Section 3

Immersion
10. The Building and Blurring of Worlds

Sound, Space, and Complex Narrative Cinema

Justin Horton


DOI: 10.5117/9789089647566/CH10

Abstract

Sound plays a crucial role in the construction of cinematic worlds by evoking spaces beyond the edges of the frame. In this essay, I examine the relationship between sound and image as it relates to offscreen space in more or less conventional narratives before turning my attention to a complex example, Christopher Nolan’s *Inception* (2010). Rather than one cinematic world, this film offers us as many as five, and sound—both at a formal and a narrative level—mediates them all. I demonstrate how the sound design in *Inception* not only helps to build its cinematic worlds but also to blur the boundaries between them.

Keywords: Sound; Cinematic Space; Puzzle Films; Narrative Complexity; *Inception*

There has been, in recent years, a pronounced shift within Media Studies away from the more time-honored concerns of plot, character, and action and towards “world,” reflecting a number of intersecting factors, among them, the increasing prevalence of transmedial narratives that feature a common universe but that “unfold across multiple media platforms” (Jenkins 2006, 334); the renewed popularity of science-fiction and fantasy, genres often noted for vast imaginary worlds and intricate mythologies (Thompson 2008); and the impact of computers on how we engage with and understand narratives and media alike (Murray 1998; Kinder 2002; Manovich 2002; Bordwell 2006; Cameron 2008). Whereas stories tend toward narrative economy by offering only information that is crucial to the plot, worlds brim over with textual detail that contributes a verisimilitude
that “absorbs” or “immerses” the reader, gamer, or viewer in them (Wolf 2012).

But what does a world sound like, and what role, if any, does sound play in its construction? Sound, like “world,” cuts across genre and medium, turning up in a diverse array of forms both new and old, from radio and television to video games and new media art installations. This remarkable heterogeneity of sound in media, however, prohibits a thorough consideration of its various incarnations in the space of a single essay. Therefore, in what follows, I will limit my focus to the context of film and how sound contributes, in ways that often escape notice, to the spectator’s mental representation of imaginary worlds and, often, to her seeming “integration” or “immersion” into them. Sound is especially well suited to this latter point, as Frances Dyson notes:

[S]ound is the immersive medium par excellence. Three-dimensional, interactive, and synesthetic, perceived in the here and now of an embodied space, sound returns to the listener the very same qualities that media mediates: that feeling of being here and now, of experiencing oneself as engulfed, enveloped, absorbed, enmeshed, in short, immersed in an environment. (Dyson 2009, 4)

But our mapping and immersion in cinematic space through sound is generally a function of hearing sounds that belongs to the world of the film. What are we to make, then, of music, noises, and speech that seem to hover around the screen, at a remove from the onscreen world? Such indeterminate sounds tend to evade the intellect or cast doubt on one’s visual certainty: after all, seeing is believing, as the saying goes.

In this chapter, I will consider first the ways in which sound aids in the construction of worlds in the minds of spectators, especially its evocation of offscreen space, which suggests parts of the world that lie outside the frame of our immediate visual apprehension. Second, I turn my attention to surround-sound technologies and how they “place” the spectator in the film world by collapsing the distance between it and the movie theater or one’s living room. Lastly, in the essay’s final third, I offer a brief analysis of Inception (Nolan 2010), a “complex” film that explicitly thematizes the processes of world building and that integrates sound in intriguing ways into its elaborate plot machinations. That is to say, Inception creates not simply a world, but worlds—plural—, using sound as a conduit between them.
Scenographic Space and the Mental Construction of Cinematic Worlds

How does the cinema construct and then place a viewer in its world? Most commercial fiction filmmaking around the globe is modelled, at least in part, on the Hollywood system of continuity, which first organizes and edits disparate shots together to create the suggestion of seamless space and time, while carefully prompting the spectator to make certain inferences and hypotheses about this cinematic world. As David Bordwell, Janet Staiger, and Kristin Thompson explain, cinematic continuity is a representational system based on various perspectival traditions as well as a narrational system. The continuity system, they explain, “turns a remarkably coherent spatial system into the vehicle for narrative causality” as “the viewer […] meets the film halfway and completes the illusion of seeing an integral fictional space” (Bordwell, Staiger, and Thompson 1985, 59). According to their study, this system was standardized by 1917, more than a decade before the coming of synchronized sound. Therefore, for the time being, we shall look to how films build worlds visually before turning our attention to sound and how it complements and, at times, complicates the seeming coherence of cinematic worlds.

At its core, the Hollywood-style system of continuity is a set of devices that, when combined, unify scenographic space and time while simultaneously effacing the evidence of its own constructedness—covering its own tracks, if you will. The result is a highly contrived yet seemingly natural presentation “organized for the spectator” who is “[kept] […] at the center of every image” (Andrew 1976, 147). One such device is the establishing shot, which, in many ways, anchors the entire system. An establishing shot is a long or medium-long shot that, as the name implies, establishes the locale of the scene and, most often, the position of characters relative to one another within it. From this base shot, filmmakers may “cut in” to closer framings.

Two principles guide these cuts: the first, which relates to a film’s narrative logic, is motivation; that is, the closer shot reveals some aspect of the scene that is pertinent to the story and that might not be readily apparent from the wider view. The second principle is what is broadly known as “matching.” One such type is the “match-on-action,” wherein a movement begun in one shot is continued in the next shot, reassuring viewers that both shots are continuous with one another and that, even though there has been a visual cut, the time and place remain consistent. The human eye tends to gravitate towards motion, and this perceptual predilection works to “hide” the cut. Another common matching convention is the “eyeline
match,” wherein we see a character glance over to something offscreen and then, in the subsequent shot, we see what she sees (though most commonly not from her “optical point of view”). Crucially, the space revealed may be of something present in the establishing shot or of a bordering space that the viewer has not yet seen. Extending the logic of the eyeline match is the shot/reverse shot, in which the viewpoint alternates between two figures. In each case, the changes in angle are not random; rather, they are motivated by a character’s look. These devices incrementally reveal parts of the imaginary world of the film (its “diegesis”), but only in a fragmentary way.\footnote{1}

The viewer’s task is to take these partial views and reconstruct them into an abstract mental image, a “cognitive map,” of a unified whole. But there are other ways in which films visually suggest spaces that are never shown directly. Noël Burch argues that there are, in fact, two distinct types of cinematic space: that which is in the frame—what he calls “concrete space”—and that which is outside it or “imaginary”. Doorways, for example, are a common means by which imaginary space is suggested beyond the concrete. If a character enters or exits a scene via a door situated on either side of the frame, then it follows that the filmic world stretches beyond that which is currently visible (Burch 1981, 17). Such an evocation of space is put to remarkably effective use in Fritz Lang’s M (1931). In its first sequence, we see young Elsie (Inge Landgut) bouncing a ball when she encounters Hans Beckert (Peter Lorre), a child serial killer who lures the young girl into his clutches with the purchase of a balloon. In one shot, we see Elsie and Beckert together before they exit the frame. Moments later, the film’s construction implies the girl’s death in two consecutive shots: first, at a previously unseen wooded location, the ball that belonged to Elsie rolls into the frame from offscreen right; and second, a shot of the balloon hung in the lattice of power lines before the wind carries it away and out of the frame to the left. The movements of objects associated with Elsie into and out of the frame here imply not only a larger world than what the screen can contain, but also, poignantly, a movement from visibility to invisibility, presence to absence, and life to death.

It bears repeating that each of the aforementioned approaches to off-screen space is conveyed visually. The question that remains is how film sound accomplishes a similar task. Evoking offscreen space, however, was not the first impulse for filmmakers early in the transition era. Rather, dialogue, even more than music and noises, ruled the day. After all, the silent cinema was never, in fact, silent, as it was often accompanied by live music, narrators, and sound effects produced in the auditorium itself (Altman 2004, 193-194). This is why sync-sound films came to be known as “talkies” rather than, say, “soundies.”
But the introduction of dialogue inaugurated its own set of challenges. First, was the concern over always visualizing the speaker, for to hear a voice absent a visualized source has been, throughout history, a source of a certain anxiety. The “disembodied voice”, as it is often called, is frequently aligned with the supernatural or otherworldly, from the voice of god to ghostly hauntings (Chion 1999, 23-26; Connor 2000, 410; Kane 2014, 150). Thus, early in the transition era, practitioners felt compelled to pair the sound of a voice with an image of the speaker’s moving lips. Secondly, this practice was abetted by the technological limitations such as microphones of limited range and cameras housed in heavy soundproof booths that masked their noisy operation. Combined, these two tendencies worked to constrain both the movement of actors and the ability to cut within the scene.

Over time, filmmakers discovered that audiences were far more adaptable to the interplay of sound and image that they might have initially assumed, and recording technology (and technician ingenuity) soon overcame these initial limitations. The lesson: sound and image, though synchronized, need not necessarily be “redundant”: the image is not duty-bound to show what the soundtrack says, and vice versa. Filmmakers thus began to exploit offscreen sounds to their advantage. One crucial innovation was the “voice-off,” wherein a character not seen in the shot can nevertheless be heard. Thus, much like the eyeline match with unseen objects or characters, the voice-off convention depends on viewer inferences about areas offscreen. Take the first few shots of Erin Brockovich (Soderbergh, 2000), wherein sound precedes the images by several seconds. Over a black screen, we hear a male voice pose a question to a female who answers back. When the image appears, we see only the titular character (Julia Roberts). One infers quickly that the situation is a job interview, but, for nearly the first minute of the film, the interviewer goes unseen, his voice and the movement of his arms in the bottom left corner of the screen are all that indicate his presence and the as-yet unrevealed remainder of the office. Here, sound provides a larger degree of spatial information about the film’s world than does the image. Such is the power of the voice-off, a device that Mary Ann Doane argues deepens the diegesis, gives it an extent which exceeds that of the image, and thus supports the claim that there is space in the fictional world which the camera does not register. [...] The voice-off is a sound which is first and foremost in the service of the film’s construction of space and only indirectly in the service of the image. It validates both what the screen reveals of the diegesis and what it conceals. (Doane 1980, 40)
Per Doane, the voice-off is as much a matter of spatial construction as it is of anxiety over bodily coherence. But Doane’s theorization makes the aural building of the diegetic world entirely a function of voice, neglecting music and noise, a position that reflects what Michel Chion calls “vococentrism” (Chion 1994, 4), his terms for the tendency of critics, filmmakers, scholars, and viewers to privilege speech over and above other aspects of film sound. Chion provides a counterexample: during interior scenes in *The Sacrifice* (Tarkovsky, 1986), birds outside are continuously heard despite them never being visualized or otherwise discussed by the characters. These “territory sounds”, as he calls them, suggest unrepresented exterior spaces at the same time that they “fill in” or enrich the image with hints of a fully realized world (Chion, 1994, 124). A more narratively pertinent example can be found in the aforementioned *M*, when Beckert realizes he is at risk through hearing the sound of sirens in the distance. In a sound film, the converging police needn’t necessarily be shown, but are merely suggested.

A guiding assumption to this point has been that the sounds we hear while watching a film are diegetic, that they are *in* and *of* the onscreen world. This is not always the case, however. We will consider *nondiegetic sound* in a later section. In the meantime, we shall turn away from how sound aids in the viewer’s mental construction of space to how it seemingly places her in the filmic world.

**Sound Technology and Immersion**

In many cases, film sound not only suggests a world but also immerses us in it. But in order to understand how, we must first attend briefly to the history of film sound technology. With few exceptions, from the 1920s to well into the 1970s, sync-sound films were presented monophonically, with the sound issuing from a single loudspeaker (or “channel”) placed behind the screen at its horizontal center (Kerins 2010, 329). Within the representational logic of the cinema, even if a sound occurs offscreen—a gunshot, say—, in mono presentation it actually reaches the viewer from straight on. No matter from where the sound is implied to originate within the film world, it always emits from the screen. The viewer therefore must ignore this incongruity in order to make the cinematic illusion cohere. Our mental representation of space is thus flexible enough to account for the disparity, willfully ignoring it so as to ensure the consistency of the imaginary space (Altman 1980, 71).
Dolby Stereo, which was introduced in the mid-1970s, quickly supplanted mono as the standard for the film sound for over two decades and utilized four distinct channels: three (left, center, right) at the screen and one “surround” loudspeaker situated away from the screen. The result, according to Chion, was that “The space of the film [was] no longer confined to the screen [and] became the entire auditorium”—what he calls the “superfield” (Chion 1994, 151). This development is a crucial one, for it makes manifest “a reversal of cinematic hierarchy”:

where historically it has been the responsibility of the image to explain the soundtrack by visually confirming the sources of sounds, it is not the soundtrack that provides the context of the image. [...] Expanding beyond the edges of the frame, the sonic world literally becomes bigger than the image, which reveals only a small fraction of the [diegetic] world. [...] The soundtrack now assumes the task of guiding the audience and creating a coherent narrative space. (Kerins 2010, 86, emphasis in original)

According to Chion, the surround channel of Dolby Stereo anchors space through the stable presentation of ambient sound, erasing, in many cases, the need for an establishing shot.

Mark Kerins argues, though, that the superfield is less apt a characterization of Digital Surround Sound (DSS), which grew to prominence in the 1990s. Though there are several proprietary loudspeaker configurations with DSS, they all share, at minimum, what is called a 5.1 arrangement: three front channels (left, center, right), two surround channels (left and right), and a low-frequency channel that operates at one-tenth the range of the other loudspeakers (thus the .1) and that is responsible for producing bass-y, low-end notes that often are palpably felt as much as heard, eliciting “the very physical sensation of being at an earth-shaking event” (ibid, 330, 134, emphasis in original). Kerins modifies Chion’s term with his own—the “ultrafield”—that accounts for the tendency of DSS sound designers to “constantly shift sounds around the multi-channel environment”. The ultrafield, writes Kerins,

seeks not to provide a continuous aural environment, but rather to continuously provide an accurate spatial environment where aural and visual space match. In short, the ultrafield is the three-dimensional sonic environment of the digital world, continuously reoriented to match the camera’s visual perspective.² (ibid, 92, emphasis in original)
When we speak of “immersion” and cinema, we have to account not simply for the world it presents visually but also sonically. Moreover, we must also be attentive to the technologies used to design and exhibit a particular film. That is to say, the degree or kind of aural immersion on offer in Tim Burton’s *Batman* (1989), a movie from the Dolby Stereo era, is quite different from a DSS-era film like *Taken* (Morel, 2008). Furthermore, we must not presume ideal conditions for every viewer: many films still carry a Dolby Stereo track in addition to their surround track for theaters that are not equipped with cutting-edge sound systems. Similarly, a viewer watching a DSS film at home on the puny speakers of a stereo television or through earbuds attached to a laptop will not experience the same sort of immersion as one in a multiplex environment. The question is thus not so much what one hears, but also from where, and in what sort of space.

The Elsewhere and the Otherworldly: Nondiegetic Sound

To this point, we have examined the various ways in which films structure space first with images, then with sounds. With regard to the latter, however, we have largely limited our discussion to sounds of and in the filmic world. But what are we to make of sounds that originate elsewhere?

A key distinction must first be drawn between diegetic and nondiegetic sounds. Diegetic sounds originate in the world of the film, and they may be divided into two categories: external and internal. External diegetic sounds belong to the world of the film and are audible to the characters within them. In the famous scene in *When Harry Met Sally…* (Reiner 1989) when Sally (Meg Ryan) loudly fakes an orgasm during lunch at a busy New York City diner, all parties hear it, to great comedic effect. This contrasts with internal diegetic sounds, which might best be thought of as entirely interior, heard only “in the head” of a single character (and, of course, the viewer). Films with a running interior monologue such as *Bridget Jones’s Diary* (Maguire 2001) are built upon this internal diegetic principle.

In contrast, nondiegetic sounds are heard only by the viewer and not by the inhabitants of the film world, and music is perhaps its most immediately recognized form. For example, as the Stallone character runs up the steps of the Philadelphia Museum of Art at the end of the training montage in *Rocky* (Avildsen 1976), no one would presume that he hears the theme music that we do. The song is thus intended “for” the viewer, priming her emotionally with its triumphant refrain. Yet the matter of music’s relation to the diegesis is not so cut-and-dried, for it frequently slides between these two levels, often
unobtrusively. In the conclusion of Richard Linklater’s *Dazed and Confused* (Linklater 1993), for instance, Mitch (Wiley Wiggins) listens to Foghat’s “Slow Ride” via headphones, which we initially hear as muddled and thin but that gradually grows louder and fuller before carrying over into the next scene, which takes place in an entirely different location. The music is no longer “sourced” within the world of the film, having become entirely nondiegetic.

But if this music does not belong to the world of the film, from whence does it come? Nondiegetic sound emerges from an uncertain place, a *somewhere else* beyond the film world that exists alongside it, a co-existence of planes, if you will. Nondiegetic sounds, then, are not merely “off” (i.e., not visualized), but “over,” at a remove from the film’s world, yet seemingly encompassing it. A perhaps familiar nondiegetic device is the voiceover, which Sarah Kozloff helpfully contrasts with the previously discussed voice-off:

“Over” actually implies more than mere screen-absence [“voice-off”]; ... [V]oice-*over* is distinguishable by the fact that one could not display the speaker by adjusting the camera’s position in the pictured story space; instead, the voice comes from another time and space. (Kozloff 1989, 3, emphasis in original)

For Pascal Bonitzer, the voiceover’s elsewhere “represents a power [...] from a place which is absolutely other (from that inscribed in the image-track)—absolutely other and absolutely undertermined” (Bonitzer 1980, 133-134, emphasis in original). Bonitzer is here referring to the particular case of documentary voiceover, where the unseen, interpreting narrator’s indeterminacy carries the weight of divine certainty. Yet, with fiction films, this is not always the case. Some voiceovers come with a clear narrative frame, such as in *Stand By Me*, where the primary story of young Gordie (Will Wheaton) and his friends’ search for the corpse of a missing boy is narrated by his adult self (Richard Dreyfuss), whom we see in the film’s bookended sections recollecting these same events as he prepares to pen his memoir. More complex deployments of the device also exist: famously, in *Sunset Blvd.* (Wilder 1950), a dead man narrates the story of his murder from beyond the grave, while in *Daughters of the Dust* (Dash 1991), a child not yet born provides the voiceover. *The Thin Red Line* (Malick 1998) utilizes voiceovers from four characters, and it is even difficult at times to discern which is speaking, hence throwing into question exactly how the sentiments expressed are to be interpreted (Millington). Moreover, it is hard to tell whether what we hear is in fact a voiceover or a momentary lapse into internal diegetic sound (i.e. a character’s thoughts as they are happening.
within the time and place of the diegesis). Such is the potentially slippery terrain of the voiceover specifically and of nondiegetic sound generally.

For our purposes, we can say that when it comes to world building, sound functions to imply spaces beyond the frame, yes, but also realms beyond the primary world. Thought of in this way, sound may be deployed in such a way as to build not simply a world, but worlds—plural. In fact, filmmakers increasingly mine the possibilities of sound to build intricate, multiplicitous worlds, especially in the terrain of “complex” narrative cinema, to which we now turn our attention.

Complex Films, Complex Worlds

We have seen how sound builds worlds in films that more or less fit within a classical paradigm: traditional, Aristotelian stories that are organized and presented in accord with Hollywood conventions. However, the 1990s saw a trend towards narrative complexity in cinema that has since been labelled “puzzle films” (Buckland 2009), “mind-game movies” (Elsaesser 2008), or “delirium cinema” (Pisters 2012, 40). Said to depart gleefully from classical norms of straightforward narrative progression in an effort to perplex spectators through an “embrace of nonlinearity, time loops, and fragmented spatio-temporal reality [...] [t]hese films are riddled with gaps, deception, labyrinthine structures, ambiguity, and overt coincidences” (Buckland 2009, 6).

An example of such narrative complexity is Inception, a cinematic Gordian Knot if ever there was one. What’s most intriguing about the film, given the context of the present essay, is that it thematizes in quite literal fashion the act of world building. Moreover, sound—both diegetic and nondiegetic—plays a critical and polyvalent role within this Baroque structure, resulting in a densely layered sound design that mirrors that of the film’s nested dream worlds.

The plot is set in motion when Cobb (Leonardo DiCaprio), an “extractor” paid to steal ideas from the minds of sleeping victims through a process he calls “shared dreaming”, is hired to do the opposite: to instill an idea into someone’s mind. Unlike more straightforward dream theft, inception requires far greater access to the mark’s subconscious, thereby necessitating not one dream world for them to share, but three. The film thus adopts a multi-tiered, dream-within-a-dream structure, as Cobb and his team delve deep into the mind of their victim, Fischer (Cillian Murphy), to plant the notion for him to break apart and sell his dying father’s empire, a move that would benefit Saito (Ken Watanabe), who is, not coincidentally, Cobb’s employer.
In order for this scheme to be successful, each of the dream levels must be sufficiently replete with detail and texture so as not to tip off Fischer of their falsity. Enter Ariadne (Ellen Page), a brilliant architecture student tasked with constructing worlds both rich and plausible that will serve as the staging ground for the group’s deception of Fischer. But Ariadne is new to the notion of “shared dreaming” and requires direction from Cobb, a tutorial that forms one of the film’s most spectacular set pieces. He follows Ariadne as she walks through a Paris-set dream world, the first of her own design. There, she reimagines not only the topography of the cityscape on the fly, with bridges and doors morphing and materializing around her as she moves about the imaginary terrain, but also the physical laws by which the world operates. For instance, she literally folds Paris upon itself at one
point, bending one half of it onto the other. *Inception* is thus an imaginary world that offers a self-referential commentary on the act of world building and, indeed, upon its own construction, adding yet another layer to the film’s Russian Doll structure.

As Ariadne investigates the contours of this Paris of her own making, she remarks, “I thought the dream space would be all about the visual, but it’s more about the feel of it.” This is an astute point, for the entire ruse hinges not on the eye, but on the ear and on bodily sensation. That is to say, in order to execute a three-tiered dream, the team must rely on heavy sedation. The tricky part, however, is waking up. To that end, Yusuf (Dileep Rao), a chemist, formulates a powerful tranquilizer that leaves “inner ear function unimpaired”. Why the ear? First, the apparatus of the inner ear—called in anatomical terms a “labyrinth,” a delicious coincidence in the current context of narrative complexity—is a key component of the body’s vestibular system, which contributes to one’s sense of balance, spatial orientation, and proprioception (one’s sense of one’s own body during movement). Thus, even when deeply asleep, the dreamers will be susceptible to the real-world sensation of falling, activated by fierce jolts or being toppled over in a chair—what Cobb and company dub a “kick”. But, in order to successfully remove the team from the nested dream world, the kick must occur in both the real world and in each dream level. To pull this off, the group relies on a piece of music, Edith Piaf’s “Non, Je ne regrette rien” as a means of synchronizing the dream worlds and their dilated temporality with the real-world kick. Put differently, the multiple kicks are timed to a song that, despite the dreamers’ intense sedation, they *continue to hear* via headphones. And the kick does not merely wake the sleeper; it also causes the dream world(s) to collapse. Such is the case in the film’s final kick, wherein Ariadne falls backwards in the fourth (unplanned) dream level before cascading in reverse through each of the other levels, which shatter and disassemble with her impact. The ear, therefore, as both an organ for hearing and for bodily comportment, is central to the film’s nested structure and how it makes (and then unmakes) its multiple worlds.

This nesting, in fact, transpires at the nondiegetic level as well. In a revelation that caused quite a stir amongst the film’s fans shortly after its release, one perceptive viewer noted that the deep, brass-y leitmotif in Hans Zimmer’s score actually corresponds to the horn intro of Piaf’s tune—what Zimmer described after the fact as his score’s “DNA” (Michaels 2009, n.p.)—if slowed down and protracted (Whitehead 2010, n.p.), a clever nod to the temporal expansion that occurs at each dream level. This note, which shifts between diegetic and nondiegetic in ways that might not be
at first apparent, is fundamental not only to the construction of the film’s score but is also the means by which its elaborately built worlds are bound together.\textsuperscript{4}

Much of the discourse surrounding narrative worlds has focused on immersion as a function of their sheer expansiveness and their abundance of textual detail. But, rather than the frequently deployed horizontal spatial metaphors of “vastness” and “extensiveness,” I propose that we might also think of complexity in terms of verticality. Indeed, this is the very architecture of the sound film: an image of a world laid over with sounds that correspond to that world, but also nondiegetic sounds suggestive of an elsewhere—or multiple elsewheres—that envelops or is layered upon the primary world. Indeed, at the level of its narrative construction, \textit{Inception} explicitly tackles the idea of worlds within worlds. But, cunningly, so too does its aural design, which traffics in nondiegetic sound’s tendency towards indeterminacy and its ability to layer and move among ever more worlds.

\textbf{Sound: World Bridging as World Building}

The question pondered in this essay—what is the role of sound in the building of cinematic worlds?—has demanded a multipronged answer, one that touches on film conventions of spatial organization, both visual and aural, and the historical development thereof; the role of offscreen space in the viewer’s mental construction of a coherent imaginary world; and how changes in film sound technology complement this cognitive image with an immersive, physical experience of this space. However, immersion and the aural evocation of offscreen space are most often functions of diegetic sound, the sounds in and of the story world. What are we to make of nondiegetic sounds that are temporally or spatially separated from the visualized world and that its inhabitants cannot hear? In many cases, nondiegetic sound bears no “existential” relation to the story world, belonging instead to the film’s “discourse.” Yet, in others, it suggests a realm, a time, or a place that is affiliated with but distinct from the “primary” world onscreen. Sound maintains a privileged connection to these indeterminate spaces, often moving subtly into and out of the diegesis, passing between realms. Thus, a film need not necessarily go to the lengths \textit{Inception} does to be “complex,” nor does it require a world as vast as Middle-earth, nor must it parcel out its narrative across multiple media platforms, for the relations of sound to image are sufficiently intricate in and of themselves.
Notes

1. The notion of diegesis originates with Plato and Aristotle and was first imported into film theory with the “filmology” movement that emerged in the 1940s. See: Souriau, 1951. See also: Introduction.

2. Several scholars have noted this break with (or modification of) traditional continuity editing, especially with regard to action films from the 1990s onward, but few apart from Kerins have integrated sound into their account of how or why the spatial organization of contemporary cinema has undergone this shift. See: Bordwell, 2002; Stork, 2011; Shaviro, 2010 and 2012.

3. By sounds of the “primary world”, I mean those that are unambiguously sourced within the diegesis and not “over” or “aside” it. Thus, my usage differs from Mark J.P. Wolf’s notion of “Primary World” as the physical surroundings of the viewer, reader, or gamer that tend to be “displaced” by the secondary, imaginary world when she is “absorbed” in a narrative. See: Wolf’s essay elsewhere in this collection.

4. This is, of course, not the only extratextual reference to Piaf in the film, as Cobb’s wife is played by Marion Cotillard, who in fact played Piaf in 2007’s La Vie en Rose (Dahan).

Works Cited


Whitehead, Cameron. 2010. “*Inception* Music Comparison.” *YouTube*. July 22. www.youtube.com/watch?v=UVkQoC4qDvM.


**Media cited**


About the author

Justin Horton is a visiting lecturer in the Moving Image Studies program at Georgia State University. His research specializations are in the areas of film theory and aesthetics, sound studies, and affect theory. His work has appeared in *Cinephile* and twice in *Cinema Journal*. He is currently at work on a book manuscript entitled *Bodies that Scatter: Sound, Cinema, Figure, Form*. 