Making Audio Visible: 
The Lessons of Visual Language for the Textualization of Sound

CHARLES BERNSTEIN

Francis Barraud, “Dog Looking at and Listening to a Phonograph” (1898).
THE TOPIC FOR THIS SECTION is “Textuality and Visual Cultures.” Marjorie Perloff and Johanna Drucker have each addressed the problem in a direct way. My own topic—the digitalization of the recording of poetry readings—would seem, on the face of it, remote, a stretch at best, perhaps even a stubborn refusal to stick to the topic. But the more I delve into my subject, the more I see how difficult it would be to extract the topic of digitalized sound files of poems from either visual culture or textuality.

The remarks here extend my work editing Close Listening: Poetry and the Performed Word. This volume remains the only overview by scholars and poets that addresses the significance of the contemporary poetry reading for literary studies and poetics. The book came out simultaneously with a related collection of essays, Sound States, edited by Adelaide Morris, that addressed the significance of acoustic technologies. I pursued my investigation in this area in “The Art of Immemorability,” an essay I wrote for Steve Clay and Jerome Rothenberg’s collection The Book of the Book, which explores the relation of the photo/electronic reproduction on modernist poetry, a theme also explored by Freiderich Kittler in Gramophone, Film, Typewriter. However, let me start with a highly compressed history of language-recording technologies.

The first great technology for recording the sound of spoken language was the phonetic scripts of about 2500 years ago. The phonetic (literally “spoken voice”) alphabet created a simple visual notation system that was used, in many instances, as a transcrip tive device: that is, as a script for performance, the letters serving as cues to prompt oral recitation. The significance of a visual technology for the notation of an aural phenomenon is profound; its implications for our way of understanding language, and indeed for our perception of both reality and its others, is incalculable, and, as a result, this relationship has been the subject of many learned treatises, very few of which were composed for the tongue. Curiously, the significance of visual language on the semantics of literary works has been less widely acknowledged. From the point of view of aesthetics and literary criticism, the work, over the past twenty years, of Drucker and Perloff, are signal exceptions. In contrast, from the point of view of bibliographic scholarship, the significance of visual coding has remained central, even as the field itself has become more marginal in an age of cultural studies and theorization.

But, back to my abbreviated history—with a two millennia jump.
In 1877, Thomas Edison made the first recording of the human voice—a short recitation of the first verse of Sara Josepha Hale’s 1830 poem “Mary Had a Little Lamb.” Edison’s recording was made just two years before the poet’s death at the age of 91. The first sound recording was of a poem, though not in the author’s own voice. Edison stumbled upon his “speaking machine” while working on a device to speed up telegraphic transmission—the telegraph itself being a device for the audio transmission of visual (alphabetic) code. Edison’s sound-recording device was variously called a talking machine, a speaking machine, and a phonograph. “Phonograph,” the word Edison chose for his company, suggests something of the visual/verbal paradox that is my topic today, meaning, as it does, something like sound writing or sound drawing. We don’t, literally, read a sound recording but play it. And while the sound recording is able to capture far more acoustic data than phonetic script, “sound writing” shares with alphabetic writing the ability to retrieve recorded verbal data. This combination of inscription and retrievability marks recorded sound as a new form of textuality, new, that is, for the past century. Indeed, as I note in “The Art of Immemorability,” sound recording is inextricably linked to modernist poetry and its specific textual forms.

The most famous image of the phonograph, “His Master’s Voice,” pictures a dog—Nipper by name—listening to a gramophone. The first version of the image was made in 1898 by the British painter Mark Barraud and is called “Dog Looking at and Listening to a Phonograph.” Later the specific machine depicted was changed and the image became famous as the logo for The Gramophone Company in England and for RCA Victor in the United States. Barraud’s work is iconic of the uncanniness of the human voice emanating from a machine unattached to a human body. Dogs, beloved by their owners for their ability to distinguish specific human voices from other sounds in the environment, are the adequate symbol of transhuman voice recognition. As Gertrude Stein put it, “I am I because my little dog knows me.” If the dog hears it, it must be Memorex. For Stein, that is, the dog wags its tail when it catches the scent of human nature but not human mind. And since dogs don’t talk, but only listen, the dog represents the ideal audient for a talking machine that, like a text, speaks but cannot hear.

Barraud’s original title, emphasizing the combination of looking and listening, is telling because one of the most striking features of the image is Nipper gazing affectionately into the Victrola’s large horn, sometimes imagined to be an ear but which more pertinently can be imagined as a mechanical throat and mouth. Nipper is looking into the horn as if this was the most natural and comforting thing in the world; as if, that is, there were something to see inside the horn, in its dark interior. In other words, the dog (I keep typing this, dyslexically, God)
is sniffing out the ghost in the machine. Indeed, with his nose almost entering into the cavity, it looks like Nipper’s head would fit snugly into the device, which might be something of a textual womb; or, alternately, as if Nipper might be swallowed up into the voice machine, as Jonah into the whale. Nipper, a dog who is the phantasy of visual culture, is looking straight into the production of a human presence that can call his name but doesn’t know he’s there.

Much has been made of the ghostly presence of the disembodied phonographic voice, the novel possibility, with this new invention, of the voices of the dead speaking to us. But I am interested in a quite different cause for the uncanniness of grammaphony. Talking machines merge two usually separate cognitive phenomena: speech-mode perception and non-speech mode perception. As Reuven Tsur explains it, we cognitively process speech sounds differently than other sounds. Consider, though, that the talking machine produces mechanical sounds that we process, against our typical automatic response, as if they were speech. With early recordings, just as with radio with bad reception, we actively search for the voice in the static, as if it were a figure to be differentiated from the ground: the mechanical semblance of voice has become the signal in a medium whose material base is sonic not vocal. In such a phonic economy, noise is sound that can’t be recuperated as voice. Tsur, following Roman Jakobson, defines the poetic mode of cognitive perception as the perception of speech as if it were sound. The grammaphone reverses the Jakobsonian definition of poetry: it incites the perception of mechanical sound as if it were speech. The grammaphone is a reverse order poetry machine.

This is what fascinated Nipper, our ideal audient. The master at whose voice he stares is not his human owner but the machine that produces a voice without body. We are all Nippers now.

The conjuring of voice unattached to body is a kind of disappearing act to which we have grown so accustomed that we need a little dog to remind us of what’s not seen. To accept a link between sound reproduction and visual culture, you would have to allow the conceit that invisibility is a part of visual culture. Voice-reproduction technologies not only undermine intuitive distinctions between sight and sound, but also between space and time.

As the story is told, the earliest playable sound recording, from 1878, is Frank Lambert’s “Talking Clock.” Lambert had the idea of marking time in a temporal medium: the new “speaking machine” would provide an alternative in real time to the visualization of time in previous clocks. The recording has four basic sections: 21 seconds of indistinct speech, 11 seconds of Lambert clearly calling out hours (five o’clock, six o’clock, seven o’clock, eight o’clock, nine o’clock, ten o’clock, eleven o’clock, twelve o’clock). There follow 36 seconds of silence and then the most extraordinary thing of all: 29 seconds of
voiced but illegible sound. Evidently, Lambert cranked the cylinder in reverse, producing, for the first time, the human voice played backward. The first distributed sound recording was not just of a human voice counting numbers but also the first sound poem. Indeed, it is impossible to imagine Lambert’s “Talking Clock” as having any practical import at all: it is nothing less than the first work of grammophonic sound poetry.

Mechanically reversing the flow of a recorded voice breaks down the perceptual distinction between space and time: it can be both disorienting and mesmerizing. The fact that, with the cranked gramophone, the recorded voice can be played back in two directions is something that puts the earliest experiments with the plasticity of the reproduced voice in conceptual tape loop with the rhythmic use of vocables in the sound performances of analphabetic cultures. Voice becomes text, just as with phonetic scripts text becomes voice.

But not so fast. Is sound—the performative dimension of the poem—really textual? We typically associate textuality with the woven texture of written language, and, indeed with visual inscription. Within the context of literary criticism or textual scholarship, the performance of the poem has not been generally recognized as part of the work, a designation that is reserved for the text understood as the scripted incarnation of the poem. Nonetheless, in alphabetic and postalphabetic visual cultures, the a/oral dimension of the poem can’t be split from the text “itself,” even if it threatens to undermine the coherence of the poem by adding possibly new and incommensurable textual layers, to echo a point of Jerome McGann’s about the multiplex nature of the textual condition. The audio text may be one more generally discounted destabilizing textual element, an element that undermines our ability to fix and present any single definitive, or even stable, text of the poem. Grammaphony is not an alternative to textuality but rather throws us deeper into it folds.

To say that we don’t literally read a sound recording, but play it, is itself to trip on the distinction between the graphic letter and the interpretive moment. Perhaps we could call this grammaphonic enunciation the sound of one hand clapping, whether that one hand is the stylus or the digital decoder. But the nonliteral sense of reading crosses the sound barrier: for to read a poem out loud is to give a reading of it just as to listen to a poem recited is to have to read it with your ears. The paradox, to say it again, is in the words we have given to these machines: for example, gramophone (or phonogram), which might be translated as texted voice, suggesting, as it does, that the lettered word—gramma—is made invisible so that it can be heard. The process of inscription and retrieval is shared by both the grammaphonic and, if we want to invent a new word for writing, the grammagraphic.
One hundred years ago, a wax cylinder inscribed with the human voice stored two minutes of sound. At present, a CD inscribed with compressed sound files can store ten hours of sound. The revolution in our ability to record and store sound has had a huge impact on many areas of the culture; not the least of these, though perhaps the least remarked on, is poetry. In the business press, it is common to hear about the disruptive effect of MP3 music file exchange on the pop-music recording industry. Dire pronouncements are made about theft of intellectual property by the very same companies that sell the hardware and software to record and play these files. Certainly, we are seeing a great change in the mobility of recording, playing, and, crucially, distributing, sound recordings. The change is quite significant, but I think the greatest significance may well be not for pop music and its financial base but for poetry and its grammaphonic status. And that’s a prediction you won’t be hearing on CNN Financial.

In the era of print, from which we are slowly emerging, the alphabetic version of a poem was supreme, unchallenged. While poems were sometimes produced in nonalphabetic contexts, within alphabetic cultures, literary production has been synonymous with writing understood as a visual system of notation. Poetry readings, while significant, have been regarded largely as extensions or supplements of the visual text of the poem. Indeed, there has been very little critical work on the poet’s performance of a poem; at least up until very recently, literary criticism has pretty much been confined to the printed text.

The reason for this is practical as much as conceptual. While archives of poetry recordings exist, they are largely inaccessible. Very few editions of poet’s sound recordings have been published. As a result, basic principles of textual scholarship have not yet been applied to the sound archive. But the times they are a-changin’.

With the advent of file compression and broad-band web connections, it is now possible for individuals to exchange sound files of recorded poetry. The digitalization of the archive of recorded poetry is just now getting started, and it is this digitalization that is my primary focus.

At the University of Pennsylvania, we have started a project we call PENNsound, which extends the work I have been doing with Loss Pequeo Glazier and Martin Spinelli at the Electronic Poetry Center in Buffalo (epc.buffalo.edu) and builds on the foundational work of Kenneth Goldsmith at Ubuweb (www.ubu.com). PENNsound is a collaboration between the Center for Programs in Contemporary Writing (Al Filries) and the Annenberg Rare Book and Manuscript Library (Michael Ryan). For this project, we intend to develop protocols for the indexing and tagging of sound files of poetry readings. At present, there are no such standards, and relative chaos exists at every basic
level, from file-naming conventions to formatting to file-format choices to relevant bibliographic coding to copyright questions to storage and preservation issues.

I believe that access to compressed sound files of individual poems, freely available via the internet, offers an intriguing and powerful alternative to the book format in collecting a poet’s work and to anthology and magazine formats in organizing constellations of poems. Imagine for a moment that you had on the hard drive of your computer a score of MP3 files of poems by 50 of your favorite twentieth-century poets. I would bet that no matter how involved or committed any of you may be to twentieth-century poetry, none of you have such a collection readily available. But in poetry’s coming digital present you will, or anyway, you easily might. What would be the implications of such a collection?

In my focus on the recording of the unaccompanied voice and more specifically on poets reading their own work, I am not unaware of the importance of sound recording technology for music; and I have been told that there has been quite a lot of material produced and distributed, over the past hundred years, with either just musical instruments or with musical instruments accompanied by singing. Quite a bit of fuss has been made about this, and it is evidently today a multibillion dollar worldwide industry. But I wouldn’t want this aspect of sound recording to overshadow what is for me, thinking in terms of the art of poetry, a crucially significant feature of sound recording technology: the reproduction of the unaccompanied human voice. In other words, I want to remain archly formalist and even purist on this issue, at the risk of seeming to direct way too much attention to what for most people is at best an epiphenomena, a footnote to the larger story.

For teachers, one obvious implication of recorded poetry becoming more available is that listening to the poem read by the poet might become a commonplace feature in any course, since such recordings could be assigned in much the same manner as the visual or alphabetic text of the poem. The sound file would become, ipso facto, a text for study, much like the visual document. The acoustic experience of listening to the poem would begin to compete with the visual experience of reading the poem. Many of us, of course, ask students to read poems aloud, or we read the poems aloud in class. But the textualization of the sound as an object of study radically changes this dynamic.

How about for readers of poetry? What effect would the collection of poetry MP3s have? For one thing, the experience of listening to poetry would be far more mobile and portable than it has been, rivaling, though not exceeding, the portability of the book. While now, those interested in poetry performance mostly access it live and in person,
where the visual element of the poet reading, the theatricality of the presentation on stage, and the social element of exchange in many ways equal, or surpass, the importance of the merely acoustic dimension. Despite the long-time availability of poetry recordings, very few people have anything more than a passing relation to this archive in comparison to their experience of reading poetry in visualized formats (writing) or going to live readings. Indeed, it is hard for me, and probably hard for any of us, to imagine that possibility—that someone would have their primary relation to poetry via audio recordings.

In poetry’s coming digital present, the poem, understood as a recorded performance of 30 seconds or 3 to 5 minutes or even 10 or 20 or 30 minutes or more, may take on an integrity that it presently does not have in the era of book collections of poems. Poems, that is, might find themselves liberated from books, free-standing, available to combine with other poems in one’s collection to play in the car or at the health club or on the plane or walking around the city or sitting at the beach or dancing on the head of needle, a virtual phonograph needle that is. Imagine digital files of discrete poems as individual leaves in an anthology of one’s own creation. Poems, set adrift from their visual grounding in alphabetic texts, might begin to resemble the songs from which, for so long, they have been divided.

A widely available digital audio poetry archive will have a pronounced influence on the production of new poetry. In the coming digital present, it becomes possible to imagine poets preparing and releasing poems that exist only as sound files, with no written text, or for which a written text is secondary. Orally composed, or nonscripted, poems have already been realized by bpNichol and Tracie Morris, among others; but greater access to digital sound editing opens other doors. In the coming digital present, poets will compose the text of their poems by dictating and editing sound files, much the way we now compose and edit alphabetic writing. In this sense, the production of poetry will come closer to film or radio—works that are produced in the cutting room. The result will not necessarily be sound poetry or poetry with accelerated variations of performance affect or number of performers, but possibly an audio–textual poetry akin to the modernist and contemporary work, extending the formal possibilities of language worked as a virtually physical material for making the poem.

Consider also the effect of easily available sound files on critical writing published in digital media. It might become as common to include a sound clip in an essay as it is now to include an image; and the sound clips could themselves be subjected to a microanalysis, via a sound editing program, that we presently do by strictly visual parsing. And I can only here mention the possible transformation of our sense of the poem when we consider the acoustic record along with the visual marking; or when we look at the visual print or sonograph of
a performed poem, as becomes possible with any sound editing program. The sound shape would not merely be a metaphor for acoustic or rhythmic analysis, but also an image created through processing the sound in ways that give it a visual trace.

Moreover, the kind of “deformative” criticism advocated by McGann and Samuels, or my own set of “wreading” experiments at the Electronic Poetry Center, which mostly involve visual mark-up and coloration and erasure or transposition / translation of visual linguistic elements of the text, could be done via acoustic processing: dropping sound levels, increasing speeds, changing timbre, repeating phrases.

For the fact is that the sound file exists not as a pure acoustic or sound event—an oral or performative event outside textuality—but as a textual condition, mediated by its visual marking, its bibliographic codes, and the tagging we give it to mark what we consider of semantic significance.

Another central issue I want address is the effect that sound files might have on scholarly editions. Or should I say the havoc that they might wreak? What, for example, is the relation of the recordings of a poem to the printed text of the poem? How might this relation be articulated in different types of editions, print and digital? Just as there are often multiple alphabetic versions of a poem, there are also multiple performed versions. Indeed it is likely that for a poet for whom we have an extensive archive of recordings, there are likely to be more variations of the poem than those produced by print forms. We might have the same poem read at very different points in the poet’s life, and there might be lexical variations in each performance (or not).

The types of variations of a poem found in audio recordings are substantially different than those in visual culture documents. To what extent to “authorize” such variations as credible variant versions of the poem is itself an issue. But what about different approaches to the performance: timbre, speed, rhythm, temporally marked lineation, emotional valences? Perhaps of greatest significance is the paratextual framing of a poem in a reading. For one thing, there are the introductory and prefatory remarks made before, during, and after the reading of the poem per se. At this year’s annual meeting of the Modern Language Association, Kenneth Sherwood presented an elaborate transcription of remarks of this kind made by Amiri Baraka in a reading in Buffalo in honor of Robert Creeley. For another, there is the context of the reading itself, which is akin to the context of publication in a magazine or by a specific publisher. Finally, there is the order of the poems presented in a reading, which will often offer a distinct and significant setting of the work.

A poem, to paraphrase Jerome McGann from a talk he will be giving this afternoon, is not a converging system of coherent signs but—now this is me again—a diverging environment of incommensurable sites
and sounds—"a dimension of mind. You’re moving into a land of both shadow and substance, of things and ideas. You’ve just crossed over into . . . the Twilight Zone."

Within the coming digital environment of a poet’s work, the sound file is at this time the orphan, something that if included would disrupt even the most expansive conception of versions, all based on different print versions.

I realize my proposal invites as many questions, and raises as many problems as it answers. This is as it should be at this stage. From the point of view of textual scholarship, it still remains to be seen what kind of tagging within the sound file would be valuable—not to say useful. One obvious issue would be search capabilities, how the conception of searching might be translated into recorded sound of poetry, beyond the obvious tagging of line and stanza breaks, titles and subtitles, key words, or even every word. Certainly, there is much to be learned from the far more advanced work in this area with various spoken word archives, from radio to oral histories. But the bibliographic issues for poems would be different. The term “bibliographic” is yet another reminder that the sound itself will be mediated by visual—graphic—tools. I might call for bibliophonic tagging, but I can’t say what that would mean.

I have emphasized the importance of the recordings of the poets’ own performance of their work. I realize that some will object that this valorizes the author function in a way apparently at odds with what many of us have argued for in terms of the active role of the reader. My choice is based on a clear aesthetic preference for poets reading their own work, but of course it is imaginable that performances by others might become significant. I address this issue in Close Listening, where I state my preference for audio recordings over video or film recordings of poets reading their work. On a practical level, the visual quality of taped readings is usually very limited in comparison to other moving-image productions; sound recordings do not have the same comparative problem. Then again, like most people, I prefer sound recordings of most music in comparison to straight "head shot" or concert video versions. But I think my claims of value for poetry audio are tied in to the particular formal significance that unaccompanied voice recording has for the grammaphonic imaginary.

"His Master’s Voice," the voice on a record, is invested with the aura of invisible presence. It speaks to us as through a wrinkle in time. The recorded voice only speaks; the possibility for dialog or response present at a reading—where the presence of an audience intimately affects what is being presented—is illusory, making our close listening across the electrostatic barrier all the more our own private affair. The recorded reading reenacts the conditions for dialog without its actual presence, unless we want to consider the presence
of the imagination. For the imaginative projection solicited by close
listening to the grammaphonic poem is the one writing has required
all along. Here we go again: it's déjà vu one more time.

*   *   *

Let me end with a brief manifesto for the PENNsound archiving of
recorded poetry:

1. **It must be free.** Ideally, all the sound material we put on the
web should be cleared for copyright to be distributed free. Users of
the site will be able to download the MP3s to their own computers
or players or play them in a streaming fashion. Teachers could make
course CDs or add the MP3s to their on-line syllabi. Other web sites
and libraries could recollect the material. Credits for digitalization and
copyright release would also be embedded into each file. One of the
advantages of working with poetry sound files is that we don't antici-
pate a problem with rights. At present, and in the conceivable future,
there is no profit to be gained by the sale of recorded poetry. There
is, however, considerable expense involved in preserving, cataloging,
and distributing such material.

2. **It must be MP3 or better.** RealAudio is a proprietary format
with sound quality that will not stand the test of time. We need to
use open formats that reproduce reasonably high quality sound, for
a listenership that is used to astoundingly good sound quality from
commercial sources.

3. **It must be singles.** At present, the vast majority of poetry record-
ing are for entire readings, typically thirty or more minutes, with no
tracking of individual cuts or poems. While these full readings have
great literary and archival value taken as a whole, few but the most
devoted listen to full recordings of readings, or if they do, fewer still
listen more than once. The more useful format is to break readings
up into individual poems and to make MP3s of each poem available.
MP3s of song-length poems could become a very appealing format for
poetry. The implications for audience, listenership, critical thinking,
poetics, and poetic production are great.

4. **It must be named.** Presently, downloaded poetry sound files
tend not to have informative names. Looking at a directory of such
files, it is impossible to determine what the file contains from the
visual information available. File sharing for music employs a simple
system of the name of the singer and the song, but the p2p system is
not compatible with FTP, especially in terms of blank spaces between
words, which need to have dashes or underlines. In addition, song
MP3 file names start with the first name of the singer, but for the poetry
files I think we need to move to the more conventional last name first
and give the date and place of the performance as well. For poetry file
names I propose: lastname-firstname—poem-title—place—date.mp3.
5. **It must embed bibliographic information in the file.** It's important that basic bibliographic information be embedded in the MP3 sound file itself, so that when someone downloads the file they get the right file name and also they get a full range of "ID3" type information—all in the same file. This is basically a consumer-oriented MP3 file exchange approach. The goal is to make these sound files easily available for users. Since downloaded files will be separated from their home library web site or catalog, information on that web site or in a catalog will not necessarily be retrievable (although the URL for the catalog can also be embedded).

6. **It must be indexed** so as to be retrievable both from a library catalog under the author’s name and via web search engines.¹

Note: PENNsound is at http://writing.upenn.edu/pennsound

**Works Cited**


