Video 5: Sounding – Resonance

[00:00:00]

Mara Mills: I'm Mara Mills, and I'm an assistant professor in Media Culture and Communication at NYU.

From the so-called sense section of the Oxford English Dictionary entry for the word *modulation*: The action of singing or making music, a tune, a melody. Variation of the quality of one's voice with respect to tone, pitch and intensity. Physics and engineering, the process of modulating, an electromagnetic wave in order to impress a signal on it, frequently preceded by a noun denoting either the characteristic of the carrier wave that is varied, as in frequency modulation, or the method by which the modulation is applied, as in grid modulation.

Is there a direct route from modulation in music and speaking to modulation in electronics? How has voice been extrapolated to signal? I'm going to give you a capsule history of this term beginning with phonetics and deaf education in the 19th century, and ending with pulse-code modulation, an early form of digital signal processing. And the telephone will serve as a hinge in this narrative.

I am placing technologies for speech transmission including the vocal tract with its sound shaping resonances -- the theme of this panel -- at the center of new media history. And in the spirit of CAST, I want to suggest that modulation is not merely a loan word from music to engineering, but a keyword, following of course, Raymond Williams.

"Keywords," he writes, "are the vocabulary we share with others, often imperfectly, when we wish to discuss many of the central processes of our common life. These are significant, indicative words," this is still Raymond Williams, "certain uses bind together certain ways of seeing culture and society. Certain other uses open up issues and problems, of which we all need to be very much more conscious."

Yesterday, we experienced moments of friction between the sciences, and art and philosophy. Here I want to think about common discourse.

Theresa Dudley began a six-month program in speech training with Alexander Graham Bell in September 1871 here in Boston. She was 17. He was 24. Dudley was one of his first students in the United States, and the very first to have been born deaf.

Her father, Lewis Dudley was a member of the Governor's Council of Massachusetts. Her mother had taught her the manual alphabet, and how to read and write in English. Thus, she already communicated by numerous means. She had also been a star pupil at the American Asylum in Hartford, where she learned sign language over the course of two years.

In a series of lessons on articulation and culture of the voice, and this page from the Alexander Graham Bell papers is a page from her lessons, much of the lessons were written down in notebooks. Bell coached her to gain quote, "command over the movements of the vocal organs." Articulation was a term from phonetics, a field newly founded in the 19th century through the work of Bell's father Alexander Melville, among others.

Articulation referred to the production of speech sounds, the meaningful elements of speech in a given language, which seemed to be buried within individual voices. Speech was studied from all sides in the 19th century. Tools and theories were exchanged between the new medical specialties of otology and laryngology, the emerging academic discipline of linguistics, with its focus on living sounds over dead letters, and acoustical research in physics, informed by calculus.

Sound was understood to be a vibratory phenomenon like heat and light. It was medium dependent, thus its media, at first, air, water, solids, also came under scrutiny. Experimental, or instrumental phonetics yielded tools for recording, transmitting, dissecting and synthesizing speech. Linguistic John Joseph states that, "experimental phonetics, the detailed measurement of speech sounds, offered the first truly positivistic approach language."

We might say that speech begins to be mechanized in this period, continuing into phases of electrification and automation with the telephone and other electroacoustic devices. And that's part of the story I'm going to tell.

In the absence of auditory feedback, Dudley worked with visual, tactile and gestural translations, or signs of speech. She might watch Bell pronounce, fff, then imitate the motions of his teeth and lips while looking in a hand mirror. He made comparative sketches of their mouths and throats, pointing out the hidden organs in their vocal tracts.

Central to articulation training, Bell employed the visible speech symbols designed by his father. You can see those symbols at the top of the page on the screen. This iconic alphabet depicted the positions of the tongue, lips and throat, what Melville called the speaking machine, for each phoneme.

[00:05:10] Melville considered his characters to be phonograms, or facsimile writing. This is how he described them in his own text. However, they portrayed the mechanism of sound production, not the acoustic wave. He too offered private lessons for quote, "all cases of impediment or defective speech, and for oratorical ineffectiveness."

Melville refined his physiological alphabet, three years of pedagogy. His students furnished, he said, "a constant variety of examples for study and corroboration." He predicted in 1863 that the science of phonetics and the allied field of deaf education would hold the key to the design of new sound technologies.

So he wrote in *Principles of Speech and Dictionary of Sounds*: "Scientific men have elaborated theories of optics, and look at the results, wonderful mechanical adaptations of optical principles, before undreamt of, and which otherwise would never have been discovered. Might not an analogous result attend the philosophical investigation of the faculty of speech, and acoustic and articulative principles be developed which would lead to mechanical inventions no less wonderful and useful than those in optics."

Working with Graham Bell, Dudley performed oral gymnastics to become familiar with different configurations of the vocal tract. Blowing on a feather through the nostrils, voicing while holding an ivory plug or a pencil between the teeth. I spent years in speech therapy in grade school. Those of you who shared this misery with me know that some of these lessons haven't changed much.

Pitch, duration, loudness and rhythm were different aspects of the culture of the voice, which Bell increasingly termed *modulation*. The very first page of this notebook, if I had shown it to you, you would see that culture of the voice was the phrase that he used for the parts of the lesson that have to do with what we would now modulation. By the time you get to page 146, modulation creeps in as the keyword.

If articulation conveyed words, modulation conveyed emotion. In the classical period that the Latin word for modulation referred to inflection of tone. When it entered English in the 14th century, it designated the action of singing, or making music. Modulation was first associated with pitch, intonation of the voice, or change in musical key, although the term soon came to be applied to timbre, rhythm and other attributes.

To explain pitch to deaf students, Bell sometimes used musical notation. But he searched for a technique that would prompt Dudley to inward self-correction. Dudley held one hand to his throat, and another to her own to get a sense of pitch and duration. Bell asked her to follow the motion of his pencil while she was inflecting a speech sound. This would be the pencil markings at the bottom half of this page.

The length of a line indicated duration. Its thickness corresponded to force, or loudness. Upward and downward motion indicated pitch.

One time, he drew a dotted line of a speech curve, and asked her to make the corresponding sound. He then penciled in her vocalization, the pitch rising off the page. If it was obvious that words were distinct from their modulation, the same words could be differently inflected. Together Bell and Dudley investigated the manifold forms, in which words could be represented, and how they could be conveyed to different senses.

Bell soon began to tutor Alice Jennings, a 21-year-old student at the Boston School for Deaf-Mutes, now the Horace Mann School. She had been deafened at eight by a bout of scarlet fever. Bell explored the anatomy and physiology of speech through his work with Jennings, although certain details remained incommunicable.

At one point, he wrote in her notebook, "I want to know what you feel. Where do you feel the pitches? In your ears? In your head, or throat, or chest? Do you have a different feeling for do, mi, sol? Try and think what the difference is."

She replied, "I cannot explain it, because I do not know how. If you were deaf, perhaps you might find it difficult to explain it when a teacher asked you. I think anyone would find more or less difficulty in explaining the feelings given them by musical intervals." End quote from Jennings.

On the other hand, in 1880, Jennings published a book of poems *Hard Echoes*, which revealed her deep intuition of sound and its uses. The poems rhymed, some employed acoustic metaphors, others the theme of acoustic exclusion. One titled "Influence" was especially thoughtful about the uncanny material semiotics of the spoken word. I'll just read you one verse. "The view-less air that every tone in swift vibrations onward sends is emblem of that power unknown, unseen that every word attends."

In the 19th century, scientists struggled to understand how the molding of speech sounds occurred. Vowels served as the primary test case. Produced by the free passage of air through the vocal tract, unlike many consonants, all the vowels could be spoken in the same pitch, and yet distinguished by aspects of frequency, timbre.

[00:10:17] Charles Wheatstone surmised that articulation was the result of resonance in the vocal tract. As the oral cavities changed position during speech, different frequencies were reinforced, and different vowels produced. Working in Germany, Hermann Von Helmholtz applied Joseph Fourier's analytic theory of heat to acoustic vibration.

A complex wave could be broken down into a series of simple waves. Conversely, sine waves could be summed to create a complex sound. Helmholtz synthesized vowel tones by coupling a series of tuning forks to resonators, metal flasks or cylinders, with apertures that could be adjusted in size to control amplification.

In turn, he theorized that the vocal cords passed a complex set of airborne vibrations through the vocal tract, two or three overtones reinforced determined the vowel sound. The vocal tract thus worked something like a filter, reinforcing certain wave components and not others.

By the winter of 1873, Graham Bell became interested in the possibilities for voice writing, or visualizing sound vibrations, as he searched for a new technical aid to enable his students to both read lips and correct their own speech. In a story that has been remembered, recorded by AT&T, and recited in almost identical terms by media theorists ever since, Alexander Graham Bell assembled ear phonautograph, here in Boston in 1874, from the tympanum bones of a corpse. A proposed aid for his deaf students, a second-generation visible speech, the phonautograph inscribed waves on sooted glass.

What has been collectively forgotten, though also recorded, are the impressions of the students who participated in Bell's speech research. And this is the alternate history that I've been narrating so far. I take it to be part of my job as a historian to tell the stories about different ways of relating to and through technology, and about the ideals that get embedded into long-lived technical systems.

This is not prediction, as Tomaso Poggio noted yesterday. And this is why I bring this up is an aside. He mentioned-- I don't know if Professor Poggio is here right now-- but he mentioned in a tone that I as a historian took to be dismissive about history and economics, that they weren't predictive.

So the kind of narration that I'm saying is central to my work, is description. And description serves many functions too. It can perform functions ranging from the integration of chaotic data, memories, feelings, to the explanation of hidden or highly complex processes. I think this is especially important for the history of electronics. I don't think it's listed online yet, but next spring the Columbia Heyman Center for the Humanities will host a cross-disciplinary conference across the sciences and humanities and arts on this topic of description across the disciplines, which may be interesting to some of you who are here at this CAST conference.

So Bell quickly decided not to employ the ear phonautograph for deaf education. Its plates, obviously represented here, represented only brief cuts from a speech stream. And he admitted that he himself found it impossible to recognize the various vowel sounds by their tracings. The waveforms did not depict the spectrum of speech, nor reveal patterns of resonance.

However, Bell would later claim that the phonautograph led him to conceive of the membrane speaking telephone which employed a diaphragm to vary an electrical current in response to sound waves. I won't recount those familiar details here, but I want to note that Bell's students followed his experiments with interest, even as he neglected his teaching for electrical research.

Jeanie Lippitt who began to work with Bell in the fall of 1872, and whose father was the governor of Rhode Island recollected, "It was not long after I began my course of training with Professor Bell in Boston that he began to tell me about his wonderful new invention. He brought it under his arm in a black box, long and narrow.

It roused my curiosity from the first, as he kept it so carefully by his side while teaching me. Afterward I found he did not dare to let it out of his sight for a moment, fearing someone would steal his precious machinery. He became so absorbed in the invention that he could think and talk of nothing else.

I soon knew all of the details of the talking machine, but it wasn't perfect, something was lacking. And it was while training my voice, as well as the voices of this other deaf pupils, that the vibration of our voices gave Professor Bell the idea, the missing link. And then his machine really talked."

Bell included Lippitt in his stage demonstrations of the telephone in Rhode Island. And he later promised her that he would build a reflector attachment for the phone so quote, "deaf people could talk over it, and see the person at the other end to answer it."

[00:15:02] The telephone spurred with a physicist Frederick Hunt called the transducing 1870s a period a burgeoning invention in the field that came to be known as electroacoustics. And as I switch into electroacoustics, I hope you'll be able to hold in your mind that I'm using the telephone as a hinge, not as an origin. Because it's all too easy to say that everything, every piece of equipment we use in electroacoustics today comes out of AT&T.

So I'm moving into sort of the electroacoustic phase, but it's a hinge in this. Historian Roland explains, "As a transducer, the telephone connected the otherwise still separate electrical and acoustic worlds."

The term modulation entered the field of electroacoustics around 1907 through the publications of Reginald Fessenden regarding his early experiments with wireless telephony, forerunner to both, and things like mobile phones and radio. And he self-consciously imports the word modulation into his writings. And then other physicists at the same time and engineers credit him with that term.

Transmitting a call via electromagnetic waves offered a solution to the problem of overseas communication. Telephoning would be faster than wireless telegraphy, which required the coding, and decoding of messages by expert telegraphers. Modulation was described by William Eccles, also a wireless researcher, as the process of molding the oscillatory currents by means of the voice.

Fessenden's work would later be understood as a type of amplitude modulation, whereby voice signals are superimposed on electrical waves of higher frequencies, in that case, radio waves. In the top two images on this slide, from a 1948 textbook on frequency analysis, modulation, and noise, we see an audio signal loaded, placed onto a carrier, and modulating its envelope or amplitudes.

This type of modulation was also made possible, initially, through the exploitation of resonance. In the late 1800s, Heinrich Hertz and other physicists began to investigate electrical waves, and in turn, the phenomenon of electrical resonance, drawing analogies to acoustics. In the 1890s, several French and American experimenters developed tuned circuits for multiplex telephony, in which a series of high frequency waves were modulated by lower frequency speech signals, and then demodulated by resonance at the receiver. Most of these I think were prototypes. Multiplexing did not become common until later in the 20th century in telephony.

Multiplexing, sending multiple signals simultaneously over the same medium, was required to accommodate an influx of callers on a limited number of lines as the telephone network expanded. Tuned circuits were incorporated into wireless transmission and eventually into radio dials where they allowed tuning in. With the invention of the electric wave filter, also based on resonance in 1922, by AT&T's George Campbell and in Germany by Carl Wagner, bands of frequencies could be selected more precisely and with less distortion from an electrical current.

Wave filters could be used to suppress noise, or for instance to transmit particular segments of speech. Filtering continues to be a central technique in audio and video engineering, and beyond. Although it is under theorized in the field of media studies. It was an old impulse newly realized with massive consequences for representation.

For those of you in sound studies, I'll just note that Veit Erlmann's influential book that has already been placed on the screen this morning, *Reason and Resonance*, posits a decline in the resonance model of hearing around 1928. I'm arguing here that resonance continued to be very influential in other domains.

In both radio and multiplex telephone systems, engineers began to weigh modulation strategies based on their susceptibility to noise and distortion. In the milieu of the Efficiency Movement in the early 20th century, the ideals for speech communication included ruggedness in the face of noise, canalization, I mean, all messages should reach their destinations, uninterrupted by extraneous signals, and compactness of the signal and speed of transmission. These latter ideals sometimes conflicted with the ruggedness imperative.

Amplitude modulation quickly began to be understood as prone to defects. According to *Transmission Systems for Communications*, an AT&T handbook first published in the 1950s, AM was subject to quote, "a number of types of impairment, noise, crosstalk, modulation distortion, and overload." In fact, all of these handbooks had a section, a chapter, until they stopped being published in the '80s titled *Transmission Impairments*.

Unintelligible interference between AM channels was known as monkey chatter. Crosstalk was the interference of other people's speech. Other people's speech becomes a problem in the telecommunications paradigm. Impairment was then, and continues to be in many branches of telecommunications engineering, the umbrella term for things like noise, distortion, attenuation. This is another shared concept between deaf pedagogically and engineering. I don't have time to give you all the evidence for this, but impairment became a favored legal and medical term in the US in the early 20th century, the rehabilitation era, because it implied injury and amenability to repair as opposed to innateness which defect implied.

[00:20:30] It enters telephone engineering around the same time, partly because Harvey Fletcher was collaborating with various physicians who worked for the New York League for the Hard of Hearing. As I've written elsewhere, telephone lines were rated using subjective listening tests to measure their articulation, and to measure the impairments of things like noise. The efficiency imperative drove the search for other forms of modulation, almost all of which I'm going to skip over, with special apologies to any radio historians or FM fans.

Experiments with the digital coding of telephone signals began with a technique called pulse-code modulation, or PCM. Although forays into PCM date to the first two decades of the 20th century, the technique was successfully developed by Alec Reeves for the French telephone system in 1937 as a means of sampling signals and transmitting pulsatile representations of the amplitude at those instance. This is some textual and an image excerpt from his US Patent.

Because PCM signals were unequivocal, either on or off, they were more resilient to noise and easier to regenerate. In the US, PCM was not commercialized until 1962 when AT&T launched the T1 Carrier System, which was the first digital voice transmission system.

This early means of digitizing began with speech coding and moved into visual applications, becoming the predominant form of digital coding by the 1970s. With PCM, encoding began to be used as a synonym for modulation. And you can see this if you sort of go through all of these transmission systems textbooks. I also want to flag that a figure-ground shift occurs in the early 20th century where the language of telecommunications is taken up by the speech sciences, and begins to be used to describe the actions of the voice, and also the actions of the ear.

After computer automation in the 1960s, the manipulation of signals became increasingly complex, and was increasingly described as processing, a category broader than that, but overlapping with modulation. So signals could be processed and then used to modulate carrier waves. Alternately, certain types of modulation, like PCM, were subsumed under the term processing.

The verb form of the term process, meaning to operate on mechanically according to a set procedure, came into use in the late 19th century in the arena of food processing. The term processing in communication was at first applied to pulse-code modulation and to the analog compression of signals, with tools such as the vocoder.

By 1993, the Signal Processing Society would state its broad mission as, and this is a quote, "the theory and application of filtering, coding, transmitting, estimating, detecting, analyzing, recognizing, synthesizing, recording, and reproducing signals by digital or analog devices or techniques. The term signal includes audio, video, speech, image, communication, geophysical, sonar, radar, medical, musical, and other signals."

And as we've seen from the biologists here-- and I should say that my original training was in the field of biology, I stopped after a master's. That is also an important-- I don't know if the Signal Processing Society claims control over the use of the term signal in the biological fields, but it's also moved there.

Using the very imprecise tool of Google NGram Viewer, it seems that there was a take off in the use of the term modulation in the mid 1920s, peaking once in 1949, and again in the 1980s. And this lines up with things like AM information theory and the commercialization of PCM.

In 1937, the term enters biology in an article by William Bloom at the University of Chicago, who credits a personal communication with, I think, maybe the better known biologist, Paul Weiss, who was originally trained in mechanical engineering. Here, modulation becomes an explanation for cellular changes that aren't programed. They were searching for a term for cellular changes they didn't fall under the category of differentiation.

So modulation means the reversible variation in the activity or form of a cell in response to a changing environment. Unlike this usage in biology, in telecommunications what we've seen in my sort of rapid capsule history is a shift from modulation as inflection, the uniqueness of the individual voice, to modulation as a mode of training and normalization, to modulation as an automated procedure for handling signals in an efficient way.

[00:25:09] This is a shift in purpose, or aesthetic, rather than definition. Each of these cases relies on the same confidence in vocal, bodily, or signal malleability. It just depends on if you interpret that as suppleness or governableness, my own speech therapy is going to fail with that.

For philosopher Gilles Deleuze, and his often cited short article postscript on the societies of control from 1992, the so-called disciplinary societies of the 18th through 20th centuries were being replaced around that time, the 1990s, by societies of control. If disciplinary societies were marked by institutions, the prison, the factory, the hospital, the school, societies of control were marked by individualized and continual regulation. Perpetual training replaces the school, in this short article. Things like pharmaceuticals and genetic engineering replace confinement in a hospital.

Drawing on the work of another philosopher, Gilbert Simondon, Deleuze argues that the central operation of disciplinary society is molding. And that of this new form of society is modulation. According to Simondon quote, "To mold is to modulate in a definitive manner. To modulate is to mold in a continuous and perpetually variable manner."

I'm less interested in this mold, modulation distinction then in the fact that modulation then has moved from, at this point in time, from electronics into critical theory. It's further evidence for my argument about the emergence of modulation as a keyword, again in the Raymond Williams sense, from the late 19th century to the present. Yes, modulation is an old term but its senses of transmission, communication, and control begin with phonetics, are directly imported into telephony, and then into an ever-expanding circle of fields.

I'll conclude with this thought regarding the claims Deleuze and Simondon. Modulation can, of course, exert a rigidly molding affect. Think of the signal processing -- and again those two terms are not identical, but overlap -- that yields auto-tune, or the software in digital hearing aids. These tend to have a standardizing effect, with little user control. Of course that's not true of all kinds of signal processing, but modulation can have rigid effects.

And the seeming molding of discipline era modulation-- in other words, its institutions and pedagogy in the 19th century and early 20th century-- was hardly definitive as a Simondon suggested. For instance, Bell's star pupils, Dudley and Jennings, never fully entered the hearing world. Dudley married a deaf man, Wallace Krause, who was a well-known metal engraver.

They became active members of the Boston Deaf and Dumb Library and Lyceum, which had been founded quote, "to provide a pleasant home to all deaf-mutes," of course this is very old fashioned language now, "where they can meet their mates, read the papers and books, enjoy the pictures, and talk as much as they please. For it is a blessed boon to do so after being deprived all day of the chance of making themselves understood."

In 1906, Alice Jennings published an article in the *Silent Worker* titled, "Is it Beneficial to the Deaf Oralist to Learn the Sign Language?" She answered, unhesitatingly, yes. The confidence of her school years at Horace Mann flagged after graduation as she found that hearing people did not accept her.

She joined a meeting group, the ABB Society -- I have not been able to find that with this acronym means, if someone could tell me, that would be great -- that brought deaf oralists and signers together. Completely contrary to the way, the history of deafness is usually told, about this extreme division between oralists and people who use sign.

This society here in Boston brought them together. She studied sign language, which broadened her social sphere. She added, "The study of signs has, I think, increased my own mental power. I have written more vividly, more pictorially, through having these pictured words constantly before my mind."

All of these ways to communicate, beyond the universal, and the efficient. Thank you.

Josh McDermott: Greetings, everyone. Thank y'all for coming. I'd like to thank the organizers for inviting me. I'm Josh McDermott. I'm a professor here in Brain and Cognitive Sciences. And I study how people hear.

[00:30:06] I'm interested in how people can do quite remarkable things with sound that we still don't know how to get machines to do. I'm interested in why those abilities break down when some of us start to lose our hearing and how we can build devices that can help us hear better.

And because we're going to be talking about how humans hear, I want us to just start out by doing some listening. What I'm going to play you here is a scene that I recorded. I just went everybody to close your eyes, and just listen to this, OK?

[RECORDED CHATTER]

OK, so that was just a scene that I recorded in a cafe on my iPhone. And, the point of playing this to you is just to point out what we do every day without really thinking about it. And that is that, you were able to listen to that and understand that there were people sitting in a restaurant. You can hear a couple different people. You could tell that they were relaxed. You could hear music in the background. You could hear dishes clattering.

And what's remarkable about that is that you were making all of those inferences from sensory input that really just, [? to the first order, ?] looks like this. There was a sound wave that was travelling from the speakers in this room, through the air. It entered your ear and it cause you're eardrum to just wiggle back and forth. And so that graph plots the pressure that would be recorded at the eardrum, which would really be equivalent to the displacement of the eardrum over time.

And so from that pattern of wiggles, your brain is able to infer all that stuff about what was happening in the world. All right, and so as a listener that's really what you care about. You're not really interested in those wiggles, per say. You're interested in whether it was a dog, or a train, or rain, or people singing, or what have you.

So that's really what I study. I'm interested in how it is that people derive information about the world from sound. And the reason that is really interesting and remarkable is that most of the things that we care about, as listeners, are not explicit in the waveform. In the sense that if I just show you that picture and I ask you what's there, well you don't know. And if we run simple machine classifiers on that waveform, they're not going to know either.

That's why you need a brain. You have to take that waveform and process it and transform it in ways that make the information that is of interest to us more easily recoverable. So, the process of inferring events in the world from sound, we call auditory scene analysis. It is one of the main things that my research group studies.

And one of the main components of auditory scene analysis is something that everyone here has encountered, and it's called the cocktail party problem. And it refers to the fact that when you're in a restaurant, or a cocktail party, or on a city street, your ear receives a mixture of sounds from a bunch of different sources. In the simplest possible case, there might be two sources, these blue waveforms here shown together.

And those sounds would originate at, say two different points in space from two people, they would travel through the air, and they would just sum together at your ear. And so what would enter your brain would be this red signal here. It's a mixture of two sources.

And the trouble is that, as a listener you're not really interested in the red thing, right? You're usually interested in one of the blue things, what some particular person was saying. So you have to somehow take the red signal and from that, extract the blue one that you're interested in. So for instance, you might be interested in hearing this.

She argues with her sister.

But what might enter your ear could be this.

[TWO INTERPOSING VOICES]

Or maybe this.

THREE INTERPOSING VOICES]

Or maybe even this.

[MANY INTERPOSING VOICES]

All right, now what's kind of amazing is that even down here at the bottom when we've added seven additional people to that one speaker, so it's an eight person cocktail party, you could probably hear, to some extent, what that target speaker was saying. Now next to each of the sounds, I've depicted what we call a spectrogram. So this is a way of taking a sound and turning into an image that you can look at it. And it depicts the frequency content of a sound, shown here on the y-axis, as that evolves over time.

So when you look at the structure here for the single speaker, you can see that it's kind of complicated and rich. And we think that your brain uses that structure in order to understand what the person was saying. And what's interesting is that as we add people to the cocktail party, we go down this column, you can see that structure that was apparent in that top panel becomes progressively more and more obscured. Until by the time you get to the bottom one, it's kind of amazing that you could do anything useful with it at all. And yet as you hopefully experienced, people do remarkably well.

So this particular problem, which is also known as sound segregation, the ability to pull out a particular sound of interest from a mixture, is what we call ill-posed. So a lot of interesting perceptual problems have this character. And what ill-posed means is that the problem, in the general case can't be solved, right? So it's not possible to take a red signal that's a mixture of two other signals and to determine what the components sources are.

And it's really akin to me telling you, well x plus y equals 17, please solve for x. And you guys may have not taken a math class in a while, but if you had been in high school and had gotten this equation on a test, you would complain. Because it's not something that you can in general solve, right?

[00:35:11] You could have 16 and 1, that equals 17. You could have 15 and 2, 14 and 3. So on and so forth. There's an infinite family of pairs of numbers that add up to be equal to 17. But that's exactly the problem that your brain is solving all the time, every day without you thinking about it, when you get mixtures of sounds and you're able to understand what one or more of them might be.

And so what makes this problem hard is that there's lots and lots and lots of combinations of sound they could have generated the red signal. So there's the blue ones in the world, but there's also lots of pairs of these green ones. So somehow or another, your brain has got to figure out what the blue ones are and not infer all the green ones.

So how on earth do we manage to do this? Well because this problem is ill posed, the only way you could possibly solve it is by making assumptions about the nature of the actual sound sources in the world. And we're able to do this because real world sound sources have some degree of regularity. And so if you know what those regularities are, you can make assumptions about that and I could help you constrain the solution.

And so just as a demonstration of the regularities of real world sounds, I'm just going to play you a bunch. So here's one.

[BEAR ROARING]

Little quiet, it's supposed to be a bear roaring.

[DOORBELL]

[HAWK SCREECHING]

It's a hawk.

[BRUSHING TEETH]

It's brushing teeth.

All right, so those are just four examples of everyday sounds, the kinds of things that you hear all the time. And what I'm going to do now is I'm going to play you three examples of sounds that are fully random. All right, so we're in the digital domain here. So a sound is really just a list of numbers, where the number tells you at each moment in time what the sound pressure would be.

And so the way that we would generate one of these random sounds is, at each of those points in time we just randomly choose a number. And then we get that string and we can just play it out as a sound waveform. And if you do that, you'll get something that sounds like this.

[STATIC]

Or this.

[STATIC]

Or this.

[STATIC]

[LAUGHTER]

OK. So those are noise, right? That's what noise is, it's random. And so the point is that, if you randomly generate sounds, you don't get things that sound like natural sounds, right? You'd have to generate an awful lot of these things until you got something that sounded like a doorbell. It wouldn't happen by chance.

So what this tells us is that the actual sounds that occur in the real world are very, very, very, very, very small subset of all possible sounds. And so, that means that there regularities to them. And we believe that your brain has internalized those regularities and then you make use of that when, for instance, you solve the cocktail party problem. All right now -- and, yeah, so we rely on the regularities of natural sounds in order hear.

So what I'm going to be talking about today is a second scene analysis problem that has received much less attention scientifically. And I think it'll be interesting for you hear about it because it relates a little bit to some of the issues that we've already heard about, and to the Lucier piece that you will hear tonight. And this second problem is due to the fact that sound sources interact with the environment on their way to your ear.

And that's depicted in this picture here. So we have is a guy sitting in a room. There's a speaker here that is producing sound. This is the person. And each of these lines depicts a path that sound could take on its way to the listener's ears.

So the green lines here shows you the direct paths, right? They go straight from the speaker to the listener's ears. But the sound can also reflect off the walls in the room. And so the blue lines here show examples of paths where the sound would make a single reflection of the wall and travel to the ear.

The red lines here show the paths where the sound would make two reflections before reaching the listener's ears. And of course, there's lots of other paths where it might reflect three times, and four times, and so forth. So in addition to the sound that is coming directly from the source, there's all these reflections that are coming from the interaction of the sound with the walls. And that's known as reverberation. And it's present in almost every environment that you'll be in.

And so what reverberation just is, is it's just the summed effect of all these reflections, right? And what are the reflections? Well they produce delayed, and usually slightly filtered copies of the original sound. So the blue waveform here would be the sound that would come directly from the sound source to the listener.

But then there would be the cyan one that would arrive a little bit later, after the sound is reflected once off of one of the surfaces in the room. And then this one would come. And this one. And this one. And this one. And you have this very, very large set, and those things all add up, and that's the sound and that enters your ears.

So reverberation, even though it's present everywhere, and poses a big challenge to the auditory system, as we'll talk about in a bit, really hasn't been studied very much scientifically. And this is kind of a cool picture that I think kind of encapsulates a lot of about the state of the field with respect to reverberation.

So this was a photo of an experiment that was done at Harvard back in the '30s. And they were trying to study sound localization. And they really wanted to study sound localization in its purest form. And so they really wanted to get rid of all of the echoes. And so what they did, is as follows, it says here, "In order to avoid possible reflecting surfaces, a tall swivel chair was erected on top of a ventilator which rises nine feet above the roof of the new biological laboratories at Harvard University. At the present time this procedure appears to be the only practical means of avoiding errors due to the reflection of sound."

[00:40:35] So this point, they hadn't invented anechoic chambers. And so this is what you had to do. You had to put people way up on the roof. And they wanted to do this because it was important to eliminate the effects of reverberation in order to measure what the brain could do.

But reverberation is ubiquitous, and it's a major challenge for the brain. One of the reasons is that reverberation profoundly distorts sound signals. And we don't realize this because our brains are very, very good at coping with this. But what I'm showing here are the sound waveforms that are produced by what I'll call a dry sound signal. So it's a sound signal without reverberation that you might measure in a very small room without many echoes, with the microphone close to the speaker. So it might sound like this.

They ate the lemon pie. Father forgot the bread.

And it's probably appropriate for me to say at this point that this is a great research topic. It's a little difficult to give presentations in large rooms though. Because the demonstrations that I will be playing you will have the effects that they're supposed to have coming out of the computer coupled with the acoustics of this room.

So this room actually has surprisingly good acoustics. I was thinking that this was going to be hard to play demos in here. I think some of them will kind of work, but just keep in mind that everything is going to sound a little bit impure. I gave a talk on this topic over the summer at a conference at Bates College and it was held in a concert hall. And it was pretty much a complete disaster. Because I was trying to play people all these different variants of reverberation, everything just sounded exactly the same because it was all being dwarfed by the room's reverb. But anyhow, you can hear the contrast, I think, with this version here.

They ate the lemon pie. Father forgot the bread.

Sounds like the person's in a train station or something like that. OK, but you're able to understand what was being said. And what's interesting about that is that if you look at the sound waveforms, these things in blue, you can see that they look very, very different. And below them, I have another spectrogram, so remember that's an image that depicts the frequency content as it changes over time. And you can again see that that looks very different in the two cases.

And the reason that's interesting is that if your brain has learned to recognize this particular pattern, if it then gets this particular pattern, it's not really obvious how you should go about determining that those two things are the same. And indeed, one finding that underscores that is that reverb is a big problem for machine speech recognition. So this is a plot here of the performance of five different state of the art algorithms, as of a couple years ago.

So we're plotting the percent errors they make as a function of the amount of reverberation. And so zero here means that there's no reverb, it's the dry case. And you can see that there are not very many errors.

But if we just introduce a little bit of reverb, and here it's being measured in units of milliseconds, and I should say this is still very modest. This is less than you would get in a classroom, right? So it's not something that would ever cause a problem to a human. You can see that the errors really jump up to like 50%, then you make it a little more, and it's getting killed.

So it's a pretty hard technological challenge, and one of the big obstacles for speech recognition right now. And it's also a challenge for people that are hearing impaired. But reverb also provides us with information. One thing that it can often tell us is something about the distance of the sound source from the listener.

So this is just a picture that's depicting a listener in a concert hall, this person here. And there's a person in back who's whispering something in the guy's ear. And then there's the performer on stage.

And so in both cases, there is sound that enters the listener's ears directly from the source, as well as after reflecting off of all the surfaces in the room. But in the case where the listener is closer, the direct sound will be higher in amplitude, because the person is right there. In the case where the performer is far away, the direct sound will be lower in amplitude relative to all of the different reflections. And so that ratio of the direct-to-reverberant energy is a cue to distance. And that's something that we think people use.

And it also tells us about the size of the room. So you can imagine if we took this room here, and we just kind of shrunk it down. Well all of the path lengths from the sound source to the ear would be a lot shorter. And so those echoes would arrive a lot closer.

And that's a difference that will be audible to people. So you can tell the difference between the sound of a small room and a sound of a big room. And reverb is, in part for this reason, a primary ingredient in music production. So it's very common, for instance, for, say different instruments to be given different amounts of reverb. Sometimes, in part, because it makes them sound like some of them are close and some of them are far.

This is an example of a song that you guys probably know, it's "For the Love of Money" by the O'Jays. And I've got just the introduction here. It's just a bass line. What the engineer did is he plays the first little riff, which is the first five seconds here. With quite a lot of reverb on it. It sounds like it's in a stadium. And then he turns it off over the second repetition.

[00:45:13] And so what I've done is I've just chopped out the second repetition, and I'm plotting it below the start of the thing. And so you can see visually, they look very different. And you'll be able to hear the difference hopefully.

[MUSIC PLAYING]

So you hear how the second version kind of -- something funny happened, right? It got a lot drier. Again, these effects are diminished in this room. But visually, you can kind of see the difference. So the fact that these things are used as tricks in music production, again, is evidence that the brain has got a lot of sensitivity to this.

OK, now to measure reverberation in a standardized way, we do something called that's called measuring an impulse response. So what's an impulse response? Well it's literally the sound of an impulse in some space. What's an impulse? Well it's a very, very short high amplitude sound. It's like what you would get if you fired a gun, as shown here.

And so the same thing that's going to happen when you fire the gun that happened with any sound source. That is, you get direct sound, straight from the source to the listener, or in this case, the microphone. And then all these delayed copies, which are the reflections off of all the different walls in the room.

And so you get a sound out, just like you would if you recorded any other sound source. But it's a sound of an impulse. And that impulse response captures all of the properties of the reverberation in the room. So it's a simple way to measure reverb.

And so here's an example.

[CLAPPING SOUND]

So that was a very quiet impulse response.

[CLAPPING SOUND]

This was recorded in a classroom. And if you plot the waveform of that, it looks like this. So you can see that before you could present the impulse, there's nothing going on, it's just silent. But then you get this big peak here, and that's the direct sound; the sound from the source of the impulse to the microphone.

And then there's all these other reflections. Early on you can kind of see them discreetly. And then there's just lots and lots of them, and they kind of blend together. And eventually it fades away as the energy gets absorbed.

OK, so the situation here is that we have this sound that enters our ears, here. And that's the combination of the sound that originated from a source and all of the properties of the environment in which that source made its sound, which we can summarize by this thing called the impulse response. Now mathematically, these two, things the sound from the source and the impulse response, they're combined f operation called convolution. It's not really important that you understand what that means. The point is just that their effects get combined, and they produce, together, the sound that enters your ears.

So the problem for the listener is that, again, as was the case in the cocktail party problem, where you're not really interested in that mixture of sounds, right? Here, you're not really interested in the convolution of those two signals, right? You're probably interested in the sound source itself. You might also be interested in the properties of the environment, if you want to understand how big the room is, or something like that. But you're not interested in their combination. And so one of the questions that my lab has been working on over the last couple of years, is whether we can view the perception of reverberation as a process of separating the sound source from the reverb.

Now we have some evidence that this might be plausible. As I showed you, humans can recognize sounds pretty well despite reverberation, at least if it's modest, and if their hearing is healthy. And I also gave you some examples of how we can derive information about the environment from reverberation. You can tell whether you're in a big room or a small room, for instance, or the sound source is near or far.

But remember, that just as was the case with the cocktail party problem, this is a problem that is what we call ill posed. So you observe this one signal, and you would like to infer one or both of the signals that actually got combined in order to generate that, right? And so in general, that's not something that you can possibly do. The only way that you could solve this is by making assumptions about the sound source and the filter.

And so what we've been working on is the notion that any assumptions about reverberation that would be made by the brain, are likely related to the acoustic regularities of the world. And so what we've been trying to do, is to just measure reverberation in lots of different real world condition, and actually to see whether it exhibits regularities, and then to test whether those regularities have maybe been internalized by human listeners. And so I'm just going to, very briefly, tell you about that work.

And this is work that was done by a postdoc in my lab by the name of James Traer. OK, so the first question is, well, what's the actual distribution of environmental impulse responses? Right, now you might imagine that these could just be wildly varied. I mean it could be any possible pattern of reflections that you could get in a space. So we wanted to just measure it and actually see what it was like.

So we built an apparatus that enabled us to measure these impulse responses. And it turns out that for various reasons, you don't actually want to go around firing guns and measuring the response. So we do something a little different, but conceptually it's the same.

So there's a speaker here that plays a sound signal, and we know what that signal is. And then there's a recording device here that records the sound that would result from that in any particular space. And here they're set up in an elevator in building 46. And so we can take the recording that we make, as well as the signal that we put in, which we know what that is. And from those we can actually derive the impulse response, which summarizes the effect of the reverberation in that particular space.

[00:50:26] So we had a bunch of people participate in a survey. So the goal here was to actually sample the reverberation that listeners encounter in their life. So people who signed up for the survey, they got sent 24 text messages a day, at randomized times. And when they would get that text message, we had their phone programed to send us back their GPS coordinates. And people were also instructed to reply to the text that they got with the photograph of where they were, as well as the address.

And so with the photograph and the address and the GPS coordinates, we could usually figure out where they were. And then we would go to the space, and measure the impulse response, and wherever it was at the time. And so we've got impulse responses now in bars and restaurants, on city streets, in the forest, in bathrooms, hallways, all the kinds of places that people spend their time. And so we just want to see what they look like, and whether they exhibit regularities.

And so at this point we've tested seven people in this for two weeks each, and so we've got about 400 locations. So remember, at each of these locations, we're going to get out something that looks like this. It's an impulse response for that space that summarizes the reverb in that particular place.

And so to analyze this, we've looked at how different frequencies within the impulse response decay over time. So what I'm showing you here is just the raw sound waveform, but we can take that and we can filter it, and we could look at what happens to particular frequencies over time. And it turns out that when you do this, you see a lot of surprisingly regular structure.

And so what I'm showing you here, is the frequency decay for two different locations. An office, which is this one here shown in this light color. And a forest, which is shown in the blue. And the important point to take away from this is that these things look a lot like straight lines. OK?

Now, for technical reasons we're plotting these things on a log scale. And that means that what we're seeing here is something called exponential decay. And it's not really critical that you know what exponential decay is. It's a very simple mathematical form.

The important point is just that there's these regularities that we see from one impulse response to another, which is that the decay has got some kind of stereotyped form. You might also note that if you look at this, you can see that the slope of the lines is not exactly the same in all the different frequency channels. And in particular, it seems to be more shallow here at low frequencies than it is at high frequencies.

And that's what's just shown in this particular graph here. We have a plot of the decay rate going from slow over here, to fast over here as a function of frequency. And then the point is that up here at the higher frequencies, the decay becomes a lot faster. That means the energy is kind of going away a lot more quickly than it was at the low frequencies.

And you could think of that as a product of the absorptive properties of typical services and of the air. So as the sound bounces around, those high frequencies are getting absorbed more quickly than are the low frequencies. And that's just something that we see.

And this is a summary graph of all of the impulse responses that we measured, showing that down the decay rate here is a function of frequency. And the point is that at lower frequencies, the decay tends to be slower. And at higher frequencies it tends to be faster. So that's just something that we see.

So the take home message of this particular project, is that the impulse responses that you observe in the lives of the average person are fairly stereotyped. So the energy pretty much always decays exponentially, and it does so faster at high frequencies than at low frequencies. And so why is this interesting? Well it raises the possibility that listeners could employ fairly strong assumptions about the nature of reverberation when they get these signals into their ear, and they're actually trying to separate the effects of the reverb from those of the sound source.

And so we've done some experiments to test whether this is the case. And what we do is, we generate synthetic impulse responses that either mimic real world reverberation or that deviate from it, for instance that don't have these particular properties up here. And so one initial simple test is to just listen to the impulse response itself. And so if your brain is interpreting that as reverberation, it should sound like an impulse that you're hearing in a space, like somebody fired a gun. And so if you do this in actual rooms, these are actual impulse responses.

[CLAPPING SOUND]

You hear stuff like this. This is 26-100 here at MIT.

[BANGING SOUND]

That's a pretty big room. And so you get this impulsive thing. It sounds like it's being played in a room. And so, if we do the same exercise with these synthetic impulse responses, if we use a case where we've mimicked the properties of real world reverberations-- so this is something where we have exponential decay that's got the correct asymmetry with frequency--

[BANGING SOUND]

You get something that sounds a lot like 26-100. Here what I'm going to play you, though, is something that's pretty cool. What we've done is the exact same thing. So we've synthesized an impulse response, but we reversed the relationship between frequency and decay.

[00:55:06] So now the low frequencies will be decaying fast and the high frequencies will be decaying slow. What's really interesting about this is that, it's very clear that your brain does not want to interpret this as reverberation. It kind of sounds like maybe somebody opened a soda can. Listen to this.

[POPPING]

So again, it's exactly the same thing. It's just noise that's decaying exponentially. But it's not decaying fast enough at the high frequencies, right? And so this is some evidence that your brain is actually-- it's incorporated this regularity into the way that it processes sound. And so when you get something that doesn't match that, it says well that's not reverb. That's actually like a source in the world, like a soda can opening.

All right, and so one interesting prediction is that the ability of a listener to extract properties of the source ought to be better if the impulse response conforms to the real world distribution that we've observed here in the world. So if listeners are actually using assumptions that are based on these regularities, well then it ought to help them when the reverberation actually is ecologically valid.

So here's just a simple experiment that we did to test that. So what we've done here, is we're presenting people with two different sounds, they're just random patterns that we synthesized. And these are spectrograms that just show that they've got particular structure.

So we will get a target sound, and then they get a probe sound. And they're just asked, well is the probe sound the same as the target, or is it different? In this particular case, it's the same, and you can just see by looking at that.

So this is a very simple case where both the target and the probe are what we call dry. So neither of them has any reverb on it, right? So you just have to say whether they are the same or different.

So the interesting case is the one that is shown down here. So now, we've added reverberation to the target. And so we're asking people, is this probe sound-- does this have the same structure as the sound source that's embedded in the reverb in the target sound? And remember, what our prediction is that if the reverberation is faithful to the properties of real world reverberation, people ought to be able to do that task. But you might expect that they would be worse if the reverberation is artificial and deviates from the properties of real world reverb.

And indeed, that's what we find. So this is just a graph that shows how good people are at matching those sounds as a function of the type of reverberation that we give it. And so if no reverb, well they're the best there. But if we look at this green bar, that's the case where we've mimicked the properties of a real world reverb, and they're actually quite good. So they're pretty good at saying whether this probe sound was the same as this thing in reverb.

And all these other four bars, here on the right, are different ways in which the reverberation deviates from the properties of actual reverb. So this is a case where it decays linearly. These are the cases where the frequency dependent decay is wrong. Here's where we actually time reversed it, and so people are always worse. And so this suggests that humans, indeed, are making assumptions about reverberation that are consistent with the regularities that we find exist in the real world.

Finally, you might ask, well can listeners estimate the impulse response properties from that reverberant signal? So we just did a test showing that they can extract some properties of the source. We'll can they extract properties of the impulse response?

And so here we played them three different sentences. Two of them were being played with the same impulse response, so like the same room. And a third one was being played in a different room that could either go first or last. And people just had to say which one was recorded in a different environment. So it would go like this.

A large piece of engine cabin vanished.

Tragedy presumes such a configuration.

She saw me, and sat down beside me three feet away.

So was it first or last?

Last.

Last, that's right, yeah. So, that's what it was. And so the question here is, is your ability to do that task going to be dependent on whether the reverberation actually acts like real world reverberation or not. And so again, we have performance here on the y-axis, this is how well people are doing.

And this blue line here is when we have exponential decay, mimicking the properties of real world reverb, and people are pretty good. And you can see here that these other lines, which are the impulse responses that don't really mimic the properties of real world reverb, people are worse at this task. So they're not as good at extracting the properties of the environment when it doesn't really behave in the way that your brain expects it to.

All right, so just to summarize here, I've argued and hopefully explained that many of the perceptual problems that our brains are so good at solving are what we call ill-posed. You can't really solve them in the general case. And they require the brain to make assumptions about the nature of the world.

And we believe that those assumptions are rooted in the statistical regularities and the fact that the world is not random, rather it's got regularities. And so reverberation is another example, like the better known “cocktail party” problem. And we've tried to study it in this framework.

So we've measured the distribution of environmental impulse responses that people encounter in their lives. And we find it they're highly constrained. They exhibit exponential decay that's frequency dependent. And we've done some experiments to show that the recognition of sound sources and the discrimination of impulse responses is better when the impulse responses are faithful to the empirical distribution.

[1:00:06] So I think this is suggestive of a separation process that uses pretty strong assumptions about reverberation in order to deal with its effects. And we think this is either learned from experience, so you grow up in environments that have lots of reverb and maybe learn their properties. It could also be built into the auditory system from birth.

So with that I'll just acknowledge my collaborators. As I said, my postdoc, James Traer, did a lot of this work. We got a lot of help making these measurements from these people. And we're supported by the McDonnell foundation.

Stefan Helmreich: We're now going to hear a commentary from Alex Rehding of Harvard Music. And I'd like to invite all of the panelists to come up and array themselves on these chairs. Alex.

Alex Rehding: Thanks so much. This is an interesting position to be in. I had a vague idea of what everybody was talking about because I know their work. I did not know what exactly they were going to say. So if you saw me, I was scribbling down notes, furiously throughout the morning. And I'm still trying to order what I heard.

We had some wonderful presentations in a variety of formats on resonance. And various ways in which this term resonates in different fields. Josh's presentation, the last one that we heard, actually provides me with a metaphor from room acoustics. One of the basic problems of room acoustics is that you really want to find an ideal medium between two properties. On the one hand, spaciousness, a kind of sense of the size of the room. For that you want a lot of reverberation. And definition, clarity, for that, you generally want little reverberation time.

And so it's often hard to find the right balance. And I see it as my task to, on the one hand, work with the spaciousness, and the kind of a multiplicity of possibilities that we've been presented with, and the kind of focus that we've been working here. The term resonance, as was already said a couple of times, was brought to our attention again recently in Viet Erlmann's book, *Reason and Resonance*.

And he thinks of it as an epistemic technique. A kind of epistemology that he juxtaposes to reason, and that interacts with reason in interesting ways, in different constellations, at different periods of time. And as Mara mentioned, Viet Erlmann perceives a decline of the epistemic power of resonance around 1928.

And that's an interesting time. I will come back to this. What we heard were three different contributions from the arts, from technology, the history of technology, and from science. We heard three different approaches. One of working with sounds by composing, one with experimenting on sounds and audition. And one about how tools are provided, how tools shape that help us in communication and in audition, in general.

What all of these approaches connect, is I think that, they're obviously all ways of understanding of audition. And I think the underlying commonality is that they all encourage us to listen to ourselves listening. What that means in each case is something slightly different.

[1:05:04] Stefan's wonderful introduction brought us back to the notion of sound as an organism, as an individual, as life. Which is something that seems irresistibly romantic. And that's something that I want to come back here. I mean romanticism here both with a capital and with a small r. But it takes us back, in some ways, to the 19th century, so in that period before 1928.

Mara highlighted, in her wonderful talk on telephony and -- let me just see -- and on the role of Alexander Graham Bell, and his notes with his deaf students. She noted the work on Helmholtz, and the analysis of sound Fourier analysis. Ways in which sound was written down. In which it was fixed, and in which it could be worked on. That is another really important aspect, I think, that has helped us come to terms with sounds and to communicate our listening.

Josh has highlighted the role of tension. Of the ways in which we filter signals into noise and meaningful sound, the cocktail effect. And the role of our prior assumptions, our prior experience, priors in general are important in making sense of the noise.

So when I say that some of this takes us back to the 19th century, perhaps the notion of -- going back to Stefan's introduction-- the notion of sound as an organism, of an individual, of life, is something that resonated really strongly with the German idealist philosophers, Kant, Hegel, and Schopenhauer, who start to latch onto music as a vehicle for the sublime. And they really think about music in ways that are different from later ways of music. Because for them, music is a strictly temporal art.

It is -- during the early 19th century, music makes a radical shift from one of the lowest art forms to the highest. And that's always puzzled commentators, how this could happen. But it seems -- and this is something that to the media theorist, Friedrich Kittler's talked about a lot -- is that the temporality of music, and the certain sonorousness of music were central to the idealist philosophers reevaluation of music because they saw it as the ideal vehicle for the sublime. And temporality is one thing that they had identified as a key issue in the sublime, especially the problem of the finitude of life.

And so they saw many important resonances between music as a sounding phenomenon. And that's in marked difference to previous philosophers, and their views of sound. And this is also the rise of organism, individual, and life. So this takes us back to the early 19th century.

And I've been thinking, this morning, about one way of bringing these various aspects together in one focal point, a prism, maybe. It's hard to come up with a sounding metaphor, here, a resonant chamber maybe. But one way in which the sciences, the arts, and technology can come together as they have done, here, in this morning session, I could find no better example than the modern siren. Which is a technological device that was invented in 1819 by a French engineer, Charles Cagniard de la Tour.

He was a major influence on Helmholtz, in his book on the sensation of tone. If you open the book, you will find that it starts with a detailed discussion of the mechanism of the siren. And what the siren did, was really to allow us to capture sound, this ephemeral phenomenon, and to analyze it in terms that would make sense, as it turned out, to scientists, to engineers, and to artists.

I'll just try to explain, in very brief terms, how the siren worked. If any of you've seen a mechanical crank siren, that's the one I'm talking about, not the modern electrical siren. So the siren did not become the warning signal off modern life that we know nowadays, because electricity first had to become widespread. And that didn't happen until the early 20th century.

[1:10:08] But in 1819, the device was invented as a radically new sound generator that was different from traditional theories of how sound worked, in terms of sine waves, and continuous phenomena standing waves that had to be set up in either string instruments, or closed tubes wind instruments. What the siren was trying to propose is something that, with the benefit of hindsight, we could describe as a digital sound generator. They didn't use that terminology then, but that's kind of how we would think about it nowadays.

Because the source of the siren was really a series of rhythmic impulses. The way it worked was that, air was blown into a bellows. And the air set a metal disc in rotation. This disc had holes punched into it in regular intervals. And as the disc rotated, air puffs would be generated. And each time an air puff got out, you would hear a little ff, ff, ff, ff, and so a rhythmic pulsation was created like that.

And as the disc speeded up, around 20 hertz, 20 impulses per second, so at a fair amount of speed, the perception of the sound would change. It would stop being heard as a rhythmic impulse, but it would become a pitch. The percept was fused and we would hear the rising pitch, the kind of, waaa that we traditionally associate with the siren.

And that was a pretty radical demonstration of what sound could be. That a continuous sounding phenomenon could be created from discontinuous impulses. And that's, because of the ff, ff, ff, on/off impulses, I think it's fair to describe it as proto-digital. And the scientific world was amazed. And we have a number of responses to this.

First of all, to stay within the realm of science, this spawned a big debate about how that phenomenon could actually be explained. And two scientists of the generation before Helmholtz, two German scientists, actually picked up this problem, Georg Simon Ohm, the one hand, and Seebeck on the other. Ohm was a theoretical physicist; Seebeck was an experimental physicist. And Seebeck believed that this was a radically new theory of sound, and that all our old theories had to be rewritten to make space for this new proto-digital sound generation.

And Ohm said, not so fast, really this is still a sine wave. It's just a much more complicated sine wave than the ones that we're used to. Ohm was also instrumental in bringing Fourier analysis to the study of acoustics. And that's precisely where this is leading; he's basically saying that even if we have on and off impulses, we can still think of that as a very complicated form of a sine wave.

And Seebeck died a little later. And so the debate was unsettled. People generally sided with Seebeck until Helmholtz came along and said; well actually, you were talking about different things. What Seebeck was talking about was what I will call clang, sonority, the kind of acoustical aspect of sound in the soundwave. Whereas what Ohm talking about is really tone, what I will call tone, the perceptual phenomenon as it arises in the ear. And so that definition became pretty significant for the 19th century.

So that so far about the scientific response to this. The psychophysical, psychoacoustical response to this was one of amazement. The siren, for them, counted as a demonstration that rhythm, the kind of ff, ff, ff, pulsation and pitch, which were treated as different dimensions of music, are we on a continuum.

And, they could have known that if they'd studied their acoustics, but that was the first time that it was demonstrated, that it became directly perceptible. And so there were some people, most of all, an astronomer from the 19th century, [INAUDIBLE], who dabbled in music theory. And he wrote this book that he was hoping would be the final word in music theory where he said, "The siren explains all aspects of music. I can you derive a complete theory of music out of the mechanism of the siren."

[1:15:07] And he proceeded to write it. There was a lot of excitement for about two months, and then he disappeared. In fact he tried to republish his book a couple of years later in 1850. And then this idea was picked up again in the 20th century by experimental composers like Henry Cowell, and later, Stockhausen.

And I should say that the whole reason I'm talking about the siren was because of one thing that Alvin Lucier said earlier about how the interference of sine waves creates a rhythmic pulsation. And in some ways, what the siren did in 1819 is the reverse of that, a rhythmic pulsation creates pitch. And so, this is just another demonstration of how the two perceptual parameters can be connected, and how they can be shown to be interlinked in ways that we don't normally expect.

And this brings me to the third area of response, of resonance, with the mechanism of the siren in the field of composition. And there were a number of poetic responses to the siren. I think one thing that we often forget is that in those days there was a keen interest in popular science. And when you go through early 19th century specialist music journals, there are regular articles that try to communicate the latest invention in science to a musically literate audience. And so, I think this is a fascinating sauce that's really not looked at very much.

But, one off the most incredible ones is by a journalist, composer, Richard Poore, who was an accolade of Richard Wagner. And that also comes out a lot in his response to the siren. Because he offers a poetic, very romantic paraphrase of the mechanism of the siren, where he talks about, he imagines, this experimental setup where the siren is placed in a darkened room.

And it just goes faster, and faster, and faster, and so we not only hear the rhythmic pulsation that turns into a pitch. But he goes beyond the audible range, and then it turns into colors, it goes into the visual realm of vibrations, and then ultimately to heat. And so you, presumably, at some point we reach microwave levels, at some point.

It's really unlikely that anything like that ever happened. Helmholtz reports in his experiments with the siren that you had a very hard time even getting beyond the audible range, so that's 20,000 hertz. Which really isn't very far, given how long Richard Poore said the vibrations continued. But there is this, hankering for a synesthetic experience that I think this is the impulse behind it. And that's obviously one of the obsessions of the 19th century. And that's also a topic that I think we have a considerable interest in at the moment.

And so, with this example from history, with the year 1819, I wonder if we can close and go back to the year 2014, to almost 200 years later. I'd ask, is this another time where the science, the arts, and technology are poised to come back together again in these exciting ways? Do we have a sense of this kind of resonance? And with that, I will close and open the floor to questions.

[1:19:18] **Question & Answer**

Stefan Helmreich: Thank you for that, Alex. And perhaps prior to opening it up immediately to questions and answers, do panelists want to respond to, riff on, any of the things about which Alex spoke?

Mara Mills: I mean, just do your very last question, I think if you work on a subject like media technology, something like video or telephones, in a way, whether it's 2014 or 1920, you're always dealing with an object that is interdisciplinary. I mean, some of the early leaders of Bell Labs immediately realized that what they were working on was something like applied phonetics. They had to deal with the production, the propagation, and the reception of sound.

So they had to study a little bit of biology and medicine, a little bit of psychology, a lot of engineering in order to create a system that could transport a sound efficiently. I mean, they wanted to know how the mouth worked, what sounds the ear could hear, physical acoustics. So it was always like interdisciplinary teams creating these large-scale technical systems around media technology.

And of course, they were creating-- there were all sorts of what they called by-products, they weren't always by-products, they often were things that inspired the telephone system in return. But that moved into domains like sound film, or experimental music, I mean a whole array of microphones and other sorts of, and loudspeakers, and other sorts of technologies. So I think in that domain, there was always interdisciplinary. But the bigger question about domains outside of media, I'm not sure I can answer. Maybe some other people have thoughts.

Stefan Helmreich: Or perhaps there are answers from the audience. So we're going to do what we did yesterday. We're going to invite people with questions to raise their hands, and we're also going to try to include folks from the overflow areas. But we'll start in this room. Caroline, you had a question.

Caroline Jones: First of all, thank you all for an incredible feast for the ears. I was interested in the ocular metaphors. I think I saw a joke, I saw an anti-Lucier I am sitting on a roof, right? I mean, so that was a wonderful little moment there. And in terms of kind of matter, antimatter, I felt that Mara and Josh both had fascinating, coming from opposite ends of a certain spectrum, a sonic spectrum that I wanted to surface.

So, if I can get this straight, Brian's tough question to Alvin Lucier about how his own experience of speech delay, and his own reading of theories about speech delay informed some of his inquiries about reverb, and repetition, and tape delay, and so on and so forth. Where the tape medium is a way of investigating his own experience of production of sound, and the hampering of sound.

I mean, that resonated incredibly well with Mara's discussion of how work on phonetics to push the deaf out of their culture of communication, through signing, toward an oral and aural tradition, and how that informed telephony, seemed to be the other side of the picture for Josh looking about how to filter out this resonance and this reverb in order to communicate through mechanical means.

So I guess if I have a question here, it is whether in the history of this spectrum of research into human hearing, and human sound production, and efficiency, and the message, and the signal, and so on, whether there are parallels to the lesion studies that informed vision research, and the kind of terrifying experiments in which David Hubel would tape kittens, he would suture kittens eyes closed. And then lo and behold, when their eyes were opened after, three months of development, they could never see.

So Josh, you ended with a question about whether this capacity to hear the environment, to use the environmental-- I was very interested in your environmental category of realistic sound. And you said, hey, you know, could be inborn, you know, could be something we experience, I think this will come back this afternoon, about consciousness being in the world, always in the world.

And so the question is whether the stupidest of experiments that were done in vision, where you just deprive the developing mammal in the most plastic period of their development from access to the environment, has that been done in hearing? And you know, Mara would have a great access to this through her historical work, so that's my convoluted question.

Josh McDermott: Sure, yeah, I can speak to that. I mean there have been-- for just methodological reasons; it's a little harder to do. I mean there have certainly been experiments where they induce hearing loss from birth, and measure the effects on the brain. And, it's very-

No, well there's cochlear implants, of course, right? Where, I mean, people will get a cochlear implant after they've been deaf for some period of time. But the cochlear implant doesn't really replicate normal hearing very effectively. And reverberation is a pretty big problem for folks in that situation. So, I mean the other issue is, well, would there be a population of say, normal hearing people, that would have an altered experience, for instance because they live in a place that doesn't really have reverberation.

[1:25:28] And we don't yet know the answer-- I mean there are populations for whom that might be the case. For instance, indigenous societies that basically live outdoors, and don't really have buildings with concrete walls, and that have different reflective properties. But we don't-- no one's really done the studies yet, that something we're interested in actually doing.

I do know that it's-- I believe it's the case the children have more difficulties with reverberation than do adults, which is consistent with the idea that it's something that you learn very gradually over the course development. But that's more anecdotal. So I think they're mostly open questions at this point.

Mara Mills: I mean, there are so many experiments in the 20th century that would be a parallel to that one, I'm not even sure where to begin. I mean, the deprivation experiments, I guess, would be something like what Georg Von Bekesy who was at Harvard, and they're not exactly deprivation. But he definitely worked with guinea pigs, damaging different parts of their ears, and then letting them live for a certain period of time, and then killing them and seeing what he could see under the microscope.

There are lots of cat experiments. I'm sure there were probably deprivation ones, but the most famous cat experiments probably start with the Wever-Bray experiment, that was done at Princeton. And many people describe this in the internal literature of psychoacoustics as the beginning of psychoacoustics. And they we're using the equipment that was loaned from Bell Labs.

And in fact, an early moment in my turn to this research, when I was a graduate student in History of Science at Harvard, I TA'd for Steven Pinker, who himself was trained in psychoacoustics. And he mentioned, sort of off hand, at some point, that psychoacoustics began with the telephone. And I thought, what does that mean? How can that possibly be true? And, you know, now I've, 10 years later, have gone down that rabbit hole.

But, you know, Wever and Bray, basically-- people thought that the ear worked just like a telephone, and they wanted to see if-- with minimal amount of processing, an analog telephone. So they thought they were placing an electrode at the auditory nerve, and then speaking into the cat's ear, and having the electrode pick up that sound. It was connected to a wire, into an actual telephone in another room. And they thought, well if the person in the other room can hear the speech, ungarbled, then the telephone, what was called the telephone theory of hearing works.

And it did seem to work, but it's a very long story, they put the electrode in the wrong place. They were picking something else up, the cochlear microphonic was what they were picking up. That inspired a ton of research at Johns Hopkins with damaging cat's ears in different ways, and trying to hear through the cats' ears what the effects of different kinds of impairments were.

I mean, and early cochlear implants were also tested on cats. I have some amazing archival photos, for people who are interested in animal experimentation. I'm a vegetarian, I don't collect these for prurient reasons, I guess. But like in Australia, Graham Clark did some; he was a wonderful photographer, as well as an experimenter. And has these amazing photos of cats that he had put very early cochlear implants into, to test them of course, before testing the on humans.

So, the cat stuff, yes. The deprivation, I don't have an exact analogue in terms of deprivation. But definitely doing damage, and then seeing what the effects are.

Brian Kane: Can I add one other, site that me might look-- these are really deprivation experiments, but it seems to me that an underutilized resource for thinking about these particular problems is the history of electronic music, and the electronic music studio. So one of the people I study very closely, Pierre Schaeffer who I argue with quite significantly in my book, is somebody who is performing these kinds of psychoacoustic experiments.

I mean, one of the things that when I hear Josh talk, there's part of me that wants to say, yeah, Schaeffer knew all this. That one of the things he gets to do by taking sounds, by playing them backwards, by adding reverberation to them, is allow us to explore the relationship between things like the perception of source, and the resonance that's, or the reverberation that's added onto it.

I mean, Schaeffer also said things like, the whole invention of the acousmatic was a way of moving away from the determinism that the signal has. So he says, for instance, that it's never on the basis of the signal that the sound object is discovered. So there's always meaning that there's this kind of underdetermination problem, right? That anything that hits the eardrum is always going underdetermine its relationship to its source and cause, and hence there has to be some other cognitive dimension that's going to be solving these problems about listening.

[1:30:01] So, and these are both research problems for him, he's working for French radio as a researcher. But they're also musical problems. These are things that show up in his own compositions, and in the compositions of the people that work in the studio. The history of musique concrete is very much a history of exploring these particular kind of perceptual problems.

How do you make a sound object? How do you obscure source? How do you obscure cause? How do you add -- you know, how do you manipulate sound such that it will start to produce these kinds of effects where we can't recognize where things come from. So, I think that that's a wonderful place where, within the electronic music studio, you start to see the relationship between science, and aesthetics, and technology.

Audience: Hi. Yeah, thank you so much. I had a question, first for Mara, and then maybe the rest of the panel might have thoughts or comments. Specifically, so you traced modulation and its history or basis in deafness. And you briefly mentioned ableism, I think when you described a passage as old fashioned.

But I wondered if you could comment more about the relation of modulation to, and that history to ableism, and Bell's relation to that, and his subjects. And then, more generally for the panel, and for Stefan to talk about vibration, and your work with disability studies, and overall that. Yeah. Thank you.

Mara Mills: Yeah, recently I was asked to write a keyword entry myself on the word deafness. And, I mean, it took me the better part of a year. It was a long entry, but it was also incur-- it just, everything I knew started to unravel as I started to research the history of this term and the present day uses of this term, in English. And, you know, first off, I mean, ableism in this specific case, relating to deafness, in the 19th century, the term deaf, you know deaf has always been a spectrum, or an umbrella. It's always been a word loosely applied to lots of different modes of either hearing, or not hearing, or lots of different modes of communication.

And, you know, and so when I started to think about Bell's students, for instance, they all described themselves as deaf, they all used multiple different kinds of communication modes. But it wasn't quite as simple as the like, sort of black and white narrative that you usually hear about Bell being completely audist, and his students being completely wrenched out of the sign language world.

Because many of them were using sign language alongside other communication techniques, all at the same time, and continued to do so after their maybe six month period of training with Bell. And in every case, the training completely failed. So it's not to say that Bell wasn't a eugenicist, he was, and he used that term about himself. And it isn't to say that, the kind of modulation that he was asking them to do was based on very restrictive ideas of what voices should be used for, and the way voices should sound.

You know, he was taking people who, especially in the case of Jennings, she could speak orally. She had become deafened when she was eight. And, you know, probably her parents or her schoolteachers felt that she had an unpleasant voice. You know, that it was modulated, it had sort of flat, the flatness of pitch, or there's lots of descriptors that were applied to deaf voices at the time. That they were mechanical, that they sounded dead, lots of sort of negative descriptors.

But in that case, I guess the ableism is really clear. Because she could articulate, she could speak orally. It was just an aesthetic practice, they were just, you know, changing her voice for the benefit of listeners who would feel comfortable by how that voice sounded. Whether it was difficult for them to understand, or whether they simply found it unpleasant.

Some of, you know, each of the people he worked with has a completely different deaf identity. And, you know some of them worked with him for a few years, and basically still signed. And I really think that's probably true of Dudley, who ended up marrying someone who had gone to a sign language school, and really being in the sign language community, and I don't think ever dropped her initial use of sign. She was this amazing signer from her school days as well. So, and I think he was working more with her on articulation.

So, I'm not sure if I'm making this too complicated to be an answer. Ableism is definitely there. And I see it most clear-- with the term modulation, it's most clearly-- modulation can be anything. If it can be singing, it could be any kind of inflection. Why have are really restrictive, narrow zone of what counts as acceptable speech?

Stefan Helmreich: I can say a tiny bit to the vibration question, by just swerving into the notion of resonance once again. I'm struck by how many of these conversations about resonance, pitch resonances, that which breaches the boundary between the subjective and the objective, which I guess I learned new words for, tone and [clang], right? In the German.

[1:35:28] But, I think what's also been useful about these talks is the way that they've given us a history for that breaching that resonance does across the subjective and the objective, and directed our attention to the place of the machine as that which produces a kind of relay, and a blurring of boundaries. And the machines are multiple. It's not just the machine, right? It's telephones, it's telegraphs, it's resonating tuning forks, it's information theoretic cyborgs. So that the particular, kind of technological history that gets written into what we imagine as embodiment, you know tracks these technological changes.

And as Mara also pointed out, the question of language is tremendously important. Following words like resonance, words like transduction, words like modulation, all those things sort of, I think, make it clear that resonance and vibration are historical terms. They're not just things in the world that come to us completely unmediated. Are there questions from the other rooms yet? Evan.

Evan Ziporyn: I think I have a question. And I think it's a question for Josh, but I have to sort of work it through to get to that point. So, it seems that what we learn from all of this, or what we're reminded of from all this, is something that we all know, that's pretty obvious. Which is, just as Cage went into the anechoic chamber and came out realizing there was no such thing as silence, and then devoted his life, in a certain sense, to exploring what that meant.

What this all comes down to is that there's no hearing in the present tense. And not just in the obvious sense that the sound travels over space, but in the sense that in order to make sense of hearing, whether in terms of pitch, you need frequency, frequency takes place over time. So you need time to hear the frequency in order to figure out what frequency it is. And as Josh told us, in order to kind of parse it in a space, you need time to figure out what kind of space you're in, and kind of decode the signal, remove the signal to the extent that you can.

So, without following that through to the point of saying, OK well, so there's no hearing in the present tense, maybe that means there's no present tense. There's no now.

Is there work being done to understand how that mechanism works, like how you take this succession of present moments that don't make sense, and kind of make a time-based decision about what you're hearing. I mean, in other words, with the impulse response tests, do you need a certain amount of time to know whether you're hearing something over time. Does that question make sense?

Josh McDermott: Yeah, absolutely, that's definitely an issue that we think a lot about. I mean, one of the questions is, you might ask, well given that, as you said, you have to somehow integrate information over time in order to really make sense of anything. Well, kind of, what's the window through which you're really looking at the sonic world? And that's something that we, and others, are actively working on. And there's probably not just one answer, it's going to depend on what exactly it is that you're trying to hear.

In specific case of reverberation, there is interesting evidence that having prior experience with a particular room actually makes you better at understanding speech, for instance, in that room. And so in the interest of time I actually didn't show you this, but in those experiments that I described where we are measuring people's ability to extract the source properties from a source that's been bathed in reverberation, we actually, on every trial, we play them about a second of some other pattern convolved with that same reverberate impulse response. Because we think that that will optimize the conditions, and help them do as best they can.

And in those experiments, we haven't actually manipulated that, but other people have done things kind of like that with speech, as I said. And you do benefit from, say, having a couple seconds of prior exposure in that particular environment. So, yeah, those are interesting effects that suggest that there is some kind of short-term learning that's going on that cause you to build a model the world over some, an amount of time that's on the order of seconds.

Brian Kane: One place that it -- so, I think these temporal questions are very, very hard, and I hope that cognitive science can give us an answer about it. But, I mean, we do have models and philosophy, as dubious as they may be to some people here that have tried to answer this. We do have Husserl and time-consciousness, which seems to be a wonderful model to deal with the question of this kind of specious present.

[1:40:25] And we also have, you know, these wonderful lines in Heidegger that have to do with hearing and being and time. You know, we don't hear the sound of the motorcycle; we hear the motorcycle. This seems me, a very important insight. He's also says that we don't hear with the ear, which seems to indicate that we hear with something else, something that goes beyond the ear.

That we have things like being in the world that allow us to understand, and train ourselves on particular spaces. So it seems to me that a lot of these various insights, that you know, these are 100-year-old insights. But there are wonderful. And it'd be great to see how they get experimentally cashed out.

Audience: I was wondering whether you could connect the dots a little bit for me between sound, emotion and memory. Why when we hear a song, or a melody, or that the cord from Tristan and Isolde. We know exactly what brings us back. Why we recognize a voice we haven't heard in 20 years. And there was mention yesterday about colors and visual sense not coming from the eyes, but the cortex. Is it the same thing with hearing? That hearing is merely a technical receptor, and the interpretation is done much in the same way as vision?

Josh McDermott: Maybe I'll start with that one. So with regard to the memory question, it's, yeah that's a great question. I mean and there's-- I think a lot of people have this sense in particular that music has got sort of an intimate relationship with memory in that you hear some song from when you were young, and it kind of instantly takes you back.

There have been a few studies done about this that I know of, which actually seem to suggest that memory for music is no more accurate than memory for other things. In particular, the study I know of compared it to faces. And I think what may be going on is there's another strand of memory research, which indicates that our memories for things that are emotional are fundamentally different for memories of things that are not emotional. They sort of, they get charged with extra juice, in some sense.

And so people have the impression that those members are more vivid, even though when you objectively test their ability to recall information, you can show they're no more accurate.

So what is really different about emotional memories is people have this illusion of vividness in some sense. And so I think that because music is so often an emotional thing for people, it makes us feel things, I think that's kind of what's going on with memory for music. Is that it's an emotional memory, and so we have the sense that there's something different about it than say, your memory of your high school classroom or something. And, I think yes, it's almost certainly the case that a lot of that stuff happens in the cortex, although the details of that are still very much being worked out.

Josh McDermott: So this is also a question mostly for Josh. So, it seems like you're-- so you're able to put these pretty strong constraints on your impulse response. Which can help you kind of deconvolve and pull out what the impulse response is based on the result signal, right?

But even with those constraints, like it's still a difficult problem to solve to figure out what that actually was. And I was wondering how much of that is like just our brain knowing what reverb kind of sounds like, and being able to pull it out. And how much of it is also coming in from our other senses, and like the context that we have already from knowing how big the room is, or like you said, if you've been in there for awhile, and you kind of like learned a little bit. But I guess mostly I'm curious about the connection with the other senses. And like the fact you can see how big the room is, and how much that helps.

Yeah, that's a great question. We are hoping to do that experiment this fall.

Yeah, so I'm not sure. I think it's plausible there could be an important influence of that, but yeah, we don't know.

Audience: Hi. This probably would be for Josh or Brian. I'm someone who's been interested in pop music, people's response to pop music. And certainly over the past 30 years, the presence of synthetic instruments has dominated how pop music is crafted and produced and heard.

[1:44:51] I wonder if, we have a different response to a drum machine then the actual thump of a drum in a room, and we can distinguish between them, or fake trumpets, or synthetic strings. Has there been any studies done into our response, or our ability to distinguish between synthetic instruments, that we have certain expectations, and real instruments that we have expectations of? In terms of, I mean, it could be a psychoacoustic or emotional response.

Brian Kane: Well I don't know the scientific literature, so I couldn't tell you about particular studies. I don't know if you know of particular studies that have done with these kinds of questions.

Josh McDermott: Well, so, I guess more broadly, I mean one of the main techniques that we use in my lab is actually sound synthesis. And we sort of, we use it as a method to actually test what people internalize about sound. And so the idea is that if you can synthesize something that it sounds like the real thing that means that your synthesis method has captured all the variables that matter to your brain.

And so, I guess I would just say, more generally, that those differences between like synthetic instruments and real instruments, I think, are really deep and interesting. Because they tell us that there's something that's being done by the real instrument that your brain is very sensitive to. And so there have been a bunch of interesting studies in, often in computer music departments that have tried to look at things like that. And showing that, for instance, in piano tones, there's usually a random component, that if you don't mimic in a synthesis method, it won't sound like a real piano, it will sound like a synth.

[106:34] And so, so yeah, there's bits and pieces of research done on that. I would be able to make some general statement other than to say that, I mean I think it's very useful and informative with respect to how we hear.

Audience: At the Venice Biennale last year, there were a number of sound pieces that were created by visual artists. There was the Mexico pavilion had a piece that seemed like that harp that we saw in the morning that was defining the space of this old church that was resonating with the movements of people walking in and out.

And there was another one in the Norwegian, in the Nordic pavilion where you would breathe carbon dioxide into an instrument that would take it to a set of leaves. And then it was mic'd so that as the carbon dioxide of your breath went in, and the oxygen came out of the stome of the leaves, it created a sound so the trees sang back to you, as the response to your breath. And it seems to me that-- so my real question is, how do you guys see the difference between what's happening in experimental music and what is being claimed by visual artists as visual art using sound?

Alvin Lucier: Well that was a good term to use, people saying this is music. This doesn't sound like music. This isn't harmony, melody, and rhythm. You'd say well if you don't like the term music, use the term sound art. And it's also meant that-- for example, we had a master's student when I was teaching who didn't know F-sharp from A-natural. He was a sound artist from England.

And we accepted him as a master's student in the same way that we'd except a Japanese musician who doesn't know the five line staff, or an Indian musician who didn't know five line staff. So the idea that sound art is a wonderful term to encompass a lot of people who are working with sound, but not necessarily music as we know it.

Brian Kane: I think that speaks a lot to, I think-- I don't know if there's any actual material difference between music and sound art, nowadays. I think there are institutional differences. And the fact that Alvin gets to work, or spend so much time, you know, building up the culture and the curriculum at Wesleyan is an institutional thing. It allows for people who are interested in sound art to work there. And where I teach, that's not so much the case.

We are very beholden to certain institutional questions about what constitutes music, and what constitutes musical knowledge. There are also, in addition to institutional differences, there are questions of histories to be written. Histories tell us a lot about the difference between sound art and music, or the supposed difference between sound art and music. One of the things that I've been deeply interested in recently is, why is it that the art historians got sounds art?

[1:50:13] Why is it that Duchamp -- that Cage is the student of Duchamp, and sound art comes out of Cage? When Cage also studied with Henry Cowell, and Schoenberg. So are there ways to give us a kind of more complicated narrative about the relationship between music and sound art? As a, kind of as an academic and intellectual, and somebody who's interested in history, I feel like that's the story that I could tell as a way to complicate this.

This is partially why I'm interested in people like David Tudor. I think Tudor's very, very important as a kind of hinge between the worlds of music and sound art. But in terms of saying looking at, are there, you know, sets of sufficient and necessary conditions that would differentiate the two, then I'm at a total loss to articulate those.

Alex Rehding: Yeah, if I could just quickly pick up on the aspect of visual culture and visual artists embarking on sonic projects, I think that's the way in which he framed it. I think the opposite is also true, that composers are quite interested in adding a visual dimension to their sound works. And that happens in various ways..

So, just to say again, what I was saying, the opposite is also true. It's not only visual artists who get increasingly interested in the sonic domain, but also, that's a trend that, I think, is discernible among certain composers. That they're increasingly interested in adding a visual domain, a visual dimension to their artwork.

But, and, jumping off from that, I want to get back to David Hubel who Caroline mentioned earlier. Because one if the really interesting thing is that, the very first experiments on the mural aspects of the visual system were, as far as I know, communicated via sounds, that mural firings were made audible. And so, I think, there is a lot of work to be done in highlighting the role of audition in scientific experiments, and I think this would be great example.

Audience: I have two short questions. One is about-- I would like to know if there are any studies conducted nowadays in the use of silent pauses, and how it's changes from like people doing normal speech, and people with like impart hearing? And the equivalence of use of silent pauses in sign language.

And my second question is if any of you have interest in attempts to translate or find equivalence for music in sign language? For example, experiments that were done by Wendy Ebsworth in London, who tries to orchestrate some kind of choreography, transforming sign language into a very complex, hybrid mode of choreography.

Josh McDermott: I can probably take the hearing impairment question. If somebody else wants to take the sign language, music question. But, with respect to the pauses in speech, I think we do have pretty good evidence that if you get rid of those, it's harder for someone with hearing impairment to deal with the signal that they're getting. And that's partly why-- and I didn't mention this in my talk, but-- one of the reasons why the cocktail party problem is so interesting and important, is that when people have damaged hearing, that's a situation where they have the most trouble. And it's the situation that we don't currently know how to fix.

So current hearing aids, you can put them in and someone has hearing impairment, if they're in a quiet room and there's just kind of one thing that's making sound, they can hear here just fine. But you then take them to a restaurant, and even though they've got the hearing aid in, and even though the sound is audible, they can't make sense of what's there. And there's sort of too many things and they overlap too much, in part because of the way that the ear changes when people lose their hearing.

[1:54:19] And one of the consequences is that the temporal resolution seems to be worse. And so you get rid of those gaps, and things get more smeared together for someone that has impaired hearing. And I don't know how that relates to sign language, I don't know much about that. Do you maybe?

Mara Mills: I mean, I've actually -- I mean I don't know about how it relates to sign language, either. But I have -- I'm doing another project on the history of text-to-speech, sort of Kurzweil and before. And I've interviewed a bunch of older users, blind users of text-to-speech machines, and all of them have pointed out to me that until artificial breaths were inserted into the speech synthesis, you couldn't read that quickly. You needed a pause, you needed-- the stream of synthetic speech would slow the ear down, would slow oral reading down. And that having synthetic breaths, basically, as part of text-to-speech-- and maybe some of you here work on that-- has made their reading faster, by ear.

Audience: I was wondering what the difference is between reverberation and resonance. I thought there was a difference, and now I'm not so sure.

Josh McDermott: Sure, so, I mean, reverberation generally just refers to the collective effect of all of the reflections that occur in an environment. And in resonances, is a more general phenomenon whereby anything, really, that you put a signal into, it could be a space, it could be a machine, will amplify certain frequencies and attenuate others. And in the case of rooms, those resonances have to do with the dimensions of the room, and whether certain frequencies constructively or destructively interfere.

And so most spaces-- it's also true of the vocal tracts. Mara talked a little bit about this, how our vocal tracts have resonances in particular places, and when we talk, we move those resonances around because we change the shape of our mouth. And when we change the shape of our mouth, the resonant frequencies change. And that's what distinguishes different vowels, it's just those resonances which are known as formants in the case of speech get moved around.

And, so yeah, reverberation usually entails certain resonances, so any environment is going to have particular resonant frequencies, and as a consequence of the reflections that those resonances will then get imparted onto speech. And so, I believe in the Lucier piece that was mentioned earlier today, what's being happened is you're essentially repeatedly exposing the sound to that same reverberant field. And so over time, those resonances really become come to dominate and that's essentially what you hear.

Audience: Yeah, I don't know if it will end up a question or not. But, I was curious about the link between; I guess it happens in signal processing where you're converting, say like a longitudinal wave of sound into a pressure waveform. And that's happening like between a sound to energy. Is that correct? I guess I'm trying, there's the notion of the -- I guess there's, I'm trying to deconvolve the vibration as something that seems to make these phenomena emerge versus like the physical substances that vibrations are happening in.

I guess the fact that it's not happening without a mediating thing. Whereas with the sirens in maybe the 1800s, we weren't quite sure of like electricity as much. We thought it was like a subtle fluid, or like some kind of magnetic attractive thing. And so I guess I'm wondering if there's like a -- I'm trying to deconvolve. I [sound] conflated, sorry.

But, there's like -- that we don't understand electricity well enough, or we use language for like, I guess the parallel between air and electrical charges, sound and energy, electrical energy, and like wind and like current flowing charges, like a wave through current charges. I guess, I, is there? It's all, there's an analogy between those spaces. Like the physical -- air is a physical substance. Like electrical charges are the physical substances. Like sound is the wave propagation through a volume of air.

[1:59:03] And electrical energy is waves traveling through a column of charges. And then like wind is the following motion of air. And current is the following motion of the charge. Like that.

Stefan Helmreich: Yeah, one way to deconvolve those analogies -- I can try to help you out -- is to delaminate them, right, is to historicize them. Is to ask the question of when those analogies come into being. When is it that electricity is analogized to water? When is it that heat is analogized to sound? Et cetera, right?

So these are not kind of pure analogies floating in a platonic space. But they are things that people have worked on through machines, through languages in order to make those cases, right? And, as with any analogy, they always fall apart, kind of at the edges, sometimes at the center as well. And I think a lot of this panel has been about the way that those analogies enter into constructive and destructive interference, if I could just totally essentialize against the grain of what I just said.