his hunting team. According to the author's own measurements, the film's average shot length (ASL) is of about fourteen seconds among a total of 599 shots. There are 27 shots that last more than one minute, the longest being the long take that covers almost the whole bear sequence, lasting 4 minutes and 58 seconds. Although *The Revenant* is cut much faster than films like *Gravity*, *Roma* or *Birdman*, it is still a long way from the norm of Hollywood's contemporary editing practices (see Chapter 3). Certainly, the ASL found in all these movies goes markedly against conventional editing, and thus cannot be simply classified by what Bordwell (2002) calls intensified continuity, the dominant editing style of the vast majority of contemporary mass-audience films.

Furthermore, Bordwell observes that films are normally shot with a broad combination of lenses. As he notes, "Francis Ford Coppola, Brian De Palma, and Steven Spielberg freely mixed long-focus and wide-angle lenses within a single film" (2002, p. 18). Joining Cuarón's latest films, *Birdman* and *The Revenant* are two other exceptions to such a stylistic trait. *Birdman* for instance, was primarily shot with wide-angle lenses (Picone, 2015), and one of the characteristics of these types of lenses is their long depth-of-field, a characteristic that Bazin

(1967) priced for its closeness to reality. As Bordwell (2015) notes in his personal blog, "[t]he wide-angle lenses used on *Birdman*, only 14mm and 18mm, don't always create wire-sharp focus in depth, but they provide enough visibility to create depth effects". In a similar way, *The Revenant* was shot with wide-lenses, ranging from 12mm to 21mm (Riley, 2015), which gives the film a highly realistic appearance. Scott Sakamoto, camera operator of the film stresses:

> Shooting with the wide lenses shows an interesting perspective and makes you feel you're part of the character. You have their panoramic POV. We're used to seeing a lot of close-ups and extreme close-ups, but in reality the view a person has is this wide POV. (Sakamoto, n.d.)

Moreover, Bordwell suggests that close framing in dialogue scenes has been seen as the norm since the 1960s, meaning that two-shots were replaced by closer shots that frame one character at a time. Bordwell notes that "[a]fter the 1960s, such two-shots were often replaced by 'singles': medium shots or close-ups showing only one player" (2002, p. 20). As analysed in Chapters 3 and 4, this is not the case for *Children of Men*, nor for *Gravity* or *Roma*; and although close framings are vastly found on *Birdman* and *The Revenant*, especially by means of camera push-ins, fuller frames are a more recurrent characteristic in these films' visual style. *Birdman* normally uses two-shots for moments of dialogue between characters. We often see Riggan having a conversation with another character, both being captured from a medium shot. *The Revenant*, on the other hand, is a film that uses a lot of wide-shots, which offer space for more than two characters to appear on the screen. It is common, for example, to see the group of fur trappers talking to one another without a close reference to their facial gestures.

Although most of the characteristics suggested by Bordwell as part of the intensified continuity style fail to describe the 'look' of Cuarón's and Iñárritu's latest films, this author highlights camera movement as a resource commonly found when long takes are employed. Indeed, the moving camera is an essential component of the visual aesthetic of the films in question. Both Cuarón and Iñárritu make an extensive use of long takes, and as argued by Bordwell, "[w]hen we do find longer takes and fuller framings, the camera is usually in motion" (2002, p. 20). The difference, of course, lies in the fact that the long take is the exception and not the rule in the context of intensified continuity, while it is a constant resource in the work of Cuarón and Iñárritu in collaboration with Lubezki. In other words, while the long take is much of the time absent in the films that conform to intensified continuity, such a visual resource is the essence of the immersive continuity style.

But apart from the visual characteristics described above, what differentiates these films from the majority of contemporary movies is their three-dimensional utilisation of sound. To a great extent, *Birdman* and *The Revenant* present the unconventional sound/image relationship initially found in *Gravity*, which helps the audience member to go through a more vivid experience. All these films' fictional settings are in a way transformed into 'lived spaces' with the intention of creating the sensation of presence in the narrative world. By means of a highly realistic representation of the way we hear in real life, these films thus succeed in giving lifelike attributes to the three-dimensional diegesis, a virtual space that as Doane notes, differs from the visible space of the screen and the acoustical space of the theatre in the sense that it "has no physical limits" (1980, p. 39); it is an imaginary space that the spectator is meant to inhabit, and through it get transported into the narrative world.

Let us analyse, for instance, the opening sequence of *Birdman*. After the title sequence, in which the crescendo of a noisy night ambience is heard together with the loudness of Antonio Sánchez's jazz drumming, the soundtrack transports us into Riggan's dressing room by abruptly cutting to the 'silence' in which we find him levitating. Just as in *Gravity*, the film begins with an aggressive contrast of loudness and silence that instantly places the spectator inside the filmic space. Moreover, the beginning of the film not only activates the process of transportation into the narrative world, it actually induces the spectator to identify with the protagonist by giving us access to his own subjectivity, and to the particular way he hears the world.

Once the sound of the opening sequence is suddenly terminated, all we can hear is the rhythmic sound of an offscreen wall clock over a distant city ambience that shape the room tone of the location. Although we do not see the clock, we can assume that it is hanging on the right wall just behind Riggan's position, as the sound emanates from the right side of the auditorium. While the camera pushes-in towards Riggan, we hear the loud and non-reverberant voice of the Birdman creature, coming from all directions, which suggest its somehow ethereal nature, as it only exists in Riggan's mind. Hence, although we do not share Riggan's subjective point-of-view nor his point-of-audition, we can hear the inner voice of his alter ego just the way he hears it inside his head, suggesting that we are some sort of invisible spectators that inhabit the protagonist's altered reality.

Riggan's state of meditation gets interrupted after the characteristic ringtone of a Skype video call is heard mainly on the rear left corner, a fairly simple spatialisation technique that

enhances the process of transportation by locating us within the fictional setting where the events take place. As Riggan gets up and walks towards the computer, the camera pans to the left, which consequently changes our audio-visual perspective in a coherent way. Slowly the moving camera frames a medium close-up of the computer, and the ringtone gradually passes from being concentrated mainly on the rear left, to be on the front centre speaker. After Riggan's daughter Sam (Emma Stone) hangs up, we are able to hear the sound of the wall clock behind us—not on the right side where it was before—keeping a coherent spatial match with the current visual perspective. As Riggan closes the laptop, we see his face on the mirror, in which we also see a reflection of the clock. The moving camera then pans to the right and frames a close-up of Riggan's face, meanwhile the filtered voice of a woman sounding through an intercom system is heard emanating from the left front: "Riggan, everyone is set for one-four. They're ready for you", she warns. The camera then frames a lateral view of Riggan's face and a red light flashing on the background. It is rehearsal time, and this time the intercom voice is heard emanating from the front centre speaker in concordance with our visual perspective.

As in Cuarón's latest films, the long take allowed Iñárritu and his sound team to pan not only sound effects, but also dialogue all throughout the exhibition space with the intention of recreating the way we experience reality. For example, after one of their conversations at the rooftop of the theatre, we follow Sam as she seduces Mike (Edward Norton) on a truth-or-dare game. Through a coherent three-dimensional synchronisation of image and sound, we are placed in the middle of the couple as they go down the stairs. We see and hear Sam in front of us, while we hear Mike (who is offscreen) closing the back door and walking behind us. His reverberant voice and footsteps are heard mainly on the left surround speaker, suggesting his position within the diegesis. Even though we do not see him, we can feel his presence right behind us, as his voice and movements emanate from the rear speakers. Mike's 'voice-off', which Doane defines as "the voice of a character who is not visible within the frame" (1980, p. 37), is in this case located in the 'exact' position where the character is located, represented in the theatre by a voice that emanates from a speaker (or group of speakers) in the back of the acoustical space. This relatively simple spatialisation technique enhances the process of transportation by locating us within a 'lived space', with sound elements being distributed in accordance to the position of their sources in the three-dimensional diegesis.

Such a realistic sound design approach is deployed throughout the movie, which differentiates the long takes used by Iñárritu (and also Cuarón) with the long takes encountered in earlier times. For example, in the famous 'Copacabana' sequence of Martin Scorsese's *Good Fellas* (1990), a Steadicam follows Henry (Ray Liotta) and his girlfriend Karen (Lorraine Bracco) as they walk towards the back corridors that lead to the club; and as this happens, the stereo soundtrack keeps all diegetic sounds in the front and all dialogue in the centre of the screen. Similarly, in the opening sequence of Brian de Palma's *Snake Eyes* (1998), dialogue is kept in the front-centre speaker, and the multichannel soundtrack does not produce the sensation of movement that is found on *Birdman* every time the camera moves through the corridors of St. James Theatre, which is achieved by attaching all kind of noises to their correspondent source of sound (e.g. the footsteps of a person who passes by from front to rear).

It is important to notice, however, that *Birdman* was not mixed nor released on Dolby Atmos; its soundtrack was produced using the traditional 5.1-based configuration, meaning that the representation of the diegetic space in this film was restrained by a limited number of audio channels. This is undoubtedly a major limitation in the sense that sounds cannot flow throughout each individual speaker, but they have to abruptly jump across large zones of speakers. Moreover, traditional channel-based setups do not include overhead speakers, which means that in *Birdman* the three-dimensionality of the sonic space had to be represented only within the 360-degrees horizontal plane.

This limitation was overcome on *The Revenant*, which included an object-based Dolby Atmos soundtrack for its theatrical release. One thing that is remarkable about the sound design of this film is the highly realistic way in which nature is portrayed. After the introduction, we are immediately transported into the location by means of an enveloping atmosphere full of '*cacayanga*', which as Randy Thom, co-supervising sound designer/editor and re-recording mixer of the film, explains, "means a kind of noise that is complex in a useful way. So, when a scene needed a certain kind of energy, [Iñárritu] would say 'we need more cacayanga in this scene', and I knew exactly what he meant" (Thom, personal interview 2019). Martín Hernández, co-supervising sound editor and sound designer of the film, elaborates:

> I think that when Alejandro calls for a little cacayanga, a little uncorrelated noise, I think that he is trying to make it more like the real world, when unpredictable things happen and it is complex, when everything is not explained or linked to what you see on screen. (Hernández in Woodhall, 2016)

These passive offscreen noises, as Chion refers to those sounds that create "an atmosphere that envelopes and stabilizes the image, without in any way inspiring us to look elsewhere or to anticipate seeing its source" (1994 [1990], p. 85), are indeed an important component of the soundtrack of *The Revenant*. However, it is the extensive utilisation of active offscreen sounds—defined by Chion as acousmatic sounds "that raise questions […] whose answer lies offscreen" (Chion, 1994, p. 85)—and their dynamic spatial characteristics, what really helps to transport us into the narrative world.

The battle sequence at the beginning of movie is an interesting example. After hunting an elk, Jim (Will Poulter) asks a group of trappers to help him hauling the meat back. Suddenly, we all—the characters and the audience—hear a distant cry for help emanating from the right side of the auditorium. The camera rapidly pans towards that direction and we see a naked man approaching from a distance. The man falls to the ground and through a wide-shot we are able to see that he has got an arrow nailed on his back. The voice of an offscreen frightened man is closely heard on our right: "Is that Coulter?", he asks. Suddenly, one of the trappers gets struck by an arrow and falls death over the wood fire. The screams, voices and whispers of the frightened men are heard all around us, which provides a clear three-dimensional representation of the despair that is felt while being in the middle of the ambush.

From a distance we see how a group of Arikaras—a Native Amerindian tribe—kill another white man and throw him down the hill. "They got Thomas", one of the fur trappers whispers, and the camera freely moves among the startled group, while the immersive soundtrack displays a variety of subtle offscreen noises. The wind, the cracking of the trees, the fauna of the location, distant shouts, and a series of silent voices are heard emanating from different directions in accordance to the position of the sources in the diegesis. This is especially noticeable when the camera shows some arrows passing through the treetops, and as the image rotates by 360-degrees, we hear the close whispering of an offscreen man moving around, matching our circling visual perspective. "I can't see'em… I can't see'em", he says, and as part of the suspense, we cannot see them either. After a few more arrows strike some trees near their position, one of the trappers decides to run away and we see him being hit by an arrow. Completely scared, Captain Andrew Henry (Domhnall Gleeson) and his men get covered behind the tree trunks, and as we see Henry and one of his men talking on screen, we hear the desperate screams of the injured man emanating from the left.

Once Glass gets into the battlefield, 3-D sound immerses us into a breath-taking combat between the American trappers and the Amerindian tribe. We experience, for instance, how a group of Arikaras arrive on their horses, and we feel them approaching from behind. As we see them appearing on the screen, we hear them passing rapidly through our position, from rear to front, not only as an effect of the coherent localization of sounds in space, but by means of the powerful rumbling produced by the LFE channel, which simulates the heavy footsteps of the cavalry. Similarly, the first Arikara warrior that attacks Glass is heard initially offscreen, approaching from behind just before he visually appears; and the Arikara chief is heard approaching from the right seconds before he decapitates one of the white men onscreen. Such an *"in-the-wings effect"*, as Chion (1994) once referred to this kind of spatialisation technique, was experimented and quickly dropped during the Dolby Stereo era (ibid); however, it is constantly employed throughout this and other recent films, which reopens the debate and challenges practitioners and academics to establish a framework under which such a treatment of sound could experience a broader renaissance. Certainly, as Chion stresses:

> [T]he in-the-wings effect created a nagging problem by violating the conventions of continuity editing and making sound matching problematic. But maybe it could have gained more permanent admittance into film practice had it been systematized along with some partial adjustments in editing conventions […] So perhaps it was a mistake to have given up so quickly. (1994, p. 84)

Such practical editing adjustments were in fact applied by Iñárritu in the form of longer shots, which facilitated the application of three-dimensional spatialisation techniques, something that is difficult to find in action sequences, as they are normally constructed out of many different shots. In relation to the long take and its benefits for the soundtrack, Randy Thom comments:

> I think the longer the take is, the more opportunity you have to do interesting things with sound. And probably the fewer cuts there are within a sequence, the more potential you have to do something with sound also, because you don't have to adjust the sound to conform with the cuts or the perspective changes. But long takes are even less common, I would say, in filmmaking now than they were five or ten years ago. The temptation is usually to make scenes and shots as short as possible because there is this idea that, especially the younger audiences' attention paying is shorter than it used to be. So, I hope that these kinds of movies and these kinds of sequences that involve long takes will persist. Obviously, every movie doesn't need to be that way, but I hope those filmmakers like Cuarón and Iñárritu, who seem to specialise in those, will continue doing them. (Thom, personal interview, 2019)

Indeed, the long take opens a big door of opportunities for sound design, something that Iñárritu supported by means of a minimalist utilisation of music. As quoted in Chapter 3, Kulezic-Wilson (2008) notes that in mainstream cinema it is common to score every action, every transition, every emotional or suspense scene. Yet, she remarks that a growing number of composers have been trying to resist this trend and convince directors to change the way music is employed, knowing that this will benefit the film by allowing the creation of more effective scores. In *Birdman*, the utilization of non-diegetic music is minimal, and in fact, music could be considered 'quasi-diegetic', as in certain moments of the film the percussive incidental music becomes part of the diegetic world. There are two specific moments when we actually see the drummer performing the score. The first time we see him is when Riggan and Mike walk towards the bar and the former throws some coins to a street musician, which happens to be the drummer (00:32:37). Right before this scene, the same rock-drumming groove is heard as non-diegetic music, but after Riggan opens the theatre's exit door, the music gradually becomes part of the diegesis, adopting its corresponding spatial signature (Altman, 1992c) and position within Iñárritu's three-dimensional representation of Manhattan. The second time we see the drummer is when Riggan walks towards the stage to deliver his final scene (01:40:14). Again, the drums gradually become diegetic as we approach the musician, and then return to the non-diegetic dimension as we walk away with Riggan. These two appearances suggest that the drum beats that we hear throughout the film are in some way part of that magic world that Iñárritu intends to portray, in which many of the things happen inside the protagonist's mind.

In addition to Antonio Sánchez's drumming score, there are a few moments that use non-diegetic music to convey certain emotions, like Riggan's depressed state of mind after his encounter with the theatre critic (01:25:15). The non-diegetic strings rise out of a diegetic piano that is heard in the interior of the bar, which connects the diegetic music with its non-diegetic counterpart. Later on, the hallucination scene that follows (01:28:07) presents Iñárritu the opportunity to employ a more elaborated orchestral score without breaking the sound concept of the film. As Rocha notes, "[t]his scene offers the audience the position of eye witnessing a hallucination whose visual and auditory content is fashioned according to the audiovisual imagery of superhero blockbusters" (2017: 145). Hence, the orchestrated incidental music that we hear throughout this particular sequence is actually part of the diegesis, but in the form of an inner world that we access through the former Hollywood star' disturbed mind. Thus, even though the three-dimensional space does not rule the physicality of music in this film, there are clever exceptions that integrate the diegetic and non-diegetic terrains to increase the sensation

of presence and the film's impression of reality, which as Bilandzic and Busselle's (2011) empirical findings suggest, can facilitate the narrative experience of transportation.

In a similar way, *The Revenant* presents a minimalist orchestral score composed by Ryuichi Sakamoto and Alva Noto (Carsten Nicolai) in collaboration with Bryce Dessner, whose main theme is used as a *leitmotif* throughout Glass's journey. Besides its emotional strings, the score is noticeable for the utilisation of long reverberation and some sort of atmospheric pads, which create the illusion of being inside a windy and cold soundscape that resembles nature. Based on his analysis of the soundtrack of *Gravity*, Mera had argued that one of the defining traits of what he calls 3-D sound

> is that music is emancipated from a fixed sound stage representation […] At the same time, music is not tied to environmental 'reality' in quite the same way as sound design, so music is, in some ways, able to move more freely. All in all, this means that the borders between what have traditionally been considered sound design and music are collapsing which, [he argues], is also primarily at the service of spatial presence (2016, p. 103).

Although Mera is describing the soundtrack of *Gravity*, his analysis resonates with Iñárritu's treatment of the score. Randy Thom argues that "there are some great places where you are hearing both sound design and music and very often you can't even tell the difference between the two" (Interview by Sound Works Collection). Martín Hernández did some processing in order to blend production sounds with the music, because he "was thinking about music more like part of the sound design" (Interview by Sound Works Collection). In respect to the position of the music in the spatial layout, Jon Taylor, *The Revenant*'s re-recording mixer explains:

> The amount of dynamics is not just sonically—but also spatially. A lot of times the music is just left and right, then sometimes it's kind of center heavy, and then other times we use all seven tracks, then sometimes it's just in the back. The film is about nature so music never just takes over, it always has spaces so that the nature can come through and co-exist [in the soundtrack]. (Taylor in Woodhall 2016)

For instance, during the breath-taking battle sequence previously analysed, the only non-diegetic sounds they use are a few drum beats that, as Taylor highlights, move across the surround setup. Hence, it is not the music that induces the sense of tension in the audience, a common device found in mass-audience films, but rather the sounds from the filmic space itself, which work together with a minimalist and three-dimensional musical score to manufacture the illusion of a world unfolding, not only in front of us, but also around us.

Immersive Continuity as a Sequence Shot Strategy

Cuarón and Iñárritu in collaboration with Lubezki have showcased the immersive potential of the long take, which combined with 3-D sound allowed these Mexican filmmakers to create films that can more easily produce an 'out of body' cinematic experience. In contemporary Hollywood, these Mexican filmmakers can be acknowledged as the pioneers of such an immersive utilisation of the sonic space, but the cinema's intention to provide a lifelike synchronisation of image and sound has a long history, which goes back to the classical Hollywood era. One of the most remarkable examples is the work of Orson Welles, and particularly the opening sequence of *Touch of Evil* (1958). For this sequence, the all-time 'master' of the long take visualised a fluid sonic effect to accompany the continuous movement of the camera. Although Universal Pictures drastically changed Welles' original idea for the 1958 original release, the restored version of the opening sequence (re-mixed by Walter Murch for the 1998 release) presents a continuous audio-visual correlation that invites us to move within a fictional town on the U.S.-Mexico border, which in a certain way resembles the tracking shots that are used in *Roma* as the characters transit within the streets of Mexico City.

Murch, acknowledged by Lastra as the father of "modern, immersive sound design" (2008, p. 123), restored the film's soundtrack based on the instructions included in a 58-page memo that Welles had written in 1957 after being fired by Universal, in which he clearly explained his idea of creating a realistic sonic environment that would follow the continuous movement of the camera. For the opening sequence, Murch explains that Welles wanted a soundtrack composed of fragments of source music (e.g. brothels, cantinas, car radios) that change and blend into each other as the camera moves slowly through the streets of the film's fictional town (Ondaatje, 2002, p.186). In other words, the idea was for the soundtrack to become "the aural equivalent of the camerawork" (ibid). In the memo, Welles stressed:

> As the camera moves through the streets of the Mexican border town, the plan was to feature a succession of different and contrasting Latin American musical numbers—the effect, that is, of our passing one cabaret orchestra after another. In honky-tonk districts on the border, loudspeakers are over the entrance of every joint, large or small, each blasting out its own tune by way of a 'come-on' or 'pitch' for the tourists. The fact that the streets are invariably loud with this music was planned as a basic device throughout the entire picture (Ondaatje, 2002, p.185)

Certainly, as Grajeda argues, statements like these prove that Welles pursued a cinematic aesthetic that included aural realism as one of its core components (2008, p. 214), a sonic

approach in which 'realistic' source music and ambient sounds would situate the spectator within the diegetic space of the film (ibid).

In fact, such a realistic audio-visual approach was idealised since the dawn of sound in cinema, when sound engineer Joseph Maxfield and others defended the idea of achieving a perfect match between image and sound (Altman, 1992a; Lastra, 2000). As Belton (1985) observes, early sound practitioners sought to produce a soundtrack that could preserve the spatial characteristics of the original scene. "Carefully positioning the microphone to blend direct and reflected sound, soundmen record not only the informational 'content' of an actor's speech or sound effects but also its spatial presence" (Belton, 1985, p. 68). Indeed, since the early years of film sound, assuring sound localization by means of a realistic spatial correlation between image and sound has been one of the big concerns for sound practitioners; and as Doane (1980) argues, sound technology has historically been innovated in response to the need of more realistic reproduction.

Nowadays, such a realistic audio-visual aesthetic seems to be experiencing a modern and more vivid renaissance in the work of Cuarón and others, with a number of films that contain moments in which diegetic music, voices, ambiences and sound effects move in a lifelike manner, following the position of the camera and the action that is happening within the three-dimensional diegesis. From science fiction, to dramas and action films, the aesthetics of immersive continuity are now more commonly found in film and media. While long takes are not a new thing, its conspiration with 3-D sound to create a more sophisticated impression of reality is a very recent implementation. Unlike the sequence shots deployed in previous eras—such as the aforementioned 'Copacabana' Steadicam-shot in Scorsese's *Good Fellas*, the opening sequences in Robert Altman's *The Player* (1992) and De Palma's *Snake Eyes*, all the long takes in Paul Thomas Anderson's *Boogie Nights* (1996), or the Dunkirk sequence in Joe Wright's *Atonement* (2007)—long shots in present-day cinema are commonly complemented by the enveloping characteristics of 3-D sound.

In fact, opening sequences continue to be a fertile field for the application of immersive sequence shots. An interesting example comes from the opening of Damien Chazelle's musical *La La Land*. This sequence shot begins with a tracking camera moving along a traffic-filled highway in Los Angeles, California. As the camera passes through a number of vehicles, the noisy horns are merged with the radio sound or music that comes from each car, which dynamically changes in accordance with what each sound system is reproducing. Once the first

musical number ends, the camera approaches Sebastian's (Ryan Gosling) car, and we gradually hear what he is hearing inside his vehicle. The camera then moves to the right, and the sound coming out of Sebastian's car is gradually lost as we begin to hear Mia's (Emma Stone) voice on our right. Throughout this sequence, Chazelle and his sound team were thus able to achieve what Welles intended for the opening sequence of *Touch of Evil*, materialising the crucial idea of sound space inherited from the very first film sound professionals.

Similarly, in Edgar Wright's *Baby Driver*—which is also in a way a musical since the story runs in synchrony with what the protagonist hears on his iPod—immersive continuity is employed during the credits sequence at the beginning of the movie, with a moving camera that follows Baby (Ansel Elgort) as he walks across the streets of downtown Atlanta. Throughout this sequence, not only sound effects, but also the voices, movements, and footsteps of background characters are positioned in accordance to where they are, thereby creating an 'inhabitable' audio-visual representation of the city. The abovementioned sequence of *Baby Driver* looks similar to, for example, the opening of Anderson's *Boogie Nights* in terms of camera movement and composition. Yet, by comparing the highly spatialised soundtrack deployed in *Baby Driver* with the screen-centric approach found in *Boogie Nights*, it is easy to distinguish a radical evolution in sound design aesthetics towards the achievement of spatial presence.

Theories of camera movement support the idea that such a visual technique helps to place the spectator within the world of the movie (i.e. Sobchack, 1992; Danto, 2006; Carroll, 2008). In an early essay, David Bordwell argues that camera movement supplies "a compelling experience of moving through space" (as cited in Morgan, 2016, p. 229); or as Morgan puts it, "camera movement makes us feel as though we are ourselves moving through the space of the film's world" (2016, p. 229). Morgan appropriately points out, however, that "sight alone is not enough to account for how we perceive the world" (ibid, p. 223); in order to represent such a phenomenon, this thesis argues, not even stereoscopic imagery would be enough. Beyond the visuals, the soundtrack ought to parallelly represent the way we hear, an approach that is distinguished in films like *Gravity*, *Roma*, *The Revenant*, *Baby Driver*, and others, with dynamic variations in sound perspective and localisation in relation to what we see on the screen.

As noted in Chapter 2, Richard Allen (1995) asserts that the illusion of presence in the cinema is alimented by the camera's ability to move and record movement, a cinematic

attribute that is enhanced by the action of diegetic sound. Today, sound design can create the sensation of movement even when the picture remains still. A good example is found in Jeremy Rush's *Wheelman* (2017), which begins with a 4-minute shot that is captured from the interior of the car that Wheelman (Frank Grillo) will use for his next job. Throughout the sequence, we hear everything from a single and three-dimensional perspective, namely the camera's immersive point-of-audition. When Wheelman surrounds the car, for instance, we hear his footsteps going around us, and as he checks the rear suspension, we hear him pushing the trunk behind us, just before entering the car through our left. This very simple panning strategy helps to create the illusion of being inside the car, it transforms the film into some sort of driving simulator and thus grabs our attention since the moment the film begins. The absence of music throughout this Atmos-mixed sequence allows all those subtle sounds to be clearly heard.

Another interesting example is found in the opening sequence of Leigh Whannell's science fiction horror film *The Invisible Man* (2020), featuring a 1-minute static shot of a cliff seen from the ocean. Throughout the shot, the sound of waves hitting the rocks is heard all around us, an approach that the director deliberately used to create the illusion of presence in the space of the movie. Referencing this particular shot, Will Files, supervising sound editor and re-recording mixer of the film, comments:

> One of my […] favourite scenes from *The Invisible Man* was actually the very first shot in the film. A locked off shot, looking at a cliff side and a big rock in the ocean. And the camera is sort of out in the ocean. It was this interesting trick the director decided he wanted to do it that way because he wanted the audience to understand that this movie would be about point-of-view and that it wasn't just sort of a random shot, it was designed to make you feel like you were there. We were able to build the sounds in such a way, where they literally went through the entire theatre and up to the front and then back through the entire theatre. It's so immersive, you feel like you're out there in the ocean and you feel the sense of the power of these waves moving through the room. (Files in Dolby, 2020b)

This shot is by all means a reminiscence of the beach scene near the end of *Roma*, a highly realistic sequence shot that has been praised for its illusory and immersive qualities. In this regard, it is worth recalling the work of Scott Richmond (2016), who theorises the illusory effects of cinema in the basis of the proprioceptive effects that some films succeed to produce on the spectator. As observed in Chapter 2, Richmond (2016) argues that *Gravity* and other films take advantage of the immersive cinematic apparatus to create the illusion of a world that unfolds "before us onscreen" (ibid, p. 6), manifested through the sensation of movement produced by bodily responses such as tension in our arms and hands, increased heart rate,

pushing into the back of our seat, etc. Movie theatres' architecture and technological affordances make the cinema a quintessential proprioceptive medium; but its illusory effects are more likely to take place when filmmakers deploy a set of aesthetics that aim towards immersion.

The proprioceptive effects that Richmond describes were actually felt by the author of this thesis while having a theatrical experience of Whannell's *The Invisible Man*, especially in those moments when immersive continuity was applied. In specific scenes throughout the film, continuous long shots and 3-D sound are used to transport us into the terrifying events that Cecilia (Elisabeth Moss) passes through after escaping from her abusive husband. Sound designer Will Files describes one of these moments:

> When she's alone in a bedroom and she thinks she hears something, she's is kind of just playing with some stuff in the corner and then the camera starts moving around the room and it settles on what seems to be an empty window and then it slowly moves back to her. We did that using Dolby Atmos, so you can actually hear the room kind of spinning around. (Files in Dolby, 2020b)

Such a 'spinning around' effect has its origins in the iconic car sequence in Cuarón's *Children of Men* (00:26:13), which was shot as a single take using a circling camera inside a vehicle. Lately, the interior of cars has in fact been an interesting playground for filmmakers in their quest to create the sensation of presence in the space of the diegesis. The opening car chase in Steve McQueen's *Widows* (2018), for instance, deploys such a 'Cuarónian' strategy with the intention of locating us in the middle of the action as if we were one of the fugitives trying to escape from the police. Here, Dolby Atmos facilitated the continuous transit of all diegetic sounds, including dialogue, through the speakers arrayed all around the space of exhibition. Once again, music is absent, so it is the sound design and not the score what induces the sensation of peril throughout the individual segments that conform this action sequence.

In fact, action sequences, which have been traditionally known for its fast-cutting aesthetics, have lately proven to be a fertile field for the utilisation of such an immersive audio-visual style. An example of this can be found in David Leitch's *Atomic Blonde*, for instance, which presents an astonishing 10-minute sequence shot (from 01:10:38 to 01:20:20). Although this film conforms to Hollywood's fast cutting style for the most part, and hence presents a screen-centric soundtrack during almost the entire movie, this specific sequence immerses the spectator with sounds that to a great extent move according to the position of the sources in the 3-D space, something that was achievable thanks to the smooth spatiotemporal continuity that

only the long take can offer. There is no music throughout the entire fight, so it is not the score what drives the emotion of the scene, but rather a sound design composed only by diegetic elements distributed across the 3-D space.

Another powerful example of immersive continuity in action sequences features in Sam Hargrave's *Extraction*, a 'Netflix Originals' production that like many other recent movies was released on the streaming platform with a Dolby Atmos soundtrack. The film is cut fast and presents the typical style found in the Hollywood action/adventure genre; yet, there is a predisposition to pan sounds three-dimensionally from the very beginning, a strategy that is deployed every time the image allows this to happen without the risk of creating a jarring distraction. In the club scene at the beginning of the film, for instance, the diegetic music that comes from the club's stage dynamically moves in accordance with the visual perspective of the moving camera, thereby creating the effect of a lived space to which the director aims to transport us.

But what truly stands out in this film is the 11-minute sequence shot (from 00:34:50 to 00:46:21) that covers the whole extraction segment. This sequence begins by showing Tyler (Chris Hemsworth) and Ovi (Rudhraksh Jaiswal) running away from their persecutors, and as the camera follows the two protagonists from behind, the distant screams of Saju (Randeep Hooda) are heard behind us, which intends to create the sensation of being there running desperately with the characters. The sequence continues with a breath-taking car chase that emulates the car sequences previously mentioned, which is followed by a highly realistic police pursuit that takes place inside an apartment building in a densely populated area of Dhaka. Although this sequence does use music during the car chases and in specific moments throughout the final battle, this section abandons the conventional scoring strategy and gives space to the silent three-dimensionality of the setting to manifest itself, and through it enhance the sensation of presence in the diegesis.

Beyond the in-the-wings effect of the bullets that pass through us, the three-dimensional soundtrack conveys our immersive point-of-audition by dynamically changing the spatial signature of each location, as well as the ratio between direct and reflected sound in relation to our perspective. For instance, radios that play the country's traditional music are heard through the corridors, which become louder once we enter into the specific room where the radio is located; the sound of a liquid being boiled increases in volume as the camera approaches the cooker, and then disappears as we move away from its source; offscreen people running

through the building directs the movement of the camera or the characters towards their direction. All these different sound elements help to create a distressing sensation; they replace the infrastructure of the room where we are physically located with a fictional universe that we are drawn into.

This segment in a way resembles the manner in which Cuarón portrayed Mexico City in *Roma*, in which the spatial identity of the Mexican capital is audio-visually represented through long takes and 3-D sound, a strategy that helps to transport us to a specific time and place in the history of the country. A similar effect occurs in *Extraction*. As we follow the characters across the different areas of the building, we see and hear the elements that shape the reality of Dhaka and its inhabitants. The impression of reality that the film conveys is capable of convincing the audience that what is being presented is real, an effect that would have been intensified if the film's producers would have given us the chance of experiencing it in a movie theatre.

Immersive sequence shots are also found in moments of Atmos-enabled TV series such as *Games of Thrones*, *Black Summer*, *The Haunting of Hill House, Dark*, among others. In *Games of Thrones*' episode 9 season 6 (*Battle of the Bastards*, 2016), for example, director Miguel Sapochnik makes use of a highly immersive long take that follows Jon Snow (Kit Harington) as he fights against Ramsay Bolton's (Iwan Rheon) men in a battle for control of Winterfell (00:38:56). Throughout the shot, the sounds of the soldiers, horses, arrows and so forth move in perfect spatiotemporal synchrony with what we see on the screen, passing from onscreen to offscreen positions and vice versa, enhancing this way our sensation of presence in the battlefield. Once again, the musical score is muted for this particular sequence, which makes the sound design alone responsible for creating an emotional impact on the spectator.

In all the sequences analysed within this section, Dolby Atmos facilitated a greater exploitation of the sonic space; yet, sound spatialisation has started to gain a more immersive disposition when filmmakers embrace the long take as the basic core of visual creation, even in films that are mixed and released on less sophisticated platforms, namely 5.1 or 7.1 systems. That is the case of *Birdman* and other American productions such as Ryan Coogler's *Creed* (2015), Barry Jenkins' *Moonlight* (2016), and Cary Murnion and Jonathan Milott's *Bushwick* (2017). The fight sequence that follows the opening of Coogler's *Creed*, for example, features a long take that follows Adonis 'Donnie' Johnson (Michael B. Jordan) as he goes to the ring. In this scene, the movement of the camera is accompanied by changes on the soundtrack, whose

sonic atmospheres vary depending on the perspective of the camera. Later on, Johnson's fight against Leo 'The Lion' Sporino (00:56:56) is covered entirely in a single shot, with instructions from each corner being heard according to where the camera is within the ring, drawing us this way to feel like being there with the boxers.

Immersive continuity has also been embraced by world cinema filmmakers, either on individual sequences, or as the film style of entire movies. That is the case, for instance, of 'one-shot' films such as the German production *Victoria* (2015) directed by Sebastian Schipper, *Utøya: July 22* (2018) directed by Norwegian filmmaker Erik Poppe, and the Canadian drama *The Body Remembers When the World Broke Open* (2019) directed by Elle-Máijá Tailfeathers and Kathleen Hepburn. Some of the characteristics of immersive continuity are also found in the long take-based Holocaust drama *Son of Saul* (2015), the debut feature of Hungarian director László Nemes, whose final sequence brings to mind the uprising sequence in Cuarón's *Children of Men*. Gaspar Noé's *Climax* (2018), is another interesting example. Noé, a devotee of the long take, deploys immersive continuity on this French/Belgian coproduction to transport us into a terrifying party. Moreover, the opening sequence of Jorge Michel Grau's *7:19* (2016) also makes use of immersive continuity to transport us into a building that collapses as a result of a devastating earthquake in Mexico City. Along a 7-minute long take, all sounds move respecting the position that the camera occupies within the three-dimensional diegesis.

In summary, the cinema's intention of offering an experience with the power of immersing the audience within the narrative world has lately become more 'tangible'. A number of films have deployed a set of aesthetics that to a great extent are reminiscent of the film style recently pioneered by Alfonso Cuarón. As such, filmmakers have showcased their appetite for achieving what Bazin (1967) once referred to as the myth of total cinema, namely a cinematic experience during which we become invisible spectators, phantasmagorical beings that simply get transported into alternative realities. It seems clear that for a very long time the long take has offered the possibility of creating realistic sonic universes, and with all the tools that are currently available, filmmakers can finally conquer the everlasting ideal of spatial realism. Although the abovementioned films are certainly a small minority, they all serve to illustrate how the Bazinian vision is still alive, even in the most commercial Hollywood blockbusters.

Dolby Atmos and the Liberation of Dialogue

As we have seen, a number of contemporary filmmakers have deployed aesthetic traits that resemble Cuarón's immersive continuity, expanding the potential of the long take with the highly immersive capabilities of 3-D sound. Yet, as a film style, immersive continuity is still rare. Cuarón and Iñárritu seem to be the only mainstream directors who have embraced such aesthetics for entire films, with a few others who have deployed it within the independent sector. Since the arrival of Dolby Atmos, however, more and more filmmakers have started to manifest their desire of using sound in a more spatialised way, even when conforming to the most typical editing conventions. It is perhaps sufficient to hear the sound of a door emanating from a speaker at the back of the auditorium—such as the one Sebastian opens after being fired in *La La Land* (00:26:12); or the one that sounds in *The Nun* when Father Burke (Demián Bichir) enters a Vatican chamber to have a meeting with some of his superiors (00:07:17)—to argue that the so-called 'exit door' effect[31] is no longer a concern. Audiences have grown accustomed to hearing all types of sound effects coming out from any speaker in the room, and with Dolby Atmos as a tool for content creators, even dialogue has escaped from its long-lasting imprisonment.

Any type of sound can now be reproduced by any speaker in the room, as long as this does not produce a jarring distraction. In *Baby Driver*, for instance, dialogue is panned to offscreen positions every time the image allows it, but the film's attachment to the fast cutting Hollywood tradition impedes this approach to sound to be deployed at all times. Keating observes that since cinema's early years, "Hollywood's predominant method of scene construction relied on master-and-coverage technique" (2019, p. 151); a filmmaking approach that consists on shooting a large amount of footage from different perspectives (e.g. master shot, medium shots, close-ups, etc.), allowing "producers and editors to reshape scenes in postproduction" (ibid). Many years have passed since the introduction of such a language, and yet the master-and-coverage technique, together with the shot/reverse-shot structure, is still Hollywood's (and cinema's as a whole) predominant method of scene construction. Today, and particularly with tools such as Dolby Atmos, there is no technical restriction that could

[31] When "the audience believes a sound in the surrounds actually originates within the theater and looks away from the screen to find out the sound's source" (Kerins, 2011, p. 72)

impede a dialogue line to be reproduced by a rear speaker. Yet, coverage-based aesthetic editorial decisions can be a barrier for such an approach to take place.

In Hollywood's narrative cinema, with its fast cutting style, dialogue remains mostly still in its position behind the screen, moving only when the image allows it. For instance, in Zack Snyder's *Batman v Superman: Dawn of Justice*, a typical superhero Hollywood blockbuster, sound effects move constantly in the 3-D space, while dialogue most of the time emanates from the centre speaker. As part of an ongoing stylistic turn, however, there are moments in which dialogue leaves its onscreen position and reinforces the cinema's intention of creating a lived space, one whose acoustic characteristics resemble the ones we know from our own world. An example of that is found in the scene where a group of policemen find a number of Asian women locked in a cell in the basement of an abandoned house, and the desperate screams of a man being tortured are heard on our back, which leads the policemen towards that direction (00:18:06).

A much more interesting example comes from the work of Darren Aronofsky, whose treatment of the soundtrack provides good opportunities for analysis. Beginning with his work on *Black Swan* (2010), Aronofsky has positioned himself as a director interested in the potential of sound to immerse the audience into the space of the story. In *Mother!*, this is evidenced throughout the film, with a Dolby Atmos soundtrack that to a great extent follows the changes of perspective driven by the movement of the camera. Although Aronofsky's treatment of sound is not as bold as Cuarón's (sonic perspective does not change with every cut), the soundtrack of *Mother!* presents several moments of three-dimensional sound spatialisation, something that is easily noticeable since the film does not include a musical score.

After Him (Javier Bardem) invites Man (Ed Harris) to stay (00:08:50), for instance, the men's voices spread from front to rear and gradually become unintelligible as we move away with Mother (Jennifer Lawrence). While she walks towards the kitchen, she looks back and observes them from a distance, and the image cuts to show the two men from Mother's perspective. As this happens, their conversation is heard once again coming out of the front-centre speaker, but it returns to the rear once the image cuts back to Mother. The three-dimensional filmic space is thus constructed through a perfect interaction of image and sound. As in *Baby Driver*, however, the high number of cuts used in this and other scenes impedes this treatment of the soundtrack to be deployed throughout the entire film.

Dialogue panning is also present in films like George Miller's *Mad Max: Fury Road* (2015), Ridley Scott's *Alien: Covenant* (2017), Andrés Muschietti's *It* (2017), Corin Hardy's *The Nun* (2018), Reed Morano's *I Think We're Alone Now* (2018), among others. Morano, for example, deploys 3-D sound throughout the whole film, even though dialogue panning is pretty much limited to the horizontal axis, with voices that move from left to right or vice versa, but not too much to the rear speakers. This works very well in the scene in which Del (Peter Dinklage) and Grace (Elle Fanning) have their first conversation through the filer of the door that separates them (00:13:12). While we see Grace, the filtered voice of Del is heard on the left side of the screen, whereas Grace's offscreen voice is heard on the right as we see Del on the other side of the door. This sonic strategy locates ourselves as invisible spectators that have access to both character's perspectives, switching our position from being inside the room to being outside and vice versa.

The tendency to pan dialogue off the screen is also found in the work of Quentin Tarantino. In his black comedy *Once Upon a Time in Hollywood*, a very bold spatialisation of voices is employed for instance when Rick (Leonardo DiCaprio) and Cliff (Brad Pitt) are watching an FBI TV show in which the former appears (01:56:00). In this scene, a long take showing the television is employed while the two characters make comments about the show. During one and a half minutes, their voices are heard coming out of the rear, as if they were sitting right behind us. The picture then cuts to a two-shot of both sitting in the living room, which is followed by their dialogue coming back to the front. But when the picture cuts once again to a close-up of the television, their voices are once again heard on our back. This treatment of the dialogue track somehow resembles the scene of Cuarón's *Roma* in which the family of protagonists are watching a comic TV show (00:15:49), and their voices and laughter are heard behind us, an approach that in the purest Cuarón style is maintained throughout the whole film.

In Tarantino's nostalgic portrait of late 60s Hollywood, however, it is incoherent to hear the character's voices moving from front to rear depending on the perspective of the camera, while the sound that emanates from the television is always kept on the front centre speaker, disregarding its actual position in relation to the camera. In any case, this is not the only scene in which Tarantino pans dialogue to the rear speakers. In the scene in which Rick delivers a poor performance while acting for the pilot of the Western television series 'Lancer' (01:12:03), the voices of the director and his assistant are heard behind us, a strategy that gives

depth to the three-dimensional setting and helps to transport us into it. The same sound design strategy can be found in the opening sequence of Spike Lee's *BlacKkKlansman* (2018), in which Dr Kennebrew Beauregard (Alec Baldwin) is recorded on camera as he delivers his racist pro-segregation speech. During this scene, the voice of his assistant giving him instructions is heard coming out from the rear-right corner, where she is located. Moreover, throughout the Black Student Union meeting scene (00:16:05), the voices of several individuals among the crowd are heard behind us, thereby enhancing the illusion of being there with the students. In these and all the films mentioned within this section, however, dialogue is most of the time kept in the centre speaker, not because it cannot move, but because the number of cuts typical of intensified continuity presents an obstacle for treating the cinematic voice in that manner.

A more extensive spatialisation of dialogue is found in Bong Joon-ho's *Okja*, the first Netflix film to include a Dolby Atmos soundtrack. In this Korean/American coproduction, the sound concept was conceived from pre-production (Lee and Stringer, 2018), and 3-D sound is used as a tool to enhance the illusion of presence in the diegesis. For instance, in the first sequences that follows the opening (from 00:05:17 to 00:34:12), 3-D sound helps to transport us into a small village in the mountains of South Korea, where a young girl named Mija (Ahn Seo-hyun) lives with her grandfather (Byun Hee-bong) and their super pig Okja. In this introductory sequences, atmospheres, sound effects and dialogue change in volume, reverberation and position according to the perspective of the camera, a sound design approach that is found throughout the movie.

As in *Baby Driver* and *Mother!*, however, such a three-dimensional synchronisation is broken in several instances of the film, especially when the shot/reverse-shot formula is used. In *Okja*, it can be argued that Bong deploys a hybrid style, one that combines immersive continuity with the more typical cutting traits of Hollywood's narrative cinema. This resonates with what Lee and Stringer point out about Bong's treatment of the soundtrack in his films, which they argue "is the result of an effort to keep in alignment realistic (Korean) sound and audio elements with dramatic and sensory impacts more readily identified with Hollywood genres" (2018, p. 150), an approach that Bong's all-time supervising sound editor Tae-young Choi refers to as "realistic with impacts" (Chio in Lee and Stringer, 2018, p. 150).

The same can be said of Bong's latest film, *Parasite*, the first foreign-language film to win the Best Picture Academy Award in 2020. Bong, who with this movie also won the

Academy Awards for Best Original Screenplay, Best International Feature Film, and Best Director, successfully employed 3-D sound to transport the spectator into the semi-basement where the Kim family resides, and into the mansion in the hills where they work for the wealthy Park family. The film's immersive soundtrack allows us to experience vividly the contrasting characteristics of the South Korean capital. For instance, an interesting scene is the one in which the Kims have to escape from the comfort of the Parks' mansion and run towards their humble semi-basement located in one of the poorest areas of the city (01:31:40). A big storm takes place as they flee back in the middle of the night, producing a catastrophic flood in their neighbourhood leaving their apartment under water. Besides the immersive atmosphere of the rain falling upon us, the voices of the characters matching their geographical position in the diegesis gives a high dose of realism to the scene, which ultimately transports us into their tragedy. The film's soundtrack thereby helps to construct "the much-vaunted sense of place" that, as Lee and Stringer notice, is typical of a Bong film (2018, p. 150).

Although Dolby Atmos has been primarily used on mainstream American films, successful films like *Roma* and *Parasite* are a demonstration of how this sonic platform has gradually expanded to international territories, within which Dolby's latest sonic platform has helped to produce content with highly immersive moments. Another example is Arregi, Garaño and Goenaga's *The Endless Trench* [*La Trinchera Infinita*] (2019), a Spanish historical drama that tells the story of Higinio (Antonio de la Torre), a Republican who at the outbreak of the Spanish Civil War hides in a hole in his house in order to survive the Falangist repression. From the very first frame, the film's soundtrack helps to transport us into the narrative world through a three-dimensional interaction between image and sound, an approach that is maintained in many sequences throughout the film. The first time Higinio runs away from the Falangists (00:06:10), for instance, the soundtrack locates us in the middle of the action by placing the guard's voices, footsteps and gunshots on the rear speakers as they follow the protagonist. Throughout the movie, the soundtrack induces us to identify with Higinio, as it offers a highly realistic representation of the way he hears from inside his hidden spot, specialising dialogue and sound effects in accordance with the perspective offered by the camera.

In summary, with the advent of Dolby's latest sonic platform, three-dimensional approaches to sound spatialisation have gradually become more recurrent, with more and more films that present moments in which dialogue and sound effects move all across the auditorium.

This aesthetic innovation only confirms that the 'exit door' effect has stopped being perceived as a problem, and that even dialogue can nowadays be panned to extreme offscreen positions, something that years ago was thought to be inadequate. However, it is important to bear in mind that three-dimensional dialogue panning can become a distraction if it is applied in moments that deploy conventional editing, as the voice would have to move repeatedly from one speaker to another in a rather jarring manner. Indeed, any sound that is jarring from one speaker to another could pull out the audience out of the story. In other words, even though this strategy is today technically feasible, it can be argued that the screen-centrality of film sound will remain unchanged as much as editing conventions do not evolve towards the achievement of the mythical total cinema.

Conclusion

Following Cuarón's *Gravity*, the film style that this thesis refers to as immersive continuity has been identified in a number of films. This would suggest that *Gravity* can no longer be considered an experimental film, and that its audio-visual aesthetics, far from being an isolated event, are becoming a growing trend across genres. Whereas Alejandro González Iñárritu has fully embraced this emergent film style in his two most recent films, other mainstream directors have deployed it on specific moments to heighten the immersive experience of their films. The long take on its own has grabbed the attention of a number of filmmakers around the world, and today it is a frequent visual strategy often complemented with the action of 3-D sound.

But even when long takes are absent, highly spatialised approaches to sound design have been observed in a variety of recent films. These findings suggest that the screen-centric sound design approach that for ages has governed film production is no longer the rule, and that audiences have grown accustomed to hear all kinds of sounds coming out from different directions across the auditorium. Virtually all films today treat the cinematic experience as a three-dimensional phenomenon; it is thus frequent to have character-related sounds, noises and atmospheres moving in the 3-D space in accordance with the position that the camera occupies within the diegesis. Even dialogue has started to move more freely, putting an end to the everlasting tradition inherited from monophonic sound.

Yet, in despite of all this, nothing has really changed. The cinematic voice is most of the times kept in the centre speaker, not because technology does not permit its movement, but because the conventional editing patterns used in the vast majority of films impede its free transit within the three-dimensional diegesis. Whereas the treatment of the visuals that Cuarón and Iñárritu have embraced is exceptional, most films are chopped in many fragments as if that was the general filmmaking rule. Although it is true that cutting is a powerful creative tool, it is also true that conventional editing has played against the cinema's nature of being, which, as suggested by Bazin (1967), was conceived by its creators with the intention of offering lifelike three-dimensional experiences. Montage, once referred by Eisenstein to as the essence of cinema, has in fact prevented the cinema from achieving its totality. Editing, as it is conventionally practiced, is today the only obstacle for the fabrication of virtually inhabitable cinematic worlds.

Although creating a fully immersive soundtrack is not always necessary, and perhaps not appropriate for particular stories, scenes or genres, the fact that sound three-dimensionality has been extensively applied in a variety of subsequent productions is a confirmation of Cuarón's 'gravitational pull'. His outstanding work has initiated a highly plausible stylistic turn, and this academic contribution aims to provide insights for a better and more effective utilisation of 3-D sound.

CHAPTER SIX

Immersive Continuity and Beyond: A Practical Methodology for Designing Films with 3-D Sound

Broadly speaking, the filmmaking process can be divided into four main stages: development, preproduction, production and postproduction. During development, the screenwriter prepares the screenplay that the director and their collaborators will transform into a motion picture following a set of creative/technical operations and workflows throughout the preproduction, production and postproduction stages. During these last three stages, the director plays a fundamental role as the person responsible for the final audio-visual product. However, to what extent do directors take responsibility for the sonic component of their films? In his seminal essay 'Designing a Movie for Sound', sound designer/supervisor and re-recording mixer, Randy Thom (1999) observes that:

> Many feature film directors tend to oscillate between two wildly different states of consciousness about sound in their movies. On one hand, they tend to ignore any serious consideration of sound (including music) throughout the planning, shooting, and early editing. Then they suddenly get a temporary dose of religion when they realize that there are holes in the story, weak scenes, and bad edits to disguise. Now they develop enormous and short-lived faith in the power and value of sound to make their movie watchable. Unfortunately it's usually way too late, and after some vain attempts to stop a haemorrhage with a bandaid, the Director's head drops, and sound cynicism rules again until late in the next project's post production. (Thom, 1999)

As Thom notes, in terms of sound, creative decisions have been traditionally relegated to postproduction, which undoubtedly limits the potential that this component could offer. Thom proposes that in order to take advantage of sound, filmmakers should not only be concerned about recording good and intelligible sound on set, or simply hiring a sound designer to manufacture sound effects, "but rather to design the film with sound in mind, to allow sound's contributions to influence creative decisions in the other crafts" (Thom, 1999). Considering the challenges and opportunities offered by the latest technologies of immersive sound, it is now more important than ever to discover new and more effective ways to manage the audio-visual contract, giving sound its full and deserved value.

Based on Cuarón's work and on the ideas of Randy Thom, Sergio Díaz and other key film sound practitioners, this chapter intends to present a practical methodology for incorporating 3-D sound within the creative decisions that take place early in the screenwriting stage during development and throughout the process of preproduction, a proposal that will be put in practice for the creation of an original screenplay and its accompanying storyboard. Ultimately, the main aim of this PhD project is to academically establish an immersive filmmaking style with its own rules of film construction; that is, the goal is to propound a set of guidelines that could facilitate the application of 3-D sound as an essential characteristic of what this thesis has termed immersive continuity.

Sound Design as the Core of Immersive Cinema

The term sound design is still as vague and controversial today as it was when it was first used. The intention of Walter Murch, the first sound professional that was credited as a sound designer for his work in Coppola's *Apocalypse Now* (1979), was to establish a supervising role for a person in charge of sound during the whole creative process. Larry Sider (2003), sound professional and co-director of the *School of Sound*,[32] argues that Murch invented the role of sound designer as a response to the need of having an individual responsible for the soundtrack during the whole creative process, a demand that is commonly ignored even though sound represents a big part of the audio-visual experience. Murch describes the sound designer as:

> Somebody who took on the responsibility of 'auralizing' the sound for the film and making definitive, creative decisions about it. Someone the director can talk to about the total sound of the film the way he talks to the cameraman about the look of the film (Murch in Kenny, 1998: p. 20).

According to Wright (2013), the concept of sound designer was conceived during the creation of *THX 1138* (1971), a film that Murch co-wrote in collaboration with George Lucas and Francis Ford Coppola. In this film, Murch's work was credited as *sound montage*, an ambiguous designation that Murch chose in order to describe his role as the person in charge of the overall soundtrack, given the fact that the International Alliance of Theatrical Stage

[32] "The School of Sound is an international annual forum to explore of the use of sound in audio-visual media. Organized by Larry Sider, it has been held yearly in the UK since 1998" (Sider, 2003: p. 8).

Employees (IATSE) labour union prohibited him to accept credit as the film's sound editor or mixer (Wright, 2013). Constantini argues that since the days of American Zoetrope, Coppola's film studio established in San Francisco at the end of 1969, we understand that there is an individual who is responsible for the soundtrack just like the Director of Photography is for the image (2018, p. 169). The idea within this independent studio was to honour the importance of sound by giving it as much as attention as was given to other crafts (Andriano-Moore, 2018: p. 546). Adriano-Moore (2018) explains, however, that it was only outside the regulations of Hollywood that such an approach to sound design was able to form (ibid). Moreover, what was known as the 'Northern California approach' to film sound, encouraged a production practice under which the sound designer was also the re-recording mixer, meaning that the sound designer was as well responsible for the final sound mix (ibid).

At the time, the sound unions working in Hollywood mandated strict divisions of labour, each with its own duties and responsibilities (Whittington, 2007: p. 55). In Murch's vision, however, having different people in charge of the different stages within the sound chain, "was as if there were two directors of photography on a film, one who lighted the scene and another who photographed it, and neither could do much about countermanding the other" (Murch as cited in Andriano-Moore, 2018: p. 550). Randy Thom (1999), whose first job in the film industry was precisely as part of the sound department on Coppola's *Apocalypse Now*, recalls:

> It was the dream of Walter Murch and others in the wildly creative early days of American Zoetrope that sound would be taken as seriously as image. They thought that at least some films could use the guidance of someone well-schooled in the art of sound in storytelling to not only create sounds but also to coordinate the use of sound in the film. This someone, they thought, would brainstorm with the director and writer in pre-production to integrate sound into the story on the page. During shooting that person would make sure that the recording and playing-back of sound on the set was given the important status it deserves, and not treated as a low-priority, which is always the temptation in the heat of trying to make the daily quota of shots. In post production that person would continue the fabrication and collection of sounds begun in pre-production, and would work with other sound professionals (composers, editors, mixers), and the Director and Editor to give the film's soundtrack a coherent and well coordinated feeling.

In the context of this thesis, the sound designer is understood precisely as that 'holistic' individual that, as Greene and Kulezic-Wilson explain, is meant to be involved since the very beginning, from preproduction to postproduction, and not only as the person that creates sound effects in postproduction (2016, p. 16). Sound practitioner Javier Quesada comments:

> In the ideal scenario, I would sit down and talk with the director from the script review […] because sometimes they don't know all the possibilities, hence the sound designer has to come and talk to them and say 'look, you can do this, here in this scene it occurs to me that this could happen', thus to improve the script in that sense so when the shooting day arrives they also consider that certain things are going to happen in sound. And more importantly, when postproduction arrives, to be all in the same frequency […] In the ideal world, I would sit down with the director and talk about the script, the possibilities, the format, the budget, and how you want to land the film […] Like the cinematographers, who are involved very early, I think that sound designers also have to be involved very early and go hand in hand clarifying to the director everything that is possible and everything that is not possible to get the most out of the film and the story. (personal interview, 2019, translated by author)

Such an aspiration, however, "has been a difficult one to realize, and in fact has made little headway since the early 1970s" (Thom, 1999). Sider claims that "[o]nly Murch, usually combining his roles as film editor and sound designer and a few others working at the high end of the film industry, can accurately call themselves sound designers, practising the kind of control over a soundtrack that Murch foresaw" (2003, p. 6). Larry Loewinger (1998), another contemporary sound practitioner, asserts that unfortunately, after Murch's debut in the 1970s, the sound designer concept never took hold. In reality, Loewinger (1998) explains:

> a soundtrack is created by a team—a production mixer, post production sound supervisor, sound effects editor, composer and re-recording mixer, to name just a few—that often resembles a patchwork rather than a unified chain of command. Unlike cinematography, there is no one person who sits at the top of the pyramid.

The concept of sound designer as invented by Murch has actually been misused, and as Thom (2019c) argues, "most people even inside the film community think that the only thing sound designers do is manufacture sound effects for science fiction and other fantasy films" (n.p). As a response to this dilemma, the title of *supervising sound editor* has gradually become the industry equivalent of the sound designer, which up to a certain point resembles Murch's original idea in the form of a supervisory role that would be in charge of sound throughout the whole process. In practice, however, not all films have a supervising sound editor (Johnson, personal interview, 2019) and if there is one, this person is generally hired in postproduction (Andriano-Moore, 2018), leaving the stages of preproduction and production without a clear path to follow.

But if that is so, how could a film take full advantage of the story possibilities offered by sound; and more specifically, how could immersive sound technologies such as Dolby Atmos be employed to their full potential? Present-day sound systems challenge filmmakers to plan, design and produce their movies acknowledging that, unlike the possibilities of previous

eras, today they have the opportunity of creating audio-visual products that not only comprise two-dimensional images, but could also include three-dimensional sounds. Today, more than ever, filmmakers need to recognise that, as *Roma*'s sound designer and supervising sound editor Sergio Díaz puts is, "sound and image are 50/50, they cannot be divorced" (personal interview, 2019, translated by author).

Although in practice most of the creative decisions related to sound are generally made in postproduction (Dancyger, 2006), there are a few exceptions worth consideration and analysis. Alfonso Cuarón, for instance, has demonstrated the efficacy of thinking in terms of sound from the very beginning. As discussed in Chapter 4, a good example of that is found in *Roma*, a film that showcases many of the aesthetic and practical strategies upon which this thesis is grounded. Cuarón (the film's screenwriter, producer, director, cinematographer, and editor) not only recruited supervising sound editor and sound designer Sergio Díaz to be involved from preproduction—a period during which he started gathering the sounds that could later give shape to the film's soundtrack—but more importantly, Cuarón wrote, planned, shot and edited the film with immersive sound in mind. As re-recording mixer Skip Lievsay points out, Cuarón created the film so that the sound team could work in a three-dimensional manner (Interview by Dolby, 2018). *Roma*'s sound editor Javier Quesada comments:

> Alfonso had it very clear that it was a 360-degree movie, where you are inside and things happen around you. So much so that in all the scenes he asked us to work on what happens onscreen, and what happens behind, outside and on the sides […] We had to fully dress the 360-degree film, almost as if you were doing something for virtual reality but made for cinema. I loved that approach; I think that all cinema should be like this, at least all cinema that intends to be experiential. (personal interview, 2019, translated by author)

Hence, beyond hiring a supervising sound designer/editor to be there during the whole process, the first and most important consideration for creating a film with 3-D sound is to write, plan, shoot and edit the picture in such a way that sound three-dimensionality could be coherently added later in the filmmaking process. In other words, in order to use the full potential of immersive sound technologies such as Dolby Atmos, the first 'sound designer' should be the screenwriter, followed by the director during pre-production and shooting, and the picture editor during the editorial stage, all of which should work in constant communication with a supervising sound designer/editor throughout the entire process. Moreover, in order to plan the film and its three-dimensional sound design in a more detailed manner, this thesis suggests that sound annotations should not only be printed on the script, but

also on the film's storyboard. In relation to this, sound designer Randy Thom (2019b) comments:

> Storyboards are simple drawings done during preproduction and production as a guide for, and to begin a conversation about, the content of shots and scenes. There should be sound design storyboards too, because it would present an opportunity (and an obligation) to actually think about, discuss, and plan the way sound will work in each scene, and in the movie as a whole. It would happen early enough so that sound ideas could have an influence on the way scenes are blocked, lit, and paced, as well as where they are shot. (n.p.)

The type of immersive cinema that this thesis aims to promote challenges filmmakers—primarily screenwriters, directors, storyboard artists and picture editors—to consider the fact that while the image is framed on the screen, the soundtrack is meant to expand in the three-dimensions according to the position of sound sources within the diegesis. In other words, in order to enhance the cinema's impression of reality, filmmakers should respect the continuum of space and time not only in terms of visual information, as traditional continuity teach us, but also in terms of sound and its arrangement in the space of exhibition. Consequently, this thesis argues that continuity editing should be replaced by a new continuity system, something that shall be proposed in the pages to come; but let us first review the well-established continuity editing and the limitations it presents for creating films with 3-D sound.

The Limitations of Continuity Editing

Bordwell, Thompson and Smith argue that an intimate understanding of continuity editing is a necessary requirement to become a director, a cinematographer, a performer, or an editor; and explain that this style intends to smoothly transmit the narrative information by means of using a series of well-organised shots (2016, p. 230). Similarly, Katz claims that "a filmmaker who learns the various framing and staging ideas of the continuity style will gain a heightened awareness of composition, editing patterns and three-dimensional design" (1991, p. 4). One of the most important components of continuity editing is the 180-degree system, under which filmmakers build the visualised filmic space around an *axis of action,* determined by what is commonly known in filmmaking as the *180-degree line* (ibid). The function of such an imaginary line is to establish the area within which the camera should be placed during shooting in order to present the action in a continuous way.

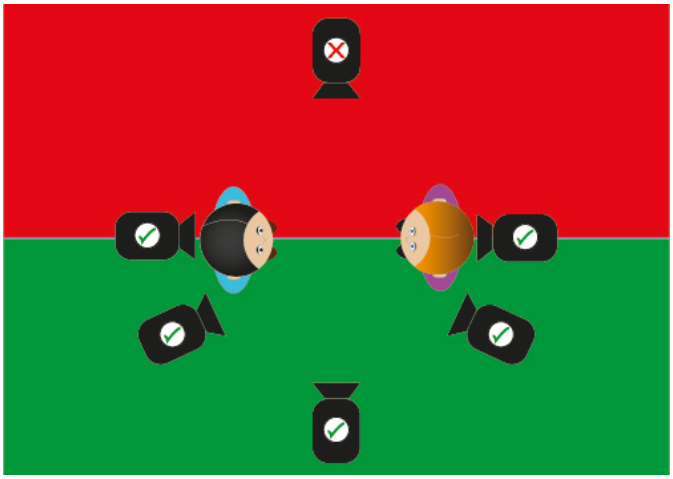


Figure 6.1 - The 180-degree system: The axis of action and the prohibited space.

In Figure 6.1, the green area represents the axis of action, which includes all the possible positions where the camera should be located to capture the action in the scene; the red area, on the other hand, represents the prohibited space. Shooting from the green area, that is from the axis of action, will maintain the relative position of the actors in the frame—Actor A seen on the left and Actor B seen on the right. However, if the camera operator captures the scene from the red area, the relative position of both actors will result in the opposite—Actor B seen on the left and Actor A seen on the right. Furthermore, the 180-degree system keeps consistent lines of sight, that is to say, when shooting from the axis of action Actor A will always be seen looking to the right, and Actor B will always be seen looking to the left. Moreover, if the actors move, when walking for example, their direction will be consistently maintained. If Actor A walks to the right, for instance, their direction will be shown as moving from left to right when being shot from the axis of action (Figure 6.2). However, if shot from the prohibited space, it would appear as if Actor A was moving from right to left.

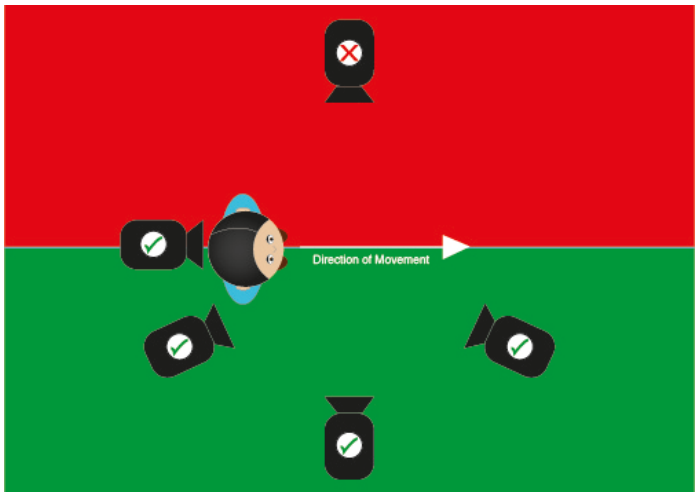


Figure 6.2 - The 180-degree system: Direction of movement.

Under such a framework, a scene can be constructed by assembling shots captured from different perspectives (commonly known as *coverage shots*), whose combination will keep continuity on the visuals as long as they were recorded from within the axis of action. This means that a director—in agreement with the director of photography (DoP)—may decide to place the camera in front of one actor at a time to frame a close-up of their face. In fact, it is quite common that if two characters are having a conversation, close-ups of both faces are captured individually in order to show one face and then the other in a shot/reverse-shot structure.

Certainly, continuity editing and the 180-degree system aim to maintain a consistent visual continuity when combining a series of shots. However, as we shall see, although such a methodology has proven to be very effective for keeping continuity on the image, it fails to consider the three-dimensional sonic relations and hence indirectly forces the sound mixer to keep dialogue and other important narrative sounds in the centre of the screen. Ultimately, the cutting practices derived from continuity editing obstruct the application of sound three-dimensionality and limit the potential of the acoustical space.

Imagine for instance a conversation between two characters on a scene shot and cut using continuity editing. Actor A and Actor B will be shown on the left and right side of the frame respectively. When they talk, their voice could be panned to the left and right following their position on the screen. However, what would happen if the camera frames a close-up of one of the characters (Actor A), while the other character (Actor B) is speaking? If the re-recording mixer wants to maintain a coherent relative position of the voices in the space of the diegesis, Actor B's voice would have to be panned to the rear. If the scene continues using close-ups in a shot/reverse-shot manner, the voices of the characters would have to repeatedly switch from the front to the rear speakers and vice versa, which most likely result on confusing the audience. As a result, the re-recording mixer is forced to rather keep dialogue on the centre speaker, sacrificing the possibility of creating the sensation of presence in an 'inhabitable' cinematic space.

For example, let us examine the diner scene in Edgar Wright's *Baby Driver* (2017), a film that, as noted in Chapter 5, deploys sound three-dimensionality in several moments throughout the movie. After a few quick shots that show the action from different perspectives, an 18-second panning shot is used to show Baby (Ansel Elgort) sitting on a table as Debora (Lily James) enters through the diner's main door.

Figure 6.3 - Two frames takes from *Baby Driver*. Devora passes from the onscreen to the offscreen space within the same shot.

Following the characteristic three-dimensional audio-visual match of the immersive continuity style, Debora's voice and footsteps are heard transiting from the front to the rear speakers as she moves form an onscreen position to a zone outside the frame (Figure 6.3). This treatment of sound connects the visible and the acoustical spaces, and ultimately helps to construct the illusion of presence in the 'unlimited' space of the diegesis, complementing with 3-D sound the camera's monocular vision and hence establishing the spectator's imaginary position as a sort of invisible auditor that is present in the narrative world. However, such a brief illusion is quickly lost in what follows due to the introduction of the familiar shot/reverse-shot formula, which, as mentioned before, is a conventional device of classical continuity and its upgraded version, referred to by Bordwell (2002) as intensified continuity.

Figure 6.4 - Four frames takes from *Baby Driver*. The typical shot/reverse-shot pattern.

After Debora returns from the kitchen, a conversation between the two characters takes place. During their chatter, a series of quick coverage shots are used to show the characters from different perspectives, within which an A-B-A-B pattern stands out. Through a combination of medium shots, over-the-shoulder shots and close-ups that comply with the 180-degree system (Figure 6.4), visual continuity is maintained; yet, the previously constructed

immersive point-of-audition is lost because all dialogues are kept in the front centre speaker, disregarding the actual position of the characters in the space, which in a way put at risk the cinematic impression of reality achieved at the beginning of the scene. This treatment of the visuals thus impedes a fluid and organic audio-visual assemble, it reduces the possibilities of sonic immersion and hence the illusory capabilities of the cinematic apparatus.

If Julian Slater—*Baby Driver*'s sound designer, supervising sound editor and re-recording mixer—had wanted to maintain the impression of reality throughout this scene by means of a coherent and three-dimensional audio-visual match, he would have had to repeatedly pan the voices of the two characters from the front to the rear and vice versa, which, as sound practitioner Richard King stresses, would have resulted in something rather jarring and distracting (School of Sound Conference, 2019). King argues:

> The idea of moving dialogue with the camera just doesn't work when there is a lot of edits. […] It's very jarring and it actually shows up, it highlights the mechanics of filmmaking in a way that I don't think is seductive to drive the audience into the emotionality of the film. […] In my mind, it works when there are maybe long shots, like in *Gravity* or *Roma*, where the audience can feel that they are in space and the camera is simply moving around. (School of Sound Conference, 2019)

Ultimately, the system of continuity utilised in the above scene forced the sound team to keep all dialogue in the centre of the screen, a limitation found in the vast majority of films. Certainly, the audio-visual approach employed in the scene discussed above may help the spectator to understand the story by accessing a series of objective points-of-view and providing intelligible dialogue lines. After all, it has been taken from granted that all the cinematic tools must work in service of the story. But this is not what cinema is all about. As Richmond points out in his analysis of *Gravity*, "the narrative is only one small aspect of a complex and integral modulation of the viewer's affects" (2016, p. 138). Indeed, for Cuarón, the story is just "like the cinematography, the sound, the acting, and the color. They are tools for cinema, and what you have to serve is cinema, not story" (Cuarón as cited in B, 2013, p. 41).

Inspired by this idea, one of the main purposes in this thesis has been to synthesise the ways in which the picture could be planned and designed in order to serve the immersive cinematic experience, whose illusory effects are only complete when the soundtrack adds to the picture the possibility of hearing in a lifelike three-dimensional manner. 3-D sound endows the image with a high level of realism, it allows each shot to have not only a specific point-of-

view, but also its own immersive point-of-audition. The idea behind this sonic strategy is precisely to represent natural human hearing; hence, if filmmakers start thinking, designing, shooting, and editing the picture with three-dimensional sound in mind, it is more likely that tools like Dolby Atmos will be used to offer a more sophisticated illusion of reality, and not simply as a commercial add-on. Sound designer Randy Thom comments:

> If a director wants to really exploit the story potential of Dolby Atmos, he or she should think about the space in which action happens when they are still writing or re-writing the script, and certainly when they are shooting the film. And if they do that and think about it seriously, I think it would have an effect on camera positions and camera moves and where actors look off screen […] and it will affect so many different kinds of creative decisions in terms of how they shoot the film. […] I look forward to a point in time when they do think in those terms. (personal interview, 2019a)

In the above example taken from *Baby Driver*, for instance, sound three-dimensionality could have been fully exploited by reducing the number of cuts, by limiting the utilisation of close-ups, by using camera movement, by adding depth to the setting with sounds that dynamically move across the space, etc. This does not mean that doing so would have guaranteed that the film would have been better, but those decisions could have helped to manufacture a more immersive scene.

Immersive Continuity and the 180-3D System

The long take is on its own a visual technique that clears the way for a fluid and three-dimensional sound spatialisation, meaning that 3-D sound works better in the absence of cuts. Although there are a number of film examples that play as single continuous shots, cutting is in general an important cinematic device that cannot be simply left behind. In this regard, this section aims to introduce a system of spatial continuity that could improve the sonic deficiencies of continuity editing by considering each shot not only as a two-dimensional visual recording of the diegesis, but as a three-dimensional audio-visual fragment of the story world. The intention of such a continuity system is to serve as a guide for the elaboration of scenes that shall be compatible with 3-D sound, something that could be achieved without sacrificing the virtues of the cut.

Walter Murch, who had the highly uncommon opportunity to work as both picture editor and sound designer in films like *Apocalypse Now* (1979) and *The English Patient* (1996),

argues that an ideal cut satisfies six different parameters: 1) it is coherent with the emotion of the moment; 2) it moves the story forward; 3) it happens at the right moment when it is rhythmically interesting; 4) it considers the audience focus of interest and their 'eye-trace' within the frame; 5) it respects what is referred to as 'planarity', that is how photography transposes the grammar of three dimensions into two; 6) and it "*respects the three-dimensional continuity of the actual space* (where people are in the room and in relation to one another)" (2001: p. 18, emphasis added).

Although for Murch the order of priority should go from parameter 1 to 6, being 1 the most important and 6 the least important one to comply with, in the context of this thesis's proposal, it is precisely the sixth parameter the one that needs to be deeply understood in order to augment the possibilities of sonic immersion. Respecting the three-dimensional continuity of the actual space is what would allow the sound department to create a fully three-dimensional soundtrack, and it is this approach to sound design what immersive continuity intends to prop up. Hence, while Murch argues that "sacrificing spatial continuity is well worth it" (ibid, p. 19) to create the right emotion, this thesis suggests that in pursuit of immersion, filmmakers should find a way to maintain emotion, story and rhythm without sacrificing eye-trace, planarity and the continuity of the three-dimensional space, something that can only occur if the scene is designed in such a way since the very beginning. In other words, the idea is to find ways to comply with all those parameters, if not in the whole film at least in those scenes where immersion is of key importance.

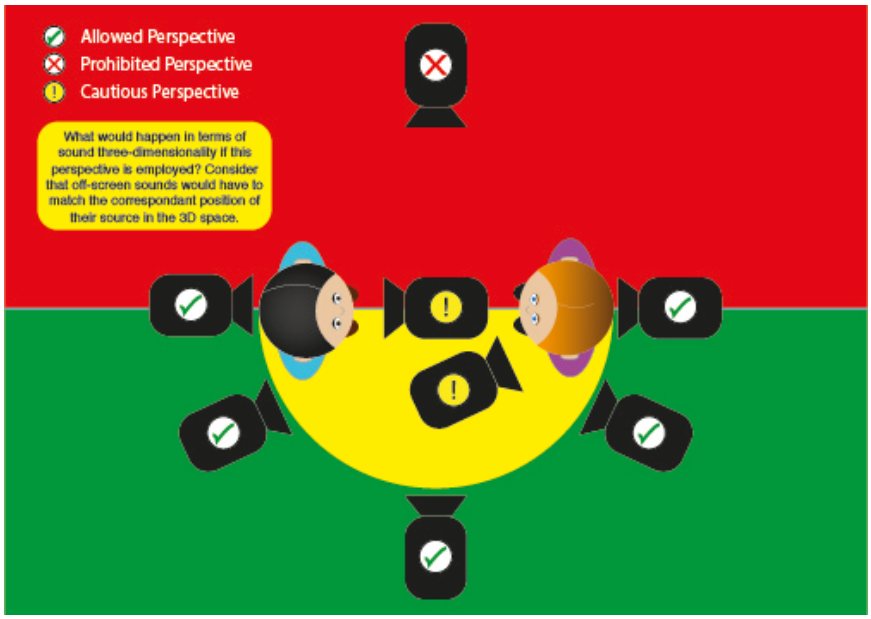


Figure 6.5 – The 180-3D System

In order to move towards a broader adoption of 3-D sound, this thesis suggests to start using an upgraded version of the 180-degree system, one that adds a few specific indications that could facilitate the application of a fully three-dimensional sound design. Essentially, this alternative system, which from now on will be referred in this thesis to as the *180-3D system*, aims to highlight that the utilisation of close-ups might, in certain circumstances, be a barrier for the deployment of sound three-dimensionality. The idea is to prevent having offscreen voices emanating from the centre speaker behind the screen, something that becomes inevitable when close shots of faces in the form of shot/reverse-shot patterns are employed. With that in mind, this thesis proposes adding an imaginary semi-circle around the actors as illustrated in Figure 6.5, which could be taken as a reference to establish a new border within the axis of action. Thus, the 180-3D system divides the possible shooting zones in three different areas: the green area, the yellow area, and the red area. Following Cuarón's approach to filmmaking, this system is based on *audio-visual perspectives* instead of monocular camera positions, meaning that the camera is understood as an invisible entity that sees partially in one direction, but hears in all directions in a three-dimensional manner.

It is important to notice that the green and red areas already existed in the traditional version of the 180-degree system, the green area representing the axis of action, and the red area representing the prohibited space. This alternative version, however, adds a yellow area, which represents the portion of the space that exists between the characters, a space that shall be used cautiously in order to respect the three-dimensional continuity of the diegetic space. Commonly, in a conversation between two characters, this area is used for shooting close-ups of both faces—the camera is placed in front of one of them, leaving the other offscreen. Therefore, when shooting or editing footage that was captured from within the yellow area, filmmakers should consider whether or not there is dialogue or other important narrative sounds that are being emitted by an offscreen source during the duration of the shot. They will have to take into account that if 3-D sound is wanted, in the end such sounds would have to emanate from a speaker located in a point in the room that matches the physical position of its source in the diegesis, which may result in something rather distracting if it not handled with care.

As explained before, cutting from a close-up of one face to a close-up of the other in a shot/reverse-shot structure, for example, could result in dialogue switching from one speaker to another, making it difficult to keep a consistent audio-visual match without distracting the

audience with dialogue that repeatedly move to opposite positions in the auditorium. Ultimately, the yellow area demands a specific treatment, it warns filmmakers to consider how the three-dimensionality of sound will work on each particular shot and within their combination.

Figure 6.6 - Two frames taken from *Roma*. A long take-driven shot/reverse-shot pattern in *Roma*.

It is worth mentioning that the 180-3D system does not forbid the utilisation of close-ups and other perspectives taken from the yellow area, but it encourages filmmakers to use them with discretion. *Roma*, for example, is a film primarily shot from the green area, resulting in an abundant number of wide-shots and a scarcity of close-ups. Yet, Cuarón does use the yellow area in a few specific scenes, and some of them even include shot/reverse-shot patterns in moments of dialogue between characters. For instance, the hotel room scene (from 00:24:59 until 00:27:28) includes a few shots that were captured from within the yellow area. During 2 minutes and 27 seconds, the scene includes a total of ten shots, resulting in an average shot length of 15 seconds. Hence, although this scene presents cuts that switch the direction of the camera from one actor to the other (Figure 6.6), each shot is maintained for a few seconds

before cutting to a reverse-shot. As a result, Fermín's voice does move from front to back on three occasions during the scene, but the length of each shot facilitates this perspective changes to be more easily assimilated.

It is worth considering that in moments like the one described above, where the shot/reverse-shot formula is needed to show action and reaction within a conversation, directors could employ over-the-shoulder shots in order to prevent having dialogue 'teleporting' from one speaker to the other. Over-the-shoulder shots have the particular quality of keeping two or more subject in the frame, which justifies having their dialogue in the centre speaker without losing the spatial coherence of the scene.

Although the example provided in Figure 6.6 only specifies a conversation between two actors, a similar imaginary framework could be established when there are more than two characters involved. In a conversation at dinner, for instance, the yellow area is the inner space that exists between the characters, which the table helps to physically represent (Figure 6.7). In such circumstances, filmmakers should carefully consider what would happen if shots were captured from within the yellow area (i.e. from within the table), and how offscreen voices would have to behave in order to respect the three-dimensional continuity of the space.

Figure 6.7 - Frame taken from *Roma*. The table in this shot represents the yellow area of the 180-3D system.

Embracing the 180-3D system as the framework of spatial continuity could certainly help filmmakers to achieve more immersive results even when using cuts; it is important to highlight, however, that beyond creating their films with three-dimensional sound in mind, the most important consideration that will allow them to maintain coherent shot-to-shot relations in the 3-D space is the utilisation of the long take as the fundamental core of the film's visual aesthetics, a characteristic that lies at the heart of the immersive continuity style. Moreover,

the utilisation of deep-focus and wide-shot driven cinematography could be beneficial for the utilisation of 3-D sound, as it allows the audience to observe foreground, mid-ground and background elements, and hence may reduce the need of cutting to particular elements shown in detail. Another fundamental strategy for being able to exploit the full potential of immersive sound is the utilisation of smooth camera movement for transiting within the filmic space, which can be done through camera panning, Steadicams, dollies, cranes, or any other moving mechanism. Of course, blocking the actors in such a way that they will be able to move appropriately within the filmic space, both onscreen and offscreen, would be also necessary. And as a final consideration, using long POV shots could also contribute for the immersive experience, as this type of shots are a mechanism to access character's subjectivity by means of representing the subject's visual perspective, which could be complemented by a three-dimensional representation of the subject's point-of-audition.

Embracing Immersive Continuity throughout the Planning of the Film

What differentiates immersive continuity from classical or intensified continuity is the three-dimensional coordination that the former maintains between image and sound. When embracing this film style, screenwriters should try to describe the events as if they were happening in a three-dimensional space, with the intention of endowing the camera (or the embodied character in the case of a POV shot) with a 'pair of ears' capable of perceiving sounds in a three-dimensional manner. Unfortunately, as Randy Thom observes, "[m]ost screenwriters either don't take sound design very seriously, haven't thought about it much, or think of it as something that happens only in post production" (2019d, n.p). Analysing the screenplay of a film like *Roma* is hence important, as it stands out as a highly uncommon case where the screenwriter took also the responsibility of being the first 'sound designer' of the film.

Figure 6.8 shows the first section of the script of *Roma*'s opening sequence. In less than a page, Cuarón describes in great detail not only the way the scene will look, but also the way it will sound. Sonic elements such as water, grimy foam, a wet floor, bubbles, a radio announcer in the distance, birds singing in their cage, a toilet flushing, and water in the sink, are all details included in the screenplay with the intention of hearing them in the final audio-visual product. Evidently, Cuarón wrote his movie with sound in mind, and although the screenplay does not provide clear indications of the position of those sounds within the geography of the film, it

offers a good mental picture of all the sound elements that will coexist in the scene. Such sonic descriptions are abundant throughout the screenplay, and in several occasions, they refer to elements that we do not see, but simply shape the environment that surrounds the characters. Many of those sounds are in fact essential for transporting the spectator into the Mexico of the 1970s, a country and a period whose soundscape has its own sonic identity.

Thursday, September 3rd, 1970

INT. PATIO TEPEJI 21 - DAY

Yellow triangles inside red squares.

Water spreading over tiles. Grimy foam.

The tile floor of a long and narrow patio stretching through the entire house: On one end, a black metal door gives onto the street. The door has frosted glass windows, two of which are broken, courtesy of some dejected goalee.

CLEO, Cleotilde/Cleodegaria Gutiérrez, a Mixtec indigenous woman, about 26 years old, walks across the patio, nudging water over the wet floor with a squeegee.

As she reaches the other end, the foam has amassed in a corner, timidly showing off its shiny little white bubbles, but -

A GUSH OF WATER surprises and drags the stubborn little bubbles to the corner where they finally vanish, whirling into the sewer.

Cleo picks up the brooms and buckets and carries them to -

THE SMALL PATIO -

Which is enclosed between the kitchen, the garage and the house. She opens the door to a small closet, puts away the brooms and buckets, walks into a small bathroom and closes the door.

The patio remains silent except for a radio announcer, his enthusiasm melting in the distance, and the sad song of two caged little birds.

The toilet flushes. Then: water from the sink. A beat, the door opens.

Cleo dries her hands on her apron, enters the kitchen and disappears behind the door connecting it to the house.

Figure 6.8 – First Section of the script of *Roma*'s opening sequence.

Aiming to take Cuarón's screenwriting approach to the next level, this thesis proposes that a screenplay could even include indications that describe the position that sounds occupy in the 3-D space, as well as their trajectory if those sounds move. In order to make this work effectively, one of the things that could be written in the screenplay is our position (the position

of the camera) as invisible spectators that 'inhabit' the diegesis. In other words, the screenplay could describe what we see at each particular moment, as well as our movement within the filmic space. An effective way to do this is evident in the 2012 version of the screenplay of *Gravity*, written by Alfonso and Jonás Cuarón. Within it, it is common to encounter indications such as "CAMERA TILTS UP WITH Matt" (p. 3), "She flies PAST CAMERA and continues in the b.g." (p. 19), "CAMERA PULLS BACK OUT of her helmet"| (p. 21), "CAMERA TRACKS WITH Ryan to reveal Matt" (p. 38), "CAMERA PANS TO the cabin window, then back to Ryan." (p. 64), and so forth. Such capitalised instructions could help to clearly indicate what exactly is meant to happen on the screen, which could be complemented by sound-related notes.

Once our visual perspective is stablished, specific sound elements could be described thinking on their position in relation to where we stand; but given that it is difficult to portray a whole three-dimensional space only with words, the storyboard artist could help to plan the film considering the position and movement of characters and objects within the diegesis.

Maurice Zuberano (production illustrator and art director), refers to the storyboard as the "diary of the film", and argues that "it is a diary about future events" (as cited in Katz, 1991, p. 23). For Katz, storyboards serve two main purposes: they permit filmmakers to visualise their ideas, and facilitate the communication of those ideas to the crew (1991, p. 23). Halligan notes that one of the first people to be involved in preproduction is the storyboard artist, who is in charge of creating the 'blueprint' for the final film (2013, p. 8). This person (or persons), Halligan observes, works closely with the director to translate the screenplay into the first vision of what the film will be. Although it is important to provide enough detail, Halligan argues that the beauty of a storyboard is only valid if it offers the director the ability to visually discover their *three-dimensional film* (ibid, p. 9). Temple Clark, storyboard artist of films like *Harry Potter and the Prisoner of Azkaban* (2004), *Children of Men* (2006), *Gravity*, and many more, comments about the importance of the storyboard:

> Why would you build a house without a plan? Or why would you do anything without thinking it through? […] In storyboards firstly you find out so many problems, that when it is written in the script you read it and you go 'yeah, yeah, that's all fine', and then you start to draw it and you go 'hang on', and you find the problems that you resolve, you can solve them beforehand. But also, not so much with Alfonso [Cuarón] because he knows what he wants a lot, but other directors like to bounce ideas backwards and forwards, and like to talk about it. The usual thing with directors is they know the opening of the scene, they know the core of the scene and they know the

> ending; and the rest we kind of talk about it and we bounce ideas about it. […] So [the storyboard] develops the ideas and enriches the ideas a lot. And then of course it is just time to give it to the producers, to the art department, to visual effects, to special effects and they would know exactly what is going on. (personal interview, 2020)

It is worth mentioning that although filmmakers understand the diegesis as a three-dimensional space, storyboards are commonly utilised to visualise each shot considering only what will be seen within the camera frame. They only provide a visual reference for what is happening onscreen, ignoring this way most of the offscreen elements of the *mise-en-scène*. This preproduction practice in a way reinforces the screen-centric convention that this thesis aims to redress, and therefore calls for changes that can adapt visualisation practises onto the immersive continuity style. Directors could ask storyboard artists to provide more information in each of their drawings. Instead of focusing only on illustrating the elements that will be projected on the screen, each drawing could include notes or indications for the most important offscreen elements that produce sound and their position in the 3-D space. For example, if at a particular moment the camera is framing a medium shot of one character, and there is another character approaching from the rear-left corner, the storyboard illustration should include an annotation about the second character as well. It would thus be useful to propound effective and perhaps standardised ways to do so, which is one of the purposes of this chapter's proposal.

But before focusing on this thesis's storyboarding methodology, let us analyse the ways in which a storyboard is commonly done. Figure 6.9 shows a segment of the storyboard of *Gravity*. As the film was conceived thinking of long takes and camera movement, Temple Clark, storyboard artist of the film, makes extensive use of arrows to represent the movement of the camera, the characters, or the objects that are part of the story world. In this particular scene, Clark uses specific arrows to indicate the direction of the camera, and the movement of both Dr Ryan Stone and her helmet. Each arrow has a word or a letter on it, which specifies whether they are representing camera movement (CAM), camera panning (PAN), the direction of the character (R as an abbreviation of Ryan), or the direction of the helmet (HELMET). In the case of the latter, the 'HELMET' arrow indicates that the object is leaving the frame and then bounces back towards the rear right corner, something that is also written in the side notes. Furthermore, by drawing curved lines that connect each frame to the next, Clark is indicating that all the action is meant to happen within a single shot.

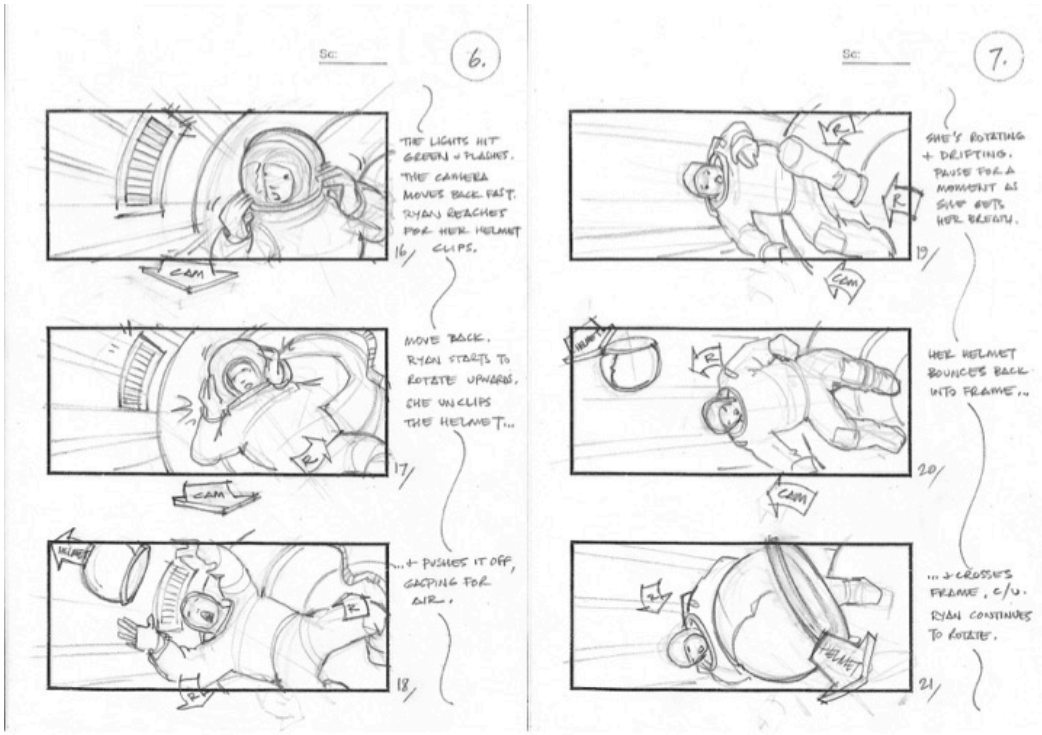


Figure 6.9 - A segment of the storyboard of *Gravity*

Although Clark's indications are very clear and useful, they do not specify what is happening beyond the limits of the screen. Once the objects are completely out of frame, we lose track of their position within the three-dimensional diegesis. Yet, they will still produce a sound, such as the helmet that clashes with a surface just after leaving the visible space. So, it would perhaps be a good idea to include an indication for each or at least the most important offscreen sound elements, which should be drawn off-frame, close to the physical position that such a sound will have once it is played back through a speaker in the auditorium. The idea is to make it clear that even though there are only a few elements being shown within the camera frame, there are some others that coexist within the same space. Those offscreen elements have their own specific position within the three-dimensional diegesis, and thus should be reproduced through a specific speaker (or group of speakers) in the cinema. By graphically visualising the offscreen elements in the storyboard, directors (as well as their collaborators, i.e. production designers, cinematographers, sound designers, editors, and other members of the creative team) would have a better understanding of what is happening beyond the limits of the screen in any given moment, and this could help them to plan and shoot each scene in a three-dimensional manner.

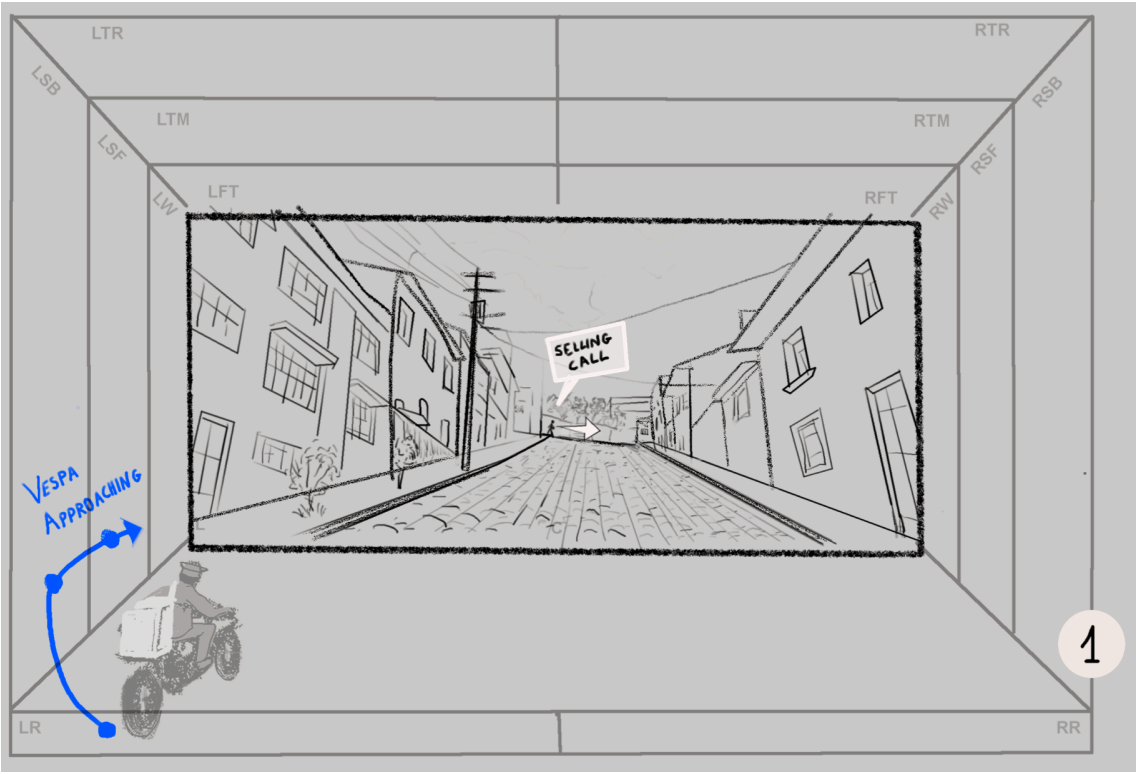


Figure 6.10 - Prototype of 3-D Storyboard

In order to schematically represent the position of those sounds within the theatrical sonic layout, this thesis proposes using a storyboard frame that includes all the different zones where the speakers are normally located, considering an 11.1.6 Dolby Atmos configuration[33] as the basic skeleton. As such, the 3-D storyboard prototype that the author propounds (Figure 6.10) uses a central frame for drawing the image, plus a set of boxes distributed throughout the acoustical space of the auditorium. Each of those boxes represents a zone of speakers, which are identified using the following nomenclature: Left (L), Centre (C), Right (R), Left Wide (LW), Right Wide (RW), Left Side Front (LSF), Right Side Front (RSF), Left Side Back (LSB), Right Side Back (RSB), Left Rear (LR), Right Rear (RR), Left Top Front (LTF), Right Top Front (RTF), Left Top Middle (LTM), Right Top Middle (RTM), Left Top Rear (LTR) and Right Top Rear (RTR).

As shown above, this storyboard schema provides the opportunity (and the obligation) of including sound-related annotations considering the position that each individual sound will have when reproduced in the exhibition environment. In this particular example taken from this thesis's original storyboard (Appendix 2), the blue dots represent three different positions where the Vespa motorcycle is heard as it approaches the camera frame, a trajectory during

[33] 11 speakers or sets of speakers surrounding the 4 walls of the room, 6 in the ceiling and 1 subwoofer.

which the sound of the motorcycle gradually emanates from specific speakers distributed across the rear and left walls. The arrow going beyond the final dot indicates that the vespa will continue moving towards the front. Ultimately, this storyboarding method can help to visualise the film with its accompanying three-dimensional sound, and hence may help filmmakers to take creative decisions with 3-D sound in mind.

A Practical Application of The Proposal

In this last section, the immersive continuity style, its system of spatial continuity, and the preproduction methodology discussed within this chapter will be used for the creation of an original screenplay and its accompanying storyboard, which in the form of appendixes shall help to practically illustrate the proposal. As previously observed, immersive continuity is a film style that intends to enhance the sensation of presence in the world of the movie by means of a lifelike recreation of the space of the diegesis. Evidently, not all films require this type of spatial treatment, yet there are some in which the interaction between characters and environment can contribute to create a more absorbing cinematic experience.

Roma can be described as an autobiographical film that was inspired by Cuarón's own personal memories. The screenplay of the film is therefore a detailed narration of moments that he or his close ones lived in the past, that is, it is a film inspired by the nostalgia of a time and a place. Following the lead of *Roma*, this thesis presents the original screenplay, "*Letters"* (Appendix 1), which takes 1970s Quito, the capital of Ecuador, as the spatiotemporal setting of the story. The 'America' neighbourhood serves as the scenario where the characters interact with each other and with the environment. In order to deploy the filmmaking methodology that this thesis propounds, the spatiotemporal continuum of the Ecuadorian capital is meant to be represented with images and sounds that conform to the immersive continuity style, with onscreen and offscreen sound elements being treated in a three-dimensional manner from the development and preproduction stages.

- ***Letters*' Screenplay and 3-D Storyboard**

EXT. "AMERICA" NEIGHBORHOOD, RIOFRIO STREET - DAY

An uphill cobblestone pavement street.

TITLES FADE IN:

SUPERIMPOSE:

"Quito, August 1972."

TITLES FADE OUT.

In the background, we see a NEWSPAPER VENDOR passing by chanting his traditional selling call.

NEWSPAPER VENDOR
(Shouting)
Mercioooooo!

In the distance, the characteristic sound of a VESPA motorcycle is heard approaching from behind. ARTURO SUASNAVAS, a postman (45), enters INTO FRAME as the sound of his VESPA is heard passing from rear to front. Arturo parks the bike and walks downhill while WHISTLING to announce his arrival.

Figure 6.11 - First section of the original screenplay "Letters".

Letters tells the story of Arturo, who in his daily working routine as a postman is confronted with a series of unexpected events that connect the inhabitants of the Quito of the 1970s with their identity, idiosyncrasy, culture and social problems. As an illustrative excerpt, the scripted sequence presented within this thesis follows one of those events, narrating Arturo's encounter with Edmundo, a jealous father who decides to bribe the postman in order to break his daughter relationship with her lover, a somewhat problematic young man who has left the country looking for work opportunities abroad. Aware of his own financial problems, Arturo agrees to stop delivering all the letters that come from the young man in exchange for a small sum of money; a corrupt act that will likely have consequences on the future of the enamoured couple. This original screenplay was written by this thesis's author considering 3-D sound and its application throughout the sequence, and therefore includes a number of sonic descriptions that intend to be the basis for what comes next: a storyboard designed using the methodology that was proposed in the previous section.

Figure 6.11 shows the first scene of the screenplay. This short sequence was envisaged as a single shot, with everything happening continuously in time and space. Inspired by the

ideas portrayed in *Roma*, the setting is described considering environmental characters that were part of Quito during that period of time, such as the newspaper vendor and his characteristic selling call. Moreover, the screenplay indicates that the sound of the postman's Vespa, a typical vehicle of that period of time, is meant to be heard passing from rear to front, creating this way the illusion of being within a lived space.

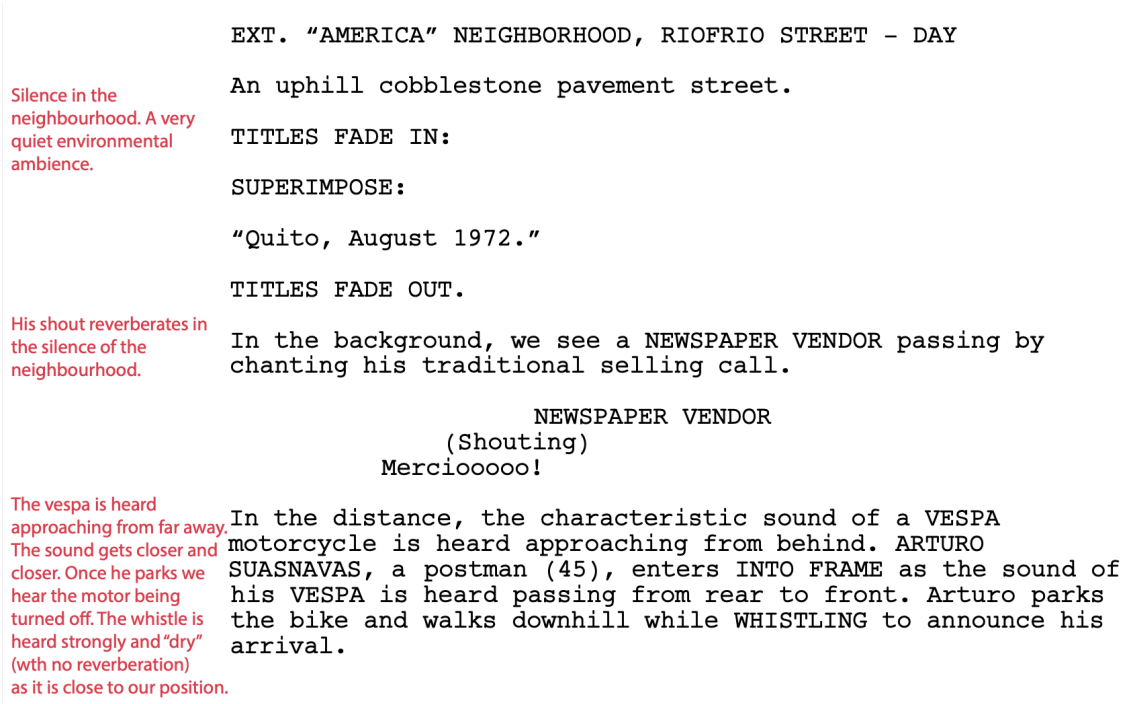


Figure 6.12 - First section of the original screenplay "Letters" with sound-related annotations.

Figure 6.12 shows the same exert from the screenplay but includes some sound-related annotations. This is explained in a graphic way in Figures 6.13a and 6.13b, which apply the 3-D storyboarding method described in the previous section. In Figure 6.13a, the blue dots and lines suggest that the motorcycle sound is meant to be heard passing from a speaker or zone of speakers positioned in the left rear (LR) of the theatre to the left side front (LSF) of the auditorium. In the case of a system like Dolby Atmos, this sonic transition could be translated onto a bigger number of speakers, which would help to make the sound movement smoother. In Figure 6.13b, the sound of the Vespa is already in the left wide (LW) area of the auditorium, and moves towards the front speakers. Similar to what happens with the newspaper selling call, the postman's whistle is annotated inside the frame, meaning that it is meant to emanate from a speaker that is close to the position of the postman on the screen.

Figure 6.13a – Frame taken from the 3-D Storyboard of "Letters".

Figure 6.13b - Frame taken from the 3-D Storyboard of "Letters"

In Figure 6.14, we can see Gardenia, one of the main characters of the story, sitting on a chair in her bedroom as the postman's whistle is heard filtering through the window, meaning that the high-pitched sound will have its own spatial signature as heard from Gardenia's location. In the screenplay, this segment is read as follows: "From inside the house, the postman's WHISTLE is heard by the beautiful GARDENIA PAREDES (17), who quickly stands up from a chair […]"[34].

[34] See Appendix 1, p. 1

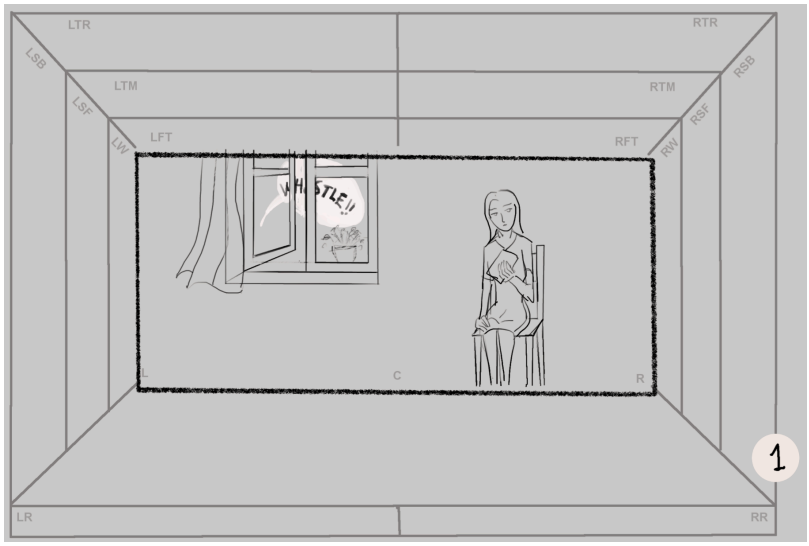
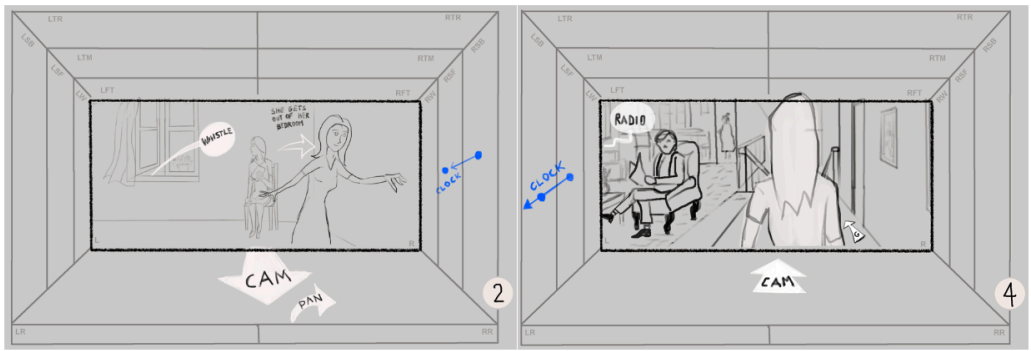


Figure 6.14 - Frame taken from the 3-D Storyboard of "Letters"

Then, Figure 6.15a shows Gardenia running as the camera moves back and pans to the right. Simultaneously, the blue dots and line in the storyboard indicate that there is a clock sounding in the right side back (RSB) of the auditorium, a sound that moves towards the theatre's rear side front (RSF) as Gardenia runs out of the room. In this case, the clock does not move but its sonic signifier does as the camera perspective changes. A similar effect is storyboarded when Gardenia is running across the corridor (Figure 6.15b), in that case, the clock is firstly located at the left side of the camera, and so it is heard coming out of the left side front (LSF) of the auditorium. But as the camera moves forward, the blue line indicates that the clocking sound shall move towards the back. This sonic element is mentioned in the screenplay by describing the character's trajectory, who "runs towards the stairs, passing by a loud grandmother CLOCK that is hanging in the left wall"[35].



Figures 6.15a (left) and 6.15b (right) – Two Frames taken from the 3-D Storyboard of "Letters"

[35] See Appendix 1, p. 1

An important aspect of this storyboarding methodology is that it graphically indicates when the voice of a character is meant to be heard in positions other than the front speakers. Figure 6.16a, for instance, shows Gardenia running down the stairs, while the voice of her father is heard behind us coming out of the left rear (LR) corner, where he is physically located. Additionally, the storyboard indicates that the sound that emanates from the living room radio shall be heard in the left side back (LSB) area of the theatre. In this case, the screenplay reads as follows: "His strong and intimidating VOICE is heard behind us as we go - DOWNSTAIRS - with Gardenia. The sound of the RADIO filters from upstairs and slowly disappears as Gardenia gets to the - GROUND FLOOR CORRIDOR – […]"[36]

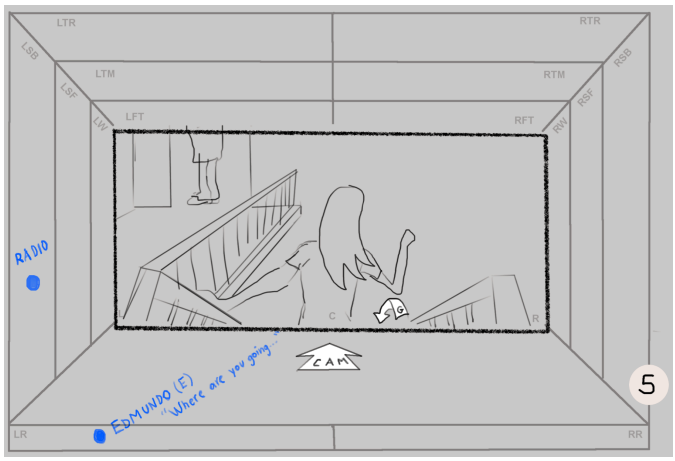
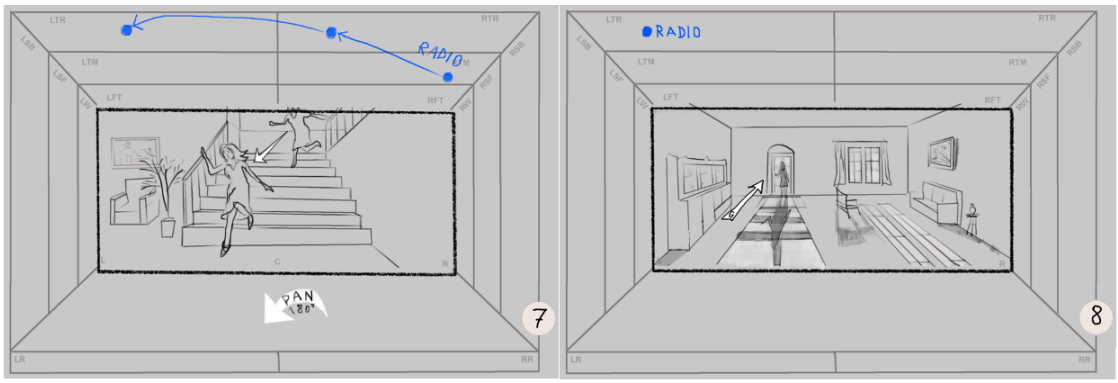


Figure 6.16a – Frame taken from the 3-D Storyboard of "Letters"



Figures 6.16b (left) and 6.16c (right) – Two Frames taken from the 3-D Storyboard of "Letters"

As Gardenia runs downstairs, the storyboard (Figure 6.16b) indicates a cut in the picture, suggesting that the camera is meant to be located in the ground floor, framing the young woman from the front. There is a 180-degree panning indication, and therefore there is

[36] See Appendix 1, p. 2

also an annotation for the sound of the radio coming from above, which given the camera panning would have to gradually move from being in the right top middle (RTM) zone of the ceiling to the left top rear (LTR) zone as the camera rotation takes place. In Figure 6.16c, the radio sound is represented only with a blue dot because the camera is fixed in that position, meaning that all sounds that do not move stay also fixed in a determined position in accordance with the visual perspective.

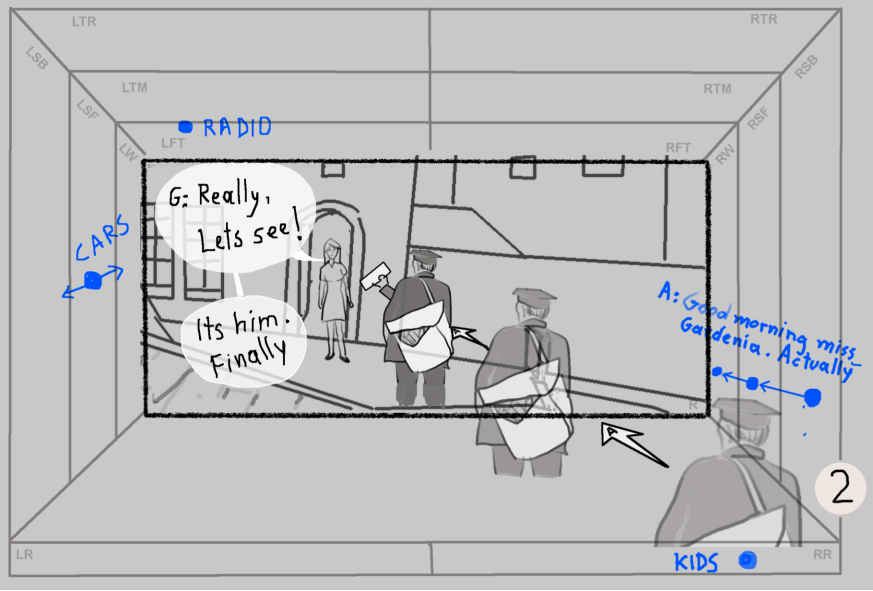


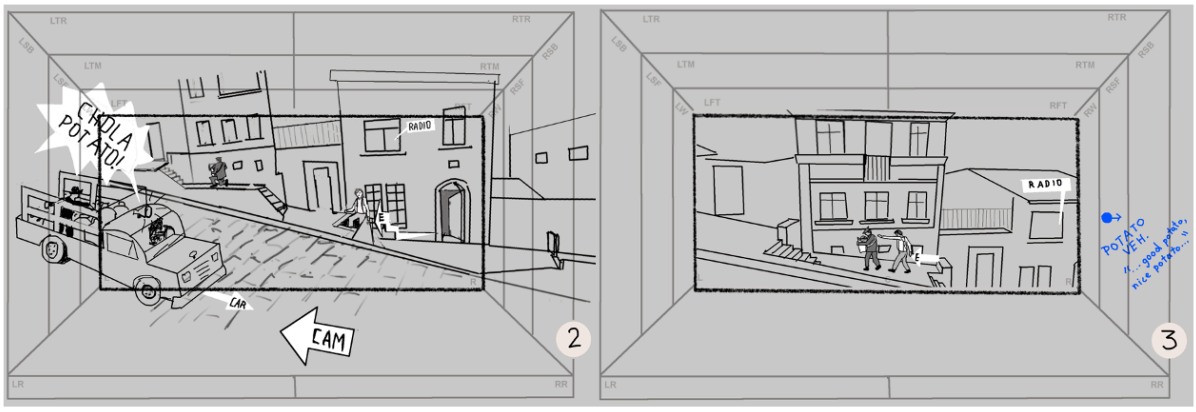
Figure 6.17 - Frame taken from the 3-D Storyboard of "Letters"

Dialogue can also move when a speaking character moves within the diegesis, and when this happens, this storyboard methodology allows the artist to illustrate the movement of voices in the 3-D space. This is illustrated, for instance, when Arturo speaks with Gardenia for the first time (Figure 6.17). In this storyboard frame, we see the postman moving from the rear right corner and entering into frame. Consequently, his voice is presented in the form of blue dots and lines that move from the right side back (RSB) zone to the right wide zone (RW) of the theatre. In this particular frame, there is also a blue dot to represent the sound coming out of the living room radio that filters from upstairs; a dot with two lines that indicates that cars are passing in both directions on our left; and a dot that represents the voices of the kids that are playing with the Vespa motorcycle behind us. This section is written in the screenplay as follows:

> As he opens the window, the sound of the living room's RADIO filters from above and is heard by Gardenia, who looks above and sees her father. Edmundo dissimulates.

Gardenia awaits at the door of their purple house, while a couple of curious kids go out running to see the postman's new bike […].

Arturo searches for The Letter in his bag while he walks towards Gardenia. He finds the letter and gives it to the young woman.[37]



Figures 6.18a (left) and 6.18b (right) - Two Frames taken from the 3-D Storyboard of "Letters"

As mentioned before, immersive continuity is based on the long take as a fundamental stylistic resource, yet, this does not mean that cutting is not permitted. Figures 6.18a, 6.18b and 6.18c show, for instance, three consecutive frames that take part in the fifth scene of the sequence. The first two frames represent a tracking wide-shot that frames Arturo and Edmundo from afar as they walk uphill, which is followed by a full shot of the two characters. In the first two frames, we see an indication that represents the living room radio that filters through the house's upper window. In the third frame, this indication disappears as the window moves offscreen, and therefore there is a blue dot positioned in the right wide (RW) zone, suggesting that such a sound should come out from a speaker (or group of speakers) in that area. During this segment, the camera tracks with Edmundo and then cuts to frame the two characters, which is described in the screenplay as follows:

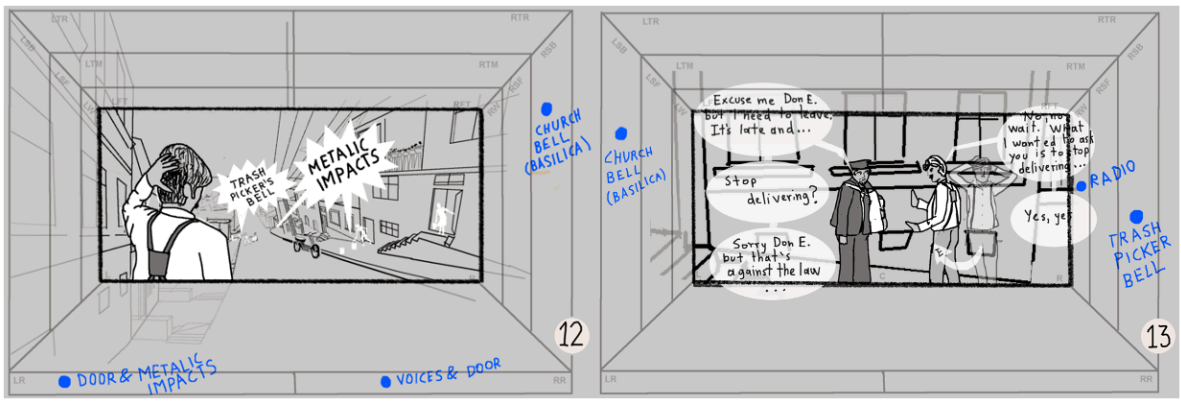
> CAMERA TRACKS WITH Edmundo as he walks uphill aiming to speak with Arturo. A POTATO SELLER passes by on a TRUCK chanting his loud SELLING CALL through a megaphone. The selling call continues from a distance as the truck goes down the whole Riofrio Street.[38]

[37] See Appendix 1, p. 2

[38] See Appendix 1, p. 3

Figure 6.18c - Frame taken from the 3-D Storyboard of "Letters"

Similarly, when the trash picker jingles his bell to announce his arrival, there is a cut to an over-the-shoulder shot (Figure 6.19a) that frames Edmundo in the front while at the same time we see the trash pickers in the far background. In this precise moment, the storyboard indicates that the sound of a distant church bell shall be heard in the right side back (RSB) corner of the theatre, while some other noises are heard behind us. What follows is a cut to a medium full shot of the two characters (Figure 6.19b), which should cause the sound of the church bell to move from the RSB position to the left side back (LSB) area of the theatre, something that is indicated on the storyboard frame.



Figures 6.19a (left) and 6.19b (right) - Two Frames taken from the 3-D Storyboard of "Letters"

Some of the sounds that conform the soundscape of the scene are described in the screenplay as follows:

Edmundo hesitates on what to do while a few neighbors go out carrying their metal trash bins, whose content the trash pickers deposits in a big metal container before carelessly dropping the NOISY METAL BINS. After a few seconds, a distant CHURCH BELL announces that it is 10am.[39]

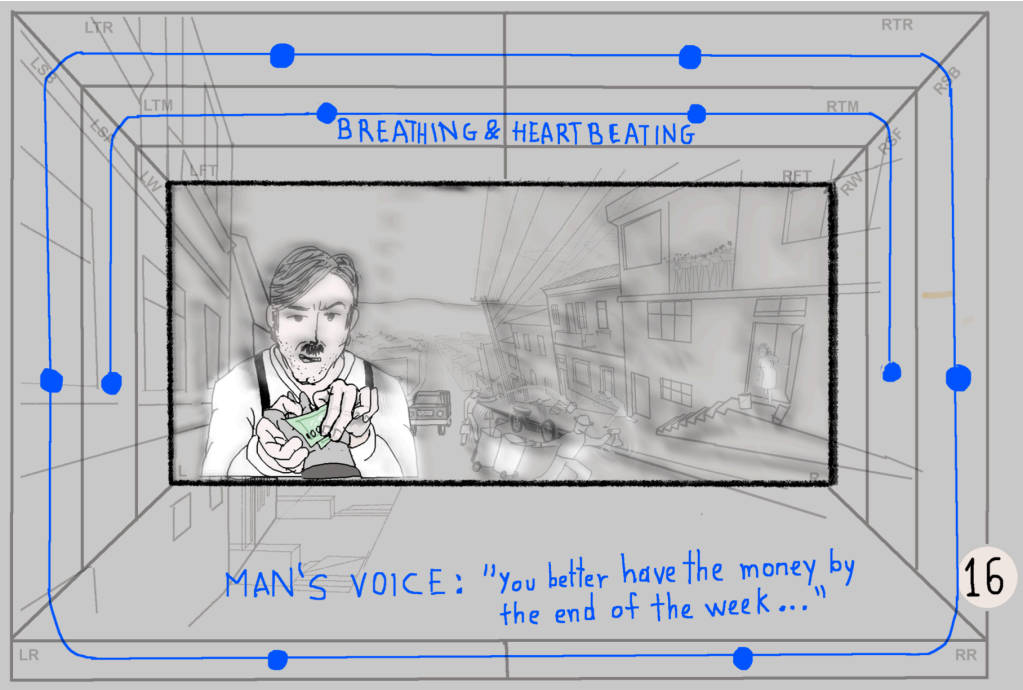


Figure 6.20 - Frame taken from the 3-D Storyboard of "Letters"

The last section of the sequence shows the moment in which Edmundo bribes Arturo, an act that is experienced from the subjective perspective of the protagonist. The idea is to see everything blurred from the postman's somehow disturbed POV (Figure 6.20), while his heartbeat is heard all around us, together with the voice of an unidentified man sounding inside Arturo's mind. In order to represent this experience on the storyboard, a number of blue dots connected with blue lines are used. They indicate that the heartbeat should come out from the speakers located in the centre of the room, while the man's voice should come out from the speakers that are positioned in the rear, creating this way the illusion of hearing an inner voice. This moment of intense subjectivity is narrated in the screenplay as follows:

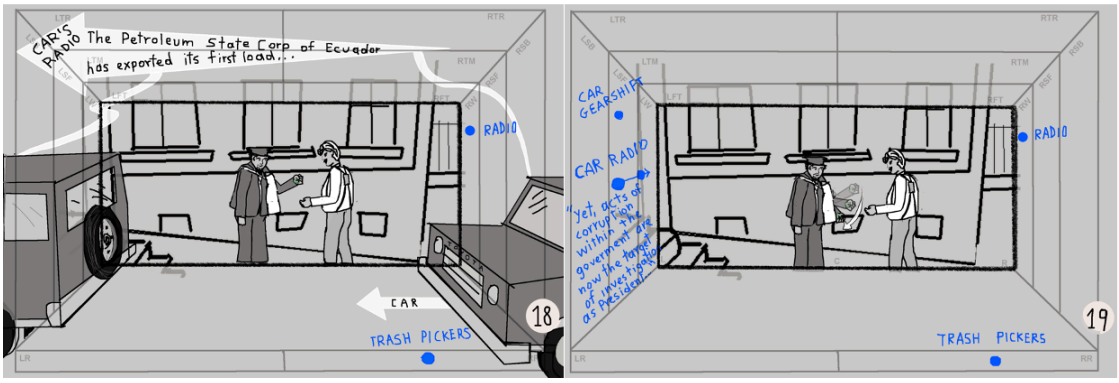
> We see Edmundo secretly handling money to the postman. The sonic atmosphere gets FILTERED as we begin to hear from Arturo's somehow altered point-of-audition. His HEARTBEAT and BREATHING put in evidence his nervousness. An echoing VOICE coming out of his own mind is heard.[40]

[39] See Appendix 1, p. 5

[40] See Appendix 1, p. 6

After this highly subjective moment, the sonic atmosphere returns to normal, and the sound of a radio coming out of a passing car takes a key role as its message is crucial for the narrative. The car is firstly seen passing through the screen carrying its sound with it (Figure 6.21a); and as the car moves off frame, the sound of its radio is represented with a blue dot that moves from the left side front (LSF) to the left wide (LW) area of the theatre, where it is meant to gradually vanish as the car turns onto that direction (Figure 6.21b). The screenplay narrates this section as follows:

> The street's SONIC ATMOSPHERE returns to normal as the voice of a NEWS ANCHOR is heard coming out from a CAR's RADIO as the vehicle stops in the corner […]
>
> The CAR leaves and the RADIO sound vanishes.[41]



Figures 6.21a (left) and 6.21b (right) - Two Frames taken from the 3-D Storyboard of "Letters"

The sequence and its storyboard finish with a wide-shot of the neighbourhood, as several sounds decorate the three-dimensional space that the spectator is invited to inhabit (Figure 6.22). The screenplay finishes with the following lines:

> The three-dimensional ATMOSPHERE of the neighbourhood surrounds us with sounds that come from all directions.
>
> A few CARS passing by, VENDORS in the distance, the trash picker jingling his BELLS in the next corner.[42]

[41] See Appendix 1, p. 6

[42] See Appendix 1, p. 6

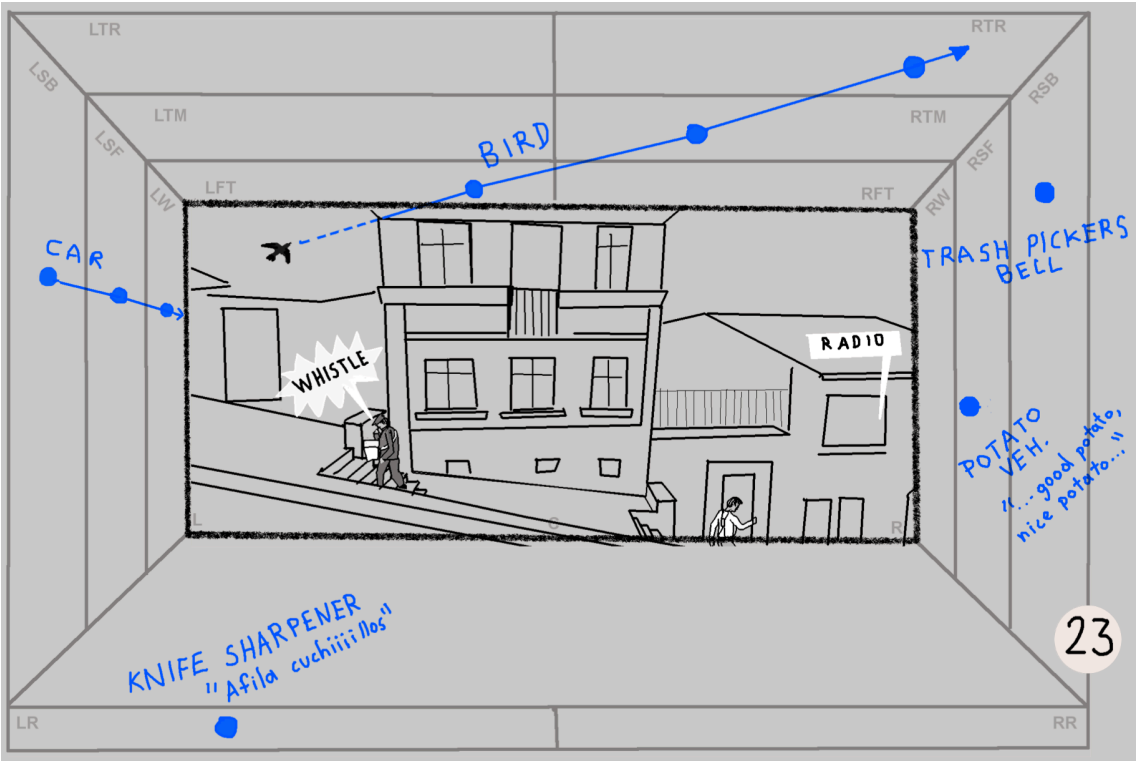


Figure 6.22 - Frame taken from the 3-D Storyboard of "Letters"

Conclusion

Sound design for film has historically been underestimated, a problem that even today, with the most advanced immersive audio technologies, continues happening. In terms of sound spatialisation, this thesis has observed that dialogue and other important narrative sounds occasionally move throughout the acoustical space; nevertheless, such an approach is still restricted due to, among other factors, visual composition and the conventional editing patterns that derive from classical continuity.

As a critique to such a screen-centric tradition, this chapter has intended to advance Cuarón's filmmaking approach by offering new mechanisms for planning the 'look' of a film considering the distribution of sounds within the diegesis; that is, the intention has been to propose a formal methodology for designing films in the basis of immersive continuity. Based on a new system of spatial continuity, referred here to as the 180-3D system, this thesis's original methodology challenges filmmakers to plan and design their films conceiving the camera not only as a unidirectional visual reference, but as an apparatus that also perceives sounds in the three dimensions.

The 3-D storyboard prototype presented in this chapter is a tool that filmmakers could use in the quest for visualising the ways in which sound and image are meant to interact at any given moment, preventing thereby the occurrence of visual treatments that could obstruct a coherent and realistic sound spatialisation. Unlike the traditional storyboarding approach, the method hear presented is capable of showing the physical position of offscreen sound elements in each specific shot, meaning that this tool can help filmmakers to take blocking and editorial decisions knowing what is happening outside the visible space. The original screenplay and storyboard (Appendixes 1 and 2), introduced within this chapter, are two documents that work as a proof of the actual applicability of the proposal.

Further empirical experimentation is needed in order to formalise the guidelines that the author proposes, but this academic introduction has aimed to shed light towards a broader adoption of the immersive continuity style.

FINAL DISCUSSION AND CONCLUSIONS

Sound technology and aesthetics have evolved in response to the need of a more realistic reproduction. With the arrival of Dolby Atmos and other immersive sonic platforms, filmmakers are now able to fabricate highly realistic three-dimensional soundtracks, replacing the space of exhibition with a fictional world in which sounds fluidly move in accordance with the position of characters and objects in the diegesis. Drawing on Green and Brock's (2000) transportation theory, this thesis argues that the impression of reality that 3-D sound can achieve may facilitate the spectator to be transported into the narrative world and thus experience affects such as persuasion and enjoyment, two fundamental motors for the film and media industries.

Alfonso Cuarón's work showcases a highly uncommon method that facilitates the deployment of the capabilities of present-day immersive audio tools. In conjunction with the cinema screen and its projector, the movie camera does a great job in representing the way we see in real life, something that is even more pronounced when stereoscopic technologies are used. But even though today's immersive sound technologies are capable of providing a highly transparent representation of the way we hear; filmmakers do not always succeed in emulating our aural sense. This is why Alfonso Cuarón's filmmaking approach is worth analysis, as it stands out as a highly uncommon case where the audio-visual representation of our world is fulfilled both visually and aurally.

It has been argued that his Bazinian long take-driven aesthetics, combined with his deployment of a lifelike 3-D soundtrack, gives shapes to a film style that this thesis denominates as immersive continuity. Initially evidenced in specific sequences of *Children of Men* (2006) and more concretely deployed in *Gravity* (2013), Cuarón consolidated his immersive film style in his most recent film, *Roma* (2018). Moreover, this thesis has observed that some of his stylistic traits are present in a number of recent films, suggesting that his approach to immersion has led to an ongoing stylistic turn.

Ultimately, based on Cuarón's work and the testimony of a number of key practitioners, this thesis has aimed to establish immersive continuity as a film style with its own aesthetic guidelines and filmmaking strategies, advancing Cuarón's proposal with new methods for

planning and designing films with 3-D sound. This was put in practice through the creation of an original screenplay and its corresponding storyboard. Personal conversations with Sergio Díaz, *Roma*'s sound designer and supervising sound editor, have served as a mechanism to test the validity of the new ideas and creative tools that the author proposes, something that will be discussed later in this concluding section.

But before moving forward, let us remind the reader the three research questions that were proposed at the beginning of this PhD project:

- What opportunities does 3D sound offer as different mode of address for filmmakers who wish to engage audiences in a more immersive fashion?

- What visual and sonic aesthetics are found in Cuarón's latest films, and have other filmmakers employed similar audio-visual traits?

- What creative strategies, technical operations and workflows have helped Cuarón and his collaborators to employ the full potential of immersive sound, and what additional screenwriting and preproduction strategies could facilitate the creation of films with three-dimensional sound design?

In order to discuss how this thesis has explored and answered each research question, this thesis's conclusion will be divided into three sections containing relevant final thoughts for each specific query, to which it will follow a short section discussing key avenues for future research. Let us begin with the theoretical aspect of the thesis, which has served as a justification for all the ideas and time spent in this PhD journey.

Theorising 3-D Sound and Audience Effects

Green and Brock's (2000) transportation theory suggests that an effective narrative has the potential of transporting the reader/listener/viewer, simply referred in cinema to as the spectator, into the narrative world. Grounded in empirical evidence, this theory claims that a transported individual may experience affects that go beyond enjoyment, such as belief and attitude change. One of the hypotheses within transportation theory indicates that the greater the realism the easier it is for the spectator to get transported into the narrative world, which in

cinema and other forms of audio-visual media can be due to the perceived illusion of being present in the space where the story takes place.

In the case of *Children of Men*, *Gravity* and *Roma*, Cuarón managed to manufacture highly realistic three-dimensional spaces, with techniques that stimulate our senses with the intention of making us feel as being present in an alternative reality. Following the premise that greater realism facilitates transportation and its effects on the spectator, it can be argued that his films are more transporting and affective than what those same narratives would have been if made using a different film style such as the conventional intensified continuity. Therefore, *Cuarón's* most recent films are in theory three examples where processes such as persuasion and enjoyment are boosted.

In terms of persuasion, all these films have a message that Cuarón aims to transmit to the audience. *Roma*, for instance, tells the story of an indigenous maid working for a wealthy family of white Mexicans; that is, the film approaches the socio-economical and racial gaps that have existed in Latin America since the Spanish colonisation. Almost every scene in the movie takes place in a space that contributes to the storytelling process, a characteristic that justifies a sound design approach that exploits the three dimensions of the world that surrounds the characters. Issues such as identity, diversity, race, gender, culture, and inequality are continually tackled in the movie, which on several occasions are perceived only sonically. Watching *Roma* within a Dolby Atmos theatrical infrastructure is thus the quintessential example of an experience that may facilitate the message to be effectively received by the transported spectator. The film has the power of placing us in the spatiotemporal continuum of the story, and thereby increasing the feeling of empathy towards the main character with whom we identify, a feeling that is enhanced by the realism of the events that we perceive through a three-dimensional audio-visual illusion. The film aims to achieve a positive change by means of promoting a message of social equality in favour of domestic workers in the neo-colonised third world. Grounded in cinematic realism, Cuarón succeeds in reviving the historical memory of a city, a nation, and a whole continent; a world which is objectively reconstructed around one of the individuals that are product of the system's social gap.

The power of persuasion of Cuarón's immersive continuity was also deployed in his film *Gravity*, which indirectly presents a message in favour of the conservation of our planet. This is evidenced, for instance, in the constant utilisation of the magnificent image of Earth seen in the background, which emphasises, not the Sun nor the stars, but the blue planet as the

centre of our universe. Moreover, the film centres its plot on the impossibility of life in the vacuum of outer space. The protagonist's quest is precisely to go back to Mother Earth, the only reachable place where life can succeed. In the vacuum, Dr Stone experiences a series of vicissitudes that rapidly turned her into the sole survivor of the mission, and it is only after overcoming a number of challenges that she is able to achieve the impossible and make her way back to the planet. Ultimately, the end sequence (01:20:25) magnifies the importance of Earth for the conservation of life. After Stone's spacecraft lands in the ocean, she craws out of the water and is welcomed by those very simple things that make our planet so special, such as the breeze, the singing birds, the little insects flying all around, all of which we perceive mainly through our ears. Floating in the water, she takes a few seconds to catch her breath while looking up to the sky, as a sort of spiritual moment that we experience vividly thanks to the power of 3-D sound. After being away in the terrifying darkness of outer space, Earth is granted with a more profound meaning. It is highlighted as a real, tangible, physical, delicate, unique place—the only 'land' we all have—a planet that, in *Children of Men*, Cuarón more explicitly warns us to sustain.

In *Children of Men*, an adaptation of P.D. James' 1992 novel *The Children of Men,* Cuarón presents a story that takes place in a dystopian near future. In 2027, humanity is facing extinction as humans have lost their ability to reproduce, and climate issues have caused an economic collapse forcing people to leave their own countries seeking asylum in the few nations with a functioning government. Through a sort of dystopian realism, *Children of Men* approaches global issues such as poverty, immigration, climate change, overpopulation, and so forth. The film's immersive continuity, employed in a number of scenes, helps to locate us there with the characters, having a vivid experience of a collapsed world. Such a realistic portrayal of a chaotic, authoritarian, and infertile society works therefore as a sort of warning that a time traveller brings for all of us. Essentially, the film claims that such a dystopian future could actually happen if humans do not take action and change their behaviours and attitude towards a more sustainable lifestyle.

Ultimately, in these and other films, filmmakers present a specific message (or at least a perspective of the world) for the audience to take away. It is plausible that such a message could achieve a more persuasive value when received within the absorbing darkness of a movie theatre, something that this thesis argues can be intensified when immersive sound is realistically deployed. Hence, immersive continuity is presented in this thesis as a cinematic

strategy that could increase the cinematic power of persuasion, which could in turn help to generate a positive impact on our society[43]. The films that have been analysed throughout this thesis have marked the way towards the consolidation of a highly affective film style. With immersive sound technologies such as Dolby Atmos gaining more momentum, the cinema is likely be more persuasive than ever before, a capability that could only be fully exploited if sound design is planned and executed in a three-dimensional manner.

But apart from increasing the persuasive power of cinema, this thesis also argues that 3-D sound can help to make a film more enjoyable. Entertainment is undoubtedly one of the primary reasons for people's desire to watch films. With Netflix and other VOD platforms gaining more and more subscribers each year, it is important to find new and more effective ways to attract audiences to go to the cinema, an exhibition space that even before the pandemic had been losing sales. Immersive continuity stands out as a purely cinematic film style, as in order to reproduce a three-dimensional soundtrack, a large number of surround (and overhead) speakers are needed. Hence, a film like *Roma*, for example, cannot produce the same affects when watched on a TV screen at home; such an immersive movie needs a cinematic sonic infrastructure in order to experience its creators' intention of transporting us into 1970's Mexico. As such, immersive continuity stands out as a feasible creative strategy that could tackle the decline that the cinema industry has been facing in the last two decades, a strategy that is grounded on the cinema's unique ability to reproduce the soundtrack through a large number of 'all-around' loudspeakers.

Ultimately, this thesis presents the hypothesis that 3-D sound may facilitate transportation to take place, and with this intensify the effects that a narrative can have on the spectator, namely persuasion, enjoyment and so forth. This, of course, needs proof, and hence stands out as an area for future research.

Cuarón's Style and a Potential Stylistic Turn

[43] Of course, this would mean as well that such an immersive film style could also be used as a weapon to persuade people with narratives that present negative behaviours or moral values.

Apart from the highly unusual deployment of 3-D sound, immersive continuity has been defined in this thesis as a film style characterised by the utilisation of long takes, a resource often complemented by wide-shots, deep-focus and camera movement. The recent appearance and consolidation of such an immersive film style can be attributed to Alfonso Cuarón, and the observation of similar audio-visual traits in filmmakers of the stature of Alejandro González Iñárritu, Bong Joon-ho, Darren Aronofsky, Edgar Wright, Leigh Whannell, Steve McQueen, Damien Chazelle, and others, is a clear manifestation of an ongoing stylistic turn. The mythical total cinema that Bazin once foresaw is now much closer, with immersive films like *Gravity* and *Roma* being the clearest but not the only examples. Although this thesis's findings are not enough to confirm Cuarón's influence on other filmmakers and on cinema as a whole, this research has certainly brought to light a certain stylistic similitude among his work and the work of a number of directors who play a key role in today's cinema. The historical evolution of the practice of film sound, which has been in constant dialogue with technological innovations, has reach to a point where filmmakers can decide to place dialogue and other important narrative sounds even in the farthest corner of the auditorium. This, which was highly unusual in previous eras, is today a fairly common exercise.

Even though intensified continuity and its fast cutting aesthetics are still the 'norm', the audio-visual characteristics of immersive continuity have gained more and more traction in current cinema, especially since Cuarón had the initiative of embracing such an experimental film style with great success in *Gravity*, one of the most commercial blockbusters of all time. Although it is true that a film set in outer space was the perfect opportunity for such a ground-breaking approach to take place, Cuarón proved himself that an intimate drama like *Roma* can also be fertile ground for the exploitation of his experimental film style. Moreover, the execution of similar aesthetic traits in action films, car chases, dramas, comedies, horror movies, and so forth, works as a proof for the actual applicability of immersive continuity in a range of genres.

As it is now, however, immersive continuity is most of the times limited to specific scenes or sequences, while the fast-cutting screen-centric aesthetics of intensified continuity predominate along most movies. Bong's latest films, for instance, present a number of moments with shots that are much longer than the average, which helped him and his sound team to pan all types of sounds in a three-dimensional manner; yet, his films also rely on the typical editing patterns derived from classical continuity, a limitation that is still found in

almost every film today. Moreover, it is worth mentioning that not all directors who choose to use long shots do also take the decision of using 3-D sound. For instance, Sam Mendes's *1917* (2020) plays as one continuous shot, yet its soundtrack conserves the screen-centric characteristics of previous eras, keeping dialogue and other narrative sounds in the centre of the screen. Woody Harrelson's *Lost in London* (2017) and Qasim Basir's *A Boy. A Girl. A Dream* (2018) are two other recent 'one-shot' movies in which 3-D sound was not deployed in despite of the opportunity that the long take presents. It can be argued, therefore, that immersive continuity is still passing through an experimental phase, and 3-D sound—specially dialogue panning—is still perceived as a risk that not all directors are willing to take.

Nevertheless, this thesis has observed that the treatment of dialogue seems to be evolving towards 3-D sound, with a number of recent examples containing moments where the spatial position of voices fit with the perspective conveyed by the camera's point-of-view. The technical affordances of Dolby Atmos and other immersive sonic platforms have facilitated such a spatialised treatment of sound to take place, with a number of individually-driven loudspeakers that offer full-frequency range distributed all throughout the space of exhibition. With such powerful audio tools now available, the only impediment for the consolidation of 3-D sound is thus tradition, not only in terms of the screen-centrality of the voice derived from the monophonic era, but more importantly, in terms of the conventional editing practices that have ruled the language of cinema for many years.

Continuity editing—namely those typical editing patterns that filmmakers use by default in all types of genres—is today the biggest obstacle to 3-D sound being more broadly adopted. A system like Dolby Atmos can thus be seen as a challenge for moving from classical or intensified continuity towards a 'less-chopped' film style. In fact, if such a radical change does not occur, it would be virtually impossible for 3-D sound to gain a broader adoption due to the confusion that a jarring sound can create. When film editing is done in the basis of shot/reverse-shot patters, Dolby Atmos can certainly serve as a commercial add-on, but not too much to create a more immersive experience. Proof of this are all those fast-cut Dolby Atmos releases that in fact sound as if they were mixed using the traditional 5.1 format. Certaliny, it can be argued that Cuarón has cast some sort of 'gravitational pull', but there are still many barriers to break in order to speak about an actual sound design evolution and thereby witnessing a major stylistic turn. Ultimately, it would be too exaggerated to argue that we are now facing a stylistic break, but, as this thesis has shown, there is indeed evidence of a desire

for a much less ‘chopped’ film style, which in turn can permit the occurrence of much less ‘tethered’ approaches to sound spatialisation.

Dolby Atmos and the Creation of 3-D Soundtracks

Apart from the analysis of Cuarón’s audio-visual aesthetics, the exploration of the creative practice behind the production of *Roma* has been helpful for determining the creative strategies, technical operation and workflows that were put in use to create what this thesis considers to be the most realistic audio work in cinema history. Years before, Cuarón was able to mix *Gravity* in Atmos, but *Roma* is the first film in which his approach to immersion was executed having the object-based functionality of Dolby’s latest sound system in mind from the very beginning. As acknowledged by Sergio Díaz, Skip Lievsay, Javier Quesada, and other sound collaborators, Cuarón had a clear plan of what he wanted to achieve sonically with the film. He wrote the script thinking of Atmos as a fundamental tool for connecting the audience with his nostalgic audio-visual tribute to Mexico City. And from there, he included sound-related activities all throughout the filmmaking workflow.

Certainly, the creative process behind the production of *Roma* is exemplary in terms of the implementation of Dolby Atmos and its object-based functionality into the workflows of sound recording and mixing. Nevertheless, it is worth mentioning that such an unconventional process took much longer than what it normally takes to create the soundtrack of a film. Hence, economically speaking, one could argue that Cuarón’s approach was less effective than the well-stablished postproduction-based sound design. Cuarón was able to allocate more time and money to sound because he had the liberty to do so as a producer of his most personal film. But this is not always the case. Hence, the question is: is it worth it to give sound more time and opportunities to contribute on the immersive experience of films? The answer to this question lies precisely in the theoretical justification discussed previously in this chapter. Creating a film with 3-D sound inevitably requires more hours of work and money than the typical screen-centric soundtrack, but it is plausible that such a sacrifice would help to enrich the cinematic experience, transforming a movie-watching spectacle into a more vivid, illusory and affective event.

In any case, knowing that increasing the budget will always be an obstacle for things to change, it would be of great use to discover new and more effective ways to achieve similar results at a more affordable cost. In the case of *Roma*, Cuarón aimed to recreate in great detail how Mexico City sounded in the 1970s. Therefore, all sound elements included in the soundtrack had to satisfy a very demanding threshold, resulting in a sound library full of original sounds recorded expressly for the movie. Every atmosphere in the film was designed using individually captured audio objects, such as the voices of background actors that were distributed independently around the 3-D space, or all the radio sounds that were attached to the vehicles that we see passing through the visual field. Undoubtedly, such a high level of perfectionism worked exquisitely on a period film with the characteristics of *Roma*, but things could be more flexible for other films if budget is a constraint.

There is not really the need of recording original sounds for each object that inhabits the diegesis, and as a matter of fact, all visual elements do not need to have its own specific sound in order to produce the illusion of three-dimensionality. If the intention is to endow the fictional setting with a lifelike corporality and thus enhance the sensation of presence in the space of the movie, all we need to do is attach discrete sound objects to the most important elements that inhabit the diegesis—such as the characters' voices, movements and footsteps, along with other key elements of the *mise-en-scène*—while background noises and atmospheres could most of the times be treated as compound and static beds. Moreover, if preserving the actual sound of the pro-filmic event is not of crucial importance, production sound recordists could then focus on capturing individualised dialogue tracks that will lend themselves for being manipulated as separate audio objects during the mix. All other sounds, as it is usually done, could be added later using conventional postproduction practices.

Nothing of this could work, however, if the film is not planned with the intention of deploying Atmos' full potential right from the screenwriting stage. Randy Thom and other sound practitioners interviewed throughout this thesis agree on the initiative that in order to take advantage of Dolby Atmos, filmmakers need to write and design their films with Atmos in mind, allowing sound ideas to intervene earlier in the creative process. Sound designer Sergio Díaz argues that having a script written with 3-D sound in mind definitely helps the sound designer as it "gives you more tools so you can perform better" (Díaz, personal interview, 2021, translated by author). He argues that the more specific the director is, "the more windows of possibilities there are to offer" (ibid). But beyond scriptwriting for sound,

Randy Thom (2019b) observes that a sound design storyboard would be useful, as it would challenge filmmakers to think about the ways sound will function in each scene and in the entire movie, giving sound the opportunity to be an influence on areas like direction, production design and cinematography.

After reading this thesis's original screenplay (Appendix 1) and revising the 3-D storyboard that accompanies it (Appendix 2), Díaz commented that he, as a sound designer, "already started to imagine things that will contribute more to it during the development process" (ibid). Díaz acknowledged the originality and potential applicability of this thesis's 3-D storyboard prototype, arguing that it is "a huge advance", because right from process of preproduction, it lets you see "the product as an immersive sound piece" (ibid). Díaz thus corroborates that the practical proposal presented in the previous chapter advances sound design practices to the next level. According to him, knowing what the director wants to achieve sonic wise is fundamental for getting the best results, hence having a script written for 3-D sound and, as he calls it, an "immersive storyboard" can help to execute sound design in a three-dimensional manner. Using these tools, filmmakers would be able to plan and communicate their audio-visual ideas to other members of the crew, visualising not only the 'look' of each specific shot, but also the distribution of sound elements within the three-dimensional exhibition space.

Together with the adjustment of screenwriting practices and the adoption of the 3-D storyboard as a method for film planning before production, this thesis also challenges filmmakers and film educators to evaluate the extent to which continuity editing and the 180-degree system keep on being appropriate references for scene blocking and editing. In film schools, students learn that it is good practice to cut from one shot to another as long as they respect visual continuity and do not cross the so-called 180-degree line. Once they become filmmakers, they follow such screen-driven guidelines and as a result produce films that contain a large number of scenes where cutting and the coverage-based shot/reverse/shot patterns are abundant, extending this way the establishment of such an old-fashioned visual language and thereby condemning sound design to its screen-centric confinement. The standardisation of Dolby Atmos calls for a radical change in the ways we manage film continuity, knowing that it is not only the visuals that need a consistent flow, but also the sound in terms of its distribution in the 3-D space.

This thesis's critique of the traditional system of film continuity is in fact a reaction to the practical failure of Dolby Atmos, a sound system that is being wasted every time a filmmaker uses it to mix the soundtrack of a fast-cut film. In other words, this PhD project has in part been an academic attempt to modify the virtually unchanged visual language that has governed film art since the silent era, a period in which Sergei Eisenstein argued that the essence of cinema lies not in images but in the relation that exists between them, namely in montage. It is totally understandable that montage was essential during a period of time when film sound did not exist, but its predominance has been maintained in despite of all the sound-related improvements that the cinema has faced since the 1920s. Editing is indeed an important aspect of cinema, but it is not its essence. Films can be made without cuts and the essence of cinema will still be maintained, which is found not in the relation that exists between images, but in the relation that exists between the narrative, the image and the sound.

Films like Iñárritu's *Birdman* and Sam Mendes' *1917* are a contemporary proof of mainstream movies that are presented without cuts, and yet no one can say that these films are not real cinema. As a matter of fact, Alfred Hitchcock had already made that clear with his 'single-take' film *Rope* (1948) in the classical Hollywood era. But there is no need to be as bold as these 'one-shot' filmmakers in order to offer sound the opportunity of behaving in a proper three-dimensional manner; all what is needed is a 'sound-conscious' treatment of the visuals; that is, in the quest of achieving a more transporting film, filmmakers should take cinematographic and editorial decisions considering the distribution of sounds in each shot and within their combination. And ultimately, they should consider employing the audio-visual characteristics of immersive continuity, if not for entire films, at least in those scenes or sequences that lend themselves to a more immersive approach. No matter the genre, immersive continuity has the power to transport us into the spatiotemporal continuum of the story, an attribute that can make cinema and specially the theatrical experience more attractive.

The 180-3D system proposed in the previous chapter aims to support such an approach by suggesting a special treatment for close-ups and all other shots that frame a subject or subjects while leaving others offscreen. Teaching filmmaking using this alternative continuity system as a guideline could be a good starting point for the consolidation of a new immersive cinema of the future, meaning that academia can be the bridge for immersive continuity to penetrate in the mind of a new generation of filmmakers. If right from film school, filmmakers start to realise that the conventional cutting patterns derived from classical continuity impede

the deployment of 3-D sound and thus the immersive experience of a film, it is more likely that more of them will join Cuarón and his creative collaborators on the journey towards the mythical total cinema. Cuarón's work has undoubtedly made an enormous contribution to the rise of a new kind of cinema, and this thesis's academic effort has aimed to transform his initial spark into a more luminous filmmaking revolution.

Future Work

This PhD has intended to formally define the immersive continuity style in terms of a series of theoretical, aesthetic and practical connotations. Theory wise, the author's hypothesis connects 3-D sound with the concept of presence, and suggests that its lifelike characteristics can facilitate the occurrence of psychological processes such as transportation and identification, which in turn can produce a series of effects on the spectator. In order to validate and give more relevance to the present contribution, it would be important to start looking for empirical evidences that could prove such a presumption, a task that could be accomplished by means of the execution of experiments with actual audiences in controlled audio-visual environments.

In terms of the study of sound technology and its relation to film aesthetics, the evidence found throughout this thesis suggests that there is an ongoing stylistic evolution towards a long take-driven cinema that in turn will facilitate a broader adoption of 3-D sound. In order to corroborate such a prediction, further film observation and analysis is needed, which should aim to find the aesthetic traits of immersive continuity in both the mainstream and the independent world cinema sectors. Special attention should be put in the films that include a Dolby Atmos soundtrack, but more traditional multi-channel formats shall also be taken into consideration. Even though the technical affordances of Dolby Atmos favour the occurrence of more spatialised sound design approaches, this sound platform is by now being used primarily on blockbusters that are less likely to break with tradition. Therefore, future research should certainly analyse the image-sound relations in upcoming Dolby Atmos releases, but the body of investigation could comprise non-Atmos films as well.

Finally, future research is needed in order to determine the most effective methods for the practical deployment of 3-D sound. Adding to the insights obtained after studying the

approach followed by Cuarón and his collaborators, studying the creative practice behind the production of other immersive films will help to determine other viable alternatives. Further conversations with filmmakers and sound designers would be useful for strengthening the ideas here presented. In terms of filmmaking education, research scholars could embrace this thesis's continuity system and pre-production methodology as a training guide for the creation of original film content, which may lead to useful student reflections and insights regarding their effectiveness and applicability. Such a practice-based methodology could as well be adopted by filmmakers that do research in the basis of original creative practice.

APPENDIX 1

Letters: The Screenplay

EXT. “AMERICA” NEIGHBORHOOD, RIOFRIO STREET - DAY

An uphill cobblestone pavement street.

TITLES FADE IN:

SUPERIMPOSE:

“Quito, August 1972.”

TITLES FADE OUT.

In the background, we see a NEWSPAPER VENDOR passing by chanting his traditional selling call.

NEWSPAPER VENDOR
(Shouting)
Mercioooooo!

In the distance, the characteristic sound of a VESPA motorcycle is heard approaching from behind. ARTURO SUASNAVAS, a postman (45), enters INTO FRAME as the sound of his VESPA is heard passing from rear to front. Arturo parks the bike and walks downhill while WHISTLING to announce his arrival.

INT. PAREDES' RESIDENCE - DAY

From inside the house, the postman's WHISTLE is heard by the beautiful GARDENIA PAREDES (17), who quickly stands up from a chair in her -

BEDROOM -

where she is looking at a young man’s picture. CAMERA TRACKS WITH her as she rapidly goes out from her room and runs through the -

UPPER FLOOR CORRIDOR -

towards the stairs, passing by a loud grandmother CLOCK that is hanging in the left wall. As Gardenia reaches the -

LIVINGROOM -

she passes by the inquisitive gaze of her father, EDMUNDO PAREDES (54), who while sitting in a sofa listens to RADIO Tarqui and reads “El Comercio” diary.

EDMUNDO
(shouting angrily)
Where are you going Gardenia?

His strong and intimidating VOICE is heard behind us as we go -

DOWNSTAIRS -

with Gardenia. The sound of the RADIO filters from upstairs and slowly disappears as Gardenia gets to the -

GROUND FLOOR CORRIDOR -

where she finally opens the main door.

EXT. “AMERICA” NEIGHBORHOOD, RIOFRIO STREET - DAY

As Gardenia opens the door, Edmundo is seen examining the event from the living room's window in the upper floor. As he opens the window, the sound of the living room’s RADIO filters from above and is heard by Gardenia, who looks above and sees her father. Edmundo dissimulates.

Gardenia awaits at the door of their purple house, while a couple of curious kids go out running to see the postman's new bike.

GARDENIA
Good morning, still no letter for me?

ARTURO
Good morning miss Gardenia! Actually, this time there is a letter for you.

GARDENIA
(smiles very exited)
Really? Let's see!

Arturo searches for The Letter in his bag while he walks towards Gardenia. He finds the letter and gives it to the young woman.

GARDENIA (CONT'D)
It's him, finally!

ARTURO
Sorry, what was that?

GARDENIA
Thank you very much. I'll see you next time!

Gardenia enters the house.

INT. PAREDES' RESIDENCE - DAY

LIVING ROOM -

The House's main DOOR being shut is heard behind us from a distance. From Edmundo's position, we see Edmundo looking at Arturo as he walks uphill to deliver other letters. Edmundo CLOSES the living room's window and walks towards the stairs. He stops as Gardenia passes by with the letter in her hands. Edmundo makes sure she enters into her bedroom before going -

DOWNSTAIRS -

towards the house's main door. The postman's WHISTLE is heard once again as Edmundo SILENTLY opens the main door.

EXT. “AMERICA” NEIGHBORHOOD, RIOFRIO STREET - DAY

CAMERA TRACKS WITH Edmundo as he walks uphill aiming to speak with Arturo. A POTATO SELLER passes by on a TRUCK chanting his loud SELLING CALL through a megaphone. The selling call continues from a distance as the truck goes down the whole Riofrio Street.

POTATO SELLER
(over a megaphone)
“Chola” potato, good potato, nice potato...!

Edmundo reaches Arturo from behind and greets him touching his back twice in sign of affection.

EDMUNDO
How are you Arturo? How is your day going?

ARTURO
Good morning, Don Edmundo. Not too bad, and you? I’m afraid I don’t have no letter for you today.

Arturo sees that one of the kids is sitting on his bike.

ARTURO (CONT'D)
(Shouting)
Hey you! Get off the bike!

The kids LAUGH while running down the street.

EDMUNDO
They love your new bike!
(Smiles)
Hmm, look, I wanted to ask you a favor.

ARTURO
Sure, how can I be of help?

Edmundo stops talking for a moment as a person's FOOTSTEPS are heard approaching from our right. He greets to Edmundo nodding his head and continues walking uphill towards his house. The newspaper vendor SELLING CALL is heard once again in the distance, together with the voice of a SCRAP RECYCLER coming out of a distant megaphone.

SCRAP RECYCLER
(over a megaphone)
We are buying old cookers, old tins, old bottles, all kinds of old metals...!

Edmundo makes sure that nobody else is nearby and continues.

EDMUNDO
You know, my daughter Gardenia was seeing this guy, Cesar Torres, do you know him?

ARTURO
Yeah... I think I know him. I'm afraid he went to work abroad, didn't he?

EDMUNDO
Yes, he did (pause). I saw you just delivered a letter for my daughter. It was from him, wasn’t it?

Arturo nods his head.

EDMUNDO (CONT'D)
You know, my daughter has become very stubborn about that relationship. And that is worrying me because he is not the type of man I want for her.
(MORE)

EDMUNDO (CONT'D)
Actually, my son told me that he saw this guy smoking marijuana or something the other day. It's simply not good for the health and future of Gardenia to keep waiting for a man like that. You know what I mean?

ARTURO
I see...

EDMUNDO
So...

Edmundo stops once again as the TRASH PICKERS arrive to the neighborhood jingling their BELL.

EDMUNDO (CONT'D)
Damn it! I forgot about the trash.

Edmundo hesitates on what to do while a few neighbors go out carrying their metal trash bins, whose content the trash pickers deposits in a big metal container before carelessly dropping the NOISY METAL BINS. After a few seconds, a distant CHURCH BELL announces that it is 10am.

ARTURO
Excuse me Don Edmundo but I need to leave. It’s late and I still have lots of work to do.

EDMUNDO
(talking softly)
No, no, wait. What I wanted to ask you is to stop delivering any letter that comes from that Cesar guy.

ARTURO
Stop delivering them?

EDMUNDO
Yes, yes.

ARTURO
Sorry Don Edmundo but that's against the law, I could lose my job because of that. I'm afraid...

EDMUNDO
(interrupting)
Here is a little something for your services.

We see Edmundo secretly handling money to the postman. The sonic atmosphere gets FILTERED as we begin to hear from Arturo's somehow altered point-of-audition. His HEARTBEAT and BREATHING put in evidence his nervousness. An echoing VOICE coming out of his own mind is heard.

MAN'S VOICE
(angrily)
You better have the money by the end of the week, you hear me?

We see Arturo's right hand receiving the money and hiding it. He looks around and remains quiet for a few seconds. The street's SONIC ATMOSPHERE returns to normal as the voice of a NEWS ANCHOR is heard coming out from a CAR's RADIO as the vehicle stops in the corner.

NEWS ANCHOR
(over a distant radio)
The Petroleum Estate Corporation of Ecuador has just exported its first load. Yet, acts of corruption within the government are now the target of investigation, as President Rodríguez Lara...

The CAR leaves and the RADIO sound vanishes.

EDMUNDO
Please, this must stay between the two of us.

Arturo remains silent. The trash pickers leave, and Edmundo goes downhill and shouts.

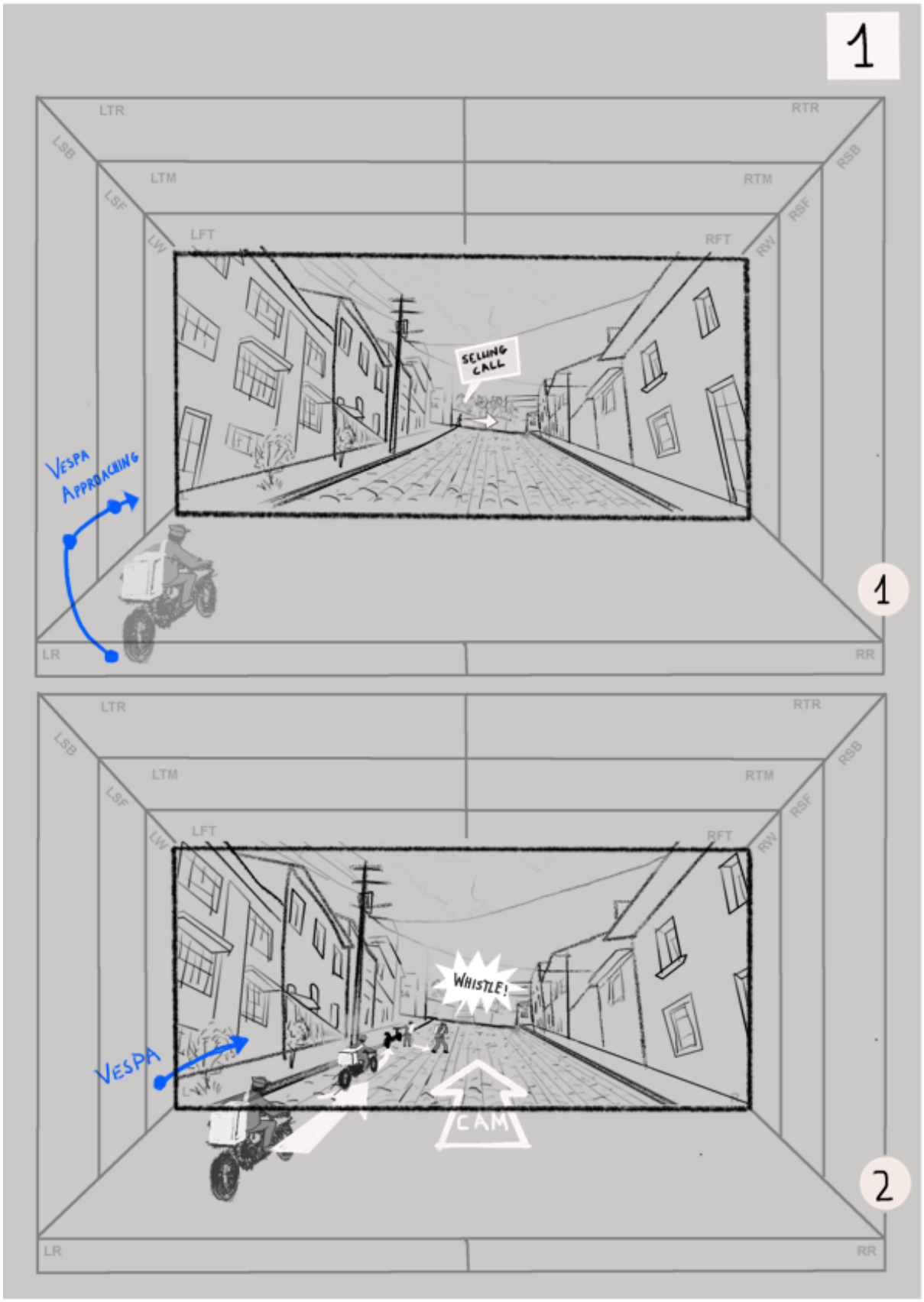
EDMUNDO (CONT'D)
(shouting)
Wait please!

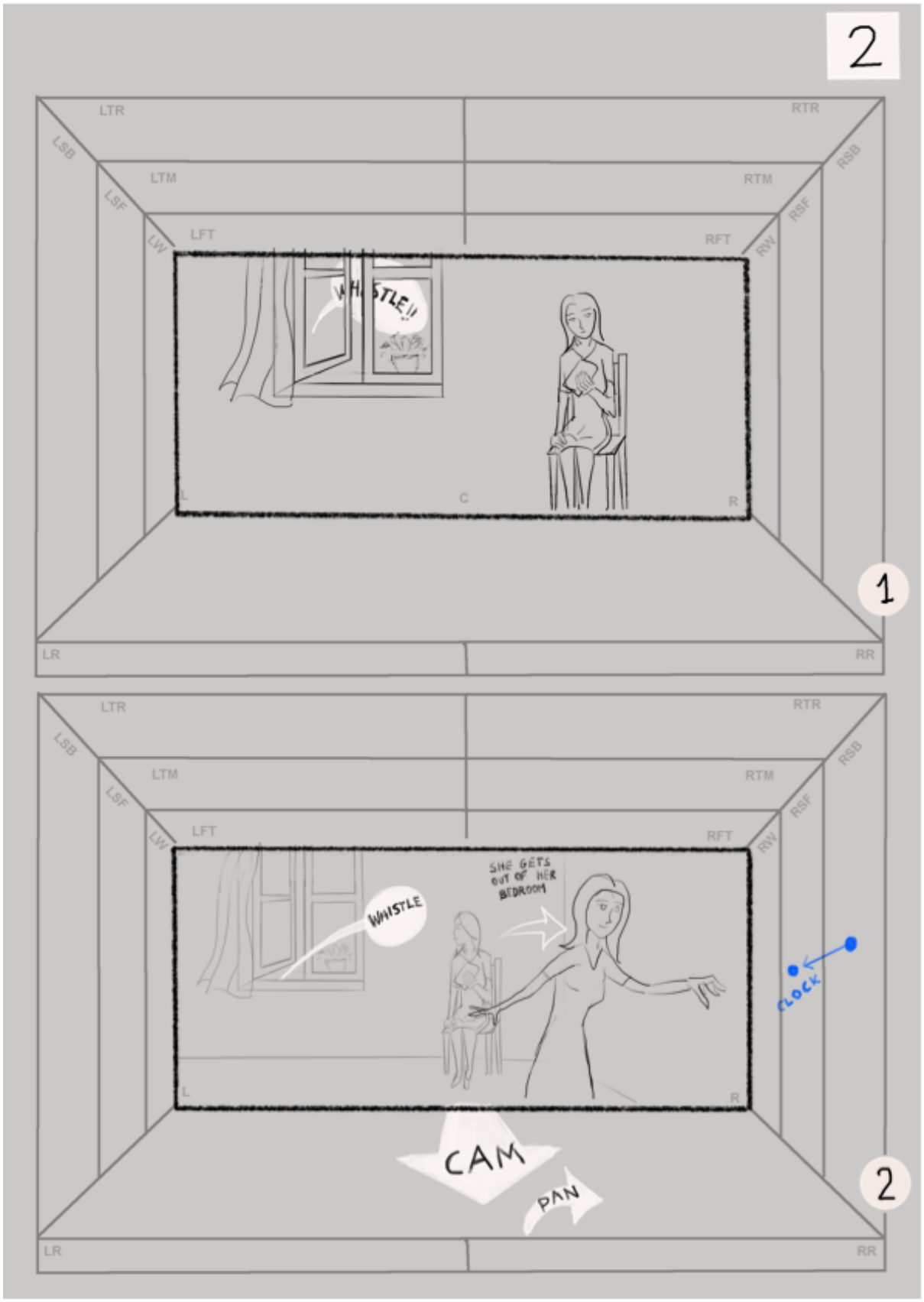
The three-dimensional ATMOSPHERE of the neighbourhood surrounds us with sounds that come from all directions.

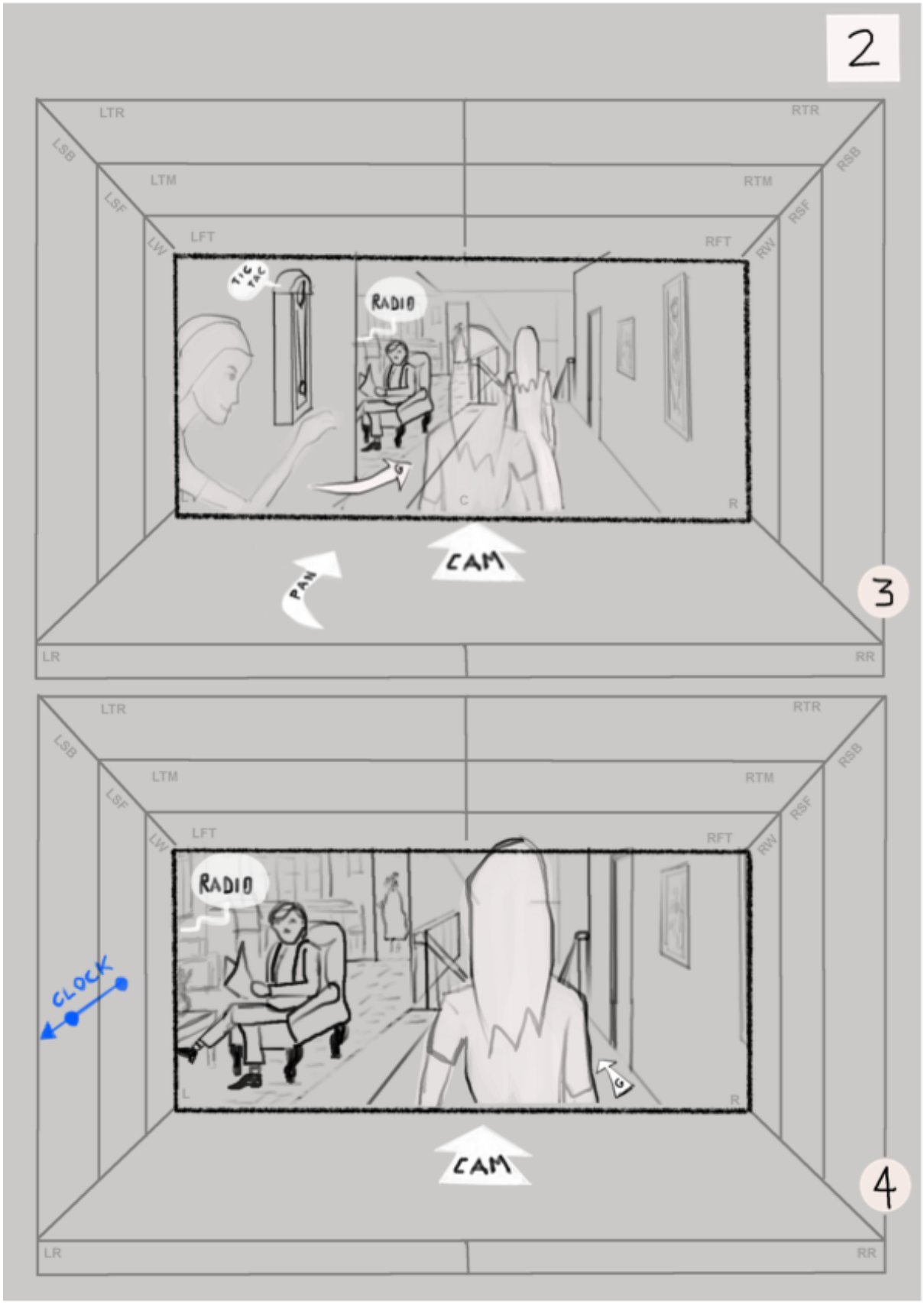
A few CARS passing by, VENDORS in the distance, the trash picker jingling his BELLS in the next corner.

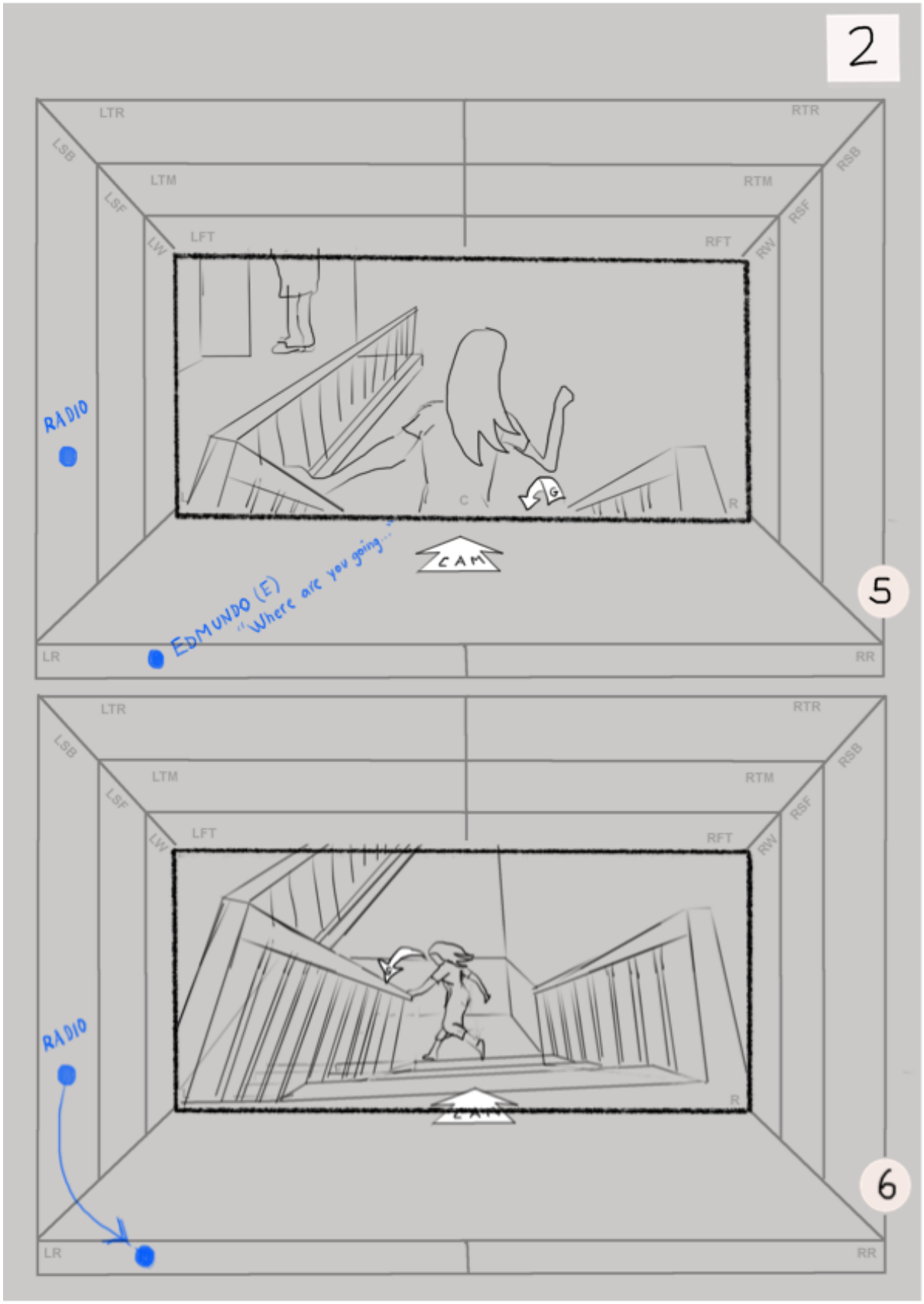
APPENDIX 2

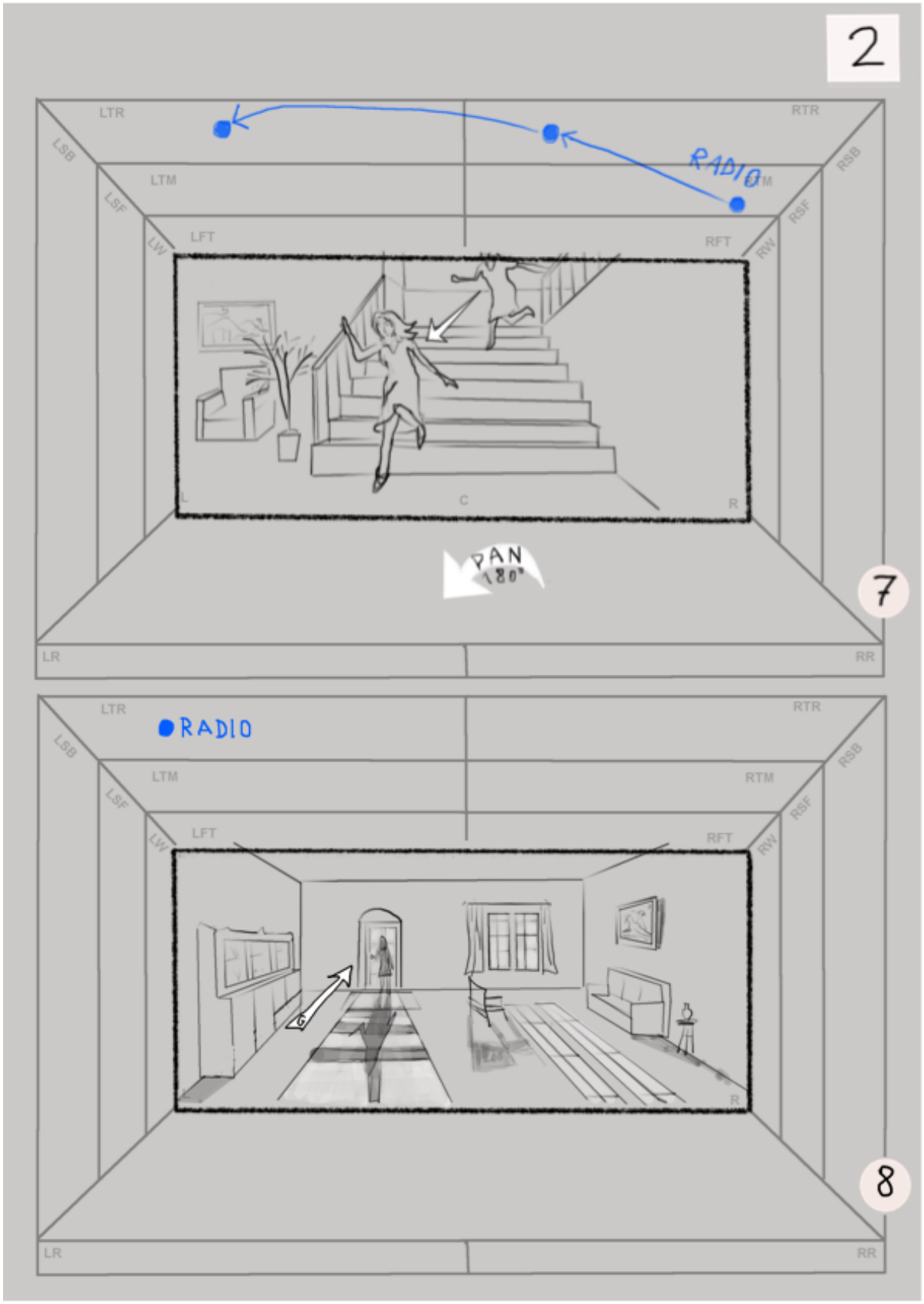
Letters: The 3-D Storyboard

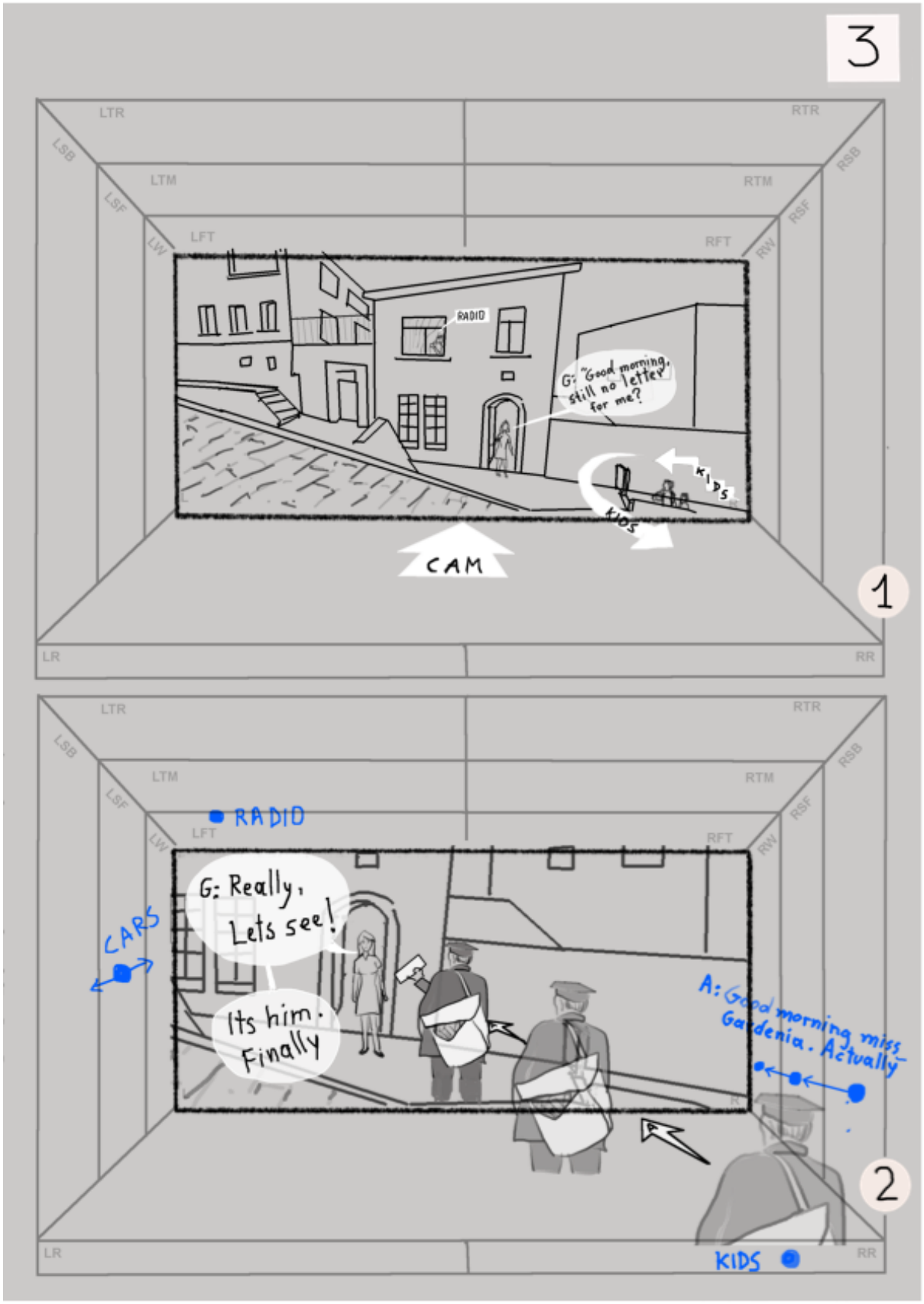

1
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
SELLING CALL
VESPA APPROACHING
1
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
WHISTLE!
VESPA
CAM
2
LR
RR

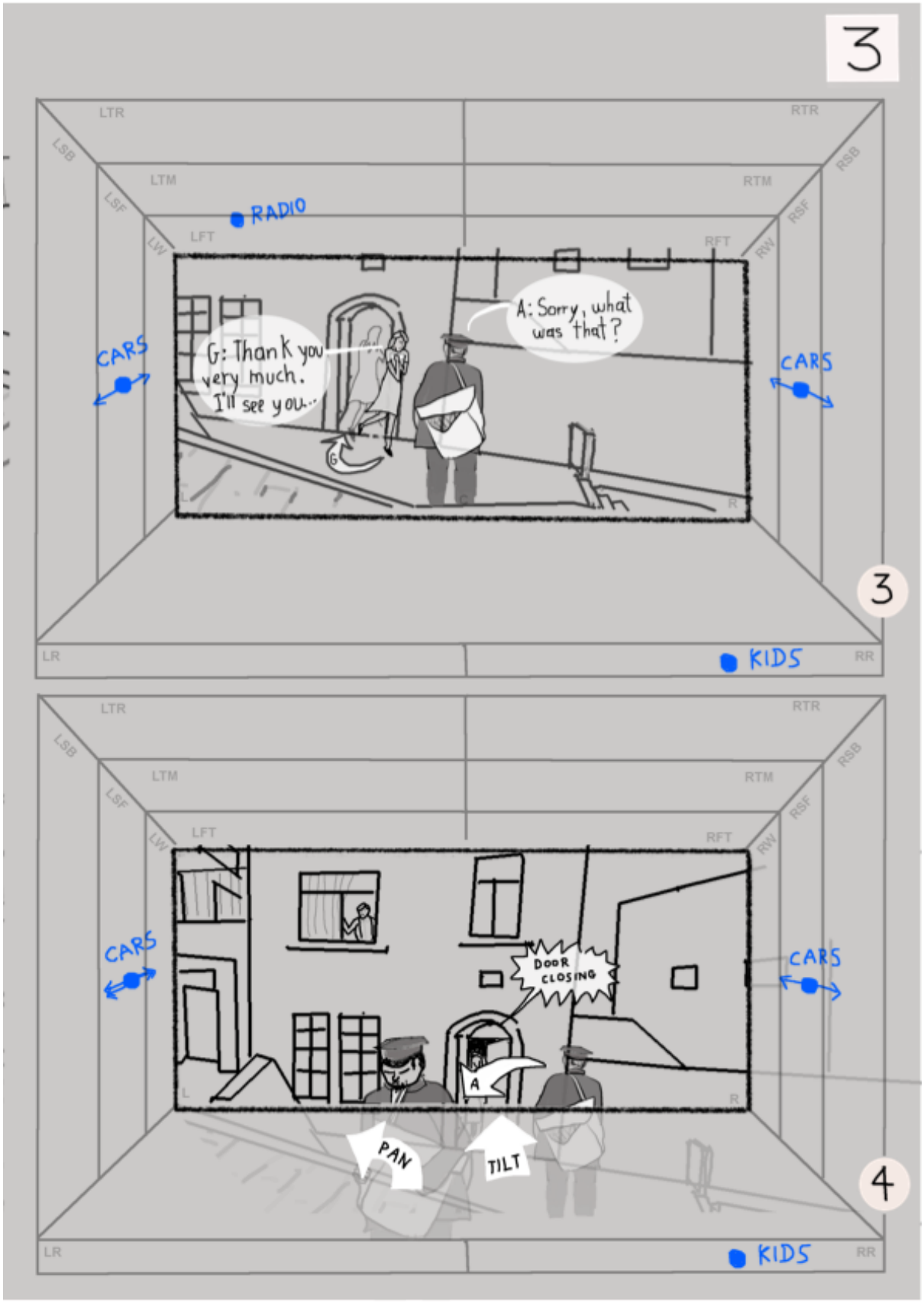

2
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
WHISTLE!!
L
C
R
1
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
WHISTLE
SHE GETS OUT OF HER BEDROOM
CLOCK
L
R
CAM
PAN
2
LR
RR

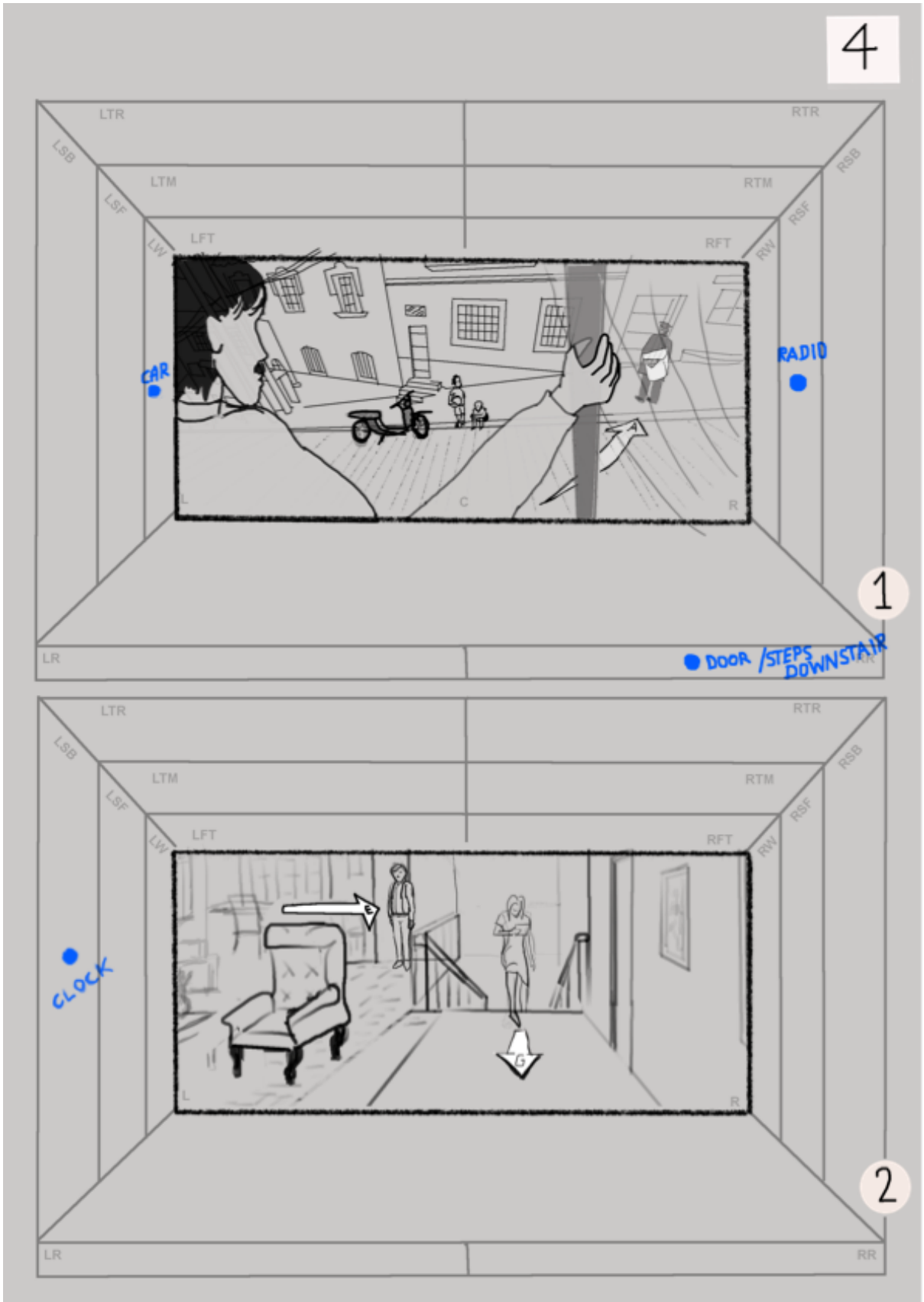

2
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
TIC TAC
RADIO
L
C
R
CAM
PAN
3
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
RADIO
CLOCK
L
R
CAM
4
LR
RR

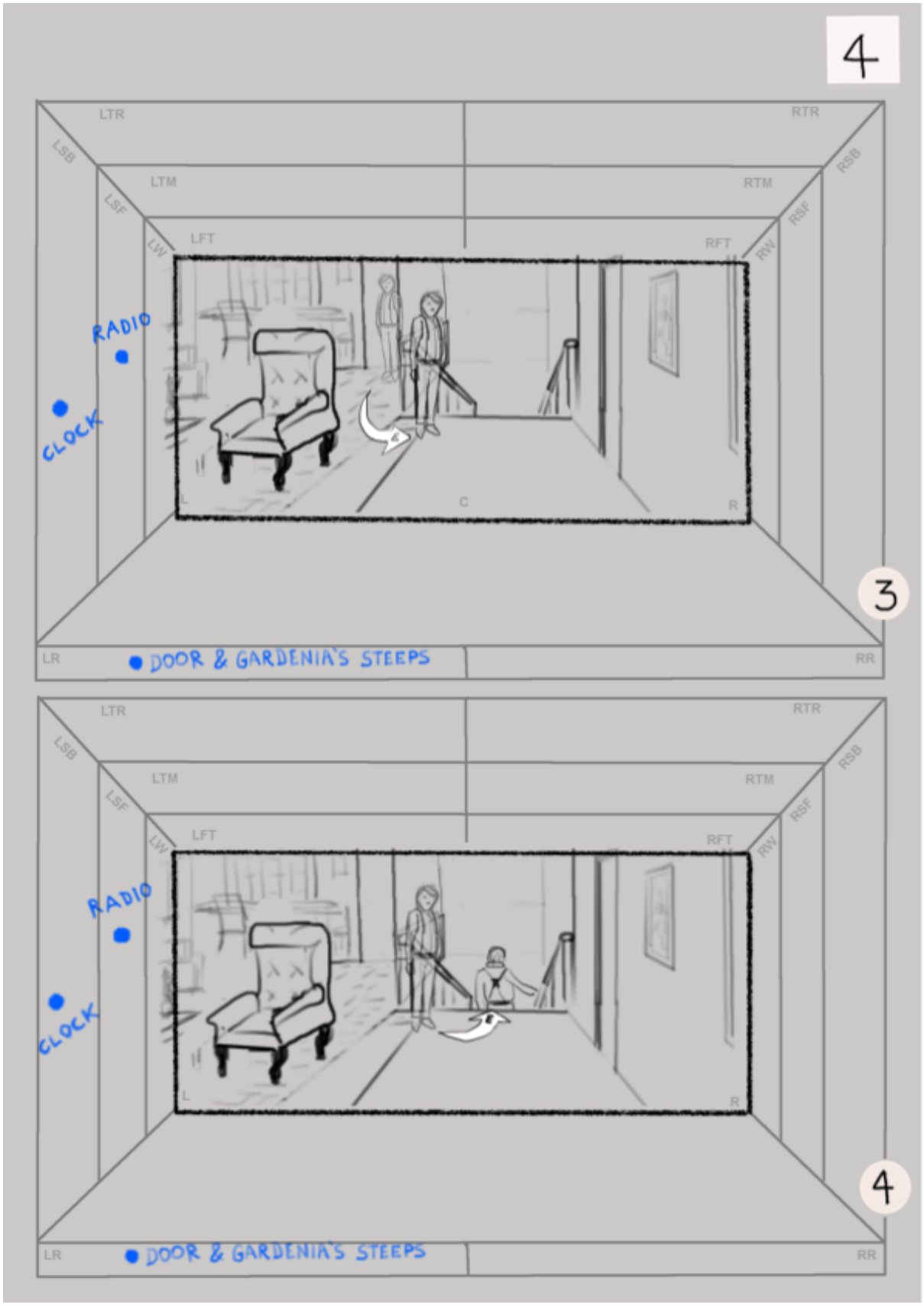

2
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
RADIO
CAM
EDMUNDO (E)
"Where are you going..."
5
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
RADIO
CAM
6
LR
RR

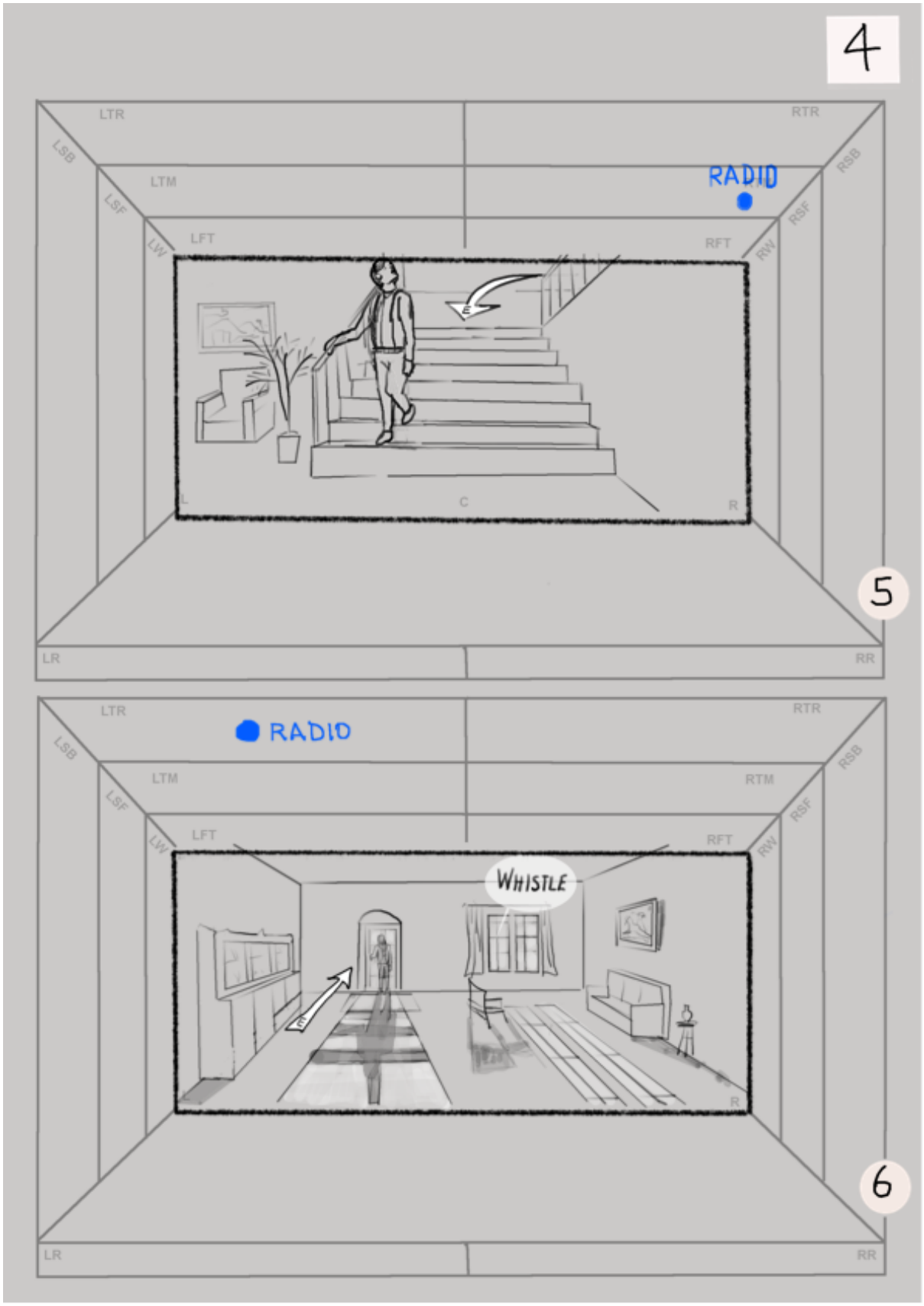

2
LTR
RTR
LSB
RSB
LTM
RADIO
RTM
LSF
RSF
LW
LFT
RFT
RW
L
C
R
PAN
180°
7
LR
RR
LTR
RADIO
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
R
8
LR
RR

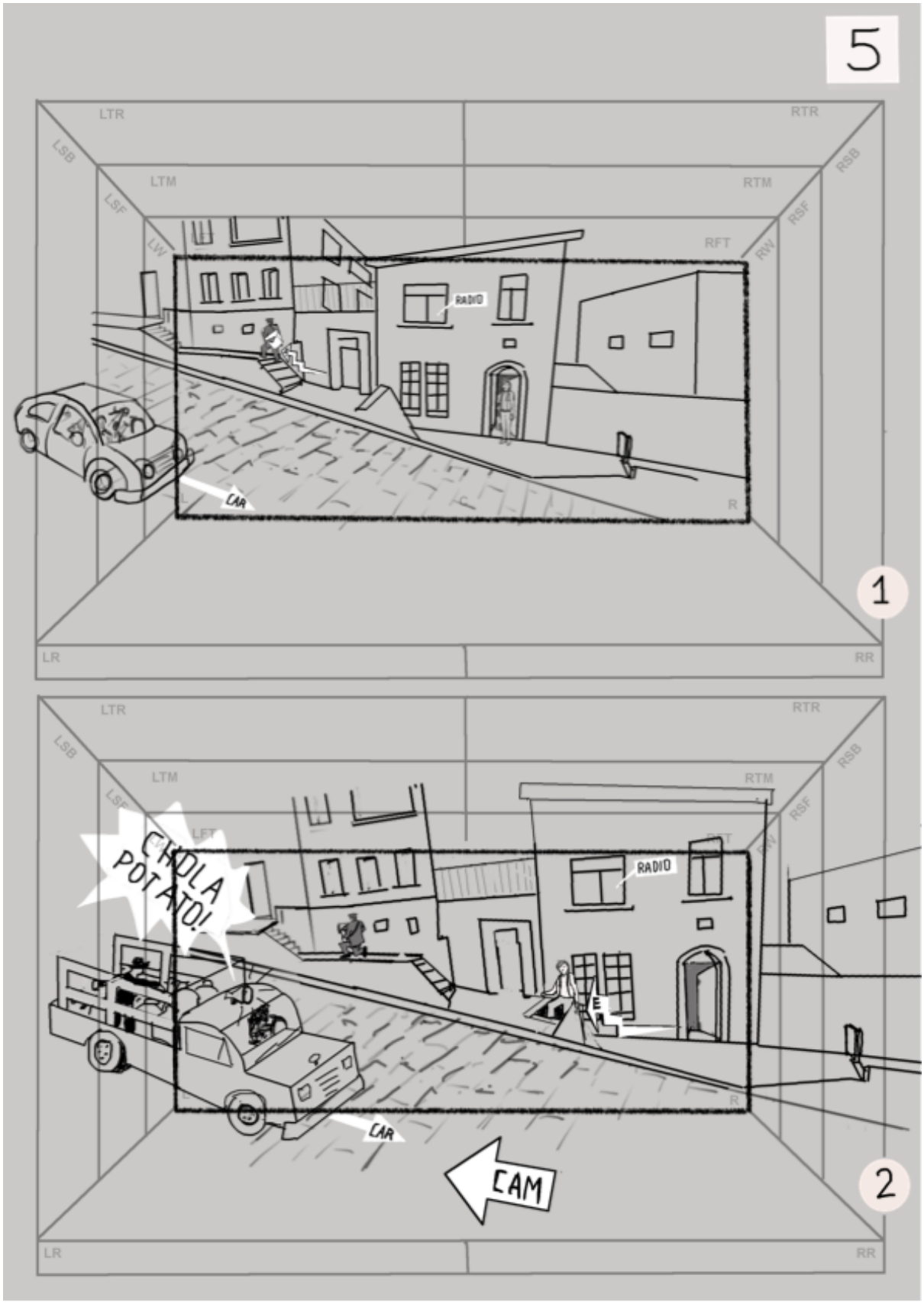

3
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
RADIO
G: "Good morning, still no letter for me?
KIDS
KIDS
CAM
1
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
RADIO
LFT
RFT
LW
RW
G: Really, Lets see!
CARS
Its him. Finally
A: Good morning miss Gardenia. Actually
2
LR
KIDS
RR

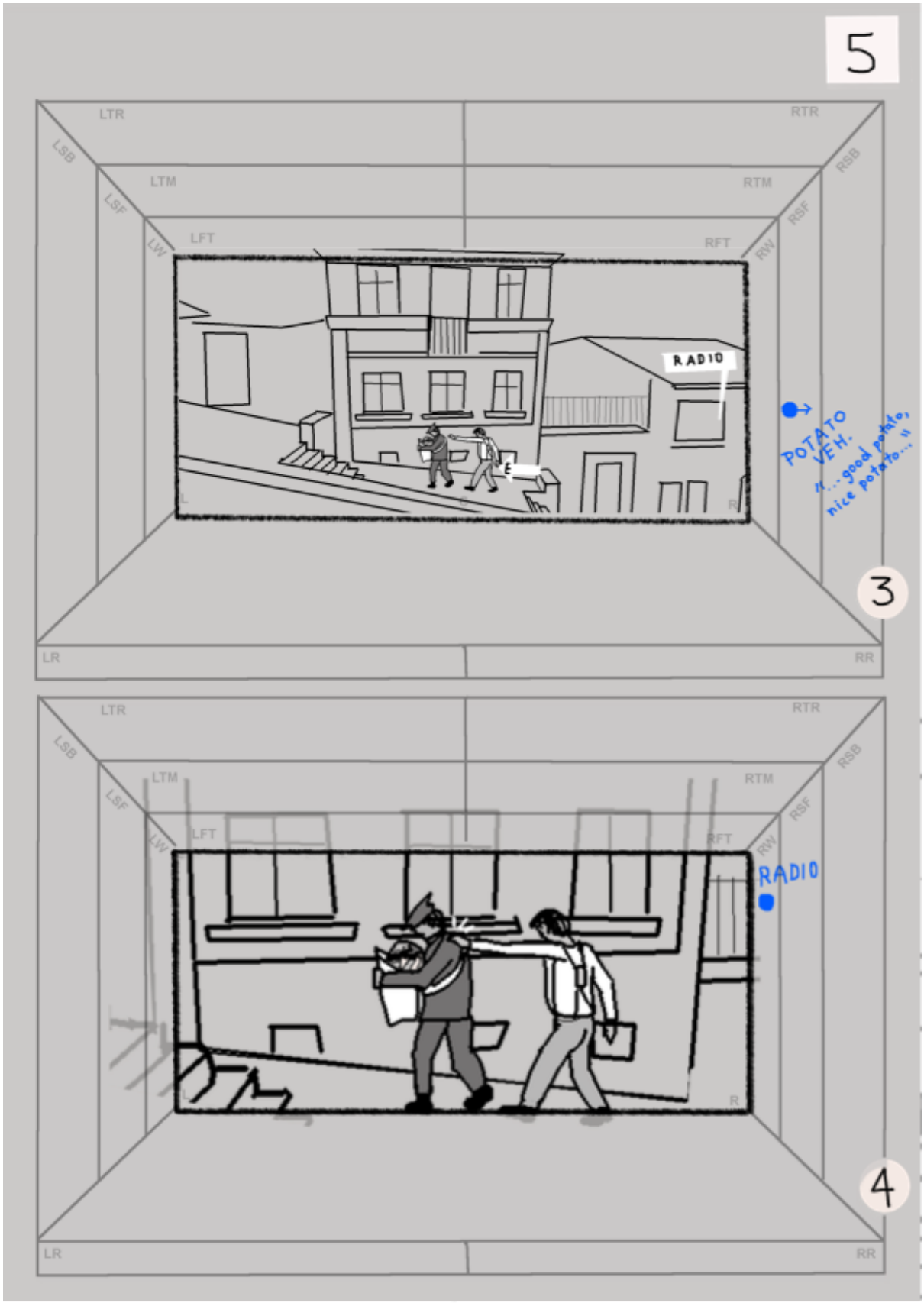

3
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
RADIO
LFT
RFT
LW
RW
A: Sorry, what was that?
G: Thank you very much. I'll see you...
CARS
CARS
3
LR
KIDS
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
CARS
DOOR CLOSING
CARS
A
PAN
TILT
4
LR
KIDS
RR


4
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
CAR
RADIO
L
C
R
1
LR
RR
DOOR /STEPS DOWNSTAIR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
CLOCK
L
R
2
LR
RR

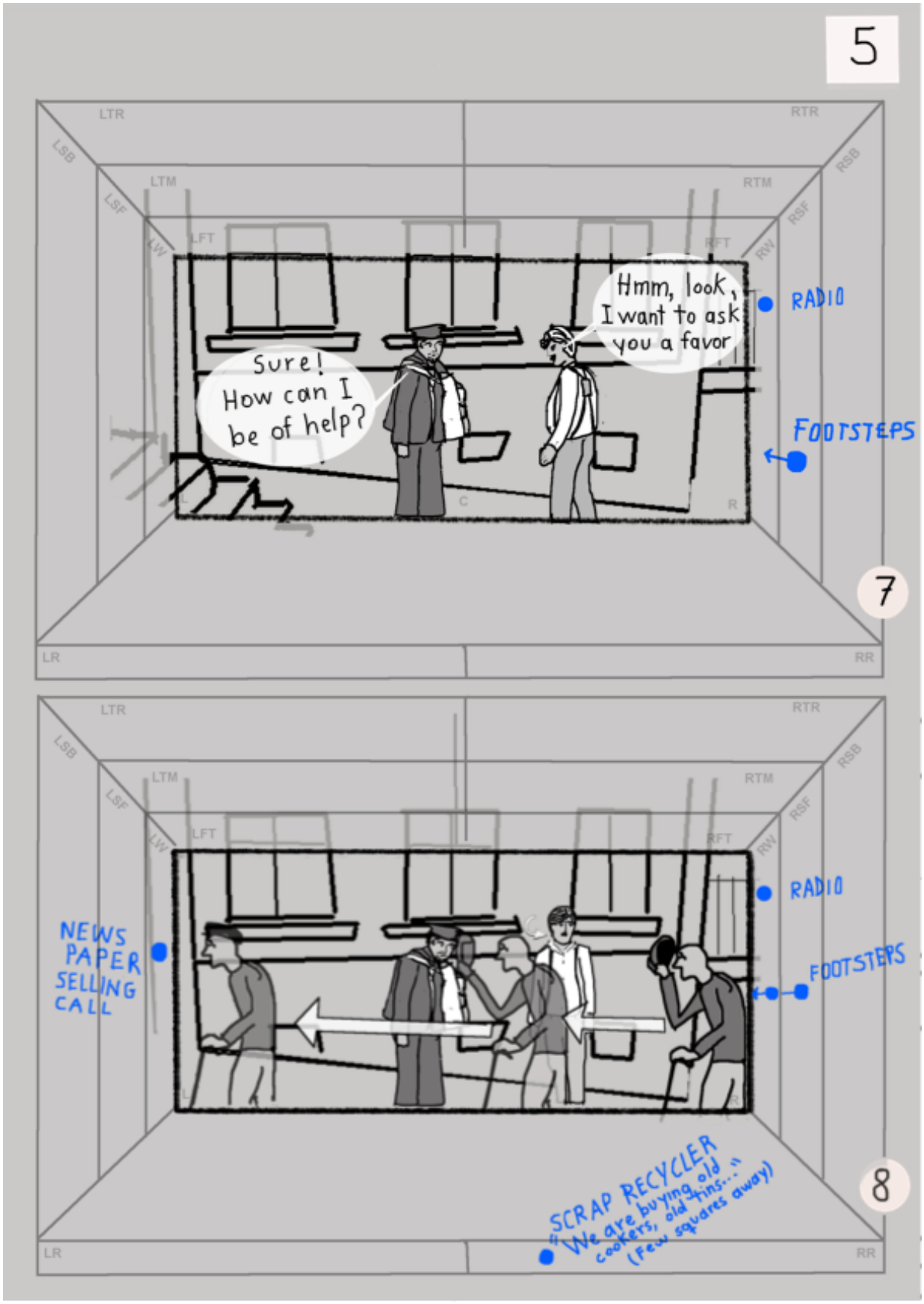

4
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
RADIO
CLOCK
L
C
R
3
LR
DOOR & GARDENIA'S STEEPS
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
RADIO
CLOCK
L
R
4
LR
DOOR & GARDENIA'S STEEPS
RR

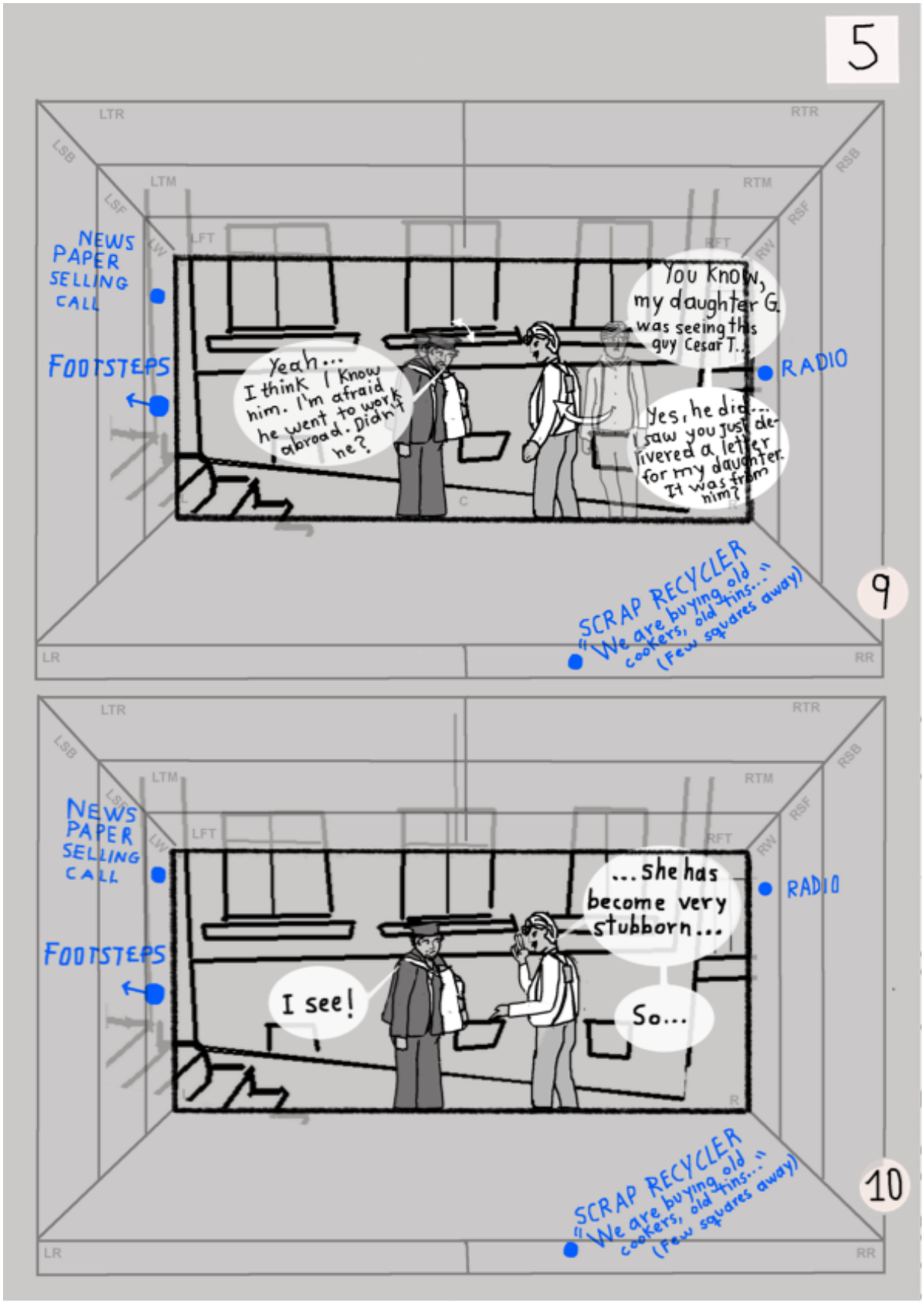

4
LTR
RTR
LSB
RSB
LTM
RADIO
LSF
RSF
LW
LFT
RFT
RW
L
C
R
5
LR
RR
LTR
RTR
RADIO
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
WHISTLE
R
6
LR
RR

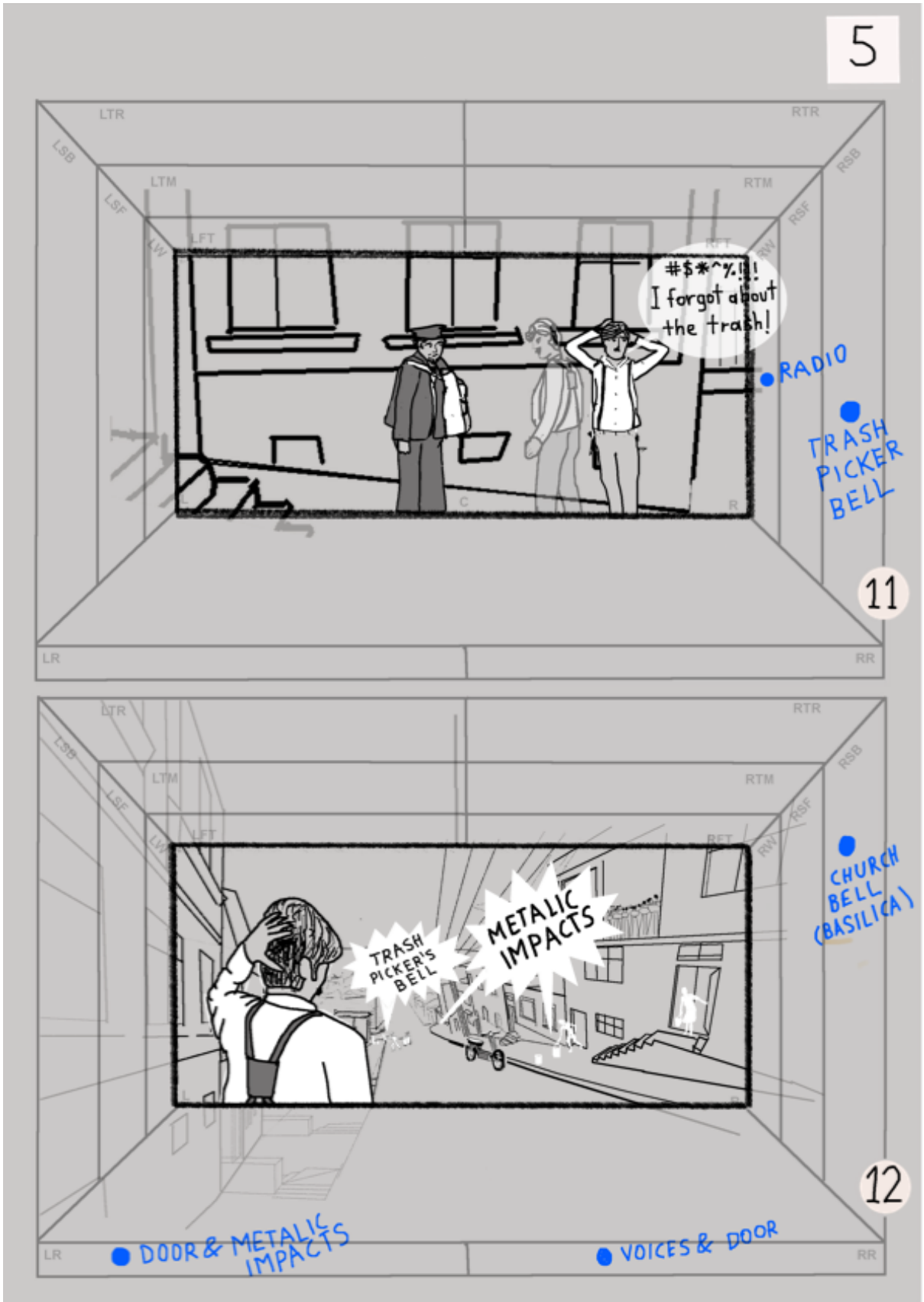

5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
RW
RFT
RADIO
CAR
R
1
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
CHOLA
POTATO!
RADIO
CAR
R
CAM
2
LR
RR

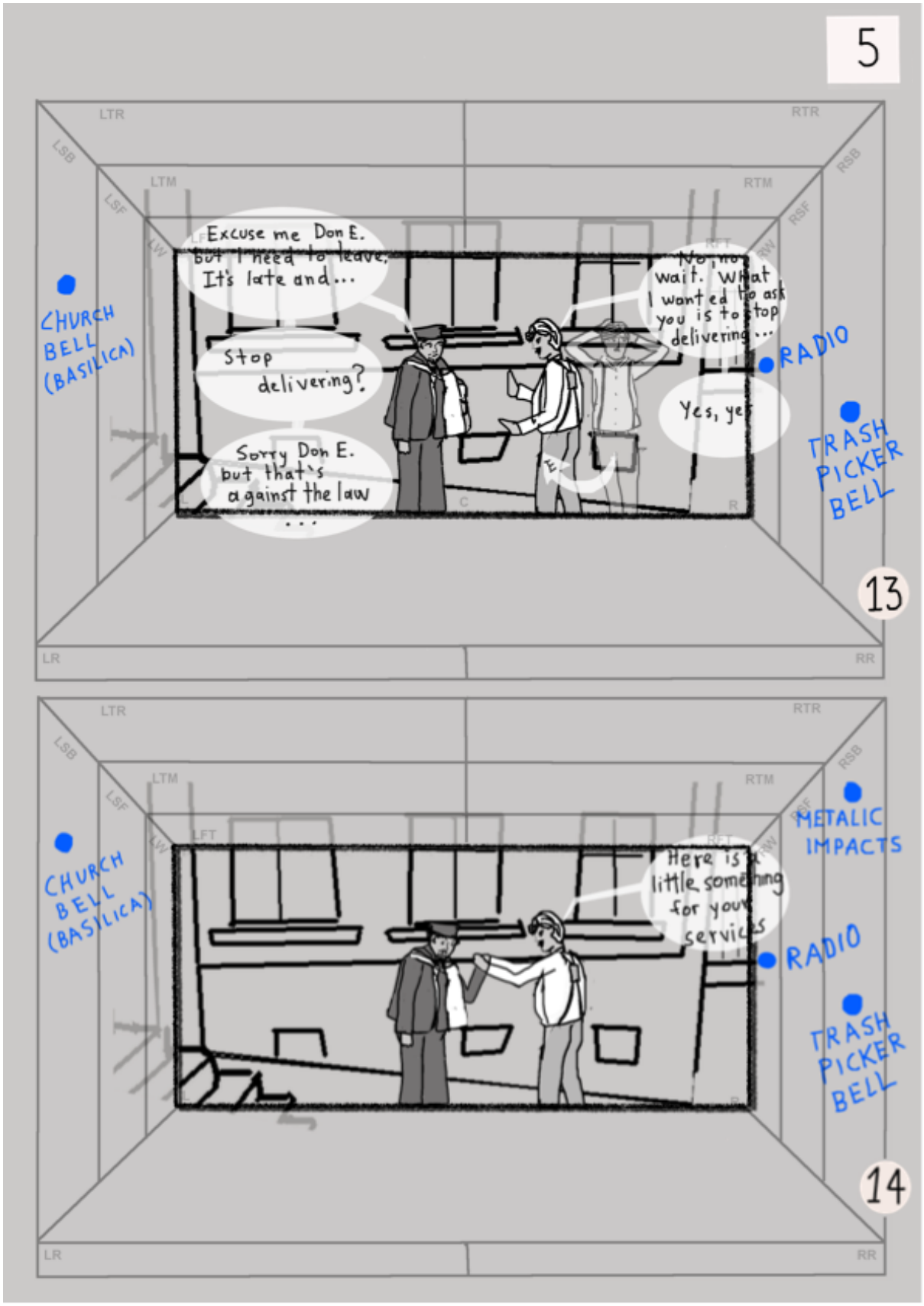

5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
RADIO
POTATO VEH.
"...good potato, nice potato..."
3
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
RADIO
4
LR
RR

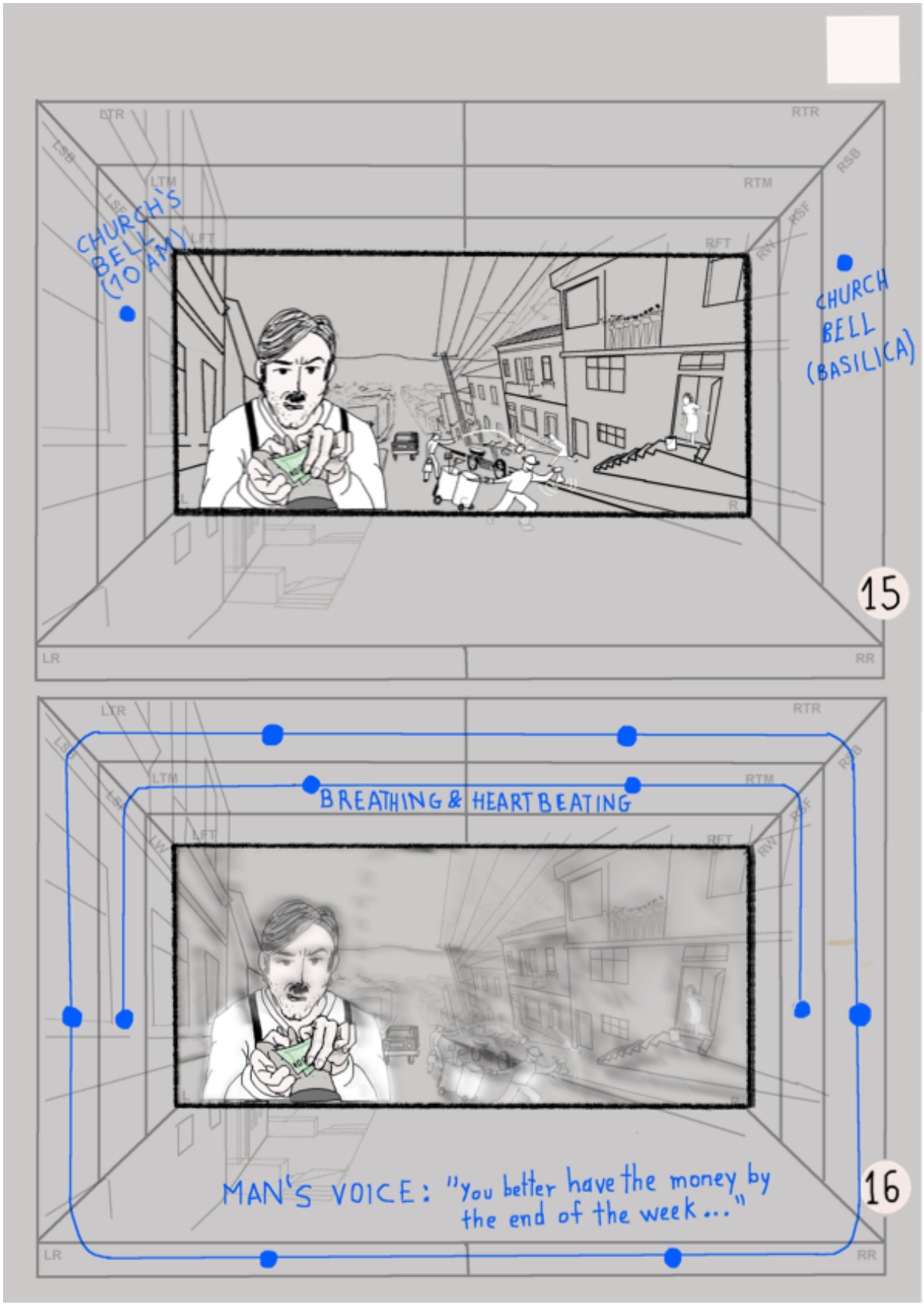

5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
Good morning, Don Edmundo. Not too bad...
How are you Arturo? How is...
RADIO
R
5
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
Hey you! Get of the bike!
They love your bike
RADIO
KIDS Laugh while running
L
R
6
LR
RR


5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
Hmm, look, I want to ask you a favor
Sure! How can I be of help?
RADIO
FOOTSTEPS
C
R
7
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
RADIO
NEWS PAPER SELLING CALL
FOOTSTEPS
SCRAP RECYCLER
"We are buying old cookers, old tins..."
(Few squares away)
8
LR
RR

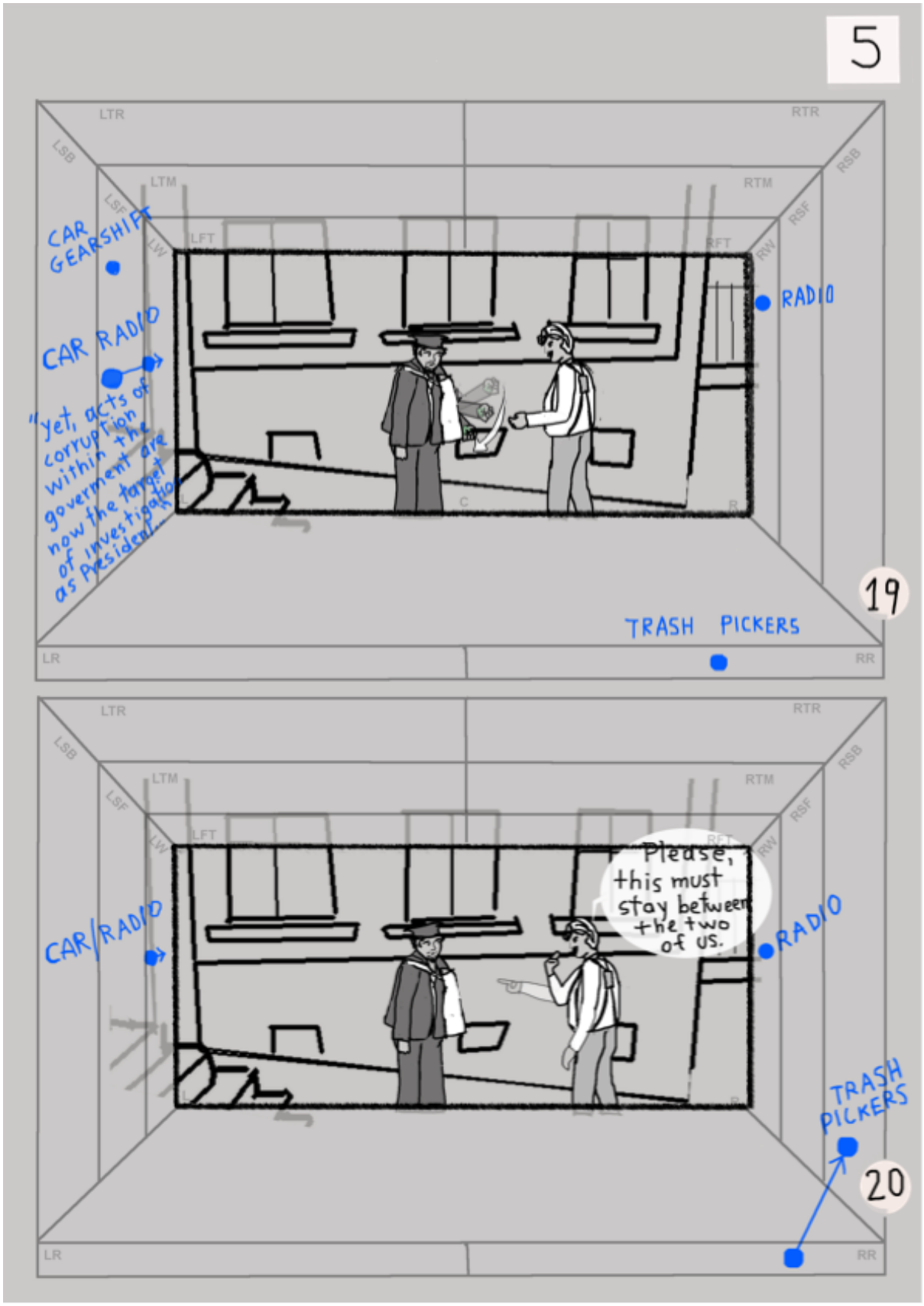

5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
NEWS PAPER SELLING CALL
FOOTSTEPS
Yeah... I think I know him. I'm afraid he went to work abroad. Didn't he?
You know, my daughter G. was seeing this guy Cesar T..
RADIO
Yes, he did... I saw you just delivered a letter for my daughter. It was from him?
SCRAP RECYCLER "We are buying old cookers, old tins..." (Few squares away)
9
LR
RR
10
...She has become very stubborn...
I see!
So...

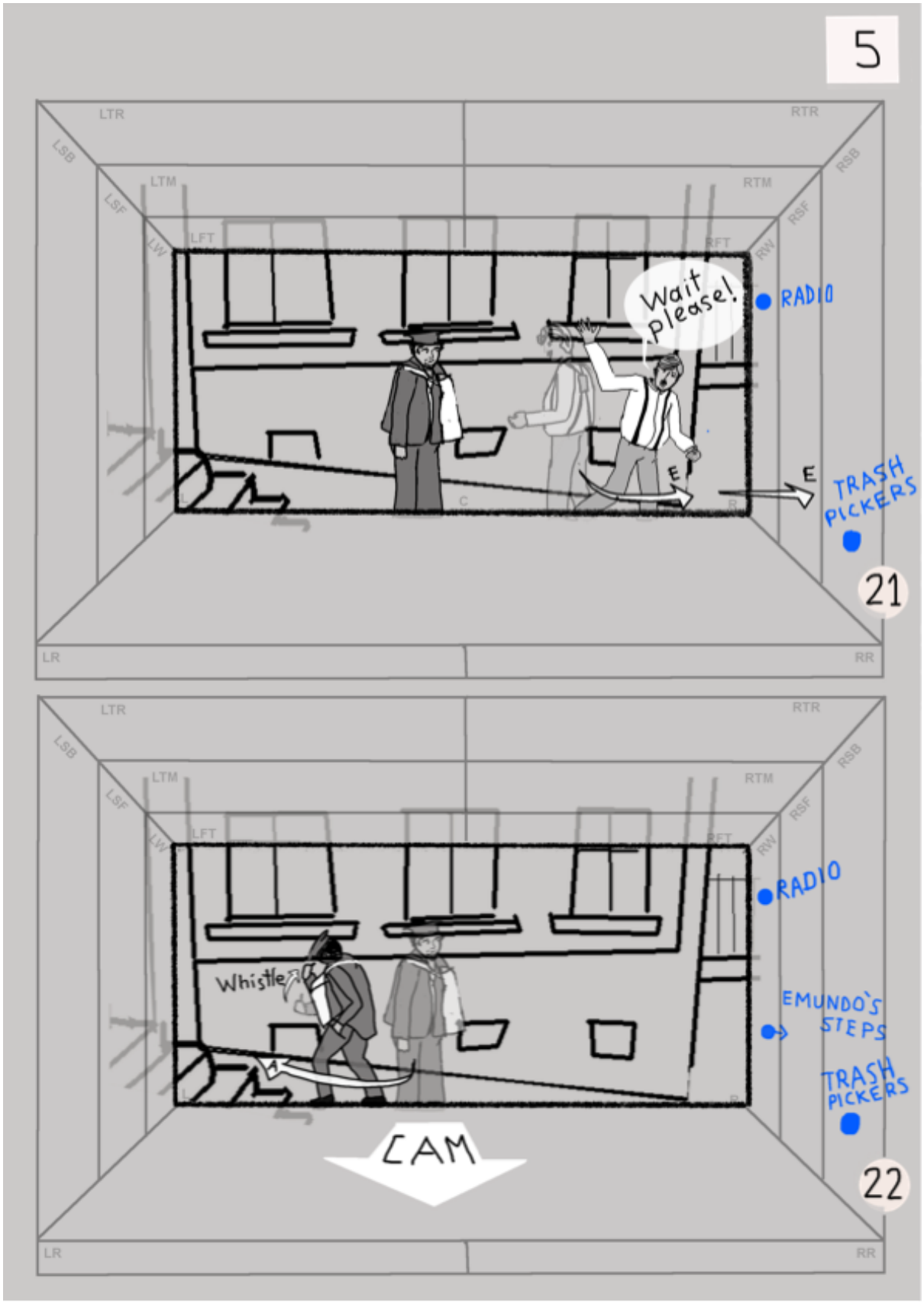

5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
#$*^%!!!
I forgot about the trash!
RADIO
TRASH PICKER BELL
11
LR
RR
TRASH PICKER'S BELL
METALIC IMPACTS
CHURCH BELL (BASILICA)
12
DOOR & METALIC IMPACTS
VOICES & DOOR

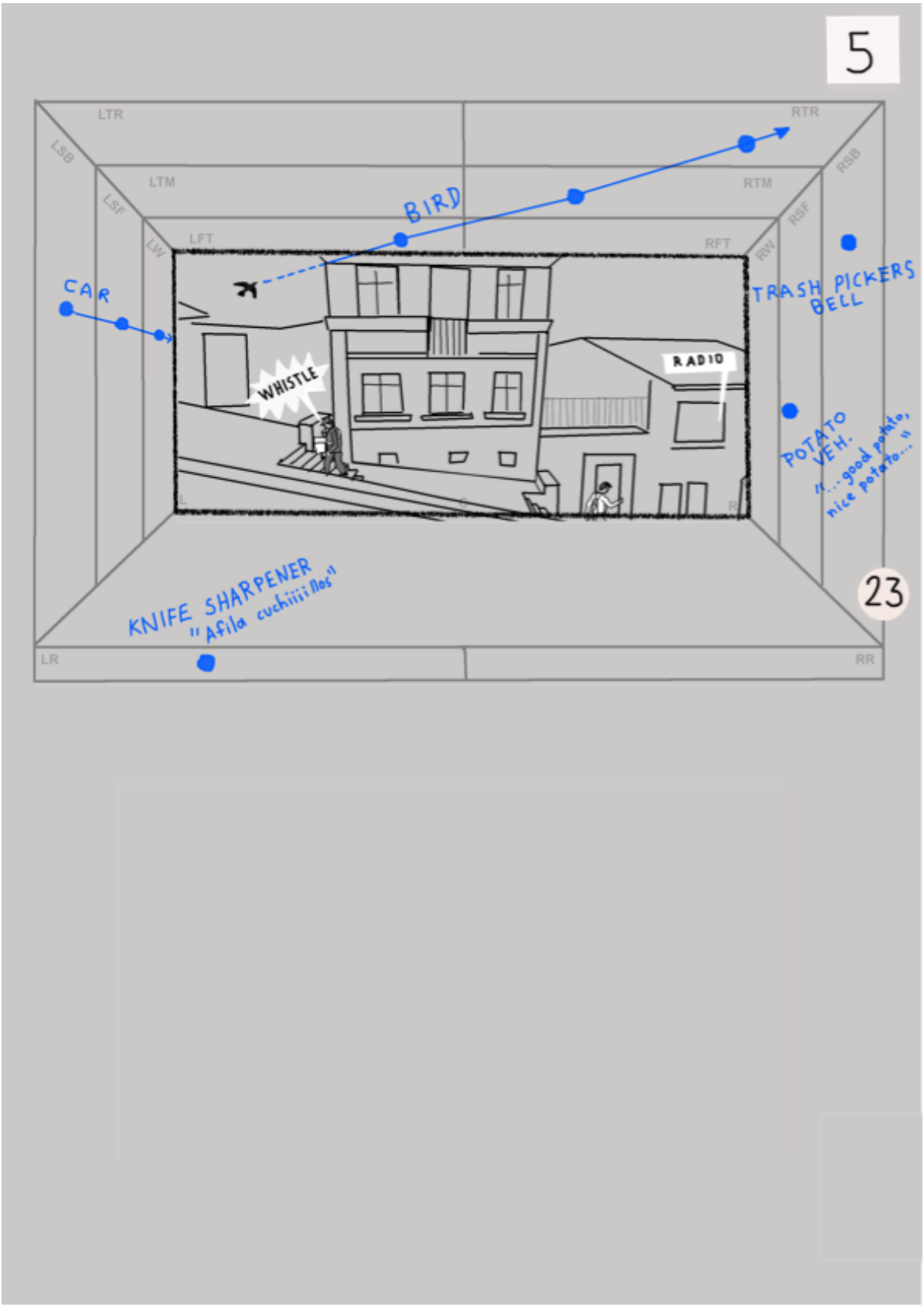

5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
C
L
R
LR
RR
Excuse me Don E. but I need to leave. It's late and...
Stop delivering?
Sorry Don E. but that's against the law ...
No, no wait. What I wanted to ask you is to stop delivering...
Yes, yes
CHURCH BELL (BASILICA)
RADIO
TRASH PICKER BELL
13
Here is a little something for your service
METALIC IMPACTS
14


CHURCH'S BELL (10 AM)
CHURCH BELL (BASILICA)
15
BREATHING & HEART BEATING
MAN'S VOICE: "You better have the money by the end of the week..."
16


5
RADIO
17
CAR'S RADIO The Petroleum State Corp of Ecuador has exported its first load...
RADIO
CAR
TRASH PICKERS
18


5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
C
R
CAR GEARSHIFT
CAR RADIO
"Yet, acts of corruption within the goverment are now the target of investigation as President...
RADIO
19
TRASH PICKERS
LR
RR
Please, this must stay between the two of us.
CAR RADIO
RADIO
TRASH PICKERS
20


5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
Wait Please!
RADIO
E
E
TRASH PICKERS
21
LR
RR
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LW
LFT
RFT
RW
Whistle
RADIO
EMUNDO'S STEPS
TRASH PICKERS
CAM
22
LR
RR


5
LTR
RTR
LSB
RSB
LTM
RTM
LSF
RSF
LFT
RFT
LW
RW
BIRD
CAR
WHISTLE
RADIO
TRASH PICKERS BELL
POTATO VEH.
"...good potato, nice potato..."
KNIFE SHARPENER
"Afila cuchiiiiiios"
23
LR
RR

GLOSSARY

3-D Sound: Three-dimensional sound. A cinematic sound design approach that deploying the capabilities of immersive audio tools spatialises sound (and music) around and even above the audience, coherently matching the camera's point-of-view with its point-of-audition.

ASL: Average shot length. The average duration of shots in a film or a group of films.

Audio Beds: Sound stems addressed as arrays to be reproduced on a group of loudspeakers, they integrate channel-based content—such as music and atmospheres—that does not require to be treated under an object-based approach.

Audio Objects: Sound elements that can be precisely localised and moved through the speakers across the auditorium, and hence are individually rendered with metadata that indicates their position for playback, their size, diffusion and other sonic characteristics.

Bass management: In cinema sound reproduction, bass management sends bass content to loudspeakers capable of reproducing it, such as the LFE channel.

Binaural Audio: Binaural is a term to describe human hearing (through two ears). Binaural audio is based on this principle to record and reproduce audio content that aims to simulate the way we hear.

CGI: Computer-generated imagery

Channel-based audio: A system of sound representation that delivers channels of audio to a specific loudspeaker (commonly a group of loudspeakers) in the cinema.

DAW: Digital audio workstation. Hardware/software used for recording, editing, mixing and producing audio files.

Deep-focus: A cinematographic technique that presents the image with a large depth-of-field.

Depth-of-field: The distance between the nearest and the farthest visual elements that are in acceptably sharp focus in an image.

DCP: Digital cinema package. A collection of digital files that stores and delivers digital cinema audio, image, and metadata.

Diegetic Sound: All elements of the soundtrack that emanate from within the narrative world of the film.

Direct Sound: The original sound waves produced by any given source, such as the sound that comes from an actor's mouth. When we are close to the source we hear mainly its direct sound, as we get farther away we hear its reverberation.

DSS: Digital surround sound. A multichannel technology that allows sound to be played back digitately through a determined number of discreate channels. Examples of DSS are Dolby 5.1 and 7.1.

Dynamic Range: In terms of sound, the dynamic range is the ratio between the largest (loudest) and smallest (quietest) amplitude, expressed in decibels (dB) that a system can assume.

'Exit Door' Effect: Occurs when audience members believe that a sound that emanates from a speaker of group of speakers located in the back of the theatre is actually a real sound that originates within the theatre, and thus they look away from the screen looking for the sound source.

HRTFs: Head-related transfer functions. In binaural recording, HRTFs are measured in order to represent the way an ear receives a sound from a point in space, which depends on the physical characteristics of the head, ears, nasal and oral cavities and torso.

Immersion: In the context if this thesis, it refers to a cinematic experience that transports the spectator into the narrative world by means of a realistic sonic representation of the three-dimensional space of the movie.

Immersive Point-of-Audition: The three-dimensional aural representation of the point within the diegesis from where we hear.

Immersive Sound: A technology that allows sound elements to be reproduced in all directions around the listener, meaning that sound mixers can position sounds around and above the audience.

LFE: Low-frequency effects. It's a channel dedicated to reproduce the low frequency content within a multitrack mix.

Monophonic Audio: A cinematic sound technology capable of reproducing sound only through one channel of audio, usually the centre speaker behind the screen.

Multichannel Sound: A cinematic sound technology capable of reproducing sound through various channels of audio, such as the typical 5.1 configuration than includes five surround channels and a subwoofer.

Object-based audio: A system of sound representation that defines sound elements—audio objects—with individual metadata that indicates how they will be reproduced during playback.

Point-of-audition: The point within the diegesis from where we hear.

POV Shot: Point-of-view shot. We see from the visual perspective of a given character.

Pro-filmic Space: The original scenario in which filming takes place.

Quadraphonic Sound: A multichannel sound configuration that uses four discreate channels of audio.

Reverberation: The sonic reflections produced when a sound wave reaches a surface, a body, or any given object. When we are close to the source we hear mainly its direct sound, as we get farther away we hear its reverberation.

Shot/reverse-shot: A conventional editing technique used in filmmaking to show one face and then another in an A-B-A-B pattern.

Signal-to-noise ratio: It is the ratio between the power of the desired sound signal and the power of the undesired sound signal (noise). A ratio greater than 0 dB indicates that there is more desired signal than noise.

SMPTE: Society of Motion Picture and Television Engineers.

Stems: Sound elements already mixed or grouped into a channel-based configuration such as stereo or 5.1.

Suspension of disbelief: A process that seduces the media user to stop counter-arguing the content of a given narrative.

Sweet spot: The zone in the theatre in which the spectator gets the optimal listening experience.

Walla: A sound effects that tries to emulate a crowd murmur in the background.

Wild Tracks: Sound effects that are recorded on location that could serve for the film's final mix.

BIBLIOGRAPHY

Andriano-Moore, S. (2018). The Rise of the Sound Designer: Northern California Film Sound in the 1960s and 1970s. *Historical Journal of Film, Radio and Television, 38(3),* 536-554.

Allen, R. (1993). Representation, Illusion, and the Cinema. *Cinema Journal*, 32(02): 21-48.

— (1995). *Projecting Illusion: Film Spectatorship and The Impression of Reality*. Cambridge: Cambridge University Press.

Allen, N. (2019). *Netflix spends record $30 million on its Oscars campaign for Roma, double what the movie cost to make* [Online]. The Telegraph. Availible at: https://www.telegraph.co.uk/news/2019/02/22/netflix-spends-record-30-million-oscars-campaign-roma-double/ [Accessed 10 Apr 2019].

Althusser, J. (1984 [1971]). Ideology and Ideological State Apparatuses (Notes towards an Investigation), *Essays on Ideology*, pp. 1-60.

Altman, R. (1992a). Sound space. In R. Altman, (Ed.). *Sound theory sound practice*. London: Routledge, pp. 46-64.

— (1992b). Introduction: Four and a half film fallacies. In R. Altman, (Ed.). *Sound theory sound practice*. London: Routledge, pp. 35-45.

— (1992c). The Material Heterogeneity of Recorded Sound. In R. Altman, (Ed.) *Sound Theory Sound Practice*. London: Routledge, pp. 15-31.

Appel, M., Koch, E., Schreier, M. & Groeben, N. (2002). Aspekte des Leseerlebens. Skalenentwicklung [Aspects of the reading experience: scale development]. *Zeitschrift für Medienpsychologie*, 14, 149–154.

AudioXpress (2016). *Top Four Emerging Trends Impacting the Global Home Theater Market 2016-2020.* [Online]. Available at: https://www.audioxpress.com/news/Top-Four-Emerging-Trends-Impacting-the-Global-Home-Theatre-Market-2016-2020 [Accessed 12 Apr 2019].

Augst, B. (1976) The Apparatus: An Introduction, Camera Obscura, 01(01), 97-101.

Auro Technologies and Barco (2015). AuroMax Next Generation Immersive Sound System: White Paper. [Online] Available at: http://www.auro-3d.com/wp-content/uploads/documents/AuroMax_White_Paper_24112015.pdf [Accessed 11 Feb 2019]

Auro Technologies (2021). *Auro 3-D* [Online] Available at: https://www.auro-3d.com/professional/industries/cinema/ [Accessed 4 Feb 2021].

Avid (2018). *Pro Tools Ultimate Features*. [Online]. Available at: https://www.avid.com/pro-tools-ultimate/features [Accessed 22 Apr 2018].

B, B. (2006). Humanity's Last Hope. *American Cinematographer - The International Journal of Film & Digital Production Techniques,* Dec 2006, 87(12): 60-75.

— (2013). Facing the Void. *American Cinematographer - The International Journal of Film & Digital Production Techniques,* Nov 2013, 94(11): 36-49.

BAFTA Guru (2019). 'Alfonso Cuarón - Did *Children of Men* Predict Populism Today? | Screenwriter's Lecture', [Online] Available at: https://www.youtube.com/watch?v=dh1EGsetbJE [Accessed: 8 Apr. 2020].

Balla, G. (2016). *The Director's Method in Contemporary Visual Effects Film: The Influence of Digital Effects on Film Directing*. Ph.D. Thesis. Univeristy of York.

Barnes, B. (2019). *Oscars Provide Stage for Participant Media's Comeback Story [Online].* Available at: https://www.nytimes.com/2019/02/22/business/media/participant-media-oscars.html [Accessed: 3 Apr 2020].

Barker, M. (2005). The Lord of the Rings and 'identification': a critical encounter. *European Journal of Communication,* 20(03), 353-378.

Baudry, J-L. (1974-1975 [1970]). Ideological Effects of the Basic Cinematographic Apparatus. Translated by Adam Williams. *Film Quarterly*, 28(02), 39-47.

— (1976 [1975]). The Apparatus: Metapsychological Approaches to the Impression of Reality in the Cinema. Translated by Jean Andrews and Bertrant Augst. *Camera Obscura*, 01(01): 104-126.

Bazin, A. (1967) *What is cinema?* Vol. 1. Translated by Hugh Gray, Berkeley, Los Angeles and London: University of California Press.

— (1971) *What is Cinema?* Vol. 2. Translated by Hugh Gray, Berkeley, Los Angeles and London: University of California Press.

Beck, J. (2003). *A quiet revolution: changes in American film sound practices, 1967-1979*. Vol. 1. Ph.D. Thesis. Iowa: The University of Iowa

— (2008). The sound of "silence": Dolby Stereo, sound design, and The Silence of The Lambs. In J. Beck and T. Grajeda (Eds.). *Lowering the boom: critical studies in film sound*. Urbana and Chicago: University of Illinois Press, pp. 68-83.

Beck, J., and Grajeda, T (2008). Introduction. in J. Beck and T. Grajeda (Eds.). *Lowering the Boom: Critical Studies in Film Sound*. Urbana: University of Illinois Press.

Belton, J. (1985). Technology and Aesthetics of Film Sound. In E. Wais and J. Belton (Eds.) *Film Sound: Theory and Practice*. New York: Columbia University Press, pp. 63-71.

— (1992). 1950s magnetic sound: the frozen revolution. In R. Altman, (Ed.). *Sound theory sound practice*. London: Routledge, pp. 154-167.

Bender, S. (2014). There is nothing to carry sound: defamiliarization and reported realism in Gravity. *Senses of Cinema*, 71. [Online] Available at: http://sensesofcinema.com/2014/feature-articles/there-is-nothing-to-carry-sound-defamiliarization-and-reported-realism-in-gravity/ [Accessed 07 Jul 2018].

Bilandzic, H. and Busselle, R. (2011). Enjoyment of films as a function of narrative experience, perceived realism and transportability. *Communications*, 36(1): 29-50.

Böcking, S., and Wirth, W. (2005). Towards conceptualizing suspension of disbelief for communication research [Online]. *Annual Conference of the International Communication Association*, New York. Available at

http://citation.allacademic.com//meta/p_mla_apa_research_citation/0/1/2/6/6/pages12669/p12669-1.php [Accessed 28 Oct 2019].

Bordwell, D. (2002). Intensified Continuity: Visual Style in Contemporary American Film. *Film Quarterly*, 55(03): 16-28.

— (2015). *Birdman: Following Riggan's orders* [Online]. Observations on film art. Available at: http://www.davidbordwell.net/blog/2015/02/23/birdman-following-riggans-orders/ [Accessed: 29 Aug. 2018].

Bordwell, D. and Carroll, N. (1996). *Post-theory: reconstructing film studies*. Madison: University of Wisconsin Press.

Bordwell, D. and Thompson, K. (2013). *Film Art: An Introduction*. 10th ed. New York: McGraw-Hill Education.

Bordwell, D.; Thompson, K.; and Smith, J. (2016). *Film Art: An Introduction*. 11th Edition. New York: McGraw-Hill Education.

Burch, N. (1982). Narrative/Diegesis-Thersholds, Limits. *Screen*, 23(02), pp. 16-33.

Butt, E. (2019) Personal Interview with author. 11 June 2019.

Carroll, N. (1988). *Mystifying Movies: Fads and Fallacies in Contemporary Film Theory*. New York: Columbia University Press.

— (2008). *The Philosophy of Motion Pictures*. Malden, MA: Blackwell.

Cass, J. (1930). The illusion of sound and picture. *Journal of the Society of Motion Picture Engineers*, 14(03), 323-326.

Chagollan, S. (2019). *Alejandro G. Iñárritu's Metaphysical World View* [Online]. Directors Guild of America. Available at: https://www.dga.org/Craft/DGAQ/All-Articles/1901-Winter-2019/DGA-Interview-Alejandro-Gonzalez-Inarritu.aspx [Accessed 07 Sep 2020].

Chion, M. (1994 [1990]). *Audio-vision: sound on screen*. Edited and translated by Claudia Gorbman. New York: Colombia University Press.

— (1999 [1982]). *The Voice in Cinema*. Edited and translated by Claudia Gorbman. New York: Columbia University Press.

Cinemetrix. *Berry Salt's Data Base*. [Online]. Available at: http://cinemetrics.lv/satltdb.php#asl [Accessed 31 Jul 2020].

Cine Premiere (2019). *Los secretos del sonido en Roma.* [Online]. Availible at: https://www.youtube.com/watch?v=ScMj2qZqcbg [Accessed 8 Apr 2019].

Clark, T. (2020). Personal interview with author. 9 September 2020.

Claypool, B., Van Baelen, W. and Van Daele, B. (N.D). Auro 11.1 versus Object-Based Sound in 3D: All Aspects Compared. *Barco* [Online]. Availible at: https://www.barco.com/secureddownloads/cd/MarketingKits/3d-sound/White%20papers/Auro%2011.1_versus_objectbased_sound_in_3D.pdf [Accessed 11 Feb 2019]

Cohen, J. (2001). Defining identification: a theoretical look at the identification of audiences with media characters. *Mass Communication & Society*, 4(03), 245-264.

— (2006). Audience identification with media characters. In J. Bryant and P. Vorderer (Eds.). *Psychology of Entertainment*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, pp. 183–97.

Cohen, S. (2019). What is Dolby Atmos Music and how can you experience it? Digital Trends [Online] Available at: https://www.digitaltrends.com/home-theater/what-is-dolby-atmos-music-and-how-to-get-it/ [Accessed 25 Jun 2019]

Connor, S. (2017). Immersive audio: Optimizing creative impact without increasing production costs. *AES Convention 143*, Convention e-Brief 370, 1-7.

Consumer Technology Association (2016). *Digital America: State of the U.S. Consumer Technology Industry.* [Online]. Available at: https://www.cta.tech/CTA/media/News/2016_DigitalAmerica.pdf [Accessed 12 Apr 2019].

— (2018). *Digital America: State of the U.S. Consumer Technology Industry 2018.* [Online]. Available at:

https://editions.mydigitalpublication.com/publication/?i=534274#{%22issue_id%22:534274,%22page%22:0} [Accessed 12 Apr 2019].

Coria, J.F. (2019). *Alfonso Cuarón o el Estilo CUEC* [Online]. Confabulario-El Universal. Available at: https://confabulario.eluniversal.com.mx/alfonso-cuaron-roma-2/ [Accessed 8 Jan 2020].

Cortés-Selva, L. (2016). En busca del VRCinema. Del cine proto-inmersivo al cine inmersivo. *Discursos Forográficos: Londrina*, 12(20), 173-204.

Costantini, G. (2018). Approaches to Sound Design: Murch and Burtt. *The New Soundtrack*, 8(2), 169-174.

Cuarón, A. (2013a). The sound of Gravity. *Sound Works Collection* [Online] Available at: http://www.soundworkscollection.com/videos/gravity [Accessed 06 Jul 2018].

— (2013b). Alfonso Cuarón – Gravity. *ICG Magazine* [Online] Available at: http://www.icgmagazine.com/web/exposure-alfonso-cuaron/ [Accessed 2 Oct 2018].

— (2014). Alfonso Cuarón and the Sound of Gravity. *Dolby* [Online]. Available at: https://www.youtube.com/watch?v=c7dKt80o0qE [Accessed 7 Feb 2019].

— (2016). *Roma's Screenplay* [Online]. Available at: https://pmcdeadline2.files.wordpress.com/2018/12/Roma-Screenplay-ENGLISH.pdf [Accessed 10 Apr 2019].

Cuarón, A. and Cuarón, J. (2012) Gravity. Screenplay [Online] Available at: https://www.retourverslecinema.com/wp-content/uploads/2013/11/scenario_du_film_gravity.pdf [Accessed 21 Sep 2020]

Dancyger, K. (2006). *The Director's Idea: The Path to Great Directing*. Amsterdam and London: Focal Press.

Danto, A. (2006). Moving Pictures [1979]. In N. Carroll, J. Choi (Eds.). *Philosophy of Film and Motion Pictures: An Anthology*. Malden, MA: Blackwell,pp. 100–112.

Díaz, S. (2019) Personal Interview with author. 2 October 2019.

— (2021). Personal email correspondence with author. 4 January 2020.

— (2021). Personal interview with author. 15 January 2021.

Dreher, C. (1931). Recording, re-recording and editing of sound. *Journal of the Society of Motion Picture Engineers*, 16(6), 756-765.

Doane, M.A. (1980). The Voice in the Cinema: The Articulation of Body and Space, Yale French Studies, 60: 33-50.

— (1985). Ideology and the Practice of Sound Editing and Mixing. In E. Wais and J. Belton (Eds.) *Film Sound: Theory and Practice*. New York: Columbia University Press, pp. 54-62.

Dolby (2014). Dolby Atmos Next Generation Audio for Cinema: White Paper. [Online]. Available at: https://www.dolby.com/us/en/technologies/dolby-atmos/dolby-atmos-next-generation-audio-for-cinema-white-paper.pdf [Accessed 14 Mar 2019].

Dolby (2016). Dolby Atmos Enabled Speaker Technology. Dolby [Online]. Available at: https://www.dolby.com/us/en/technologies/dolby-atmos/dolby-atmos-enabled-speaker-technology.pdf [Accessed 24 Jun 2019].

Dolby (2018a). *Dolby Atmos.* [Online]. Available at: https://www.dolby.com/us/en/brands/dolby-atmos.html [Accessed 18 Apr 2018].

Dolby (2018c). *Home Theatre Speaker Guide* [Online]. Available at: https://www.dolby.com/us/en/guide/7.1.2-dolby-atmos-enabled-speaker-setup-guide.pdf [Accessed 22 Apr 2018].

Dolby (2019). *Global Leader on Immersive Audio.* [Online]. Available at: https://www.dolby.com/us/en/technologies/dolby-atmos/cinema-exhibitor-testimonials.html [Accessed 27 May 2020].

Dolby (2020a). *Dolby Digital 5.1* [Online]. Available at: https://professional.dolby.com/tv/dolby-digital/ [Accessed 14 Oct 2020].

Dolby (2020b). *Bring Back the Magic: Will Files, The Invisible Man*. [Online]. Available at: https://www.youtube.com/watch?v=97DRePyy9tc [Accessed 13 Oct 2020].

Dolby (2021). Dolby Atmos: Global Leader in Immersive Audio [Online] Available at: https://professional.dolby.com/cinema/ [Accessed 04 Feb 2020].

DTS (2012) The First Step Toward 3D Audio: DTS Neo:XTM. White Paper. [Online] Available at: https://www.stormaudio.com/media/dtsneoxwhite_paper__019028000_1625_04032013.pdf [Accessed 24 Jun 2019].

DTS (2020b) Production Tools. *DTS* [Online] Available at: https://dts.com/professional/production-tools [Accessed 14 Oct 2020].

DTS (2021) DTS Movies at Home: Blu-ray Releases Featuring DTS:X and DTS-HF Audio. *DTS* [Online] Available at: https://dts.com/discover/dts-movies-at-home [Accessed 14 Oct 2020].

Fazio, G. (2016). *Lubezki achieves the extraordinary long shot* [Online]. The Japan Times. Available at: https://www.japantimes.co.jp/culture/2016/04/20/films/lubezki-achieves-extraordinary-long-shot/ [Accessed 09 Sep 2020].

Filmmaker Magazine (2018). *Skip Lievsay and Craig Henighan on Designing the Sound of Alfonso Cuarón's Roma*. [Online]. Available at: https://filmmakermagazine.com/106491-skip-lievsay-craig-henighan-alfonso-cuaron-roma/#.XKIgFutKiRs [Accessed 01 Apr 2019].

France, R. (2019) Media Production Show Conference. London. 12 July 2019.

Freud, S. (1927). Fetishism. In *Miscellaneous Papers, 1888-1938*, Vol.5 of *Collected Papers*, (London : Hogarth and Institute of Psycho-Analysis, 1924-1950), 198-204.

Füg, S., Marston, D. and Norcross, S. (2016). The Audio Definition Model – A Flexible Standardised Representation for Next Generation Audio Content in Broadcasting and Beyond. *AES Convention 141*, 9626, 1-10).

Gabriel, G. and Sonnenschein, D. (2016). Inner and Outer Worlds in the Film Gravity: A Multidisciplinary Approach. In L. Greene and D. Kulezic-Wilson (Eds.). *The Palgrave Handbook of Sound Design and Music in Screen Media.* London: Palgrave Macmillan UK, pp. 90-96.

Gerrig, R. (1993). *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. New York; New Haven: Yale University Press.

Getino, O. and Solanas, F. (1969). Toward a Third Cinema. *Tricontinental*, 14, 107-132.

GoldDerby (2018). *Sergio Diaz ('Roma' sound editor) on evoking the Mexico City of Alfonso Cuaron's past*. [Online]. Available at: https://www.youtube.com/watch?time_continue=23&v=s-5nBxB51nM [Accessed 5 Apr 2019].

Grajeda, T. (2008). "A Question of the Ear" Listening to Touch of Evil. in J. Beck and T. Grajeda (eds). *Lowering the Boom: Critical Studies in Film Sound*. Urbana: University of Illinois Press, pp. 201-217.

Grajeda, T. (2016). Listening to Violence: Point-of-Audition Sound, Aural Interpellation, and the Rupture of Hearing. In L. Greene, D. Kulezic-Wilson (Eds.). *The Palgrave Handbook of Sound Design and Music in Screen Media.* London: Palgrave Macmillan, pp. 91-111.

Green, M. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes*, 38, 247–266.

Green, M. and Brock, T. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of personality and social psychology*, 79(05), 701-721.

Green, M., Brook, T. and Kaufman, G. (2004). Understanding media enjoyment: the role of transportation into narrative worlds. *Communication Theory*, 14(04), 311-327.

Green, M. and Clark, J. (2013). Transportation into narrative worlds: implications for entertainment media influences on tobacco use. *Addiction*, 108(03), 477-484.

Green, M. and Sestir, M. (2017). Transportation Theory. In P. Rössler, C. Hoffner and L. Zooner (Eds.). *The international encyclopedia of media effects*, Chichester, West Sussex; Malden, MA: John Wiley & Sons, Inc.

Greene, L. and Kulezic-Wilson, D. (2016) Introduction. In L. Greene and D. Kulezic-Wilson, (Eds.). *The Palgrave Handbook of Sound Design and Music in Screen Media : Integrated Soundtracks*. London: Palgrave Macmillan. Available at:

http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1231407&site=ehost-live [Accessed: 17 January 2019].

Greiner, E. (2019). Personal Interview with author. 20 September 2019.

Halligan, F. (2013). The Art of Movie Storyboards: Visualising the Action of the World's Greatest Films. East Sussex: ILEX.

Hernández, M. (2016). *The Sound of The Revenant* [Online]. Sound Works Collection. Available at: http://soundworkscollection.com/videos/the-sound-of-the-revenant [Accessed 08 Oct 2018].

Hoberman, J. (2013). *Drowning in the Digital Abyss.* [Online]. The New York Review. Available at: https://www.nybooks.com/daily/2013/10/11/gravity-drowning-digital-abyss/ [Accessed 07 Apr 2021].

Holman, T. (2010). Sound for Film and Television. New York and London: Focal Press - Taylor and Francis Group.

Hövenne, M. (2017). Editing Space as an Audio-visual Composition. In J. Wildfeuer and J. Bateman (Eds.). *Film Text Analysis: New Perspectives on the Analysis of Filmic Meaning*, New York: Routledge, pp. 46-65.

Hunt, F. (1938). Sound pictures in auditory perspectives. *Journal of the Society of Motion Picture Engineers*, 31(4), 351-357.

Hunt, A. (2020). 'Alfonso Cuarón Reveals His Ghost-Perspective, His Influences, and a Theory He Hates'. No Film School. [Online] Available at: https://nofilmschool.com/alfonso-cuaron-interview [Accessed 7 Sep 2020].

IMax (2013). Gravity: *IMax Behind the Frame*. [Online]. Available at: https://www.youtube.com/watch?time_continue=65&v=39w4Ow5QDGQ [Accessed 27 Mar 2019].

IMDb (2019). *Alfonso Cuarón*. [Online]. Available at: https://www.imdb.com/name/nm0190859/?ref_=nv_sr_1?ref_=nv_sr_1

International Society for Presence Research (2000). *Presence Defined.* [Online]. Available at: https://ispr.info/about-presence-2/about-presence/ [Accessed 5 Oct 2020].

Isaacs, B. (2016). Reality Effects: The Ideology of the Long Take in the Cinema of Alfonso Cuarón. In S. Denson and J. Leyda, (Eds.). *Post-Cinema: Theorizing 21st-Century Films*, Falmer: Reframe Books, pp. 474-513.

Johnson, D. (2019). Personal interview with author. 24 July 2019.

Johnson, S. (2019). *The Politics of "Roma": A Journey Through Mexico City's Turbulent Past and America's Divided Present* [Online]. Hollywood Reporter Availible at: https://www.hollywoodreporter.com/features/politics-roma-alfonso-cuaron-revisits-mexicos-turbulent-history-1176330 [Accessed 7 Jan 2020].

Jones, N. (2013). Quantification and Substitution: The Abstract Space of Virtual Cinematography. *Animation: An Interdisciplinary Journal*, 8(03), 253-266.

Karlsson, R. (2019) Personal Interview with author. 24 April 2019.

Katz, S. (1991). Film Directing Shot by Shot: Visualizing from Concept to Screen. Study City: Michael Wiese Productions.

Keating, P. (2019). *The Dynamic Frame: Camera Movement in Classical Hollywood*. New York: Columbia University Press.

Kenny, T. (1998). The Search for Order in Sound & Picture. *Mix*, April, 12-24.

— (2013) DTS Headphone:X: An 11.1 Experience That Brings the Consumer Into the Mix Room. *Mix*, December, 2013, pp. 24-28.

Kerins, M. (2006). Narration in the cinema of digital sound. *The Velvet Light Trap*, 58, 41-54.

— (2011). *Beyond Dolby Stereo: Cinema in the digital sound age*. Bloomington: Indiana University Press.

— (2015). The Modern Entertainment Marketplace, 2000-Present. In K. Kalinak, (Ed.). *Sound: Dialogue, Music and Effects*. New Brunswick, New Jersey: Rutgers University Press

King, R. (2019). School of Sound Conference, London. April 2019.

Kim, T., and Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-Mediated Communication*, 3(2).

Kiser, G. (2019). Personal Interview with author. Translated by the author. 10 September 2019.

Knudsen, E. (2016). The Total Filmmaker: thinking of screenwriting, directing and editing as one role. *New Writing*, 13 (01), pp.109-129.

Kohn, E. (2017). ‘‘Okja’: How One Visit to a Slaughterhouse Turned Bong Joon Ho Into a Vegan’. *IndieWire* [Online] Available at: https://www.indiewire.com/2017/06/okja-bong-joon-ho-vegan-1201839076/ [Accessed: 19 May 2020].

Kulezic-Wilson, D. (2008). Sound design is the new score. *Music, Sound and The Moving Image*, 2(2), 127-131.

Lacan, J. (1966 [1949]). The Mirror Stage as Formative of the *I* Function as Revealed in Psychoanalytic Experience. In *Ecrits, The First Complete Edition in English,* Translated by B. Fink. London: W.W. Norton, pp. 75-81.

Lastra, J. (1992). Reading, writing and representing sound. In R. Altman, (Ed.). *Sound theory sound practice*. London: Routledge, pp. 65-86.

— (2000). *Sound technology and the American Cinema: perception, representation, modernity*. New York: Columbia University Press.

Lee, J.Y. and Stringer, J. (2018). From Screenwriting for Sound to Film Sound Maps: The Evolution of Live Tone’s Creative Alliance with Bong Joon-ho. *The New Soundtrack*, 8(2), 145-159.

Lefebvre, M. (2006). Between Setting and Landscape in the Cinema. In M. Lefebvre, (Ed.). *Landscape and Film*. London and New York: Routledge.

Lewin, G. (1931). Dubbing and its relation to sound picture production. *Journal of the Society of Motion Picture Engineers*, 16(1), 38-48.

Liang, D. (2016). Sound, Space, Gravity: A Kaleidoscopic Hearing (Part I). *The New Soundtrack*, 6(1): 1-15.

Lievsay, S. (2013). How the sound masters of '*Gravity*' broke the rules to make noise in a vacuum. *The Verge* [Online] Available at: https://www.theverge.com/2013/10/10/4822482/the-sound-design-of-gravity-glenn-freemantle-skip-lievsay [Accessed 06 Jul 2018].

— (2018). "Conversations With Sound Artists: Roma - Skip Lievsay". *Dolby*. [Online] Available at https://www.youtube.com/watch?v=6MUwqTmdb8Y [Accessed 31 Jul 2019].

— (2019). Personal Interview with author. 24 September 2019.

Lima, K. and Cruz-Santiago, A. (2019). *Roma: how Alfonso Cuarón's movie is spurring Mexico to treat domestic workers more fairly.* [Online]. The Conversation Available at: https://theconversation.com/roma-how-alfonso-cuarons-movie-is-spurring-mexico-to-treat-domestic-workers-more-fairly-112338 [Accessed 8 Jan 2020].

Loewinger, L. (1998). A Sound Idea: The Rationale Behind the Position of Sound Designer And Why It Never Took Hold [Online]. *The Independent*, October, Availible at: http://www.larryaudio.com/aSouIde.html [Accessed: 9 Jan 2019].

Lombard, M., Lee, S., Sun, W., Xu, K. and Yang. H. (2017). Presence Theory. In P. Rössler, C. Hoffner and L. Zooner (Eds.). *The international encyclopedia of media effects*, Chichester, West Sussex; Malden, MA: John Wiley & Sons, Inc.

Lossius, T. and Anderson, J. (2014). ATK Reaper: The Ambisonic Toolkit as JSFX plugins. *Proceedings ICMC|SMC,* 1338-1345.

MacCabe, C. (1974). Realism and the Cinema: Notes on some Brechtian Theses. *Screen*, 15 (02), pp. 7-27.

Manolas, C. and Pauletto, S. (2009). Enlarging the Diegetic Space: Uses of the Multi-channel Soundtrack in Cinematic Narrative. *The Soundtrack*, 2(01): 39-55.

Maxfield, J. (1930). Acoustic control of recording for talking motion pictures. *Journal of the Society of Motion Picture Engineers*, 14(01), 85-95.

Maxfield, J. and Harrison, H. (1926). Methods of high quality recording and reproducing of music and speech based on telephone research. *Bell System Technical Journal*, 5(03), 493-523.

McClintock, P. (2018). *In an unprecedented move by Netflix, Alfonso Cuaron's acclaimed film is opening in select cinemas more than three weeks before it debuts on the streamer* [Online]. Hollywood Reporter. Available at: https://www.hollywoodreporter.com/news/roma-details-netflix-movie-theaters-1163315 [Accessed 10 Apr 2019].

Mera, M. (2016) Towards 3-D Sound: Spatial Presence and the Space Vacuum. In L. Greene, D. Kulezic-Wilson (Eds.). *The Palgrave Handbook of Sound Design and Music in Screen Media.* London: Palgrave Macmillan, pp. 91-111.

Metz, C. (1974). On the Impression of Reality in the Cinema. In C. Metz. *Film Language: A Semiotics of the Cinema*. Translated by Michael Taylor. New York: Oxford University Press, 1974, pp. 3-15.

— (1982 [1975]). *The imaginary signifier: psychoanalysis and the cinema*. Translated by Celia Britton, Annwyl Williams, Ben Brewster and Alfred Guzzetti. Bloomington: Indiana University Press.

Morgan, D. (2016). Where are we?: camera movements and the problem of point of view. *New Review of Film and Television Studies*, 14(02), 222-248.

Murch, W. (2001). *In the Blink of an Eye: A Perspective on Film Editing*. 2dn Ed. Beverly Hills: Silma-James Press.

Nájar, A. (2018*). "Roma", de Alfonso Cuáron: quiénes eran Los Halcones, el brutal grupo paramilitar mexicano que aparece en la aclamada película*. [Online]. BBC Mundo. Availible at: https://www.bbc.com/mundo/noticias-america-latina-46562349

Netflix (2018a). *Roma: The Sounds of Real Life.* [Online]. Available at: https://www.youtube.com/watch?v=t7umEMFXxM0&feature=youtu.be [Accessed 19 Mar 2019].

— (2018b). *Alfonso Cuarón's Roma to be Released Theatrically in Over 600 Theaters Worldwide*. [Online]. Available at: https://media.netflix.com/en/press-releases/alfonso-cuar%C3%B3ns-roma-to-be-released-theatrically-in-over-600-theaters-worldwide [Accessed 19 Mar 2019].

— (2018c). *MyROMA Q&A with Writer/Director Alfonso Cuarón*. [Online]. Available at: https://www.youtube.com/watch?v=atZ0dx6pUno [Accessed 5 Apr 2019].

Oldfield, R., Shirley, B., Spille, J. (2013). Object-based audio for interactive football broadcast. *Multimed Tools Appl,* 74, 2717–274

Ondaatje, M. (2002). *The Conversations: Walter Murch and the Art of Editing Film,* London: Bloomsbury Publishing.

Oumano, H. (2011). *Cinema Today: A Conversation with Thirty-nine Filmmakers from around the World*. New Brunswick, New Jersey and London: Rutgers University Press

O'Falt, C. (2019). *Roma: Alfonso Cuarón Proved Dolby Atmos is the Perfect Tool for Intimate Films, not Blockbusters* [Online]. Indiwire. Available at: https://www.indiewire.com/2019/02/roma-sound-mixing-alfonso-cuaron-skip-lievsay-dolby-atmos-1202041828/ [Accessed 01 Apr 2019].

O'Meara, J. (2012). Inside the outsider's ear: hyperacousia and marginalized character identification. *The Soundtrack*, 5(02): 121-129.

Palenchar, J. (2015). DTS:X Positioned as Alternative to Atmos. *Twice,* 26 January 2015, p. 8.

Picone, J. (2015). *The best cinematography: a look at Birdman* [Online]. New York Film Academy. Available at: https://www.nyfa.edu/student-resources/best-cinematography-look-birdman/ [Accessed: 29 Aug. 2018].

Plato, *The Republic*, vii. Translated by Francis M. Cornford (1976). Camera Obscura, 01(01), 103.

Quesada, J. (2019). Personal Interview with author. 29 November 2019.

Rabiger, M. and Hurbis-Cherrier, M. *Directing: Film Techniques and Aesthetics*. New York and London: Focal Press: Taylor and Francis Group.

Recuber, T. (2007). Immersion Cinema: The Rationalization and Reenchantment of Cinematic Space. *Space and Culture*, 10(3): 315-330.

Richmond, S. (2016). *Cinema's Bodily Illusions: Flying, Floating, and Hallucinating*. Minneapolis, London: University of Minnesota Press.

Richter, F. (2021). Netflix Growth Re-Accelerates to Round Off 2020. *Statista* [Online] Available at: https://www.statista.com/chart/21465/global-paid-net-subscriber-additions-by-netflix/ [Accessed 16 Feb 2021].

Riley, J. (2015). *'Revenant' cinematographer Emmanuel Lubezki used only natural light* [Online]. Variety. Available at: https://variety.com/2015/artisans/production/the-revenant-cinematography-emmanuel-lubezki-1201661435/ [Accessed: 29 Aug. 2018].

Robertson, A. (2017). *Alejandro González Iñárritu's incredible VR experience is getting a special Oscar award* [Online]. The Verge Available at: https://www.theverge.com/2017/10/27/16562434/alejandro-gonzalez-inarritu-carne-y-arena-oscar-special-award-vr [Accessed: 7 may. 2020].

Rocha, H. (2017). Cartographic Soundtrack: Voice, Sound and Music in Birdman. *The New Soundtrack*, 7(2), 137–151.

Roginska, A. and Geluso, P. (2018). Introduction. In A. Roginska and P. Geluso (Eds.). *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*. New York: Routledge, pp. 1-4.

Rumsey, F. (2013). Cinema Sound in the 3-D Era. *Journal of The Audio Engineering Society*, 61(05), 340-344.

— (2018b). Surround sound. In A. Roginska and P. Geluso (Eds.). *Immersive sound: The art and science of binaural and multi-channel audio*. New York: Routledge.

Rushton, R. (2002). Cinema's double: some reflections on Metz. *Screen*, 43(02): 107-118.

Rushton, R., and Bettinson, G. (2010). *What is Film Theory: An Introduction to Contemporary Debates*, Berkshire: Open University Press.

Russell, K. (2001). *Directing Film: The Director's Art from Script to Cutting Room*. Barcelona: Bookprint.

Ryan, M. (2010). *Producer to Producer: A Step by Step Guide to Low-Budget Independent Film Producing*. Studio City: Michael Wiese Productions.

Sakamoto, S. (n.y.). *The Revenant: Shooting in the elements* [Online]. Society of Camera Operators. Available at: http://soc.org/project/the-revenant-shooting-in-the-elements/ [Accessed: 29 Aug. 2018].

Salt, B. (2009). *Film Style and Technology: History and Analysis*. 3rd Ed. London: Starword.

Sergi, G. (2013) Knocking at the door of cinematic artifice: Dolby *Atmos*, challenges and opportunities. *The New Soundtrack*, 3(02): 107–121.

Sestir, M. and Green, M. (2010). You are who you watch: Identification and transportation effects on temporary self-concept. *Social Influence*, 5(4): 272-288.

Sobchack, V. (1992). *The Address of the Eye: A Phenomenology of Visual Experience*. Princeton, NJ: Princeton University Press.

Sider, L. (2003). If You Wish to See, Listen: The Role of Sound Design. *Journal of Media Practice*, 4(1), 5-16.

Sims, D. (2018). *Netflix Finally Changed Its Biggest Rule* [Online]. The Atlantic Available at: https://www.theatlantic.com/entertainment/archive/2018/11/netflix-roma-ballad-of-buster-scruggs-theatrical-release-rule/574888/ [Accessed 11 Apr 2019].

Sippell, M. (2018). *Alfonso Cuaron Says Roma Is Better in Theaters* [Online]. Variety Available at https://variety.com/2018/scene/news/alfonso-cuaron-seeing-roma-in-theaters-home-netflix-1203086472/ [Accessed 11 Apr 2019].

Smith, J. (2013) The Sound of Intensified Continuity. In J. Richardson and C. Gorbman (Eds.). *The Oxford Handbook of New Audiovisual Aesthetics*. Oxford: Oxford University Press, pp. 331-356.

Society of Motion Picture and Television Engineers (SMPTE) (2018). *SMPTE Publishes Immersive Audio Standards for Cinema* [Online] Available at: https://www.smpte.org/news-events/news-releases/smpte-publishes-immersive-audio-standards-cinema [Accessed 12 March 2019]

SoundWorks Collection (2015). *Conversations with Sound Artists: Writing For Sound.* [Online]. Available at https://www.dolby.com/us/en/dolby-institute/podcasts.html [Accessed 9 Jan 2019].

Stam, R. and Spence, L. (1983). Colonialism, Racism and Representation. *Screen*, 14 (02), pp. 2-20.

Statista (2019a). *Preferred movie watching locations among adults in the United States as of February 2018*. [Online]. Available at: https://www.statista.com/statistics/264399/preferred-place-of-movie-consumption-in-the-us/ [Accessed 15 Apr 2019].

— (2019b). *Netflix - Statistics & Facts*. [Online]. Available at: https://www.statista.com/topics/842/netflix/ [Accessed 15 Apr 2019].

— (2021a) *Number of Netflix paid subscribers worldwide from 3rd quarter 2011 to 3rd quarter 2020* [Online] Available at: https://www.statista.com/statistics/250934/quarterly-number-of-netflix-streaming-subscribers-worldwide/#:~:text=Netflix%20had%20195.15%20million%20paid,Netflix's%20total%20global%20subscriber%20base [Accessed 20 Jan 2021].

— (2021b) *Number of households subscribing to Netflix in the United Kingdom (UK) from 1st quarter 2014 to 3rd quarter 2020* [Online] Available at: https://www.statista.com/statistics/529734/netflix-households-in-the-uk/#:~:text=The%20number%20of%20households%20in,subscribers%20every%20quarter%20since%202014 [Accessed 16 Feb 2021].

— (2021c) *Number of Netflix paid subscribers worldwide from 1st quarter 2013 to 4th quarter 2020* [Online] Available at: https://www.statista.com/statistics/250934/quarterly-number-of-netflix-streaming-subscribers-

worldwide/#:~:text=Netflix%20had%20195.15%20million%20paid,Netflix's%20total%20global%20subscriber%20base. [Accessed 16 Feb 2021].

Stern, M. (2017). *Alfonso Cuarón Explains the Darwinian Ending of Gravity*. Daily Beast [Online]. Available at: https://www.thedailybeast.com/alfonso-cuaron-explains-the-darwinian-ending-of-gravity?ref=scroll [Accessed: 6 Apr. 2020].

Sweney, M. (2018). *Netflix to overtake Sky's satellite TV subscriptions by end of year* [Online]. The Guardian. Available at: https://www.theguardian.com/media/2018/dec/23/netflix-to-overtake-sky-satelite-tv-subscriptions-by-end-of-year [Accessed 15 Apr 2019].

Tal-Or, N., and Cohen, J. (2010). Understanding audience involvement: conceptualizing and manipulating identification and transportation. *Poetics*, 38, 402–418.

Technavio (2018). *Global Home Audio Equipment Market 2018-2022*. [Online]. Available at https://www.technavio.com/report/global-home-audio-equipment-market-analysis-share-2018 [Accessed 11 Apr 2019].

The Numbers (2021). *Domestic Movie Theatrical Market Summary 1995 to 2020.* [Online] Available at: https://www.the-numbers.com/market/ [Accessed 16 Feb 2021].

Thom, R. (1999). *Designing a Movie for Sound* [Online]. Filmsound.org. Availible at: http://filmsound.org/articles/designing_for_sound.htm [Accessed 28 May 2019]

— (2015). "Conversations with Sound Artists: Writing For Sound". *Sound Works Collection*. [Online] Available at https://www.dolby.com/us/en/dolby-institute/podcasts.html [Accessed 9 Jan 2019].

— (2019). Personal Interview with author. 10 August 2019.

Thom, R. (2019b). *Sound Design Storyboards* [Online]. Randy Thom Blog: Thoughts on Film Sound Design. Available at: https://randythomblog.wordpress.com/2019/03/14/sound-design-storyboards/ [Accesed 02 Oct 2020]

Thom, R. (2019c). *Screenwriting for Sound* [Online]. Randy Thom Blog: Thoughts on Film Sound Design. Available at:

https://randythomblog.wordpress.com/2019/04/11/screenwriting-for-sound/ [Accesed 02 Oct 2020]

Thom, R. (2019d). *Script Analysis for Sound Design, Raymond Chandler's Script "Playback," First Set of Scenes* [Online]. Randy Thom Blog: Thoughts on Film Sound Design. Available at: https://randythomblog.wordpress.com/2019/12/23/script-analysis-for-sound-design/ [Accesed 28 Oct 2020]

Thompson, A. (2018). *Roma Drives Netflix to Break Its Own Rules About Theatrical Release* [Online]. Indiwire. Available at: https://www.indiewire.com/2018/10/roma-netflix-rules-theatrical-release-oscars-1202016876/ [Accessed 11 Apr 2019].

Thompson, K. (2013). *GRAVITY, part 1: two characters adrift in an experimental film* [Online], Observations on film art. Available at: http://www.davidbordwell.net/blog/2013/11/07/gravity-part-1-two-characters-adrift-in-an-experimental-film/ [Accessed 24 Sep 2018].

Udden, J. (2009). Child of the Long Take: Alfonso Cuaron's Film Aesthetics in the Shadow of Globalization. *Style*, 43 (1): 26-44.

Vincie, A. (2013). *Preparing for Takeoff. Preproduction for the Independent Filmmaker*. New York and London: Taylor & Francis.

Whittington, W. (2007). *Sound Design & Science Fiction*. Austin: University of Texas Press.

Whittington, W. (2011). Sound Design for a Found Future: Alfonso Cuarón's Children of Men. *New Review of Film and Television Studies,* 9(01): 3-14.

Wilkinson, C. (2013). *The Working Film Director: How to Arrive, Survive & Thrive in the Director's Chair*. 2nd Ed. Studio City: Michael Wiese Productions.

Woodhall, W. (2014). *Pre-Production Audio for Post Production Sound: Make a Plan to Capture Great Sound* [Online]. ProVideo Coallition. Availible at: https://www.provideocoalition.com/pre-production-audio-for-post-production-sound/ [Accesed 11 Jan 2019]

— (2016). *Creating the sounds for The Revenant* [Online]. Provideo Coallition. Available at: https://www.provideocoalition.com/creating-the-sounds-for-the-revenant/ [Accessed: 05 Sep. 2018].

Wright, B. (2013). What do we Hear? The Pluralism of Sound Design in Hollywood Sound Production. *The New Soundtrack*, 3(2), 137-157.

Wurtzler, S. (1992). "She sang live, but the microphone was turned off:" The live, the recorded and the subject of representation. In R. Altman, (Ed.). *Sound theory sound practice*. London: Routledge, pp. 87-103.

Zhang, W. et al. (2017). Surround by Sound: A Review of Spatial Audio Recording and Reproduction. *Applied Sciences*, 7(05), 1-19.