

David Trippett

Human Sounds: the Obscenity of Information

9

In 2017, Alexander Payne’s film *Downsizing* pursued an old thought experiment: what if humans could be shrunk, their perceptual worlds miniaturized, their bodies made Lilliputian all of an instant? How would they reexperience the space of the environment, of sound and light? In Payne’s film, the lead character, played by Matt Damon, opts in to a government program whereby he is to be shrunk by a factor of 2,744, along with thousands of others, to form an experimental “miniature” community in an attempt to solve the climate crisis. His wife—emerging as a science-fiction skeptic—opts out of the downsizing process at the last minute, leaving Damon’s character to ponder the wisdom of his irreversible choice.

With its comedic focus on relationships, the film skirts questions of realism, that is, the unfamiliar, “real” perceptual world such a shrunken “human” might experience.ⁱ

Miniaturized people, including Damon, retain their deep voices, their perceptual ranges and acuity, and their sensory proprioception; they feel no change of atmospheric pressure, gain no insight into the newly massified world around them, and can still engage with their unminiaturized human interlocutors at will. Within such narrow dimensions, the only “loss” is body mass; a narrative fact rather than an experiential postulate.ⁱⁱ All other parameters of existence remain stable—in keeping with a genealogy of minihuman films from *The Incredible Shrinking Man* (1957) to *Ant-Man* (2015)—presenting a shallow fiction that leaves perceptual questions unasked.

In abstract terms, body morphology can be taken as a contested object, an assertion of matter in its relation to identity, one that places into relief the character of the relationship between perception, self-perception, and objects. “My ‘own’ body is material,” Jane Bennett

asserts, “and yet this vital materiality is not fully or exclusively human,” for it depends on myriad microscopic bacteria (“swarms of foreigners”) that neither are human nor can be perceived by the naked eye (“a nested set of microbiomes”) (2010: 112–13). In this context, the relation between media and realism is anchored by human perception, which in turn is rendered idiosyncratic by virtue of an individual’s unique sense apparatus, acuity, plasticity, individual history, and training. Historically, the body has been figured as an object of difference in contexts as divergent as Thomas Aquinas’s Dominican theology, where physical things become individualized through “matter signed with quantity” (rather than form; see Funkenstein 1986: 135), to Jean Baudrillard’s poststructuralist critique of media, with its three orders of simulacra—a classical order of “counterfeit,” an industrial order of “indefinite reproducibility,” and a digital order of “simulation”—where this third order denies the possibility of counterfeiting an original body matter, flattening out uniqueness in favor of fractal bodies, or “models from which all forms proceed according to modulated differences” (Baudrillard [1976] 2017: 77). In the context of digital film, duly enmeshed in Baudrillard’s narrative of simulation, modeled perception (simulation of what it is like to hear or see as a different being) brings the argument full circle, where body morphology itself—that whose uniqueness becomes endangered by means of its simulation—demands individuality of matter via the very medium that denies it this, whether as a gnat, a crocodile, or a miniature Matt Damon. “Everything began with objects,” Baudrillard once remarked, in a parody of *Genesis*. “Yet there is no longer a system of objects” (Baudrillard 1988: 11). Understanding of the body both as unique in its perceptual apparatus and as a unique configuration of matter no longer has any meaning in digital representation, he infers, which makes the question of realism in film redundant by definition: “real” is forever referable to a subject position constructed by the audiovisual technology. First published in 1987, this statement about the role of digital information in society formed part of Baudrillard’s

submission for the “habilitation” at the Sorbonne. Originally titled *L’Autre par lui-même* and translated into English as *The Ecstasy of Communication*, it bore the mildly sarcastic title “Habilitation” and was nearly rejected—perhaps because of what some have described as the misogynistic overtones of his concept of seduction and the argument’s recurrent allusions to pornography and the sexual body, perhaps because of the brevity of its claims (the original French edition is barely 92 small pages in length). For present purposes, what is remarkable about it is the book’s articulation of an enduring confrontation between the human body and digital media at a time when digital screen media was—by today’s standards—in its infancy. Two decades into the 21st century and the replication of voices through neural networks, the concept of the deepfake and of the legal ownership of singing holograms from Hatsune Miku to Maria Callas registers a further shift in the relation of mediatized appearance to the putatively real, a shift whose technological details Baudrillard could hardly have envisaged. Across this divide, the injured concept—the possibility of a unique identity—remains recalcitrant, perhaps because most witnesses to these simulations would still regard themselves as individuals.

To be sure, it now seems unsurprising for a postmodern philosopher in the 1980s to signal as casualties the principles of a reality beyond the play of appearances, the existence of unique individual subjects, claims for a truth or a metaphysics that persists. But for Baudrillard in 1988 the loss of these concepts is positioned historically in relation to the radical increase of *matterless data* that accompanied screen media, summarized in the metaphor of the screen’s flat surface. It was superficiality made literal. On the face of it, it was as though these older cultural tropes had somehow been given up *recently* in response to the proliferation of cathode-ray TVs and the digital audio of Sony’s PCM-1 encoder: “*Today* the scene and the mirror have given way to the screen and the network. There is no longer any transcendence of depth,” he writes defiantly, “but only the immanent surface of

operations unfolding, the smooth and functional surface of communication” (1988: 12). If these were technological affordances, they were thoroughly unwelcome. Trapped on this infinite surface, we learn, Baudrillard’s subject reflexively inhabits an overly transparent world in which aesthetic experience becomes entirely soluble in information streams, a world saturated in digital signs and their instantaneous networks; this environment creates a historically unprecedented identity that rebounds on the subject, who—unable to “produce the limits of his being,” unable to produce her- or himself as a mirror—becomes “a pure screen, a pure absorption and re-absorption surface of the influent networks” (27). More than a transformation, this represents a cold loss of identity, whose familiar grain of tangibility is no longer valid.

Dystopian rhetoric aside, Baudrillard’s claim that simulation is built on a world of code has proven influential.ⁱⁱⁱ It asserts that contemporary culture can be coded into ones and zeros, that “digitality is among us. It haunts all the messages and signs of our society” (Baudrillard [1976] 2017: 82). Alongside a darkening worldview in which knowledge and the very processes of thought are subsumed within data flows, the social control implied by a coded environment sets up a formidable political adversary that—for Baudrillard—must be resisted: “You can’t fight the code with political economy, nor with ‘revolution’...can we fight DNA?...Perhaps death and death alone, the reversibility of death, belongs to a higher order than the code” (25). Such rhetoric betrays a concrete reality in the late 1990s, one that fed the anxiety of identity loss implied by digital media. While the rhetoric of the deepfake was decades away, genetic cloning was a crisp, new technology. Across multiple essays Baudrillard rails specifically against human cloning, which he posits as code applied to no-longer-unique bodies: “The Father and Mother have disappeared...in the service of a *matrix called code*” (Baudrillard 2010: 96). The genetic formula inscribed into cells undoes the body’s physical reality by virtue of its potential for infinite replication; hence the simple fact

of DNA (“the prosthesis par excellence”) transforms it into a simulation. In this guise, the body becomes an assemblage of virtual quantities, “a stockpile of information and of messages, a fodder for data processing... The individual is no longer anything but a cancerous metastasis of its base formula” (99–100). In its multiple iterations, this verdict passes through sarcasm (“it allows complex beings to achieve the destiny of protozoas”—96) to protest (“without the Other as mirror, as reflecting surface, consciousness of self is threatened with irradiation in the void”—140), turning finally to moral outrage, where the “subtle death” of doubling constitutes an innate self-destructiveness or “the transparency of evil” to which—*pace* Jean-Paul Sartre—even “the hell of other people would have been preferable” (Baudrillard 1993: 139). Denuded of aura, the individual is fatally cheapened in the process of simulation. With a nod to Sophocles’s *Oedipus*, this coding of bodies is “still incest, but without the tragedy” (138).

It is the question of identity that forms the red thread in the discussion of digitality or quantification that follows. After a critique of Baudrillard’s historical technologies, this chapter identifies realism as a philosophical proposition that, in the digital age, has become synonymous with the relation of quantity to technology, a relation that has deep historical roots from the microscope in the mid-17th century to chronophotography in the 1890s and “imperceptible” pixels-per-inch in the 2010s. The calculation of perceptual difference, by applying the ratio of different body sizes to frequencies (whether in fiction or in acoustics), offers an explanation as to why frequency resolution has become a central parameter for realism in speech synthesis, technologies that synthesize the spoken voice, in a closing case study that indicates the extent to which identity and voice are no longer uniquely bonded nor primarily referable to physical bodies.

Obscenity, or dots on a line

Before pursuing this critique in the context of speech synthesis and its modes of simulation, two concepts that shape Baudrillard's understanding of digital media bear some consideration. The first is *obscenity*; the second, *communication*. Both are central to his short book *Ecstasy of Communication* and undergird the anxieties provoked by mediatized information circa 1987.

The primary definition of obscenity in the *Oxford English Dictionary* is that which is “offensive or grossly indecent, lewd.”^{iv} Etymologically, it is often linked to the Latin *caenum* (“filth”). But the etymology of the term is unclear. A disputed reading of the Latin grammarian Marcus Terentius Varro's *De lingua Latina* from the first century BC led to the notion that *scaena* (from *ob-scaenum*) could refer to the stage, where the indecent content in classical plays, content that is offensive to the eyes of the gods, should take place offstage or concealed from public view. This would include the classical acts of moral outrage: Medea's enraged murder of her sons sired by Jason, after he abandons her for a younger princess, in Euripides's *Medea*; Tarquin's rape of Lucretia in Ovid's poem; and the aforementioned *Oedipus*, where at the end of the play the protagonist must gouge out his own eyes. In a literal sense, obscene—in this “folk” etymology—came to mean that which eradicates our gaze. While this etymology is almost certainly wrong,^v it is precisely the definition that Baudrillard inverts in his critique of screen media.

Obscenity begins where there is no more spectacle, no more stage, no more theatre, no more illusion, when every-thing becomes immediately transparent, visible, exposed in the raw and inexorable light of information and communication. / *We no longer partake of the drama of alienation, but are in the ecstasy of communication.* And this ecstasy is obscene. Obscene is that which eliminates the gaze, the image and every representation. (Baudrillard 1988: 22)

Here, far from concealment, the gesture of the obscene is that of zooming up infinitely close, of seeing so clearly as to be indistinguishable with what is taking place. In erasing the gap

between what is taking place and acts of witnessing, it abolishes all schemes of representation, all space for mystery or hermeneutics. Significantly, it can accomplish this only through declarative, technological means: code and pixel density. For each time Baudrillard mentions obscenity, he loops back to the agencies of digital “information and communication.” Admittedly, these terms remain undetermined, and with the benefit of hindsight, we might simply read his statements as an undeveloped intellectual position, a way station en route to his more mature work on simulation and the hyperreal. But against such teleology, it also reverts to the fantasy of encoding personal experience and aspects of personal identity—our “private universe” that hitherto had been a secretive matter. It is in this spirit that, riffing on common associations of obscenity, he clarified:

Obscenity is not confined to sexuality, because today there is a pornography of information and communication, a pornography of circuits and networks, and objects in their legibility... It is no longer the obscenity of the hidden, the repressed, the obscure, but that of the visible, the all-too-visible, the more-visible-than-visible; it is the obscenity of that which no longer contains a secret and is entirely soluble in information and communication.

(Baudrillard 1988: 22)

If this metaphor emanates from a technical capacity for visual close-ups that decontextualize for screen viewers explicit views of what is not supposed to be seen, it sits within an established genealogy of technological affordances for sensory perception, including in the electronic manipulation of sound.^{vi} Of course, the visual affect of extreme close-ups is multivalent, and can equally be enlisted to deny the obscenity of a totalizing, data-rich sensation (‘more-visible-than-visible’) that Baudrillard has in mind. For as Lisa Marks reminds us, the electronic effects of pixelation in close-ups often draw attention to texture rather than realism, creating perceptions of haptic or tactile images through blurring or other manipulation of the underlying bitmap (Marks 2000: 176).

But it takes a moment to remind ourselves that in the 1980s, when these sentences were being written, virtual artifacts—putatively immaterial objects like digital holograms, 3D modeling, and cyberspace—had barely been invented; according to Martin Hilbert and

Priscila López, less than one percent of the world's media storage capacity was digital in 1986; by 2007 it would become 94 percent (2011: 60–65). And the first practical video coding format, discrete cosine transform (DCT)—initially proposed as an image compression technique in 1972—was only adopted for compressing online video in 1988, the year of Baudrillard's English translation. If we step outside Baudrillard's philosophical frame, then, the question arises as to what form of data is at issue when a writer such as this refers to “the raw and inexorable light of information.” In the context of realism, one answer is *resolution*. Here, realism is posed as a quantitative proposition: what density of pixels or bit depth is needed to fully dissolve the secrets of aesthetic experience in “information”? Bits, unlike atoms, have no mass, color, or size and can travel at the speed of light. They are *symbols*, commonly considered as ones and zeros, to be set and reset as declarative assertions with no capacity for ambiguity, and once famously described by Nicholas Negroponte as “a state of being: on or off, true or false, up or down, in or out, black or white” (1995: 14). For sound, they are synonymous with audio sample rates: a 12-bit sampler outputs 12 bits of data for every sample, and the sonic resolution relates to the number of samples per second. As such, they are units of a symbolic sonic existence—capable of higher and lower “definition”—that challenge the singular authenticity of acoustic sound.

Since its first use in 1936, the term *high definition* has shifted continuously from its origins in the number of lines of an analog TV screen to audiovisual media, from 8K imagery to so-called “lossless” sound. It is unnecessary to rehearse this history; suffice it to say that the upper limit of high-resolution screen technology has reached a surface of at least 220 million pixels on supercomputers in San Diego and at least 192,000 samples per second for audio recordings at 24-bit depth.^{vii} While this represents a significant increase in relation to what preceded it in 1987, there is no reason to assume it has reached an absolute limit. For a philosophy of perception, however, it is indicative that the quantitative argument has

prevailed: detractors to recent marketing strategies, such as Apple's "retina" display (a branding tool for screens of putatively higher pixel density than the cellular organization in the eye at given viewing distances), have implicitly accepted the quantitative realism on offer. For Raymond Soneira, President of DisplayMate Technologies, 477 pixels per inch would be needed at a viewing distance of 12 inches (Steve Jobs had asserted circa 300 ppi).^{viii} Technology journalist John Brownlee judged similarly: "Apple's Retina Displays are only about 33% of the way there" (2012). By protesting against the degree required, both tacitly accepted the notion that ever greater density will ultimately render media screens indistinguishable from vision, what Jonathan Sterne once called 'the dream of verisimilitude' (Sterne 2012: 4). But given the limits of perceptual mechanisms in human eyes and ears that were determined in the 19th century by the likes of Thomas Young and Rudolf König based on a wave theory of light and sound, increases of resolution do not lead to a hypothetical endpoint, the fabled hyperreality, where it is impossible to distinguish sound samples from real voices at the level of sensation. A multimodal sensorium combined with nondeterministic cognitive processing cannot be accounted for in a mathematical mapping of physiology, of cell onto pixel, cochlear hair onto audio sample. As early as 1994, Michel Chion argued that the definition of a sound signal, rather than any correspondence to reality, is what creates a 'hyperreal effect' for listeners, citing the habit in sound recording of using 'more treble than would be heard in a real situation' (Chion 1994: 98-99). And as the editors of a more recent volume on stereo observe, stereophonic listening and fidelity were never "synonymous or even fully coterminous," reliant on logics and practical needs that construct listener positions with no obligation to what is taken to be quantitatively "true" (Théberge, Devine and Everett 2015: 27). My claim, insofar as this discussion permits one, is that a historical perspective indicates that the sensation of the real, like the unreal or simulation, has no technological correlate. It remains a philosophical proposition.

If we cast a glance back in history, the principle is essentially that of *dots on a line*. As such, it can be explained more clearly in relation to its original formation in the paradoxes of the pre-Alexandrian philosopher Zeno of Elea (c. 490–c. 430 BC), whose argument against pure motion is equally applicable to that against a realism defined by quantity. Recorded in Aristotle’s *Physics*, four paradoxes ascribed to Zeno concern the relations of time and motion. Zeno’s arrow—the third paradox—characterizes apparently opposed states of existence, both of which are exclusively true yet mutually contradictory: An arrow in flight is always traveling; yet at any given point in time it is stationary. Although it is always in motion, it cannot have time to move unless it is permitted more than one instant—that is, permitted to occupy at least two successive positions. At any given moment, therefore, the arrow is at rest, motionless at each point in its swift course. No matter how many individual, static moments accrue, even an explosion of pointillist speckles could never equate to pure motion as perceived by the eye witnessing the arrow in flight.

This paradox was co-opted in the early 20th century by another French philosopher enthralled by the implications of the screen to explain the illusion of understanding duration, or any linear process of becoming. For Henri Bergson in 1907, parallel technologies resulting from developments in chronophotography in the 1880s, such as Edison’s kinetoscope and the Lumière brothers’ cinématographe, provided a new visual basis for interpreting the mind and its cognitive processes for relating sense acuity to an environment. Preoccupied above all by the necessities of action, Bergson asserts, “The intellect, like the senses, is limited to taking, at intervals, views that are instantaneous and by that very fact immobile of the becoming of matter” (Bergson 2005: 224). In Bergson’s view, this illusion is embodied in the paradigm of moving-image technology, whose rapidly successive still pictures gave the impression of motion to audiences through the agency of the revolving mechanism. [Figure 9.1](#), taken from Etienne-Jules Marey’s study *Cycliste* (c. 1894), illustrates the principle of chronophotography

that Bergson had in mind. The paradox illuminated by such seemingly mobile images, as Bergson recognized, was that of Zeno:

In order that the pictures may be animated, there must be movement somewhere. The movement does indeed exist here; it is in the apparatus. It is because the film of the cinematograph unrolls, bringing in turn the different photographs of the scene to continue each other, that each actor of the scene recovers his mobility...Such is the contrivance of the cinematograph. And such is also that of our knowledge. Instead of attaching ourselves to the inner becoming of things, we place ourselves outside them in order to recompose their becoming artificially. We take snapshots, as it were, of the passing reality, and, as these are characteristic of the reality, we have only to string them on a becoming, abstract, uniform and invisible, situated at the back of the apparatus of knowledge, in order to imitate what there is that is characteristic in this becoming itself. Perception, intellection, language proceed so in general. Whether we would think becoming, or express it, or even perceive it, we hardly do anything else than set going a kind of cinematograph inside us.
(Bergson 2005: 251–52).

This practice of interpreting physical or mental functions in terms of tools or technical apparatus may resonate with more recent narratives of technogenesis (that humans coevolved with tools and technologies, where interior thought relates dynamically to exterior technicity) that we associate with thinkers such as Katherine Hayles (2012) and Bernhard Siegert (2003), particularly in light of recent re-evaluations of Ernst Kapp's pioneering *Elements of a Philosophy of Technology* (1877). But historically, interpreting cognitive functions in terms of new mechanical paradigms was widespread in the wake of the *Exposition universelle* of 1889 and 1900. Bergson's insight into the cinématographe mechanism was, however, to highlight not our ability but our failure truly to understand motion as a metaphor for becoming or duration. He regarded as foolish attempts to realize pure motion simply by intensifying the artifice, that is, by increasing the speed of cylinder rotation (or "resolution"), thereby making the intervals between states infinitely smaller: "Before the intervening movement you will always experience the disappointment of the child who tries by clapping his hands together to crush the smoke" (Bergson 2005: 254). As this quip illustrates, the principle of pure continuity was not simply a matter of quantity, or intensity of "information." Continuity, or real lived experience, is of a different order to concatenated

instants (spatialized as pixilation), and hence simulations. In 1907, just as now, these were not mathematically relatable to human perception.

Accepting such “foolishnesss,” our instinctive trust of sensory feedback ensures that one enduring definition of realism is the degree to which a simulated object *resembles* its real-world object for sentient perception.^{ix} For digital visual media, this indexical relation pertains for games and computer-generated images.^x For sonic media, putatively authentic sound samples are indexical to real-world sounds, but, by implication, to their sound sources also. This implies an interface between sensation and response, whereby listeners become aware of some level of cognitive response beyond cognizing the bare sensory stimulus. That is to say, I am aware of my reaction to the sound *of* the voice calling me—what the American semiotician John Deely has termed the *species expressae*: “the cognitive response of the organism to the cognitive experience of a stimulus” (1994: 134).^{xi} Discourses on realism in auditory gaming samples and sonic immersion have become well established since the millennium (Jørgensen 2006; Grimshaw 2008). But the claim that a sound signal can be replicated precisely enough that its output fools the cognitive response of listeners on both of Deely’s levels—as cognizing organism and as raw stimulus—remains untested. We might call it the “perfect speaker” hypothesis: where a speaker not only is indistinguishable from the sound source but also inspires the same emotional reaction. To be sure, in the context of audiophile culture, speaker manufacturers’ reliance on indexical relationships between sound source and sound sample have been applied to advertising for decades—and, as the British firm Mordaunt-Short’s phono-realist advertisement from 1988 implies, not only for humans. **Figure 9.2**, from *What Hi-Fi?* magazine, validates the speaker’s sonic authenticity by a Humboldt penguin’s confusion over the object of its amorous affection. In the absence of a firm definition for high-resolution audio for ears, comparable to the pixels-per-inch for retinas, perception must remain the register of realism.^{xii}

Speeds of existence: multiples of 1,000

A debate coeval with that of cinematographic realism is how to calculate empirically the difference in sense acuity and cognition between the perception of different humans. Just after Bergson, the philosopher Wilhelm Dilthey argued that poetry had always used language to produce “the impression and illusion of reality” ([1910] 1985). To that end, it too becomes an analogue technology that transports its readers’ cognitive sense into alien realities, not unlike a trip to the flicks. The reader, writes Dilthey:

finds himself in a world of appearance not subject to the necessities of his actual existence (*Existenz*). But the [poetic] work heightens the reader’s feeling of his human existence (*Daseingefühl*). For the person confined by the course of his own life, it satisfies the longing to experience possibilities which he himself cannot realize. It opens up to him a view into a higher and more powerful world.
(250)^{xiii}

Far from a dusty play of signs, the pleasures afforded by such a new experience are sensory, we learn: “pleasure in sound, rhythm, and visual clarity” that contributes to a sequence of psychic processes culminating in “a genuine understanding of an event on the basis of its relations to the whole scope of life” (251). Since research by figures like Rudolph König and Carl Stumpf proved empirically that auditory worlds existed beyond the range of human hearing, visual worlds beyond human sight, the very concept of a single reality appeared naively anthropocentric, on this empirical basis, perhaps for the first time.

One contemporary example of such an attitude was Dilthey’s teacher, the German entomologist Karl Ernst de Baer, who asked the corresponding question: “Which version of nature is the right one?” This was the title of a lecture he gave in May 1860 (to mark the founding of the Russian Entomological Society in St. Petersburg), where he presented listeners with a thought experiment concerning sensory perception. It goes as follows: if a human life span of 80 years consists of 29,200 days, and if this were to pass by one thousand times faster giving a compressed life span of 29 days, which could again be sped up by a factor of one thousand, it would result in a total life span of 41 to 42 minutes (41m 46s), and

the corresponding rate of perception would be a million times faster than usual. For such a person, de Baer suggests, the organic world would probably appear disappointingly static, but other experiences currently unavailable to us would be accessible. “All the sounds we hear would certainly be inaudible to such people, if their ears remain morphologically similar to ours; but perhaps they would perceive sounds that we do not hear, indeed perhaps they would even *hear* light that we *see*” (de Baer 1862: 30, author’s translation). Returning to the first temporal compression (one-thousandth of a full life), he calculates that the highest sounds we perceive, vibrating at 48,000 times between two pulsations,^{xiv} would vibrate only 48 times between pulsations for people of shortened life span, hence they would sound low (30–31). At the upper end, even the second compression, resulting in a 42-minute life, he argues, would not quite open up our perceptual apparatus to an ether vibrating at “several hundred billion times a second.”

But we could take the idea of shortening a real life further, until these vibrations of the ether, which we currently experience as light and colour, actually become audible. And might there yet be in nature quite different vibrations which are too fast for us to experience as sound, and too slow to appear to us as light?...It is not at all preposterous to believe so...The planets, our earth among them, move through the ether with quite considerable speed and must set this speed. Is there not perhaps a sounding of outer space, a harmony of the spheres, that is audible to ears quite different to ours?
(31–32)

The quasi-scientific postulate of alien auditory realities and the apparently simple manner of calculating their relation to lived experience suggest the degree of fascination that limited perception held for those curious about the biological underpinning of human nature.

As an entomologist, de Bear had insects in mind when comparing perceptual realities of beings of compressed and uncompressed lifespan. Fully half a century later, this animal-human underpinning would receive perhaps its most enduring articulation, one whose abandonment of categorical differentiation between sentient organisms has proven attractive for posthumanist discourse in our century. In 1909, the Baltic German biological Jacob von Uexküll published *Umwelt und Innenwelt der Tiere*, in which he formalized his theory of

Umwelt, whereby each sentient organism creates its unique environment by its capacity to receive only signals that register on its peculiar sense organs. It inhabits a bubble, its individual *Umwelt*, which is determined by what is perceived sensorially based on sense acuity (*Merkwelt*) and the uses to which these senses are regularly put, their habits or training (*Wirkwelt*). As a result, the world is different, sensorially speaking, even for members of the same species. Uexküll's theory has been recounted many times; here—following de Bear's compressions by one thousand—I'll mention only his example of the tick that, upon smelling the butyric acid of passing prey, must drop down from a tree onto the animal and begin boring for blood. It will feed only once before dying, so it can neither learn nor refine the procedure. According to Uexküll, an experiment at the Zoological Institute in Rostock determined that the tick could survive up to 18 years without nourishment, that is, 18 years on a tree branch before falling onto a passing animal. During this time, Uexküll hypothesizes, the animal goes into a kind of hibernation, unaware of time passing, as the perceptual moment is lengthened far beyond that of human perception:

The tick can wait 18 years; we humans cannot. Our human time consists of a series of moments, i.e. the shortest segments of time in which the world exhibits no changes. For a moment's duration, the world stands still. A human moment lasts one-eighteenth of a second...The duration of a moment is different in different animals...During its waiting period, the tick is in a state similar to sleep...Time stands still in the tick's waiting period...and it starts again only when the signal "butyric acid" awakens the tick to renewed activity.
(von Uexküll [1934] 2010: 52)

Given the variant speeds of existence contemplated above, his conclusion that "the subject controls the time of its environment" bears a striking relation to the crank-driven technology of cinematography—which, as Inga Pollmann has argued, "played a key role in Uexküll's development of his theory of *Umwelt*" (2013: 779). How, we might wonder, did he settle on the comparison of 18 years to one-18th of a second, a tick's perceptual "moment" to a human's?^{xv} This is perhaps nothing but a multiplication of the most common frame rate for the cinematograph—18 frames per second becomes 18 years; the one, the speed at which

human perception experiences “motion picture” from still photograms, the other, the length a tick can suspend consciousness without noticing a perceptual gap. Such a comparison is in keeping with Uexküll’s inclination to draw on contemporary technology (rather than make purely speculative leaps) to overcome the theoretical impasse of accessing other sensory worlds, worlds that remain empirically unknowable to individual humans, or—put differently—to determine externally an experience that is internally subjective and autonomous.

If cinematography is here figured as only the most contemporary form of a *deus ex machina* that achieves what human perception cannot, questions over speeds of cognition, perception of frequencies, and sensory resolution have found expression across motley contexts concerned with the discursive proposition of human realism. Sampling these contexts serves to uncover a behavioral habit whereby recent technological apparatuses, rather than logic or imagination, are co-opted by writers to overcome the above theoretical impasse between external and internal determination of subjective perception. H. G. Wells’s short story *The New Accelerator* (1901) gave expression to the fantasy of time axis manipulation in the realm of fiction; it postulates an elixir that, when taken, speeds up the taker’s cognitive and physiological processes so that the subject feels identical in him- or herself but the external world is radically slowed down: “My heart...was beating a thousand times a second, but that caused me no discomfort at all.” The illustration accompanying a magazine reprint in 1926 (Figure 9.3) depicts the two leading characters casually observing the “statuesque” modern traffic flying past, as though in freeze frame (Wells 1926: 60). Unsurprisingly, perhaps, playful thinking about speeds and resolutions of perception has taken root within the scientific imagination over centuries. During the Great Plague of 1665 the English natural philosopher Robert Hooke first presented the idea of thinking oneself into new perceptual worlds through the new technology of the microscope. The illustration of a

blue fly, as Hooke peered at it through Christopher White’s microscope (Figure 9.4), indicates the invisible, miniature world made visible by his device. With an ear for *microsonic* realities, his text *Micrographia* relates how the sound of bees’ wings, understood in relation to the vibration of a musical string (“tun’d unison to it”), vibrates “many hundreds, if not some thousands” of times per second, and may be “the quickest vibrating *spontaneous* motions of any in the world.”^{xvi} This auditory extrapolation, from an optical fascination with fluttering wings, is indicative of how vision became only the first sense modality to be treated to changed acuity. “Mechanical inventions” to enhance hearing in comparable ways are “not improbable,” he speculates; they could result in the ability to hear ten furlongs away, we learn, or hold a conversation “through a wall a yard thick” by propagating “auditory” vibrations not through air, but along wire or via light.^{xvii}

The quasi-literary imagination behind Hooke’s ideas is set in relief by comparison to later, deliberate borrowings from fiction, such as those of the Leipzig Cantor and theorist Moritz Hauptmann, who in 1863 speculated on the hearing of alternatively sized bodies and the proportional relations between their new sensory realities. With reference to Swift’s *Gulliver’s Travels* (1726), he explained that Lilliputians, at half a foot tall, are 12 times smaller than Gulliver’s six-foot stature; Brobdingnagians are to Gulliver as he is to the Lilliputians: 12 times larger. Hence, the relations may be characterized as follows (Hauptmann 1863: 25):

<i>Lilliputians</i>	<i>Gulliver</i>	<i>Brobdingnag</i>
1/12	:1	:12/1
1	:12	:144

On this basis, the longest organ pipe of 32 feet (16Hz) would give the lowest C₂ to Gulliver (and us), but—according to Swift’s ratios—the same would be only 2⅔ feet for the

Lilliputians, but 384 feet for the Brobdingnagians. As [Figure 9.5](#) shows, the equivalent low C₂ pitch for Lilliputians would therefore be Gulliver's g (196Hz) below middle C (c¹).

Likewise, if a Lilliputian oboe tunes the orchestra to A = 440Hz, the 1:12 pitch ratio would result in an e⁴ for Gulliver, while the lowest pitch of the Lilliputian double bass, the E of a 16-foot organ pipe, would be equivalent to Gulliver's b¹. Hauptmann's illustrations extend to the Brobdingnagian orchestra: "From the double basses, bass trombones, ophicliides, and everything that produces a deep tone, we [and Gulliver] would only see the movements of the players and feel the aerial vibrations." With poetic infrasound in mind, an oblique reference to George Berkeley's immaterialism was perhaps inevitable: "Sound, like color, is merely subjective. Neither exists without a listening ear and a seeing eye" (Hauptmann 1863: 26, author's translation).

While morphology of the ear can be assumed to behave according to simple ratios, and a quantitative approach to realism and corporeal difference might appear to have prevailed in this historical context, Hauptmann ultimately cautions that temporality is not so simple:

We cannot claim that a symphony that lasts...60 minutes for us, must last 5 minutes in Lilliput, 12 hours for the Brobdingnags. Other temporal dimensions may certainly be supposed...Since [Lilliputians and Brobdingnagians] inhabit our world, were conceived and warmed by our sun, so their year is the same, their day, their hours are just as long; but their metronome, pendulum, heartbeat, the movement of their accompaniment remain in relation to their body size. In short, conflicts and doubts arise everywhere, which we will soon leave behind, and we will have to be satisfied with the assumption that they are human conditions as they are to us, as befits humans of five to six feet tall.
(Hauptmann 1863: 27)

Here, the theoretical impasse identified by Uexküll, between external determination of a sensory experience and internal subjective autonomy, remains recalcitrant as the poetic fictions of smaller and larger people are made to inhabit the putatively singular world with its singular mass and speed of rotation around the sun.^{xviii} Different hypothetical life expectancies would further complicate the multiple temporalities, so ultimately Hauptmann's thought experiment, alongside de Bear's ratio-adjusted vibrations, already begins to

undermine the quantitative approach to realism that technological apparatuses afford, whether microscopes, organ pipes, or cinématographes.

Voice resolution at 1:1,000

If, finally, we time-travel to the present, a more contemporary context for quantitative realism indicates that the discourse's reliance on emergent technology remains undimmed in the third decade of the 21st century. Until 2010 speech synthesis, from digital assistants like Alexa and Siri to simulations that ventriloquize our own voices, typically functioned by sampling large amounts of recorded speech fragments from one individual so words can be reassembled into an utterance appropriate to the message being conveyed. The component sounds were simply concatenated into theoretically endless chains of human-like utterances, dubbed "concatenative text-to-speech" synthesis. While these remain rooted in phonemic sounds recorded in the real world, cobbled together by algorithm, a more recent approach sees synthetic voices emanate from the generation of raw waveforms, assembled one waveform at a time and densely combined. That is, synthetic sound samples are pieced together to form waveforms at high resolution to mimic a real voice. Harking back to Baudrillard's terms of reference, this constitutes a third-order simulation. An example is DeepMind's WaveNet where, like melodies generated by Markov chains, a predictive distribution for each audio sample is conditioned on all previous ones, rising to at least 16,000 samples per second, a remarkable level of artifice in pursuit of what the WaveNet engineer Aäron van den Oord has called "subjective naturalness." This artificial approach to natural voices aims to "directly model the raw waveform of the audio signal, one sample at a time" (2016). Given Uexküll's ratio of *18 years:1/18 of a second*, the sample rate is not arbitrary, as we shall see.

A similarly synthetic process of voice simulation, at a resolution 1,000 times lower, is the Austrian composer Peter Ablinger's *Deus Cantando* (God, singing) (2009). This is only one

of the most recent spectral analyses of recorded speech that form the basis of his aptly named “speaking piano,” a computer-controlled player piano that replicates on the instrument’s 88 keys the decomposed sound spectrum of recorded human speech. As Ablinger explains:

Using...16 units per second (about the limit of the player piano), the original [sound] source approaches the border of recognition within the reproduction. With practice listening the player piano can even perform structures possible for a listener to...understand as spoken sentences. (Ablinger, n.d.)

That is, you can “hear” the piano pronounce words only when you simultaneously see its words or know them in advance. The speaking piano’s sample rate, 1,000 times lower than WaveNet’s simulation, teeters on the brink of comprehensible phonemes (i.e., far removed from a “perfect speaker”), and a visual analog might be the differently pixelated screens that Uexküll uses to imagine the different visual worlds for a human, a fly, and a mollusc, based on the cellular density in their retinas, where visual objects become progressively harder to make out (Uexküll [1934] 2010: 64–65). For Ablinger, comprehensibility is secondary to investigating the liminal space between phonorealism and the innately musical medium of the 19th-century piano—or, as he puts it, “the observation of ‘reality’ via ‘music’” (Ablinger, n.d.). Faced with the question of what the “reality” of the sound of a human voice might be, technological innovation forces quantitative, frequency-based answers of the kind we’ve just sampled. While frequency rates vary, none is any the less artificial. Accepting the split of auditory perception into an infinite plurality, and with a continuing reliance on technological affordance, what we understand to be spoken sounds may become defined more by what can be simulated, rather than what any individual perception make take to be “real.” This would seem the tacit assumption behind more common assertions that a synthetic sound, when heard, becomes real in its own right.

Coda

It is a truism that, phenomenologically, voice and identity become interdependent over time for the subject; the timbre, intonation, and cadence of your vibrating physiology become, in part, your identifying sound. As Steven Connor famously put it: “Voice is not simply an emission of the body; it is also the imaginary production of a secondary body, a body double: a ‘voice-body’” (2000: 35). Beyond this monist coupling, the sensation of hearing oneself talk, the feel of our resonating throat, is characteristic of self-identity in a genetic sense. It is the first sound we hear in the outside world. So it seems unsurprising that it was the early materialists of the late 18th century who would recognize its self-identifying agency as such. The poetic preface to Erasmus Darwin’s *Zoonomia* in 1794, for instance, speaks of the moment a child first perceives sound in the world, before it incrementally becomes less alien:

’Erewhile, emerging from its liquid bed,
It lifts in gelid air its nodding head;
The light’s first dawn with trembling eyelid hails,
With lungs untaught arrests the balmy gales;
Tries its new tongue in tones unknown, and hears
The strange vibrations with unpractis’d ears.
(Darwin [1794] 1809: v)

In this historical context, self-recognition also works at the level of the species; hearing a voice in the desert announces to you “a being like yourself,” explained Rousseau in *On the Origin of Languages*. Vocal signs “are, so to speak, the voice of the soul” (Rousseau [1781] 1986: 63–64). And as Michel Serres reminds us, such sentiments exceed the narrow dualism they imply, for “all *real* bodies shimmer like watered silk. They are hazy surfaces, mixtures of body and soul” ([1985] 2008: 35, emphasis added). This is perhaps the reason why voices cannot be relinquished in the miniaturized characters in films such as *Downsizing*, to return to

the reference point with which we started. A size-modulated voice would imply a change of underlying identity incommensurate with the film's narrative continuity. The lack of discursive treatment around realism in the context of filmic miniaturization would seem beside the point, then. The medium of digital cinema embodies this discourse in a cultural technique of quantitative sampling and the history of perception this implies. Historically, the voice's condition has been perennially technologized, but with high-frequency speech synthesis, voices have seemingly become fractal for the first time, a data set with the potential for infinite replication, and as such are subject to the very critique that Baudrillard leveled at cloned DNA, with all the rhetorical excess and intellectual violence this implies.

Beyond this paired critique of realism and identity, the ethical quandary implied by "cloned" voices raises a further question: do you own this identifying sound as a composer "owns" a composition, a performer "owns" a recording, and humans own their DNA; or is the simulation autonomous on its own terms? Lyrebird, a voice-cloning company in Montreal, uses generative speech synthesis technology similar to Wavenet, but specializes in drawing on human voice samples to ventriloquize those voices in words and statements they never uttered. The technology is susceptible to "deep fake" media, and Lyrebird has taken a public stand on its ethical responsibilities:

In many use cases, the results [of generative media] are already indistinguishable from real media. This technology has exciting applications...but it also holds the potential for misuse...We are committed to modeling a responsible implementation of these technologies, unlocking the benefits of generative media while safeguarding against malicious use.

We believe you should own and control the use of your digital voice. [We use] a process for training speech models that depends on real-time verbal feedback, ensuring that individuals can only create a text-to-speech model of their own voice. Once created, the user is the owner of their voice and has the sole authority to decide when and how it is used.^{xix}

Here, the ethical ground is guaranteed by the participation of the voice-owner (who must offer verbal feedback to generate the simulation), but in other hands the technology could proceed without consent. If simulations of the human voice are already attaining hyperreal heights through 16,000 samples per second, it may be necessary to define a real voice

according to its origins in a human body, rather than any *a priori* sonic principles. Adapting Baudrillard: “It is no longer a question of a false representation of [a *real* voice] but of concealing the fact that the real [voice] is no longer [singularly] real” (Baudrillard 2010: 12–13).

In other words, the phenomenon of synthetic speech and artificial generation raises the underlying issue of how we might define the “real” of sound itself in the digital age, whether this in fact has any validity in the absence of a single (perceiving) subject position, or warrants status in our critical thinking. To the extent that *sound* can be considered an object, and therefore something that can be possessed as a digital quantity, do we have a right to own sounds arising from our congenital biological frame? Or might we consider these an accident or corollary of evolutionary history? Whether taken as an ontological or a historical matter, this topic—arising from perceptual realism—has occupied commentators long before digital speech synthesis, and points to a philosophical instability at the heart of sound studies, namely: the notion of sound itself as a contested object.

References

1. Abbate, Carolyn. 2016. “Sound Object Lessons.” *Journal of the American Musicological Society* 69: 793–829.
2. Ablinger, Peter. n.d. “Quadraturen.” <http://ablinger.mur.at/docu11.html#principles>.
3. Bains, Paul. 2006. *The Primacy of Semiosis: An Ontology of Relations*. Toronto, ON: University of Toronto Press.
4. Barton, Ruth. 2003. “‘Men of Science’: Language, Identity and Professionalization in the Mid-Victorian Scientific Community.” *History of Science* 41: 73–119.
5. Baudrillard, Jean. (1976) 1993. *Symbolic Exchange and Death*. Rev. ed., translated by Ian Hamilton Grant. London: SAGE.
6. ———. 1988. *The Ecstasy of Communication*. Translated by Bernard and Caroline Schutze. New York: Sylvere Lotringer.
7. ———. 1993. *The Transparency of Evil: Essays on Extreme Phenomena*. Translated by James Benedict. London: Verso.
8. ———. 2010. *Simulacra and Simulation*. Translated by Sheila Faria Glaser. Ann Arbor, MI: University of Michigan Press.
9. Bennett, Jane. 2010. *Vibrant Matter: A Political Ecology of Things*. Durham, NC: Duke University Press.
10. Bergson, Henri. 2005. *Creative Evolution*. Translated by Arthur Mitchell. New York: Barnes & Noble.

11. Brownlee, John. 2012. "Why Retina Isn't Enough." *CultOfMac*, June 15, 2012. <https://www.cultofmac.com/173702/why-retina-isnt-enough-feature/>.
12. Chion, Michel. 1994. *Audio-Vision*. Translated by Claudia Gorbman. New York: Columbia University Press
13. Connor, Steven. 2000. *Dumbstruck: A Cultural History of Ventriloquism*. Oxford: Oxford University Press.
14. Damböck, Christian, and Hans-Ulrich Lessing, eds. 2016. *Dilthey als Wissenschaftsphilosoph*. Freiburg: Karl Alber.
15. Darley, Andrew. 2000. *Visual Digital Culture: Surface Play and Spectacle in New Media Genres*. London: Routledge.
16. Darwin, Erasmus. (1794) 1809. *Zoonomia*. Boston, MA: Thomas and Andrews.
17. de Baer, Karl Ernst. 1862. *Welche Auffassung der lebenden Natur ist die richtige?* Berlin: August Hirschwald.
18. Deely, John. 1994. *New Beginnings: Early Modern Philosophy and Postmodern Thought*. Toronto, ON: University of Toronto Press.
20. Der Deriam, James. 2001. *Virtuous War*. Boulder, CO: Westview.
21. Dilthey, Wilhelm. (1910) 1985. "Poetry and Lived Experience." In *Poetry and Experience*, edited by Rudolf A. Makkreel and Frithjof Rodi, 250–53. Princeton, NJ: Princeton University Press.
22. Funkenstein, Amos. 1986. *Theology and the Scientific Imagination from the Middle Ages to the Seventeenth Century*. 2nd ed. Princeton, NJ: Princeton University Press.
23. Grimshaw, Mark. 2008. *The Acoustic Ecology of the First-Person Shooter: The Player Experience of Sound in the First-Person Shooter Computer Game*. Saarbrücken: Mueller.
24. Hauptmann, Moritz. 1863. "Klang." In *Jahrbücher für musikalische Wissenschaft*, edited by Friedrich Chrysander. Leipzig: Breitkopf & Härtel.
25. Hayles, Katherine N. 2012. *How We Think: Digital Media and Contemporary Technogenesis*. Chicago, IL: University of Chicago Press.
26. Hilbert, Martin, and Priscila López. 2011. "The World's Technological Capacity to Store, Communicate, and Compute Information." *Science* 332: 60–65.
27. Hooke, Robert. 1665. *Micrographia*. London: printed for John Martin.
28. Jørgensen, Kristine. 2006. "On the Functional Aspects of Computer Game Audio." *Proceedings of Audio Mostly Conference*, October 11–12, 2006, Piteå, Sweden. <http://hdl.handle.net/1956/6734>.
29. Kapp, Ernst. 1877. *Grundlinien einer Philosophie der Technik* [Elements of a philosophy of technology]. Brunswick: Westermann.
30. Locke, John. (1689) 2008. *An Essay Concerning Human Understanding*. Abridged by Pauline Phemister. Oxford: Oxford University Press.
31. Marks, Lisa. 2000. *The Skin of the Film*. Durham NC and London: Duke University Press.
32. Negroponte, Nicholas. 1995. *Being Digital*. New York: Knopf.
33. Perry, Nick. 1993. *Hyperreality and Global Culture*. London: Routledge.
34. Pollmann, Inga. 2013. "Invisible Worlds, Visible: Uexküll's Umwelt, Film, and Film Theory." *Critical Inquiry* 39: 777–816.
35. Roads, Curtis. 2004. *Microsound*. Cambridge, MA: MIT Press.
36. Rousseau, Jean-Jacques. (1781) 1986. *On the Origin of Languages*. Translated by John H. Moran and Alexander Gode. Chicago, IL: University of Chicago Press.
37. Serres, Michel. (1985) 2008. *The Five Senses: A Philosophy of Mingled Bodies*. Translated by Margaret Sankey and Peter Cowley. London: Continuum, 2008.
38. Siegert, Bernhard. 2003. *Passage des Digitalen*. Berlin: Brinkmann & Bose.

39. Sterne, Jonathan. 2012. *MP3: The Meaning of a Format*. Durham NC: Duke University Press.
- Strachan, Robert. 2017. *Sonic Technologies: Popular Music, Digital Culture and the Creative Process*. New York: Bloomsbury.
40. Théberge, Paul, Kyle Devine, and Tom Everett, eds. 2015. *Living Stereo: Histories and Cultures of Multichannel Sound*. New York: Bloomsbury.
41. Trippett, David. 2018. "Music and the Transhuman Ear: Ultrasonics, Material Bodies and the Limits of Sensation." *Musical Quarterly* 100: 199–261.
42. van den Oord, Aäron, and Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Viuyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. 2016. "WaveNet: A Generative Model for Raw Audio." September 19, 2016. <http://deepmind.com/blog/wave-net-generative-model-raw-audio/>.
43. Varro, Marcus Terentius. 2006. *Varro on the Latin Language [De lingua Latina]*. Translated by Roland Kent. Loeb Classical Library. Cambridge, MA: Harvard University Press.
44. von Uexküll, Jacob. (1934) 2010. *A Foray into the World of Animals and Humans*. Translated by Joseph D. O'Neil. Minneapolis, MN: University of Minnesota Press.
45. Wells, H.G. (1901) 1926. 'The New Accelerator.' *Amazing Stories* 1 (April): 57–61, 96.

Figure 9.1 Etienne-Jules Marey's chronophotographic study "Cyclist" (c. 1894). Image from "La Collection des appareils," Iconothèque, Cinémathèque Française, Paris.

Figure 9.2 Modaunt-Short's advert for *What Hi-fi?* (1988), illustrating the desirability of (undefined) perfect fidelity in audio reproduction.

Figure 9.3 Unsigned illustration of H. G. Wells's 'The New Accelerator', depicting Professor Gibberne and the narrator observing a radically slowed-down environment, for the inaugural issue of *Amazing Stories* 1 (April 1926), 57.

Figure 9.4 Robert Hooke's large-scale illustration of a blue fly as seen through magnifying glasses, and reproduced in Hooke's *Micrographia* (1665), the first major work on microscopy. British Library Collection 435.e.19.

Figure 9.5 Moritz Hauptmann's musical examples for the same sonic frequencies heard by differently sized bodies, modulated by Swift's ratios given above.

ⁱ It is perhaps indicative that the deceptive realism of filmic effect has made it a favored vehicle for exploring vicarious perception. Small wonder, then, that as early as 1901, Georges Méliès's *The Dwarf and the Giant* depicts a single man who splits into two versions of himself, one who grows tall, one who shrinks to an eighth in size. Contributions to this subgenre of sci-fi films on the topic of size alteration would include those from *The Incredible Shrinking Man* (Jack Arnold, 1957), *Darby O'Gill and the Little People* (Robert Stevenson, 1959), and *Devil-Doll* (Lindsay Shonteff, 1964) to *Honey, I Shrunk the Kids* (Joe Johnston, 1989) and its ensuing franchise with the Walt Disney Company, as well as *Ant-Man* (Peyton Reed, 2015).

ⁱⁱ The technical challenge of counterpointing tiny and "normal" humans was itself sufficient for the film to be nominated for the American Visual Effects Society's award for Outstanding Supporting Visual Effects in a Motion Picture. See <https://visualeffectssociety.com/portfolio-items/2017-16th-annual-ves-awards/?portfolio-Cats=29>.

ⁱⁱⁱ Two examples of critiques that respond to Baudrillard's claims would include Perry 1993, which explores cultural contexts where original cannot be distinguished from copy, and Der Deriam 2001, which situates the theory of virtuality in warfare.

^{iv} See <https://www.oed.com/view/Entry/129823?redirectedFrom=obscene#eid>.

^v It appears to be a misunderstanding of Varro, who in fact argues just the opposite, that anything shameful is called *obscenum* because it ought not to be said openly *other than* on stage. See Varro 2006: VII: 351.

^{vi} A recent summary is given in Strachan 2017.

^{vii} See <https://www.sciencedaily.com/releases/2007/08/070823122253.htm>.

^{viii} See <https://www.pcmag.com/archive/analyst-challenges-apples-iphone-4-retina-display-claims-251638> and <https://www.npr.org/sections/alltechconsidered/2010/06/07/127530049/live-blogging-apple-s-developers-conference>.

^{ix} For the modern period, the touchstone for placing trust in sensation remains John Locke's anteriority of sensation to reflection ([1689] 2008).

^x This definition of realism has been explored by Darley 2000.

^{xi} A thoughtful critique of Deely's framework is given in Bains 2006: 49ff.

^{xii} A joint definition for high-resolution audio, agreed between the Recording Industry Association of America, the Consumer Electronics Association, the Digital Entertainment Group, and the Recording Academy Producers & Engineers Wing, remains technologically open, and rooted in intentionality: "lossless audio capable of reproducing the full spectrum of sound from recordings which have been mastered from better than CD quality (48 kHz/20-bit or higher) music sources which represent what the artists, producers and engineers originally intended." See "High Resolution Audio Initiative Gets Major Boost with New 'Hi-Res MUSIC' Logo and Branding Materials for Digital Retailers," *The Recording Industry Association of America (RIAA)*, June 23, 2015, <https://www.riaa.com/high-resolution-audio-initiative-gets-major-boost-with-new-hi-res-music-logo-and-branding-materials-for-digital-retailers/>.

^{xiii} Reading perceptual mechanisms into words would seem more than just another term for literary realism. Its boldness may be taken as indicative of the multidisciplinary outlook afforded by that generation of 19th-century scientists who lived through the professionalization of different branches of the sciences, human and natural, within universities. See Damböck and Lessing 2016 and Barton 2003.

^{xiv} It was not uncommon during the second half of the 19th century for scientists to propose a higher upper limit for the aerial frequencies human ears could hear, now commonly accepted to be 20,000Hz. See Trippett 2018: 202–7.

^{xv} More recent theorists posit the smallest unit, or "grain," of audible sound at between a thousandth and a tenth of a second. See Roads 2004: 86–97.

^{xvi} Robert Hooke, *Micrographia* [1665], "Observation 38 'on the structure and motion of the wings of flies'."

^{xvii} Hooke, *Micrographia*, "Preface." On the origin of the microphone, see Abbate 2016.

^{xviii} It is precisely the singular world of Classical biology that Uexküll would reject. See Trippett 2018: 208–ff.

^{xix} See <https://www.descript.com/ethics>.