



## Voices without bodies 2

### Description

I'm reprising the post of 30 November 2013 called [Voices without bodies](#). The blog continued the reflections in my book [The Tuning of Place: Sociable Spaces and Pervasive Digital Media](#) published by MIT Press in 2010. Here's the original with some updates assisted by ChatGPT.

Question to **Siri**:

What's *The Wizard of Oz* about?

Siri: It's about some Dorothy, her intelligent assistants, and her little dog too. Some are not so intelligent, I guess.

Update: The development of speech recognition and speech synthesis on smartphones at the time exemplified by Siri brought into focus how important the voice is in helping people feel engaged with the devices they carry. Even then, this had less to do with convenience than with fantasy and science fiction, which in turn is about the city and film.

What now seems quaint is not the detachment of the voice, but its simplicity. Today's large language models generate extended dialogue, sustained tone, and context-sensitive responses. The voice is no longer merely reactive; it appears conversational, reflective, and at times uncannily poised. Yet the fantasy remains the same. (That ends the update comment.)

All the world's a stage, wrote Shakespeare. Baroque architects arranged the urban environment as if it were a stage set. Now city inhabitants are more likely to experience the world as if through a cinematic frame. The miniature screens we carry around, along with animated billboards, urban screens, and projections, reinforce the sense that the world is mediated and assessed as if it were a film at least some of the time. Alongside the visuals, we carry soundtracks and voices.

The sound theorist **Michel Chion** bases his studies of voice in cinema on the concept of the *acousmêtre*: the acousmatic being, a voice whose source is invisible and unknown. In film this is the voice of the narrator, someone off-stage, or behind the curtain. For Chion, the *acousmêtre* conveys

ubiquity, panopticism, omniscience, and omnipotence – in other words, authority.

The **acousmêtre** is all-seeing, all-knowing, and all-powerful. Voice-over narration in documentaries and news broadcasts appears to inherit its authority from this tendency. As a hapless acousmêtre, the fated Wizard of Oz concealed his modest human frame behind a screen, using voice alone to project authority he otherwise lacked.

Chion also refers to the “already visualised acousmêtre”: a voice whose face we know, but do not currently see, and which often projects reassurance. The voice of a friend on the telephone has this function and, in 2013, so did the soothing, indefatigable quips of Siri.

Update: Today, this reassurance is no longer tied to a single assistant or device. Voices generated by large language models are cloned, customised, and multiplied. They may sound like friends, presenters, or even oneself. The **acousmêtre is no longer singular; it is reproducible.** (Back to the original.)

The “commentator acousmêtre” is another category: a disembodied voice with no personal stake in the action. Comparable to the narrator of a news item or nature programme, it is also evident in public address announcements, whether live or recorded. This is close to what Chion calls the “radio acousmêtre”: a voice that does not have the option of showing its face.

The “complete acousmêtre” is the voice to which we cannot yet attach a face, but which remains liable to appear in the visual field at any moment. This is a powerful cinematic device. Directors play with the possibility of revelation, and with the suspense this creates – what Chion calls the “epiphany of the acousmêtre.” Alfred Hitchcock’s *Psycho* is exemplary here: the acousmêtre brings disequilibrium and tension.

There is an amusing scene in *The Big Bang Theory* in which Raj appears to meet Siri – an impossibility, of course. The tension lies in whether Raj can speak back (considering his pathological shyness).

Update: In retrospect, the joke is revealing. What was once impossible – speaking back fluently, at length, and with apparent reciprocity – has now become routine. The acousmêtre not only answers; it elaborates. (End of update.)

## The grafted voice

To interact with Siri in 2013 required an up-to-date phone, a current operating system, and a live network connection. The voice was grafted on, and sometimes dropped out. We are accustomed to this with voices, which are frequently detached from the bodies they serve.

Cinema has long exploited this separation. Film rarely uses voices recorded at the same moment as images. Voices are added later, adjusted, spatialised, and synchronised in the sound studio. This post-synchronisation reveals the disjunction between voices and bodies, sounds and images. In some cases the connection is deliberately loosened, as in Fellini’s use of voices that “hang on the bodies of actors in only the loosest and freest sense, in space as well as in time.”

Update: Digital cinema cameras are now effectively silent, eliminating the constant background noise that once made clean on-set recording difficult. Directional microphones, radio lavaliers, and compact multichannel field recorders have improved dramatically, allowing dialogue to be captured with high signal-to-noise ratios even in complex environments. Post-production tools can now remove background noise, isolate voices, and repair damaged audio in ways that were previously impossible or prohibitively expensive.

As a result, **many films now prioritise live production sound** and design their shoots around capturing usable dialogue on the day. Actors are often encouraged to treat the take as vocally final rather than provisional. (That ends the update.)

Cinema depends on false attributions of vocal agency: dubbing, Foley editing, simulated vocal apparatus. Siri's adjustable gender and accent, and its occasional muteness, already hinted at this cinematic treatment of the voice.

Update: Today, voice cloning, synthetic presenters, and AI-generated narration extend this logic. Voices are no longer merely detached from bodies; they are transferable, iterable, and temporally displaced. (Back to the original.)

## Siri-alism

Algorithms simulate voices. Digital devices read documents aloud, announce nearby events, report on software processes, and provide navigational instructions. The detached voice is also menacing. In *2001: A Space Odyssey*, Stanley Kubrick presents the breakdown of the computer **HAL 9000** as an acousmatic event. HAL is an all-seeing, insistent voice. His dismantling is accompanied by a regression into childish speech — a nursery rhyme — revealing his digital substrate. For Chion, this is — a strange death, leaving no trace, no body, and no echo. •

Update: The contemporary twist is that today's systems do not die so neatly. Voices persist across updates, platforms, and deployments. They fade, fork, and reappear. The acousmatre no longer vanishes; it lingers. (End of update.)

The vehicle for this speculation remains the acousmatre — a concept that recognises the voice as already disembodied, floating, cut, looped, muted, and grafted.

Question to Siri:

— What's *2001: A Space Odyssey* about? •

Siri: — It's about an assistant named HAL who tries to make contact with a higher intelligence. These two guys get in the way and mess it all up. •

Update: In 2025, the joke turns back on us. The assistants are no longer merely in the way; they are now participants in the telling. (End of update.)

## Notes

- It looks as though I am stuck with this test avatar recorded outdoors via handheld smartphone over 20 seconds until I subscribe to the next level of HeyGen avatar creation. I'm also committed to sounding too much like an Australian hey ho!
- The HeyGen agent also made recommendations and sourced related slide material. It didn't comply with my request to remove these. It seems that HeyGen inferred the relevance of such B-roll content from the script I supplied.
- This post was originally published at the conclusion of the symposium *What is sound design?* at the University of Edinburgh.
- It adapts material from a chapter written with Martin Parker on voice and space.
  - Update: At the time, neither Siri nor contemporary generative AI systems existed. The conceptual apparatus, however, remains apt. (End of update comment.)
- Page references are to Chion's *The Voice in Cinema*.
- Slavoj Žižek observes that voice does not simply accompany the image, but points toward a gap in the visible toward what eludes the gaze. The voice marks absence. Update comment: **That gap has not closed with AI voices; it has multiplied.**

## References

- Michel Chion, *Audio-Vision: Sound on Screen*.
- Michel Chion, *The Voice in Cinema*.
- Richard Coyne, Richard Coyne, *The Tuning of Place: Sociable Spaces and Pervasive Digital Media*.
- Richard Coyne and Martin Parker, *Sounding Off: The Place of Voice in Ubiquitous Digital Media*.
- Richard Coyne and Martin Parker, *Voices Out of Place: Voice and Non-Place in Mobile Communication*.
- Richard Coyne and Martin Parker, *Voice and Space: The Agency of the Acousmêtre in Spatial Design*.
- Slavoj Žižek, *Gaze and Voice as Love Objects*.

## Category

1. Artificial Intelligence
2. Podcast
3. Voice and text

## Tags

1. acousmetre
2. voice

## Date Created

January 24, 2026

---

**Author**  
rcoyne99

*default watermark*