Revisiting the Canon of Sound Theory

Michel Chion, *Audio-Vision: Sound on Screen*. Second Edition, trans. Claudia Gorbman (New York: Columbia University Press, 2019).

Audio-Vision, Sound on Screen was first published in French in 1990 and translated to English in 1994. For being one of the first attempts to address sound in film beyond the technological perspective, the book became widely referenced among scholars and practitioners. Twenty-five years later, it has been republished by Columbia University Press.

The author Michel Chion (1947), a French film critic and composer, began his career as an assistant to Pierre Schaeffer in 1970, at the *Groupe de Recherches Musicales* (GRM).[1] This led him to collaborate with the French journal *Cahiers du cinéma* between 1982–87 which, together with the arrival of videotape, triggered a book trilogy about sound in film: *La voix au cinéma* (1982; *The Voice in Cinema* in 1999)[2] in which he analyzes the hierarchy of voice in the soundtrack ("vococentrism" — p. 6); *Le son au cinema* (1985; *Film, a Sound Art* in 2003),[3] where he continues to access the language in film-sound; and *La toile trouée. La parole au cinéma* (1988),[4] in which he accesses the development of sound-film through its dialogues.[5]

Audio-Vision: Sound on Screen is a compilation of this trilogy. It describes the relationship between sound and image and establishes the core terminology of sound for film, which has been frequently cited since then. At the time, there was no antecedent, which is partly due to a general difficulty in speaking *objectively* about sound. Thus, as formulated in the foreword by Walter Murch, the terminology very often recurs to ordinary language and therefore is self-explanatory. For example, "points of synchronization" (p. 36), which are the moments in which image and sound meet in time; or "extension" (p. 84), which is sound's capacity to extend the perception of the space outside the visual frame.

Very often, Chion extends his explanations to other subjects, grasping different fields without going beyond the surface. For example, when he attempts to analyze a certain psychology of hearing in film, by discussing "influences of sound on perceptions of movement and speed" (p. 11), or "on perception of time in the image" (p. 12). Or, when he states that "textual speech has been considerably dis-

---

1) The *Groupe de Recherches Musicales* (GRM) is a musical research center created by Pierre Schaeffer in 1958. It became official part of the Research Service of French Radio and Television in 1960, and then in 1975 integrated The *Institut national de l'audiovisuel* (INA).

2) Michel Chion, *The Voice in Cinema*, trans. Claudia Gorbman (New York: Columbia University Press, 1999). *La voix au cinéma* has been translated into Czech by Josef Fulka – Michel Chion, *Hlas ve filmu*, přel. Josef Fulka (Praha: NAMU, 2020).

3) Michel Chion, *Film, a Sound Art*, trans. Claudia Gorbman (New York: Columbia University Press, 2003).

4) Michel Chion, *La toile trouée — la parole au cinéma* (Paris: Editions de l'Etoile, 1988).

5) Parts of this review are derived from an interview I did with Michel Chion for the journal *ArteActa*.

cussed and theorized, often considered as an annex or outgrowth of literature" (p. 157). Where, when or by whom — it is not said. He speaks from a position in which he rarely makes any reference to other studies (often saying "many scholars" or "some theorists"), making it hard to understand his contribution.

Whereas that was already somewhat critical in the first edition of his book, it is certainly not addressed in the new edition. Most of the time, Chion does not theorize about sound per se but about elements that happen to be audible: voice, speech, text, music. He focuses on the impact such usage has in the overall language of film, rather than sound as a language in itself. His attention belongs to the filmmakers rather than the sound-makers (Walter Murch being the exception).[6]

In this line, *Audio-Vision* proposes an analysis of "soundtrack" based on three elements (speech, noise and music) in two parts: "The Audiovisual Contract", which is the foundation of "audio-vision" itself; and the segment "Beyond Sounds and Images", which comprises the concepts of "superfield", considers noise and sound effects, proposes the "audio-logo-vision" and explains the audiovisual analysis itself.

The Audiovisual Contract

"The Audiovisual Contract" claims that sound and image are mutually influenced and that "the audiovisual relationship is not natural but rather a sort of symbolic pact to which the audio-spectator agrees when she or he considers the elements of sound and image to be participating in one and the same entity or world."[7] But this "pact" is the core definition of fiction itself: etymologically, fiction implies the making or manufacturing of something; it presupposes a creator and it requires "the willing suspension of disbelief", which suggests that it is not completely detached from reality despite being a construct.[8] This "Contract" is, then, a principle that goes beyond audiovisual fiction.

The first chapter launches Chion's understanding of the relationship between sound and image. "The projection of sound on image" creates the "added value", which is the "expressive and informative value with which a sound enriches a given image so as to create the definite impression, in the immediate or remembered experience one has of it, that this information or expression 'naturally' comes from what is seen, and is already contained in the image itself" (p. 5). Later on, *Who Framed Roger Rabbit* (Robert Zemeckis, 1988) is posed as an example of such: the sounds of Jessica make her believable to the audience (p. 116).

In this line, the second chapter proposes "three listening modes": causal, semantic and reduced. Chion's listening modes are themselves a variation of Schaeffer's original listening modes in *Treatise on Musical Objects* (2007).[9] Schaeffer enumerates four listening modes: écouter — to listen; *ouïr* — to

6) An excellent counterpoint is Vincent LoBrutto, *Sound-on-film: Interviews with Creators of Film Sound* (Santa Barbara: Greenwood Publishing Group, 1994).

7) This definition is in the original version of the glossary and it is not included in the revised version. Michel Chion, *Audio-Vision: Sound on Screen*, trans. Claudia Gorbman (New York: Columbia University Press, 1994), 222.

8) Samuel Taylor Coleridge, *Biographia Literaria; or Biographical Sketches of My Literary Life and Opinions* (London: Rest Fenner, 1817).

9) Pierre Schaeffer, *Treatise on Musical Objects: An Essay across Disciplines*, trans. John Dack and Christine North (Berkeley: University of California Press, 2017).

perceive aurally; *entendre* — to hear, and *comprendre* — to understand; which Chion had already attempted at clarifying in *Guide des objets sonores: Pierre Schaeffer et la recherche musicale* (1983).[10]

If Chion's first two modes are self-explanatory (one understands the cause of what one hears, and one understands the meaning of what one hears — both even if not seen), the idea of reduced listening in an audiovisual context is quite elaborate. The "reduced listening mode" is imported from Pierre Schaeffer's early studies, where theories of acousmatic music were first drafted.[11] In its turn, the term acousmatic is borrowed from a long mythology that may or may not have started with Pythagoras.[12] Either way, the definition of "reduced listening" is entangled with the concept of "acousmatic" sound: in the same way that an acousmatic sound is a sound of which one does not see the source, reduced listening focuses on the sonic traits of the sound itself, regardless of its source. In other words, both emphasize the listening experience.

The third chapter is dedicated to "(…) perspectives on audiovisual relations" (p. 35), referring to the audiovisual counterpoint (or dissonance) as the possibility to avoid linear interpretation or redundancy between audio and image. Tarkovsky's *Solaris* (1972) is given as one such case (*free counterpoint* — p. 38). This chapter also explores the lack of "unity" in sound editing, by comparing the visual frame to the (im)possibility of an "auditory shot" (p. 41). While images establish a frame and they cannot be superimposed or cut without notice, sound can be over layered, edited and manipulated without notice. Again, this statement can be refuted because this is not absolutely true in all circumstances. First, it is a choice of the makers: one can choose to superimpose two sounds that clearly do not belong together in the same time or space within the narrative. By the same token, the sound designer decides what is included in the "auditory frame" or not. Therefore, the sound boundaries are set just as much as the camera sets the frame.

In Chion's opinion, Godard is one of the few cases in which sound jump-cuts are part of cinema's rhetoric. In this line, he also mentions the "elements of auditory setting", which considers some sound conventions established along the years in sound-film. These are sounds that punctuate the scenes without the audio-viewer's awareness; they *naturally* belong there (dogs barking at a distance, for example). Finally, he introduces "synchresis", which is "a word forged by combining *synchronism* and *synthesis*" — a phenomenon dependent on "contextual determinations" (p. 64). In fact, upon Chion's visit to FAMU (Prague, 2019) he mentioned *Daisies* (Věra Chytilová, 1966) as an example of "synchresis". The first scene in the movie (after the opening credits) illustrates this phenomenon quite accurately: a squeaking (wood) sound in synchronization with the arms of the two girls moving; a forged relationship between sound and image portraying their *puppetness*.

The fourth chapter continues to explore the idea that sound lacks the possibility of a frame. He calls this *spatial* relationship "the audiovisual scene". The analogy is based on the fact that the image frame is a container in itself, claiming there is no such comparison in sound (for sound knows no boundaries). Although true in a first instance, this idea is the subject of widespread debate and largely

10) John Dack and Christine North have offered an unofficial translation to English available online since 2009. Michel Chion, *Guide to Sound Objects*, trans. John Dack and Christine North, 2009, accessed March 19, 2021, https://monoskop.org/images/0/01/Chion_Michel_Guide_To_Sound_Objects_Pierre_Schaeffer_and_Musical_Research.pdf.

11) Pierre Schaeffer, *In Search of a Concrete Music*, trans. John Dack and Christine North (Berkeley: University of California Press, 2013).

12) See Brian Kane, *Sound Unseen: Acousmatic Sound in Theory and Practice* (Oxford: Oxford University Press, 2014).

discussed in theories of sound, space and amplification, amongst many other fields.[13] Nevertheless, this is the point of departure for discussing "localization" of sound and, consequently, acousmatic theory. For this, Chion borrows a concept from *musique concrète* mostly to make a distinction between "visualized" and "not visualized" (acousmatic) sound. In this way, he divides "onscreen, offscreen and nondiegetic sound" (p. 73). To these categories, he then adds "ambient, internal and on the air" sounds (pp. 75–78). These considerations then lead to another widely cited term: "the point of audition" (p. 88). Although he claims image and sound are not comparable, the whole concept of "point of audition" is developed as an analogy to the image (point of view): "the analysis of this complex question shows that *point of audition* is not parallel to the notion of *point of view*, because of profound differences between sound and image and between seeing and hearing" (p. 209). In that line, the point of audition is "the point in space from where we hear the sound" (p. 209).

Following that analogy, the fifth chapter arrives at yet another concept that became standard in sound studies, that of "rendered sound" (p. 98). It recurs again to the idea that the audiovisual relationship is not a natural one. Accordingly, this is because sound is manipulated and artificially added to the image. It is also due to a disjunctive ideology: the illusion of unity. And why is unity an illusion? Because there is a "philosophical dualism of body and soul: the voice on the soundtrack, the visual aspect in the image" (p. 251). Instead of elaborating on this, Chion offers a few considerations about the technical evolution of the theaters and sound amplification. He considers "definition" vs. "fidelity", in order to arrive at the idea of "reproduction" vs. "rendering". That is, "the film spectator recognizes sounds to be truthful, effective, and fitting not so much if they reproduce what would be heard in the same situation in reality, but if they render (convey, express) the feelings associated with the situation" (p. 108). In other words, render is the ability of cinema in general to convey an experience, rather than reporting it. But in any case, the experience is not an experience of the *real* but a *fictional* one. Later on, Chion mentions E. H. Gombrich's *Art and Illusion* as "wholly relevant" to rendering but does not elaborate further (p. 180).

Last, in the seventh chapter the "phantom body" finalizes this path through the audiovisual contract. The phantom body itself is an analogy to the concept of "acousmêtre", which had been proposed in *The Voice of Cinema*. The acousmêtre is a specific character whose existence relies specifically on sound. It is not just a narrator, it is a body deliberately hidden, invisible and on the "verge of appearing" (p. 127). This voice not only illustrates the power of sound but also a "*primary identification* with the camera" and a power of "*textual speech*" (p. 150). The acousmêtre also shelters the concept of "suspension", which revolves also around absence. In short, suspension combines expectation and suspense.

Beyond Sounds and Images

The second part of the book is much shorter and approaches the subject matter in a less thorough and systematic way. It launches on considerations about direct sound, noises and sound amplification, towards the concept of "superfield" — "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 loud-

---

13) See for instance Denis Smalley, "Space-Form and the Acousmatic Image," *Organised Sound* 12, no. 1 (2007), 35–58.

speakers outside the physical boundaries of the screen" (p. 143), which is not so different of the previously mentioned "extension".

The second chapter of this part (eighth of the whole book) is heavily dedicated to text (speech, voice). First, throughout three modes of speech: *theatrical, textual* and *emanation* (pp. 148-166). In the second case, we have the exception "wandering text" (p. 154) — a relativized *verbal chiaroscuro* (p. 159). Additionally, Chion considers the "endless integration", which is basically cinema's capacity to keep up to date and integrate novelty into its own conventions (p. 166). Although this chapter is called "toward an audio-logo-visual poetics", the concept of "audio-logo-vision" is never explained.

Last, Chion proposes "an introduction to audiovisual analysis", which "aims to understand the ways in which a sequence or whole film works in its use of sound combined with its use of images" (p. 172). This practice is vastly used in his sound courses, often posing two simple questions: "what do I see of what I hear" and "what do I hear of what I see" (p. 181). Throughout this introduction, the exercises also lead to specific concepts, which helps understanding their implementation in a slightly more practical way. Following this, the segment finishes with an excerpt analysis of *La dolce vita* (Federico Fellini, 1960) and *The Silence* (Ingmar Bergman, 1963).

2019 Audio-Vision

The new edition consists of a light language revision, in order to avoid a certain vagueness but probably also resulting from new findings through the several years in which his lectures were repeated. Here and there we denote a little contextualization, such as a larger introduction to "methods of observation in audiovisual analysis" (p. 174), or referring to multiple screen sizes, smartphones and internet streaming. A few times he also addresses previous critiques of the book (e.g. p. 74). There is a slight update on the discussions, as for instance the spatial treatment concerning multitrack recording (although not completely incorporating contemporary deeds, such as 5.1. or Dolby ATMOS). There are also new film references, for those that did not exist at the time of the first edition but do fit the argument. For example, Spike Jonze's *Her* (2013) provides a clear understanding of a "disembodied" voice, the "acousmêtre" (p. 126), but also *Taxi Driver* (Martin Scorsese, 1976) is now contextualized as an example for *phantom sound/deacousmatized* voice (p. 181).

While the core structure of the book is the same, the three listening modes have been reformulated. Now, causal listening is subdivided into "figurative" and "detective" (p. 22). Respectively, they concern *diegetic* and *profilmic reality*. This distinction is not really developed any further, nor it is explained why it is necessary. On the same note, Chion opts for the term "codal listening" instead of semantic (p. 25), evincing that the meaning is not in the sound itself but in the symbols of the narrative. Finally, to reduce listening he adds the "pivot-dimensions among sound categories", a property shared by sound elements of different categories (p. 30) — in which he fits a paragraph about sound design and another on phenomenology of perception.

The segment about the "point of audition" has also been developed further. Analogous to the image, "point of audition" pertains to spatial or subjective designations. However, Chion argues that it is not possible to establish such positioning for sound's omnidirectional nature (both in the source that is projected in all directions and the listener who can listen from any position in space). To this end, this part has been extended and reorganized with a particular focus on a list of "telephemes" (p. 91). Accordingly, "point of audition" is not really possible because there is no such thing as a *symbolic mi-*

crophone; that is, a subjective and metaphoric perspective on sound (p. 96). And yet, point of audition can be *objective* or *subjective* (p. 209).

In the second part of the book "Beyond Sounds and Images", the whole chapter on "television, video art and music video" has been excluded, possibly due to its expanding theories since Chion first published *Audio-Vision*. However, a segment was added to the textual categories: "the opposite noniconogenic narration (…) literally narration that does not create images" (p. 156). Additionally, there is also a whole new segment in which Chion approaches the difference between "said and shown" (p. 167), accounting to six of them: *scansion, contrast, contradiction, counterpoint, c/omission* and *sensory naming*.[14] Finally, in this new revision Chion offers a "Chronology — Landmarks of the Sound Film" (p. 215).

The legacy

Needless to say, "sound studies" as a field has blossomed immensely in the past two decades with an increasing number of studies and publications from different perspectives.[15] In comparison, Chion's vision of sound film is very limited. For instance, when stating that "many people consider location sound (direct sound) not only the sole morally acceptable solution in filmmaking but also the simplest, since it eliminates the problem of having to make choices" (p. 105). I would argue that location sound is not simple at all: it is technically very demanding and full of challenges. It is also not clear who these "many people" are, and what is moral about it. For example, ADR (Automated Dialog Replacement) is an extremely common technique and many times the director's first choice. Most importantly, it is necessary to understand that choosing a microphone is a complex chain of multiple variables that will definitively initiate the whole path of sound design.

Chion claims that "a cinema-specific vocabulary for sound isn't available" (p. 175) and hence his attempt to create one; but he often recurs to terms that actually have a specific meaning in sound terminology, and quite differently from what he is trying to convey (e.g., rarefaction, masking…). The same happens, for example, when he re-adjusts the listening modes. Now, the *causal-detective* listening mode comprises a brief mention of foley, but it seems completely disconnected from the practice or technique itself: foley should never be heard as such. Good foley work means the sounds blend in with the action and therefore seem recorded at the same time as the image. So, who would be watching a camp-fire scene and wondering how the sounds of the fire were produced (p. 21)? It should go by unnoticed.

---

14) Although not mentioned, "Show vs. Tell" is a concept initially proposed by Anton Chekhov, see Percy Lubbock, *The Craft of Fiction* (New York: Charles Scribner's Sons., 1921) and Anton Pavlovich Chekhov, *The Unknown Chekhov: Stories and Other Writings Hitherto Untranslated*, trans. Avrahm Yarmolinsky (New York: Noonday Press, 1954).

15) On sound for film, see for example: Jay Beck and Tony Grajeda, *Lowering the Boom: Critical Studies in Film Sound* (Champaign: University of Illinois Press, 2008), or Vanessa Ament, *The Foley Grail: The Art of Performing Sound for Film, Games, and Animation* (Waltham: Focal Press, 2009). On a broader approach see for example Seth Kim-Cohen, *In the Blink of an Ear: Toward a Non-Cochlear Sonic Art* (New York: Bloomsbury Publishing, 2009) or Brandon LaBelle, *Background Noise, Second Edition: Perspectives on Sound Art* (New York: Bloomsbury Publishing, 2015).

Having said that, sound for film owes him much of the attention it got since he started writing about it, and the crossover between *musique concrète* and sound-design is still much needed. *Audio-Vision: Sound on Screen* is a legacy to all sound theorists since its first edition.

Sara Pinheiro

Bibliography

Ament, Vanessa. *The Foley Grail: The Art of Performing Sound for Film, Games, and Animation* (Waltham: Focal Press, 2009).

Beck, Jay, and Tony Grajeda. *Lowering the Boom: Critical Studies in Film Sound* (Champaign: University of Illinois Press, 2008).

Chekhov, Anton Pavlovich. *The Unknown Chekhov: Stories and Other Writings Hitherto Untranslated*, trans. Avrahm Yarmolinsky (New York: Noonday Press, 1954).

Chion, Michel. *Hlas ve filmu*, přel. Josef Fulka (Praha: NAMU, 2020).

Chion, Michel. *Audio-Vision: Sound on Screen*, trans. Claudia Gorbman (New York: Columbia University Press, 1994), 222.

Chion, Michel. *Film, a Sound Art*, trans. Claudia Gorbman (New York: Columbia University Press, 2003).

Chion, Michel. *Guide to Sound Objects*, trans. John Dack and Christine North, 2009, accessed March 19, 2021, https://monoskop.org/images/0/01/Chion_Michel_Guide_To_Sound_Objects_Pierre_Schaeffer_and_Musical_Research.pdf.

Chion, Michel. *La toile trouée — la parole au cinéma* (Paris: Editions de l'Etoile, 1988).

Chion, Michel. *The Voice in Cinema*, trans. Claudia Gorbman (New York: Columbia University Press, 1999).

Kane, Brian. *Sound Unseen: Acousmatic Sound in Theory and Practice* (Oxford: Oxford University Press, 2014).

Kim-Cohen, Seth. *In the Blink of an Ear: Toward a Non-Cochlear Sonic Art* (New York: Bloomsbury Publishing, 2009).

LaBelle, Brandon. *Background Noise, Second Edition: Perspectives on Sound Art* (New York: Bloomsbury Publishing, 2015).

LoBrutto, Vincent. *Sound-on-film: Interviews with Creators of Film Sound* (Santa Barbara: Greenwood Publishing Group, 1994).

Lubbock, Percy. *The Craft of Fiction* (New York: Charles Scribner's Sons., 1921).

Schaeffer, Pierre. *In Search of a Concrete Music*, trans. John Dack and Christine North (Berkeley: University of California Press, 2013).

Schaeffer, Pierre. *Treatise on Musical Objects: An Essay across Disciplines*, trans. John Dack and Christine North (Berkeley: University of California Press, 2017).

Smalley, Denis. "Space-Form and the Acousmatic Image," *Organised Sound* 12, no. 1 (2007), 35–58.

Taylor Coleridge, Samuel. *Biographia Literaria; or Biographical Sketches of My Literary Life and Opinions* (London: Rest Fenner, 1817).