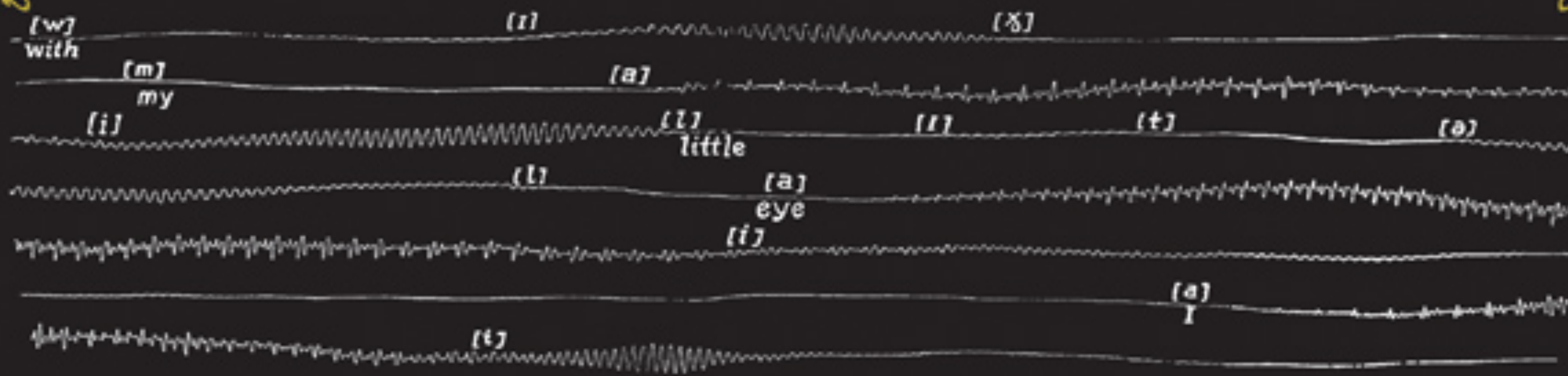
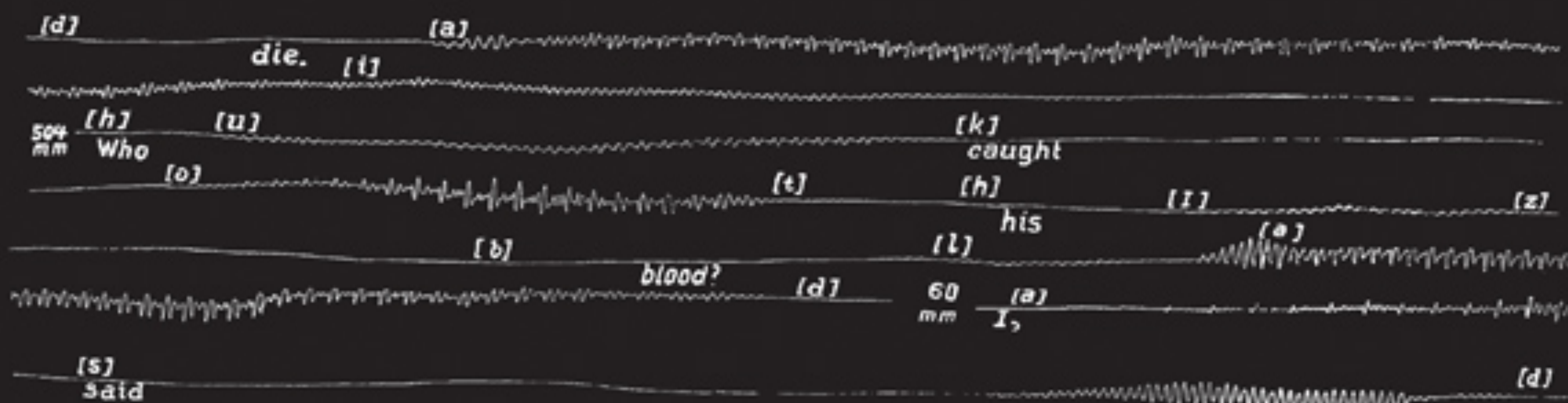
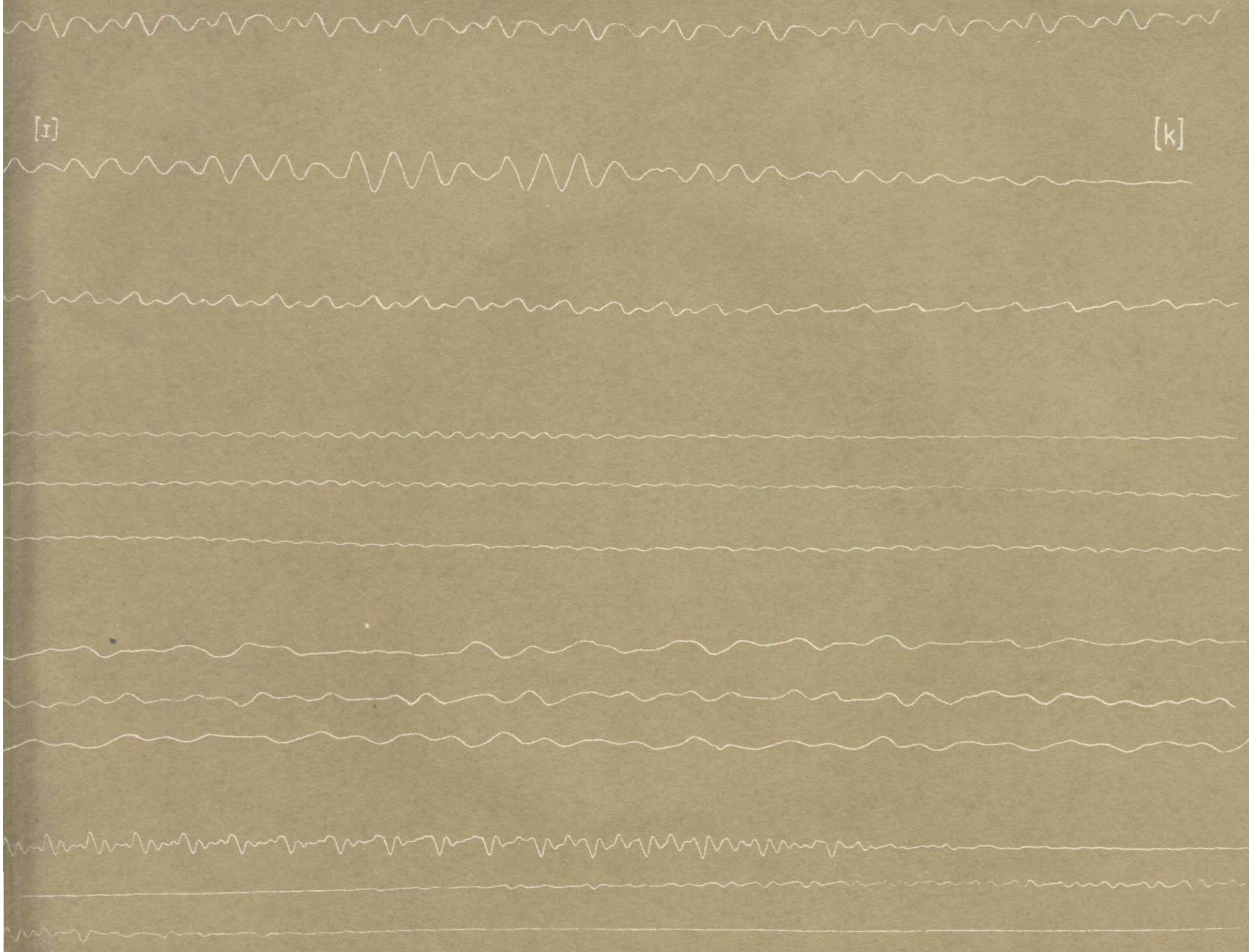


Pictures of Sound

One Thousand Years of Educed Audio: 980–1980



by Patrick Feaster


[ɪ]
[k]

fla gi

as tum ri

bet is be

Semi beas flagitans ua li ta

re

mis be lis

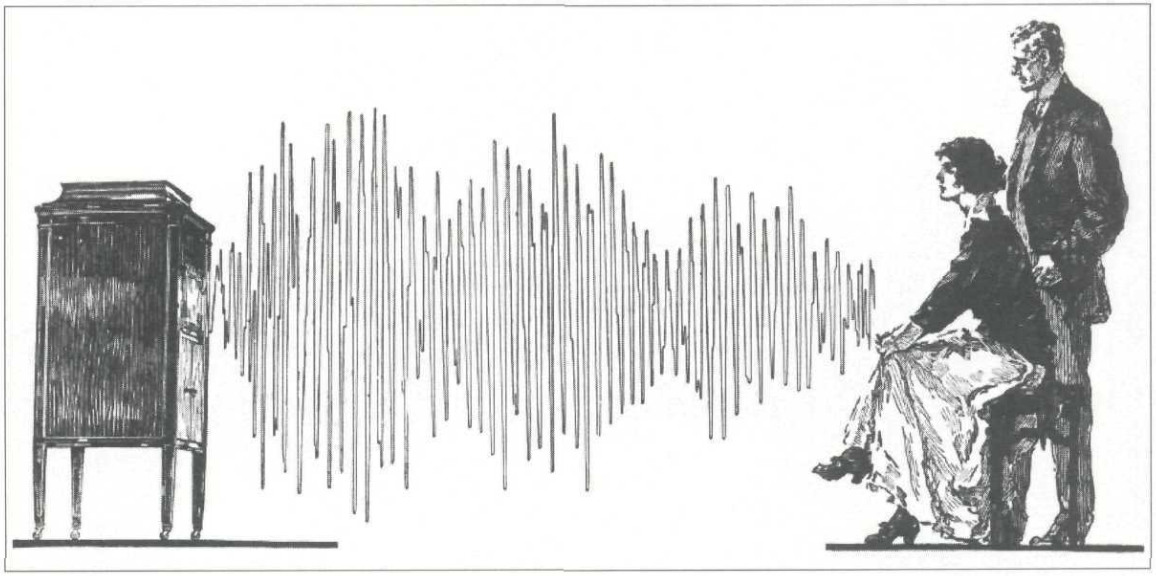
li ta ma

re

Pictures of Sound

One Thousand Years of Educed Audio: 980–1980

Dust-to-Digital
2012



"The Miracle of Sound," Aeolian-Vocalion advertisement, *New York Tribune*, 8 March 1915.

(((((((INTRODUCTION)))))))

What you have before you is a collection of historical audio, but it's more than that. It's a collection that seeks to challenge existing assumptions about what historical audio itself is. Some of its contents will be easily recognizable as "sound recordings," but the status of others is less obvious, and that's exactly the point. The boundary between what is and what isn't a "sound recording" is less rigid and clear-cut than you might imagine.

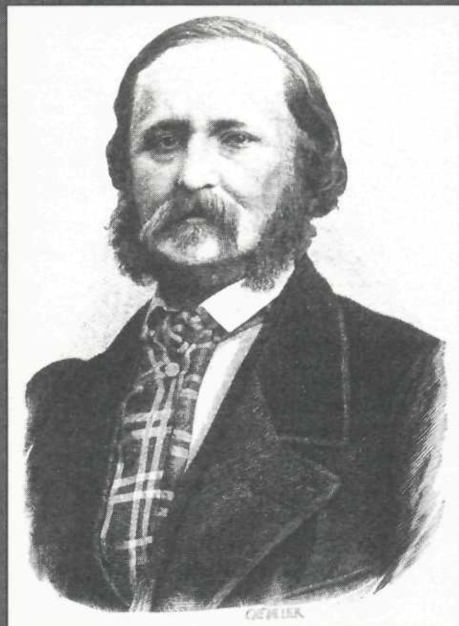
Many people, if asked for an off-the-cuff definition of *sound recording, recorded sound,* or some equivalent term, would probably come up with something like "a record made by documenting sound waves over time in order to reproduce them automatically as sound." At first glance, this seems to be a perfectly reasonable definition—after all, it accurately describes the scenario that originated with Thomas Edison's invention of his speaking phonograph in 1877, which has long been regarded as the foundation of modern sound-recording culture. However, it also leaves out much of what actually passes for "recorded sound" today.

Say I record myself saying "hello" into a computer microphone so that I can open the resulting sound file in a program that will display the data visually on the screen for me to study—researchers in various fields do this sort of thing all the time. And then say I delete the sound file after I've studied it in this way, since its purpose has been served. It was never played back, and nobody ever had any intention of playing it back, but I'm pretty sure most people would concede that it was still a "sound recording." It was certainly a record of sound waves, just as a seismogram is a record of an earthquake, and it *could* have been played back at any point, even though that wasn't its intended purpose. But if we accept that the hypothetical sound file I've been describing was a legitimate "sound recording," then we need to drop the second part of the definition cited above, since it apparently doesn't matter whether or not the sound waves were recorded with the intention of playback. That leaves us with "a record made by documenting sound waves over time."

And yet many of the things we classify in practice today as "sound recordings" aren't actually "made by documenting sound waves over time." Some of them are created through entirely synthetic means, while others combine "recorded" elements such as a human voice with synthetic elements such as electronic backing music. If we accept this kind of material as "recorded sound" too, then we need to jettison the first part of the definition cited above as well, since it apparently doesn't matter whether what we're dealing with is in fact a "record" of some prior reality.

In contemporary practice, then, a *sound recording* or *recorded sound* isn't necessarily intended for playback, and it isn't necessarily made by documenting preexisting sound waves. The only necessary condition is that it should *represent* sound in a certain very specific way—namely, by expressing it in terms of amplitude or frequency as a function of time. Amplitude and frequency are interchangeable here, since the one can be mathematically converted into the other through what's known as a Fourier transform—that's how modern sound-editing software can allow users to toggle back and forth at will between "waveform" displays of amplitude and "spectral" displays of frequency.

What does all this mean for *historical* audio? Until very recently, people working in the field of early recorded sound limited themselves to materials that fit the off-the-cuff definition we started out with above—that is, "records made by documenting sound waves over time in order to reproduce them automatically as sound." But that situation changed dramatically in 2008, when the First Sounds initiative I'd co-founded with David Giovannoni, Richard Martin, and Meagan Hennessey released audio recovered from a phonautogram of the song "Au Clair de la Lune" as recorded by Édouard-Léon Scott de Martinville in the year 1860, seventeen years before Edison had invented his phonograph. Although this phonautogram was unquestionably "a record made by documenting sound waves over time," it had been created for visual apprehension rather than for playback, which sparked some debate about what exactly it was. One school of thought was summed up nicely by a headline in the *Los Angeles Times*: "Sound Recording is Older than You Thought."[1] Another was spelled out in the comments section of a YouTube video:



Édouard-Léon Scott de Martinville.

> "This was NOT a recording. This is a graphic representation of sound waves. There was never any intent to record audible sound for the purpose of listening to it. [Scott] was studying what sound waves LOOKED like." [2]

I fell into the camp that believed Scott's phonautogram *was* a sound recording, as the arguments I made earlier would imply. However, I was also intrigued to learn that some people disagreed. One might think there would be some straightforward way of answering the question, "What's the world's oldest known sound recording?" But it turns out that there isn't, not so much because we don't know what's out there as because there's no general consensus as to *what a sound recording is*. If we define it one way, Scott's phonautograms count; but if we define it another way, they don't.

Let's pursue this line of thought a step further. If we define a *sound recording* or *recorded sound* in terms that are truly broad enough to encompass the full range of current practices—as an inscription that expresses sound in terms of amplitude or frequency as a function of time—and then look back through the historical record for inscriptions that fit this definition, what will we find?

The present collection aims to answer that question in as provocative a manner as possible—and, I might add, by moving in a direction that's entirely different from the usual line of speculation into ways in which primeval sound vibrations might have been inscribed by chance onto the surfaces of ancient clay pots, into the brushstrokes of old paintings, or into prehistoric lava flows.

The source materials I've assembled here share two features in common. First, they're all *pictures*, made in order to be apprehended visually, but they don't represent subjects in the visual world; instead, they were created specifically to depict sound, music, speech, or rhythm. Second, we can *play* all of these pictures, not as a musician plays a guitar, or even as a musical box plays a pinned cylinder, but in fundamentally the same way as a record player plays records or an iPod plays mp3s; and when we do, we hear something that resembles—to a greater or lesser degree—the sound, music, speech, or rhythm the pictures were originally intended to represent.

Because my sources are two-dimensional images on paper or parchment rather than vinyl LPs, magnetic audiotapes, or other audio formats for which standard playback equipment exists, I've had to convert them into sound in some unconventional ways. However, nothing I've done has involved the use of any custom hardware or software, and I've tried to explain all my techniques in enough detail that you should be able to duplicate them yourself if you wish. I've also given some account of the context, purpose, and format of each piece, although in many cases this only scratches the surface—there's much more that could be said about this material, and my purpose here is only to introduce you to it. Unless otherwise stated, translations from French, German, and Latin sources are my own.

Regardless of whether the various "pictures of sound" presented here change the way you think about historic audio forever or merely strike you as whimsical new forms of experimental music, I hope you'll find them consistently thought-provoking—and perhaps even enjoyable.

— Patrick Feaster,
May 2012

(((((((1. BALBASTRE'S "ROMANCE")))))))

(2:51)

The fourth and final volume of Dom François Bédos de Celles' monumental treatise *L'art du facteur d'orgues* [*The Art of the Organ-Builder*], published in 1778, contains a section on the programming of automatic musical instruments such as barrel organs, accompanied by a number of engravings that illustrate sample musical programs. One of those programs is the subject of our first track.

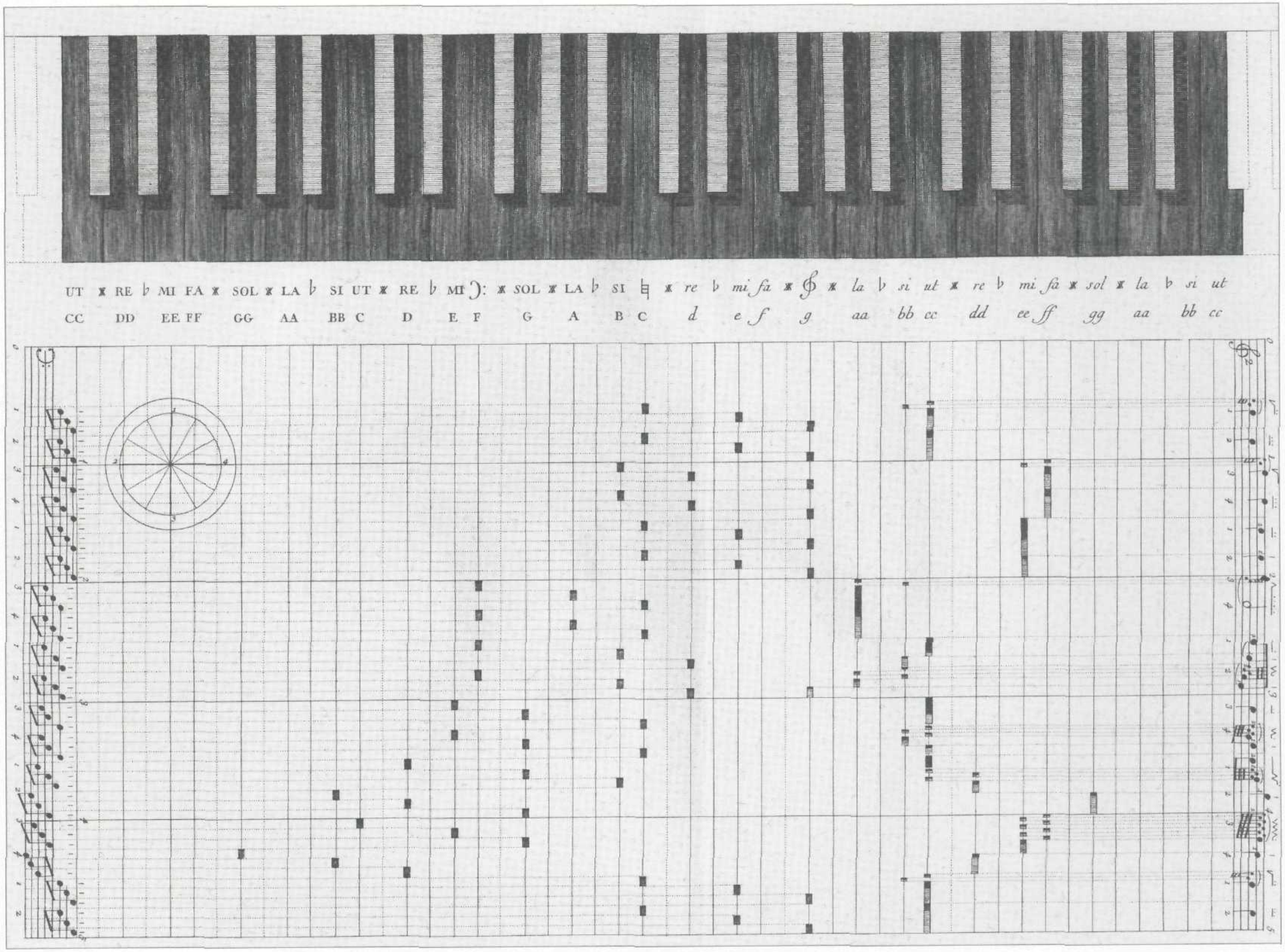
As it happens, Dom Bédos hadn't prepared the section on musical programming himself. Instead, he'd turned for assistance to Père Marie Dominique Joseph Engramelle, author of a recent book on the subject called *La Tonotechnie ou l'art de noter les cylindres*. A reasonably accurate translation of that title would be *Tonotechny, or the Art of Noting Cylinders*, but the French *noter* lends itself to a range of translations, all of which I've seen used in connection with Engramelle's work: "notating," "marking," "inscribing," "writing," "registering," and "recording." Translators seem especially to like the last of these, "recording," with its cagey implication that Engramelle's process aimed to achieve something in the spirit of later recording technologies such as the phonograph.

At first, Dom Bédos seems to have envisioned a collaboration with Engramelle, but in the end, he simply inserted a text which Engramelle had furnished him into his book under the heading "De la Tonotechnie ou Notage des Cylindres";[3] that is, "on tonotechny or the *notage* of cylinders." The technical term *notage* refers in French exclusively to the "setting of pins in barrel organs,"[4] but this, too, has sometimes been translated into English during recent years as "recording."[5]

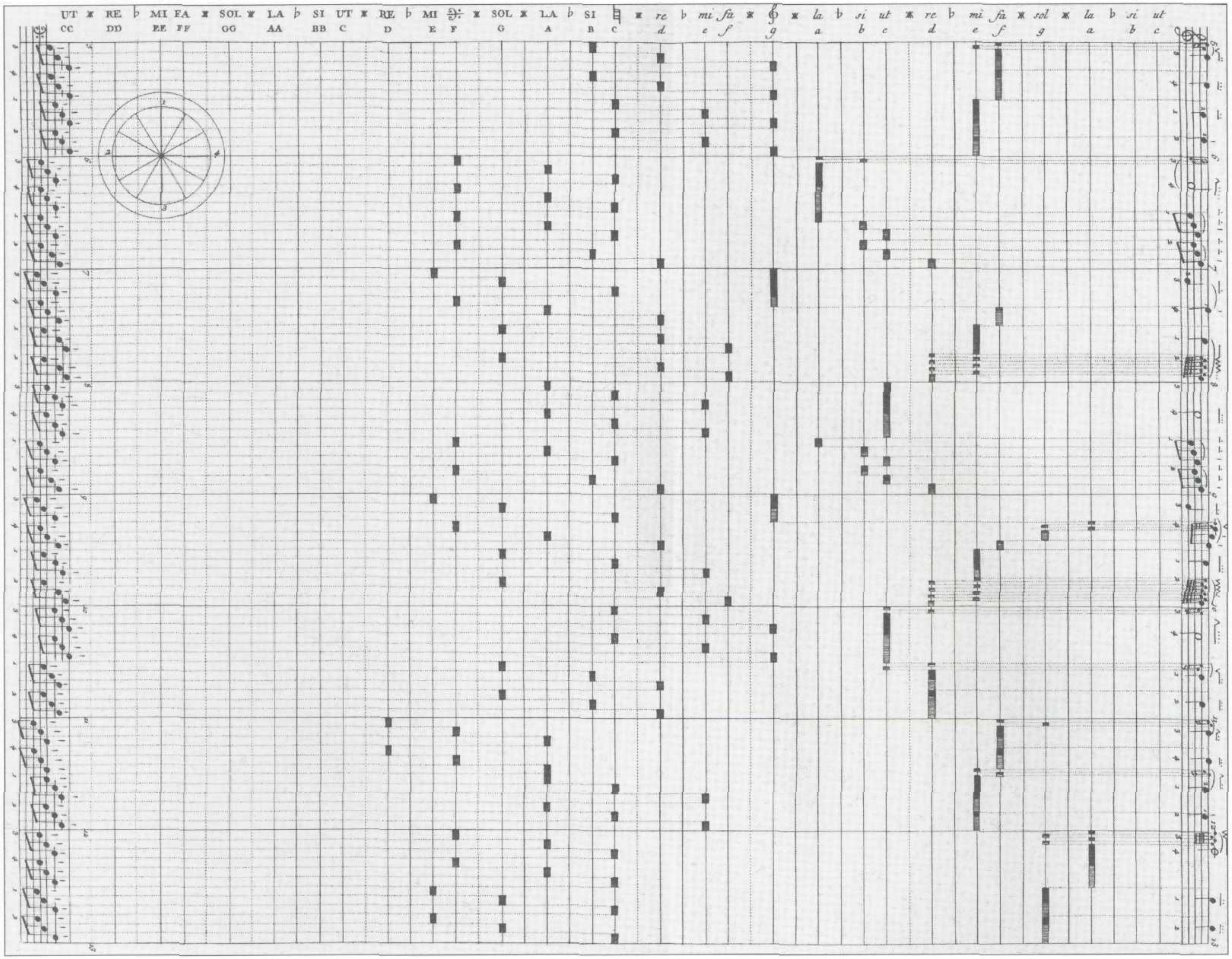
Of course, there's an important difference between programs for automatic musical instruments and what we ordinarily think of today as "sound recordings." Helmut Kowar spells it out nicely as follows:

> Mechanical instruments and sound recordings store music, but the encoding of the musical data is different: with musical automata, the carrier of information is a program especially designed to activate mechanisms which play a musical instrument. The carrier of sound recordings represents an image of the acoustic events depicted as waves in analogue or digitized form.[6]

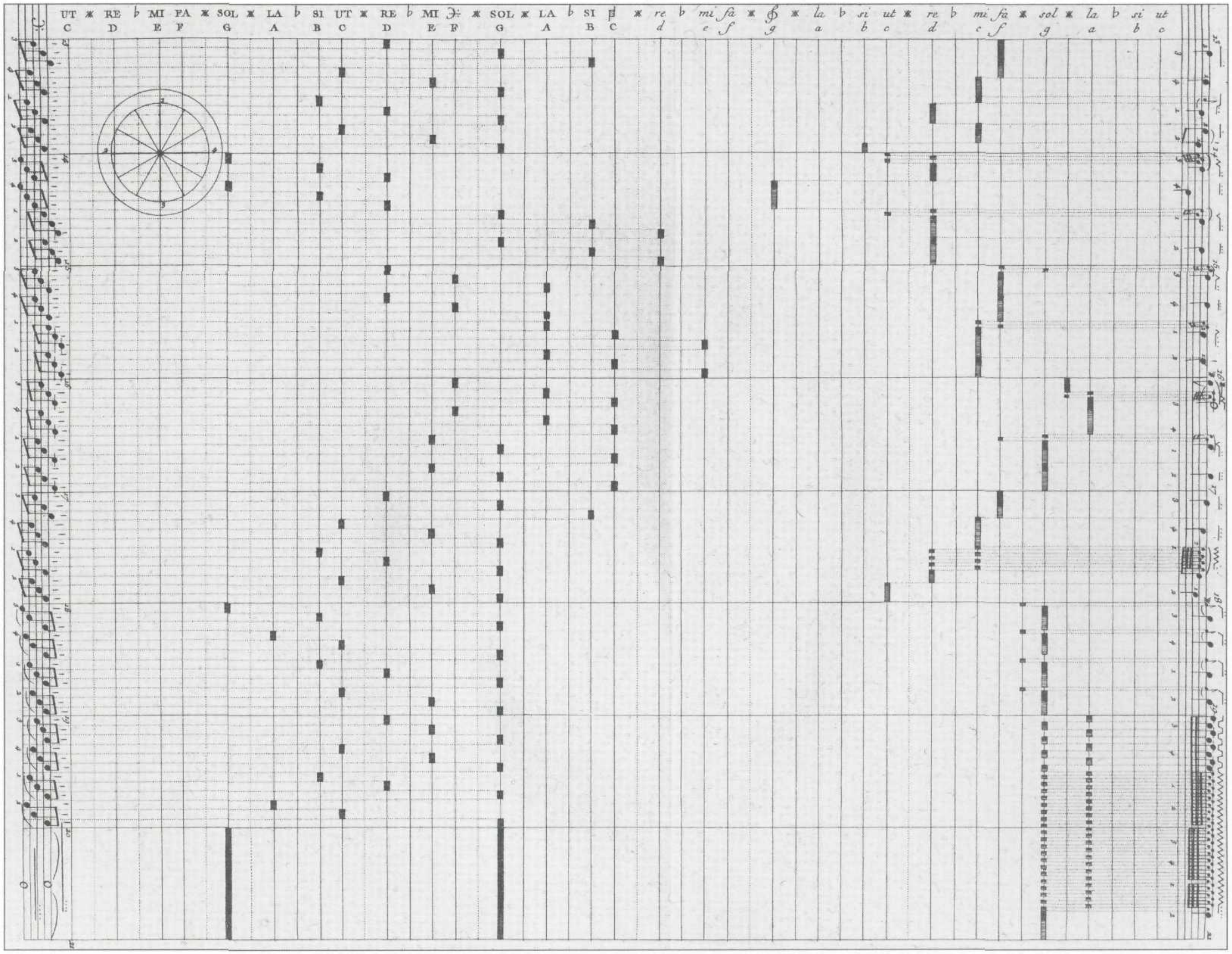
Still, the art of tonotechny did bear at least a superficial resemblance to "sound recording" as people typically construe that term today: it entailed inscribing pieces of music with unusually great detail and precision, and in a format intended to be played automatically. In Engramelle's view, those two characteristics necessarily went hand in hand.



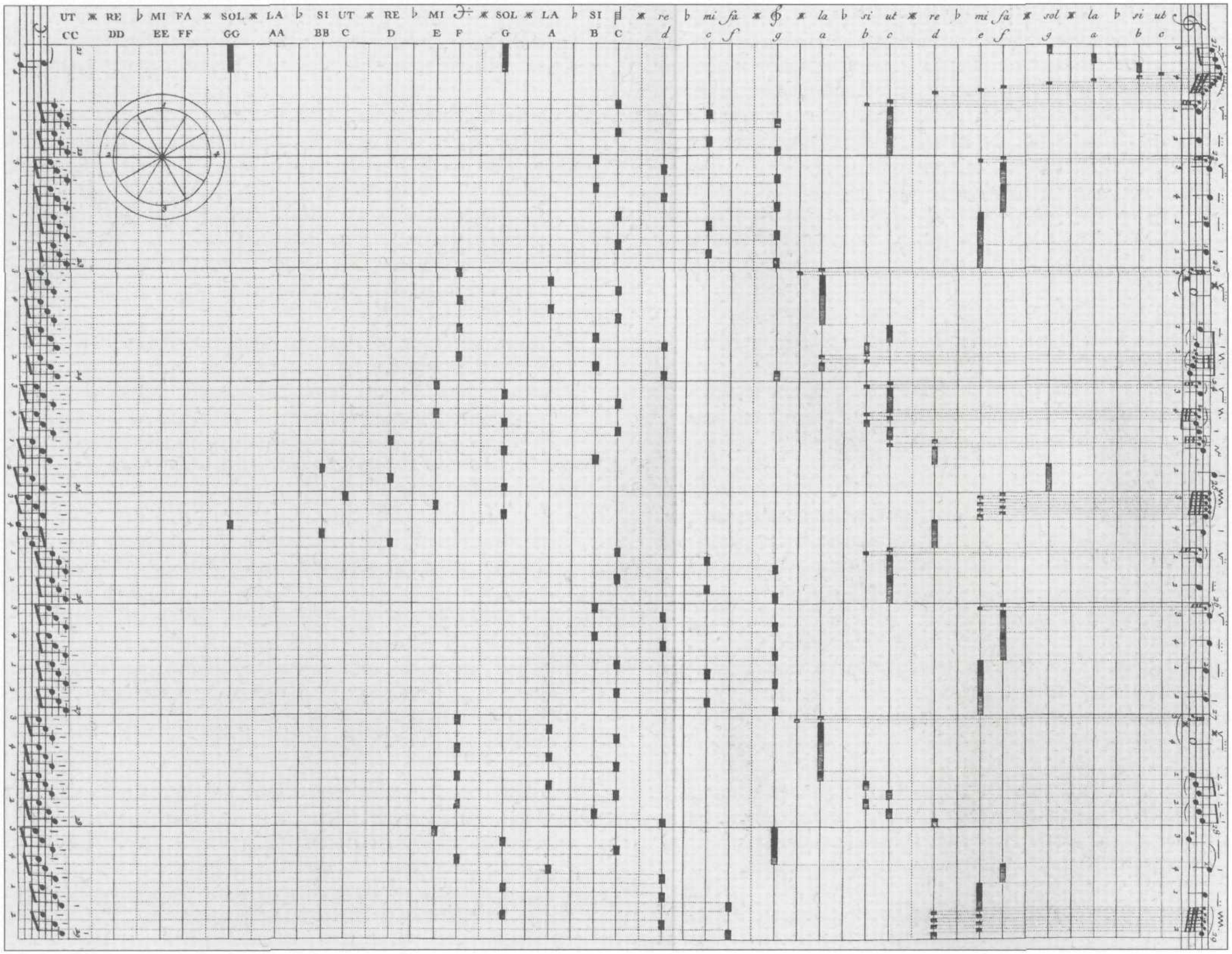
Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 120.*



Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 121.*



Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 122.*



Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 123.*

It all came down to fundamental differences between man and machine. When live musicians played musical compositions read from conventional staff notation, they always injected some degree of subjective artistry into their performances, whether or not they were consciously aware of what they were doing. For example, a performer at the organ didn't just play the notes shown on the page, but also inserted unwritten silences between the notes—"silences of articulation," as Engramelle called them. By contrast, mechanical musical instruments didn't have the ability to make such subjective interpolations; they played *exactly* what they were programmed to play. That reality demanded a new way of thinking about music and how to inscribe it, which, as Engramelle explained, was where tonotechny came in:

> There is a method of conceiving of music entirely different from the one taught in all the treatises on this art; it is based on performance itself.
>
> Cylinders need to express this performance in the greatest detail; that is, to represent not only the notes, but all their constituent parts, of which the aggregate produces the whole effect in performance.[7]

By "constituent parts," Engramelle seems to have meant those juxtapositions of held notes and silences which conventional musical notation left to the discretion of live performers, but which automatic musical instruments needed to have spelled out for them correctly by means of pins and spaces if their playing wasn't to sound ghastly.

The word Engramelle used for performance, *exécution*, could refer simply to the mechanical execution of a note or a piece of music, but when he wrote that tonotechny was "based on performance itself," it's evident that he meant performance in an artistic sense as well. This comes across clearly in his discussion of one of the specimens of *notage* he provided to his readers by way of illustration, extending over nine fold-out engravings: a "Romance" composed and performed by Claude Balbastre, one of the most famous French musicians of the latter half of the eighteenth century (among other things, he taught Marie Antoinette how to play the harpsichord).

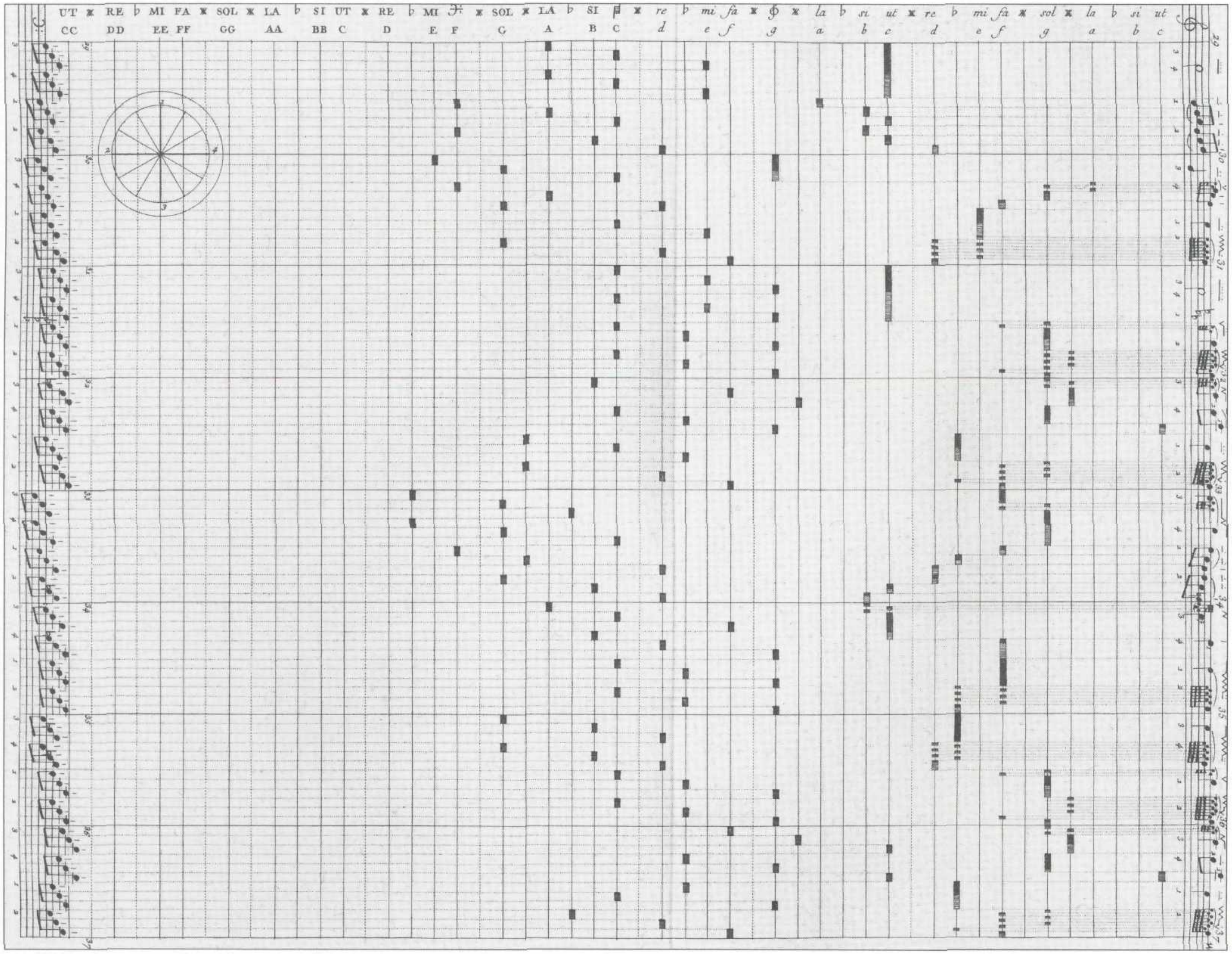
Engramelle understood his *notage* of Balbastre's "Romance" as a crowning masterpiece of his art, an ideal example of tonotechny at its best and most impressive. And what made it so impressive, he believed, was that it embodied a specific live performance style with impeccable accuracy—that of Claude Balbastre himself, who had cooperated with him in its preparation:

> Monsieur Balbastre, one of the most famous organists of Paris, submitted himself to all we asked of him in this regard; not only did he go to the trouble of noting his piece on paper himself, such as is seen engraved; but he performed it several times, and his performance was monitored with a second-watch in hand; this is why one is in a condition to assert that his entire piece should not exceed 165 seconds in duration; and that, having been reviewed by himself, it is in his true style [*genre*] of performance.[8]

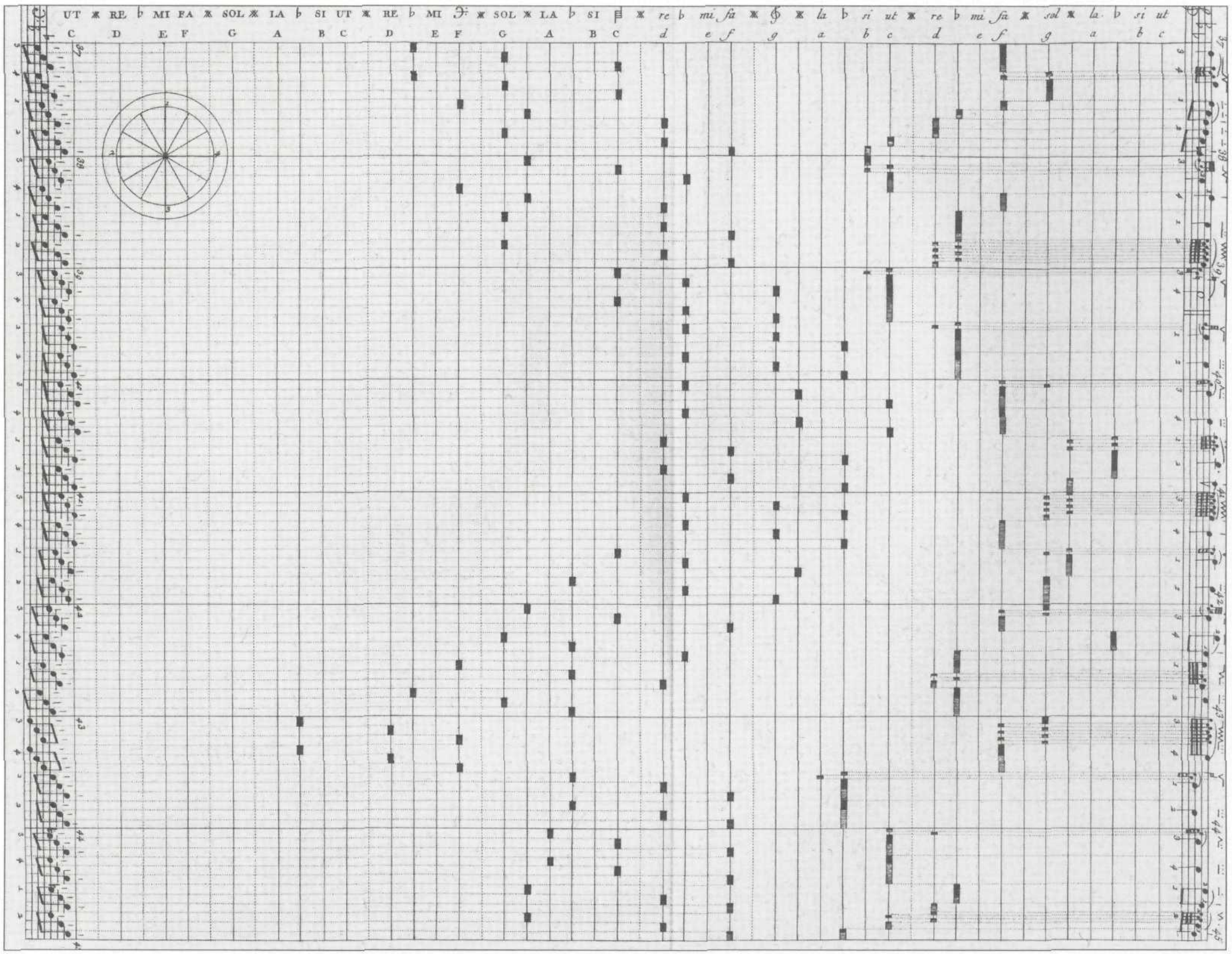
Balbastre's willingness to repeat his "Romance" several times for Engramelle's benefit must have been helpful when it came to verifying subtle details of performance style, but it also presented a dilemma: the different performances weren't quite identical to each another, so Engramelle sometimes had to choose among variant renditions when noting a particular passage. This was true, for instance, of the trill heard approximately 48 seconds into the track:

> In noting this trill thus, it will be in Monsieur Balbastre's most usual style; for it happens sometimes that he prolongs this graduated increase of speed a bit more, and at other times that he diminishes it: the detail I've given here of this increase for the total duration of a half-note will suffice to show the means of doing it for the duration of five eighth-notes, of three quarter notes, of seven eighth-notes, and even of a whole note if one wishes, the one being no more difficult than the other; the important thing in *notage* is to catch as much as one can the true style of the author.[9]

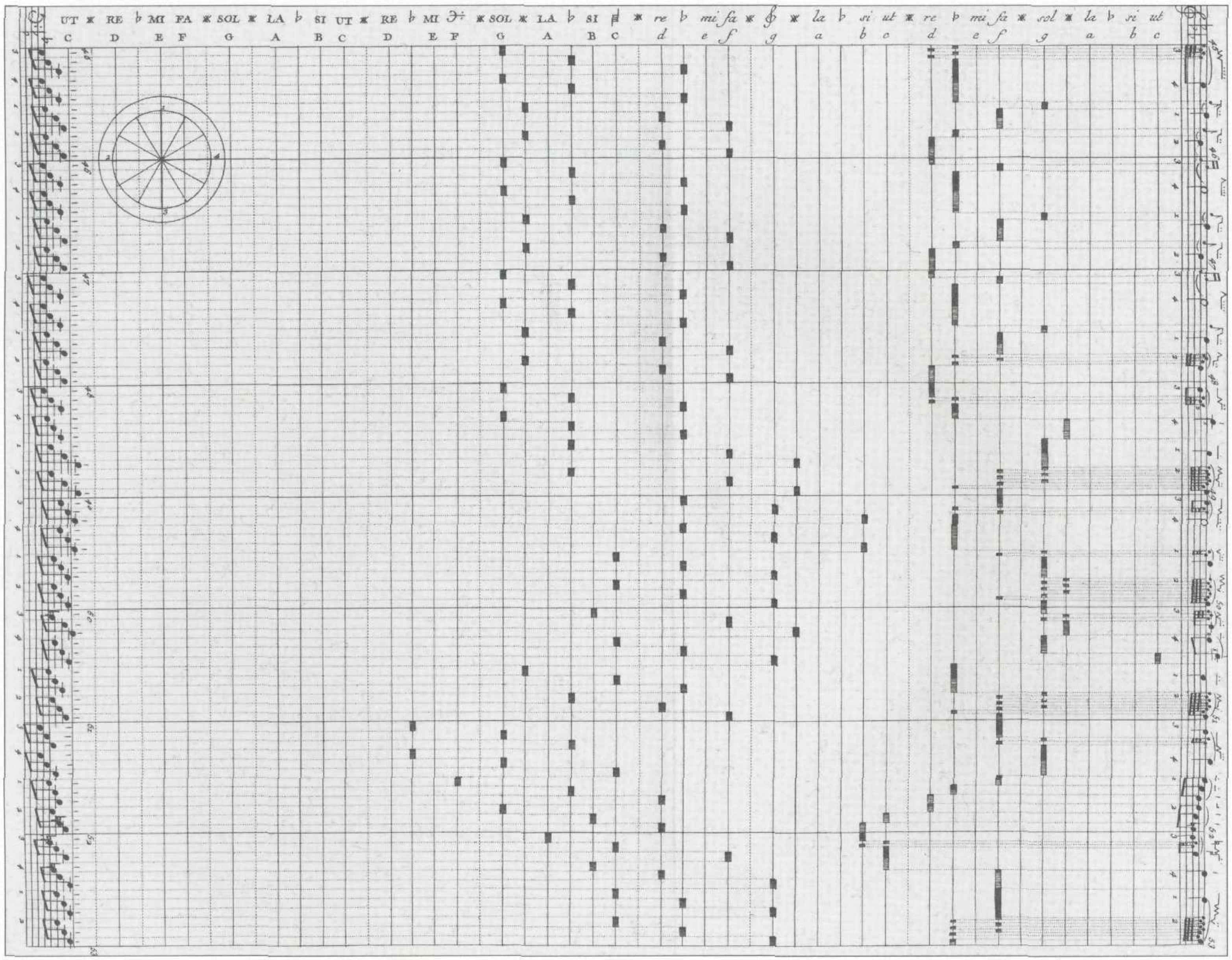
Engramelle's meticulous efforts paid off in the long run: today, scholars prize his *notage* of Balbastre's "Romance" as a unique document in the history of eighteenth-century musical performance practice. David Fuller identifies it as, to the best of his knowledge, "the *only* piece on a mechanical instrument which was set up under the direction of the composer and the only one chosen with the express purpose of illustrating the ability of a cylinder organ to reproduce the ornamentation, articulation, and rhythmic nuances



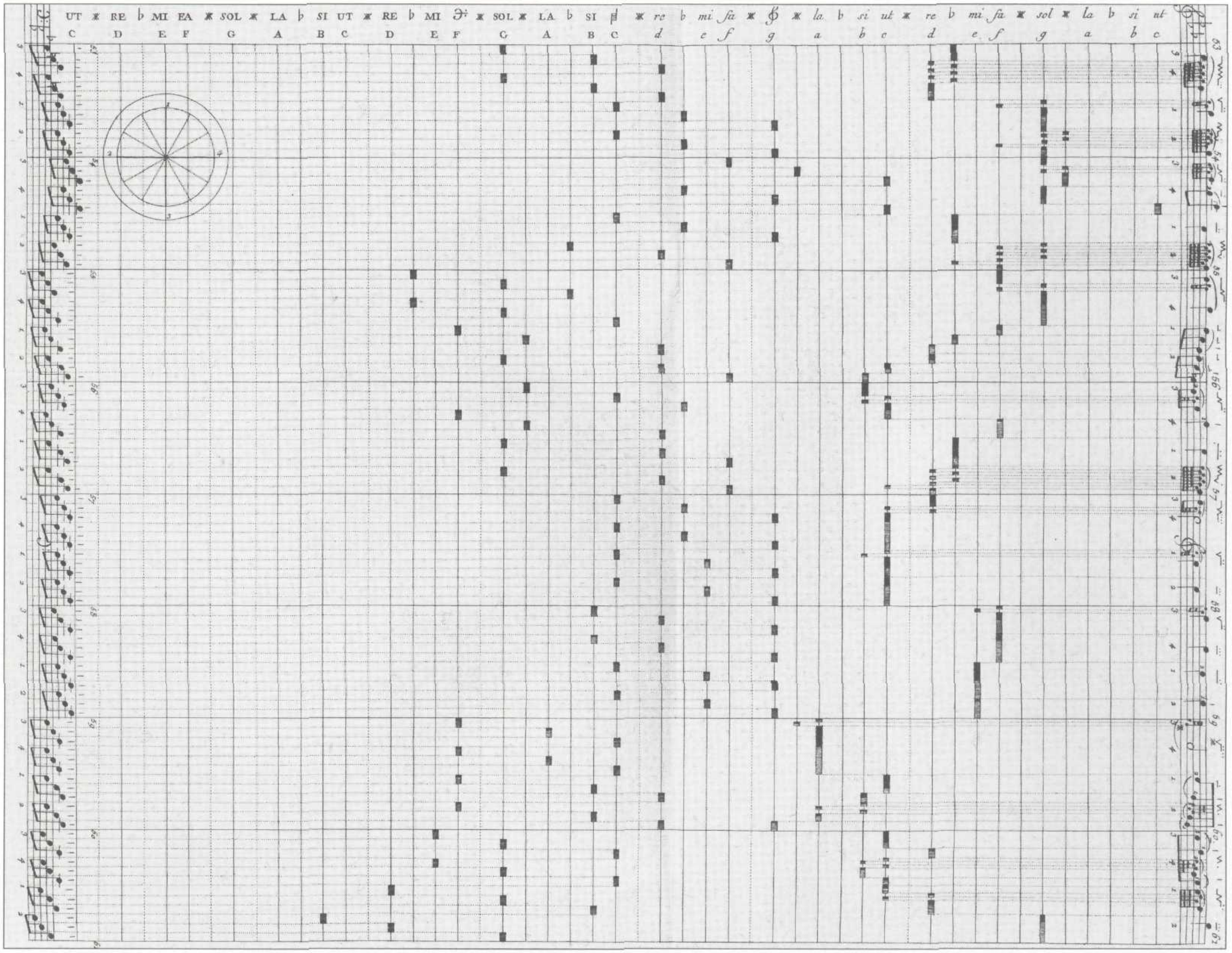
Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 124.*



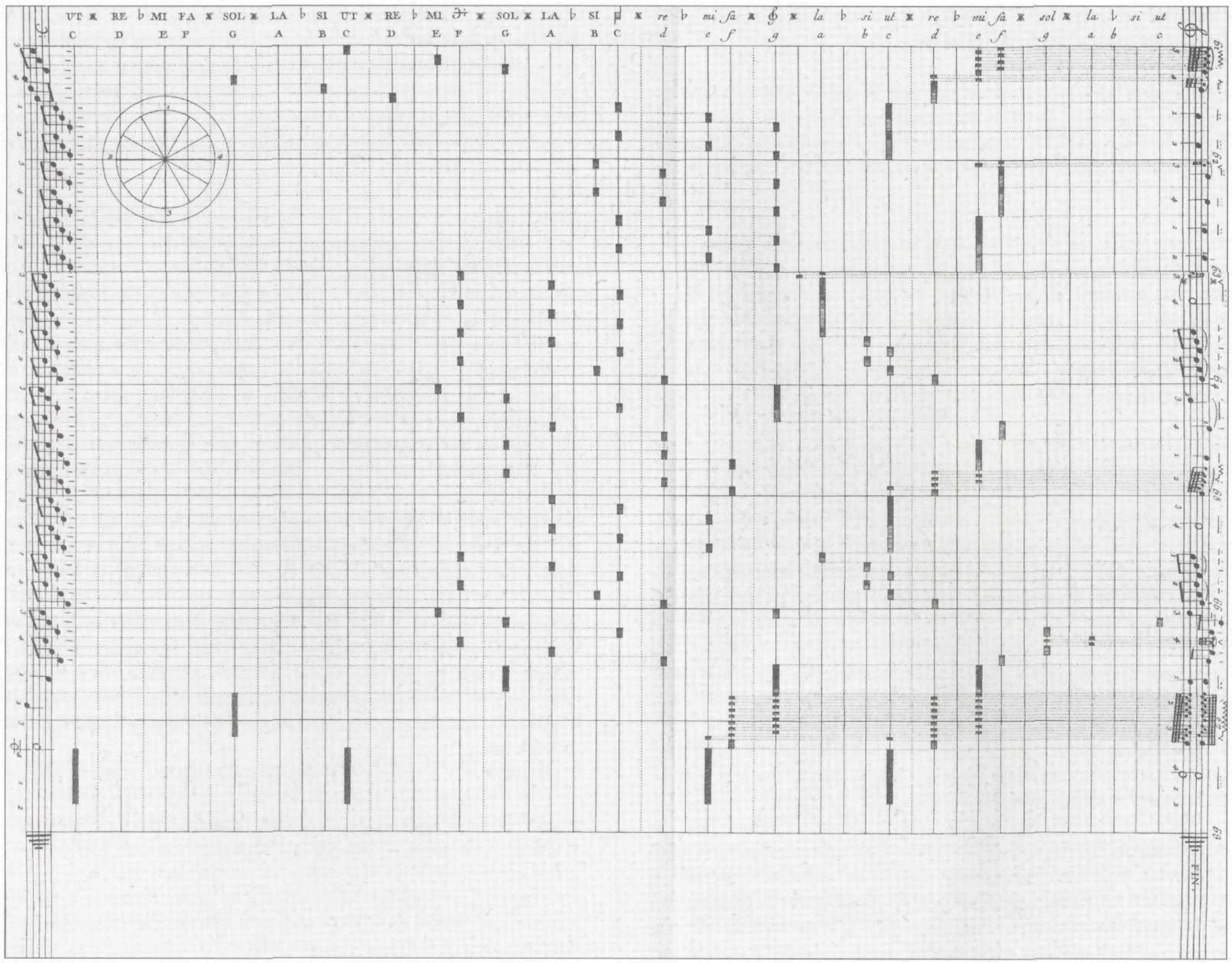
Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 125.*



Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 126.*



Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 127.*



Dom François Bedos de Celles, L'art du Facteur d'Orgues, *Vol. 4 (Paris: L. F. Delatour, 1778), Plate 128.*

of an ideal live performance."[10] It has been repeatedly studied as a source of insight into the musical aesthetics of its age.

Several attempts have been made in the past to play Engramelle's *notage* of Balbastre's "Romance" in all its rich detail—quite successful attempts, I might add. It's been played by an actual mechanical organ from a perforated sheet, for example, and it was first played synthetically by computer in the early 1970s:

> Dr. Peter Gena...developed a computer sound-synthesis language called ELMUS which he used to produce the *Romance*. A short program was prepared so that the data for each note and rest could be entered into the computer correctly.... ELMUS converted the frequency and duration for each note into a wave form of a stopped flute with five partials at correct proportional intensities. Each note had a quick attack and a decay time of .03667 second.[11]

Judging from the account just quoted, Gena must have measured where each note began and ended in Engramelle's diagrams and then entered this information manually into a program that could be used to synthesize the corresponding musical notes in a "stopped flute" voice. This approach doesn't differ much from taking standard musical notation and creating a MIDI file from it by keying in the data by hand, note by note—except, of course, that the data for Balbastre's "Romance" happens to be unusually detailed and precise.

The version of the "Romance" presented in track one was likewise created using a computer, but in a very different way. I didn't assign any voice to the notes, and I didn't enter any information by hand about which individual notes should be played, or for how long—in fact, I didn't pay any attention whatsoever to the music as such. Instead, I merely played the sequence of images *as though it were a sound recording.*

Before you ask, no, I didn't just slap the book down on a turntable and drop a needle on the page—but my technique isn't much different from that in spirit: I took the actual marks printed on paper in the year 1778 and converted them into sound using a straightforward digital method that we can also use to play back ordinary recordings of speech and other sounds made centuries later. And the result isn't just random noise. To the contrary, we hear the very piece of music the inscription was drawn up to embody back in the eighteenth century, with every subtle aural nuance based directly on the details of the original inscription (including the occasional mistake). In other words, the technique I'm discussing isn't just a handy way of producing something audible from these pictures as a novelty; it's actually *appropriate* to the material.

Now, I don't mean to claim that the method demonstrated here is better, more accurate, or more authentic than the others that have been used in the past to play the barrel-organ notation for Balbastre's "Romance" and other, similar inscriptions on paper from long ago. Each of these methods has its advantages and disadvantages. However, I believe mine is the least *intrusive* method available for converting these inscriptions into sound—I don't do anything subjectively to them beyond specifying time and frequency scales, a step roughly equivalent to deciding on a playback speed for a gramophone disc. Quite apart from the question of whether such inscriptions might count as "recordings" in the first place, then, we can treat them as full-fledged sound recordings in practice if we choose, enabling ourselves to listen to them in an uncannily direct way. In track one, nothing gets between your ear and the original eighteenth-century ink print; you experience Engramelle's inscription as sound with an absolute minimum of twenty-first-century intervention.

If you're wondering how all this is possible, don't worry—I'll explain everything in the notes that follow, and we'll take Balbastre's "Romance" up again for consideration towards the end of our program. For now, I'd like to move on with this teaser: the sounds you'll hear in track two were converted into audio using the very same technique I used to play Engramelle's *notage* in track one.

Source: Dom François Bedos de Celles, *L'art du Facteur d'Orgues,* Vol. 4 (Paris: L. F. Delatour, 1778), Plates 120-128.

((((((2. THIS IS A SOUND SPECTROGRAM))))))

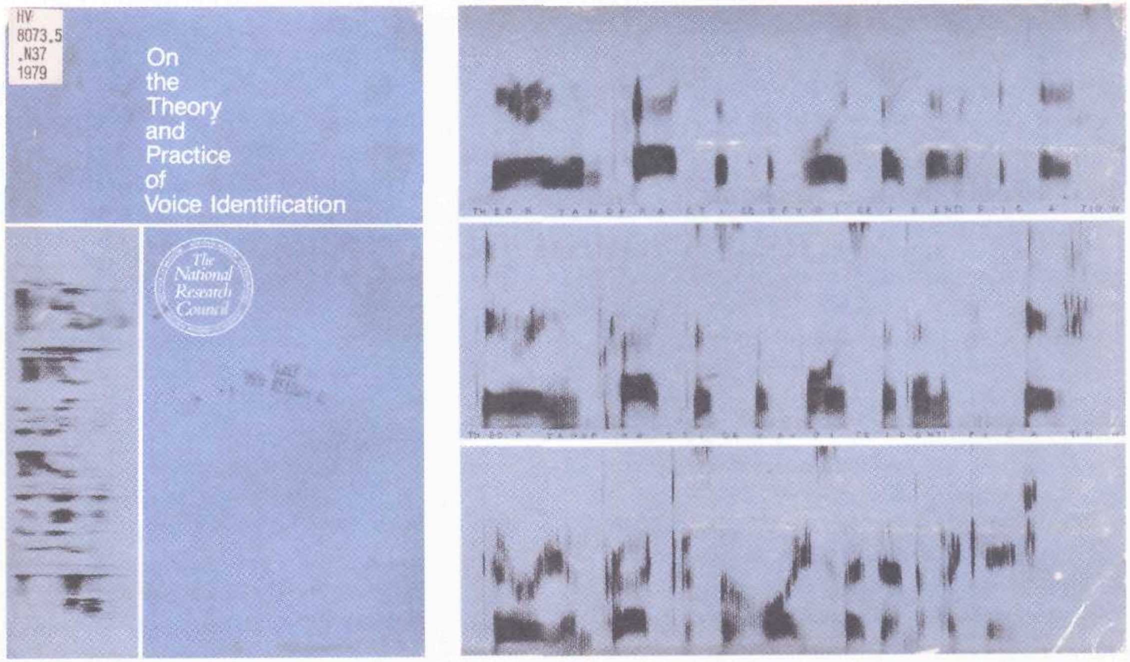
(2:21)

The automatic sound spectrograph emerged from Bell Laboratories after World War II as a powerful new tool for visualizing recorded sounds—it's the analog ancestor of the digitally rendered spectral displays in modern sound-editing software. How it works is less important for our purposes here than the fact that it does work, automatically decomposing complex audio waveforms into their constituent frequencies through a Fourier transform and displaying these graphically so that the trained eye can make sense of them. That said, different spectrographic settings yield somewhat different results: a *narrowband* spectrogram has higher resolution in the frequency domain than a *wideband* spectrogram does, but lower resolution in the time domain. With this trade-off in mind, researchers who work with spectrograms choose between narrowband and wideband spectrograms depending on what they want to study.

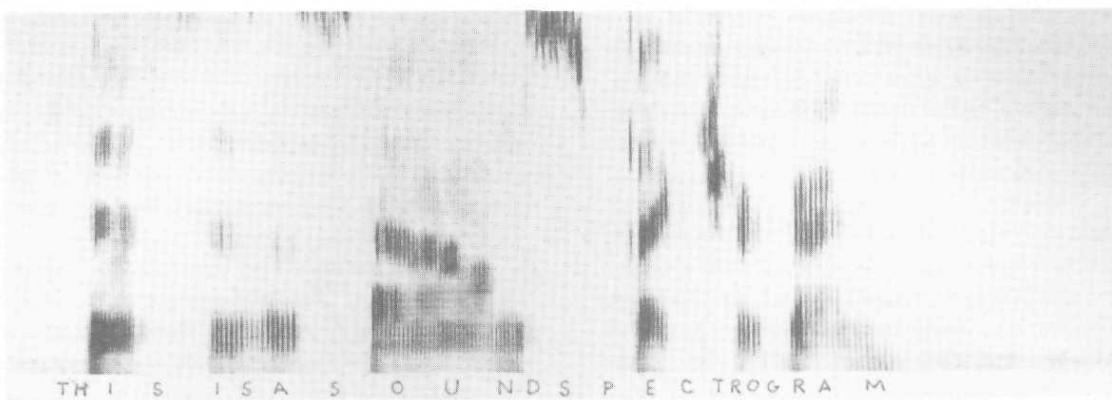
With sufficient effort, people can actually learn to "read" spectrograms of recorded speech, deciphering spoken words just by looking at their spectrographic patterns. Wideband spectrograms are easier to "read" in this sense than narrowband ones are because they more clearly show the resonances of the human vocal tract that characterize different phonemes, the aural building blocks of language. Many of the principles involved in "reading" spoken language spectrographically were first set forth in a book called *Visible Speech* (1947), which is also an excellent source of early sound spectrograms if you'd like to study the subject further.[12]

But the speech spectrograms featured in track two weren't made just to be deciphered into words. Instead, the National Academy of Sciences prepared them in the 1970s in connection with a study carried out at the request of the Federal Bureau of Investigation to explore whether the law enforcement community might be able to use sound spectrograms to help identify the voices of individual people. All four repetitions of the phrase "theory and practice of voice identification" are taken from the cover of the resulting book, *On the Theory and Practice of Voice Identification*—given the subject matter, someone must have thought it would be clever to present the book's title on the cover spectrographically, as spoken by four different people, and not just in the form of conventional text. But here's what the authors had to say inside about the three repetitions of "this is a sound spectrogram":

> The three samples come from male speakers saying, "This is a sound spectrogram." The spectrograms on page 4 come from Speaker A; that on page 5 comes from Speaker B. These samples show that two utterances of the same speech material spoken by the same person can produce distinct differences in the spectrograms and that larger differences can appear when the same material is spoken by a different person.[13]



"Theory and practice of voice identification (4×),"
On the Theory and Practice of Voice Identification, front and back covers.



"This is a sound spectrogram." *On the Theory and Practice of Voice Identification*, page 4,
Figure 1, "Three Samples of Speech Sound Spectrograms," first sample: "Speaker A: Day 1."

In other words, spectrograms of the same phrase spoken twice by the same person looked different from each other, but spectrograms of the same phrase spoken by different people looked even *more* different, hinting that they might indeed be useful in forensic voice identification. The authors also included two spectrograms of the phrase "I can see you" to illustrate the respective merits of narrowband and wideband spectrograms:

> Figures B-1 and B-2 are voicegrams of the phrase "I *can* see you," with the word *can* emphasized. An uncertainty exists when making a frequency/time examination of a signal. We may either measure frequency more accurately at the expense of time accuracy—the result of using a "narrow band" analyzing filter—or we may measure time more accurately at the expense of frequency accuracy—the result of using a "wide band" filter. Figure B-1 displays a narrow band analysis and Figure B-2 displays a wide band analysis. Note that in Figure B-1, during the vowel or voiced sounds, the fundamental frequency and the harmonics are well displayed. In Figure B2, the vertical lines or striations that occur during the vowel sounds represent the time delineation of the individual puffs of air through the glottis.[14]

Overall, the authors summed up what they were hoping to accomplish as follows:

> Visual displays such as the speech sound spectrograms shown here permit examination in a manner different from that afforded by listening to the sound. The eye can wander around the picture and across the time and frequency dimensions in an unconstrained way, and thereby can seek and examine small details in the physical features of the speech sounds represented graphically. In a different way, listening to the speech offers a natural and long-practiced ability to assimilate information about the meaning, nuances, dialect, and identity of the speaker. Of course neither the spectrogram examined nor the sound heard contains any voice information that was not present in the original speech signal.[15]

A reader looking at these spectrograms back in 1979 might have questioned whether they could truly embody the details of speech thoroughly enough to aid FBI investigations—how much information was really in there? By

Bald Eagles. *Identifying Individual Bald Eagles with Voiceprints: a feasibility study*, Figures 1-5.

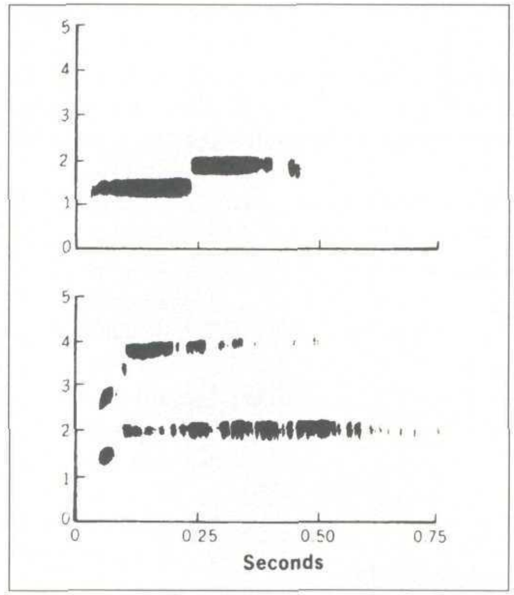


Figure 1: Sonagrams of 'wail' calls of two bald eagles.

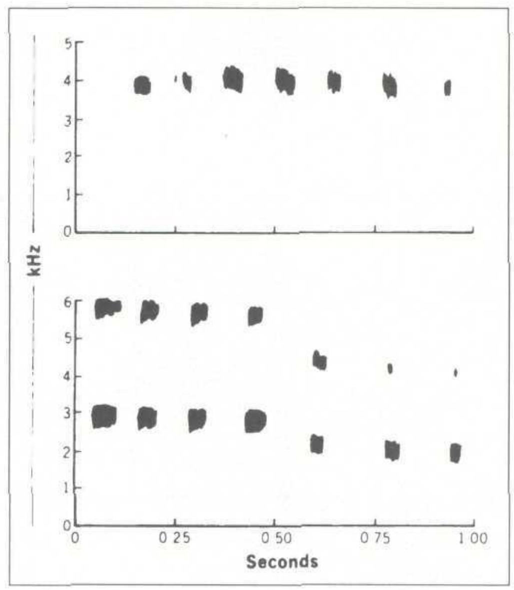


Figure 2: Sonagrams of 'peal' calls of two bald eagles.

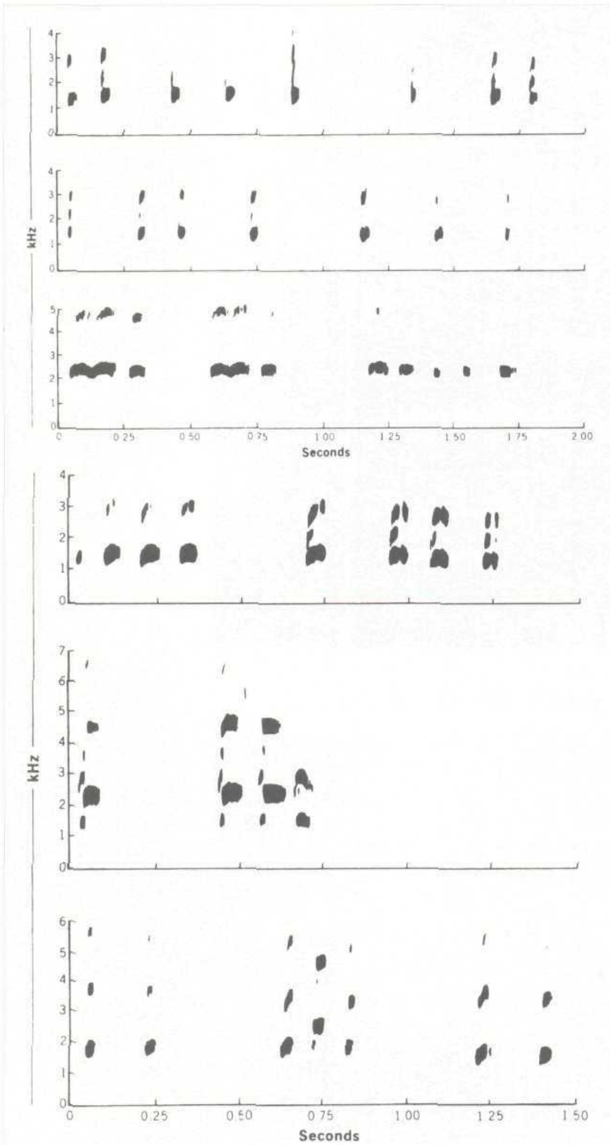


Figure 3—Sonagrams of 'chatter' calls of six bald eagles.

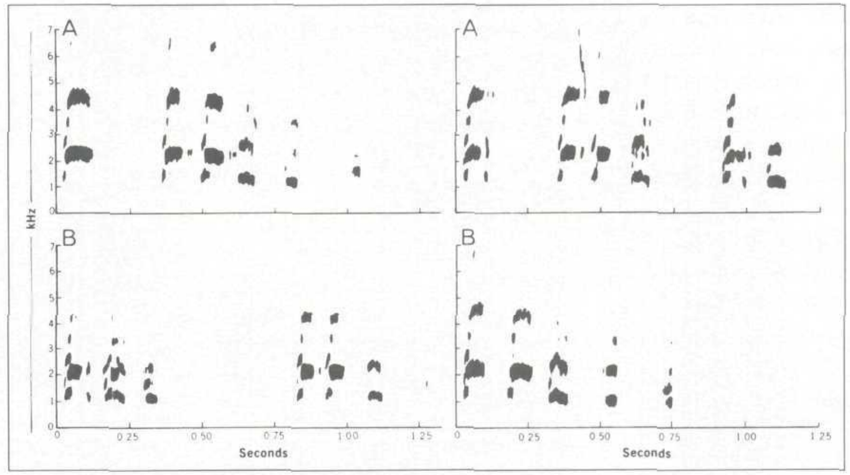


Figure 4—Sonagrams of four different chatter calls of one bald eagle in the same year. (*A*) May 19, 1980, recordings; (*B*) June 24, 1980, recordings.

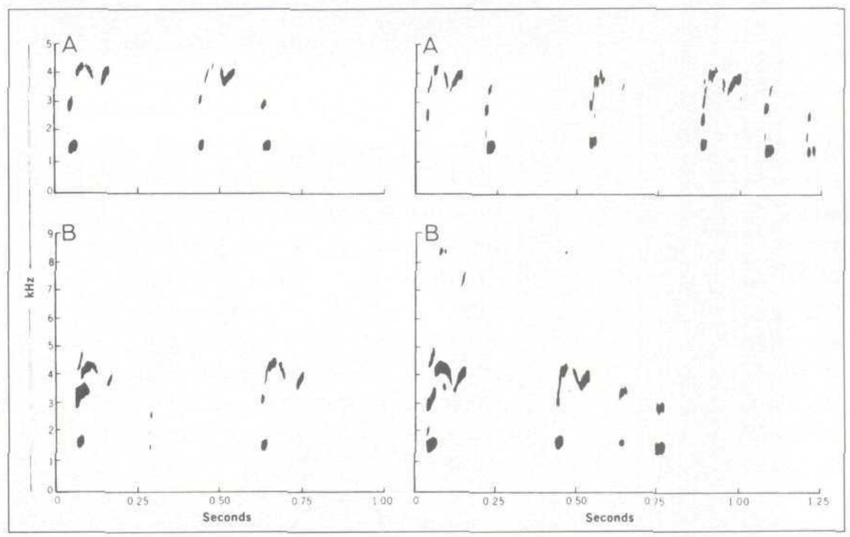


Figure 5—Sonagrams of four different chatter calls, presumably of the same bird in two consecutive years. (*A*) May 19, 1979, recordings; (*B*) June 14, 1980, recordings.

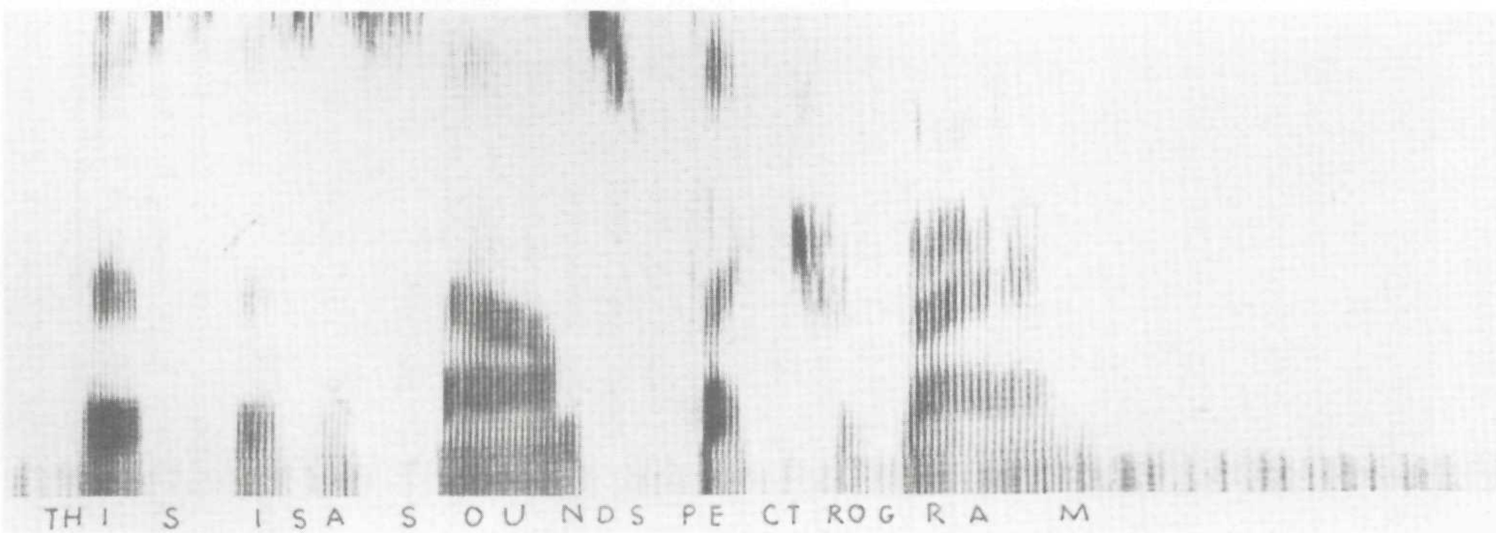
playing back the spectrograms today, we can finally answer that question in an exciting and aurally compelling way.

And it's surprisingly easy to do this. Many readily available pieces of software can now convert sound spectrograms back into sound, even though that's not how they're usually explained or advertised. They work by means of *additive synthesis*—that is, they generate sine waves at all the various amplitudes and frequencies indicated by the positions and intensities of pixels in an image interpreted as a graph of frequency against time, *adding* them together into a composite waveform. The developers of such software generally target it at creators of experimental music who want to translate random digital images into audio as a means of producing novel effects—for instance by "playing" a photograph of a tree's branches, the Mona Lisa, or a piece of abstract visual art. But the same software also works on actual sound spectrograms, reversing the Fourier transform that had decomposed the complex wave into its component frequencies in the first place. In other words, we can repurpose it as a *playback technology* for sound spectrograms, and that's exactly what I've done throughout track two. The specific program I've put to this use is AudioPaint by Nicolas Fournel.[16]

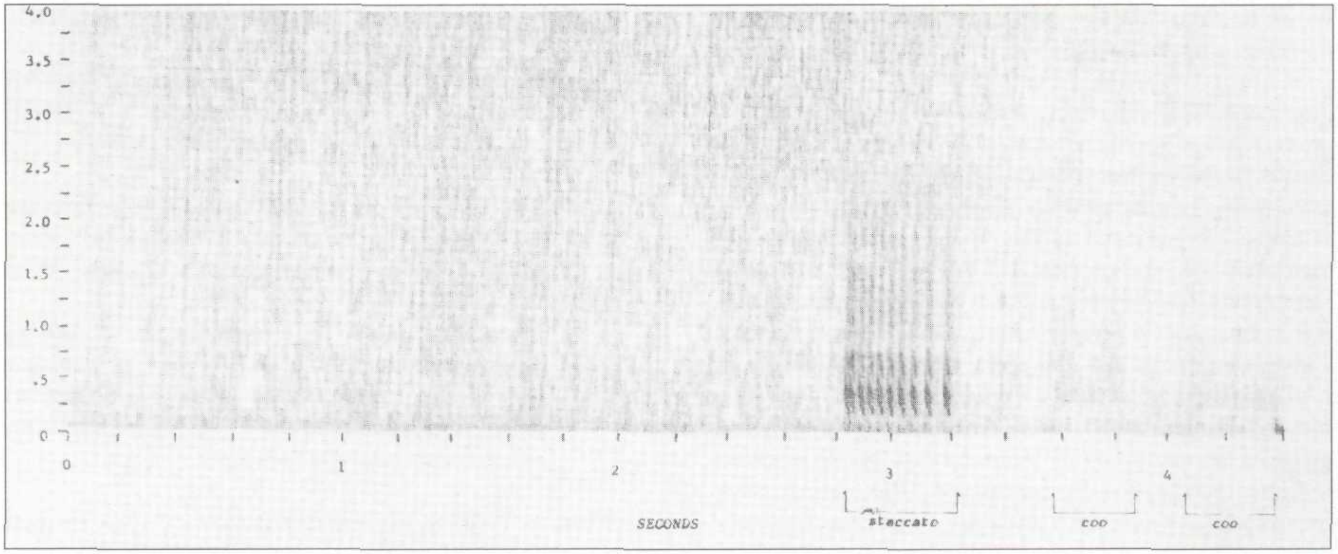
To play back the spectrograms in *On the Theory and Practice of Voice Identification*, all we need to do is scan them and run the digitized images through the additive synthesis software, setting two parameters:

1. **Time Scale:** What should the duration be from beginning to end? Answer: the text states that "these samples show a segment of speech lasting about two seconds," so we set the duration to two seconds.
2. **Frequency Scale:** What frequencies should correspond to the top and bottom of the image, and should the scale be linear or logarithmic? Answer: the text specifies that "these samples show a range of about 100 to 4000 Hz (hertz or cycles per second)," and that they employ a linear scale.[17]

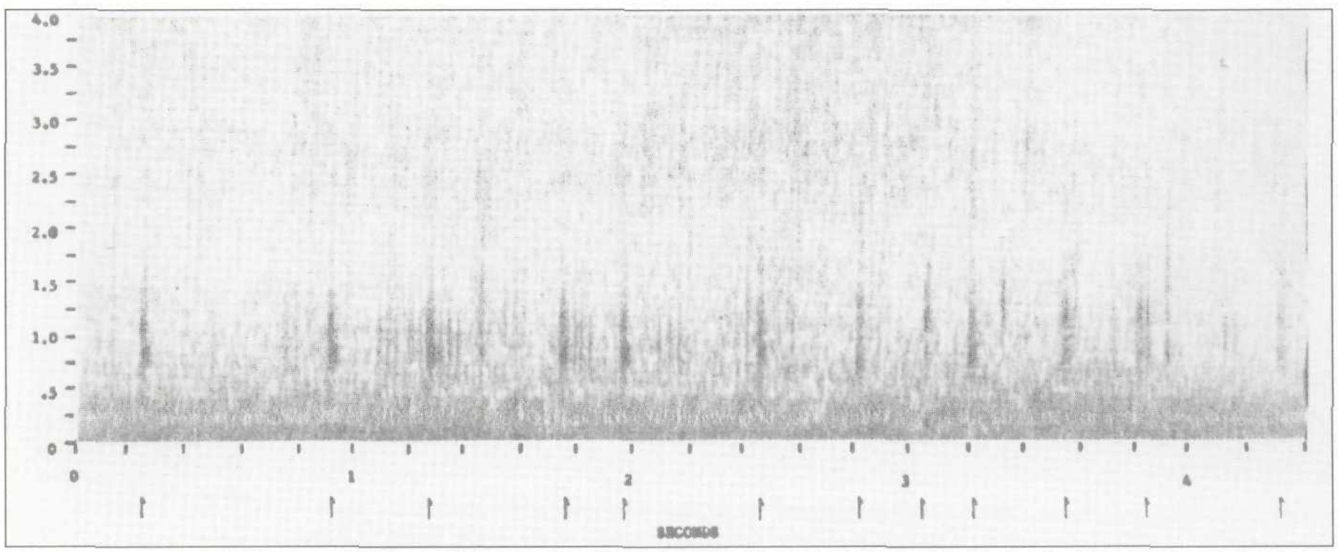
When we listen to the results, the recorded speech is perfectly intelligible, although the voices sound oddly whispery. From the perspective of



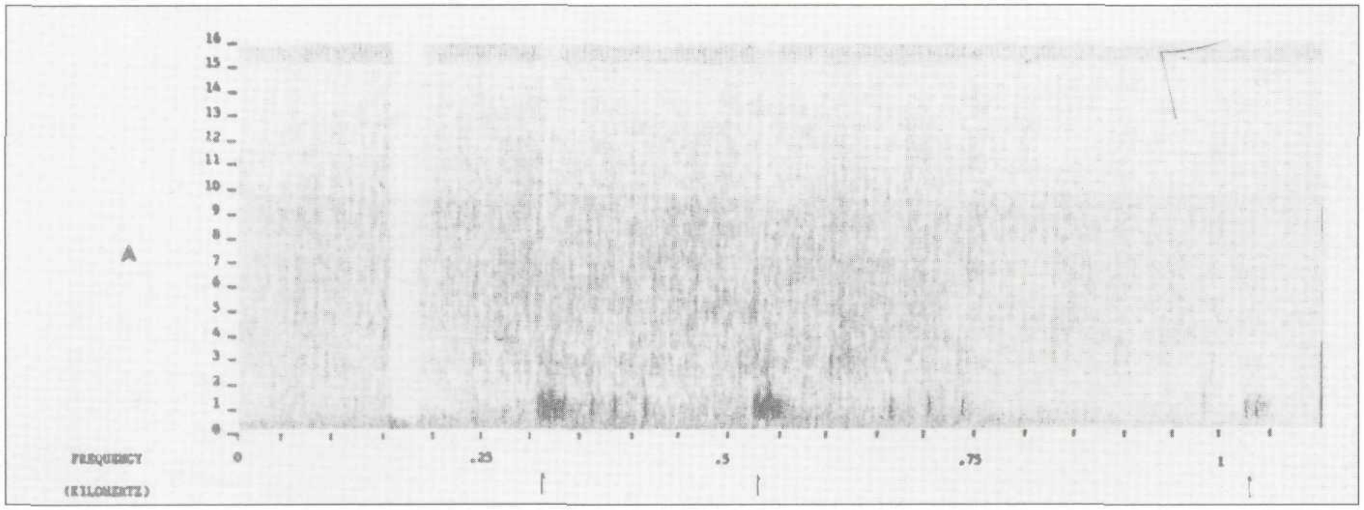
"This is a sound spectrogram." *On the Theory and Practice of Voice Identification*, page 4, Figure 1, "Three Samples of Speech Sound Spectrograms," second sample: "Speaker A: Day 4."



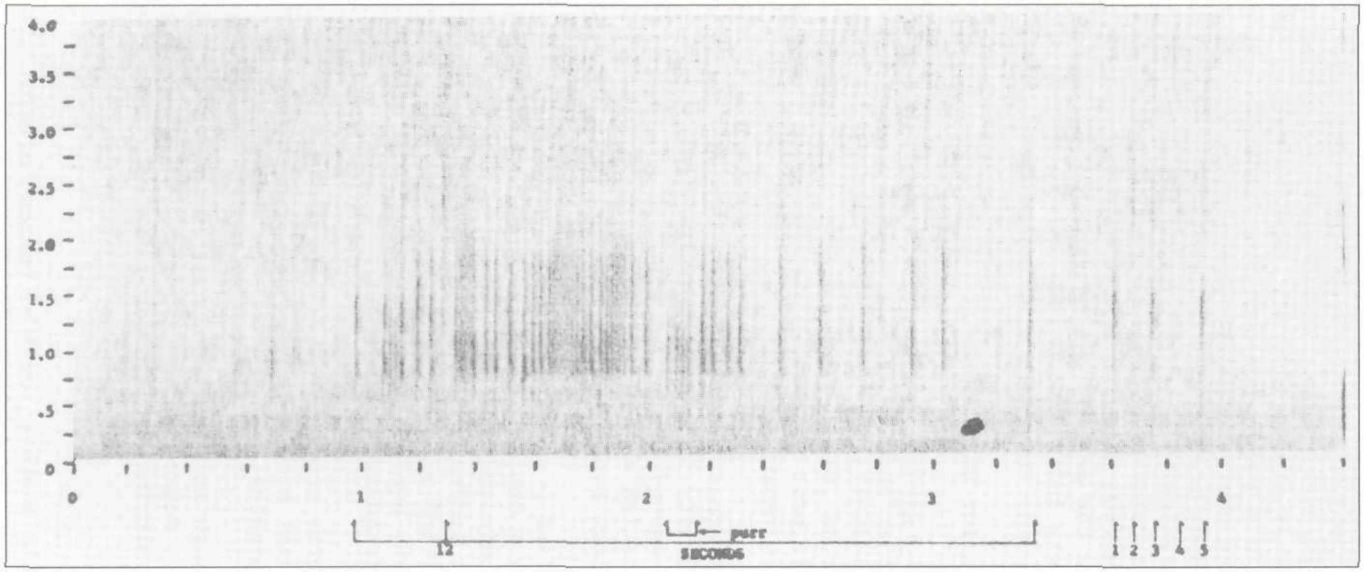
Marine Organisms.
Figure 3B, Staccato of Holocentrus ascensionis.



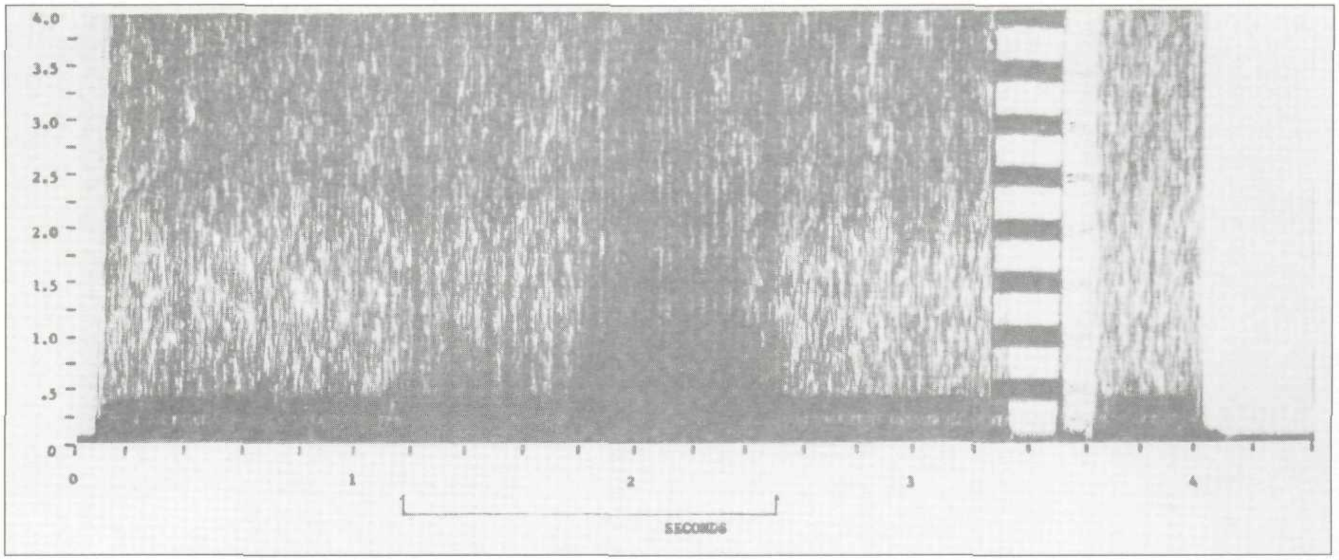
Marine Organisms.
Figure 6, Quacks of Flammeo marianus.



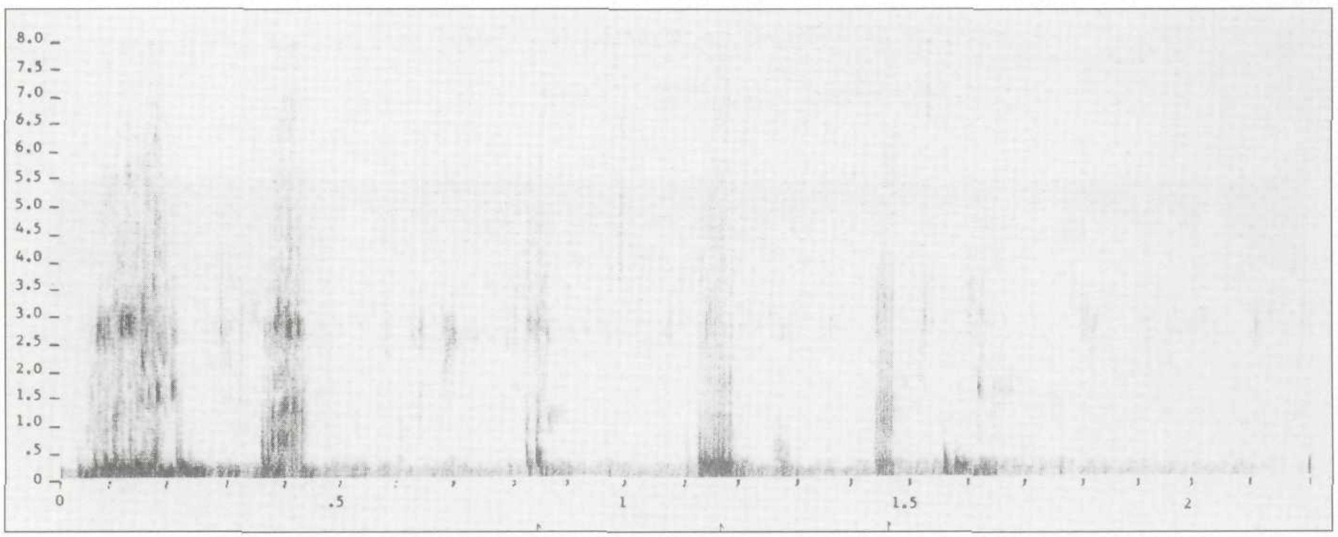
Marine Organisms.
Figure 7A, Quacks of Flammeo marianus.



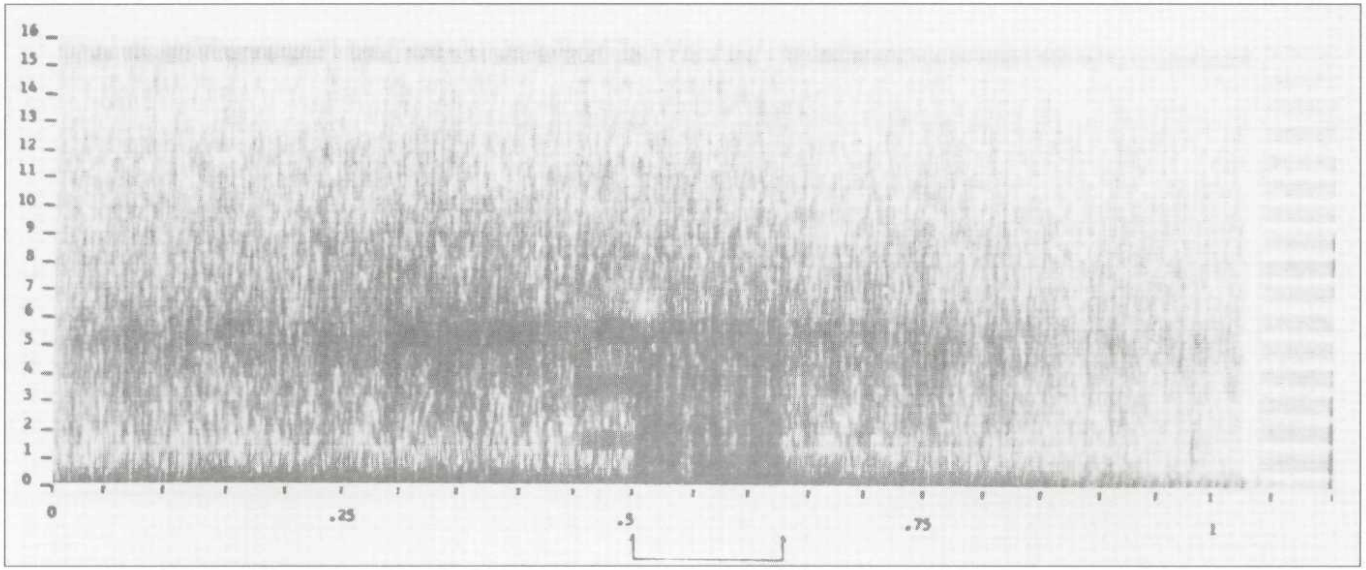
Marine Organisms.
Figure 8B, Squeaking door sound.



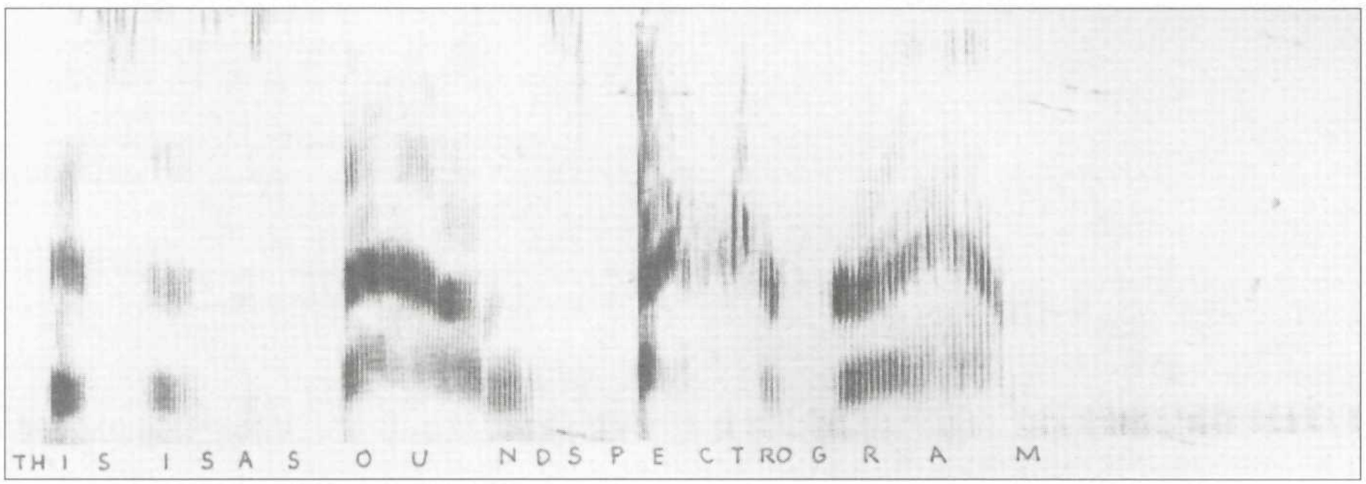
Marine Organisms.
Figure 9B, Roar.



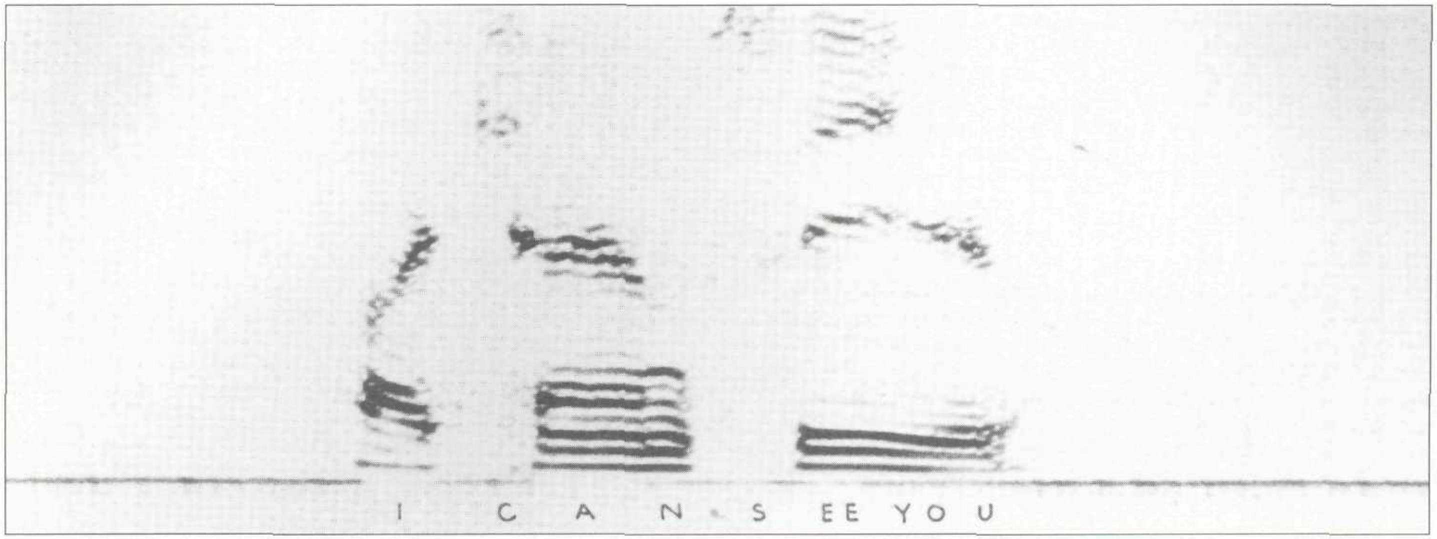
Marine Organisms.
Figure 11A, Rasping of Panulirus argus.



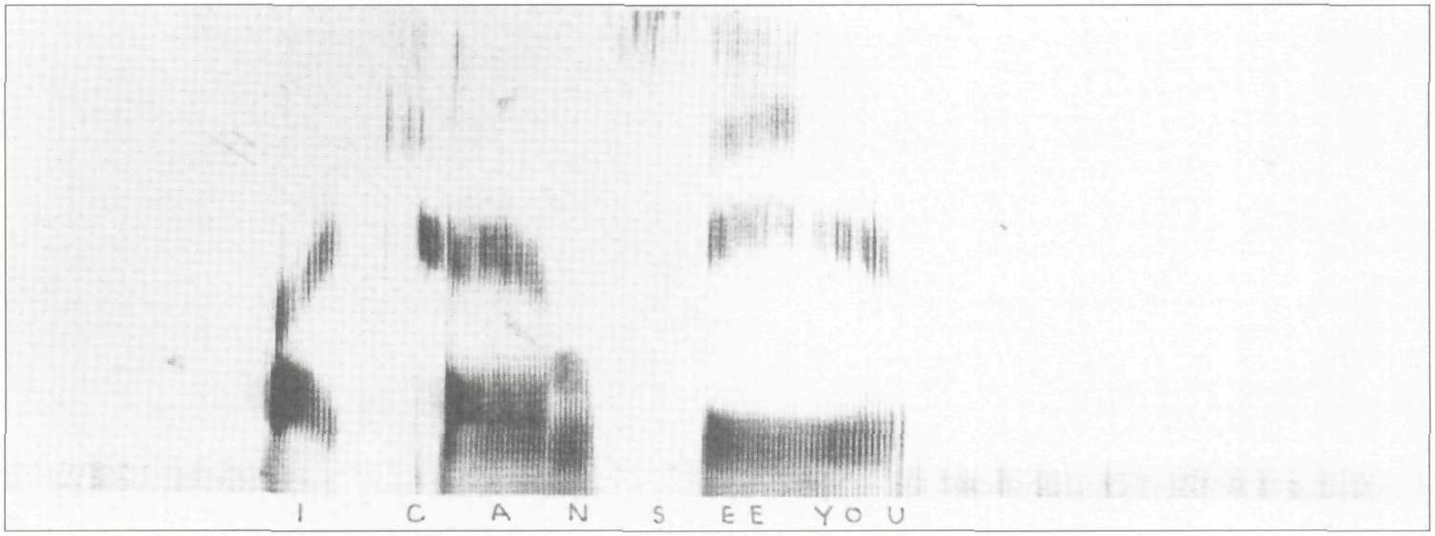
Marine Organisms.
Figure 12A, Crunch of Parrotfish.



"This is a sound spectrogram." *On the Theory and Practice of Voice Identification*, page 5,
Figure 1, "Three Samples of Speech Sound Spectrograms," third sample: "Speaker B: Day 1."



"I can see you," *On the Theory and Practice of Voice Identification*, page 82,
Figure B-1, "Voicegram displaying a narrow band analysis."



"I can see you," *On the Theory and Practice of Voice Identification*, page 82,
Figure B-2, "Voicegram displaying a wide band analysis."

the original study, of course, the crucial question is whether there's enough information present to distinguish the voices of different speakers. After listening, do you think there is?

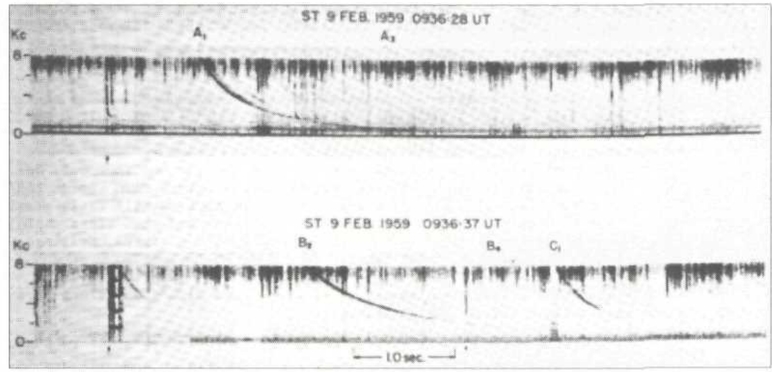
Either way, the fact that we can recover *intelligible* speech from images like these demonstrates vividly that the sound spectrogram is a legitimate sound recording format that contains the same kind of information as other "sound recordings." In the fall of 2008, I coined the term *paleospectrophony* ("old-spectrum-sounding") to refer to the use of reverse Fourier transform software to "play" old inscriptions of sound that can be read as graphs of frequency against time, as these can. The software had been around for years at that point, but to the best of my knowledge nobody else had ever tried to apply it to *historic* sound inscriptions before—so until someone tells me otherwise, I'll feel I'm entitled to give the technique its name.

Over the past seventy years, automatically generated sound spectrograms have been used to illustrate many types of sound apart from human speech. Our next example comes from a spectrographic study of the calls of bald eagles recorded in northern California during 1979 and 1980; the goal was to see whether spectrograms like these could be used to track individual eagles for wildlife management purposes as an alternative to the more intrusive practice of capturing and marking them. The authors summarized:

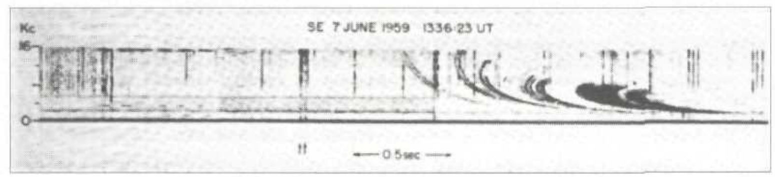
> The study sought to determine whether individual distinctiveness among calls can be used to investigate aspects of the eagle's life history, including seasonal or yearly dispersal between nest sites or territories, long-term survivorship, and perhaps the source of birds aggregating in winter roosts. Although preliminary results are inconclusive, they suggest that the technique of using sound spectrographic analysis of calls to identify individual birds shows promise.[18]

For our purposes, what's important here is that when we play back the spectrograms from the publication describing the study, we hear something that sounds like the calls of bald eagles.

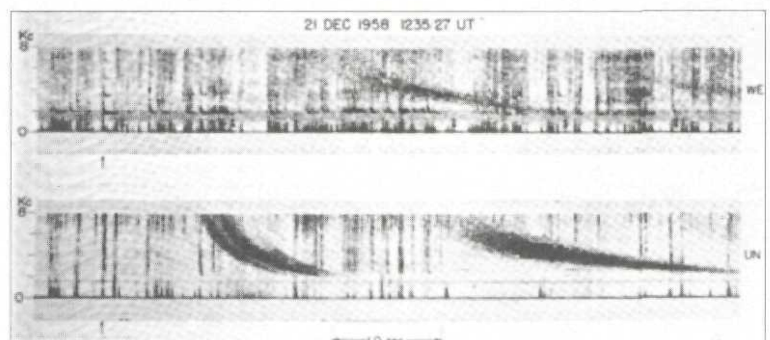
The two remaining examples probe what I take to be the outer limits of automatic sound spectrography. First comes a sequence of brief recordings of marine organisms made in 1970 in the Tektite habitat, an innovative



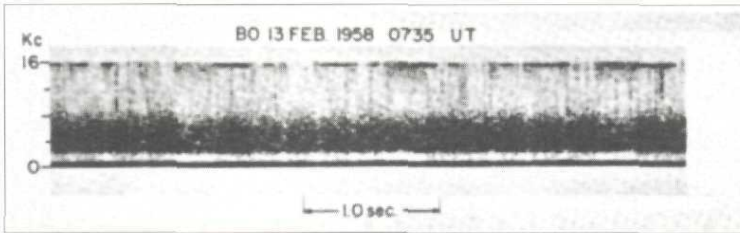
Whistlers. *IGY General Report Number 21*, November 1965, page 281, Figure 3, Short and Long Whistlers [*two strips*].



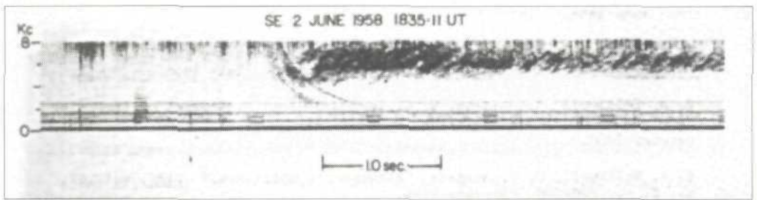
Whistlers. *IGY General Report Number 21*, November 1965, page 281, Figure 4, Nose Whistler.



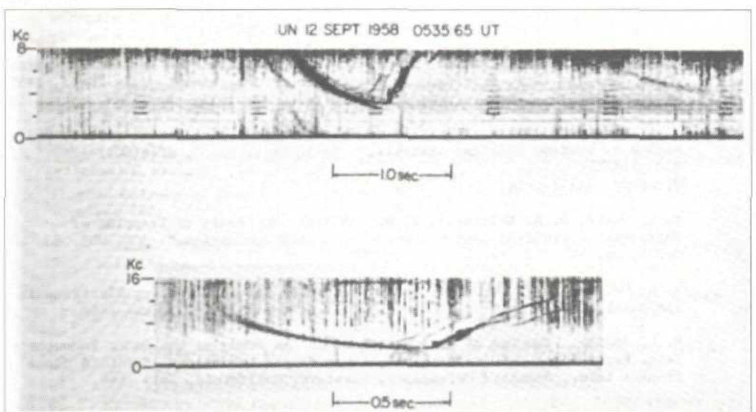
Whistlers. *IGY General Report Number 21*, November 1965, page 281, Figure 5, Whistler Echo Train [*two strips*].



Whistlers. *IGY General Report Number 21*, November 1965, page 283, Figure 6, Hiss.



Whistlers. *IGY General Report Number 21*, November 1965, page 283, Figure 7, Dawn Chorus.



Whistlers. *IGY General Report Number 21*, November 1965, page 283, Figure 8, Discrete Events [*two strips*].

underwater laboratory in the Virgin Islands—so these are representations of sounds passing not through air, but through water. And finally we have a series of spectrograms made in the 1950s of "whistlers" and other natural radio waves, generated by lightning strikes and by other processes less well understood. In their original form, these phenomena are electromagnetic waves, and not sound waves at all. However, they happen to fall within audible frequency ranges, and they're frequently picked up and translated into sound by radio receivers, which makes them fair game for audio recording equipment.

All of this goes to show that any "sound recording," no matter how abstruse its subject matter, can be displayed as a spectrogram; and that the spectrogram can, in turn, be converted back into sound.

a. "Theory and practice of voice identification (4×). This is a sound spectrogram." Source: *On the Theory and Practice of Voice Identification* (Washington DC: National Academy of Sciences, 1979), front and back covers and page 4 (Figure 1, "Three Samples of Speech Sound Spectrograms," first sample: "Speaker A: Day 1").

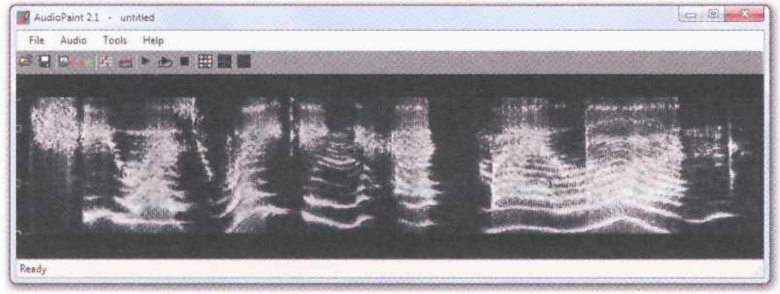
b. Bald Eagles. Source: Jared Verner and Robert N. Lehman, *Identifying Individual Bald Eagles with Voiceprints: a feasibility study*, United States Department of Agriculture Research Note PSW-359 (December 1982).
i. "Figure 1—Sonagrams of 'wail' calls of two bald eagles."
ii. "Figure 2—Sonagrams of 'peal' calls of two bald eagles." "Figure 3—Sonagrams of 'chatter' calls of six bald eagles."
iii. "Figure 4—Sonagrams of four different chatter calls of one bald eagle in the same year. (*A*) May 19, 1980, recordings; (*B*) June 24, 1980, recordings." "Figure 5—Sonagrams of four different chatter calls, presumably of the same bird in two consecutive years. (*A*) May 19, 1979, recordings; (*B*) June 14, 1980, recordings."

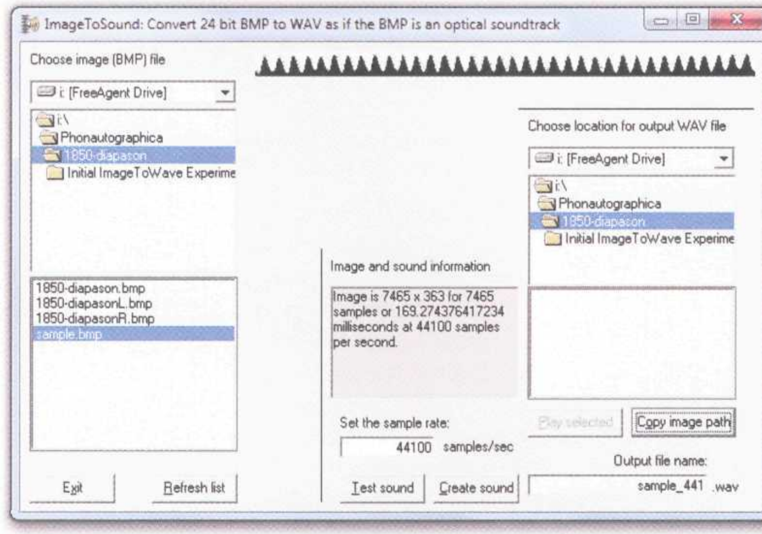
c. "This is a sound spectrogram." Source: *On the Theory and Practice of Voice Identification* (Washington DC: National Academy of Sciences, 1979), page 4 (Figure 1, "Three Samples of Speech Sound Spectrograms," second sample: "Speaker A: Day 4").

d. Marine Organisms. Source: "Passive and Experimental Bio-Acoustical Studies on Marine Organisms in Their Natural Habitat," Thomas J. Bright and William W. Schroeder, in *Tektite 2: Scientists in the Sea*, ed. J. W. Miller, J. G. VanDerwalker, and R. A. Waller (Washington DC: U. S. Department of the Interior, 1971), 82-105.
i. Figure 3B, Staccato of Holocentrus ascensionis
ii. Figure 6, Quacks of Flammeo marianus
iii. Figure 7A, Quacks of Flammeo marianus
iv. Figure 8B, Squeaking door sound
v. Figure 9B, Roar
vi. Figure 11A, Rasping of Panulirus argus
vii. Figure 12A, Crunch of Parrotfish

e. "This is a sound spectrogram. I can see you. I can see you." Source: *On the Theory and Practice of Voice Identification* (Washington DC: National Academy of Sciences, 1979), page 5 (Figure 1, "Three Samples of Speech Sound Spectrograms," third sample: "Speaker B: Day 1"); and page 82 (Figure B-1, "Voicegram displaying a narrow band analysis"; Figure B-2, "Voicegram displaying a wide band analysis").

f. Whistlers. Source: IGY General Report Number 21, November 1965: *Report on the U. S. Program for the International Geophysical Year, July 1, 1957—December 31, 1958* (Washington DC: National Academy of Sciences—National Research Council, November 1965), pp. 281, 283.
i. Figure 3, Short and Long Whistlers [*two strips*]
ii. Figure 4, Nose Whistler
iii. Figure 5, Whistler Echo Train [*two strips*]
iv. Figure 6, Hiss
v. Figure 7, Dawn Chorus
vi. Figure 8, Discrete Events [*two strips*]





Modern software applied to historic pictures of sound.

(((((((3. SCRIPTURE'S CURVES   )))))))

(1:25)

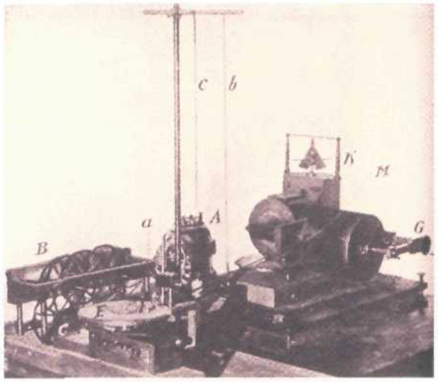
Before the invention of the automatic sound spectrograph, it was more challenging for researchers to study recorded sounds visually than it became afterwards—but that doesn't mean nobody tried. Edward Wheeler Scripture was a pioneer in the study of what he called "speech curves": wavy lines that document the rapid back-and-forth motions of membranes over time under the influence of aerial sound waves. Whereas sound *spectrograms* are graphs of time versus *frequency*, Scripture's "curves" are sound *oscillograms*, graphs of time versus *amplitude*—meaning, in this case, a membrane's degree of displacement from its rest position. Other examples of such "curves" include the grooves of LPs and the waveforms displayed by modern sound-editing software.

Scripture had no doubt that "speech curves" documented their subject matter richly and accurately. "The record itself is at least as good as the sound which it can be made to give," he wrote, meaning that the proof of its accuracy lay in aural reproduction: whatever details you could *hear* by playing a gramophone disc were necessarily also *there* in the groove. The challenge lay in making sense of them. Scripture admitted that "speech curves" were difficult to read, characterizing them as "at first sight no more intelligible than a line of Chinese ideographs"—and he doesn't seem to have been writing for a Chinese audience. With all this in mind, he expressed the ultimate goal of his research as follows:

> The entire intellectual and emotional impression conveyed by the voice from the speaker to the hearer is contained in the speech vibration and registered in the speech curve. Hardly any problem of greater interest could be proposed than that of discovering the manner of getting from a voice curve the data concerning the action of the vocal organs in such an exact and minute form that conclusions can be drawn concerning the variations in the voice as depending on every emotion, on every condition of health, on every step in voice culture, on every difference in vowels and consonants, on each change in dialect, etc.[19]

Scripture did his best-known work using material that was available at the time on commercially manufactured gramophone discs (the direct ancestors of the vinyl LP). The "curves" on the discs themselves were too tiny to be easily studied, and they were also coiled inconveniently into spirals, but Scripture rigged up some equipment to copy them mechanically from the discs onto sheets of paper as straight lines and at much larger size. It was in this form that he studied them, and it was also in this form that he published some of them in print to illustrate his findings.

So how well did Scripture's published "curves" actually represent the sounds he was seeking to investigate? As we've seen, he himself believed that the true test of a record's accuracy lay in its aural reproduction, and if we want to test his "curves" in this way, all we need to do is convert the printed waveform shapes into a playable digital form. So that's what I've done here, though I've had to use a more roundabout method than I did with the spectrograms in track two (see description on right).



One of the instruments Scripture used for transferring waveforms from gramophone discs onto paper.

Our first two specimens of Scripture's "curves" are among his longer published examples, even though they last only a few seconds each. One is taken from Berliner 6015, "The Sad Story of the Death and the Burial of Poor Cock Robin," recited by William F. Hooley,[20] and contains the words: "With my little eye I saw him die. Who caught his blood? I, said the fish...." The other is an excerpt from Berliner 693, a "Speech on Forefathers' Day" delivered by the well-known politician and after-dinner speaker Chauncey Depew: "without regard to race or creed, I can...." In order to save space, Scripture cut some silent passages from these two traces and replaced them with bracketed notations of how many millimeters he'd removed. These omissions stand out when we play the traces back today because the surface noise that's present elsewhere drops out, leaving conspicuous silences. Apart from that, the two recordings sound pretty respectable for the late 1890s. As was the case with the speech spectrograms in the previous track, we can definitely recover intelligible spoken language from images like these: the wavy lines on the printed page are indeed "sound recordings," just as surely as were the original gramophone discs from which Scripture copied them. The examples that follow are much shorter, and more typical of early pub-

WAVEFORM IMAGE TO WAV FILE:

HOW IT'S DONE

To the best of my knowledge, there isn't any software readily available to the public for translating waveform images directly into equivalent sound files. However, I've found a workaround using ImageToSound by Stopmotion Software, a program that converts images into sound files as though they were optical motion picture sound tracks.[21]

Optical sound tracks on motion picture film consist of long, narrow bands that vary either in width or in opacity; when they're moved past a light source, they obstruct it to varying degrees so that the fluctuating intensity of light can be transduced into an audio signal. Back in the 1920s, experimental artists sometimes painted optical film sound tracks by hand as a means of generating novel synthetic sounds, and ImageToSound was designed to enable people to do the same thing in the digital realm—basically, it calculates the average brightness of successive columns of pixels in a digital image and assigns the values to successive samples in a digital sound file. Like AudioPaint, this program was created with experimental sound art in mind, but it also lends itself handily to the playback of recorded sounds.

Of course, Scripture's "curves" aren't bands of varying width—they're wavy lines. However, they can be converted into bands of varying width using commonly available graphics editing software, in three steps:

1. Isolate each wavy line in a long, narrow, rectangular strip, adjusted as needed to make it a bright trace against a dark background.
2. Check for any breaks or flaws in the line and repair them (this is the time-consuming part).
3. Create two new images using a paintbucket tool to fill the area *above* the line and the area *below* the line respectively with white. This gives us a pair of white bands of varying width against black backgrounds.

Run the two "filled" images through ImageToSound, place the resulting sound files side by side in stereo, reduce that to mono, and voilà—we have a digital sound file that displays the same "shape" as the original wavy line in Scripture's book. I've used some version of this same technique to play all the waveform images included in *Pictures of Sound*.

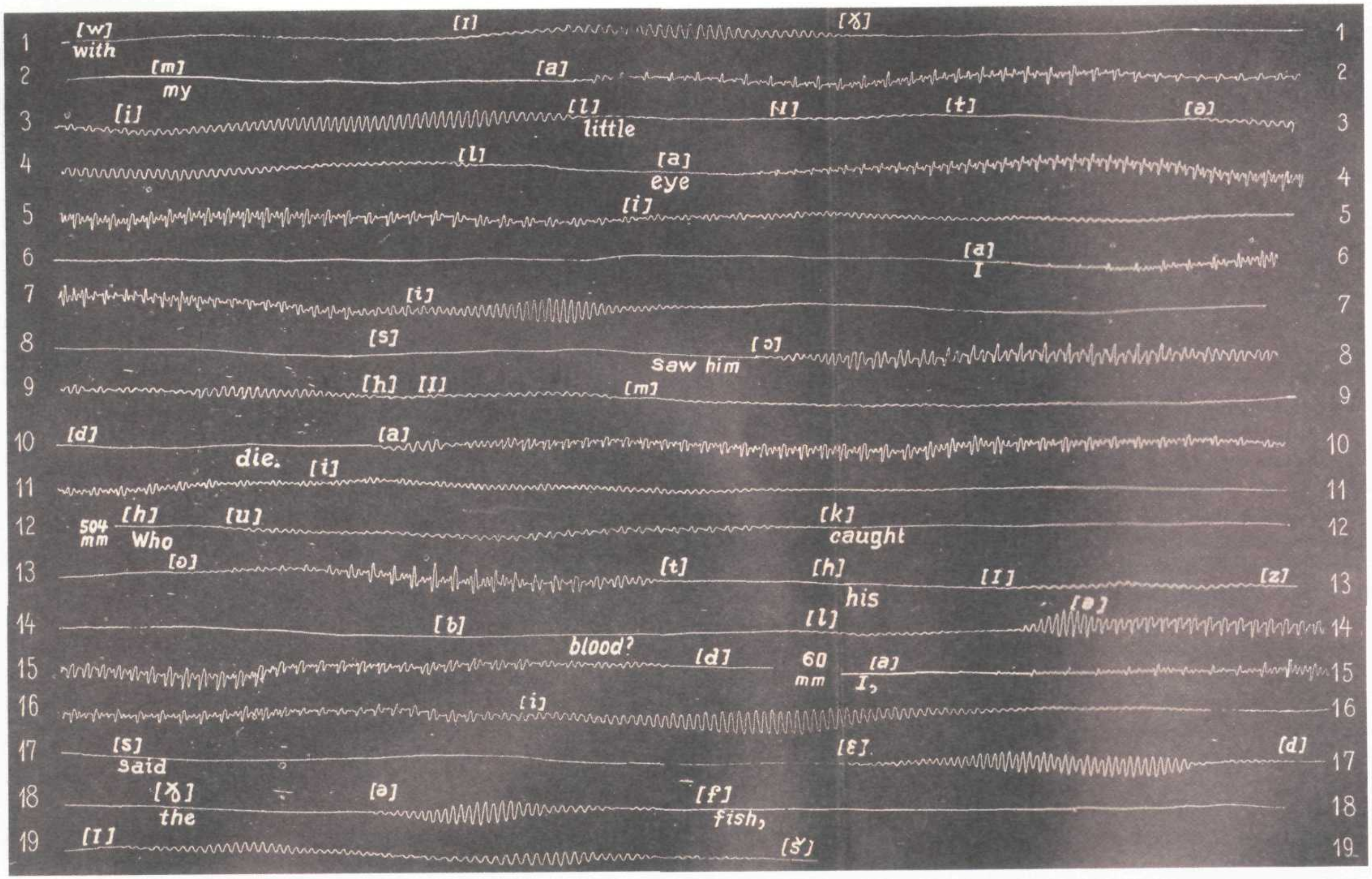


Plate IX. Curves from "Cock Robin."—"With my little eye I saw him die. Who caught his blood? I, said the fish...."
E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves.*

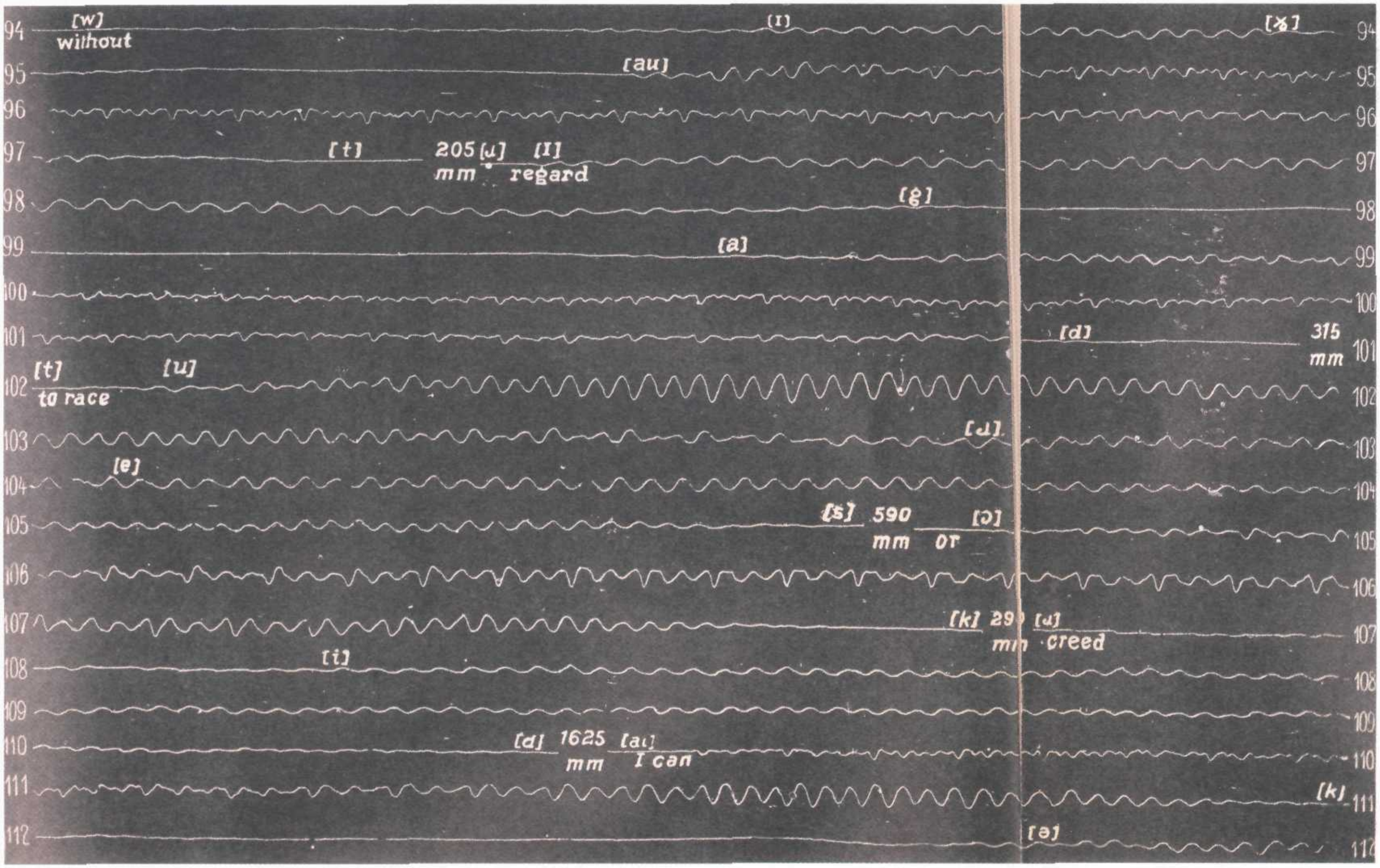
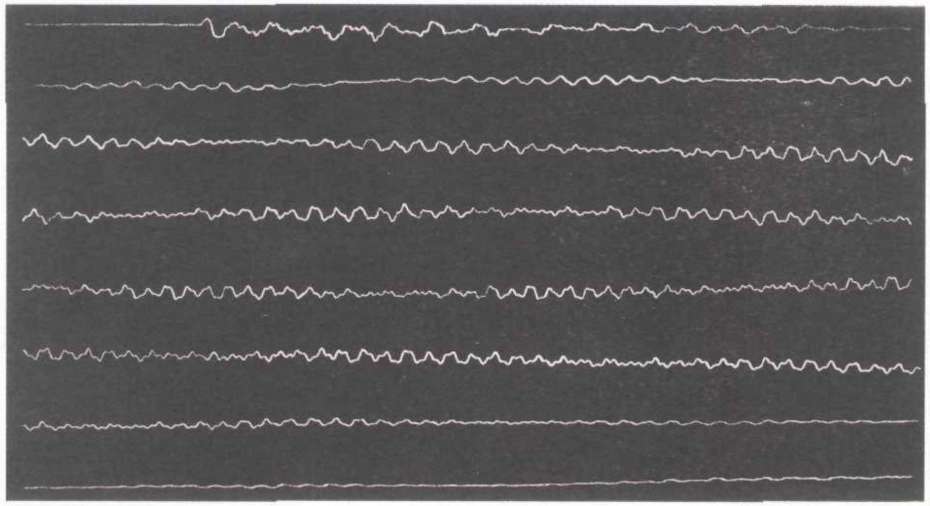


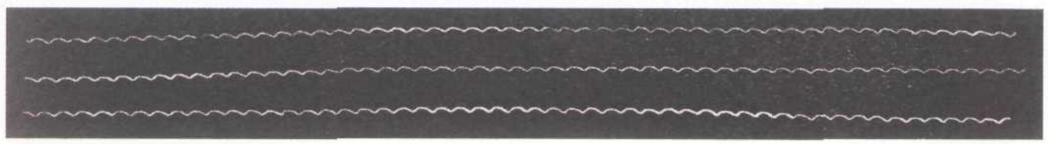
Plate VIII. Portion of "Speech on Forefathers' Day," by Chauncey M. Depew.—"Without regard to race or creed, I can...."
E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves.*

Fig. 32—Orchestra. "The curve in figure 32 is from the record of a note from an orchestra."
E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves.*

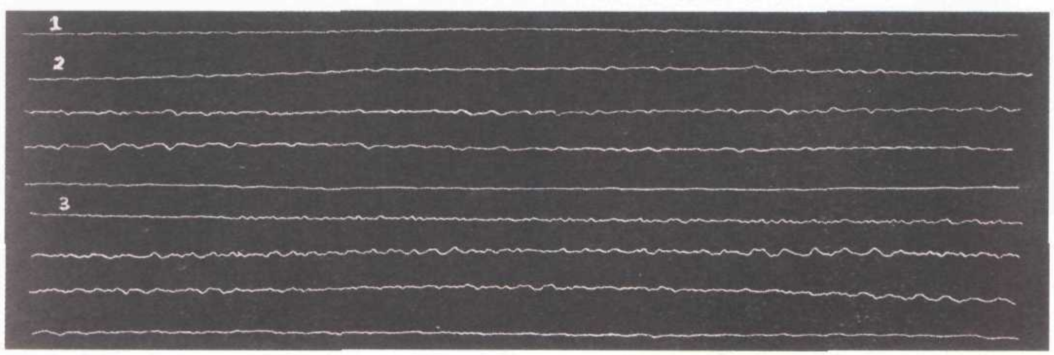


33.

34.

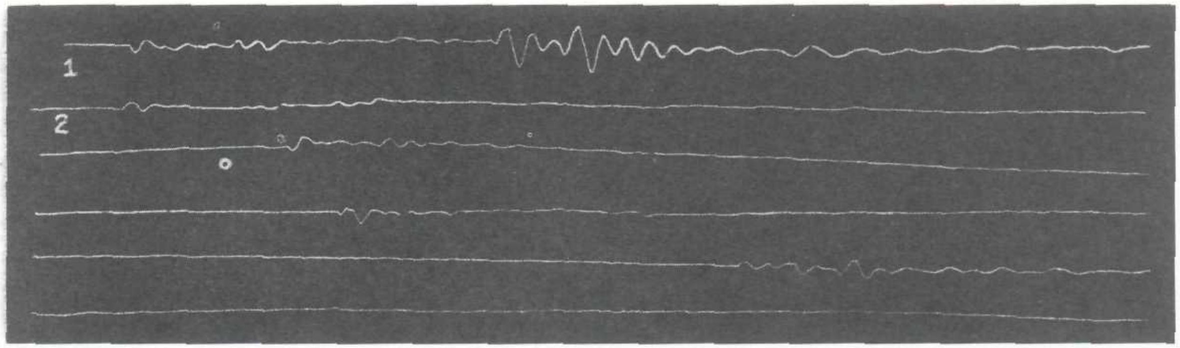


35.



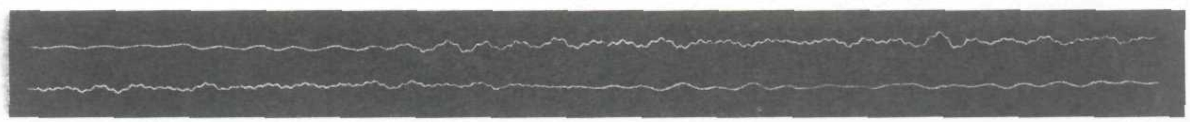
36.

Fig. 33—Gong. **Fig. 34**—Conductor's whistle. **Fig. 35**—Locomotive whistle. **Fig. 36**—Three locomotive puffs.
E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves.*



37.

38.



39.

40.

Fig. 37—Blocks [two separate segments]. **Fig.** 38—Whistling. **Fig.** 39—Whistling with piano. **Fig.** 40—Plucked string.
E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves.*

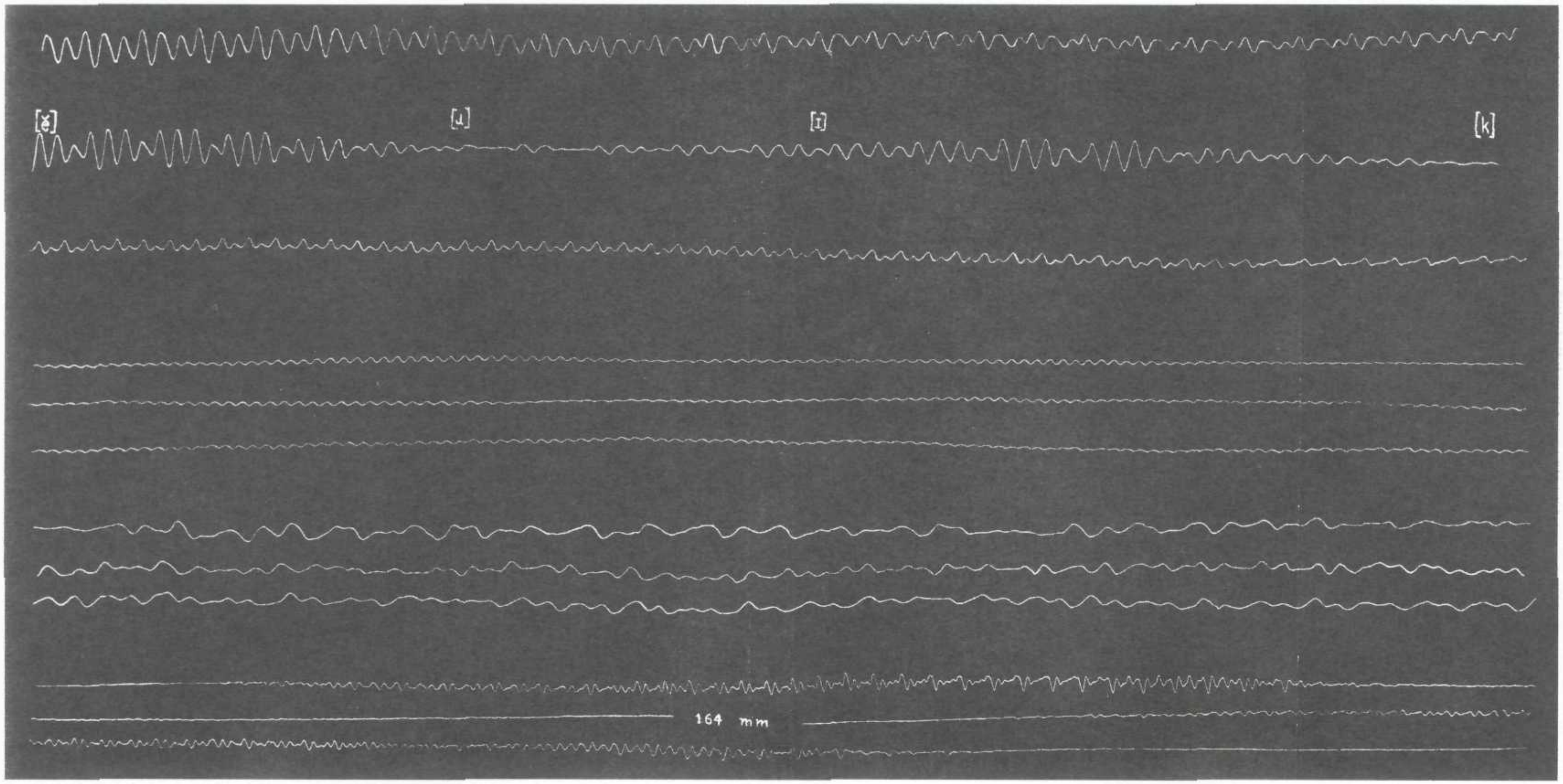


Plate VII. Part of curve of [o]. Part of curve of "America." Part of a Chinese vowel. Part of a trill (tremolo). Part of a chord from a piano. Laughter.
E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves.*

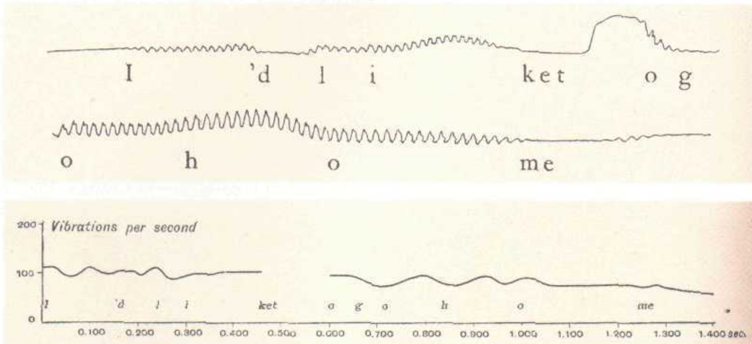
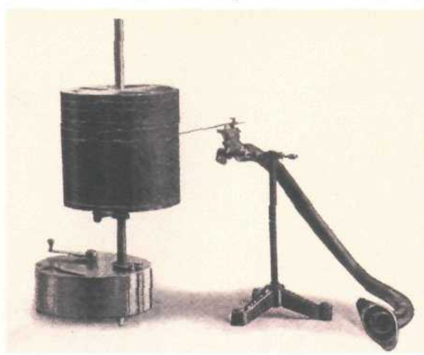


Fig. 5—Record of "I'd like to go home," by a normal voice."
and **Fig.** 6—Melody plot to Fig. 5.

lished waveform images—mere split-second snippets of "curve," intended only to show distinctive wave shapes. Because they're so short, I've played each one three times in rapid succession to help your ear get a grip on it. Scripture identifies his subject matter in each case, but you might find it interesting to listen first and to try to guess what the various sounds are before you look at the captions. You'll probably be surprised by some of them.

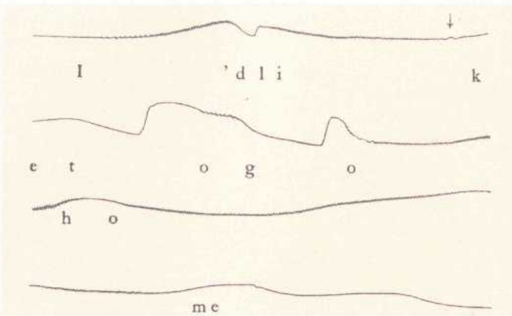
All the "curves" we've considered here so far come from Scripture's well-known book *Researches in Experimental Phonetics* (1906), but our final examples are instead excerpted from a scholarly article published ten years later, in 1916—a study of the speech of patients with multiple sclerosis. This time, Scripture wasn't working with preexisting recordings copied from gramophone discs; instead, he'd made his own speech recordings using an instrument called the *phonautograph*, designed specifically to record "curves" on paper for visual analysis. In both of the recordings I've presented here, a speaker says "I'd like to go home"—probably a standardized test phrase Scripture had chosen. But Scripture didn't just give us the "curves" in their original form. He also measured them to calculate how the fundamental pitch of the speaker's voice had varied over time, and he displayed this variation in turn by means of hand-drawn "melody plots," or graphs of frequency versus time.



The phonautograph Scripture used to trace "speech curves" for his research into the speech of patients with multiple sclerosis.

Several decades later, the developers of the automatic sound spectrograph would use a striking analogy to contrast waveform and spectrographic displays of aural data: "This latter pattern [the spectrogram] is analogous to the pattern of a rug, spread out so that it is clearly visible. The former [the 'curve' or waveform] is analogous to seeing all the threads of a rug unravelled and bundled together. The pattern material is there, but in no shape to be visualized as a meaningful whole."[22] By creating "melody plots," Scripture was making a similar effort to reveal patterns in his data that weren't apparent when it was displayed in the form of raw "curves." Of course, his "melody plots" aren't as detailed as later sound spectrograms: they're what we'd now call *pitch contours*, showing only the fundamental with none of the overtones. However, sound spectrograms and pitch contours are both graphs of frequency versus time, and because of that similarity, we can convert Scripture's "melody plots" back into sound using the same technique we used to play back the spectrograms in the previous track—the technique I'm calling paleospectrophony. In other words, when it comes to playback, hand-drawn "melody plots" like these are interchangeable with automatically generated sound spectrograms. We can play and hear them both in the same way, according to the same logic of actualization, regardless of the difference in how they came into being.

So at the end of this track, you'll hear both recorded specimens of "I'd like to go home" not once but twice, played first from the "speech curve," and then from the "melody plot." The two versions of each recording don't sound identical; after all, the translation from "speech curve" to "melody plot" entailed a considerable loss of information. But they're certainly comparable—you can hear that the one is based on the other, even if only imperfectly.



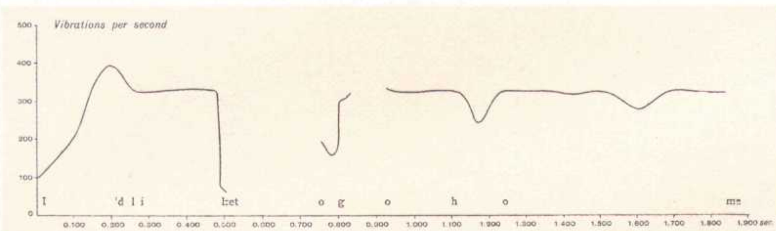


Fig. 17—Record of "I'd like to go home," by E. A. E.
and **Fig. 18**—Melody plot to Fig. 17.

a. Source: E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves* (Washington DC: Carnegie Institution of Washington, 1906).
i. "Plate IX. Curves from "Cock Robin."—"With my little eye I saw him die. Who caught his blood? I, said the fish...."
ii. "Plate VIII. Portion of 'Speech on Forefathers' Day, ' by Chauncey M. Depew."—"Without regard to race or creed, I can...."
iii. Plate facing page 32. "Fig. 32.—Orchestra." "The curve in figure 32 is from the record of a note from an orchestra" (33).
iv. Plate facing page 33. "Fig. 33.—Gong. Fig. 34.—Conductor's whistle. Fig. 35.—Locomotive whistle. Fig. 36.—Three locomotive puffs." "The record of the gong in figure 33 shows the effect of a blow. It begins with a sudden upward deviation due to the blow of the hammer and an almost complete vibration of smooth form before the gong begins its special vibrations.... The record of the conductor's whistle in figure 34 shows the very short vibrations for the high note and also the pseudobeats (periodic weakenings of the sound) that are produced by the rattling of the little ball in the whistle.... Figure 35 gives a small piece out of the record of a European locomotive whistle.... The first of the three locomotive puffs in figure 36 was soft and long.... The second puff was louder, hollower, and of lower pitch. The third was still louder, but again of higher pitch" (33-34).
v. Plate facing page 34. "Fig. 37.—Blocks [two separate segments]. Fig. 38.—Whistling. Fig. 39.—Whistling with piano. Fig. 40.—Plucked string." "The first line of figure 37 gives the curve produced by striking two blocks in rapid succession to imitate the gallop of a horse.... The last four lines of figure 37 give four blows of the same blocks in rapid succession, with increasing intensity, the last blow being specially emphasized.... The whistling curve in figure 38 is from a professional whistler.... Figure 39 gives a note whistled to a piano accompaniment; they eye readily selects the portions where the piano vibrations occur alone and those where the high note of the whistling is imposed upon them.... Figure 40 gives the vibrations produced by a plucked string of a Chinese musical instrument; the pitch of the string is high" (34). Note: the "plucked string" trace doesn't include the attack and so ends up sounding rather like a whistle.
vi. "Plate VII. Part of curve of [o]. Part of curve of 'America.' Part of a Chinese vowel. Part of a trill (tremolo). Part of a chord from a piano. Laughter." "The first two specimens in plate VII are from a record by S. Weir Mitchell. The former gives one-fifth of the exclamation 'oh' spoken sorrowfully..... The second specimen gives the last half of [ĕ] (not [ɛ]) followed by [ɹɪ] from the word 'America'.... The third specimen on the plate is from a Chinese vowel.... The next curve on the plate is a small portion of a trill by an Italian voice (Mme. Chalia) on a high note.... The next specimen on the plate is the curve of the first half of a piano chord with one note more prominent than the others.... The curves of laughter give the last two of a set of laughs which may be indicated by 'Ah, hah, hah, hah'" (35-36). Note: the first three speech curves sound deceptively like plucked strings, and I think I recognize the laughter of the well-known phonograph performer Billy Golden in the final example.

b. Source: E. W. Scripture, "Records of Speech in Disseminated Sclerosis," *Brain: A Journal of Neurology* 39 (1916): 455-477.
i. "Fig. 5.—Record of 'I'd like to go home,' by a normal voice." "Fig. 6.—Melody plot to fig. 5." (p. 458)
ii. "Fig. 17.—Record of 'I'd like to go home,' by E. A. E." (p. 470) "Fig. 18.—Melody plot to fig. 17." (p. 471)

(((((((4. PHONOPHOTOGRAPHY)))))))

(1:22)

Edward Wheeler Scripture wasn't the only investigator who measured the "curves" traced by sound recording equipment and then drew graphs of his measurements using a spectrogram-like format. In the 1920s, Milton Metfessel developed a technique called *phonophotography*, which entailed recording sound as wavy lines on moving strips of photographic film, measuring the waves by hand, and then displaying their frequencies graphically in "pattern notation," with conventional musical notes thrown in for good measure. Metfessel was affiliated with the psychological laboratory headed by Carl Seashore at the University of Iowa, and he wanted to depict the human singing voice with heightened detail and accuracy specifically as a means of studying emotional expression. "The deviation from the exact is, on the whole, the medium for the creation of the beautiful—for the conveying of emotion," Seashore claimed in the introduction to Metfessel's book *Phonophotography in Folk Music* (1928). "The exact is cold, restricted, and unemotional, and, however beautiful in itself, soon palls upon us."[23] From this perspective, tiny deviations in pitch and timing—such as occur in vibrato—were crucially important for researchers to detect and analyze if they ever hoped to understand in a scientifically rigorous way what it was that made musical performances artistic. And Metfessel's phonophotographic inscriptions of singing certainly captured plenty of vibrato, representing this graphically as wavy lines instead of straight ones. "This study," he boasted, "describes music as it has never been described before."[24]

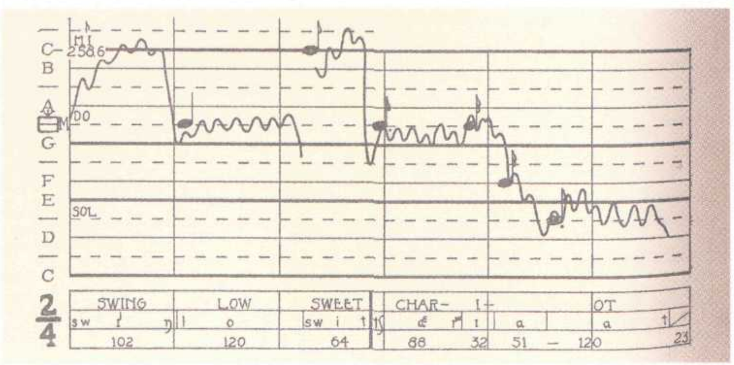
Metfessel gave his book the subtitle *American Negro Songs in New Notation*, and most of his examples document singing by African American performers in the southern United States. He claimed that he had chosen this particular culture group for study only because it was "expedient"; the real goal of the work was to illustrate the value of phonophotography for the investigation of so-called "primitive music" more broadly. Even so, Metfessel made an effort to justify the technique's application to African American vocal music in particular as follows:

> Much of the charm and distinctiveness of the singing of Negroes lies in queer pranks of their voices, but these twists and turns occur too quickly. One seeking to analyze them or find out anything about them is only bewildered. As a consequence, valuable detail and necessary accuracy have been lacking in studies of folk music.... [But] by using the phonophotographic technique, the word "unnotatable" no longer need be used with reference to music or speech. It has been possible to notate *all* the twists, quavers, trills, breaks in the voice, quick slurs, erratic tempi, and other similar features so often a part of folk singing.[25]

Most of the transcriptions found in Metfessel's book represent solo voices, including our second example, "I Heard the Voice of Jesus Say." Metfessel described this as "a good illustration of uncultured Negro singing" by a "man of low intelligence but rather high emotional responsiveness," adding: "The ornaments of the song are excessive, but fit into the general emotional intoxication of the singer."[26] There's no mistaking Metfessel's condescending attitude towards his subject, but at the same time this passage exemplifies the kind of conclusion he was most eager to draw, linking the heavy use of "ornaments" to emotional excess. Apparently, greater emotion revealed itself in Metfessel's notation as greater wiggliness.

The same book also includes two of Metfessel's attempts to transcribe polyphonic singing. Our track opens with one of them—"Swing Low, Sweet Chariot," sung by the Hampton Quartet—and closes with another: "Little David," sung by the quartet of the North Carolina State College for Negroes in Durham, North Carolina. The nature of Metfessel's equipment prevented him from actually recording all four voices singing simultaneously in either case, so he had to use a makeshift approach which he described as follows:

> The individual voices of the quartet...were recorded one at a time, and a separate analysis made for each voice. The graph-curves were then brought together on one staff. The singers went through their song four times, each member of the group stepping up to the mouthpiece in his turn. The three not being recorded were a short distance away, their singing being sufficiently loud to give the singer his bearings but not strong enough to affect the camera... Relatively,



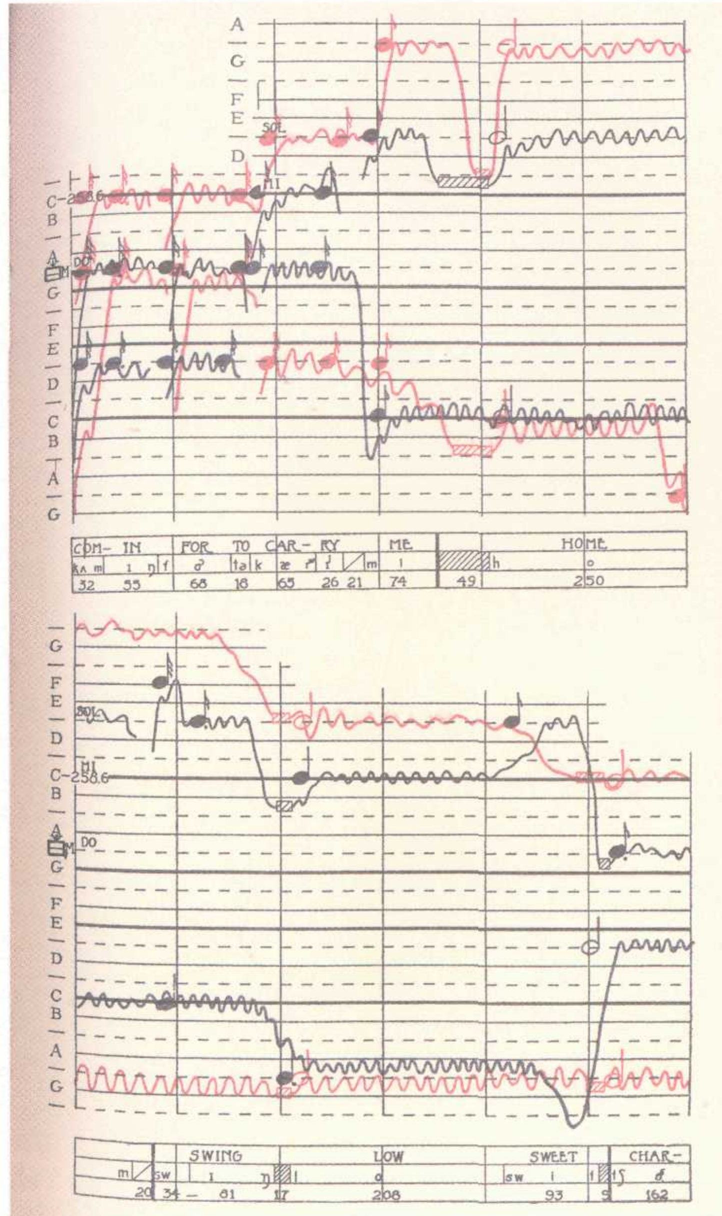


Fig. 29—"Swing Low, Sweet Chariot," Graphs 1, 2, 3, *Phonophotography in Folk Music: American Negro Songs in New Notation*, pages 82-83.

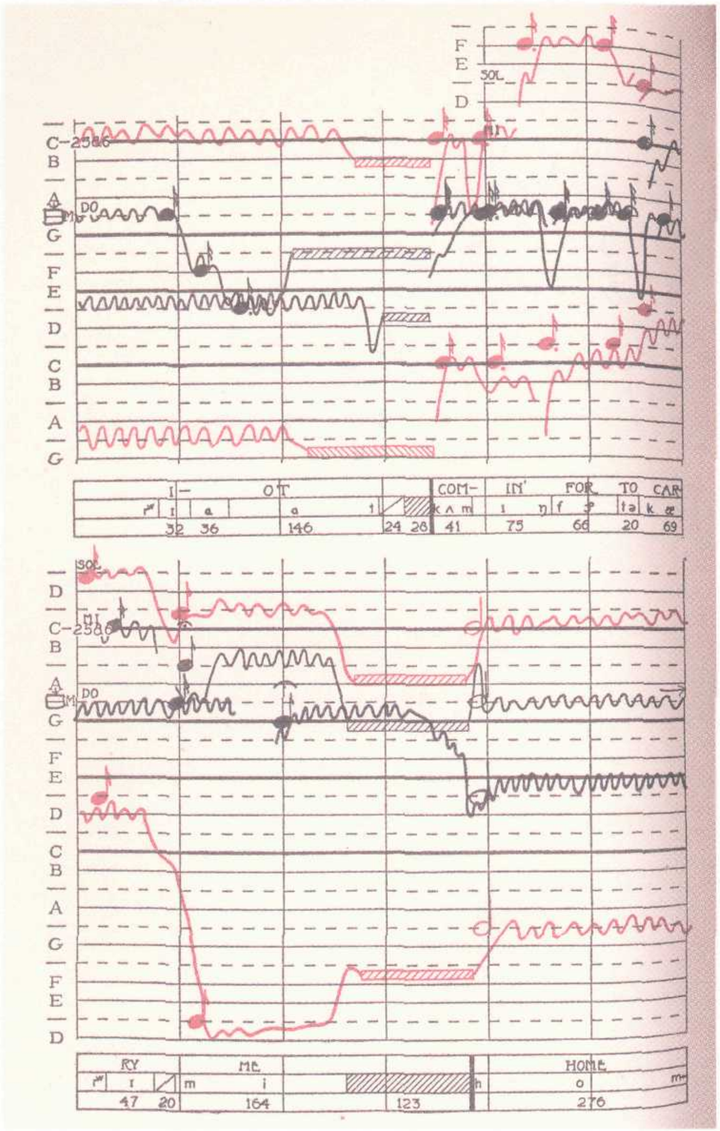


Fig. 29—"Swing Low, Sweet Chariot," Graphs 4, 5,
Phonophotography in Folk Music: American Negro Songs in New Notation, page 84.

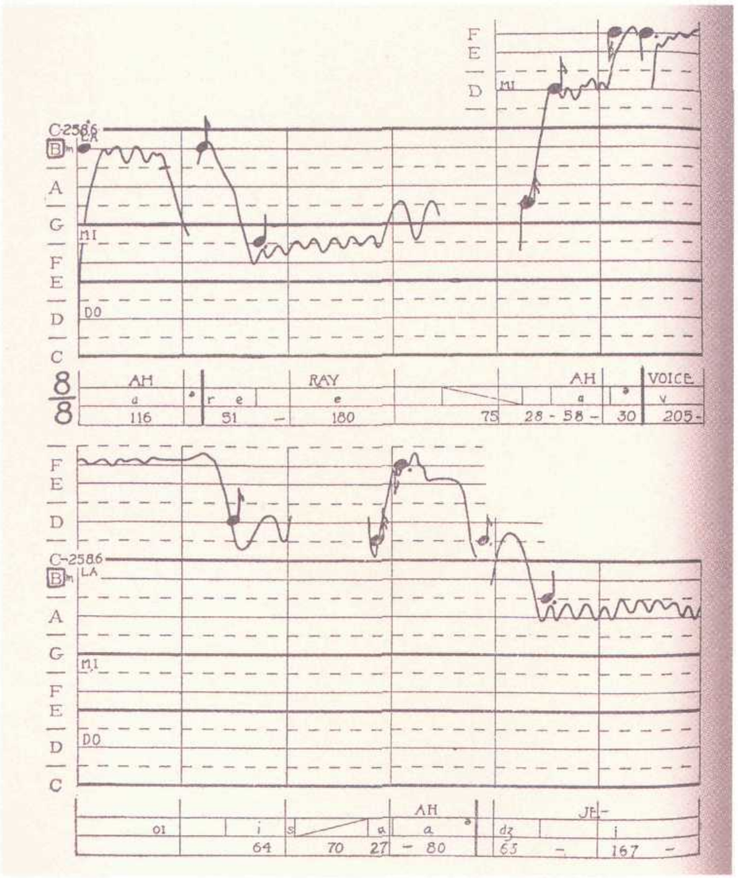
> the frequency representations of the graph-curves were not tampered with, and changes in duration are signified on the notation by obliquely-ruled rectangles. The actual singing of each part is therefore left intact in keeping the four together.... Blue, red, and black inks are used to keep the voice lines of each part distinct.[27]

Metfessel's original recordings on film don't seem to survive anywhere today, which is an unfortunate loss. But we can still listen to his published "pattern notation" in much the same way we listened to Scripture's "melody plots"—by playing the inscriptions paleospectrophonically, just as though they were automatic sound spectrograms. We need to proceed a little differently this time around, though. Scripture's "melody plots" can be played as is, straight off the page, because they don't contain any extraneous marks. By contrast, Metfessel's "pattern notation" requires us to isolate the marks that graphically represent the sounds from other marks that serve other purposes: the conventional musical notes drawn in for the benefit of musicians and the solid and dotted guidelines that make up the background grid. I refer to the process of erasing such extraneous marks as *degridding*, since it most often involves the removal of a grid of some kind, and I've had to do this frequently in the tracks that follow. It's picky, time-consuming work, since even a tiny fleck of color that can barely be seen can still be heard, and despite my best efforts, I'm sure occasional glitches of this sort have made it through into the final tracks.

Once we've completed the degridding step, we can play Metfessel's transcriptions paleospectrophonically as soon as we've worked out how to set our two parameters:

1. **Time Scale.** Metfessel states that "a 'graph' consists of one staff of six seconds' duration, with the legend below it,"[28] so each of the six bar-like sections we see marked off in each staff presumably corresponds to one second of time.
2. **Frequency Scale.** The vertical axis is divided into evenly-spaced semitones, which correspond to a logarithmic scale (each octave up doubles the frequency). Metfessel also identifies his middle C as 258.6 Hz. By measuring the vertical distance between any two notes separated by an octave, we learn how much vertical space corresponds to a doubling of frequency—call this value y, measured in

pixels. Then we extend each image vertically just far enough beyond its existing boundaries so that the distances above and below middle C are both multiples of y. In this way, we end up with an image that has a height of $4y$, corresponding to a range of exactly four octaves, with middle C (258.6 Hz) in the exact center. To calculate the bottom frequency, we halve 258.6 Hz twice (64.65 Hz), and to calculate the top frequency, we double it twice (1034.4 Hz).



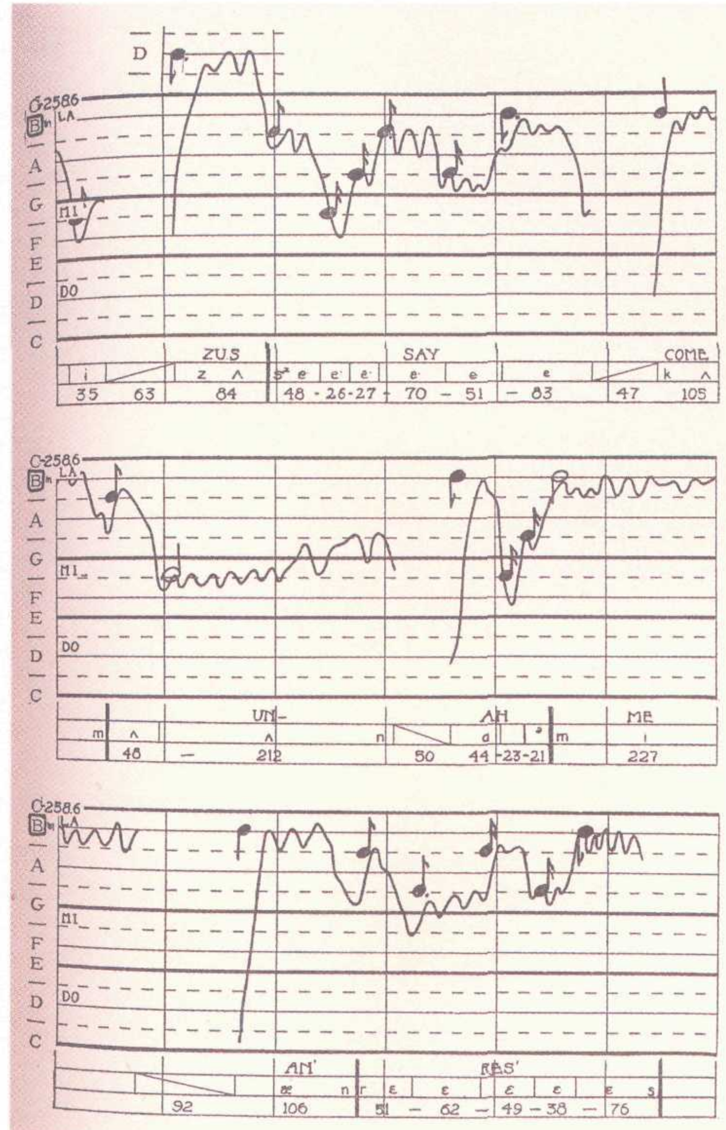


Fig. 10—"I Heard the Voice," Graphs 1, 2, 3, 4, 5,
Phonophotography in Folk Music: American Negro Songs in New Notation, pages 48-49.

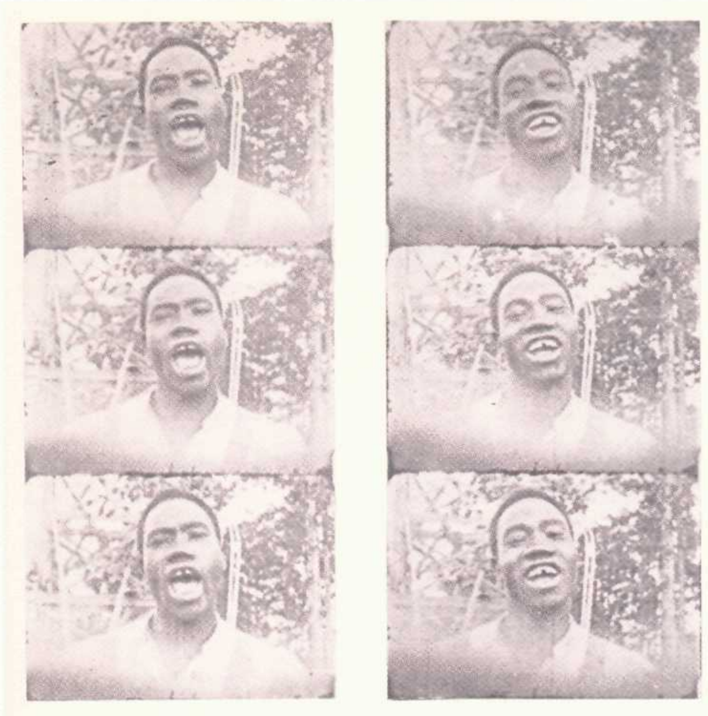


Fig. 11—"I Heard the Voice," page 50,
Phonophotography in Folk Music: American Negro Songs in New Notation.

I've made one concession to listenability here which you'll encounter from time to time on the CD: the left channel contains Metfessel's "pattern notation" set to the frequency range he specified, while the right channel is set one octave up from that range (i.e., 129.3 Hz to 2068.8 Hz). This stereophonic octave-doubling seems to make inscriptions like Metfessel's easier to follow and more pleasant to listen to, but purists are free to listen to the left channel alone by itself.

Metfessel's transcription of "I Heard the Voice of Jesus Say" represents a specific solo vocal performance that could theoretically be "reproduced," so we might be justified in claiming that we've "played it back" here by means of paleospectrophony. But the status of the other material heard in this

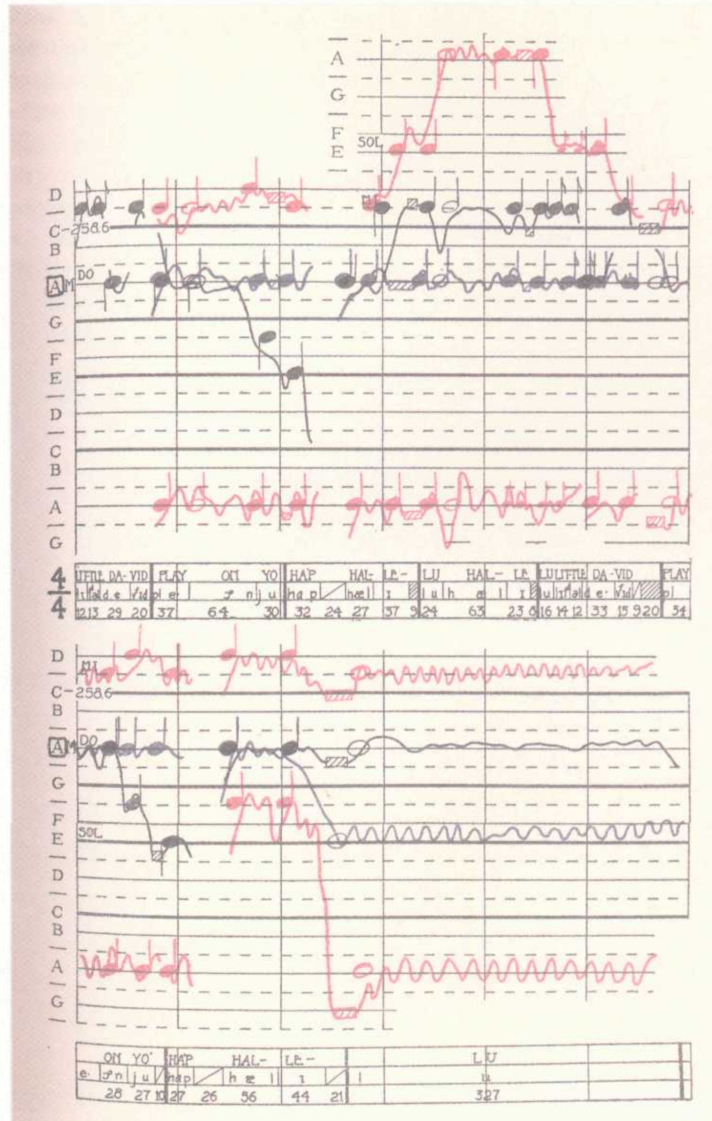


Fig. 27—"Little David," page 79,
Phonophotography in Folk Music: American Negro Songs in New Notation.

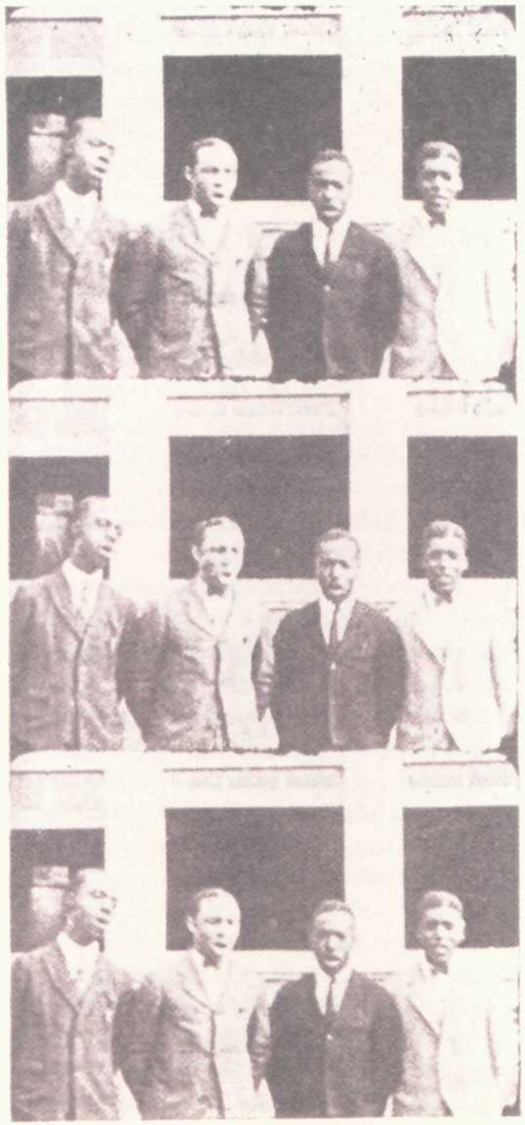


Fig. 28—"Little David," page 80,
Phonophotography in Folk Music: American Negro Songs in New Notation.

track is less clear-cut. As we've seen, Metfessel's technique for transcribing polyphonic singing involved cobbling together multiple separately-recorded vocal parts, expanded with "obliquely-ruled rectangles" as needed to get them to line up with each other along the time axis. What we're hearing in these cases is an edited and idealized representation of the subject matter, not a "reproduction" of any single historical performance. Granted, the same thing could be said of most studio recordings produced by today's commercial music industry. Nevertheless, I raise this point to suggest that "playback" may not be the best way for us to think about what we're accomplishing here—and that's even truer of the material in the next track.

Source: Milton Metfessel, *Phonophotography in Folk Music: American Negro Songs in New Notation* (Chapel Hill: University of North Carolina Press, 1928).

a. "Swing Low, Sweet Chariot," Fig. 29, pages 82-84.
b. "I Heard the Voice," Fig. 10, pages 48-9; Fig. 11, page 50.
c. "Little David," Fig. 27, page 79; Fig. 28, page 80.

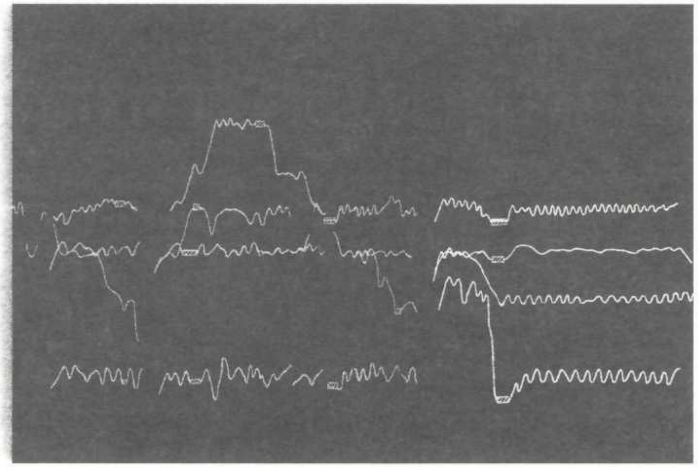


Image prepared for eduction.

(((((((5. PEASANT SONGS OF GREAT RUSSIA)))))))

(7:35)

Like Milton Metfessel, Yevgeniya Linyova made field recordings of vernacular musical performances and used an innovative graphical system to represent them on paper. In *The Peasant Songs of Great Russia* (1905), she introduced her project as follows:

> The collection submitted now is the first attempt at recording the part-songs of Russian peasants *by means of a phonograph*. The chief object of this collection is to give *the most accurate possible record of the peasant part song, without any alterations and improvements, just as it is sung by the people,* and thus to contribute to an elaboration of a correct method for writing down specimens of the popular genius.... Once recorded on the cylinder, the melody is repeated without any possibility of a mistake. It can be repeated until a perfectly correct record on paper is obtained.[29]

Unlike Metfessel, Linyova didn't make physical measurements of the wave periods inscribed by her equipment; instead, she simply played the cylinders over and over as many times as was necessary for her to transcribe the melodies from them by ear. "The work was of a new kind and one to which I was not used," she confided. "A great deal of labour has been spent on it. My patience got sometimes exhausted in trying to note down some rebellious song that was constantly eluding me; the shrill sound of the phonograph nearly drove me mad."[30] Although she used conventional musical notation for most of the transcriptions she made of Russian part-songs, she found this system to be an inherently poor fit for the material, as she explained:

> In view of the peculiar structure of the song in harmonic respect, it is very difficult to bring it under the rules of modern music.... On the one hand, the folk-song is based on the intervals of the natural scale without temperament, for the exact notation of which there are no corresponding signs in music. On the other hand—the rhythmical accent in a folk-song, connected with the varying accent of the verse, can be with great difficulty brought under the uniform metrical accentuation of our time system.[31]

With this in mind, Linyova transcribed a few representative songs strictly as graphs of time against frequency to show their intervals and rhythms in a format that wasn't based on "rules of modern music":

> In order to examine the music of the folk-song from another point of view and to compare records obtained from phonograms with those written down by ear, the music was represented in a graphical form shown on Tables I and II. These diagrams give simultaneously the melodic and rhythmic design of the song as well as the correla-

tion of the secondary parts with the principal melody. Each vertical division of the blue net corresponds with the interval of a semi-tone and on the left side of each diagram are marked various steps of the scale. The horizontal divisions of the net correspond to the duration of each note and one division is assumed as equal to one eighth note or quaver, and two divisions—to one quarter note (crotchet). The vertical black lines represent the bars. Separate vocal parts are represented with lines of different colour, and thus each part appears distinctly in the counterpunctual design. The red lines represent the part of the leader or solo singer.[32]

The first three graphical transcriptions in *The Peasant Songs of Great Russia* show different versions of the same song, "Gory" or "Ye Hills": first, a rendition Linyova had recorded herself using the phonograph, and then two other versions which she'd transcribed from published sources and which, for various reasons, she found incompatible with the "usual peasant style"—that is, she was offering them as negative examples in contrast with her own, and as inadequate or misleading representations of folk singing.[33] Her fourth transcription, "Kak ou nas na Sviatoy Roussi" ("As among us in Holy Russia"), was based on another of her own phonograph recordings of part-singing, so it's presumably intended as another "good" example.

Linyova also included graphical transcriptions of two monophonic songs: first "a very old *chorovod* or play-song (Vesnianka) which was used to accompany games," taken from a published song collection, with a text alluding "to the ancient custom of 'buying the bride'"; and second, "one of the typical contemporary songs that have sprung up among the factory workers."[34] Finally, she graphed out the parts to "A Mighty Fortress is Our God" by Martin Luther "as a typical example of West-European harmonization" which "may, by way of comparison, help us to a true appreciation of the harmonization of peasant songs."[35]

Many of Linyova's original field recordings on phonograph cylinder are preserved today at the Phonogram Archive of the Institute of Russian Literature (Pushkin House) of the Russian Academy of Sciences in Saint Petersburg, possibly including the two recordings from which she made transcriptions one and four in *The Peasant Songs of Great Russia*. As far as I'm aware, however, those particular cylinder recordings have never been published as such, so the graphical transcriptions may offer our only ready access to their content.

Today we can play Linyova's graphical transcriptions paleospectrophonically in the same way we played Metfessel's "pattern notation": by measuring the vertical distance corresponding to an octave, degridding the images, and running the results through reverse Fourier transform software—we don't have exact frequency or time values in this case, but we can choose reasonable approximations. When we do this, some of the peculiarities of the inscriptions yield distinctive sounds that I personally find quite interesting and pleasant. For example, Linyova often represented one of the vocal parts in her transcriptions with a dotted line that comes out in the sound files as an intermittent note, creating an accidental cross-rhythmic effect. Moreover, her practice of linking the successive horizontal lines comprising a vocal part with vertical lines produces a sudden burst of sound at each transition between notes which reminds me vaguely of listening to a handbell choir. Linyova seems to have made a few mistakes in her transcriptions besides, but I find that these add to the charm of the resulting sound files, and I haven't made any effort to correct them.

We might arguably speak of "reproducing" sound from the graphical transcriptions Linyova made from her own phonograph cylinders, since they're derived from specific recorded performances. On the other hand, she transcribed some of her other examples from works published in conventional musical notation, and in those cases there would seem to be no specific original performance there to "reproduce." And yet both types of inscription employ the same format. We can convert both into sound using the same exact method, so if we're not consistently "reproducing" sound, then how should we describe what we're doing?

The word I like to use for this process is *educe*, meaning "to bring out, elicit, develop, from a condition of latent, rudimentary, or merely potential existence."[36] When we apply paleospectrophony to each of Linyova's graphical transcriptions in turn, the result might not be a *reproduction* of any sound that ever existed before, but we can still refer to it without hesitation as an *eduction* of sound. That is, we aren't "playing" it as a live musician might play a composition from a piece of sheet music or even as an automatic musical instrument might play a given piece from a pinned barrel, a perforated

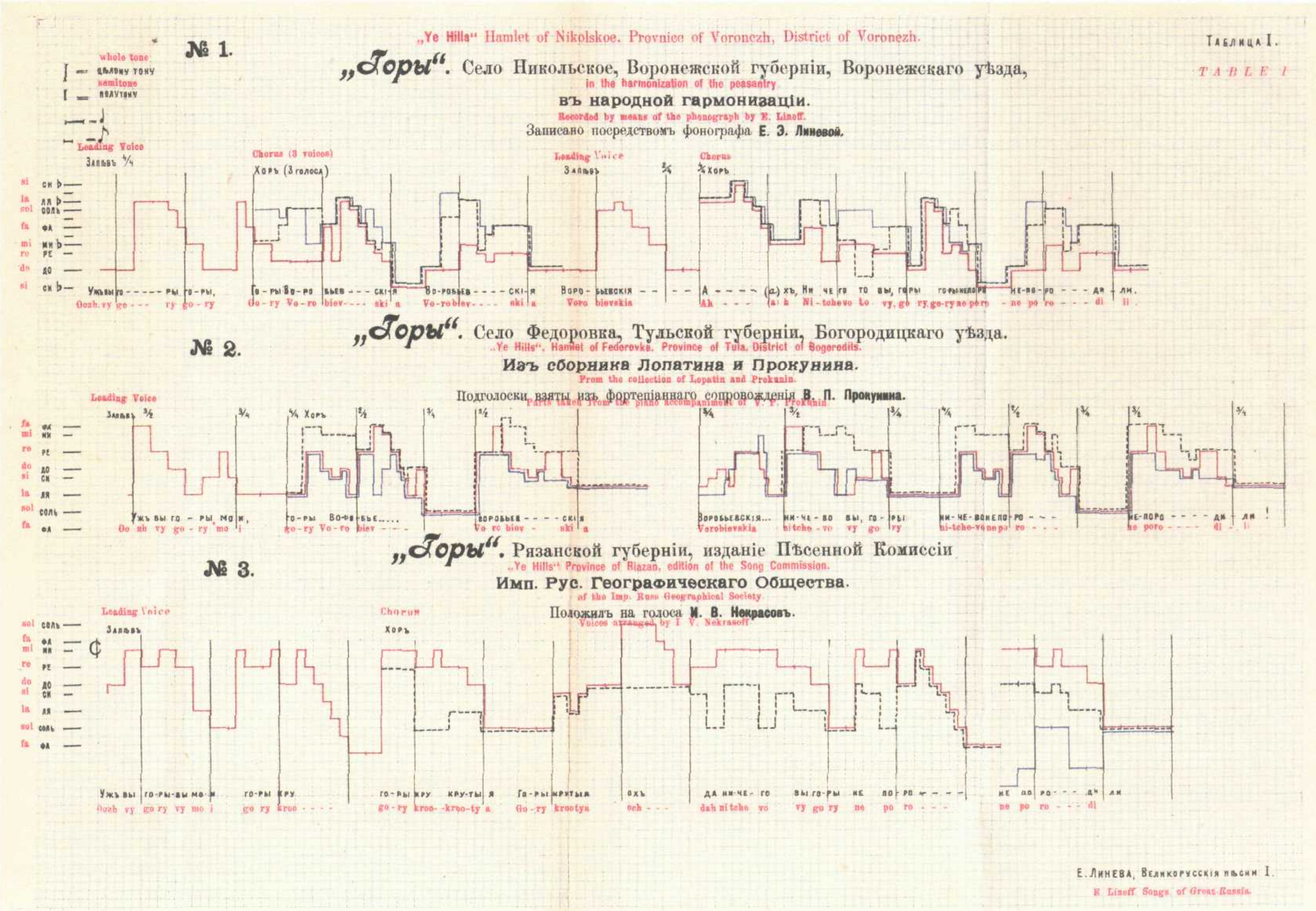


TABLE I—"Ye Hills"

No. 1: "Gory," Parish Nikolskoe, Voronej district, Province Voronej, from Linyova's own collection transcribed from phonogram.

No. 2: "Gory," (District Valdaj, province Novgorod), from the collection of Lopatin and Prokounin.

No. 3: "Gory," (District Rannenburg, province Riasan) from the collection of the Song Commission of the Imperial Geographical Society.

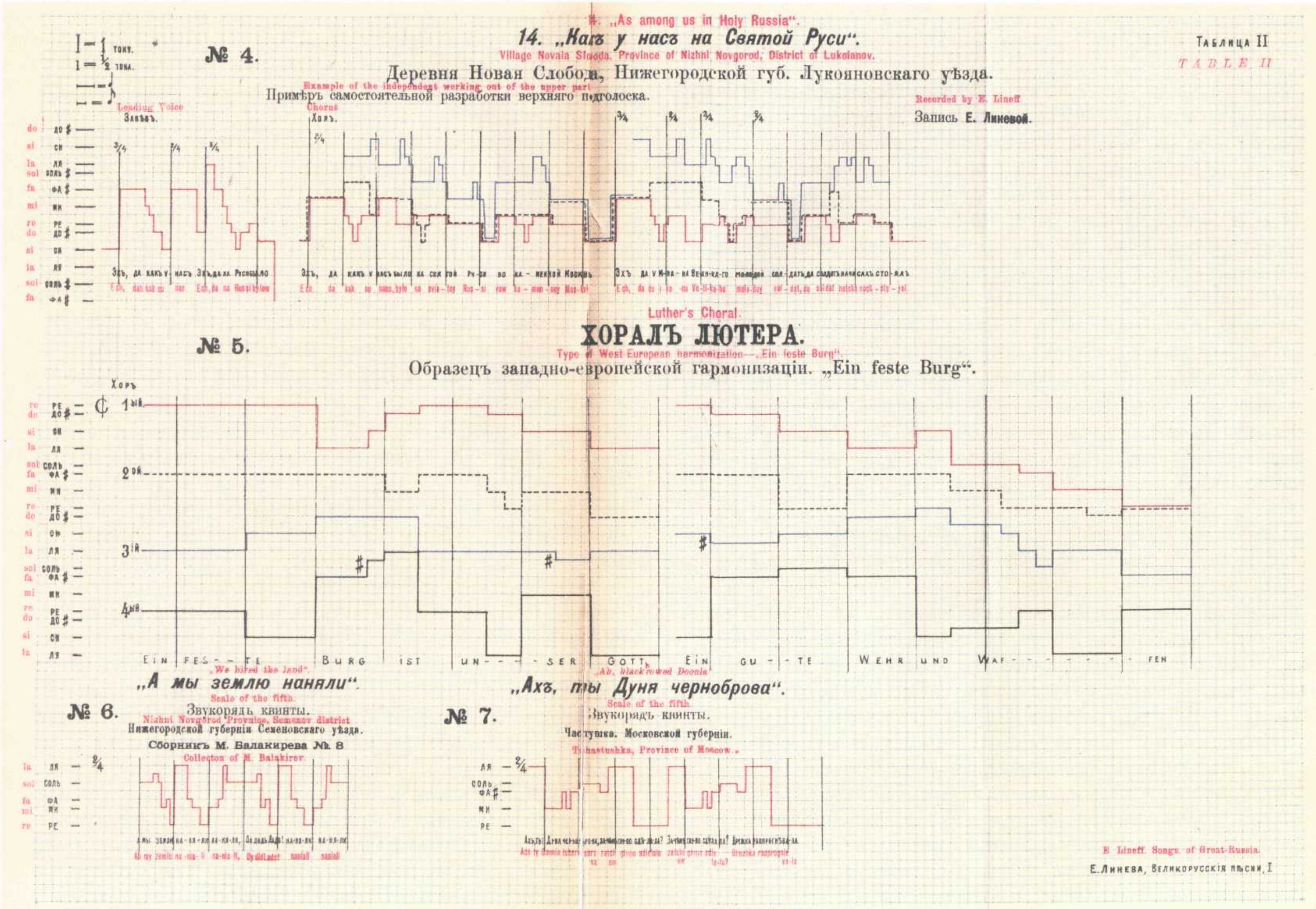


TABLE 2

No. 4: "As among us in Holy Russia," "Kak ou nas na Sviatoy Roussi."
No. 6: "We hired the land," "A mi zemliou naniali," from the collection of M. Balakireff.
No. 7: "Ah, black rowed Doonia," "Ach ti Dounia tchernobrova," a factory dance song (*tchastouchka*).
No. 5: "Luther's Choral, Ein Feste Burg," "Choral by Luther 'Ein feste Burg.'"

sheet, or a MIDI file—we're simply actualizing the raw aural data that lies latent within the image itself.

I'm not sure it's literally appropriate to call Linyova's transcriptions "sound recordings," since some of them weren't "recorded" in the sense we ordinarily associate with that term. Nevertheless, they might well count as sound recordings under international law. The relevant term here isn't *sound recording*, but *phonogram*, which doesn't have the same literal implications and therefore isn't limited by them. The WIPO Performances and Phonograms Treaty of 1996 formally defines a phonogram as a "fixation of the sounds of a performance or of other sounds, or of a representation of sounds." That second clause was thrown in specifically to encompass "phonograms that are not fixations of sounds," as the framers of the treaty noted: "Phonograms may be produced for instance using digital technology that fixes data which can be used to generate sounds even though no 'real' sounds have yet been produced."[37] The bottom line: for the purposes of international law, a phonogram doesn't need to have been "recorded" in any particular way so long as it's a "fixation of a representation of sounds" that "can be used to generate sounds."

According to the WIPO treaty definition, then, it would seem that any inscription we can convert into sound is a phonogram, regardless of how it came into being. But that interpretation seems too broad to be meaningful in practice. As I mentioned earlier, we can use reverse Fourier transform software to turn any image whatsoever into sound if we want to, including a photograph of a tree's branches or the Mona Lisa, but it would be absurd to conclude from this fact that every image in the world is *ipso facto* also a phonogram. The important distinction here, I believe, lies in whether or not an image was originally created to function as a representation of sound. The photograph of the tree's branches and the Mona Lisa were presumably created as representations of phenomena in the visual world, not the auditory world. There's a word for taking non-auditory data of this sort and expressing it aurally: *sonification*. On the other hand, images like Linyova's were originally created as representations of phenomena in the auditory world, not the visual world. When we educe them, we're not linking visual parameters to auditory ones arbitrarily at our own whim. Instead, those correlations are already present, inherent in the logic of the inscriptions themselves: "right" represents motion forward in time, "up" means an increase in frequency, "down" means a decrease in frequency, and so forth. The inscriptions themselves embody a phonographically meaningful data structure, just as a synthetically generated mp3 file does. We can legitimately *educe* them as sound because the sound is already implicit as such within them.

Track five contains each of Linyova's graphical transcriptions repeated twice, raised by one octave the second time for variety, with two stereo channels separated by one octave throughout.

Source: Eugenie Lineff, *The Peasant Songs of Great Russia as they are in the folk's harmonization* (St. Petersburg, David Nutt, 1905).
a. "Ye Hills," Table 1, No. 1: "'Gory' Parish Nikolskoe, Voronej district, Province Voronej, from my collection transcribed from phonogram."
b. "Ye Hills," Table 1, No. 2: "'Gory', (District Valdaj, province Novgorod), from the collection of Lopatin and Prokounin. The second and third voices are taken from the pianoforte accompaniment of B. P. Prokounin."
c. "Ye Hills," Table 1, No. 3: "'Gory' (District Rannenburg, province Riasan) from the collection of the Song Commission of the Imperial Geographical Society. Arranged for part singing by I. V. Nekrassoff."
d. "As among us in Holy Russia," Table 2, No. 4. "Kak ou nas na Sviatoy Roussi". (No. 14 of the collection by the author, sung and taken down by phonograph in the parish of Novoya Sloboda, district Loukoyanoff, province Nijni-Novgorod.) She writes: "From beginning of the 4-th bar all the lines of the diagram No. 4 representing the parts should be lowered one division to read correctly" (XXI, n.)—but I've played the inscription as is without making this adjustment.
e. "We hired the land," Table 2, No. 6: "'A mi zemliou naniali', from the collection of M. Balakireff."
f. "Ah, black rowed Doonia," Table 2, No. 7: "'Ach ti Dounia tchernobrova'. A factory dance song (*tchastouchka*)."
g. "Luther's Choral, Ein Feste Burg," Table 2, No. 5: "Choral by Luther 'Ein feste Burg'."

(((((((6. SPEECH ON FOREFATHER'S DAY)))))))

(2:03)

As I mentioned in the notes on track three, Edward Wheeler Scripture devised special machinery for copying "speech curves" from commercially available gramophone discs onto sheets of paper in straight lines and at larger size so that he could study them and incorporate them into his books as visual illustrations. But that wasn't the only way early gramophone recordings made it into print.

For example, a print of one gramophone disc appeared as part of a Zonophone advertisement in the March 1898 issue of *Cosmopolitan*, just as flat disc records and the machines to play them were first gaining a secure foothold in American households as a source of home entertainment. The text of the ad reads: "Place in the Zonophone and this is what you hear spoken," followed by a transcription of the speech supposedly represented by the recording. Of course, anyone who had obeyed the written instructions and actually tried to play the inked spiral on a Zonophone with a steel needle would have got nothing out of it but a chewed-up piece of paper. But today we can accept the ad's challenge with a better chance of success by educing sound from the image digitally. To create track six, I took a high-resolution scan of the spiral, converted it into a series of parallel lines through a polar to rectangular coordinates transform in Photoshop (something I've done with all the gramophone disc images in *Pictures of Sound*), and then proceeded just as I had with Scripture's "speech curves."

The disc shown in the advertisement contains some writing in its center area: "Chauncey M. Depew / From His Speech on Forefather's Day / 693Z." Chauncey Depew was a well-known attorney, railroad executive, and politician—he'd stumped New York for Abraham Lincoln in 1860 and went on in 1899 to become a United States Senator—but he was also highly regarded as an after-dinner speaker, and this particular recording showcases him in that capacity. The gramophone discs of the late 1890s often displayed the signatures of performers, as this example does, providing a visual mark of authentication: the purchaser of such a disc could be sure of owning the actual voice of Chauncey Depew, and not a mere imitation. The "Z" after the catalog number identifies a specific take—in other words, selection number 693 was recorded more than once, but "693Z" refers uniquely to just one of those recordings. As it happens, no actual shellac pressings of take 693Z have been reported to exist, so it's possible that this particular recording survives only thanks to its use in the *Cosmopolitan* ad, much as many important early motion pictures have come down to us only in the form of paper prints deposited for copyright registration purposes with the Library of Congress.

By coincidence, Chauncey Depew's "Speech on Forefathers' Day" also happens to have been one of the gramophone selections Edward Wheeler Scripture studied and presented in print as a visual illustration—but Scripture

had approached it very differently, reflecting his own very different goals. His published Depew "speech curve" takes up more space than the picture of the disc in the *Cosmopolitan* ad does, but it represents only a three-second excerpt (as heard in track three), greatly enlarged for ease of study. By contrast, the image in the *Cosmopolitan* ad contains the entire two-minute speech, but in the form of "speech curves" coiled into a tight spiral, and so tiny they can barely be seen. After all, the ad's purpose was only to give potential Zonophone buyers a sense for what a gramophone disc looked like, and not to enable them to study the form of the wavy line in any serious way. Instead of enlarging the waveforms as Scripture had done, the designer of the ad had actually reduced them to two-thirds of their original size—a scale at which minor imperfections of ink and paper produce extreme visual noise.

Even so, the speech still comes out clearly enough in playback (at 75 rpm) that we can detect a mistake in the accompanying transcription: Depew says "on the other in New England," not "on the other side in New England," as stated in the ad. Given the dicey source of the audio, I think that's rather impressive.

(((❋)))

Source: Advertisement, *Cosmopolitan* 24:5 (March 1898).

(Speech text from image at left:)

"My ancestors having arrived in this country among the early settlers, on the one side in New York, on the other side in New England, and having fallen in love and married in the old fashioned way, without regard to race or creed, I can claim membership of and attend all the national celebrations. First come the Scotch, whose dinner is made digestible by the bagpipes and indigestible by haggis, and whose glory in literature and philosophy none can dispute. The Scotchman keeps the Sabbath and everything else he can lay his hands on. Next come my own Dutch Knickerbocker brethren, who believed that Holland kept alive the spark of civil and religious liberty and are happy in the wisdom of their far sighted ancestors who pre-empted all the land on Manhattan Island. Then the sons of St. Patrick revel in wit and eloquence while the Welshmen display the intellect of Gladstone and the obstinacy of an army mule. But for real, honest, double-breasted claiming of all that there is in this country, and much that there is in the world of which the nineteenth century can boast and the twentieth century hope for, the Yankee takes the palm," etc., etc.

(((((((7. NUMBERS AND LETTERS)))))))

(1:10)

Emile Berliner's gramophone was first placed on the market not in America, but in Europe, where the first commercially manufactured gramophone discs became available for purchase in mid-1890. Even before that time, however, a couple images of complete discs had appeared in print so that curious readers could see what they looked like, and one of these is the subject of track seven. It's taken from an issue of the German illustrated periodical *Prometheus* which originally came out around 8 January 1890, according to my colleague Stephan Puille, who also made the high-resolution scan on which our playback is based.[38] As in the previous case, no conventionally playable pressings of this recording are known to exist; only the paper print version seems to survive. The text of the accompanying article reads as follows:

> In our figure 3 we present to our readers...a phonautogram in the original. This is not an imitation, but the impression directly effected in the printing press by a gramophonic original plate which the inventor [Emile Berliner] was good enough to prepare specifically for this purpose. The plates ordinarily used by him differ from the one printed only in considerably greater, approximately doubled diameter. The phonautogram represented on our plate contains the numbers from 1 to 10 and all letters of the alphabet in their order. The rotational speed of the plate was repeatedly changed on purpose during recording.[39]

The image itself is twelve centimeters in diameter, the same size as the first commercially manufactured gramophone discs, and as a negative relief print, it displays its white spiral trace against a black inked background. During this period, original gramophone recordings were prepared by tracing a wavy line in a waxy coating on a zinc disc and then etching it into the metal with chromic acid. Because the original zinc was used in this instance directly as a printing plate, the writing inscribed in its center area appears in mirror image, but it's still legible, giving us Berliner's own written notation of the content, "Zahlen, A, B, C, usw" ["Numbers, A, B, C, etc."], as well as the recording date, 14 December 1889. That makes it the oldest dated twelve-centimeter gramophone recording known to exist, and the layout of its written information differs enough from that seen on the first commercially manufactured discs (hardly any of which have dates inscribed on them) to suggest that it might have preceded them as a sort of experimental prototype.

The text of the accompanying article states that Berliner had intentionally varied the speed of rotation while recording, but the disc itself bears the inscription "R 50," which Stephan and I believe specifies the number of rotations per minute to be used when playing it, in line with a statement from 1888 that each of Berliner's discs had "engraved at its centre...the number 40, 50 or 60...so that whenever a disc may be used for reproduction people

Prometheus 1:14 (1890).

will know at what velocity to revolve the disc."[40] I've accordingly presented the content here at a consistent 50 rpm. First, we hear Berliner counting in German from one to twenty-five (not one to ten as stated in the article), and the 50 rpm speed sounds about right for that portion of the recording. Then, as Berliner recites the alphabet from A to Z—skipping over J—his speech decreases steadily in pitch and rapidity until his parting word, "ade" ("bye"), is scarcely recognizable; to produce this result, he must have gradually increased the speed of rotation while recording to about 140 rpm. If Stephan and I are right about the 50 rpm designation, this might well be the world's oldest surviving trick recording, intended to show the novel effect of playing a recorded voice at the "wrong" speed[41]—so here, too, I believe it's useful for us to think of what we're doing *in terms of eduction* rather than *reproduction*.

Source: *Prometheus* 1:14 (1890), 212, scanned by Stephan Puille.

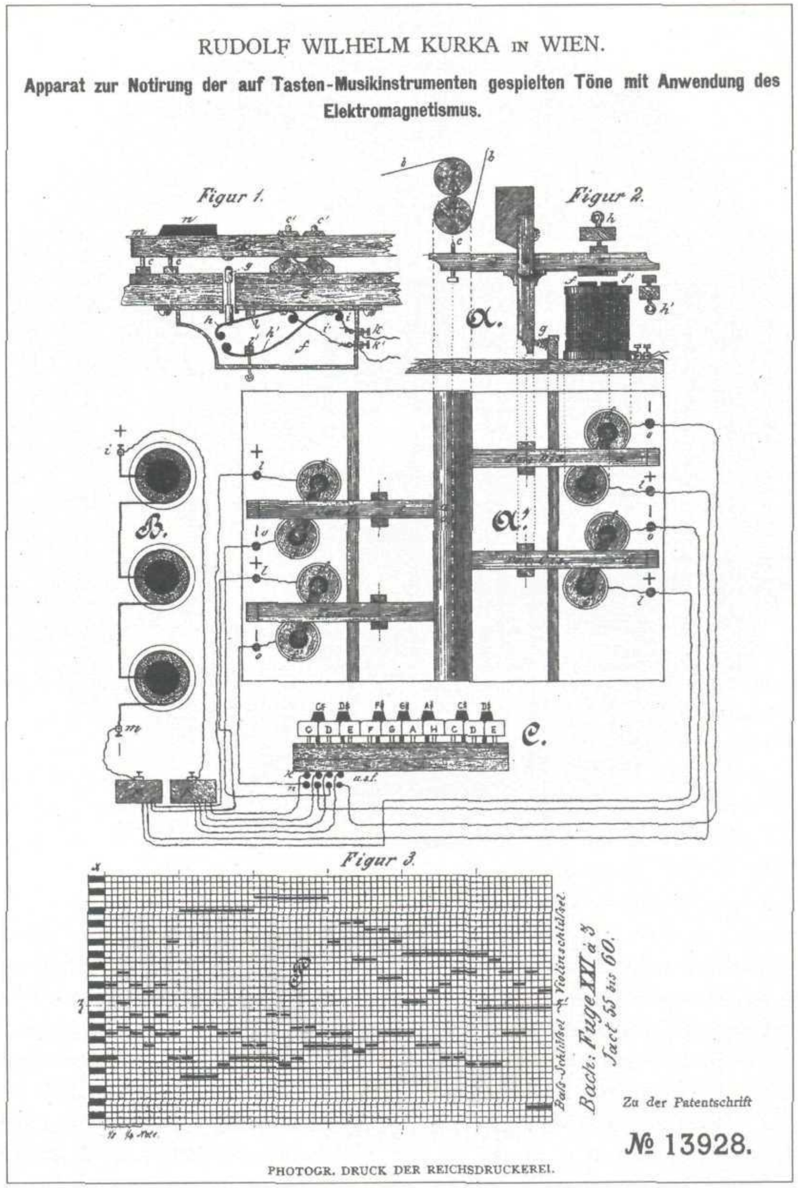
(((((((8. MELOGRAPHY)))))))

(1:39)

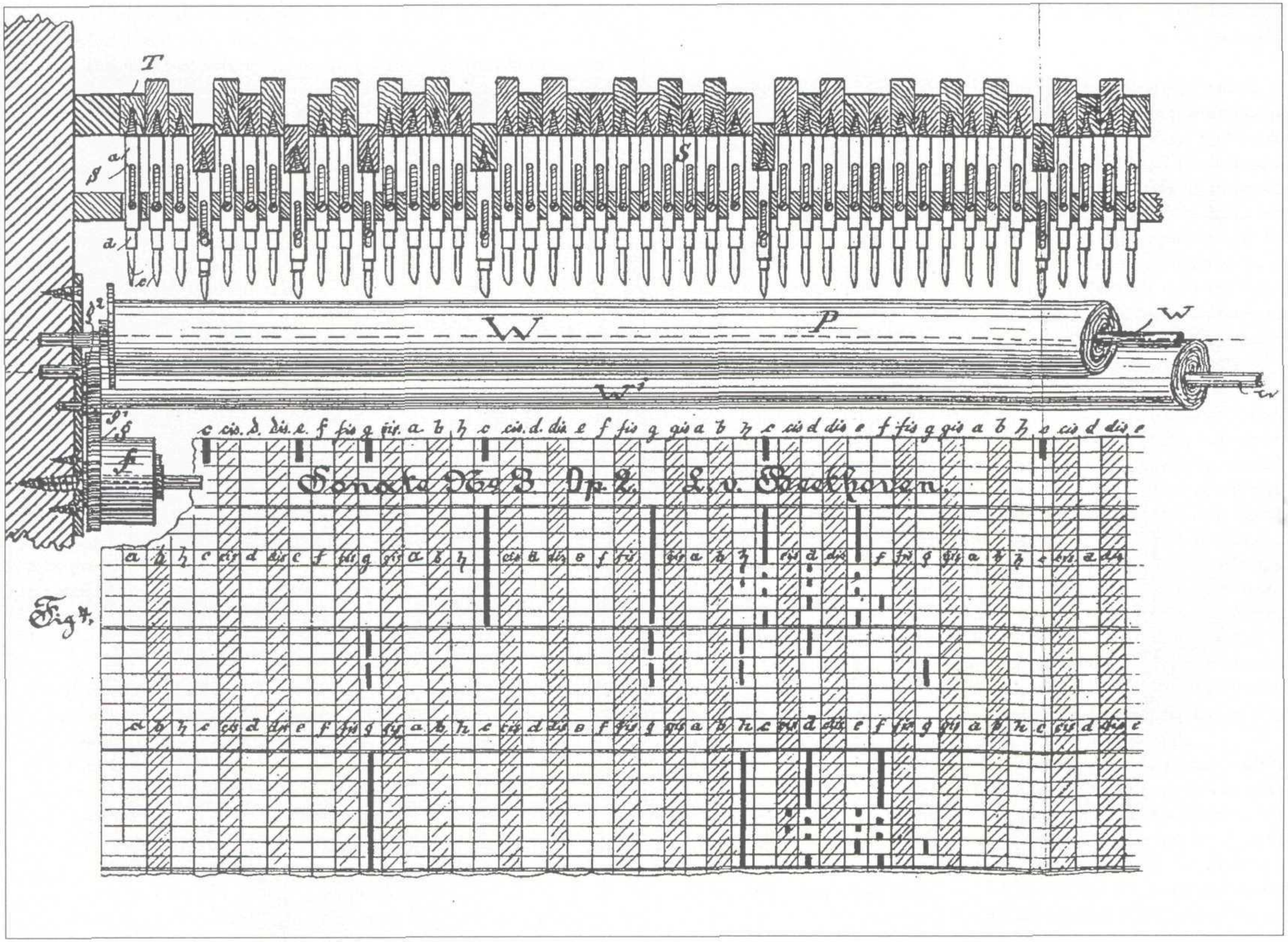
A musical almanac for the year 1776 included the following anecdote about the same Père Engramelle whose *notage* of Balbastre's "Romance" was featured in our opening track:

> Musicians often regret that there is no means of fixing thought and noting the pieces which a skillful composer performs on the spot and which he no longer remembers a moment later. We announce that this discovery has been made, and here is how it happened.
>
> An Italian virtuoso was in Lorraine at the court of King Stanislas. He had performed some harpsichord pieces which were much admired, but which he did not wish to give to anyone. Baptiste, musician of the King of Poland, spoke of this to Père Engramelle, who thought he saw a means of getting these pieces, and who pledged Baptiste to bring his harpsichordist to him a few days later.
>
> In the interval Père Engramelle placed under his harpsichord a large cylinder covered with white paper and overlaid with paper blackened with oil. He made a recording keyboard, the keys of which answered to those of the harpsichord, such that everything performed on the harpsichord was marked on the cylinder with the aid of the blackened paper. This cylinder was set in motion by a crank placed at the tip of the harpsichord and carried upon pivots with screws, such that it advanced a little to the side at each turn so that the different marks would not become confused. Its total revolution was fifteen turns, and it lasted around three quarters of an hour.
>
> The whole mechanism was concealed in the cleverest manner. The harpsichordist came to Père Engramelle on the agreed day, and he performed his pieces. As soon as he had left, Père Engramelle exposed his cylinder where not a note was missing. The Italian having returned a few days later, he was made to listen to a canary-organ [*serinette*] which repeated his pieces, and which imitated unto the charms of his playing. His surprise could scarcely be described, and he could not refrain from applauding himself at a theft made in so ingenious a fashion.[42]

I can't vouch for the truth of this story, but it does a nice job of explaining a method of recording keyboard music that saw fairly wide use in later years, for example in the creation of "hand-played" rolls for automatic pianos. This method is often referred to as *melography*, so that's what I'll call it here. The fact that melographic inscriptions are recorded from live performances has led some people to think of them as "sound recordings," and Scott Joplin's hand-played piano rolls were actually inducted into the National Recording Registry in 2002 as "outstanding examples of a less-familiar, now nearly-obsolete sound recording format."[43] But I'm less interested here in



Excerpt from Bach, Fugue XXI. Rudolf Wilhelm Kurka, "Notirung der auf Tasten-Musikinstrumenten gespielten Töne mit Anwendung des Elektromagnetismus," German patent 13928, issued 13 October 1880.



Excerpt from Beethoven, Sonata No. 3. Paul Boehm and Fedor Juliusberger, "Neuerung an Vorrichtungen zum selbstthätigen Aufzeichnen gespielter Musikstücke," German patent 37081, issued 13 November 1885.

how melographic inscriptions came into being than I am in how we can play them today.

It should be possible to read many types of melographic inscription as straightforward graphs of time versus frequency, which means that we can also educe sound from them by means of paleospectrophony. In the anecdote about Engramelle and the Italian harpsichordist, for example, the direction in which the cylinder rotated would give us our time axis, and the length of the keyboard would give us a viable frequency axis, presuming all available keys were represented in order by evenly spaced columns. So if we managed somehow to get our hands on Engramelle's surreptitiously-made record of the harpsichord performances, we could play it paleospectrophonically, and we'd probably hear a passable rendition of the recorded music—although it wouldn't sound anything like a harpsichord *per se*.

Unfortunately, I haven't been able to find any authentic melographic recordings from before the era of the hand-played piano roll, even though I'm sure some must exist out there somewhere. Nevertheless, patent specifications for melographic recording systems sometimes included drawings of sample recordings, and I've taken a couple of those as the basis for this track. The first is an excerpt from Johann Sebastian Bach's Fugue 21 in B flat major as found in Rudolf Wilhelm Kurka's German patent specification of 1880 for a method of "notation of notes played on keyboard instruments with the use of electromagnetism." The second comes from another German patent specification, this one from 1885: Ludwig van Beethoven's Piano Sonata No. 3 in C major, as used by Paul Boehm and Fedor Juliusberger to illustrate their "innovation in devices for the automatic recording of played pieces of music." Both are repeated twice with stereo channels separated one octave, and with both channels raised one octave the second time around.

If these inscriptions were originally meant to represent music played specifically on a piano, then we're hearing the intended notes and timing, but not the intended timbre—our results sound more like an organ. At the same time, the organ-like timbre isn't something I've imposed on the material arbitrarily; it's *educed* from the inscriptions, just as the notes and timing are, giving us an indexical aural representation of the precise thickness and layout of the graphical marks themselves.

The two examples heard in track eight may not be actual melographic "recordings"—in fact, they're almost certainly not—but if the real thing were ever to turn up, I suspect we would be able to educe it in the same way.

a. Excerpt from Bach, Fugue XXI. Source: Rudolf Wilhelm Kurka, "Notirung der auf Tasten-Musikinstrumenten gespielten Töne mit Anwendung des Elektromagnetismus," German patent 13928, issued 13 October 1880.
b. Excerpt from Beethoven, Sonata No. 3. Paul Boehm and Fedor Juliusberger, "Neuerung an Vorrichtungen zum selbstthätigen Aufzeichnen gespielter Musikstücke," German patent 37081, issued 13 November 1885.

(((((((9. SCHALLDRUCK)))))))

(2:33)

Emile Berliner invented his gramophone in the United States, but he returned to his native Germany in the latter part of 1889 to develop it further and to exploit it commercially for the first time. One of his main collaborators in this period was Louis Rosenthal of Frankfurt, and the recording featured in this track, made at Berliner's workplace in Hanover, might well document their first meeting. It survives today only in the form of a negative intaglio print, twenty-one centimeters in diameter, inserted into one of Berliner's scrapbooks preserved at the Library of Congress. The trace appears as a black ink spiral printed on a white paper ring mounted on red cardboard with some handwritten text in the center area: "Schalldruck in Kupferdruck Manier von Grammophon-Zink-Platte 1889" ["sound-pressing in copper-plate manner from zinc gramophone plate 1889"]. For many years, that inscription provided the only clue about the nature of this recording. But in 2010, I played the recording itself back from a high-resolution digital scan, at 50 rpm, and Stephan Puille, Norman Bruderhofer, and I prepared a transcription and translation of its contents shortly thereafter, which I copy here with only minor revisions:[44]

[Emile Berliner:]
Also hier ist eine Platte die Herr Louis Rosenthal mit nach Frankfurt nehmen wird. Er will dieselbe vervielfältigen. Vielleicht braucht er verschiedene Prozesse um das zu tun. Ein Prozess ist zum Beispiel, von der Platte einen Abdruck auf Falschpapier zu machen und dann über Falschpapierkopie eine, eine Photogravüre auf Zink oder andere Metalle auszuführen.

Also heute ist der, äh, 11. November 1889, und ich bin hier auf der Andreaestraße Nummer 2A, eine Treppe hoch, in Hannover. Jetzt wollen wir mal also irgend etwas sprechen:

[Emile Berliner:]
So here's a plate which Mr. Louis Rosenthal will take with him to Frankfurt. He wants to duplicate it. Maybe he'll use different processes to do that. One process is, for example, to make a print of the plate on tracing paper and then via tracing-paper copy to carry out a—a photogravure on zinc or other metals.

So today is the, uh, 11th of November 1889, and I'm here on the Andreaestraße, number 2A, one flight up, in Hanover. So now let's recite something:

Fest gemauert in der Erden
Steht die Form, aus Lehm gebrannt.
Heute muss die Glocke werden,
Frisch, Gesellen! seid zur Hand.
Von der Stirne heiß
Rinnen muss der Schweiß,
Soll das Werk den Meister loben,
Doch der Segen kommt von oben.

Fast walled up within the earth,
The model, baked in clay doth stand.
Today the bell must have its birth!
Ho! my comrades, be at hand!
From each burning brow
Drops of sweat must flow,
Ere to the Master praise be given—
Yet the blessing comes from Heaven.[45]

Eins, zwei, drei, vier, fünf, sechs, sieben, acht, neun, zehn, elf, zwölf, dreizehn, vierzehn, fünfzehn, sechzehn, siebzehn, achtzehn, neunzehn, zwanzig. Un, deux, trois, quatre, cinq, six, sept, huit, neuf, dix, onze, douze, treize, quatorze, quinze, seize, dix-sept, dix-huit, dix-neuf, vingt.

[*Counts from one to twenty in German, then in French.*]

Das war also französisch. Das kann ich aber auch in Englisch. One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty.

So that was French. But I can do that in English too. [*Counts from one to twenty in English.*]

Ha ha ha ha ha ha ha. Jetzt wollen wir mal was singen.

Ha ha ha ha ha ha ha. Now let's sing something.

Morgenrot! Morgenrot! Leuchtest mir zum frühen Tod!
Bald wird die Trompete blasen, dann muss ich mein Leben lassen,
Ich und mancher Kamerad.

Morning sky! Morning sky! Light'st thou my way soon to die!
Soon the trumpets they will blow, then from my life I must go,
Many a comrade and I.

Na, könnte er selbst, Herr Rosenthal, was singen. Sprechen Sie mal was hinein.

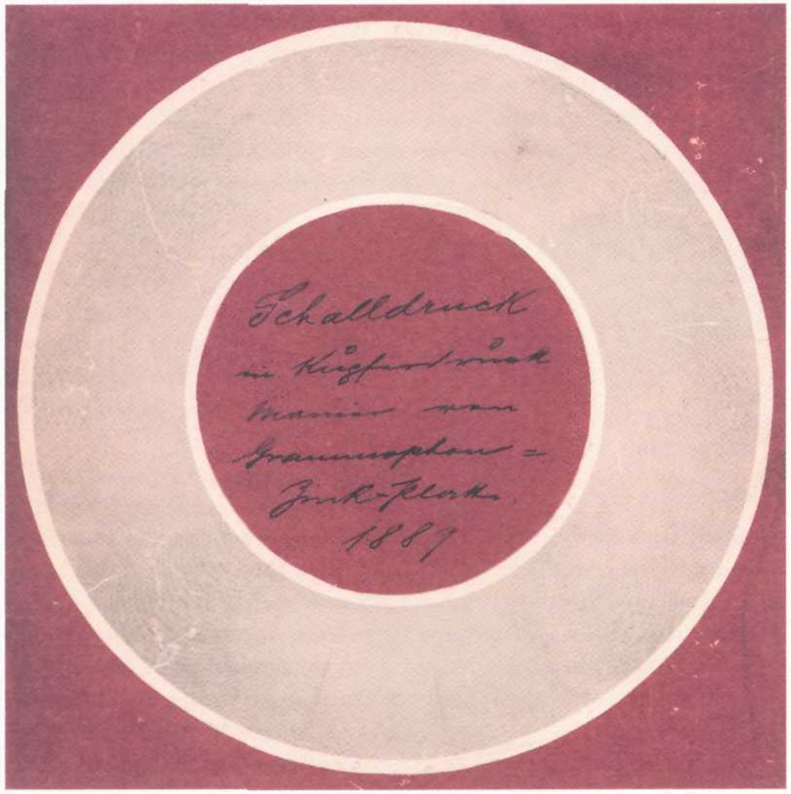
Well, let's have him, Mr. Rosenthal, sing something himself. Speak something into it.

[Louis Rosenthal:]
Ich weiß nicht ob's gelingen wird, ich werde aber mein Möglichstes tun um diese Geschichte zu fördern.

[Louis Rosenthal:]
I don't know if it will succeed, but I'll do my utmost to further this affair.

[Emile Berliner:]
Adieu für heute, ade. Das war Rosenthal der zuletzt gesprochen hat, aus Frankfurt.

[Emile Berliner:]
Adieu for today, bye. That was Rosenthal who last spoke, from Frankfurt.



Gramophone scrapbook, #14120101, Emile Berliner Papers, Library of Congress, item number 36.

There's a lot of significant content here—in fact, Berliner's performance of the excerpt from "Reiters Morgengsang" is, to the best of my knowledge, the oldest recording of singing in the German language currently known to exist. What I find most striking about the recording, however, is its explicitly articulated place within the experimental history of the gramophone itself. Berliner made it, first of all, to demonstrate his recording process for Rosenthal so that he could witness a record being made at first hand; and second, to create a specimen for Rosenthal to take back with him to Frankfurt for use in experiments with different methods of duplication. Moreover, the form in which the recording survives today indicates that Rosenthal must actually have used it for its intended purpose—namely, by making a paper print from it. Berliner mentioned this line of experimentation during a lecture he gave before the Elekrotechnischer Verein in Berlin on 26 November 1889, just fifteen days later:

> From this etched plate, three-dimensional copies of all kinds can be produced with ease. Furthermore, it can be used as a printing plate, and from such sound-pressings [*Schalldrucken*] on ordinary or also tracing paper—of which I have a few here—discs can be produced by means of photogravure that exactly agree with and sound like the original plate. The pressed sound-registers can, as is now being done in Frankfurt, be photographically enlarged, and from this enlargement enlarged sound-plates can be made in turn through photogravure, which should then be louder than the original plate.[46]

Maybe the "Schalldruck" Berliner later pasted into his scrapbook was one of the very specimens he had with him when he gave this talk.

(((❋)))

Source: Gramophone scrapbook, #14120101, Emile Berliner Papers, Library of Congress, item number 36.

(((((((10. DER HANDSCHUH)))))))

(2:20)

In track seven, we listened to a sample gramophone recording published in *Prometheus* at the start of January 1890. About a month later, a second gramophone recording appeared in print, this time in the German illustrated magazine *Über Land und Meer*—a black spiral trace on a white background, fifteen centimeters in diameter. The accompanying article introduces it as follows:

> The gramophonic plate (sound-plate, tone-plate) is just as imperishable as a copperplate engraving. It can be duplicated easily through galvanoplasty; each copy sounds exactly like the original. Furthermore, prints on paper can be prepared from the sound-plates in any desired number by means of the copperplate printing press, and sound-plates that agree with the original plate can be made after the copies (sound-pressings [*Schalldrucken*]) through heliogravure (photogravure, phototypy, heliotypy) in turn. Thus anyone who receives by post in an ordinary letter a phonogram in the form of a sound-pressing and has a sound-plate made from it by photomechanical means has material for his listening-gramophone. There is also the possibility of enlarging a sound-pressing with the help of photography and in this manner to amplify the sound of the original plate.
>
> Figure 3 shows what a sound-pressing looks like. It is a print made firstly with the help of the copperplate printing press of a sound-plate which contains Schiller's ballad "Der Handschuh" recited by E. Berliner, 320 words, in a belt a few centimeters in breadth... The outer ring of the spiral, somewhat reduced in our illustration, has a diameter of about 20 centimeters on the original plate, and thus a circumference of about 63 centimeters; at the inner ring the corresponding measurements are 12 and 38 centimeters. Since the disc was moved during recording at a constant speed of 50 revolutions per minute, and thus 5/6 of a rotation per second, so for each second of speaking time there results 53 centimeters of spiral length in the outer ring and 32 centimeters of spiral length in the inner.

As we've seen, Berliner stated that he had "a few" sound-pressings with him during his Berlin lecture of 26 November 1889, and the "Schalldruck" featured in the previous track has the number three written at the top and bottom of its outer rim, suggesting that it might have been the third in a series. Several clues suggest that the "Handschuh" recording published in *Über Land und Meer* might have been one of the others. Its original twenty-centimeter diameter, 50 rpm recording speed, and intaglio format closely resemble those of the other "Schalldruck," while the fact that it's a positive (counterclockwise) print rather than a negative (clockwise) one, and reduced to three quarters of its original size, hints at the use of the same photomechanical techniques Rosenthal was experimenting with during this period. For these reasons, "Der Handschuh" seems likely to represent another zinc original which Rosenthal had taken back with him to Frankfurt after his visit to Berliner in Hanover in November 1889, and which he used to make another sample sound-pressing that formed the basis for the print in *Über Land und Meer* a few months later.

The accompanying article discusses the visual appearance of the "Handschuh" recording at some length, implying that readers should have found it interesting to look at:

"With the help of a magnifying glass one recognizes already with a one-time enlargement, much more clearly than is possible with the naked eye, the rings of the spiral running tightly within one another. In many places one catches sight of sharply pronounced waves of the most varied forms. Each wave is the picture of an aerial sound wave, caused by a certain sound, which the gramophone had automatically recorded on the sound-plate.... The form of the waves conforms to the character and timbre of the sound. Already the inventor of the gramophone has succeeded in identifying the characteristic waveform of a few sounds (the vowels a, e, i, o, u). It will yet require immense mental work, but the goal will be reached in the foreseeable future of reading a sound-pressing as one reads writing and sheet music. If we get there: what a triumph of the human mind!"

However, the article also spells out a method readers could have used to *listen* to the recording in front of them, if they felt it was worth the trouble:

"The well-disposed reader will be astonished when I reveal to him that a single person can coax out the content from a sound-plate even without a listening-gramophone. It sounds like a fairy tale.

Let an engraved plate be made through heliotypy from the sound-pressing in figure 3 and lain (center upon center) on a disc which is turned steadily at about fifty times per minute. Now, anyone who holds a bamboo stick about 15 centimeters long and ¾ centimeter thick, to the end of which a stout darning needle is attached such that 1 centimeter protrudes, at an angle between the teeth and presses the point of the needle with the finger softly into the groove, starting from the edge, hears clearly the "Handschuh" in the speech of the inventor, especially if both ears are stopped with cotton. An ordinary fountain pen with a steel nib from which one of the two points is broken off, a knitting needle, or simply a hard toothpick likewise guarantee the success of the experiment. With hard material, for example a knitting needle, it is advisable to attach a rubber tube to the upper end to be held by the teeth, in order to avoid a harsh ringing sound. An easily-operated lathe is sufficient for holding the plate."

I suppose anyone who had actually followed these instructions back in the year 1890 might have had some chance of hearing Berliner's recitation of "Der Handschuh" in the way described, however fanciful it might seem. Together with the earlier speculation that ink prints of gramophone recordings could be sent through the mail and then reconverted into playable form at their destination, this scenario hints at a vision of a future world in which *pictures* of sound recordings and *playable* sound recordings would be readily interchangeable and interconvertible.

Today, of course, you can listen to "Der Handschuh" without having to hold a bamboo stick between your teeth simply by playing track ten on the CD.

Der Handschuh

Vor seinem Löwengarten,
Das Kampfspiel zu erwarten,
Saß König Franz,
Und um ihn die Großen der Krone,
Und rings auf dem Balkone
Die Damen in schönem Kranz.

Und wie er winkt mit dem Finger,
Auftut sich der weite Zwinger,
Und hinein mit bedächtigem Schritt
Ein Löwe tritt
Und sieht sich stumm
Rings um,
Mit langem Gähnen,
Und schüttelt die Mähnen
Und streckt die Glieder
Und legt sich nieder.

Und der König winkt wieder,
Da öffnet sich behend
Ein zweites Tor,
Daraus rennt
Mit wildem Sprunge
Ein Tiger hervor.
Wie der den Löwen erschaut,
Brüllt er laut,
Schlägt mit dem Schweif
Einen furchtbaren Reif
Und recket die Zunge,
Und im Kreise scheu

The Glove

Before his lion-yard,
awaiting the fighting match,
sat King Francis,
and around him the great ones of the crown,
and round about the balcony
the ladies in a beautiful circle.

And as he signals with his finger,
the wide cage opens,
and in with measured stride
steps a lion,
and looks mutely
around,
with a long yawn,
and shakes his mane,
and stretches his limbs
and lies down.

And the king signals again;
now swiftly opens
a second gate;
from it
with a wild leap
runs forth a tiger.
As he sees the lion,
he roars loudly,
whips with his tail
a frightful ring
and extends his tongue,
and in a cautious circle

Umgeht er den Leu
Grimmig schnurrend,
Drauf streckt er sich murrend
Zur Seite nieder.

Und der König winkt wieder,
Da speit das doppelt geöffnete Haus
Zwei Leoparden auf einmal aus,
Die stürzen mit mutiger Kampfbegier
Auf das Tigertier;
Das packt sie mit seinen grimmigen Tatzen,
Und der Leu mit Gebrüll
Richtet sich auf, da wird's still;
Und herum im Kreis,
Von Mordsucht heiß,
Lagern die greulichen Katzen.

Da fällt von des Altans Rand
Ein Handschuh von schöner Hand
Zwischen den Tiger und den Leun
Mitten hinein.

Und zu Ritter Delorges, spottender Weis',
Wendet sich Fräulein Kunigund:
»Herr Ritter, ist Eure Lieb' so heiß,
Wie Ihr mir's schwört zu jeder Stund',
Ei, so hebt mir den Handschuh auf.«

Und der Ritter in schnellem Lauf
Steigt hinab in den furchtbaren Zwinger
Mit festem Schritte
Und aus der Ungeheuer Mitte
Nimmt er den Handschuh mit keckem Finger.

Und mit Erstaunen und mit Grauen
Sehen's die Ritter und Edelfrauen,
Und gelassen bringt er den Handschuh zurück.
Da schallt ihm sein Lob aus jedem Munde,

he walks round the lion
purring grimly;
then he stretches himself out grumbling
at the side.

And the king signals again;
now the doubly opened house spits
two leopards out at once;
they fall with daring battle-lust
upon the tiger-beast;
it seizes them with its fierce paws,
and the lion with a roar
stands up—now it gets quiet;
and around in a circle,
hot with thirst for murder,
the dreadful cats encamp.

Now falls from the edge of the balcony
a glove from beautiful hand
between the tiger and the lion
right into the middle.

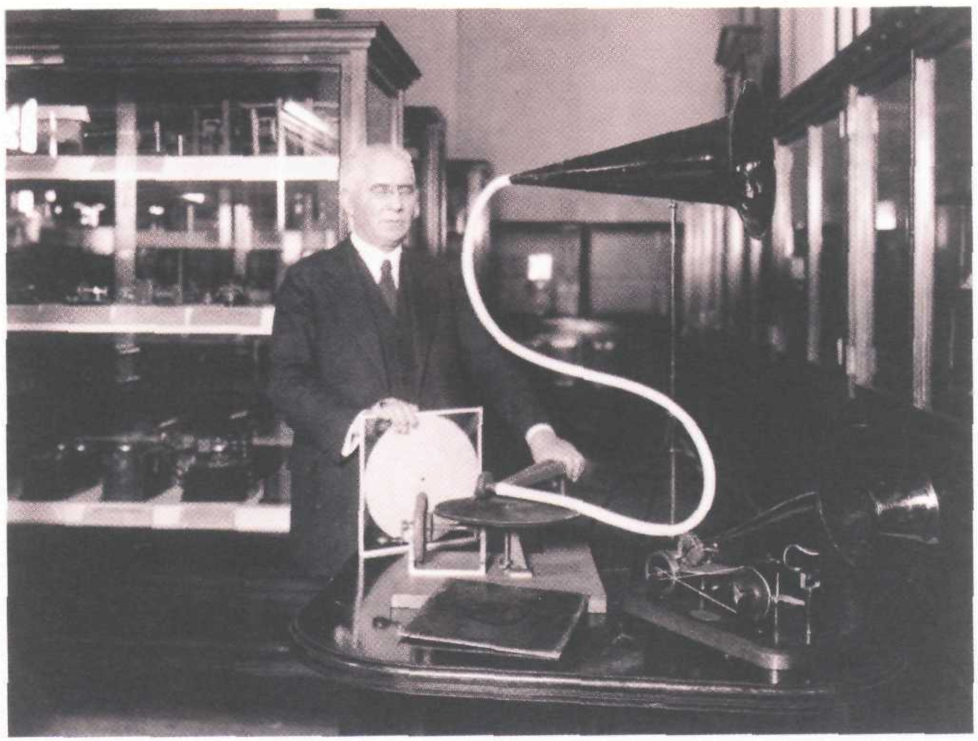
And to the knight Delorges, mockingly,
Miss Kunigund turns:
"Sir Knight, if your love is as hot
as you swear to me at every hour,
well, then pick up the glove for me."

And the knight in swift motion
climbs down into the frightful cage
with determined stride
and from the monsters' midst
he takes the glove with a bold finger.

And with amazement and with horror
the knights and noblewomen see it;
and calmly he brings the glove back.
Now his praise echoes from every mouth,

Aber mit zärtlichem Liebesblick —	but with a tender look of affection—
Er verheißt ihm sein nahes Glück —	it promises him his near happiness—
Empfängt ihn Fräulein Kunigunde.	Miss Kunigund receives him.
Und er wirft ihr den Handschuh ins Gesicht:	And he throws the glove in her face:
»Den Dank, Dame, begehr' ich nicht!«	"Lady, I don't want the thanks!"
Und verlässt sie zur selben Stunde.	And leaves her that same hour.

Source: R. Raab, "Das Grammophon, eine neue Schallwiederholungsmaschine," *Über Land und Meer* 63 (1890), 395.



Emile Berliner with some of his early gramophones and recordings.

(((((((11. BALANCE TO CORNERS)))))))

(0:49)

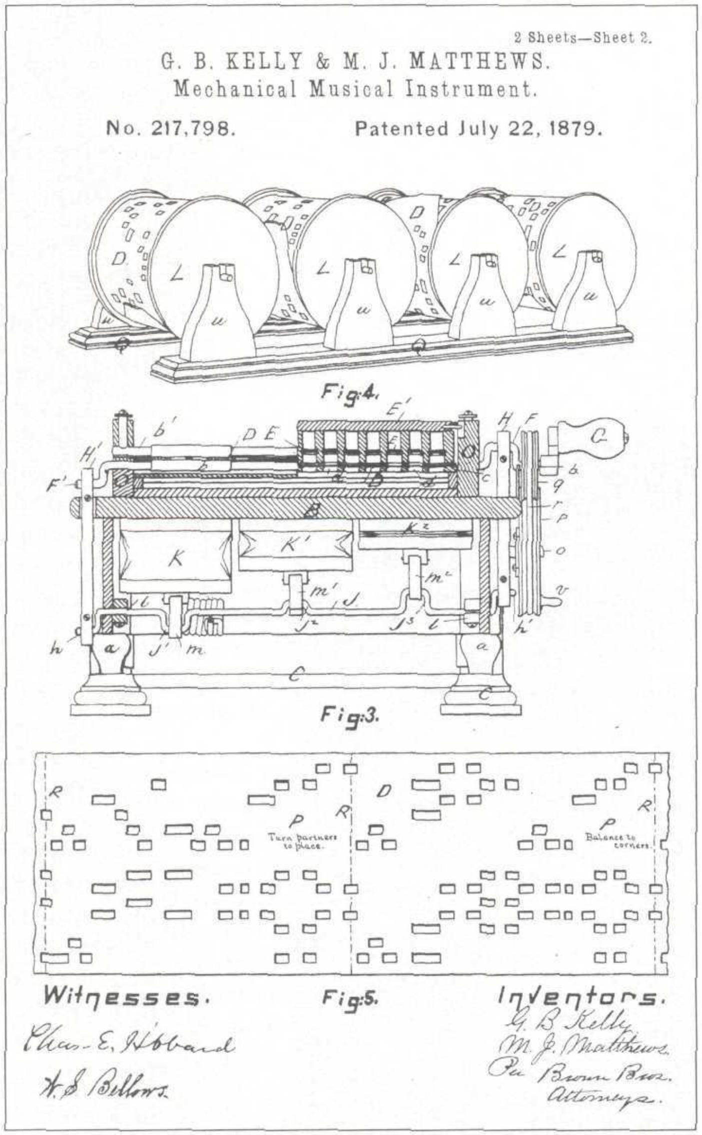
In the notes on track nine, I pointed out that melographic recordings of keyboard performances can often be read as graphs of time versus frequency, which means that we can educe them paleospectrophonically. The same is also true of many other types of automatic musical media.

But a data format needs to fulfill two conditions in order for this to work. First, one dimension has to correspond to time, which in the case of automatic musical media it almost always does. Second, the other dimension needs to correspond to frequency, and that's where we most often run into difficulties. Sometimes we find the notes comprising the scale positioned entirely out of order, which is an obvious deal-breaker. But even if the notes are arranged in order by pitch, we can still be out of luck if their physical spacing doesn't correspond well to their actual frequency intervals. The problem is that not all automatic musical instruments can play a complete chromatic scale, and media formats intended for those instruments that can't tend to divide the available space equally among whatever notes they *can* play. The octave ends up divided spatially into fewer than twelve units, which has unfortunate implications for paleospectrophony.

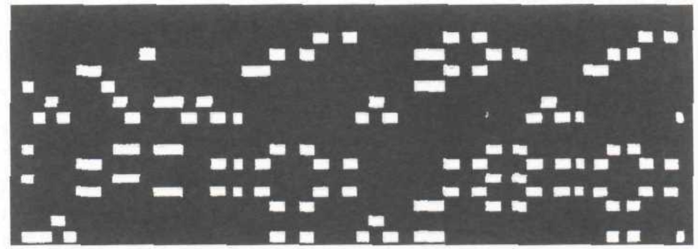
Take for example the device pictured in an American patent specification of 1879 for an "improvement in mechanical musical instruments": an "orguinette" in which rectangular perforations in a moving paper strip controlled the flow of air from a bellows to a rank of reeds. One of this instrument's distinctive features was that it had dance calls printed on its paper strip alongside the perforations, so that someone could call the figures for a dance correctly by reading them aloud as they passed a stationary pointer. A sample musical program pictured in the specification—oriented to play from right to left—contains a couple representative dance calls to show how this was supposed to work: "balance to corners" and "turn partners to place."

But this instrument was designed to play only a limited range of notes: in the sample inscription, evenly spaced rows were assigned starting from the bottom to G, A, B, C, D, E, F, and then, following a "blank" row, G, A, B, C, D, and E. If we count the "blank" row, the octave was divided spatially into just eight units, so if we try to educe this inscription paleospectrophonically in its original, unaltered form, the results come out in an eight-tone equal-tempered scale, which plays havoc with the harmonies. You can hear what this sounds like in the first part of track eleven. (Once again, stereo channels are separated by one octave. Since the rectangles representing the notes are drawn only in outline, I also had to fill them in before processing—a minor detail.)

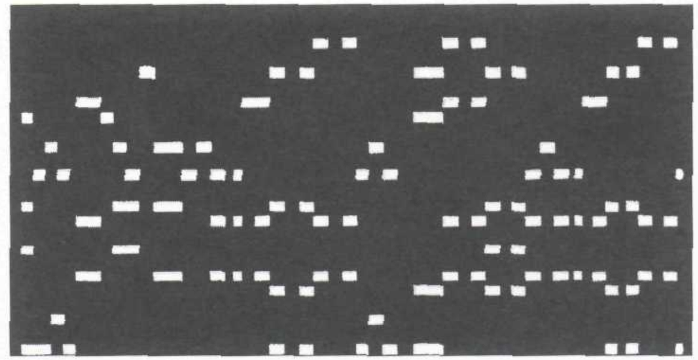
We can fix the problem if we wish by inserting extra blank rows into the image, corresponding to G#, A#, C#, and D#, so that its octave gets divided



"Improvement in Mechanical Musical Instruments,"
U. S. Patent 217,798, sheet 2.



Track in original form.



Track spaced out with extra rows.

spatially into twelve units instead of eight. Then, when we educe the inscription paleospectrophonically, this gives us a twelve-tone equal-tempered scale that does justice to the harmonies. That's what you hear in the second part of track eleven. Personally, I feel that this tactic ought to be considered cheating under the rules of paleospectrophony, and I've done it here just this once only to illustrate a point. Fortunately, many automatic musical media formats *do* divide the octave spatially into twelve evenly-spaced units, and we can educe those paleospectrophonically to good effect without tampering with the frequency axis. We'll hear some examples of this kind in the next two tracks.

Source: George B. Kelly and Mason J. Matthews, "Improvement in Mechanical Musical Instruments," U. S. Patent 217,798, filed 10 January 1879, issued 22 July 1879.

((((((12. SAILING, SAILING))))))

(1:17)

Track twelve is based on a pair of sample musical programs found in an American patent specification filed in 1885 for a new type of "perforated music-sheet." What made this "music-sheet" distinctive and patentable was the fact that it was divided into two sections, with the treble notes from the lower section strategically duplicated in the upper section. The patent text explained:

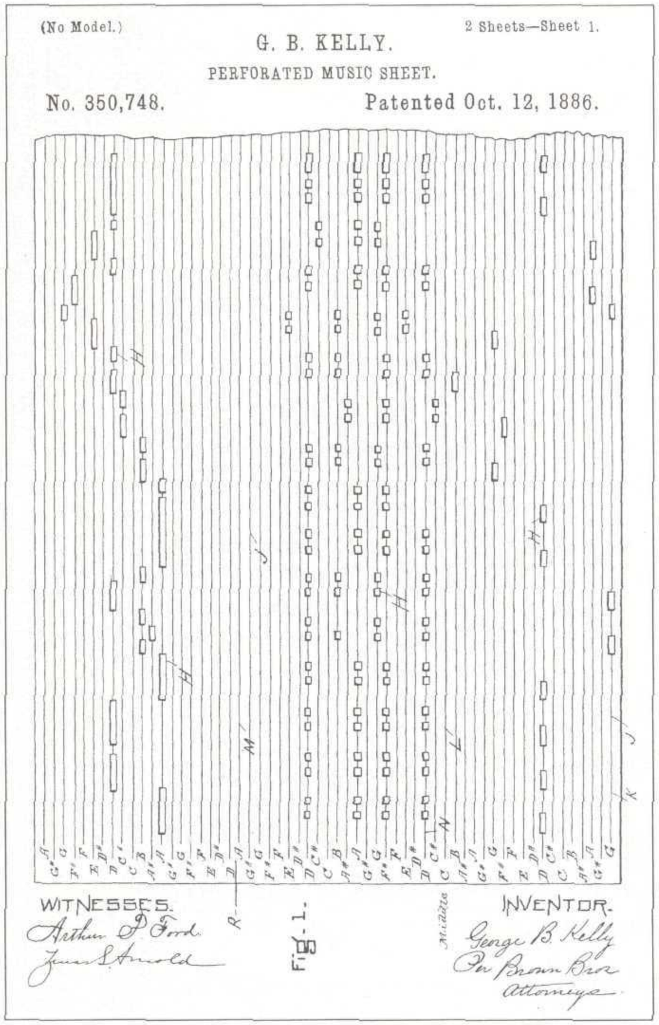
> A music-sheet for a mechanical musical instrument having its upper part or treble-notes duplicated, as described, and located or arranged on said sheet above the regular scale of the sheet...has many and important advantages over the present manner of arranging the scale on said sheets, where there is only one line or row of perforations for each note. For example, a separate and independent line or row of perforations is had for each note of the melody, the accompaniment or harmony in its higher notes also having its separate and independent notes, so that each part can be arranged thereon separately and independently and irrespective of the other; also, the melody-notes can have their sounding devices of a different quality of tone, or can be arranged an octave higher or lower, if desired.

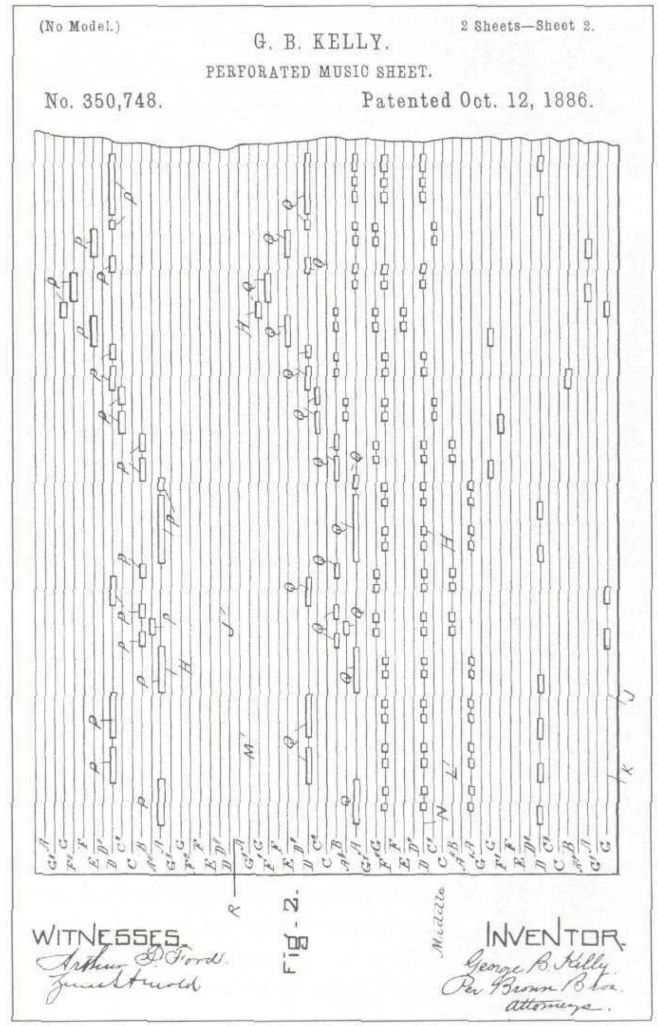
The first sample plate shows an arrangement in which the "upper series or duplication of treble notes is to be used for the melody of the musical composition, and all below the said duplicated notes for the accompaniment and harmony." The second sample plate instead shows an alternative arrangement "where P represents the melody in the notes that are duplicated, and Q the same melody in the scale proper, the accompaniment being arranged in accordance therewith. Both notes of the melody are of the same pitch, although, as is obvious, the duplicated series of notes could have its sounding devices an octave higher." In other words, the first plate uses the two sections to *separate* the melody from the accompaniment, whereas the second plate uses them to *double up* the melody.

In track twelve, the left channel corresponds to the upper sections of the plates, while the right channel corresponds to the lower sections of the plates. I've repeated each plate twice, first with the upper section set to the same frequencies as the lower section, and then with the upper section set an octave higher—taking as my cue the patent text's own recommendations about how the two sections of the "music-sheet" might be put to use.

Both plates represent different arrangements of the song "Sailing, Sailing, Over the Bounding Main," which had first been published in 1880. At the time it was chosen as the subject for these sample "music-sheets," then, it was only a few years old—a popular tune of the day that had yet to withstand the test of time.

"Perforated Music-Sheet," U. S. Patent 350,748, Sheets 1 and 2.





Source: George B. Kelly, "Perforated Music-Sheet," U. S. Patent 350,748, filed 7 December 1885, issued 12 October 1886.

(((((((13. FROM CHOPIN'S "BERCEUSE")))))))

(1:34)

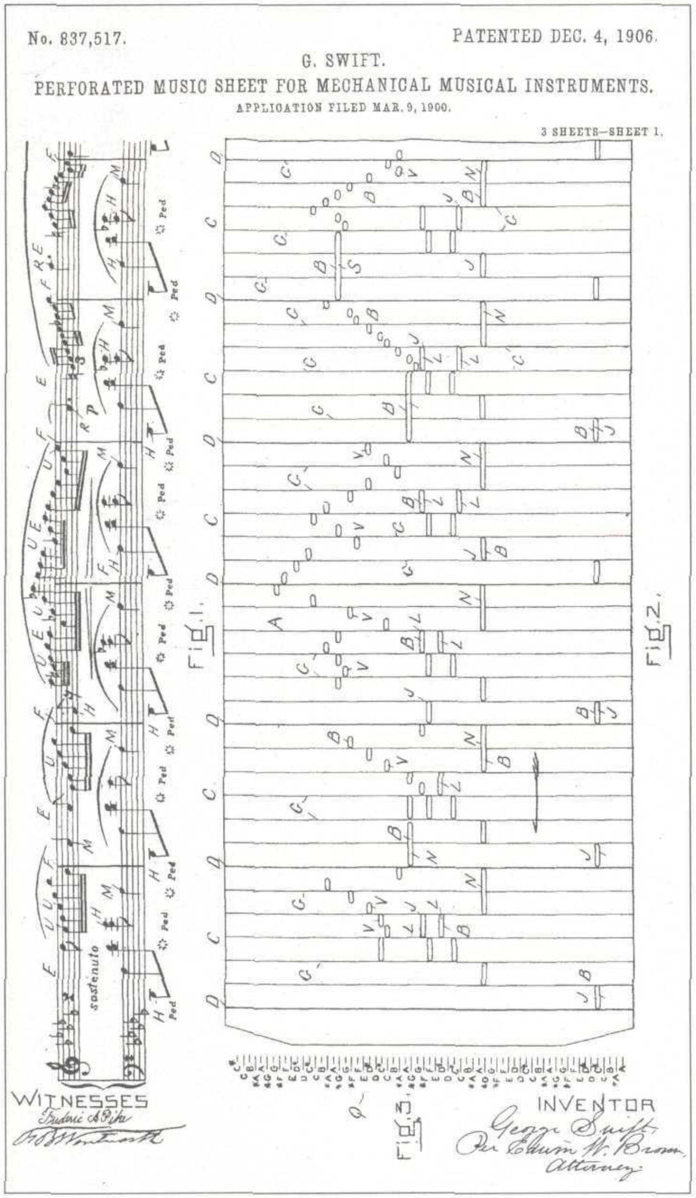
In 1906, George Swift of Meriden, Connecticut, patented a method of encoding music onto piano rolls that was intentionally imprecise—"I make no claim whatever," he insisted, "to music-sheets wherein the note-perforations are of the exact length of the notes and measures to correspond with the indications of the musical score." This might seem like an inauspicious approach to the design of automatic musical media, but Swift presumably knew what he was doing—he worked preparing master piano rolls for Æolian, the leading company in the field, which also ended up acquiring this patent from him. The patent's stated goal was to make the operation of automatic pianos "less mechanical in effect and more like perfect manual playing," and Swift's solution was to increase the lengths of the perforations so that they would overlap each other by a certain amount:

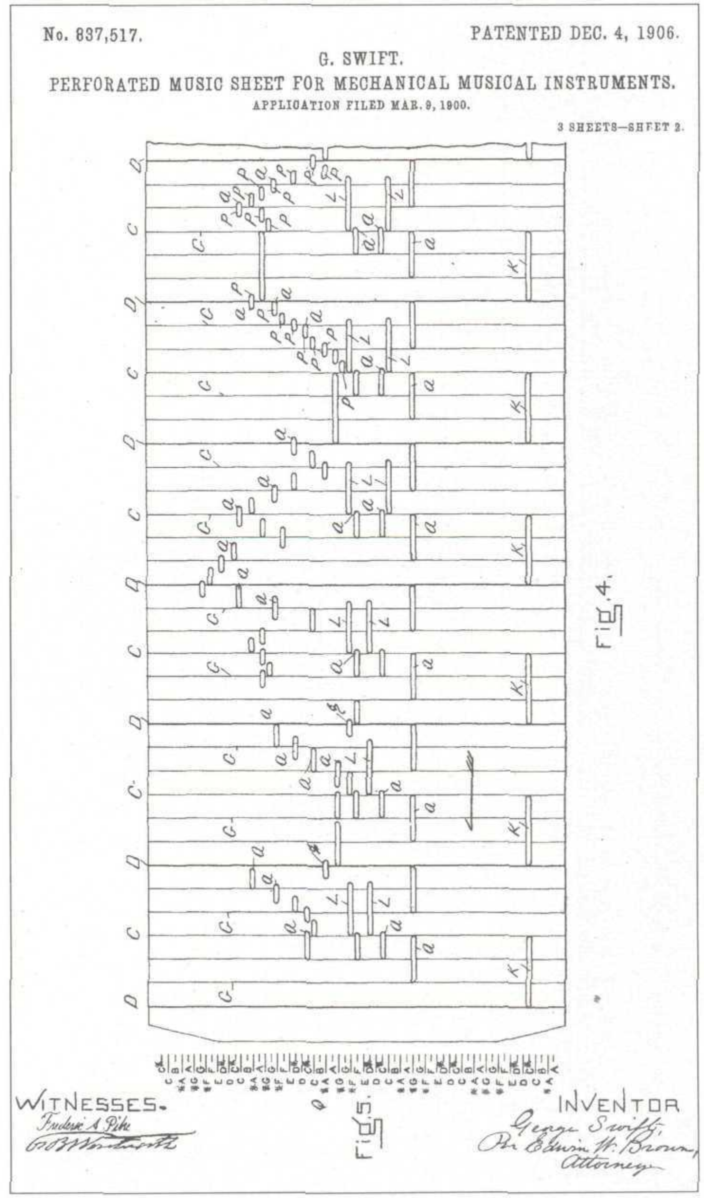
> My discovery which, following on long experiment, produced my present invention was as follows: The prolongation of the several note-perforations of the legato melody in any particular part of the composition or in all legato parts of the composition produces a smoothness of sound and melodiousness without any such dissonances as might be inferred from the resulting overlapping of succeeding note-perforations.... By this simple difference the defect in the music otherwise produced is at once both eliminated and also demonstrated.

Swift illustrated his idea with a pair of plates showing Chopin's "Berceuse" encoded onto piano rolls in two different ways: first, with the usual precision; and second, according to his new technique with its overlapping perforations. In track thirteen, both plates are presented twice each with stereo channels separated one octave, and with both channels raised one octave the second time around. The two versions do indeed sound very different from each other, with the second having a kind of dreamlike quality about it—perhaps enabling us to experience something of the aesthetic effect Swift had in mind.

((( )))

Source: George Swift, "Perforated Music Sheet for Mechanical Musical Instruments," U. S. Patent 837,517, filed 9 March 1900, issued 4 December 1906.





"Perforated Music Sheet for Mechanical Musical Instruments," U. S. Patent 837,517, Sheets 1 and 2.

(((((((14. AU CLAIR DE LA LUNE—S'IL FAUT QU'À CE RIVAL—VOLE, PETITE ABEILLE (1860))))))))

(5:34)

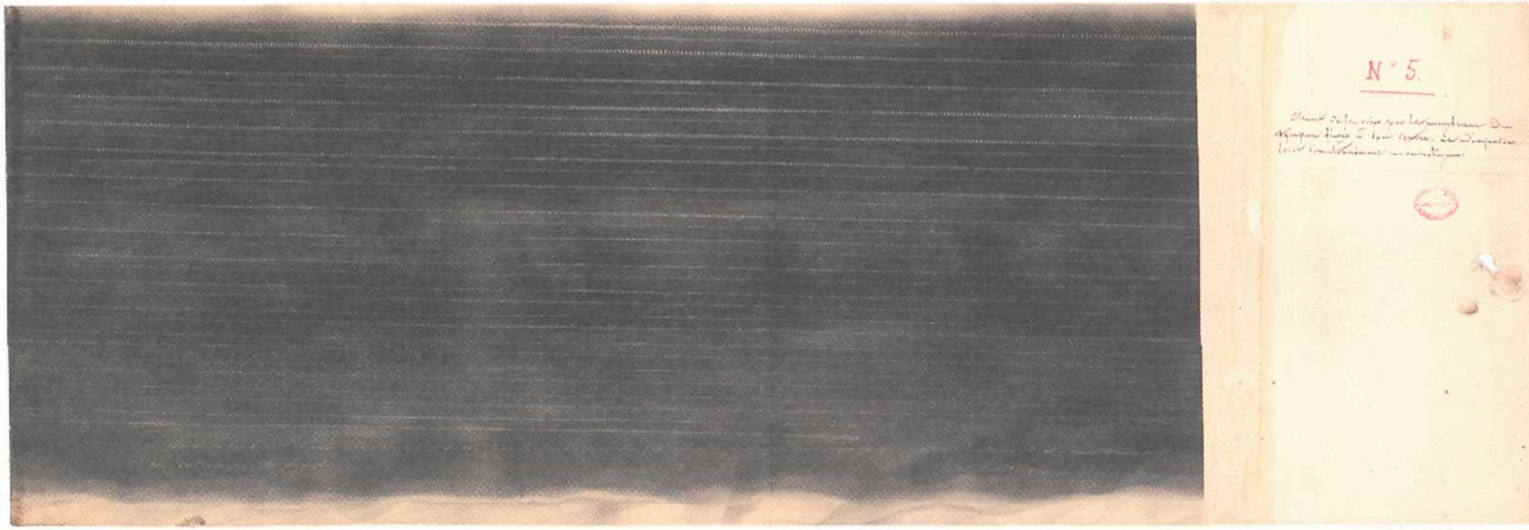
In January 1857, a typographer named Édouard-Léon Scott de Martinville deposited a sealed packet with the Académie des Sciences in Paris describing a revolutionary technique for making automatic records of speech and other sounds. He opened his account with a string of rhetorical questions:

> Is there a possibility of reaching, in the case of sound, a result analogous to that attained at present for light by photographic processes? Can it be hoped that the day is near when the musical phrase, escaped from the singer's lips, will be written by itself and as if without the musician's knowledge on a docile paper and leave an imperishable trace of those fugitive melodies which the memory no longer finds when it seeks them?.... Will one be able to preserve for the future generation some features of the diction of one of those eminent actors, those grand artists who die without leaving behind them the faintest trace of their genius?[47]

Scott's answer to all these questions was *yes*. He knew that the human ear picks up sound vibrations from the air and passes them along to the auditory nerve so that we can hear them, so he reasoned that a suitably constructed artificial ear ought to be capable of *recording* those same vibrations—and through them, any audible sound whatsoever. His plans underwent a number of technical modifications over the next few years, but in their simplest and most consistent form, they involved using a funnel to concentrate sound waves onto an eardrum-like membrane with a stylus attached to its underside that trailed against a moving surface blackened with soot from an oil lamp. In this way, sound waves passing through the air caused the membrane to move back and forth, and the stylus left behind a visible trace of these motions in the soot. Scott's invention, which he called the *phonautograph*, was the first mechanism in the world that recorded aerial sound waves over time. Its inscriptions were known in turn as *phonautograms*.

Scott's ultimate goal in recording sound was not to play it back, but to represent it on paper in a visually intelligible form as an improvement on ordinary written language:

> Our current writing expresses but one only of the modes according to which the voice represents thought; it is suited to representing nothing but articulation. Natural writing or stenography, of which here are the first rudiments, returns the rhythm, the expression thereof: it is a function of tonality, of intensity, of timbre, of measure. For this reason it is called to play a new and unforeseen role in the relations of intellectual life; it will be living speech; our manual or printed calligraphy is nothing but dead speech.[48]



"Au Clair de la Lune."

Top image: Scott #36, 9 April 1860, First Sounds Facsimile 6, p. 10; *Bottom image:* Scott #44, 20 April 1860, First Sounds Facsimile 5, p. 6.

According to Scott's vision, the people of the future would learn to read phonautograms by eye, perhaps enabling them to "hear" the recorded content in their minds' ears. He supposed that much of the value of recorded sound would lie in its ability to capture the nuances of great performances of speech and music, and this particular part of his dream closely resembles what commentators would say about Edison's phonograph a couple decades later. Nevertheless, Scott's assumptions about how people would *access* such material differed greatly from those of later periods, and they also had some significant repercussions on the form his actual recordings took.

Today, we're used to audio media having consistent speeds: vinyl discs that rotate at 33⅓ or 45 revolutions per minute, magnetic audiotape that runs at 7½ or 15 inches per second, digital sound files that play at 44,100 or 48,000 samples per second, and so forth. It's that consistency that enables us to play such media at a "correct" speed. But Scott turned the drum of his phonautograph directly by hand while recording, so its speed of rotation varied erratically from moment to moment. At first, this left him with no reliable means of measuring time or frequency: he could study the amplitude and shape of the recorded waveforms, but that was about it. This technical limitation was finally overcome in 1859 through the addition to the phonautograph of a 250 Hz tuning-fork that could record its own vibrations as a second trace parallel to the main one, furnishing timecode that could be used to measure how much time had elapsed between different events documented in the main trace.

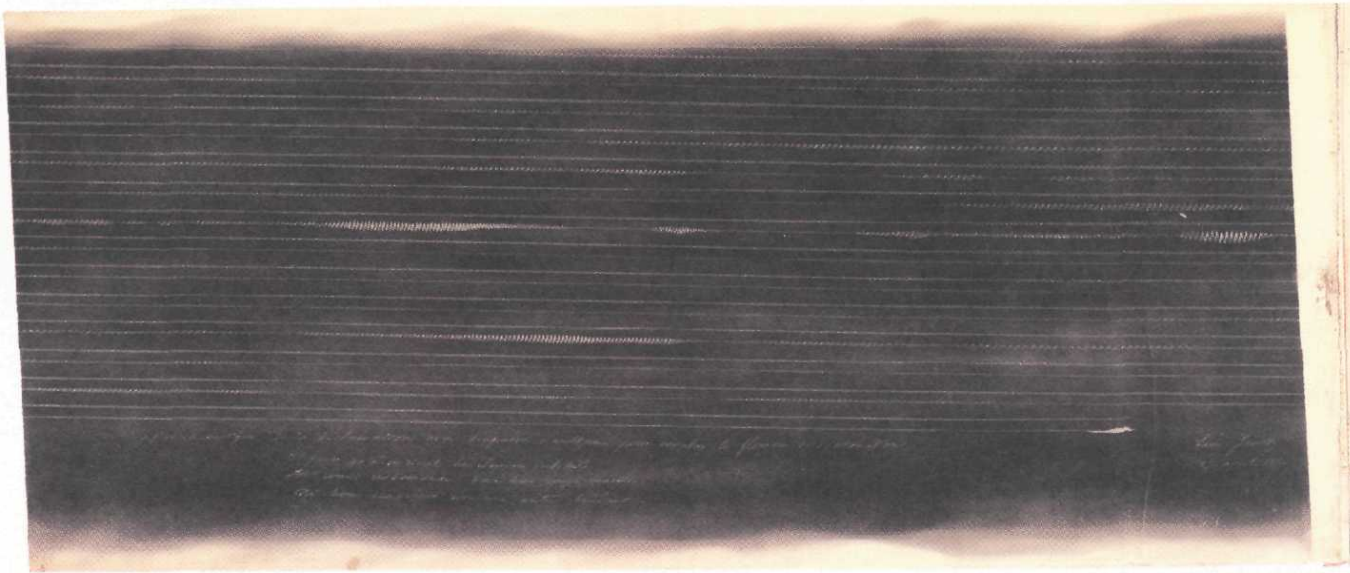
Scott's phonautograph provided an important source of inspiration for later developments in recorded sound, including the invention of the gramophone by Emile Berliner, who referred to his own recordings at first as "phonautograms." It also retained its value as a scientific instrument for many years, though its form varied considerably over time. As I mentioned in my notes on track three, Edward Wheeler Scripture used a phonautograph to record "speech curves" for his article of 1916 on the speech of patients with multiple sclerosis. However, he also entertained some doubts about its accuracy:

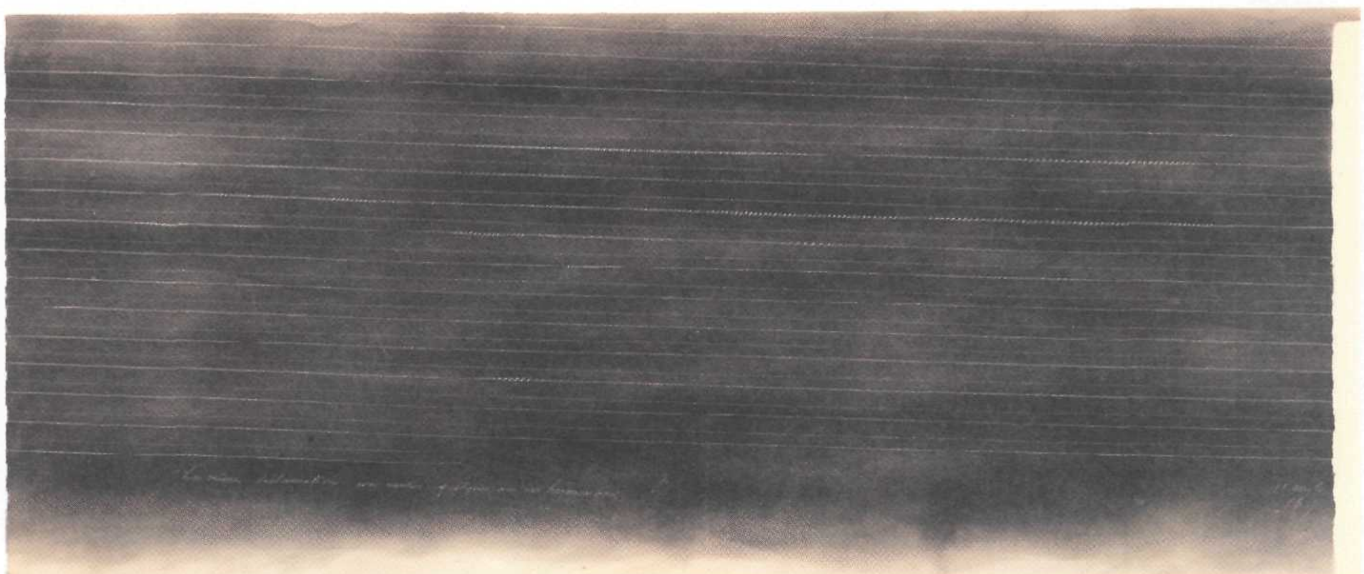
> [T]here is no guarantee whatever that the apparatus records correctly. The curve can not be turned back into sound so that the success of the record can be judged by the ear.... A phonautograph recorder could be trusted only if it could be proved that it gives the proper curves. This might be done by turning its curves back into speech.[49]

The proof Scripture had in mind was a long time in coming. In 2008, however, recognizable sound was recovered from one of Scott's own phonautograms for the first time in history through the efforts of the First Sounds initiative. David Giovannoni traveled to Paris to make high-resolution scans of several specimens of Scott's work in the archives of the Académie des Sciences, and he chose one of them—a seemingly well-recorded phonautogram of the song "Au Clair de la Lune," dated 9 April 1860—as the most promising candidate for playback. After some laborious graphic processing, he handed off a set of images to Carl Haber and Earl Cornell at Lawrence Berkeley National Laboratory, who carried out the crucial step of converting the digital images into digital sound files by using a "virtual stylus" to "track" the waveforms just as though they were grooves on a vinyl LP. Then, as a final step, I used the tuning-fork timecode to speed-correct the raw audio into a recognizable vocal rendition of the song. For the first time in living memory, we could listen to an actual rendition of vocal music performed before the outbreak of the American Civil War! That was big news, and the audio we released was picked up by broadcast media around the world.[50] (Only some months later did we realize that we'd played it back at twice the correct speed, turning a deep male voice—which we've identified as Scott's own—into that of a "young girl.")

Tracks fourteen, fifteen, and sixteen consist of audio I've recovered from Scott's phonautograms since then using my "optical film sound track" technique. One advantage this technique enjoys here over the "virtual stylus" is that it doesn't actually need to "track" the waveforms, which makes it more forgiving of certain types of distortion, such as instances where the trace loops back on itself along the time axis. In tracks fourteen and fifteen, I've presented each phonautogram three times:

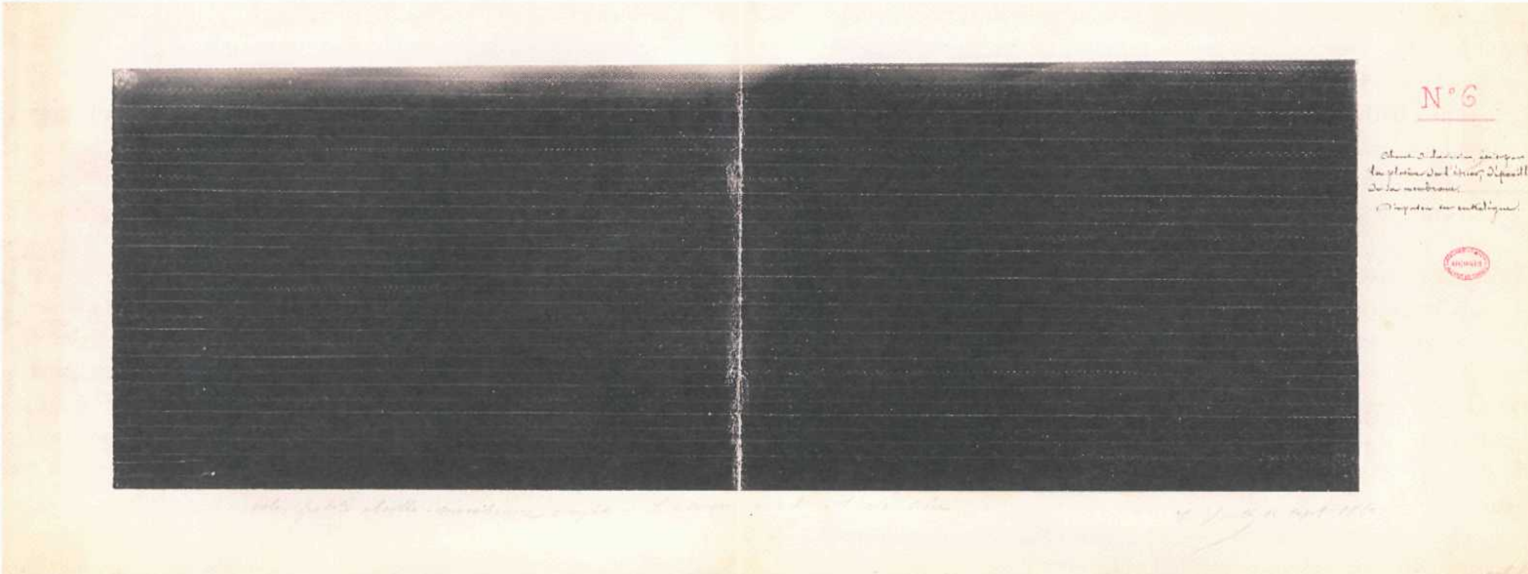
1. **Without speed correction, with voice and tuning fork in stereo.** What you hear first was educed straight off the page at a constant linear speed of 18.375 inches per second—that's simply what we get when we take scans at 2400 dots per inch, translate each column of dots into an audio sample, and play those samples at the audio CD standard rate of 44,100 per second.
2. **With speed correction, with voice and tuning fork in stereo.** Now

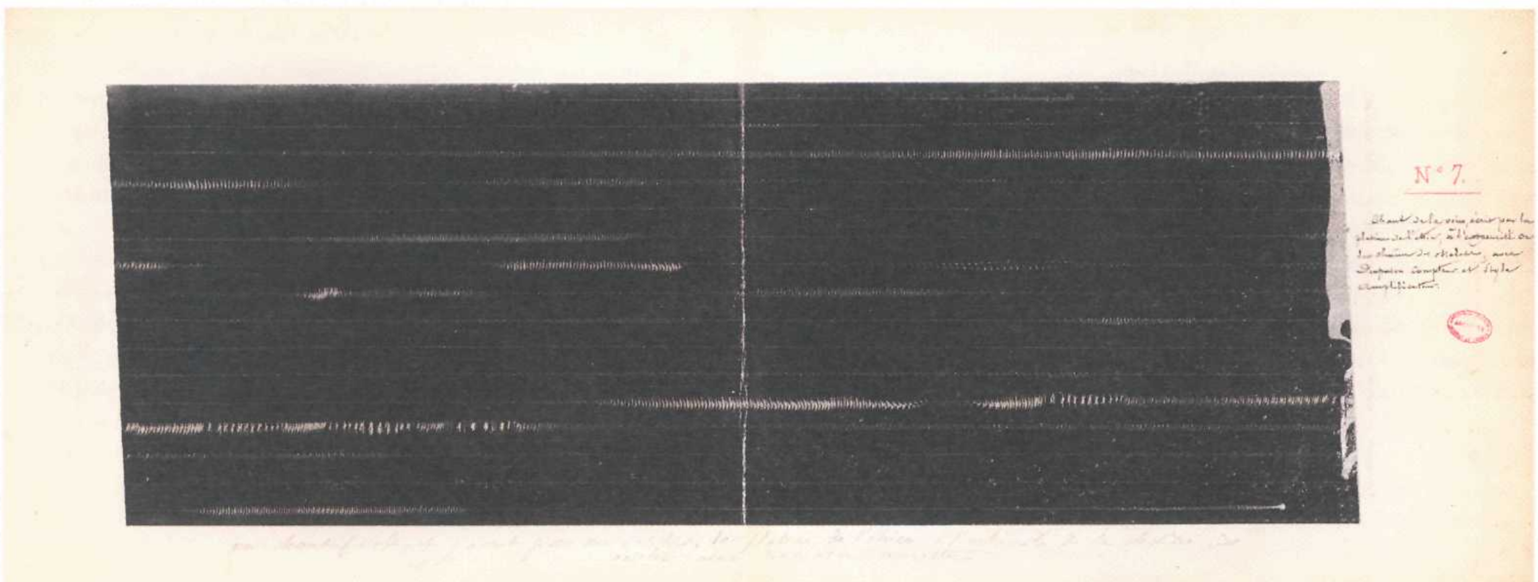




"S'il faut qu'à ce rival."

Top image: Scott #38, 17 April 1860, First Sounds Facsimile 5, p. 8; *Bottom image:* Scott #41, 18 April 1860, First Sounds Facsimile 5, p. 9.





"Vole, petite abeille."

Top image: Scott #49, 15 September 1860, First Sounds Facsimile 6, p. 11; *Bottom image:* Scott #50, ca. late September 1860, First Sounds Facsimile 6, p. 12.

I've manually assigned a duration of twenty milliseconds to each group of five wave periods in the tuning-fork track, restoring the tuning fork to its original constant frequency of 250 Hz and thereby correcting both tracks for fluctuations in recording speed simultaneously. First, you hear both speed-corrected tracks together.

3. **With speed correction, voice only, in mono.** Finally, you hear the speed-corrected voice track by itself, without the tuning fork.

Track fourteen contains three examples of phonautograms that exist in pairs—that is, instances in which we have two different phonautograms of essentially the same subject matter. As you listen to each example in turn, try to make sense of the first version you hear of each phonautogram—the one without speed correction. You can take advantage of the repetition in subject matter to help you do this. Of course, you'll be listening "cold" to the first phonautogram in each pair, but you'll have some idea what to expect when you hear the second one because you'll have heard the speed-corrected audio from a very similar phonautogram just a moment before. See if you can use these examples to train your ear to compensate for the speed fluctuations, and to listen through the "wow" to the spoken or musical content lurking behind it.

Our first two phonautograms feature the song "Au Clair de la Lune," one dated 9 April 1860 and the other dated 20 April 1860. Scott sings very slowly both times, probably in order to create a record that would be visually easy to parse—although this is also the speed at which the song would have been sung as a lullaby. The take dated 9 April is Scott's oldest known recording with timecode, which means that it's the oldest example of his work which we can reliably speed-correct today.

The subject of our second pair of phonautograms is a dramatic recitation from Jean-François Ducis's *Othello*, a French play based rather loosely on Shakespeare's original:

> *S'il faut qu'à ce rival hédelmone infidèle*
> *Ait remis ce bandeau | dans leur rage cruelle*
> *Nos lions du désert sous leur antre brûlant—*

> So it must be that to this rival faithless Hédelmone
> Gave this diadem | in their cruel rage
> Our lions of the desert under their burning lair—

Scott used this passage repeatedly as a test recitation over the course of his experiments, but these two examples date respectively from 17 and 18 April 1860. The take dated 17 April is Scott's oldest known recording of spoken language with timecode.

In our third pair of phonautograms, Scott sings a melody he identifies in his inscriptions as "Vole, petite abeille" ["Fly, little bee"], but which is more properly the "Chanson de l'Abeille" ["Song of the Bee"] from the comic opera *La reine Topaze* [*Queen Topaz*] by Victor Massé, first performed in 1856. These are Scott's last known phonautograms, dating from mid-to-late September 1860. They differ technically from the other phonautograms in this track in that Scott recorded them using an artificial chain of ossicles in addition to the artificial eardrum. The second example has two other noteworthy features besides. First, Scott had added an "amplifying lever" to the end of the artificial chain of ossicles, which might explain why this recording sounds so different from the other one. Second, the tuning fork traced its vibrations mostly not as a wavy line, but as a series of dots—presumably because it was vibrating *vertically* rather than *laterally* with respect to the surface of the cylinder. In practice, a row of dots can be interpreted as a band of varying width with a width of zero in the spaces between the dots, so I ran the trace through ImageToSound as is, without the usual "fill" step described in the notes on track three. In the notes that follow, I'll refer to this strategy as *vertical* eduction, as opposed to *lateral* eduction.

((( )))

Source: Archives of the Académie des Sciences; Library of the Institut de France; scans by David Giovannoni et al., FirstSounds.org.
a. "Au Clair de la Lune," Scott #36 (Académie, plate 5), 9 April 1860
b. "Au Clair de la Lune," Scott #44 (Institut, leaf [5r]), 20 April 1860
c. "S'il faut qu'à ce rival," Scott #38 (Institut, leaf [6]), 17 April 1860
d. "S'il faut qu'à ce rival," Scott #41 (Institut, leaf [7]), 18 April 1860
e. "Vole, petite abeille," Scott #49 (Académie, plate 6), 15 September 1860
f. "Vole, petite abeille," Scott #50 (Académie, plate 7), ca. late September 1860

(((((((15. GAMME DE LA VOIX—ÉPELLATION DU MOT "RIRA"— LE JOUR N'EST PAS PLUS PUR—CHI CREDERIA CHE SOTTO FORME UMANE—ET INCARNATUS EST (1860))))))))

(4:26)

Track fifteen is structured in much the same way as track fourteen, but this time each selection is represented by just one phonautogram rather than two.

First we hear Scott singing an ascending vocal scale, *do, re, mi, fa, sol, la, ti, do*, probably at the beginning of September 1860. As with the final example in the previous track, the timecode takes the form of a series of dots, so I've educed it laterally.

In the second example, Scott spells the word *rira* ["will laugh"] three times in a row: *R, I, RI, R, A, RA, RIRA.* The breakdown by syllables and the use of the name "re" for the letter *r* is consistent with *la nouvelle épellation* ("the new spelling"), a proposed educational reform that was never widely adopted. This phonautogram was recorded on 18 April 1860, and Scott's stated purpose in making it was to study the combination of vowels with the *r grasseyé* or uvular *r*.

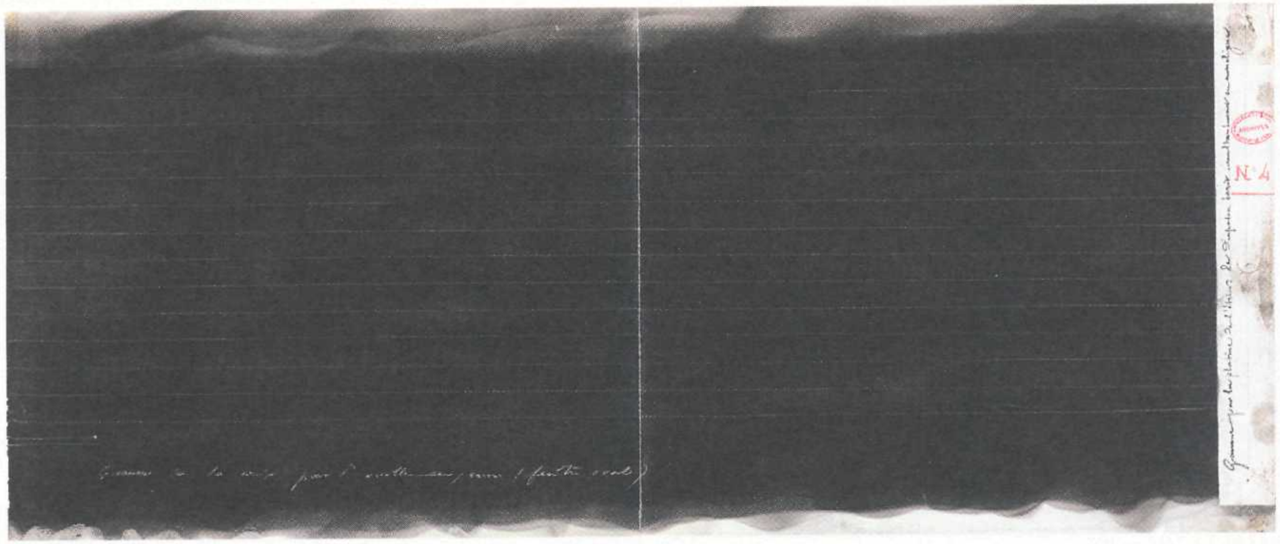
Our third example features Scott reciting a line from Jean Racine's *Phèdre* on 19 April 1860: "Le jour n'est pas plus pur que le fond de mon cœur" ["The day is no more pure than the depth of my heart"]. Unlike Scott's dramatic rendition of the passage from Ducis's *Othello*, this recording appears to have been intended merely as a phonemic study of syllables uttered slowly and deliberately.

Next comes a recitation in Italian: the opening lines of Torquato Tasso's pastoral drama *Aminta*, probably recorded in April or May 1860. Scott wrote the following notes at the bottom of the phonautogram:

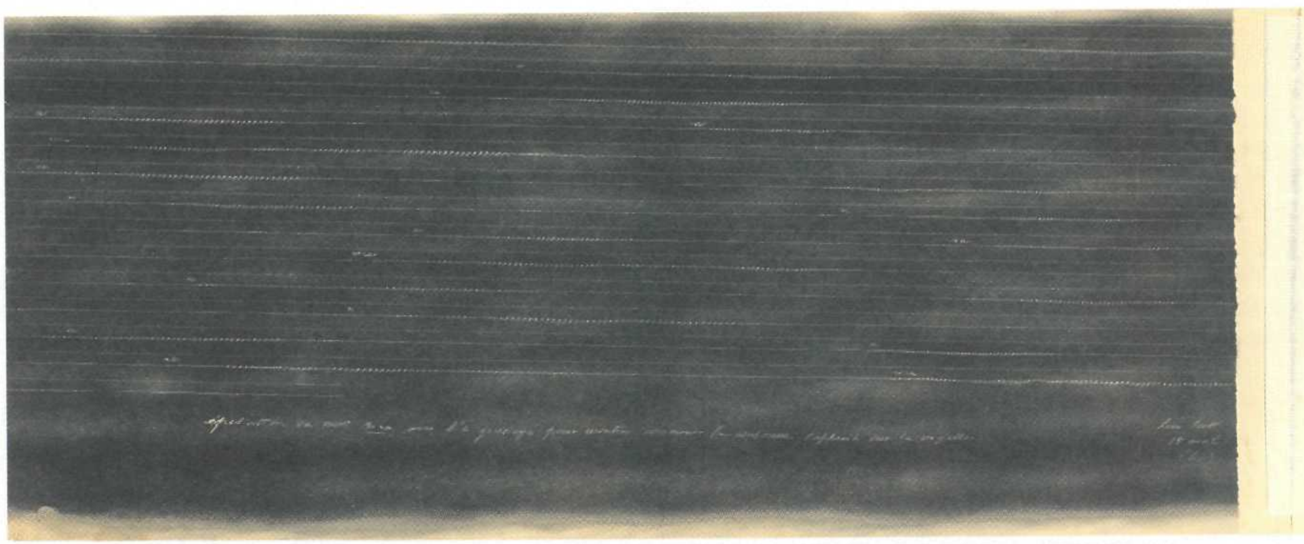
> *étude sur l'accent tonique à la requête de M. le professeur Regnault*
> *Chi crederia che sotto forme umane* e sotto queste pastorali spoglie fosse nascosto un Dio? non mi—*
>
> **je me suis trompé | il faudrait umane forme*

> Study of the tonic accent at the request of Professor Regnault. *Who would believe that under a form human* and under this pastoral garb would be found a God? Not onl—*
>
> *I was wrong | it should be *human form*

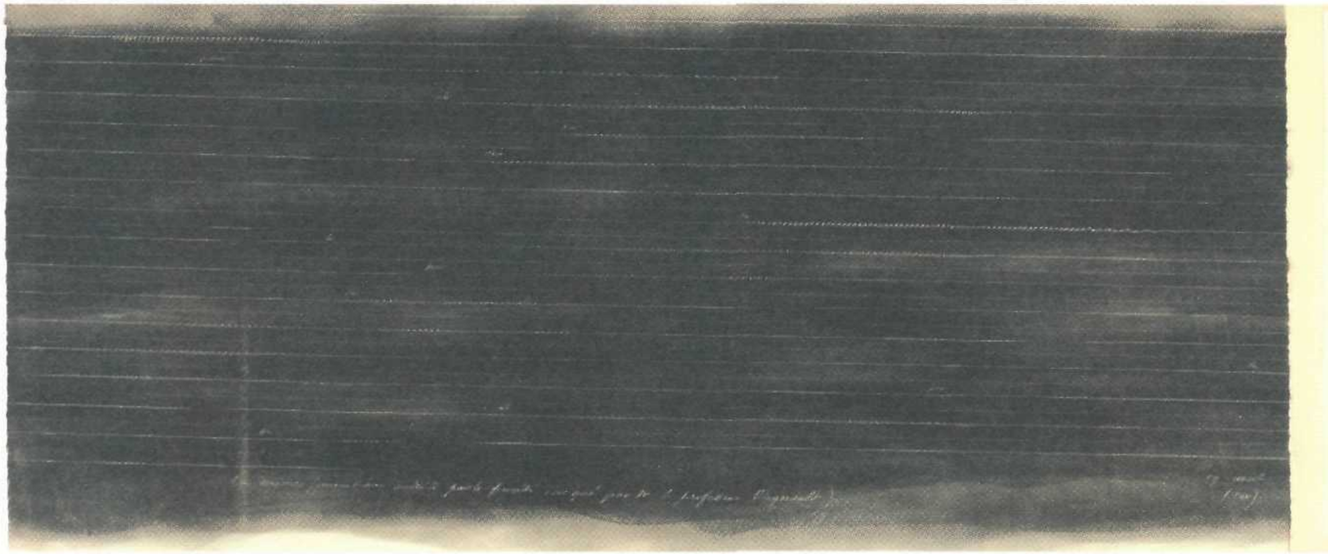
The speaker in the recording can be heard saying "forme umane" rather than "umane forme," so by taking responsibility for this mistake in his note at the bottom, Scott indirectly identifies himself as the speaker. And because this sounds like the same voice heard on Scott's other phonautograms, we



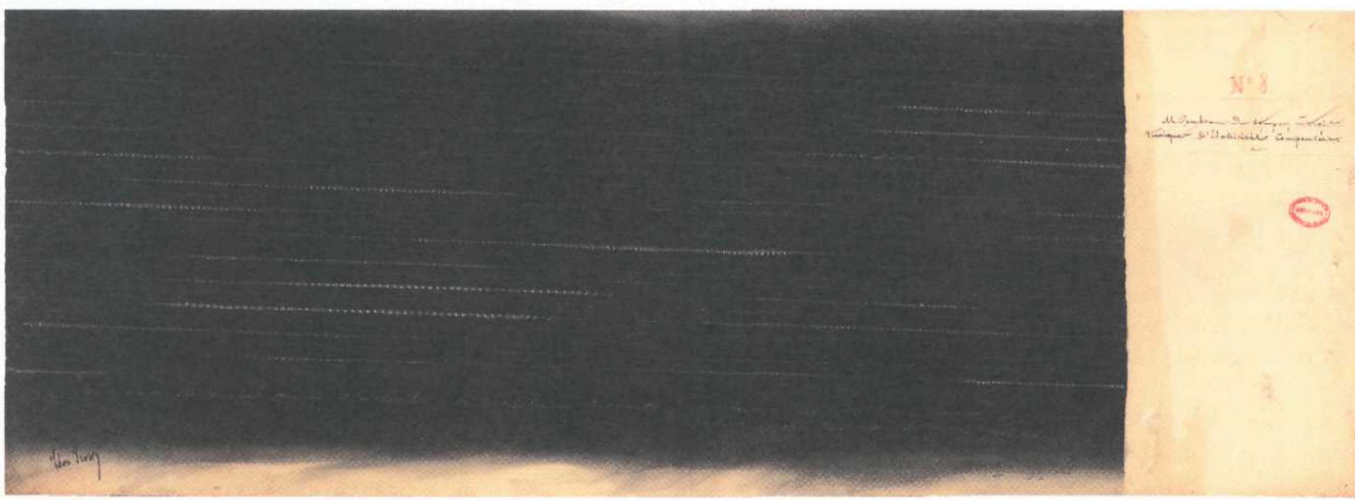
"Gamme de la Voix,"
Scott #48, ca. early September 1860, First Sounds Facsimile 6, p. 9.



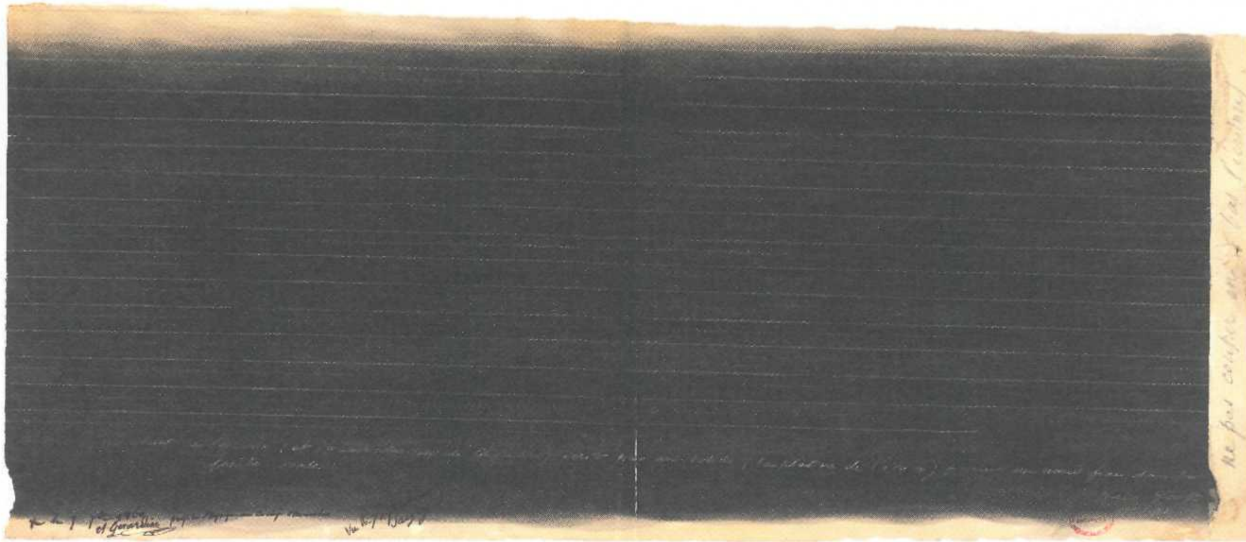
"Épellation du mot 'rira',"
Scott #42, 18 April 1860, First Sounds Facsimile 5, p. 10.



"Le jour n'est pas plus pur,"
Scott #43, 19 April 1860, First Sounds Facsimile 5, p. 11.



"Chi crederia che sotto forme umane,"
Scott #45, ca. April-May 1860, First Sounds Facsimile 6, p. 13.



"Et incarnatus est,"
Scott #47, 1 September 1860, First Sounds Facsimile 6, p. 15.

believe we're hearing Scott himself in those other cases as well.

The subject of our final example is a vocal rendition of "Et Incarnatus Est" by Luigi Cherubini, probably taken from the *Missa Solemnis* in D minor, as recorded on 1 September 1860.

Source: Archives of the Académie des Sciences; Library of the Institut de France; scans by David Giovannoni et al., FirstSounds.org.

a. "Gamme de la Voix," Scott #48 (Académie, plate 4), ca. early September 1860
b. "Épellation du mot 'rira'," Scott #42 (Institut, leaf [8]), 18 April 1860
c. "Le jour n'est pas plus pur," Scott #43 (Institut, leaf [9]), 19 April 1860
d. "Chi crederia che sotto forme umane," Scott #45 (Académie, plate 8), ca. April-May 1860
e. "Et incarnatus est," Scott #47 (Académie, plate 9), 1 September 1860

(((((((16. PHONAUTOGRAMS FROM 1857)))))))

(4:28)

The phonautograms we've been listening to so far all date from the year 1860, and they all contain tuning-fork timecode, which is why we can restore them to their original speeds and "reproduce" the sounds they document with at least some degree of fidelity.

But these aren't actually the oldest extant records made by aerial sound waves. Over thirty phonautograms which Scott recorded during the year 1857 survive as well, in the archives of the Société d'Encouragement pour l'Industrie Nationale.[51] Unfortunately, they lack the all-important timecode, so we can't speed-correct them—at least, not in any objective way. Track sixteen presents several of these earlier phonautograms without speed correction, and this is where your ear-training in tracks fourteen and fifteen might pay off. Can you make out what you're hearing here, even just a little?

We begin with a phonautogram of singing which Scott made in July 1857—the oldest recording he's known to have created as a long, continuous trace around a cylinder rather than as a split-second pass across a flat glass plate. It contains two separate parallel recordings, both educed here vertically.

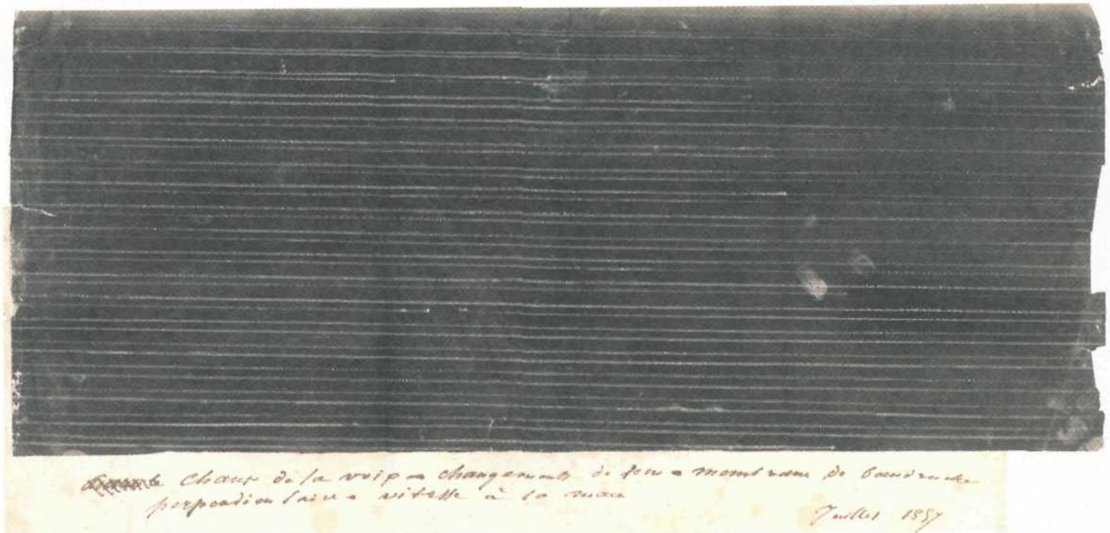
The second and third phonautograms featured in this track date from sometime in the fall of 1857, and both likewise contain two separate parallel recordings each. The first of these two phonautograms is distinctive for having been made not with a membrane as usual, but by a stylus attached directly to a pipe made of ash wood.

Example number four is a phonautogram of the complete Lord's Prayer, dated October 1857. Scott's stylus seems to have gone utterly haywire while he was recording it, but this is his oldest known recording of continuous spoken language, so I've tried to educe it anyway—laterally in the left channel, and vertically in the right.

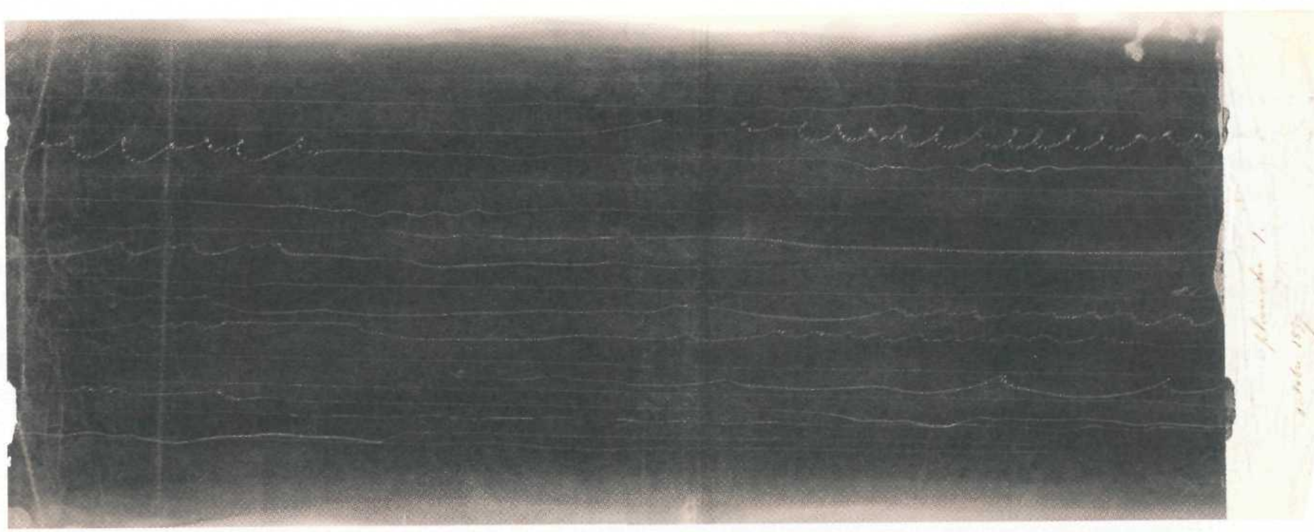
Next comes a "study of the timbre of the voice" recorded around November 1857 (I've left this eduction in the original stereo so that you can hear what that sounds like), followed by a brief recording of a "deep voice."

The final example in this track is dated 17 August 1857, which makes it the earliest phonautogram we can pinpoint to a specific day. Scott identified its contents in writing as "song at a distance," and two enigmatic inscriptions also appear at the beginning and end of the trace itself: "jeune jouvencelle" ["young little girl"] and "les échos" ["the echoes"].[52]

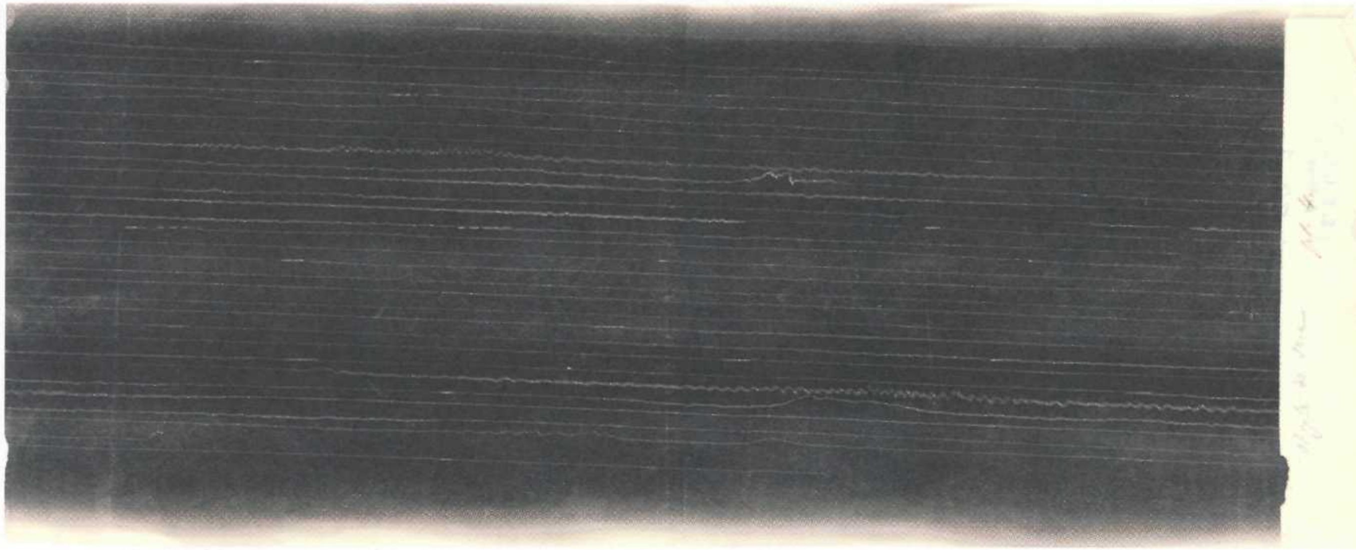
The phonautograms in this track are all technically records made by sound, and they can all be educed. But do they enable us to "reproduce" sounds from the past in any meaningful way? You'll have to judge that for yourself.



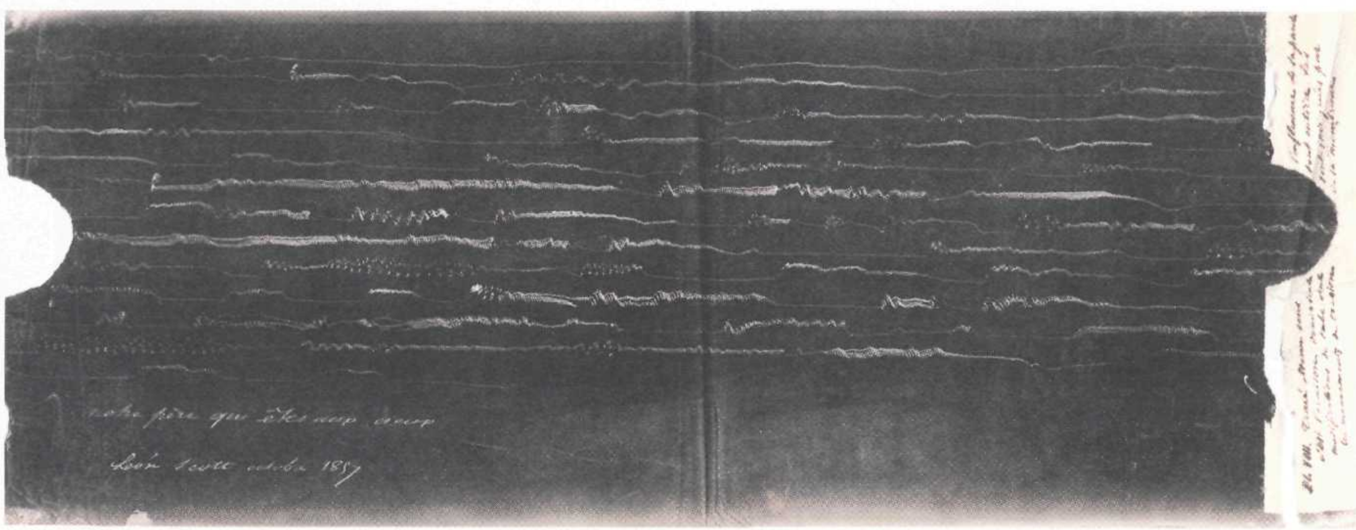
"Chant de la voix—changements de ton" ["Song of the voice—changes of tone"],
Scott #4, July 1857, First Sounds Facsimile 4, p. 3.



"Tuyau de frêne" ["ashen pipe"],
Scott #10, ca. August-October 1857, First Sounds Facsimile 4, p. 9.

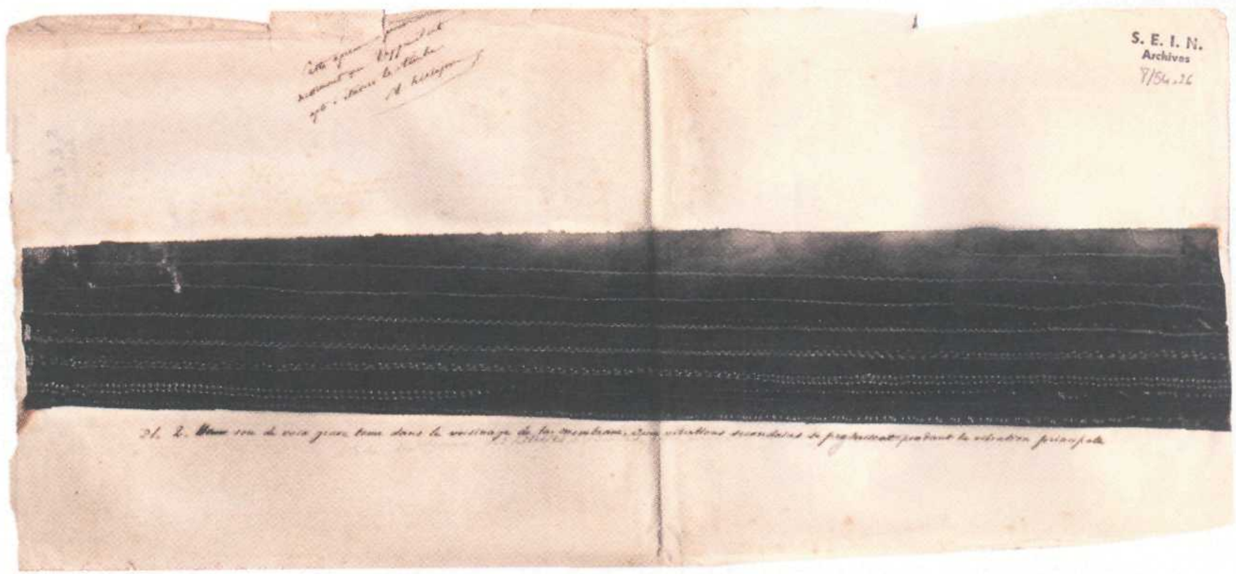


"Style de soie" ["stylus of bristle"],
Scott #12, ca. August-October 1857, First Sounds Facsimile 4, p. 10.

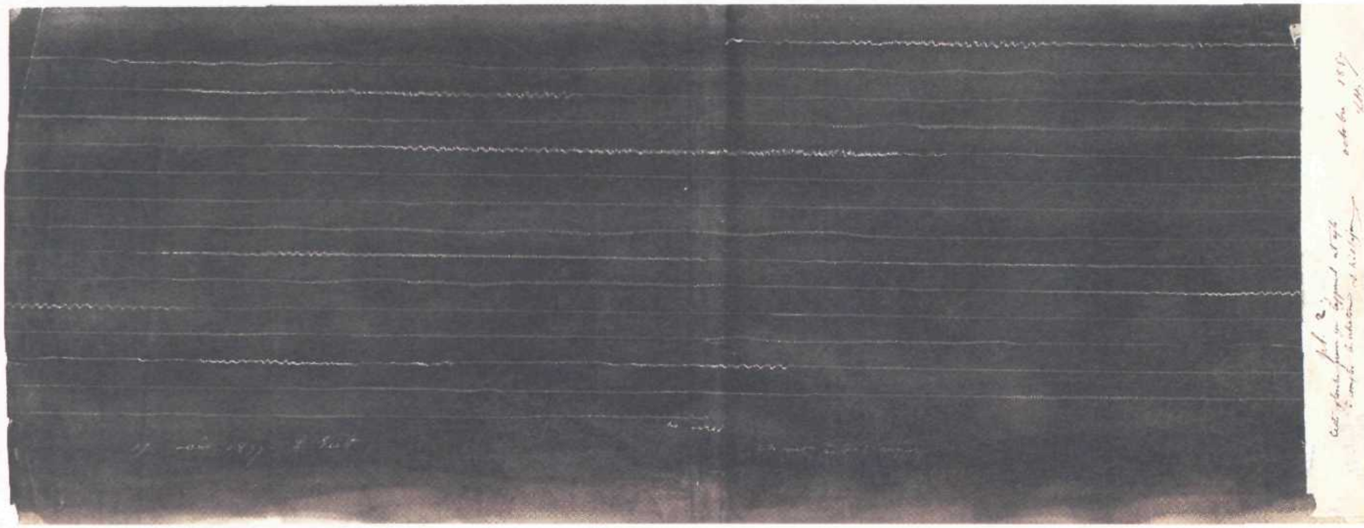


"Notre père qui êtes aux cieux" ["our father who art in heaven"],
Scott #20, October 1857, First Sounds Facsimile 4, p. 12.

"Étude sur le timbre de la voix" ["study of the timbre of the voice"],
Scott #25, ca. November 1857, First Sounds Facsimile 4, p. 27.



"Un son de voix grave" ["a sound of deep voice"],
Scott #14, ca. October 1857, First Sounds Facsimile 4, p. 31.



"Jeune jouvencelle" ["**young little girl**"],
Scott #9, 17 August 1857, First Sounds Facsimile 4, p. 32.

(((✺)))

Source: Archives of the Société d'Encouragement pour l'Industrie Nationale; scans by FirstSounds.org

a. "Chant de la voix—changements de ton" ["Song of the voice—changes of tone"], Scott #4 (SEIN 8/54-3), July 1857, two separate traces, educed vertically
b. "Tuyau de frêne" ["ashen pipe"], Scott #10 (SEIN 8/54-7), ca. August-October 1857, two separate traces, educed laterally from 600 dpi scan (1:13)
c. "Style de soie" ["stylus of bristle"], Scott #12 (SEIN 8/54-8), ca. August-October 1857, two separate traces, educed laterally from 600 dpi scan (1:50)
d. "Notre père qui êtes aux cieux" ["our father who art in heaven"], Scott #20 (SEIN 8/54-10), October 1857, left channel lateral, right channel vertical (2:48)
e. "Étude sur le timbre de la voix" ["study of the timbre of the voice"], Scott #25, ca. November 1857 (SEIN 8/54-22), lateral, stereo (3:13)
f. "Un son de voix grave" ["a sound of deep voice"], Scott #14 (SEIN 8/54-26), ca. October 1857, educed laterally but vertically at very end (3:42)
g. "Jeune jouvencelle" ["young little girl"], Scott #9 (SEIN 8/54-27), 17 August 1857, lateral (3:57)

(((((((17. MAGIA PHONOTACTICA)))))))

(3:58)

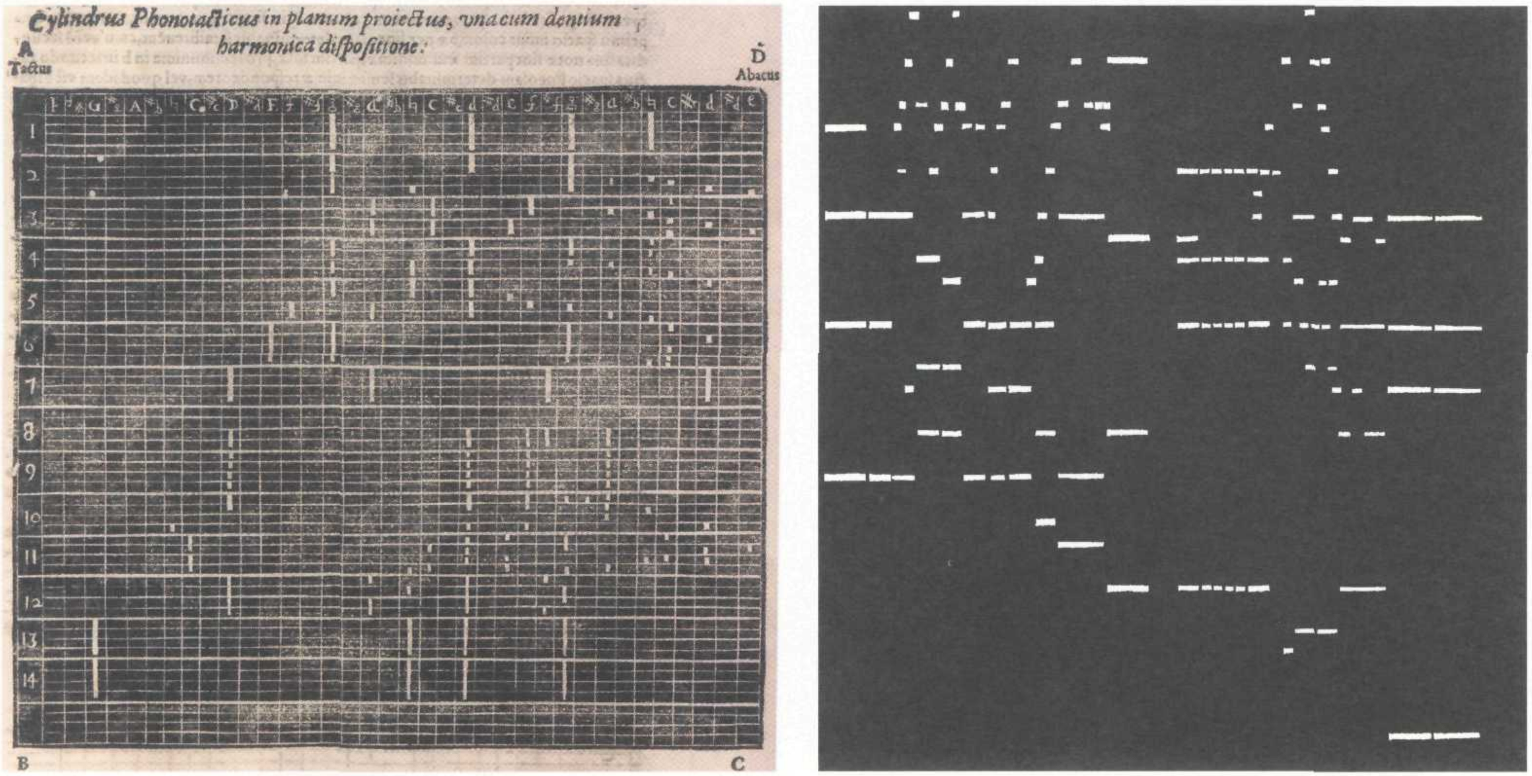
The seventeenth-century Jesuit scholar Athanasius Kircher was a man of wide-ranging interests, equally at home in the investigation of Egyptian hieroglyphs, magnetism, geology, technology, mathematics, and medicine, to cite just a few of the areas of study to which he contributed; the subtitle of one recent book refers to him as "the last man who knew everything."[53] He was also a prolific author, and his books were richly illustrated with engravings that tend to receive more attention from scholars today than his actual writings do—he wrote in Latin, and few of his works have been translated into modern vernacular languages. It's the engravings that will be of greatest interest to us here as well.

Kircher's *Musurgia Universalis,* published in 1650, is an eclectic treatise on music that's said to have influenced such composers as Johann Sebastian Bach and Ludwig van Beethoven. Among other things, it contains a section on automatic musical instruments, including instructions on how to program a *cylindrus dentatus* or toothed cylinder—a medium designed to rotate at constant speed while its teeth engaged a mechanism that controlled the playing of notes on an organ over time. Kircher also referred to this medium as a *cylindrus phonotacticus,* meaning a cylinder for the arrangement of sounds, and he described two ways of encoding a song onto it:

> You can either transfer this song directly onto the cylinder or sketch it out separately on paper beforehand; the latter is not only easier but better places the whole composition before the eyes as though displayed on a plane surface of the cylinder.[54]

In other words, Kircher thought it best to plot out a piece of music in advance on a sheet of paper which represented the cylinder's curved surface "projected into a plane" for ease of viewing. He provided illustrations of five such "projections" in his book, each one representing a different musical program, and all five of these "projections" may be heard educed paleospectrophonically in track seventeen. Each one is presented twice: once in mono, and then once in stereo with the right channel raised up an octave.

Our first three examples are taken from "projections" that closely resemble each other in format and purpose. As to format, Kircher's vertical axis represents time, and his horizontal axis represents frequency, dividing the octave spatially into twelve evenly-spaced units to yield a workable chromatic scale. As to purpose, the first example uses a four-part "cantilena" to introduce the phonotactic method in general; the second uses a "triphonium syncopatum" to demonstrate the encoding of syncopation, in the sense of notes held across bar lines; and the third shows the encoding of a piece in triple time—Kircher identifies it as a composition by Simplicio Todeschi published in 1627. Collectively, these examples aimed to show how to translate various



"Cylindrus Phonotacticus in planum proiectus, vnacum dentium harmonica dispositione"
[Phonotactic cylinder projected into a plane with one harmonious arrangement of teeth.]
(p. 2:315); the same image prepared for eduction (at *right*).

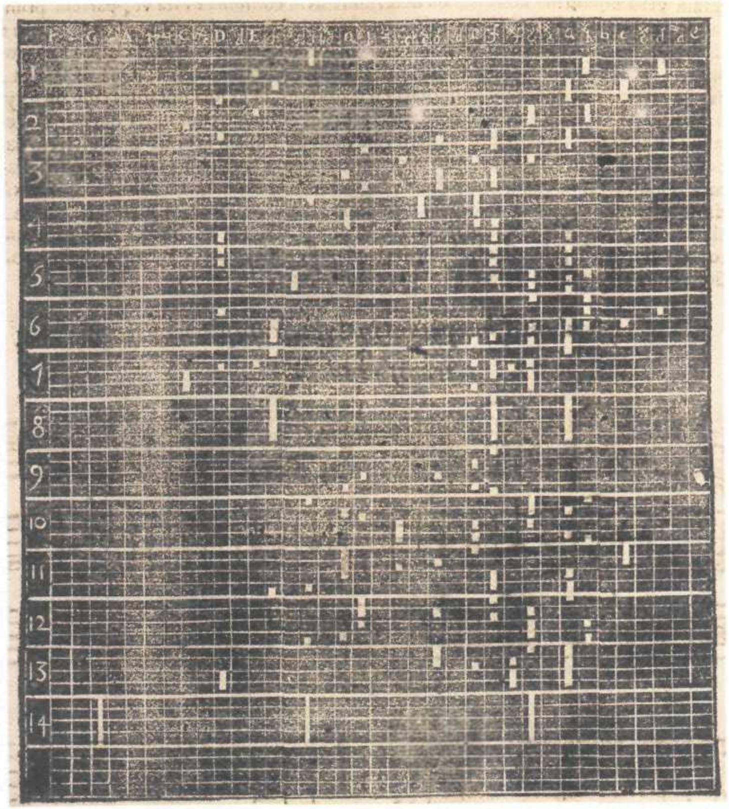
familiar musical features into a machine-readable format.

But our next two examples reflect somewhat grander ambitions for what Kircher's book sometimes called "phonotactic magic" [*magia phonotactica*].[55] Example number four was intended to support the claim that automatic instruments could be made to play impressively rapid passages of music with superhuman accuracy:

> This geometrical division has so great an advantage that no skillful manual playing can approach it; for it can if you wish display all imaginable diminutions of time so perfectly that not a note should escape being heard, not even a sixty-fourth note [*quadricaudata*].[56]

This time, Kircher's frequency axis contains just *eleven* units—it's missing B flat, perhaps by mistake—but when we educe the engraving paleospectrophonically, it still yields something close enough to a chromatic scale to sound reasonably acceptable; and, of course, we can play the trill however rapidly we want.

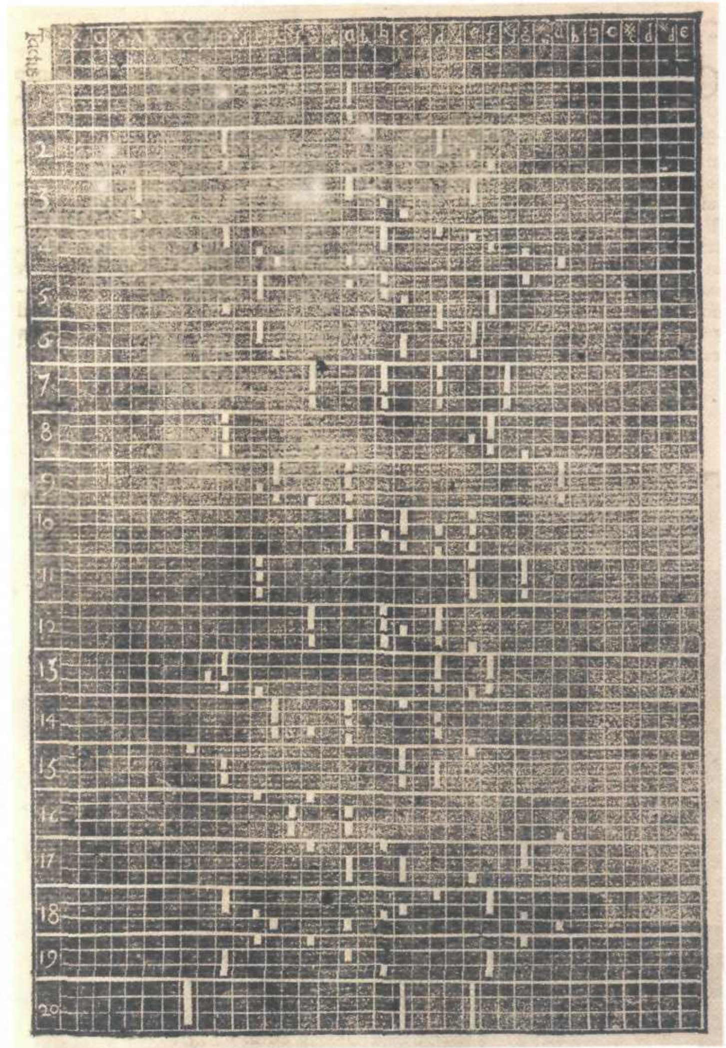
In our fifth and final example, we find the octave divided not into the usual



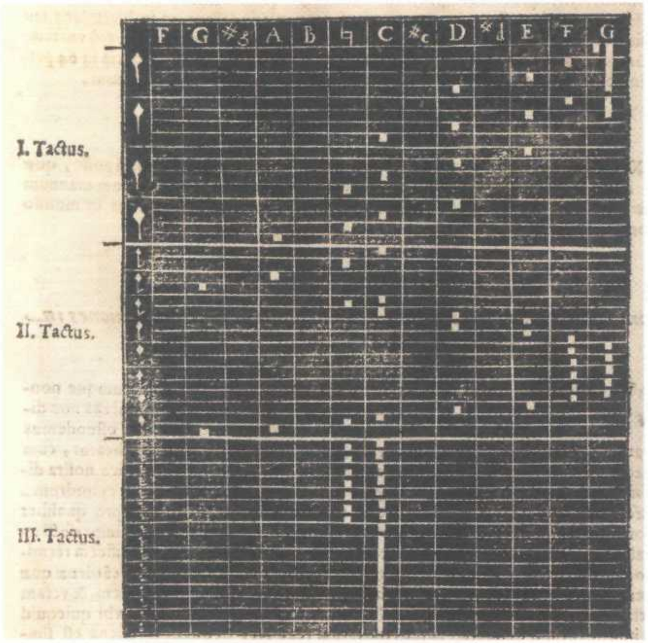
"Cylindrus Phonotacticus in planum proiectus, vna cum dentium harmonica dispositione, qua præcedentis triphonij, voces, notarum situs, & temporum mensura exacte indicantur." [Phonotactic cylinder projected into a plane with one harmonious arrangement of teeth by which the voices, places of notes, and measure of times are exactly indicated.] (p. 2:322).

twelve units, but into *twenty-eight* so that it could accommodate a piece of experimental microtonal music. As Kircher observed, it was difficult for human beings to perform music of this kind on instruments, and nearly impossible for them to perform it vocally:

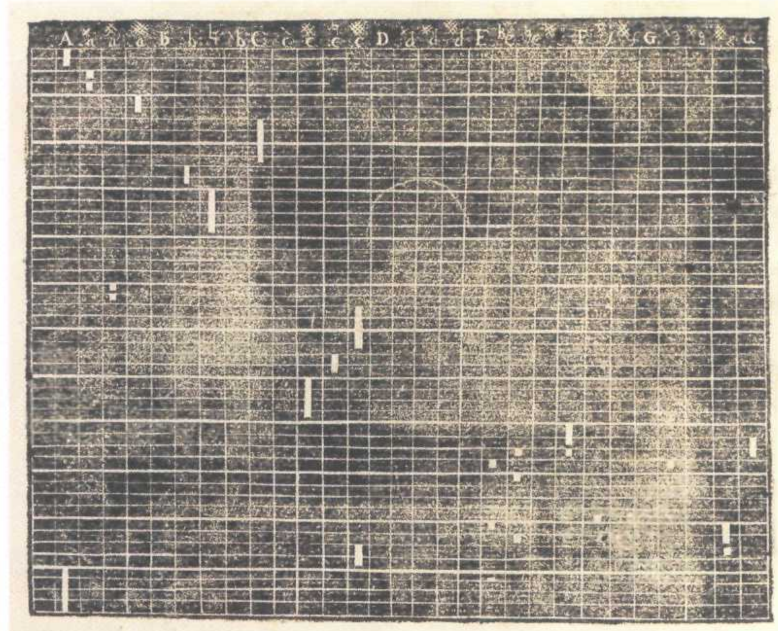
> But on our phonotactic cylinder the thing is entirely easy, to the extent that I think it possible for such genres [*genera*] to be presented



"Tripla in Cylindrum transferenda."
[Triple to be transferred onto cylinder.] (p. 2:327).



"Quadratum Phonotacticum. Teretismos seu Diminutiones notarum, quas vulgò Trillos & Grupos vocant exacta diuisione continens." [Phonotactic Square. Containing *teretismos* or diminutions of notes, which are commonly called *trillos* and *gruppos*, in exact division.] (p. 2:323).



"Monodium Diatonico Chromatico-Enharmonicum in Cylindro phonotactico exhibendum." [Diatonic-chromatic-enharmonic monody to be presented on phonotactic cylinder.] (p. 2:328).

perfectly, exactly, and distinctly in no other way than by automata of this kind, in which all intervals, however small you wish, can be arranged in such exact divisions that a faultless performance might be obtained. In order, therefore, that you might see a particular specimen of this thing, we will show how to transfer this chromatic-enharmonic monody onto a cylinder, as truly nothing abstruse and recondite can be devised within the musical universe that our phonotactic art might not present.[57]

The last two examples both involved dividing one axis of a phonotactic graph into unusually small units, enabling it to represent tiny gradations of time in the one case and tiny gradations of frequency in the other. It was these capabilities that led Kircher to boast that "phonotactic magic" could handle all imaginable forms of music, including some which he believed would exceed the capacity of any human performer. But what was true in principle wasn't necessarily true in practice. Although it was easy enough for Kircher to graph out impossibly difficult pieces of music on paper, he would have found some obstacles in the way of actualizing them as sound: limits to the speed at which the mechanism of an organ could be made to respond, for example, or the inconvenience of constructing a specially-tuned rank of organ pipes to play microtonal music. The technology of his day wouldn't have been able to do justice to the full range of phonotactic inscriptions he could imagine and create.

But paleospectrophony can handle even the smallest gradations of time and frequency which Kircher was capable of graphing out on paper. In that sense, the technique used here to create track seventeen might constitute the ultimate realization of Kircher's own vision.

(((☼)))

Source: Athanasius Kircher, *Musurgia universalis sive ars magna consoni et dissoni in X. libros digesta* (Roma: Grignani Corbelletti, 1650).

a. "Cylindrus Phonotacticus in planum proiectus, vnacum dentium harmonica dispositione" [Phonotactic cylinder projected into a plane with one harmonious arrangement of teeth.] (p. 2:315)

b. "Cylindrus Phonotacticus in planum proiectus, vna cum dentium harmonica dispositione, qua præcedentis triphonij, voces, notarum situs, & temporum mensura exacte indicantur." [Phonotactic cylinder projected into a plane with one harmonious arrangement of teeth by which the voices, places of notes, and measure of times are exactly indicated.] (p. 2:322)

c. "Tripla in Cylindrum transferenda." [Triple to be transferred onto cylinder.] (p. 2:327)

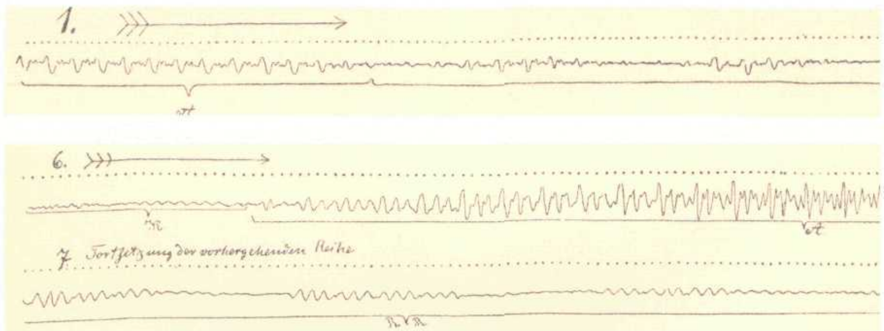
d. "Quadratum Phonotacticum. Teretismos seu Diminutiones notarum, quas vulgò Trillos & Grupos vocant exacta diuisione continens." [Phonotactic Square. Containing *teretismos* or diminutions of notes, which are commonly called *trillos* and *gruppos*, in exact division.] (p. 2:323)

e. "Monodium Diatonico Chromatico-Enharmonicum in Cylindro phonotactico exhibendum." [Diatonic-chromatic-enharmonic monody to be presented on phonotactic cylinder.] (p. 2:328)

Athanasius Kircher.

(((((((18. SPRACHZEICHNER)))))))

(1:06)



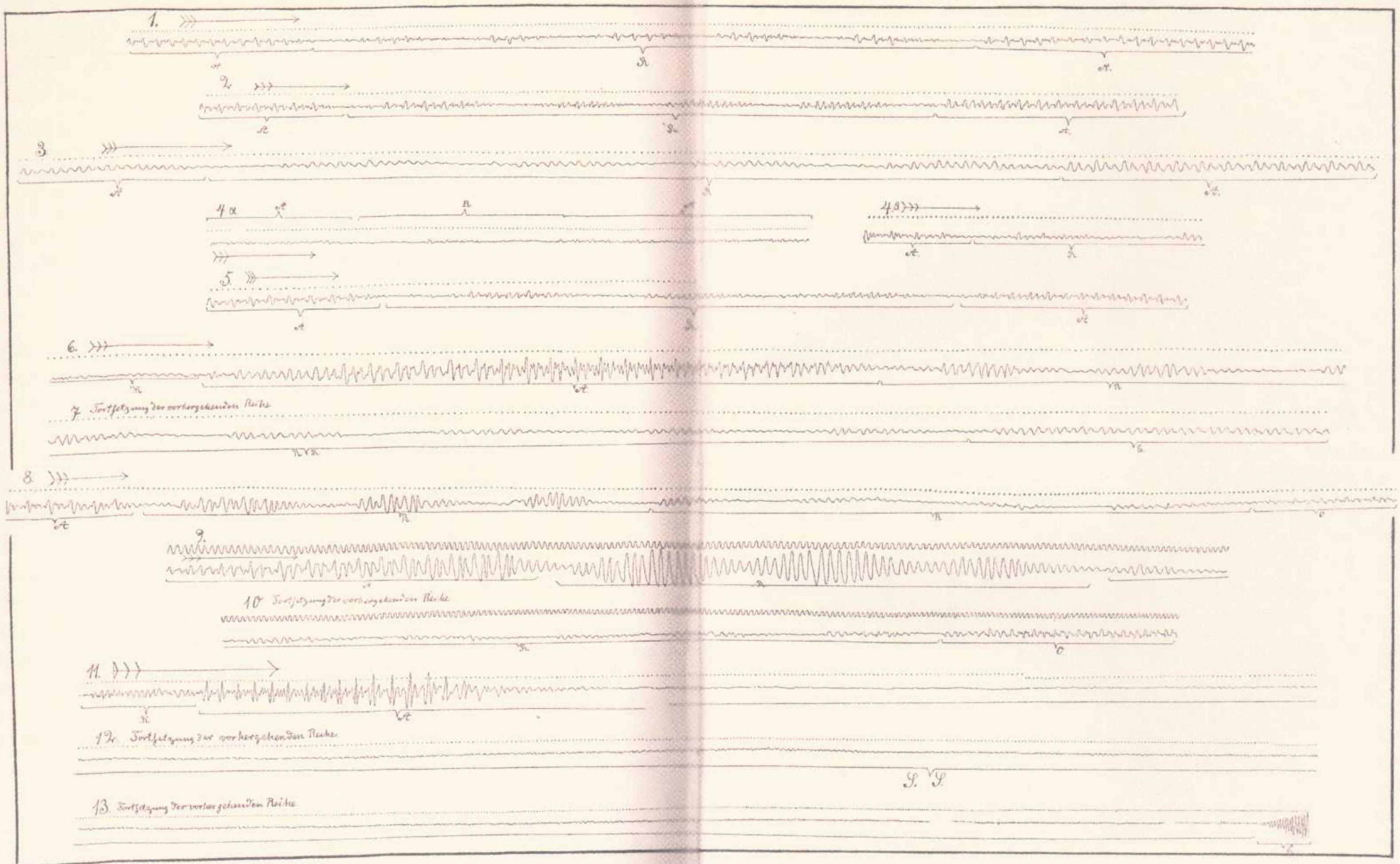
Top image: (K)ara [Plate I, line 1]; *Bottom image:* Karre [Plate II, lines 6 and 7].

In the mid-1880s, a student of medicine, surgery, and obstetrics at the University of Kiel named Paul Wendeler chose an innovative project for his dissertation research: he set out to obtain visual records of the sound waves corresponding to consonants. There had already been a number of comparable studies of vowels by this time, but nobody had yet tried to do much with consonants, so even the most basic findings in this area had the potential to break new ground. The sound-recording instrument Wendeler used for his project was a modified phonautograph called the *Sprachzeichner* or "speech-depicter" designed by his mentor, Victor Hensen, a professor of physiology better known for his pioneering work in marine biology (he coined the term "plankton," among other things). The *Sprachzeichner*'s most distinctive feature was that the waveforms it traced on lampblacked glass plates were so tiny they had to be viewed under a microscope. "Wendeler's curves are mostly only a few hundredths of a millimeter tall," Hensen observed; "the greatest elongations...rarely exceed 0.2 mm."[58]

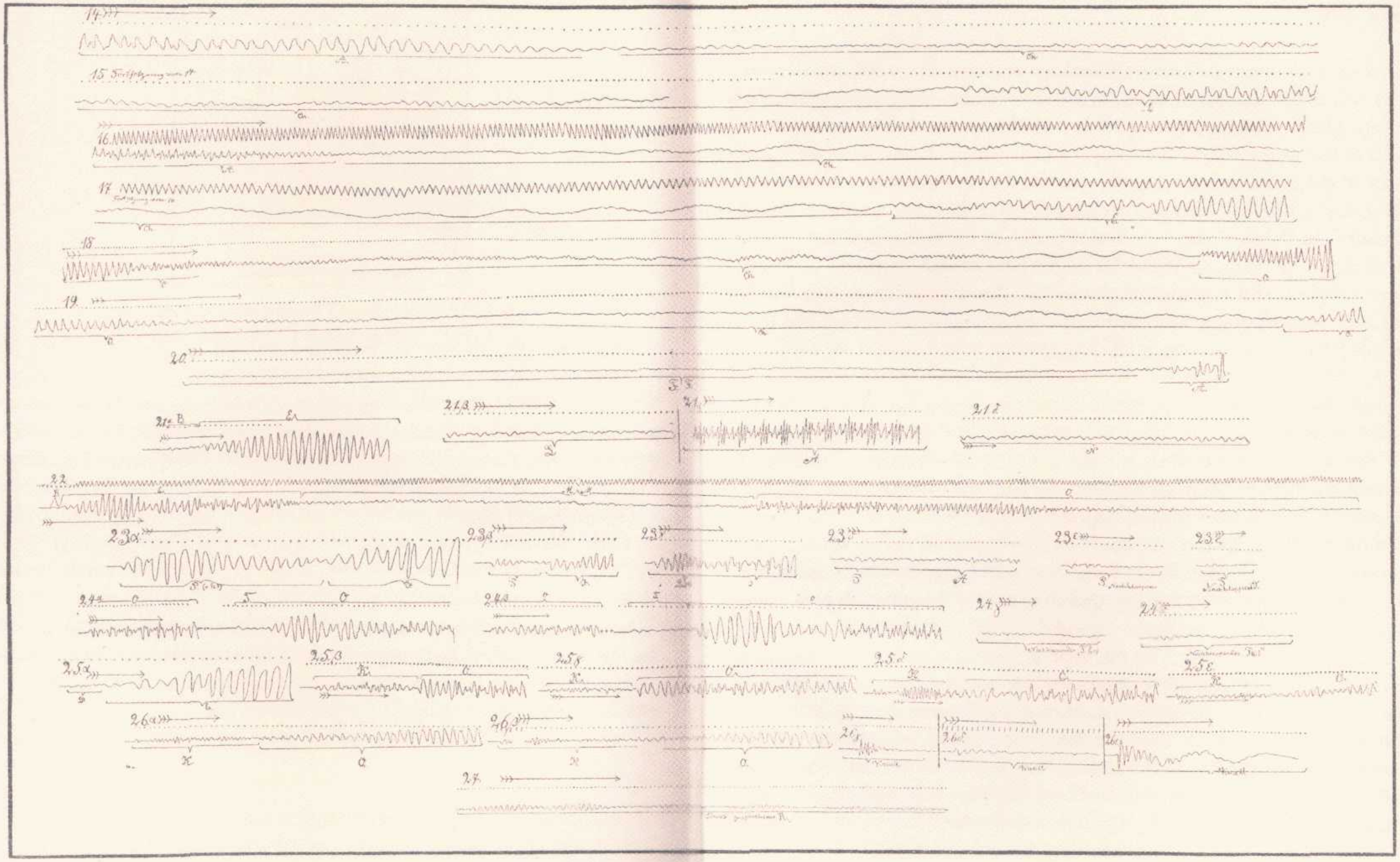
The original glass plates containing these microscopic traces were supposedly deposited with the Physiological Institute in Kiel and available there for inspection during the 1880s, but Wendeler also prepared some sample waveforms for wider print publication—a task that involved placing his originals under a Leitz microscope, copying them by hand with the help of a Zeiss drawing-prism, reducing the drawings photographically to one-third their original size, and then enlarging them slightly when transferring them onto the printing plates. Track eighteen is based on several of the images published in this way.

Wendeler seems to have moved his glass plates by hand at varying speed while recording on them, but he also recorded the vibrations of a 690 Hz tuning fork alongside the traces made by the voice—often just as a row of dots—and we can use this second trace as timecode to correct for irregularities in recording speed. In track eighteen, each example is accordingly

Zeitschrift für Biologie Band XXIII. Tafel II.



Kasse [Plate II, lines 11-13].



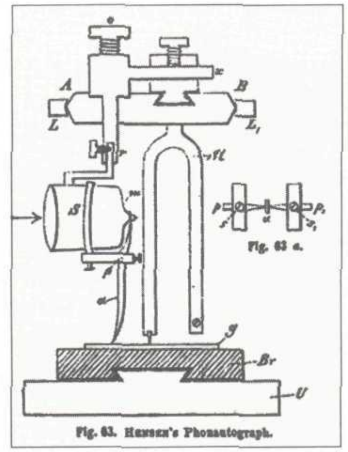
Bemmo [Plate III, line 22].

presented: (1) three times without speed correction, with voice and tuning fork in stereo; (2) three times with speed correction, with voice and tuning fork in stereo; (3) three times with speed correction, voice only, in mono. Finally, all the speed-corrected versions of the voice by itself are repeated once each.

Most of the sample waveforms Wendeler prepared for publication represent only short excerpts of words or syllables chosen to illustrate particular points. Our first example is consistent with this approach: its subject matter is identified as the word *Kara* sung at a constant pitch, but with the *k* and most of the *a* omitted, leaving the *r* as the focus of attention. It's unclear, incidentally, whether *Kara* was even supposed to be a real word—it means "penalty" in Polish, "black" in Turkish, and so forth, but it doesn't mean anything in German, so Wendeler might have chosen it merely as an experimentally useful sequence of phonemes. The same is true of our fourth example, identified as a "complete curve" of *Bemmo*—which could be Italian, or Japanese, or nothing at all.

By contrast, examples two and three clearly represent actual German words: *Karre,* meaning "cart" or "wheelbarrow," and *Kasse,* meaning "cash-desk." To the best of my knowledge, these are the oldest extant recordings of identifiable spoken words in the German language. Nevertheless, Wendeler recorded them with specific experimental goals in mind, which had a significant impact in turn on *how* he chose to record them. Thus, he acknowledged that the double *s* in *Kasse* "was held for a very long time, hence the great prolongation of the corresponding part of the curve." It seems he'd found it was "difficult to obtain a good S-curve," so he had drawn out the *s* in *Kasse* to unusual length in the hope of getting it to record, something like *Kassse*.[59] In the published version of the trace, he also omitted most of the *e* from the end of the word as uninteresting. Because the *s* didn't end up recording very well in spite of Wendeler's best efforts, we actually end up hearing something today that sounds more like *Ka....e*. On the other hand, *Karre* turned out pretty well, so we have at least that one respectable recording of a spoken German word from the mid-1880s.

(((☼)))



An earlier model of Hensen's Sprachzeichner. L. Hermann, Handbuch der Physiologie *(Leipzig: F. C. W. Vogel, 1879), 1:2:187.*

Source: Paul Wendeler, "Ein Versuch, die Schallbewegung einiger Consonanten und anderer Geräusche mit dem *Hensen*'schen Sprachzeichner graphisch darzustellen," *Zeitschrift für Biologie* 23 (1887), 303-320, Plates II and III.

a. (K)ara [Plate I, line 1]
b. Karre [Plate II, lines 6 and 7]
c. Kasse [Plate II, lines 11-13]
d. Bemmo [Plate III, line 22]

This work had first appeared in print the previous year as Paul Wendeler, *Ein Versuch, die Schallbewegung einiger Konsonanten under anderen Geräusche mit dem Hensen'schen Sprachzeichner graphisch darzustellen* (München: R. Oldenbourg, 1886).

(((((((19. MANOMETRIC FLAMES)))))))

(1:19)

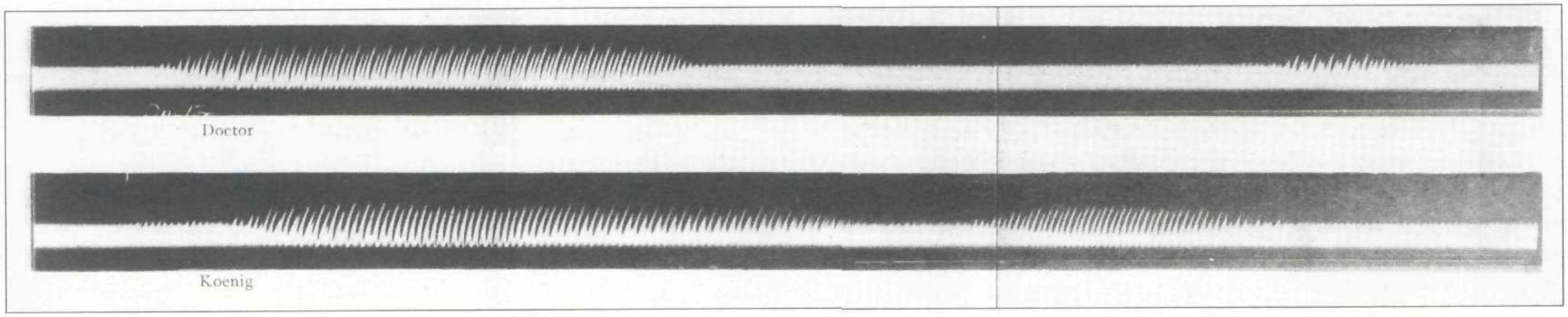
The phonautograph wasn't the only apparatus available during the nineteenth century for creating graphical images of the human voice. Another instrument used for the same purpose was the flame manometer invented by Rudolph Koenig. This worked by having a membrane cover a space into which lighting gas was fed on its way to a gas-burner, such that the membrane's in-and-out movements in response to changes in air pressure controlled the amount of gas reaching the burner and thereby caused the height of the resulting flame to fluctuate. A rotating mirror could then be used to display the changes in the flame over time as a bright band of varying height. Unlike the phonautograph, Koenig's flame manometer didn't lend itself easily to the creation of permanent images, and the first published illustrations of its output had to be drafted laboriously by hand.

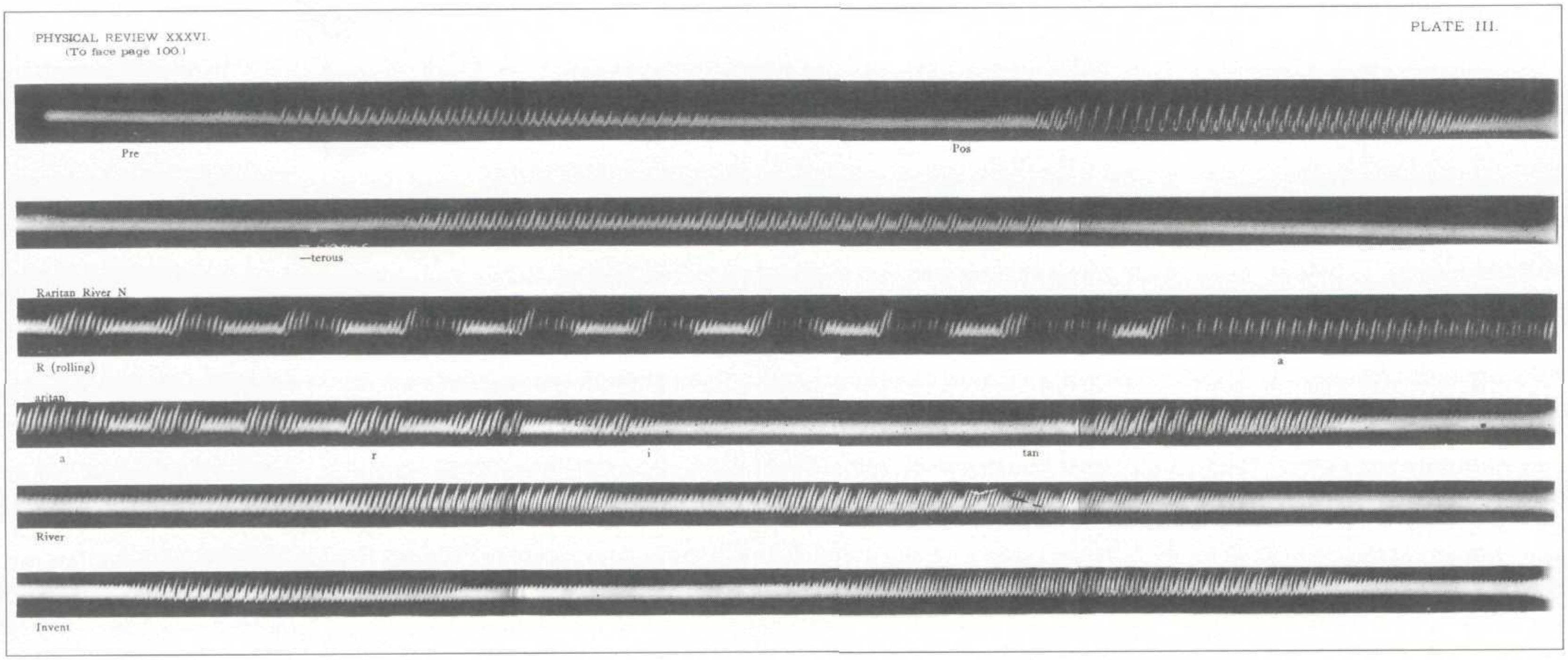
In 1872, Koenig published a table of images he'd prepared using his flame manometer to depict vowel sounds sung at different musical pitches. He wrote:

> In order to bring forth the images of the vowels, I sing them into a small funnel-shaped mouthpiece which is connected to the hollow space in front of the membrane by means of a short rubber hose.... I had already created, and caused to be painted, images for the vowels *u o a e i* sung on the notes of the two octaves C to c̄ in the year 1867.... Unfortunately, the five painted tables weren't ready in time for the [Paris] Exhibition, but I showed them in 1868 at the meeting of natural scientists in Dresden. If I have put off their publication until now, the reason for it is that I wanted to subject them to a more thorough revision beforehand from which, however, the diseased state of my throat unfortunately kept preventing me, which didn't permit me to undertake these strenuous experiments. Now, since I can no longer hope for a recovery, I have checked these images once again as well as could be done and give them now, even if not perfectly correct, then at least as accurately as I was able to succeed in drawing them.[60]

Koenig's manometric flame sketches have much in common with other graphs of time versus amplitude we've considered so far, except that in this case each successive point in time corresponds not to a straight vertical column, but to a curved arc. In order to educe sound from Koenig's vowel pictures, we first need to straighten those arcs out into vertical columns as best we can. You can hear the result of my efforts in the first half of track nineteen. I've presented the images first in order by column (the vowel *u* sung at all the various pitches, then the vowel *o*, then the vowel *a*, and so on) and then in order by row (the vowels *u, o, a, e, i* sung at one pitch, then at the next pitch, then at the next pitch, and so on).

"Flammenbilder der Vocale U, O, A, E, J
gesungen auf die Noten der zwei Octaven von C bis c̄."





"*The Photography of Manometric Flames,*"

Doctor Koenig [Plate I, bottom]; **Invent** [Plate III, strip six]; **Preposterous** [Plate III, strips one and two]; **Raritan River** [Plate III, strips three through five].

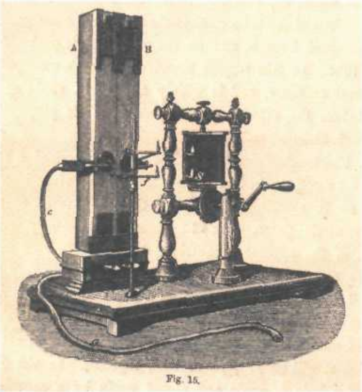
What we hear doesn't sound much like vowel sounds, but that's probably due to the inaccuracy of the drawings, and not the fault of the flame manometer itself. I say this because it turns out that we can educe perfectly intelligible speech from other manometric flame images created a quarter of a century later.

Edward L. Nichols and Ernest Merritt both belonged to the faculty of Cornell University, where they co-founded the *Physical Review*, the first American journal devoted exclusively to physics, in 1893. Five years later, they published an article in it to report the results of an investigation they had made into the potential value of the flame manometer for studies of the human voice. As they explained:

> It was decided first of all to study whole words and sentences rather than mere isolated sounds; for it is at least questionable whether a continuously sounded vowel has the same character as it would have if spoken naturally as a part of some word. It is clear that whole words cannot be studied without the aid of photography; and it is equally clear that some special form of apparatus is necessary, in order that photographs may be obtained conveniently and with certainty. Attempts to photograph whole words upon a sliding plate were usually unsuccessful, even when ten inch plates were used. Long films were finally employed, these being mounted upon a continuously moving drum.... The films used in the final experiments were 48 in. long and 3 in. wide. They were mounted upon the cylindrical surface of a large drum of tinned iron. The width of the drum was the same as that of the film, while the circumference was slightly less than 48 in. The film therefore extended completely around the drum, lapping over slightly at the ends.... [T]he exposure could be made to begin at any time desired and to automatically cease at the end of one revolution of the drum.[61]

Nichols and Merritt accompanied their article with three fold-out plates of manometric flame photographs. Four examples taken from these plates can be heard in the second part of track nineteen, presented three times each and then repeated once each at the end. The oscillations were originally recorded at an angle rather than straight up and down, but we can correct for this by skewing the images back in the opposite direction—a step equivalent to azimuth adjustment on later magnetic tape machines. Then, because the traces already take the form of bands of varying width, we can run them through ImageToSound without further graphic processing.

The test utterances heard in our four examples are "Doctor Koenig," presumably chosen in homage to the inventor of the flame manometer; "invent"; "preposterous"; and "Raritan River." In this last case, Nichols and Merritt intentionally contrasted the rolled *r* at the beginning of *Raritan* with the smooth *r* at the beginning of *river*, "in the production of which care was taken to avoid any tremulous motion of the tongue."[62]



Koenig's flame manometer. Pietro Blaserna, The Theory of Sound in its Relation to Music *(New York: D. Appleton and Company, 186), 24.*

(((✻)))

a. Source: Rudolph Koenig, "Flammenbilder der Vocale U, O, A, E, J gesungen auf die Noten der zwei Octaven von C bis c̄," *Annalen der Physik und Chemie* 146:2 (1872), Table III.

b. Source: Edward L. Nichols and Ernest Merritt, "The Photography of Manometric Flames," *Physical Review* 7 (1898), 93-101, Plates I-III.

i. Doctor Koenig [Plate I, bottom]

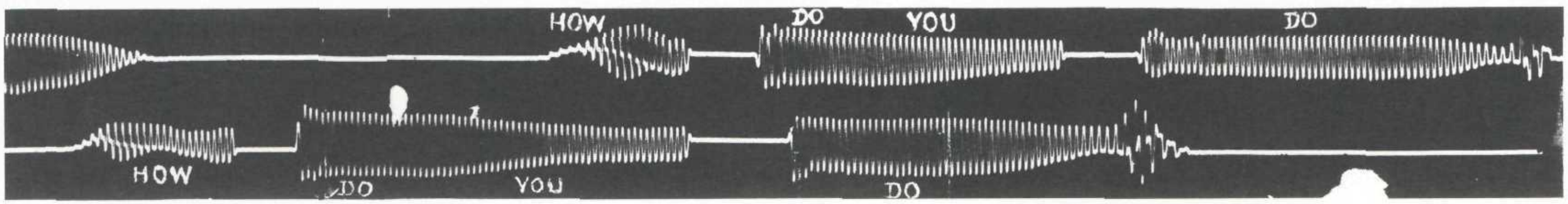
ii. Invent [Plate III, strip six]

iii. Preposterous [Plate III, strips one and two]

iv. Raritan River [Plate III, strips three through five]

(((((((20. BROWN UNIVERSITY, HOW DO YOU DO?)))))))

(1:28)

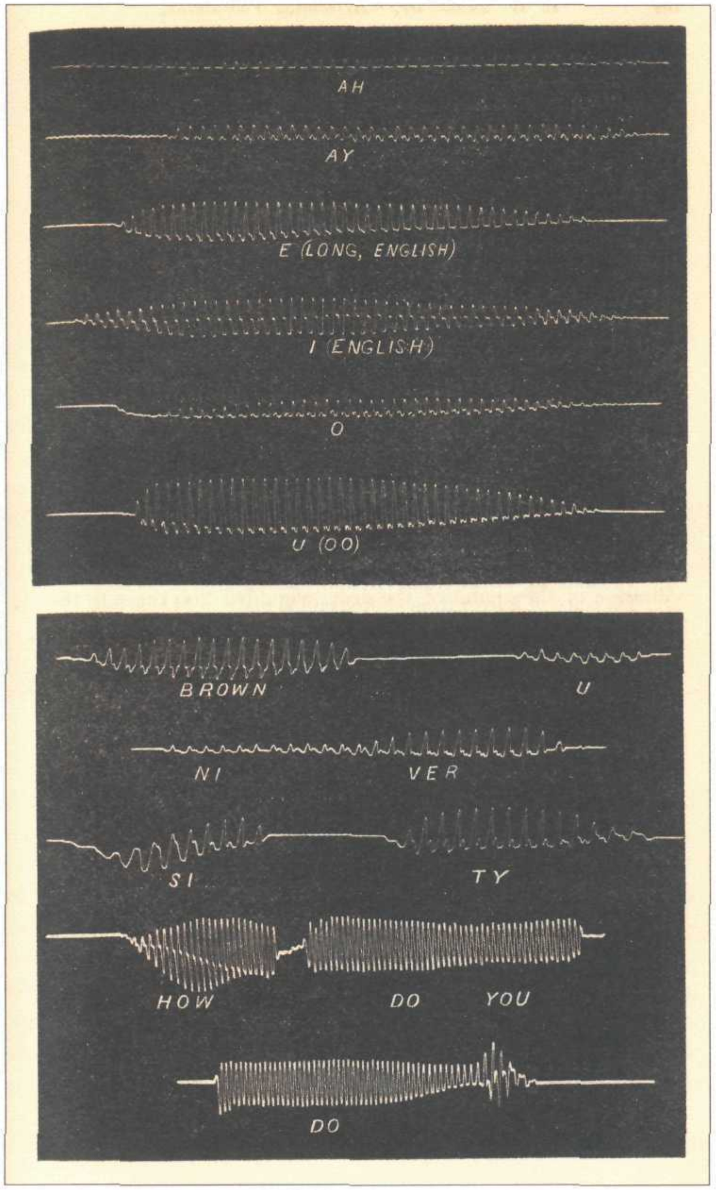


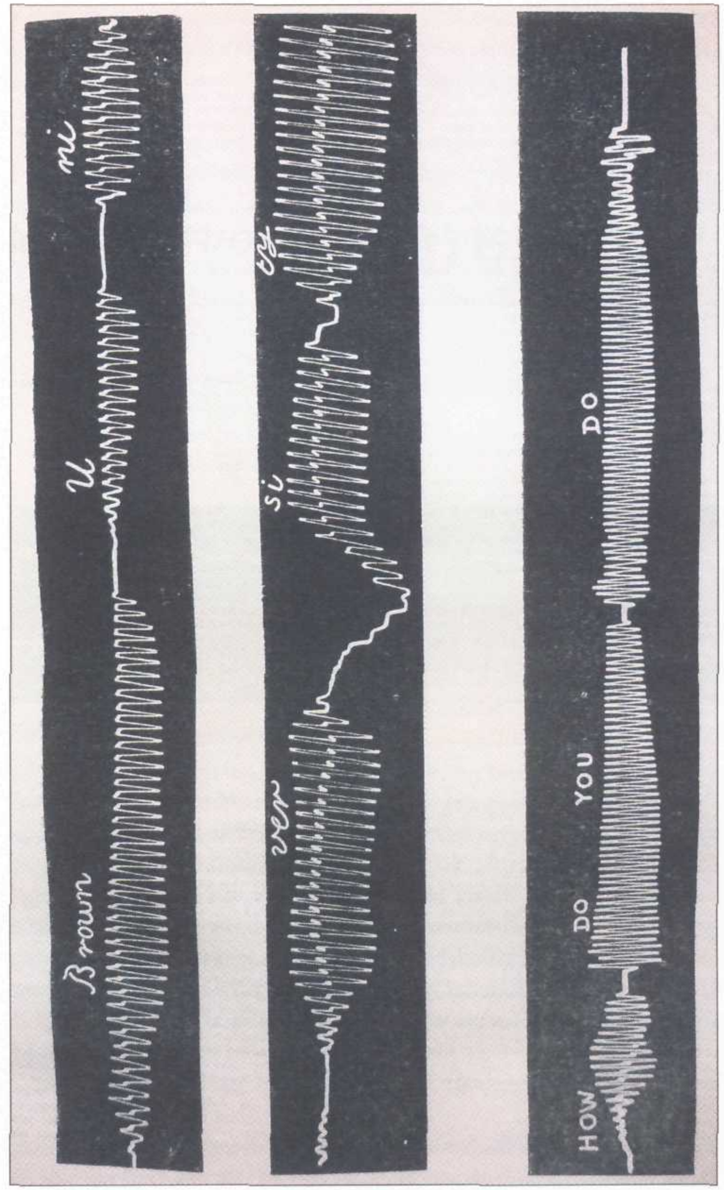
"How do you do?" (2×). Alexander Graham Bell Family Papers, Library of Congress.

Two decades before Nichols and Merritt took the manometric flame photographs featured in the previous track, another American researcher had succeeded in using photography to record speech vibrations for the first time in history. Eli Whitney Blake, Jr., a great-nephew of Eli Whitney of cotton gin fame, was Hazard Professor of Physics at Brown University, where he conducted a number of telephone exhibitions and experiments during 1877. Building on this foundation the following year, he pioneered a new technique of recording sound which entailed using a glass photographic plate to record deflections in a beam of light bounced off a mirror attached to the membrane of a telephone mouthpiece.

An original print made from one of Blake's plates survives today among Alexander Graham Bell's papers at the Library of Congress, and we hear its contents repeated twice at the beginning of track twenty. The print displays two separate traces, each representing the spoken words "How do you do?" The trace at the top also seems to contain a chopped-off piece of some other syllable before "how," suggesting that Blake might have had some trouble getting the timing of his experiments right at first. "How do you do?" is a typical American English greeting, of course, but it was also a common test utterance for telephones in this period, which might explain why Blake used it in this offshoot of his telephonic activities. In any event, its significance here is considerable: these two traces constitute the oldest known recordings of audibly recognizable speech in the English language.

In July 1878, Blake published an account of his accomplishment in the *American Journal of Science and Arts* and the British journal *Nature*, illustrat-





(*Left:*) "Ah, ay, e, i, o, u / Brown University / How do you do?" *American Journal of Science* 16:91 (July 1878), 57.
(*Right:*) "Brown University / How do you do?" *Nature* 18 (July 25, 1878), 339.

ed in both cases by images of sample recordings he had prepared. One of the images was identical both times: a plate showing the six vowel sounds *ah, ay, e, i, o,* and *u*. The other images differed from publication to publication but still represented the same two spoken phrases: "Brown University," where Blake was a professor of physics, and where he likely made the recordings; and "How do you do?"

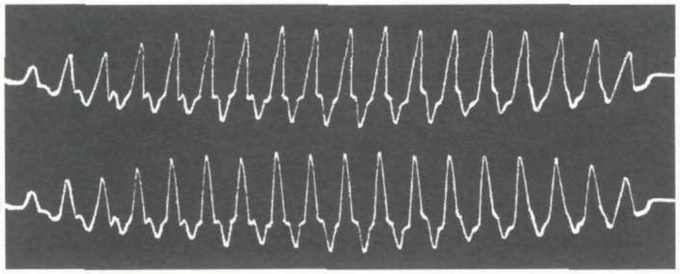
By comparing these recordings with the original print at the Library of Congress, we learn that they had been edited significantly for publication. The version of "How do you do?" published in *Nature* was copied from the upper trace displayed in the original print, but with shorter gaps left between the words, indicating that parts of the silences must have been edited out. The version published in the *American Journal of Science and Arts* is divided into two separate lines, one for "how do you" and the other for "do"; and although the "do" is copied from the end of the second trace in the original print, the "how do you" doesn't match the beginning of that trace and must have been copied from some other plate altogether—so we're dealing here with a composite spliced together from multiple takes.

Blake cites recording speeds in inches per second for all his examples, and he also tells us that some of his images have been reduced to 0.56 their original size (the exceptions are the full-page images unique to *Nature*, reproduced at full size). Based on this information, we can play these recordings at their exactly specified speeds. However, "Brown University" sounds much too fast at the stated 40 inches per second, and I suspect the correct speed in this case is actually half that. Blake's waveforms are also tilted at a slight angle, so it's worth our while to skew the images back in the opposite direction—as noted earlier, this process is equivalent to azimuth adjustment.

The second part of track twenty presents each of Blake's published recordings four times in a row, educed in a slightly different way each time so that you can hear how the differences in processing affect the sound:

- In stereo, without azimuth adjustment.
- Summed to mono, without azimuth adjustment.
- In stereo, with azimuth adjustment.
- Summed to mono, with azimuth adjustment.

Finally, you'll hear the final versions of all Blake's published recordings strung together at the end of the track as a grand finale: "Ah, ay, e, i, o, u- Brown University, how do you do? Brown University, how do you do?" I think that would make a terrific college yell.



Azimuth correction:
before (top) *and after* (bottom).

(((❁)))

Source:
a. "How do you do?" (2×) Source: Print found in Alexander Graham Bell Family Papers, Library of Congress, Box 317.

b. Source: E. W. Blake, Jr., "A method of recording Articulate Vibrations by means of Photography," *American Journal of Science* 16:91 (July 1878), 54-59, at page 57.
i. "Ah, ay, e, i, o, u" [21½ inches per second]
ii. "Brown University" at stated speed [40 inches per second]
iii. "Brown University" at half that speed [20 inches per second]
iv. "How do you do?" [14 inches per second]

c. Source: E. W. Blake, Jr., "A method of recording Articulate Vibrations by means of Photography," *Nature* 18 (July 25, 1878), 338-340, at 339.
i. "Brown University" at stated speed [40 inches per second]
ii. "Brown University" at half that speed [20 inches per second]
iii. "How do you do?" [14 inches per second]

(((((((21. CLEMENS RECTOR AETERNE)))))))

(10:01)

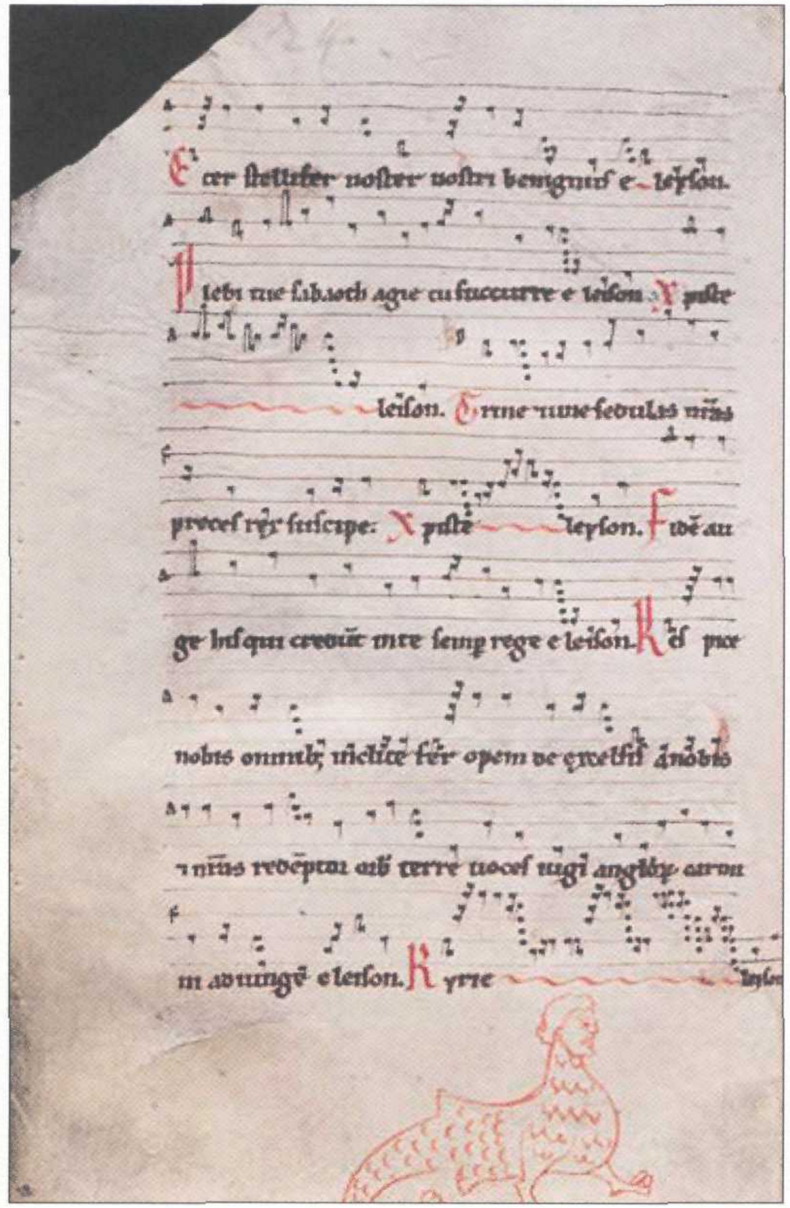
What would happen if we were to take an ordinary piece of sheet music and try to play it paleospectrophonically? The idea isn't as crazy as it might seem. We could read the horizontal dimension as a time axis, insofar as movement from left to right corresponds to motion forward in time. And we could read the vertical dimension as a frequency axis, insofar as movement from bottom to top corresponds to increase in pitch.

But the results would sound disappointing, or at least very strange. The problem is that modern staff notation is filled with arbitrary conventions that paleospectrophony can't accommodate. Even after a thorough degridding (to remove staff and bar lines and other extraneous marks such as stems, beams, and flags), we'd still face serious obstacles. For instance, the differences in appearance between a whole note, a half note, and a quarter note wouldn't correspond in any usable way to their intended differences in duration. Staff notation divides the octave spatially into just seven units, counting both lines and spaces, so the notes would come out in an unfamiliar seven-tone equal-tempered scale. Sharps, flats, ties and so forth would have to be ignored—or, worse, treated as acoustic data in their own right. Overall, modern staff notation violates the assumptions of paleospectrophony on too many levels for this to be a meaningful exercise.

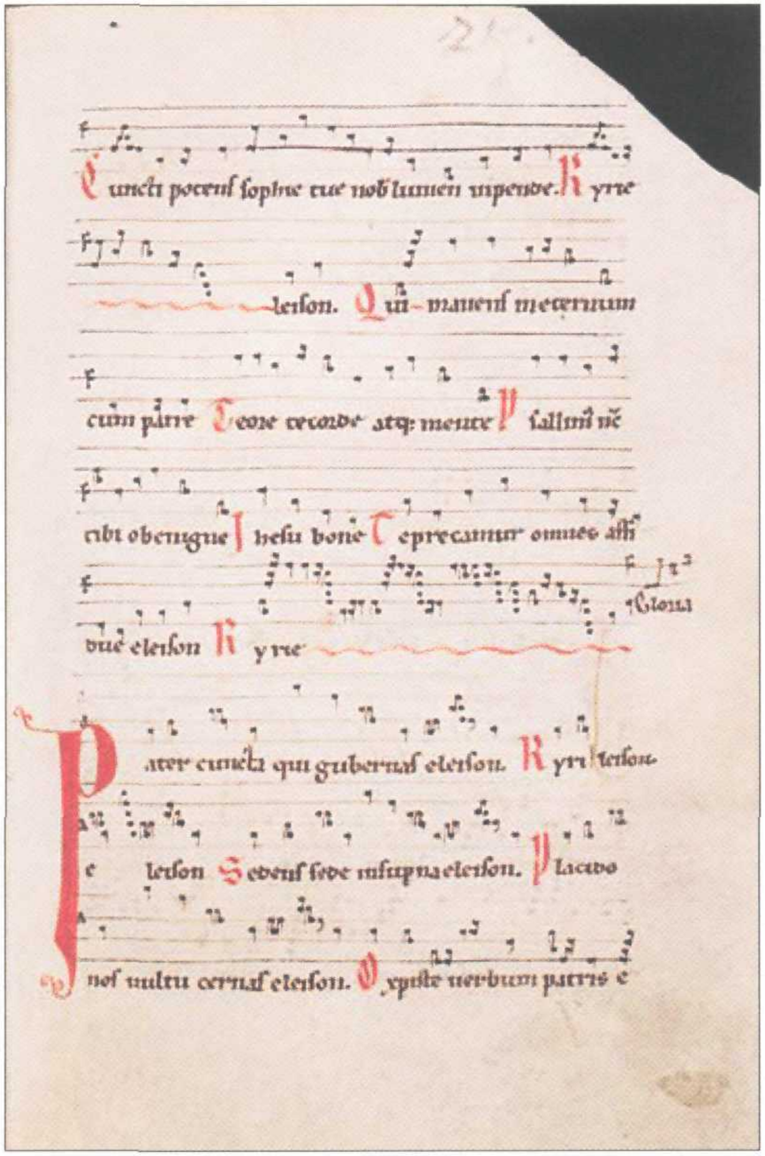
But it hasn't always been that way. Many forms of musical notation in use during the Middle Ages resembled graphs of time versus frequency more closely than today's staff notation does, so musical manuscripts dating back several hundred years are often much better candidates for paleospectrophonic eduction than today's sheet music is. True, even these forms of notation weren't *exactly* graphs of time versus frequency, so the most we can hope to hear in this way is a loose approximation of the intended musical content. But we're talking about *playing medieval manuscripts automatically, just as though they were sound recordings*—so even a loose approximation would still be downright mind-bending.

The example I've chosen for this track is the musical notation for "Clemens rector aeterne," a monophonic liturgical prayer centered on the Greek phrase *kyrie eleison* ("Lord, have mercy"), as found in a manuscript thought to have been created at the Cathedral of Lausanne sometime before the year 1250. Like modern staff notation, it divides the octave vertically into seven units, which should theoretically give us a seven-tone equal-tempered scale; in practice, irregularities in the spacing of the lines of the staff and in the vertical thickness and placement of notes actually result in something much less consistent than that. But even if the scale isn't quite right, the intervals we hear are correct as to direction and *approximately* correct as to distance.

The spacing of notes along the horizontal axis depends largely on the pres-

St. Gallen, Stiftsbibliothek, Cod. Sang. 383, 23 and 24.



St. Gallen, Stiftsbibliothek, Cod. Sang. 383, 25.

ence and length of words written underneath the staff. Some passages of the music are mostly syllabic, meaning that one syllable of text corresponds to one note, and in these passages the written notes end up spaced out with large gaps between them. But other passages are melismatic, meaning that one syllable of text spans multiple notes. In these passages, the written notes don't need to accommodate words written underneath, so their spacing is more closely analogous to their intended pace—and also much tighter than in the syllabic sections. In choosing a speed for this piece, I aimed for a compromise in which the syllabic sections wouldn't sound too slow and the melismatic sections wouldn't sound too fast.

And I've made one other concession to listenability besides. The right channel contains a straightforward paleospectrophonic eduction of the images from the manuscript, but the left channel contains the same content raised one octave *plus a synthetic echo derived from both channels*. I was somewhat reluctant to add the echo because it's not an "authentic" part of the eduction and so might be considered cheating. Without it, however, I find that the syllabic passages sound unpleasantly choppy; and I think the echo might be philosophically justified as a means of masking the fact that the width of the written notes wasn't originally intended to encode any information about their duration. But purists who don't want the echo are certainly welcome to listen to just the right channel by itself.

Source: St. Gallen, Stiftsbibliothek, Cod. Sang. 383, 23-25.

(((((((22. ALLELUIA)))))))

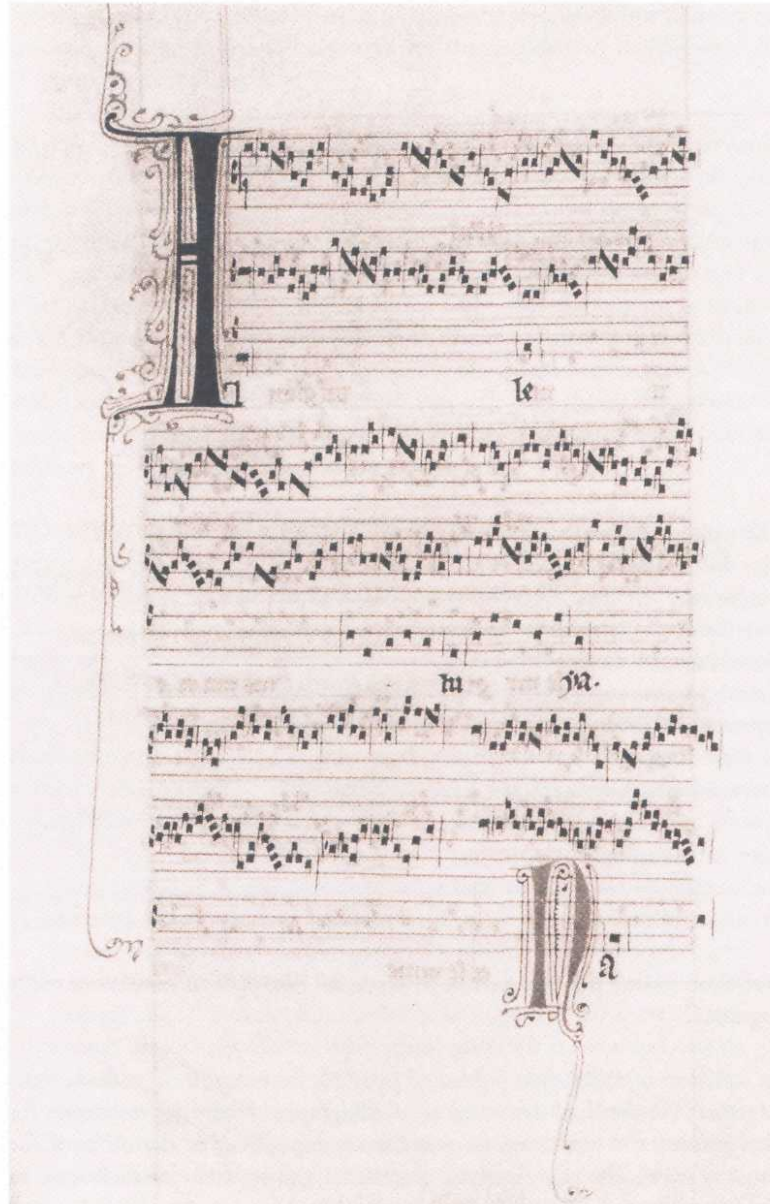
(2:03)

So now we've educed some medieval *monophonic* notation—but what might a comparable example of medieval *polyphonic* notation sound like?

Track twenty-two is based on a *quaternio* found in a French manuscript dating from roughly the middle of the thirteenth century, a richly melismatic rendering of the single word "alleluia." I looked through a good many manuscripts of medieval polyphonic notation hoping to find one in which the notes representing the different parts lined up reasonably well with each other along the time axis. Usually they don't line up at all; this particular example is somewhat exceptional. But even here, the parts don't line up as well as I'd like, and the seven-tone equal-tempered scale produces some very strange-sounding harmonies. Frankly, I don't think the result is very meaningful as a representation of medieval music—but it might do well as the soundtrack for a horror film. Stereo channels are used to keep the three parts maximally distinct.

(((☼)))

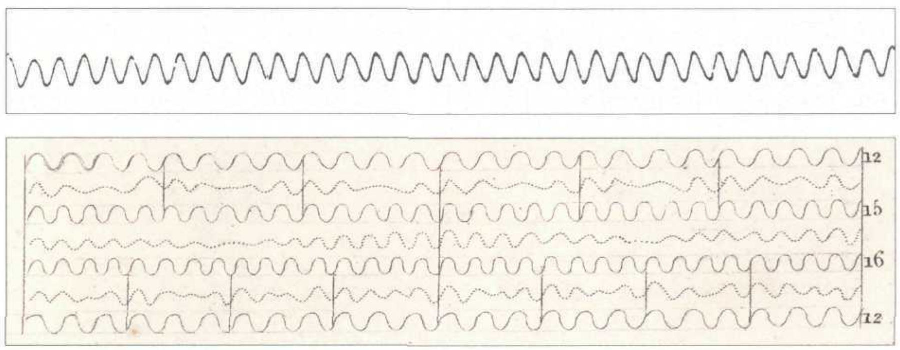
Source: MS Wolfenbüttel 1099 (1206), f. 16r. Facsimile in *Publications of Mediaeval Musical Manuscripts No. 2: Facsimile Reproduction of the Manuscript Wolfenbüttel 1099 (1206),* with introduction by Luther Dittmer (Brooklyn: Institute of Mediaeval Music, 1960).



(((((((23. DOTS BEFORE SCOTT'S)))))))

(1:13)

The phonautograph invented by Édouard-Léon Scott de Martinville in the mid-1850s was the first instrument in the world capable of recording aerial sound waves over time. And it represented these sound waves in a distinctive format which it shares in common with the LP groove and the waveform display in sound-editing software: the *oscillogram,* a graph of time versus amplitude.

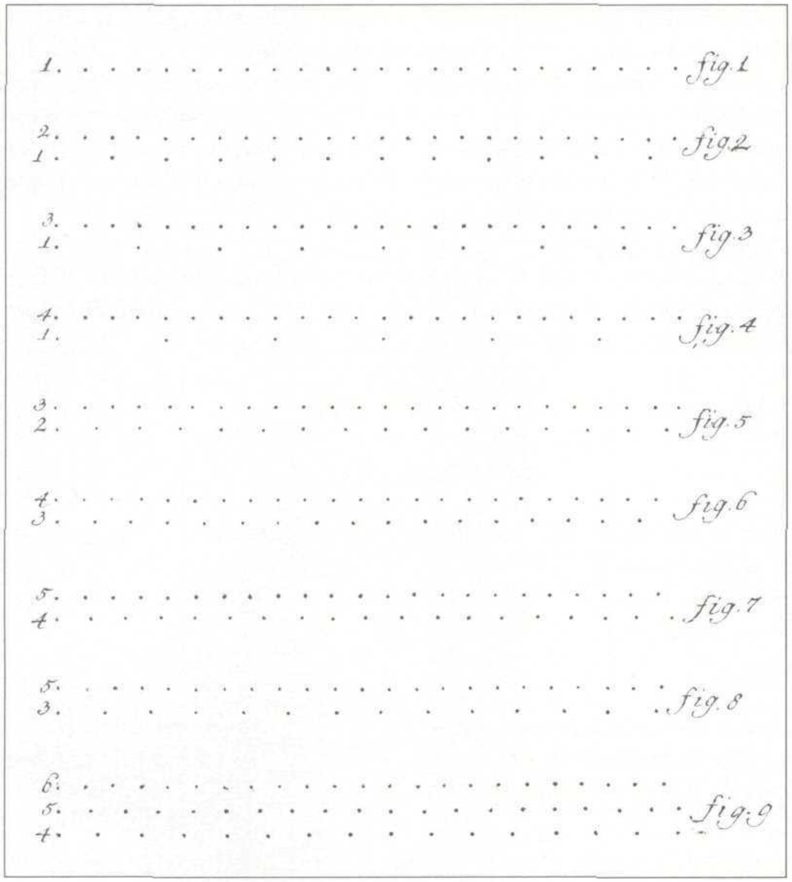


(*Top:*) "Notions générales de physique et de météorologie à l'usage de la jeunesse," 413; (*Bottom:*) "A Course of Lectures on Natural Philosophy and the Mechanical Arts," Plate XXV, fig. 353.

But Scott wasn't the first person to represent sound by means of an oscillogram.

In a physics textbook first published in 1850, for example, Claude-Servais-Matthias Pouillet had described an oscillographic recording technique for determining the numerical frequencies corresponding to the notes of the musical scale. The experimenter obtained a tuning fork for each note in an octave, equipped it with a stylus, set it in vibration, let the stylus trail against the ink-coated circumference of a wheel rotated at a uniform speed by electromotor, and counted the number of vibrations it traced in the space of one second. Pouillet reported the results as 528, 594, 660, 704, 792, 880, 990, and 1056 (in simple vibrations per second). More importantly for our purposes, he also published an enlarged picture of a sample waveform supposedly traced in this way.[63] At the beginning of track twenty-three, you hear Pouillet's sample waveform educed at each of his designated frequencies in turn, with the whole sequence repeated three times.

Over forty years before that, the famous English polymath Thomas Young had sketched out a group of waveforms to illustrate the phenomenon of beats, periodic changes in volume that result from the interference between waves of two slightly different frequencies. He explained the reasoning be-



Tentamen novae theoriae musicae,
Tab. I, page 36.

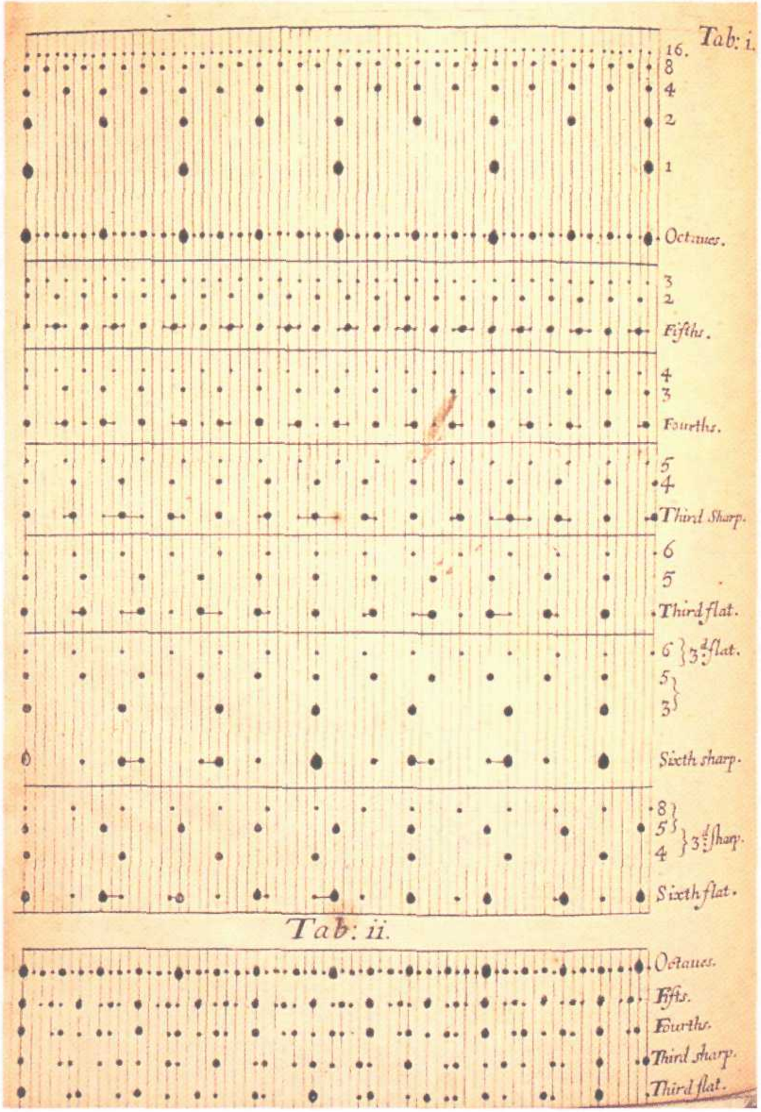
hind his oscillographic format as follows:

> The situation of a particle at any time may be represented by supposing it to mark its path on a surface sliding uniformly along in a transverse direction. Thus, if we fix a small pencil in a vibrating rod, and draw a sheet of paper along, against the point of the pencil, an undulated line will be marked on the paper, and will correctly represent the progress of the vibration. Whatever the nature of the sound transmitted through any medium may be, it may be shown that the path thus described will also indicate the situation of the different particles at any one time.[64]

To prepare his illustration of beats, Young first drew four parallel sine waves with frequencies in the ratio 12 : 15 : 16 : 12. Then, between each adjacent pair of curves, he constructed another curve out of points lying midway between in order to represent their "joint effect." The resulting images that appeared in Young's *Course of Lectures on Natural Philosophy and the Mechanical Arts*—officially published in 1807, but the engraving itself is dated 1806—clarified his point nicely.[65] Young had drawn these particular waveforms by hand rather than recording them automatically, but they're still oscillographic in format, and we can still play them: they comprise our second example in track twenty-three, educed several times at multiple arbitrary speeds (as are the third and fourth examples).

To the best of my knowledge, Young's illustration of beats is the oldest extant drawing of sound as a *waveform*—that is, a graph of time versus amplitude in the form of a continuously undulating line. But an oscillogram can also take the form of a row of dots; remember, Scott and Wendeler both had some of the records they made of the vibrations of tuning forks turn out this way. The eminent mathematician Leonhard Euler had already made use of this format in his *Tentamen Novae Theoriae Musicae* (1739), which is the source of our third example in track twenty-three, educed vertically;[66] and he used it again in a letter dated 29 April 1760, part of his well-known *Letters Addressed to a German Princess*:

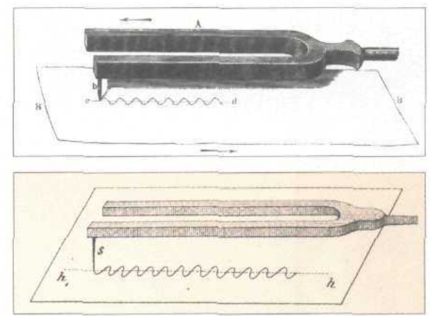
> The perception of a simple musical sound may...be compared to a series of dots equidistant from each other, as . . . . If the intervals between these dots be greater or smaller, the sound produced will be lower or higher. It cannot be doubted, that the perception of a simple sound is somewhat similar or analogous to the sight of such a series of dots equidistant from each other: we are enabled thus to represent to the eye what the ear perceives on hearing sound. If the distances between the dots were not equal, or were these dots scattered about confusedly, they would be a representation of a confused noise, inconsistent with harmony.[67]



A Philosophical Essay of Musick,
Tab. I.

An even older oscillographic representation of sound—again, as rows of dots—appears in a work by Francis North, published in 1677: a table described as showing "the pulses comprehended in the several chords...first set down in several lines, and then in one line with couplings, for the more easie observing the regular distances of the by-pulses on each side of the coincidence."[68] This table is the source of our fourth and final audio example in track twenty-three, again educed vertically.

In short, playable sound oscillograms may be found dating back long before Scott's invention of the phonautograph, even if they end up sounding more like video game music from the 1980s than anything else.



Method of recording the vibrations of a tuning fork directly.
(*Top:*) H. Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (Braunschweig: Friedrich Vieweg und Sohn, 1865), 33.
(*Bottom:*) Franz Melde, *Die Lehre von den Schwingungscurven* (Leipzig: Johann Ambrosius Barth, 1864), Taf. IV, Fig. 45.

(((❂)))

a. Source: M. Pouillet, *Notions générales de physique et de météorologie à l'usage de la jeunesse* (Paris: Béchet Jeune, 1853), 413; the same plate had been used for the 1850 edition.
b. Source: Thomas Young, *A Course of Lectures on Natural Philosophy and the Mechanical Arts* (London: William Savage, 1807), Plate XXV, fig. 353.
c. Source: Leonhard Euler, *Tentamen novae theoriae musicae* (Petropolis: Academia Scientiarum, 1739), Tab. I facing page 36. Educed vertically.
d. Francis North, *A Philosophical Essay of Musick* (London: John Martyn, 1677), Tab. I. Educed vertically.

(((((((24. SCOLICA ENCHIRIADIS — 25. MUSICA ENCHIRIADIS)))))))

(2:46), (5:53)

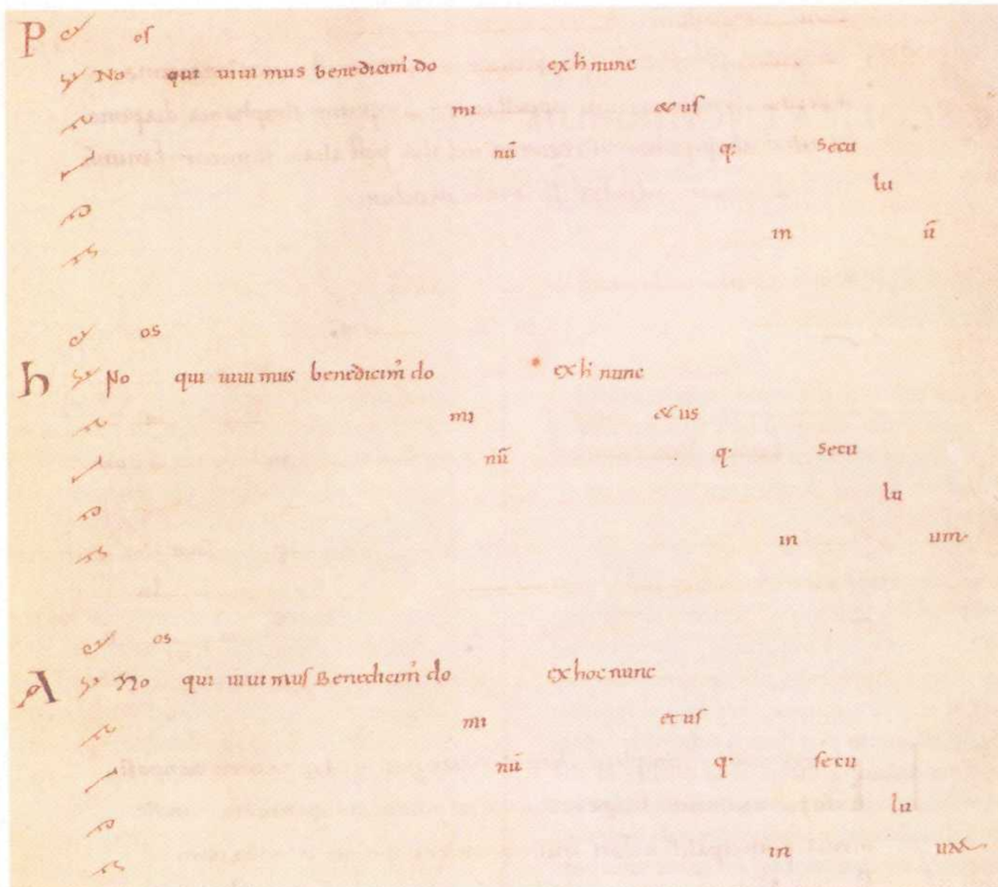
The oldest known written polyphonic music in the world exists in a distinctive format known as Daseian notation that was used only in a handful of ninth-century musicological works, the best-known being a pair of treatises called *Scolica enchiriadis* and *Musica enchiriadis*. It turns out to be an excellent candidate for paleospectrophonic eduction.

In Daseian notation, the vertical axis corresponds to frequency and is generally labeled on the left-hand side with a column of symbols identifying different pitches. According to the theory laid out in the accompanying treatises, the scale represented by these symbols is highly idiosyncratic—it's supposed to go up two semitones, then one semitone, then two semitones, then two semitones, and then repeat, a pattern that doesn't subdivide evenly into octaves. However, the rows representing the various pitches in the notation itself are evenly spaced, so from a paleospectrophonic standpoint, any theoretical variation between full-tone and semitone intervals is a moot point. Whenever two notes of the scale *are* separated by an octave, the space between them is divided into seven units—so if we equate that distance along the vertical axis with a doubling of frequency, we end up for purposes of eduction with a seven-tone equal-tempered scale.

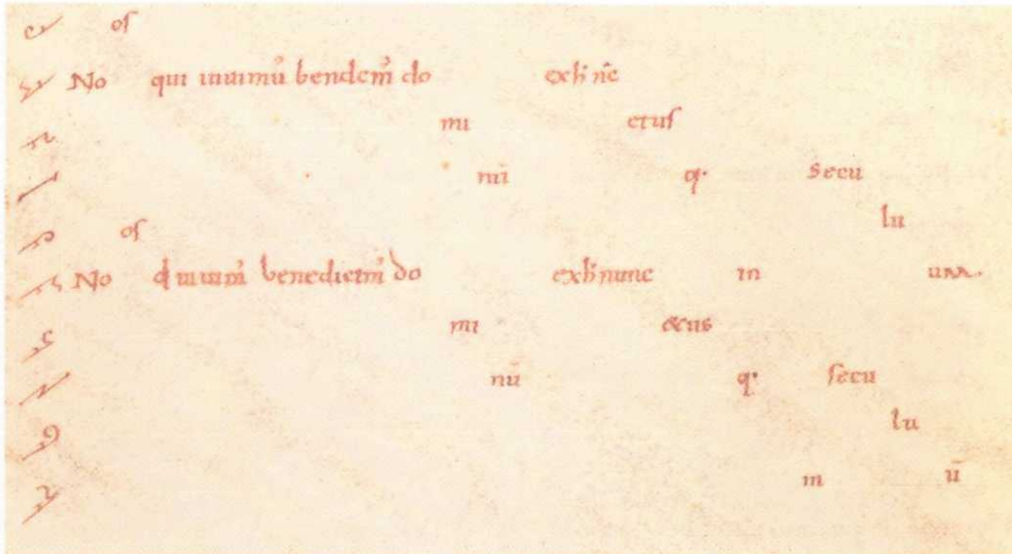
The most striking feature of Daseian notation is that each piece of the text being set to music is written in the row corresponding to the pitch on which it is supposed to be sung. As a result, the horizontal axis does a reasonable job of representing time for purposes of eduction: the more words there are, and the longer they are, the longer the note sung to them ends up being sustained. Moreover, the rows are spaced out widely enough along the vertical axis that the different shapes of the characters comprising the text don't significantly affect the pitch when they're educed.

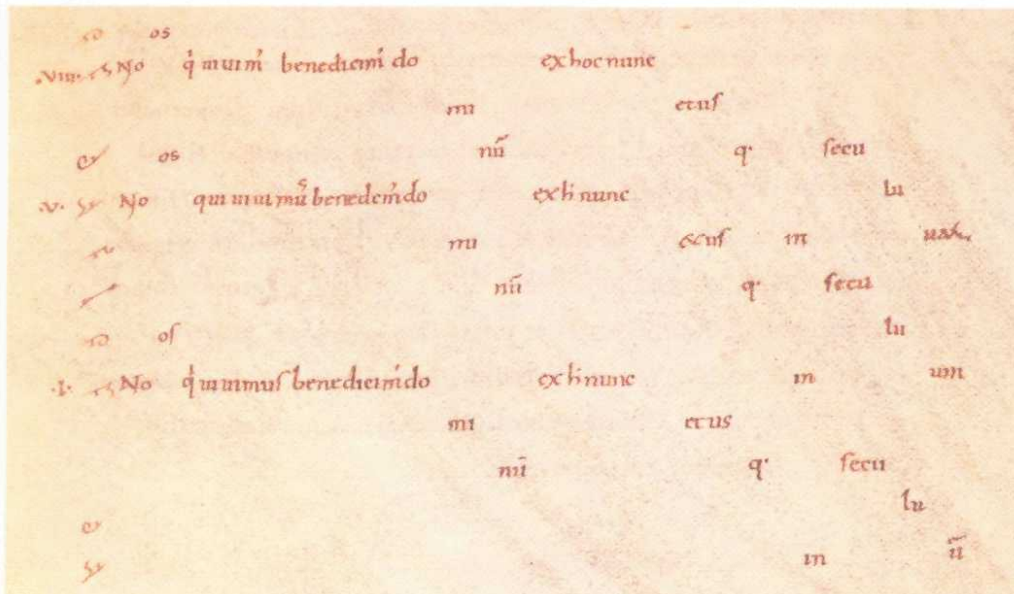
Tracks twenty-four and twenty-five are both based on parts of the same manuscript, thought to have been created in Lorraine towards the end of the tenth century. Track twenty-four presents the musical notation from the *Scolica enchiriadis*, which consists of eleven variations on the sung text "Nos qui vivimus benedicimus dominum ex hoc nunc et usque in seculum" ["We who live praise the Lord, now and forever"]. Track twenty-five then does the same thing with the *Musica enchiriadis*. In two cases (folios 4r and 9r), multiple lines plotted adjacent to one another weren't intended to represent parts sung simultaneously, so I've educed each line separately. Also, folios 7r and 7v use a familiar letter scale rather than the arcane Daseian symbols. For the most part, however, the eduction was pretty straightforward.

In these two tracks, you're listening to what might be the oldest automatically playable inscriptions in the world, dating back over a thousand years—and what you hear seems to represent the intended musical content of the

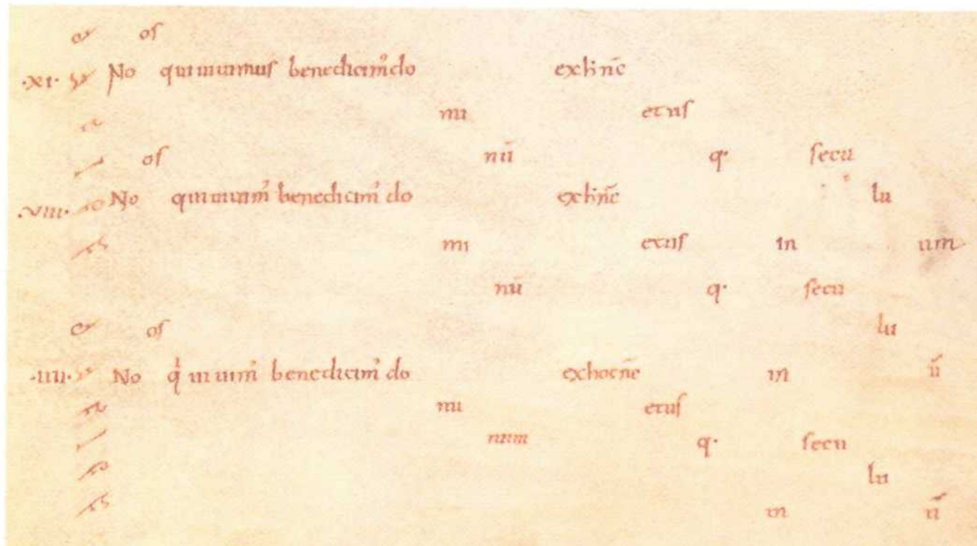


Track 24a [26r].

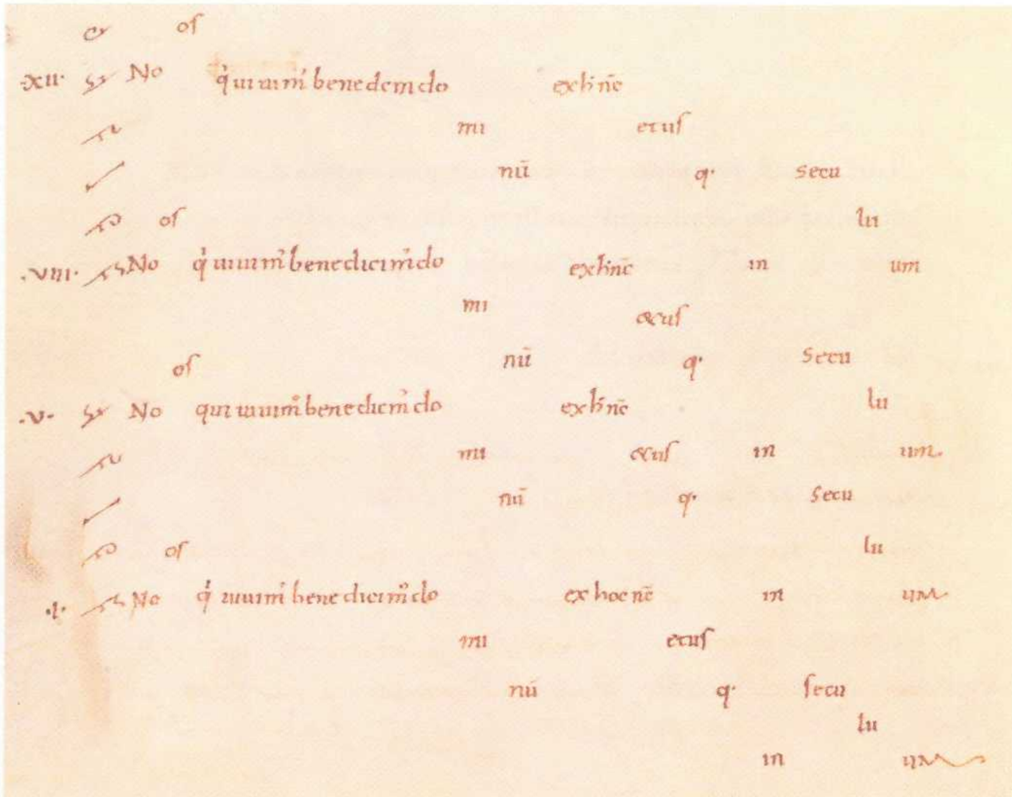




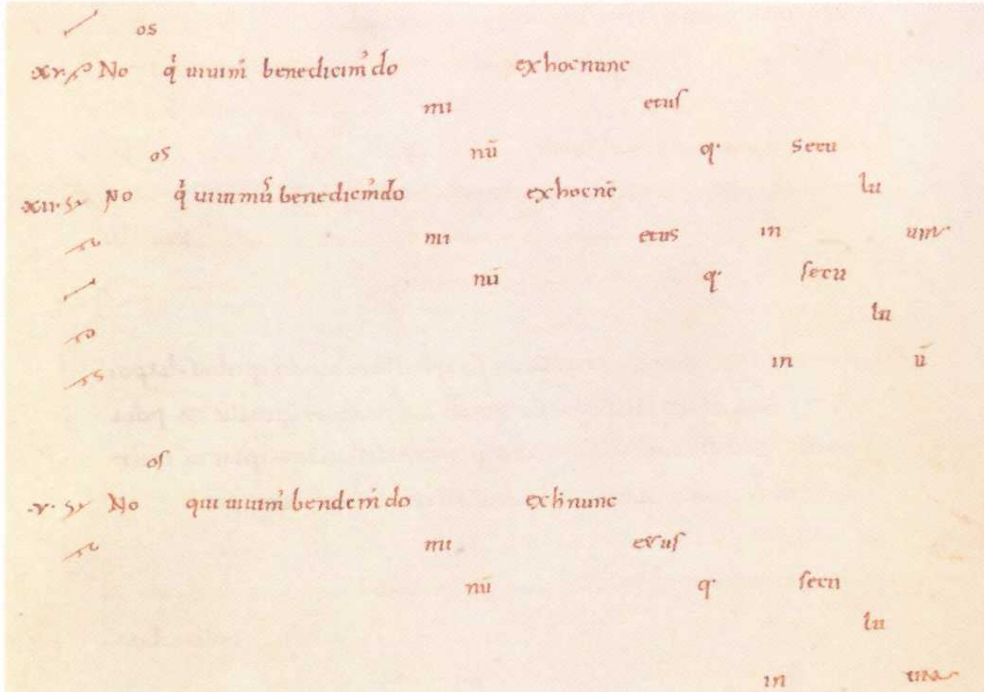
(*Top:*) Track 24b [26v],
(*Bottom:*) Track 24c [27r].



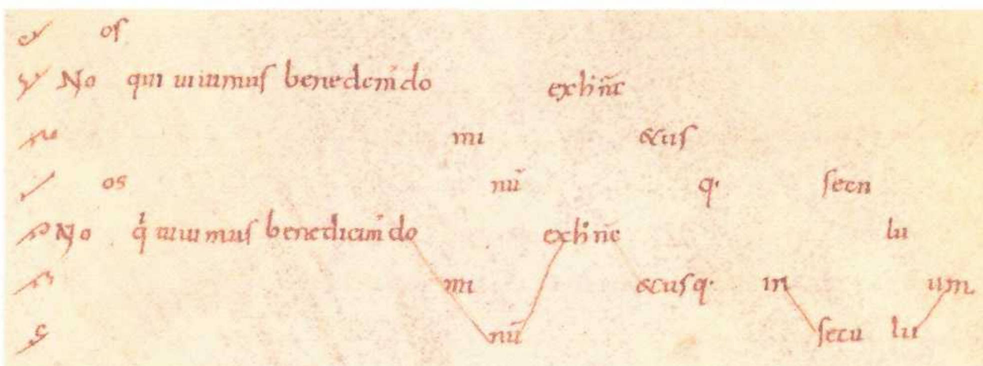
Track 24d [27r].



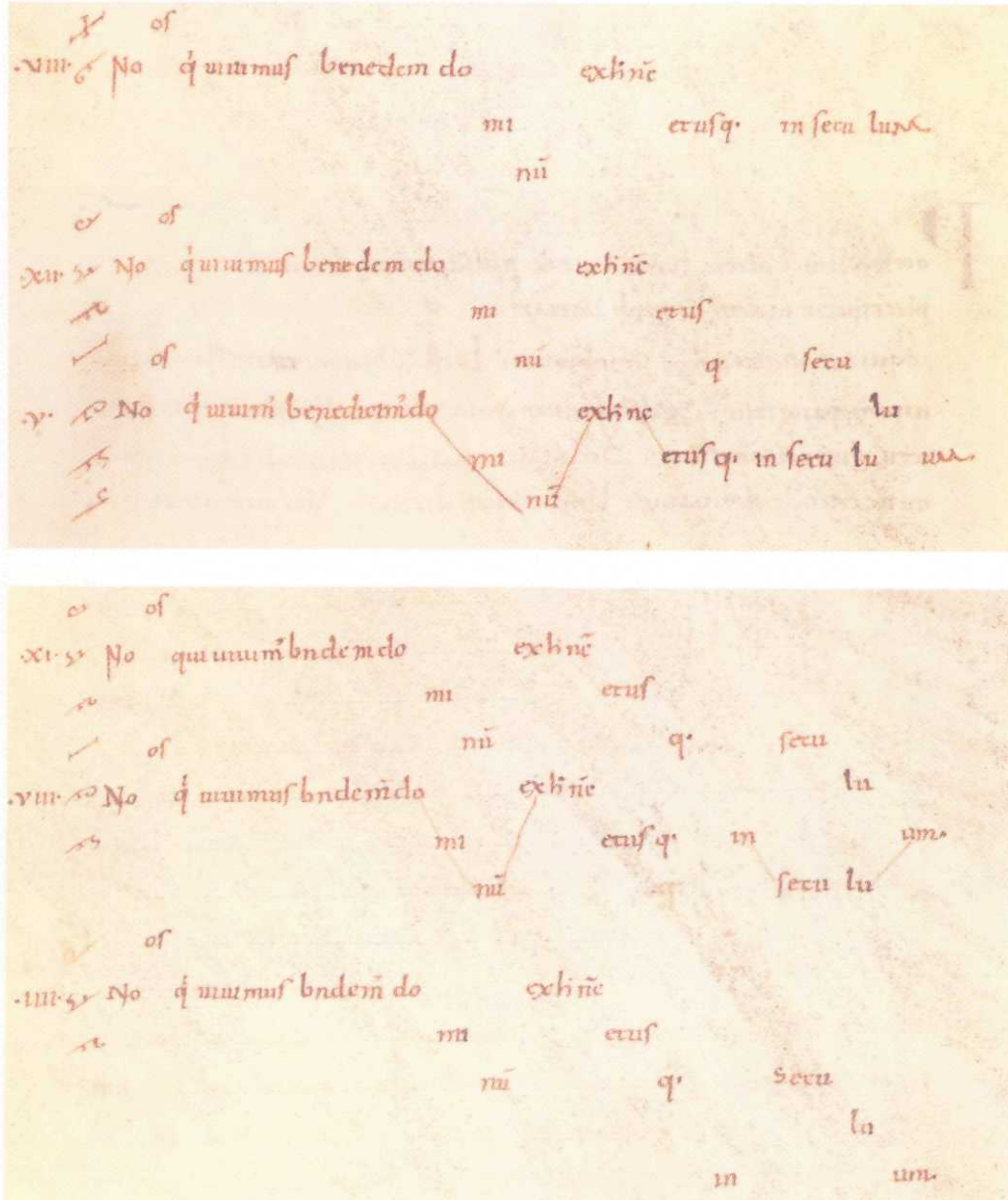
Track 24e [27v].



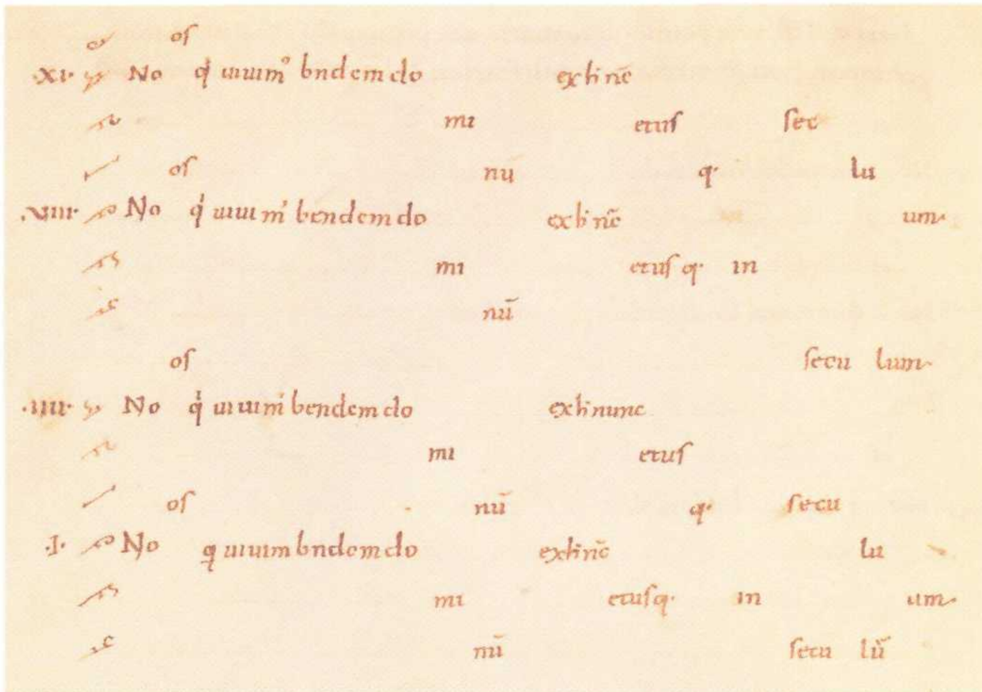
Track 24f [28r].



Track 24g [28v].

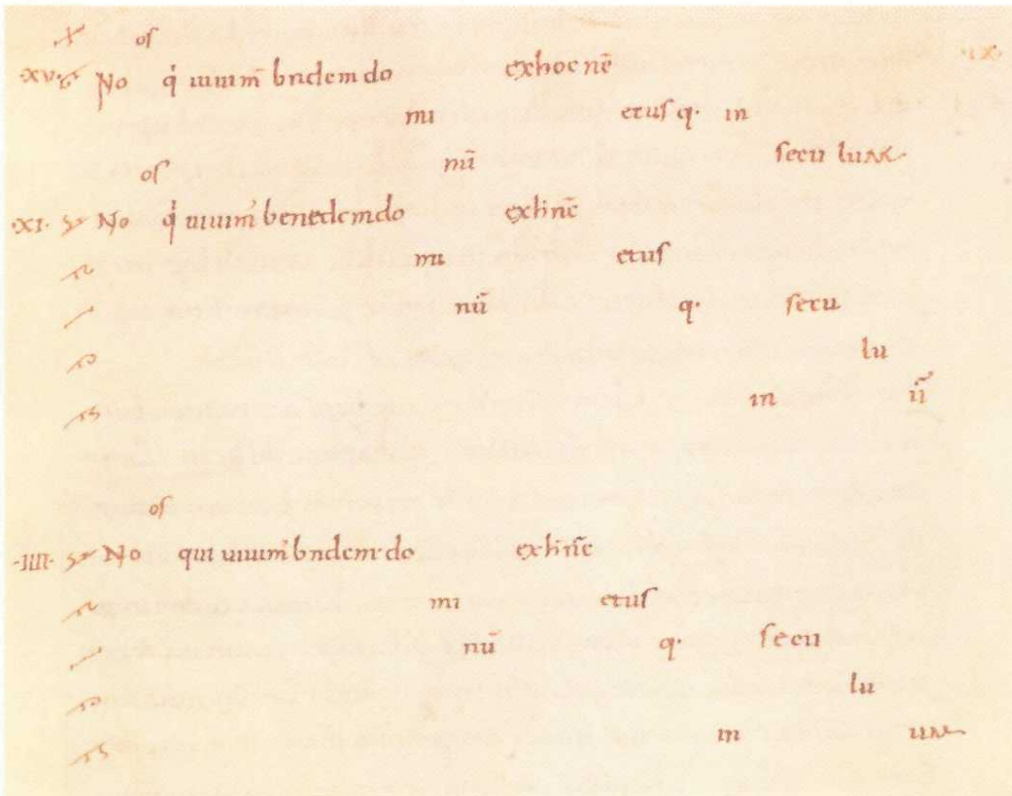


(*Top:*) Track 24h [28v]. (*Bottom:*) Track 24i [29r].

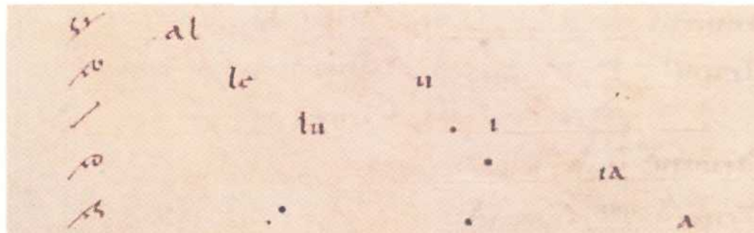


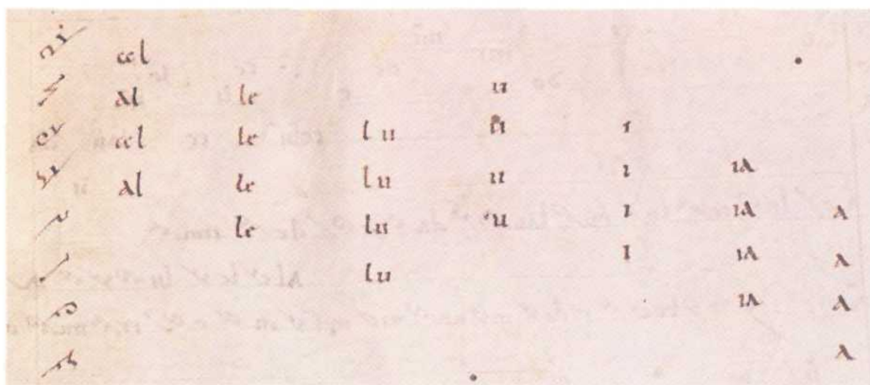
of
·xv· No q̄ uiuum bndem do ex h nc
mi etus sec
of nu q· lu
·viii· No q̄ uiui m bendem do ex h nc um·
mi etus q· in
nū
of secu lum·
·iiii· No q̄ ui uimi bendem do ex h nunc
mi etus
of nū q· secu
·i· No q uiuim bndem do ex hinc lu
mi etus q· in um·
nū secu lū

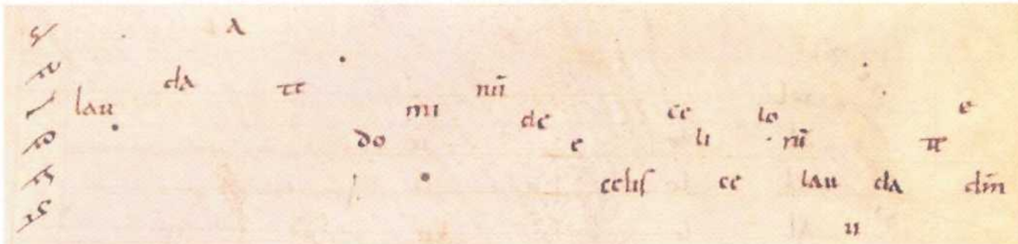
Track 24j [29v].

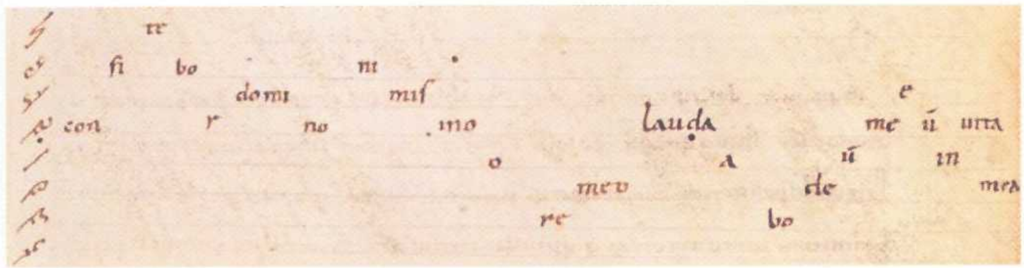


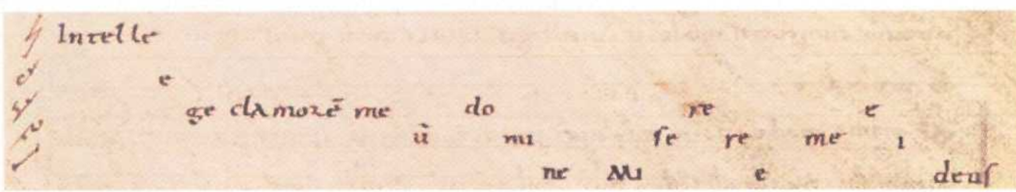
Track 24k [30r].

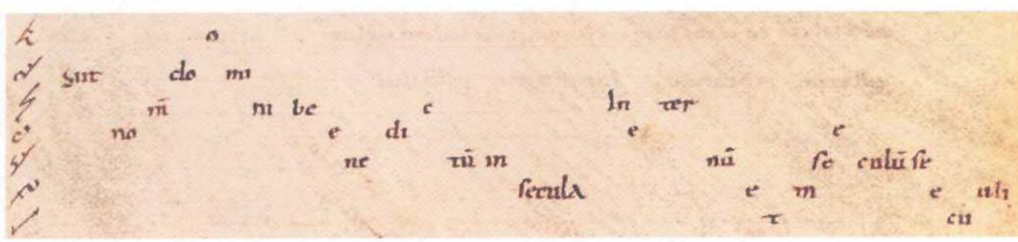

Track 25*a* [3v].


Track 25b [4r].

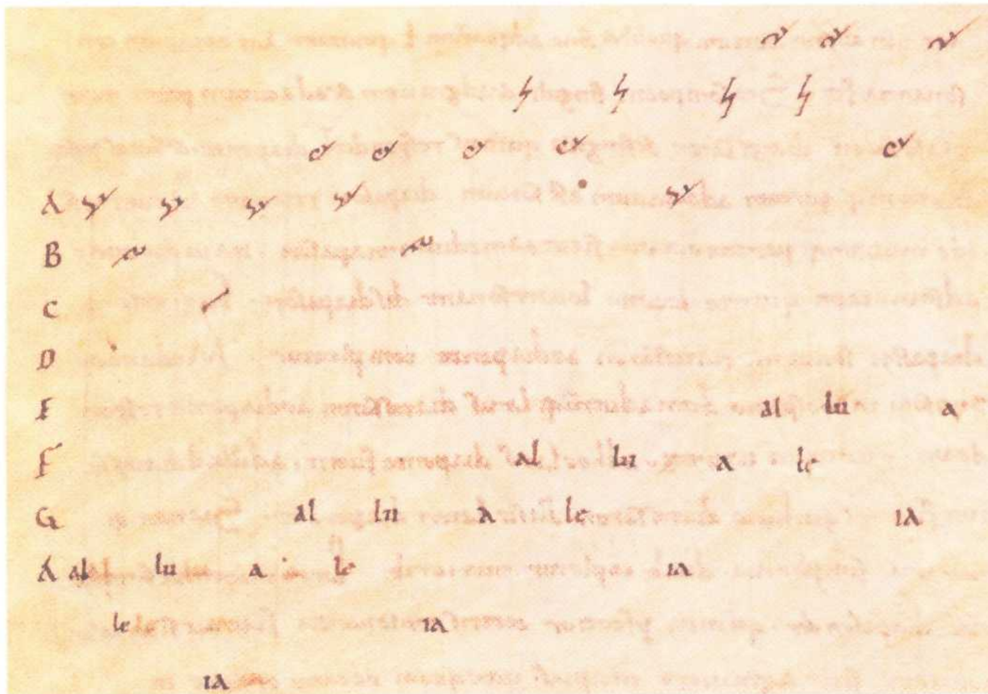




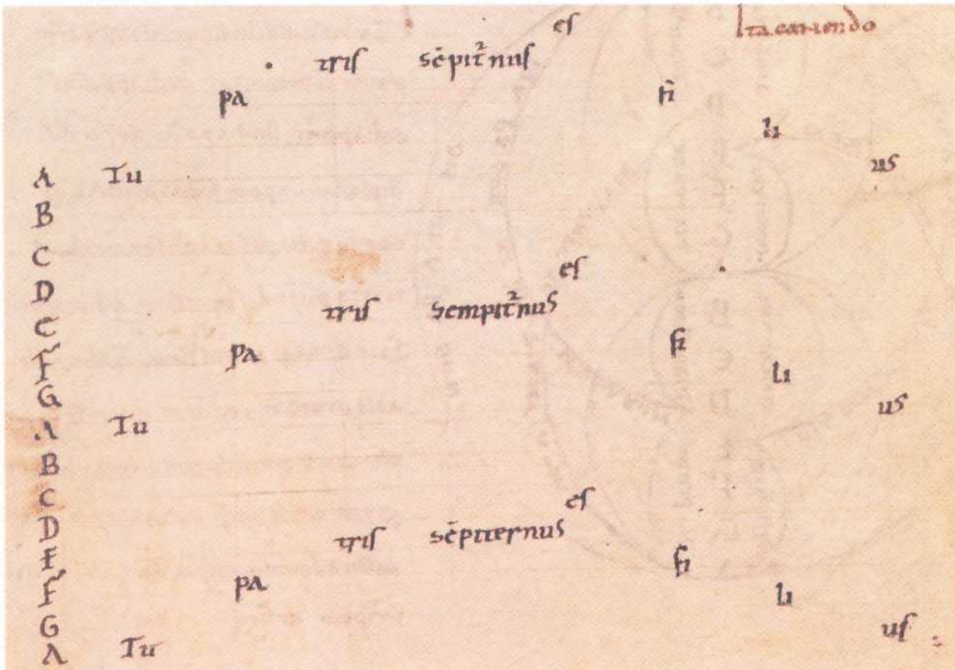


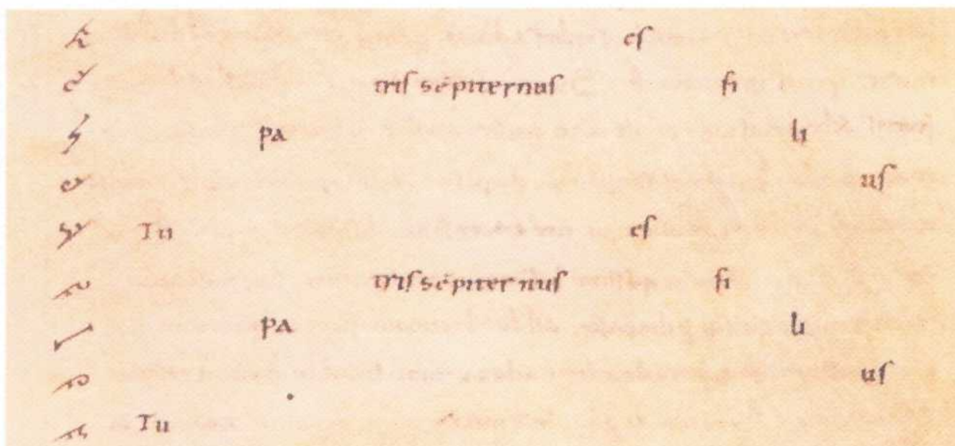


Track 25c [4v].

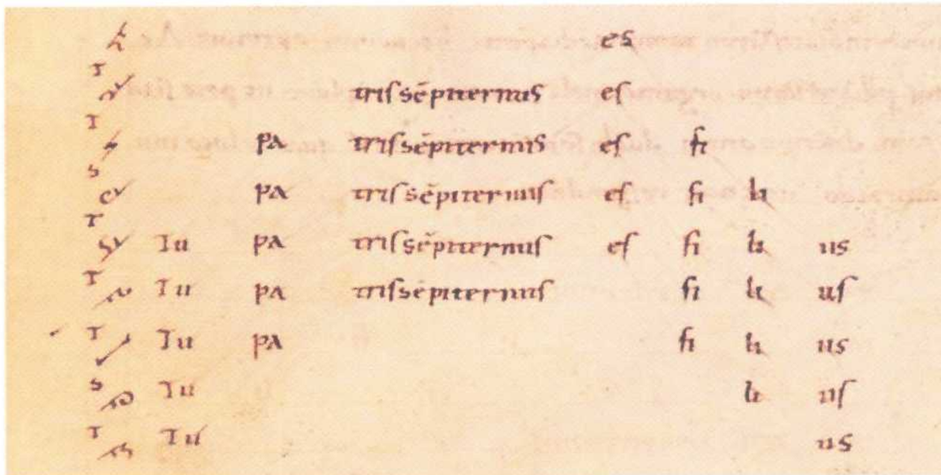


Track 25d [7r].

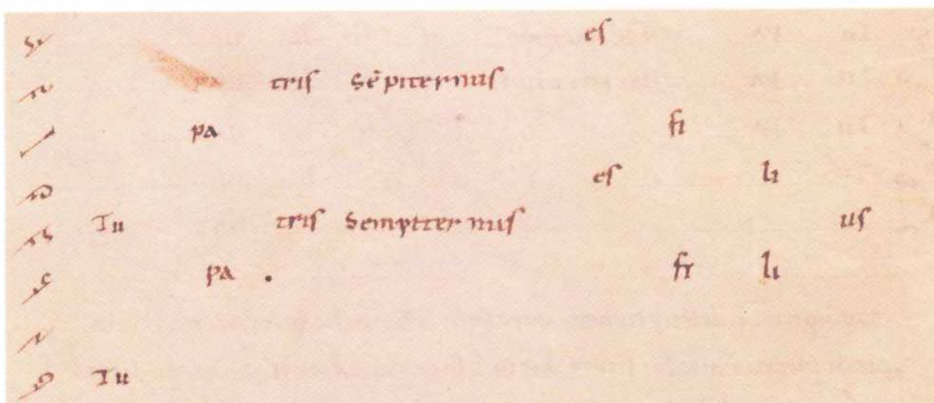




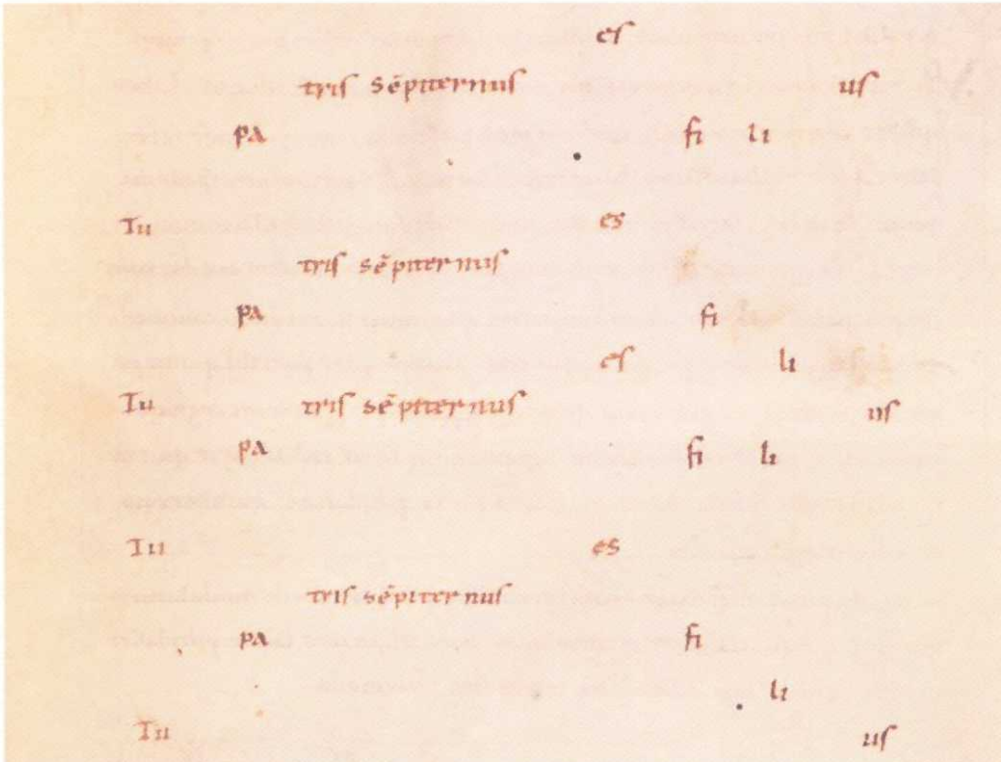
(*Top:*) Track 25e [7v].
(*Bottom:*) Track 25f [8v].



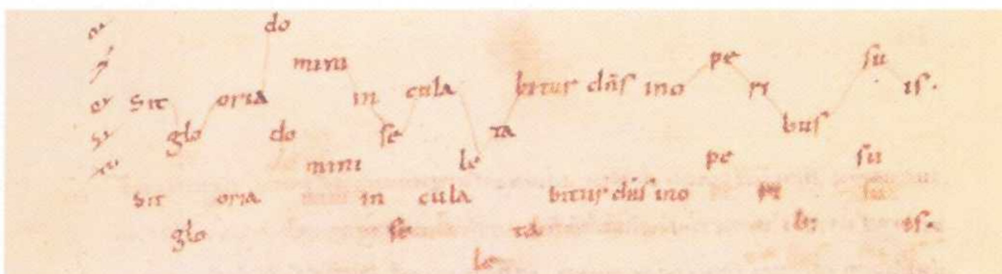
Track 25g [9r].



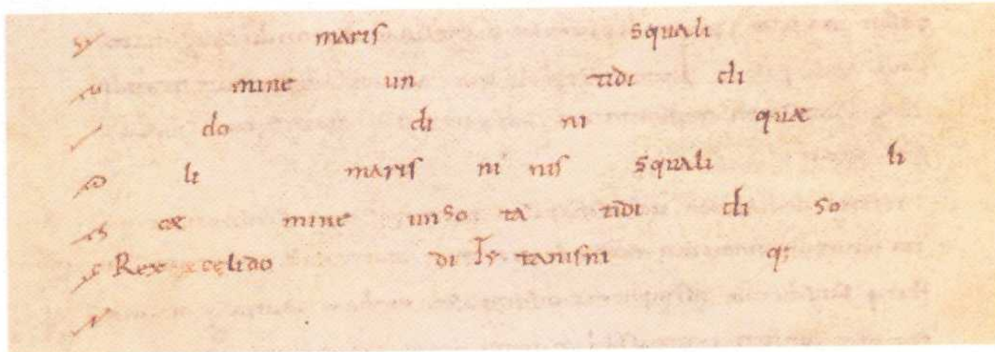
Track 25h [9v].



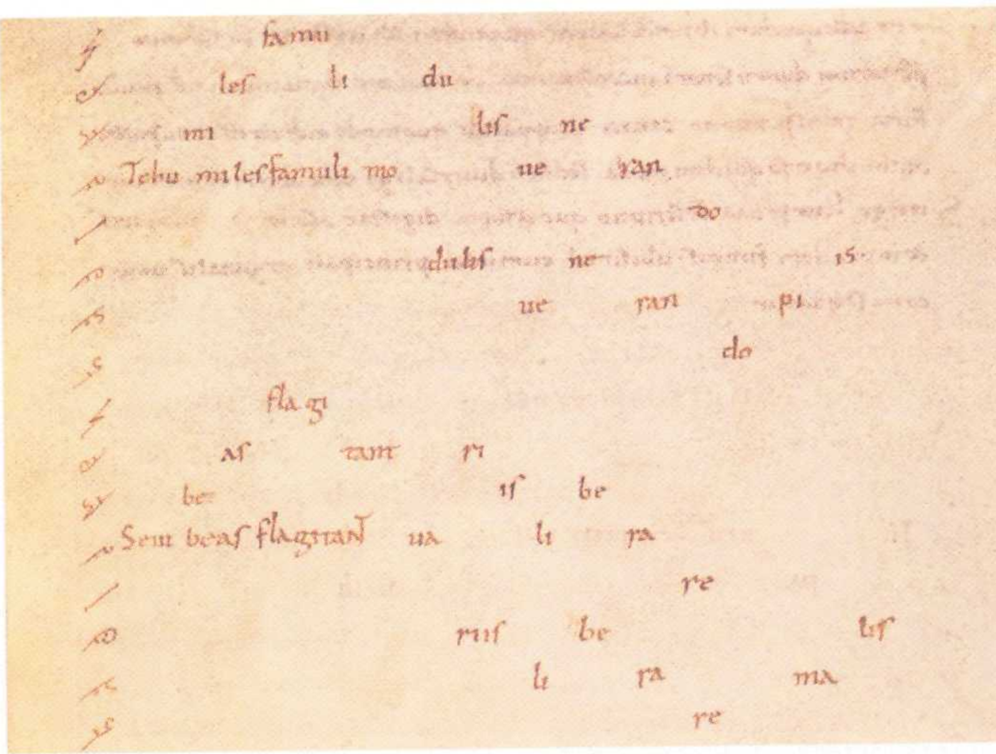
Track 25i [10r].



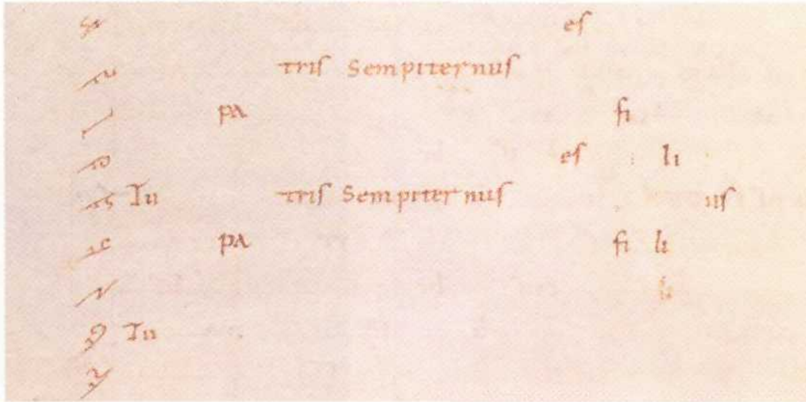
Track 25j [10v].



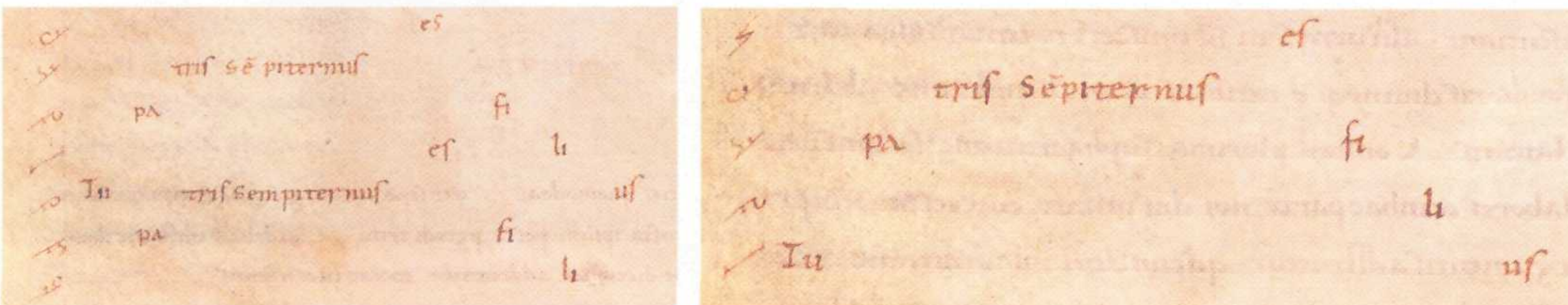
Track 25k [12v].



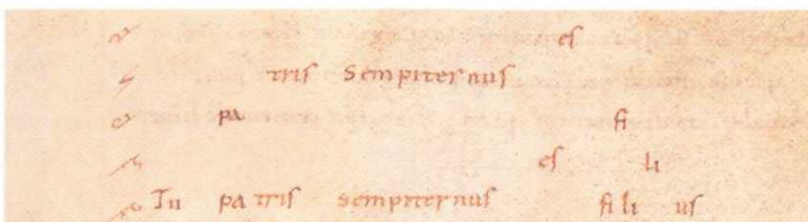
Track 25l [13r].



Track 25m [13v].



Track 25n [14r].



Track 25o [14v].

inscriptions surprisingly well. I've compared some of my eductions to well-informed modern choral performances of the same pieces, and the melodies, the harmonies, and even the rhythms sound remarkably similar. So not only can we play these inscriptions automatically, just as though they were sound recordings, but we can also do so *meaningfully*.

(((☼)))

Track 24

"Nos qui vivimus benedicimus dominum ex hoc nunc et usque in seculum." (11×)

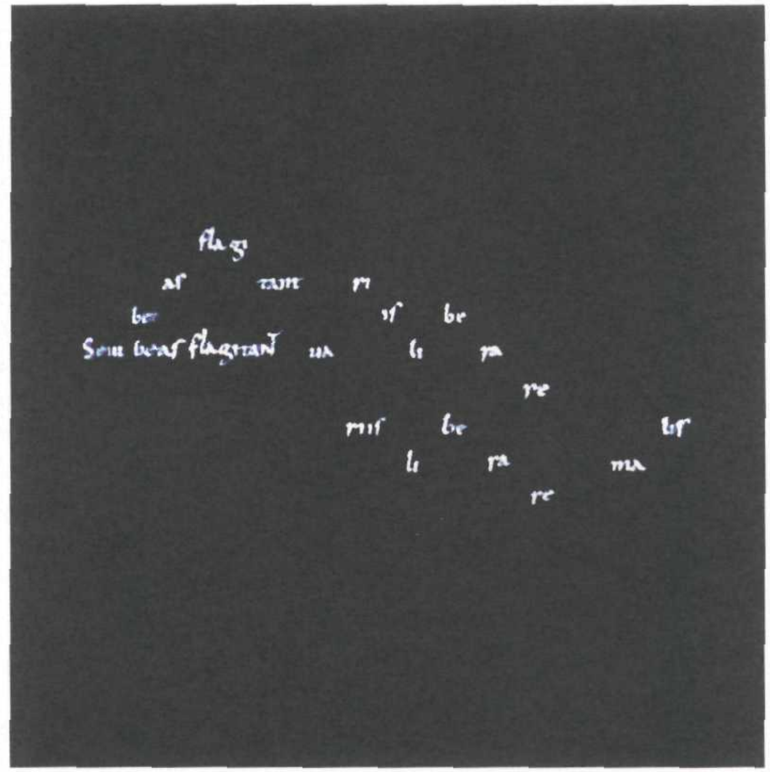
a. [26r] **b.** [26v] **c/d.** [27r] **e.** [27v]. **f.** [28r], **g/h.** [28v], **i.** [29r], **j.** [29v], **k** [30r]

Source: Staatsbibliothek Bamberg Msc. Class. 9.

Track 25

a. Alleluia [3v]
b. Alleluia (4×, separated out) [4r] (0:10)
c. Laudate dominum de cælis.... [4v] (1:02)
d. Alleluia, alleluia, alleluia, alleluia [7r] (2:03)
e. *Tu patris sempiternus es filius* [7v] (2:17)
f. Tu patris sempiternus es filius [8v] (2:31)
g. Tu patris sempiternus es filius (5×, separated out) [9r] (2:44)
h. Tu patris sempiternus es filius [9v] (3:44)
i. Tu patris sempiternus es filius [10r] (3:57)
j. Sit gloria domini in secula.... [10v] (4:10)
k. Rex cæli domine.... [12v] (4:24)
l. Te humiles famuli.... [13r] (4:38)
m. Tu patris sempiternus es filius [13v] (5:04)
n. Tu patris sempiternus es filius (2×) [14r] (5:16)
o. Tu patris sempiternus es filius [14v] (5:40)

Source: Staatsbibliothek Bamberg Msc. Class. 9.



Track 25l, image processed for eduction.

(((((((26. MELODY AND MEASURE OF SPEECH)))))))

(1:36)

In 1857, the scientific journalist Louis Figuier wrote a piece about Scott's phonautograph for a Paris newspaper in which he ruminated as follows:

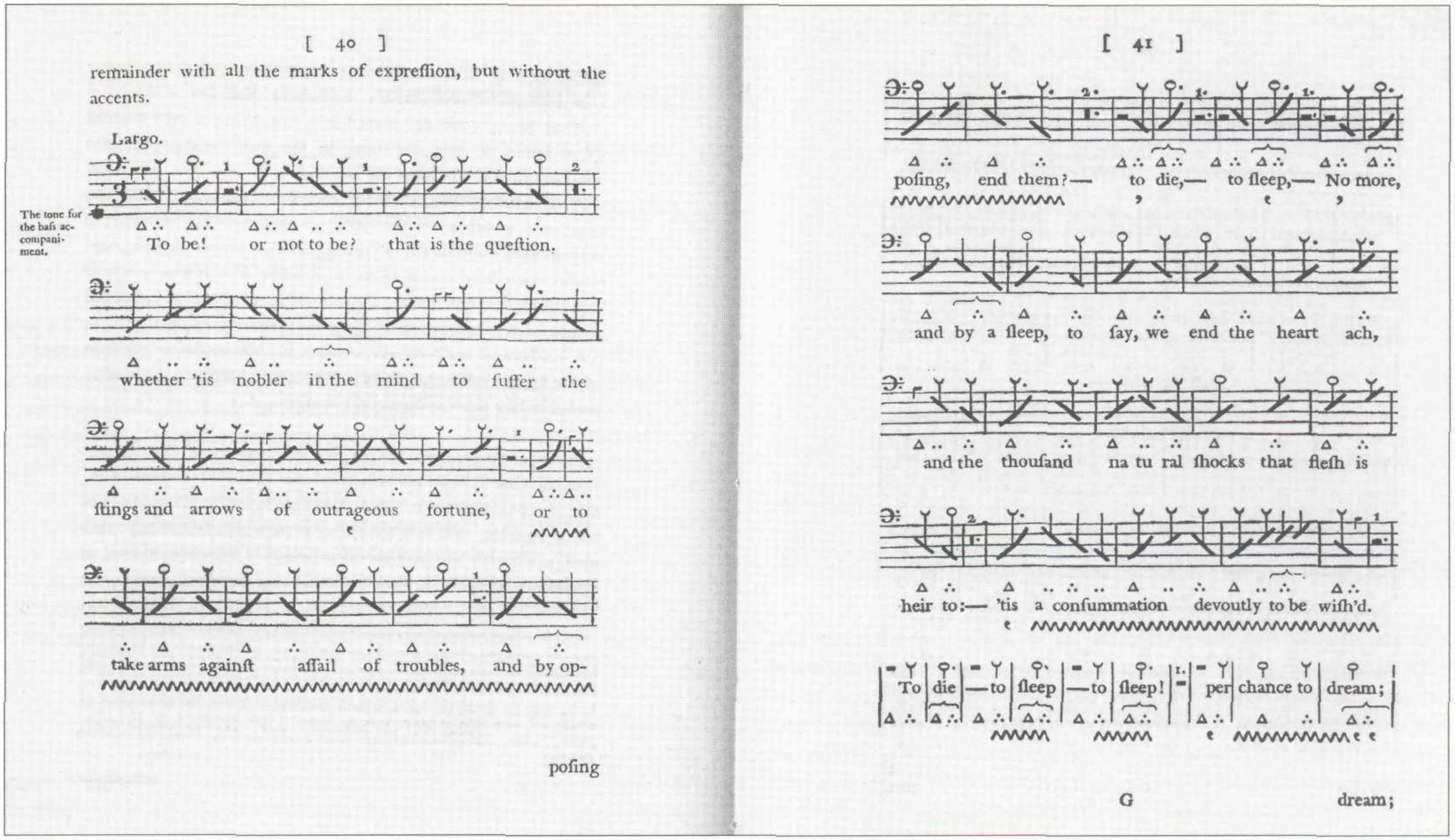
> Writing and print express speech, it is true, but speech dead and faded. You chance to hear recited some beautiful verses by Ligier [*a celebrated actor*]: write them down and give them to a child to read, you will no longer recognize them. To restore life to them, it would have been necessary to accentuate them, to note them as in music; still, the goal would only have been very imperfectly reached.[69]

Of course, Figuier was only setting the stage for his announcement that the phonautograph would be able in the future to record all the expressive nuances of dramatic oratory with ease. But what interests me here is his remark about the prospect of "noting" verses "as in music"—the original verb is *noter,* the same one Engramelle had used to refer to "noting" music for automatic instruments on cylinders. As it happens, there had in fact been a few efforts in the past to "note" dramatic oratory using systems based on musical notation. The best known was probably the scheme outlined by Joshua Steele in his *Essay Towards Establishing the Melody and Measure of Speech to be Expressed and Perpetuated by Peculiar Symbols* (1775):

> We have heard of Betterton, Booth, and Wilks, and some of us have seen Quin [*these were all famous British actors of the late seventeenth and early eighteenth centuries*]; the portraits of their persons are probably preserved, but no models of their elocution remain; nor any proofs, except vague assertions and arbitrary opinions, to decide on the comparative merits in the way of their profession, between them and the moderns. Had some of the celebrated speeches from Shakespeare been noted and accented as they spoke them, we should be able now to judge, whether the oratory of our stage is improved or debased. If the method, here assayed, can be brought into familiar use, the types of modern elocution may be transmitted to posterity as accurately as we have received the musical compositions of Corelli.[70]

Steele's method involved drawing the "melody" of speech on a musical staff in the form of curved lines. These curves were also attached to arbitrary markers of duration, with additional marks drawn in underneath to represent emphasis and force, but by themselves the curves closely resemble modern pitch contours in appearance and meaning.

One of the dramatic texts Steele used to illustrate his system was Hamlet's Soliloquy. "It is so many years since I saw the tragedy of Hamlet performed," he wrote, "that I have no remembrance of the expressions sufficient to enable


[40]

remainder with all the marks of expreſſion, but without the accents.

Largo.

The tone for the baſs accompaniment.

To be! or not to be? that is the queſtion,

whether 'tis nobler in the mind to ſuffer the

ſtings and arrows of outrageous fortune, or to

take arms againſt aſſail of troubles, and by op-

poſing

[41]

poſing, end them?— to die,— to ſleep,— No more,

and by a ſleep, to ſay, we end the heart ach,

and the thouſand na tu ral ſhocks that fleſh is

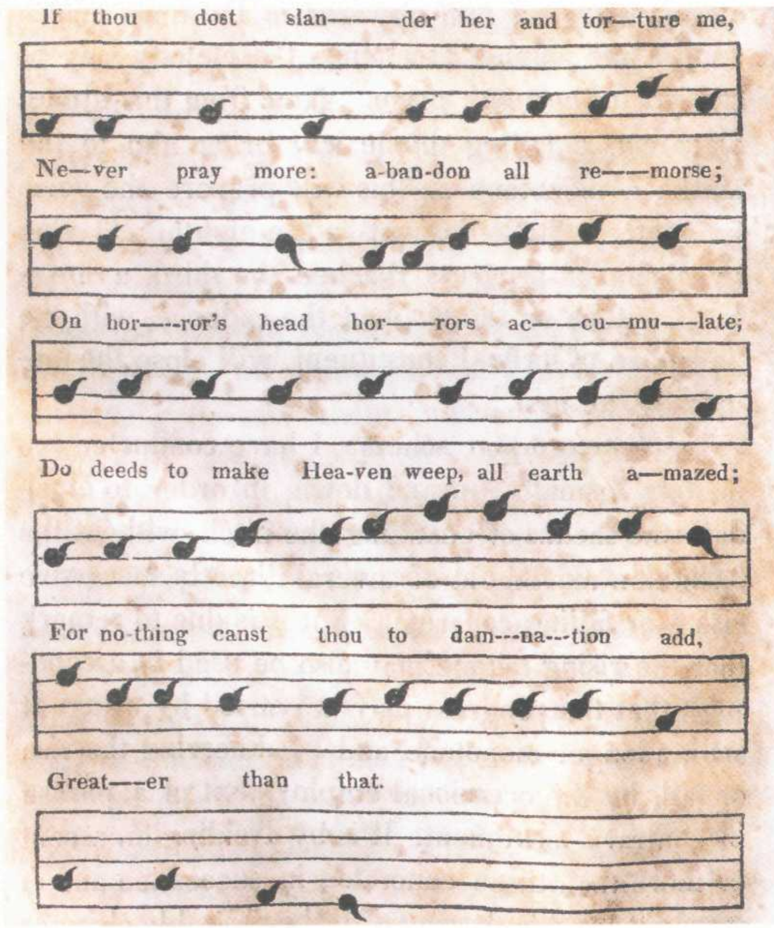
heir to:— 'tis a conſummation devoutly to be wiſh'd.

To die —to ſleep —to ſleep! per chance to dream;

G

dream;

"To be, or not to be...." *An Essay Towards Establishing the Melody and Measure of Speech to be Expressed and Perpetuated by Peculiar Symbols*, pgs. 40-41.



"If thou dost slander her...."
The Philosophy of the Human Voice, page 134.

me to set the following speech in the manner of any great actor." Instead, he transcribed it as he had delivered it himself during some of his experiments.[71] In the first part of track twenty-six, I've educed the curves exactly as they appear on the page in Steele's notation of the Soliloquy, disregarding the additional markers of duration, emphasis, and force. At best, I'm sure the results correspond only very loosely to the style of delivery Steele had in mind—but where else are you going to find an automatically playable representation of theatrical speech from late eighteenth-century England? The American elocutionist James Rush took a similar approach to noting speech in *The Philosophy of the Human Voice* (1827). In his modified staff notation, each line or space was supposed to represent a full tone, corresponding nicely to a logarithmic frequency scale. However, Rush made no effort to show timing, and the shapes of his notes weren't supposed to function like pitch contours to the same extent as Steele's: "The whole of this notation being mere metaphor," he wrote, "there is no meaning in the curve given to the sign of the vanish [*i.e., the tail*]."[72]

The second part of track twenty-six is based on Rush's notation of a passage from *Othello* which, as he stated, "shows the progress of the voice, through a compass of nine diatonic degrees: the rule of the rise and fall being observed, and the melody being therein agreeably diversified." Once again, I don't want to claim too much for what we're hearing. Apart from its other deficiencies, Rush himself stressed that his notation of this passage was "restricted to an exemplification of the means for moving through the compass of the voice," and that it didn't begin to address "the emphatic expression of this forcible passage."[73] But it's still exciting to have access to an automatically playable representation of American theatrical speech from the 1820s, however imperfect it might be.

((( )))

a. "To be, or not to be...." Source: Joshua Steele, *An Essay Towards Establishing the Melody and Measure of Speech to be Expressed and Perpetuated by Peculiar Symbols* (London: J. Almon, 1775), 40-41.

b. "If thou dost slander her...." Source: James Rush, *The Philosophy of the Human Voice* (Philadelphia, J. Maxwell, 1827), 134.

(((((((27. WHAT HATH GOD WROUGHT?)))))))

(1:01)

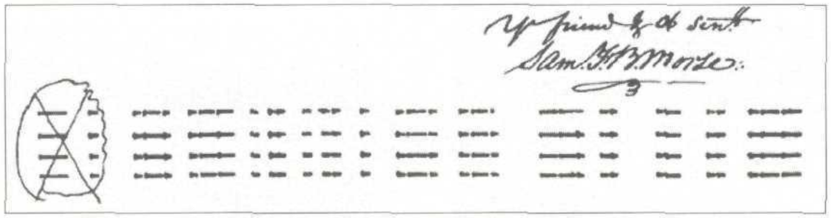
The inventor Samuel Morse sent one of his business partners a letter on 13 April 1838 to report his latest progress in developing a new system of electromagnetic telegraphy. At the bottom, he included an actual record of a telegraphic transmission, which he explained as follows:

> The pen which is so divided as to make 4 copies of the intelligence at the same moment, is also newly invented; you have a specimen of the first trial of the machinery, below; you have the alphabet, so I shall let you decypher it, only remarking that the dots are a little elongated and have, (as well as the dashes,) a blunt beginning which will not belong to the sign when properly adjusted. You will also take care in reading the signs to observe the difference, (4 kinds) of spaces or intervals.[74]

The sample message—SUCCESS TO YOU—employs an early version of Morse Code that differs greatly from the one later put into practical use. At this stage in the evolution of the telegraph, it would also have been transmitted using a "port rule" to run a wire along a previously composed line of moveable type with raised teeth, and not by tapping a key.

On 24 May 1844, Morse more famously transmitted the words WHAT HATH GOD WROUGHT to Baltimore from the Capitol Building in Washington, officially opening the telegraph line linking the two cities. This was the first message that had ever been sent on a completed working telegraph line in the United States, so its transmission arguably marks the birth of modern American telecommunications. The text was a Biblical quotation (Numbers 23:23) that had been chosen for the purpose by Annie Ellsworth, the daughter of Henry Leavitt Ellsworth, first Commissioner of the United States Patent Office. The code itself had taken on a more familiar form since 1838, and the recording apparatus in Baltimore now used a steel point to emboss messages in three parallel lines. I haven't seen any conclusive evidence as to how Morse actually composed the famous message for transmission, but he later claimed to have used a "finger-key" in Washington "previous to June, 1844,"[75] and most sources seem to take for granted that he tapped the message out in real time. The original embossed paper strip is preserved today at the Library of Congress.

At first, Morse assumed telegraph operators would decode transmitted messages visually by reading the dots and dashes recorded on paper strips, but within a few years some operators had learned to decode the letters by ear as they came in, which eventually became standard practice. Thus, although the telegraph soon became a "sound medium" of sorts, it was initially perceived and treated as a visual medium.



Letter from Samuel F. B. Morse to Francis O. J. Smith,
13 April 1838.

First Telegraphic Message,
Samuel F. B. Morse Papers, Library of Congress.

But the recorded telegraphic dots and dashes of the earlier period still document *rhythms unfolding over time*—whether the rhythmic striking of a wire against the teeth of a "port rule" or the manual tapping of a telegraph key. And today we can play those recorded rhythms back.

In track twenty-seven, I've educed the two Morse Code records described above (including a crossed-out mistake at the beginning of the first example) using a modified form of paleospectrophony. Of course, the recorded dots and dashes don't contain any information corresponding to frequency; they tell us only whether or not an electric current was present at any given moment. So instead of basing the additive synthesis on sine waves in this case, I've based it on randomly generated noise. The duration is set according to the average transmission speeds I've seen cited for early Morse telegraphy: ten words per minute in 1838 and thirty letters per minute in 1844.

Morse's transmission of the words WHAT HATH GOD WROUGHT by telegraph from Washington to Baltimore was a watershed moment in the history of modern telecommunications, but I believe it's also significant for another reason. To the best of my knowledge, it was the first event of such historical magnitude to be recorded automatically in a form we can "reproduce" today.

(((✺)))

a. Source: Letter from Samuel F. B. Morse to Francis O. J. Smith, 13 April 1838, reproduced in facsimile in J. Cummings Vail, ed., Early History of the Electro-Magnetic Telegraph (New York: Hine Brothers, 1914), facing pages 16 and 17, at 17.

b. Source: "First Telegraphic Message," Samuel F. B. Morse Papers, Library of Congress. The handwritten inscription at the top reads: "This sentence was written from Washington by me at the Baltimore Terminus at 8h 45 min A. M. on Friday May 24th 1844, being the first ever transmitted from Washington to Baltimore by Telegraph and was indited by my much loved friend Annie G. Ellsworth. Saml F B Morse, Superintendent of Elec. Mag. Telegraphs."

((((((28. BARCELONNETTE))))))

(1:23)

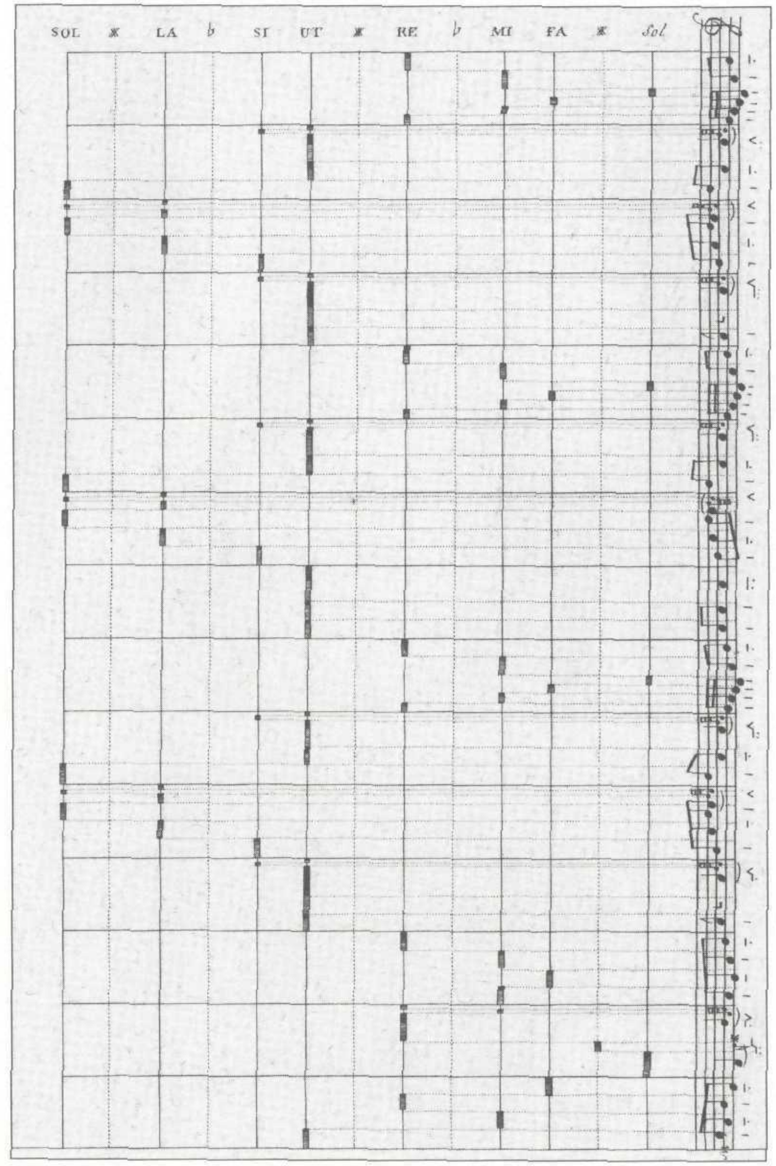
In my notes on track one, I claimed that I'd played Engramelle's *notage* of Balbastre's "Romance" just as though it were a sound recording. By now, I'm sure you've long since figured out what I meant by this: namely, that I'd educed it paleospectrophonically (and with stereo channels separated one octave). But I also had to carry out one special processing step that I haven't yet described, and I'd like to explain what it was in connection with a shorter and simpler piece of *notage* taken from the same book—this one representing a tune called the "Barcelonnette."

In programming this tune, Engramelle was torn between two conflicting goals. He wanted to show how to note music for the serinette (or canary-organ), which had a maximum duration of twenty seconds; but the complete "Barcelonnette," played with all its repeats, was twice as long as a serinette could accommodate. He could have cut it to the right length by leaving out all the repeats, but he also wanted to illustrate the art of *notage* in as rich and varied a manner as he could, which he realized he could accomplish better by leaving the repeats in and ornamenting them in different ways. In the end, he furnished a twenty-second version of the "Barcelonnette" in conventional staff notation for use in actual serinettes, but he extended the *notage* itself out to twice that length, stressing that this was intended only for purposes of instruction. He wrote further:

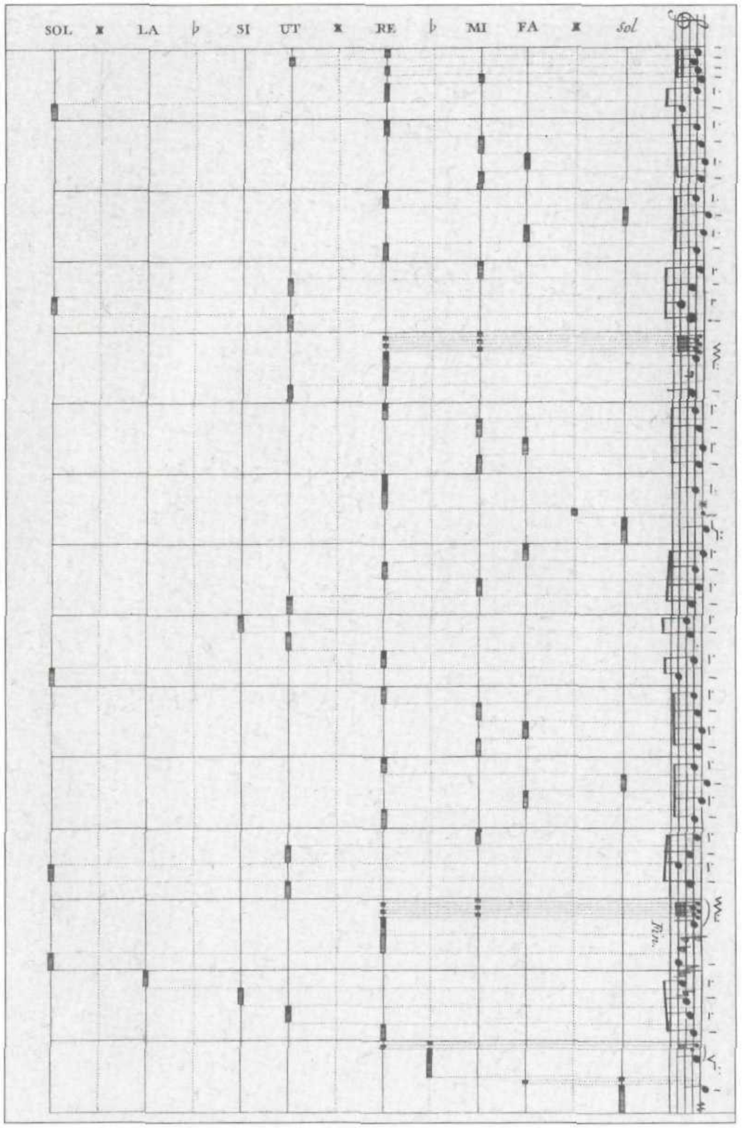
> Plates 114, 115, 116, 117 and 118 represent the whole air marked or pointed on the cylinder, of which they are so many fragments. All the black parts show the *holding* or sound of the notes, and the parts hatched with a simple stroke show the *silences* which are in their turn to complete the value thereof.[76]

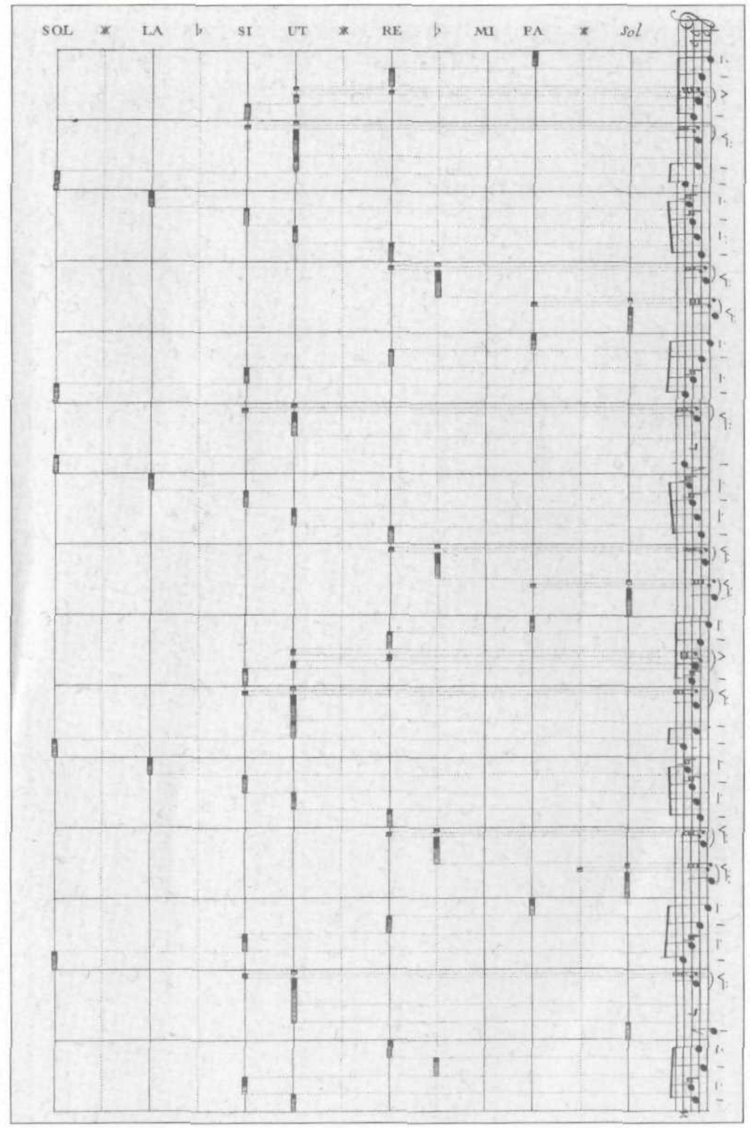
That second sentence is important. According to Engramelle, the total duration of each musical note was divided in performance into two separate parts: a sounded part followed by a silent part. He included both of these parts in his *notage* when graphing it out on paper, but with different shadings to show which parts were to be sounded and which parts were to be silent. In his view, a musical performance without silences would have been wholly unacceptable, as he asserted in connection with Balbastre's "Romance":

> Assuredly, if this Romance were noted on a cylinder, or even played in this fashion by a musician, on whatever instrument it might be, it would not be bearable, and it would repulse all those who heard it.
>
> The lack of *silence of articulation* would leave no separation between the notes, which would confuse everything to the point of making it nothing but a muddled hum: the sixty-fourth notes for

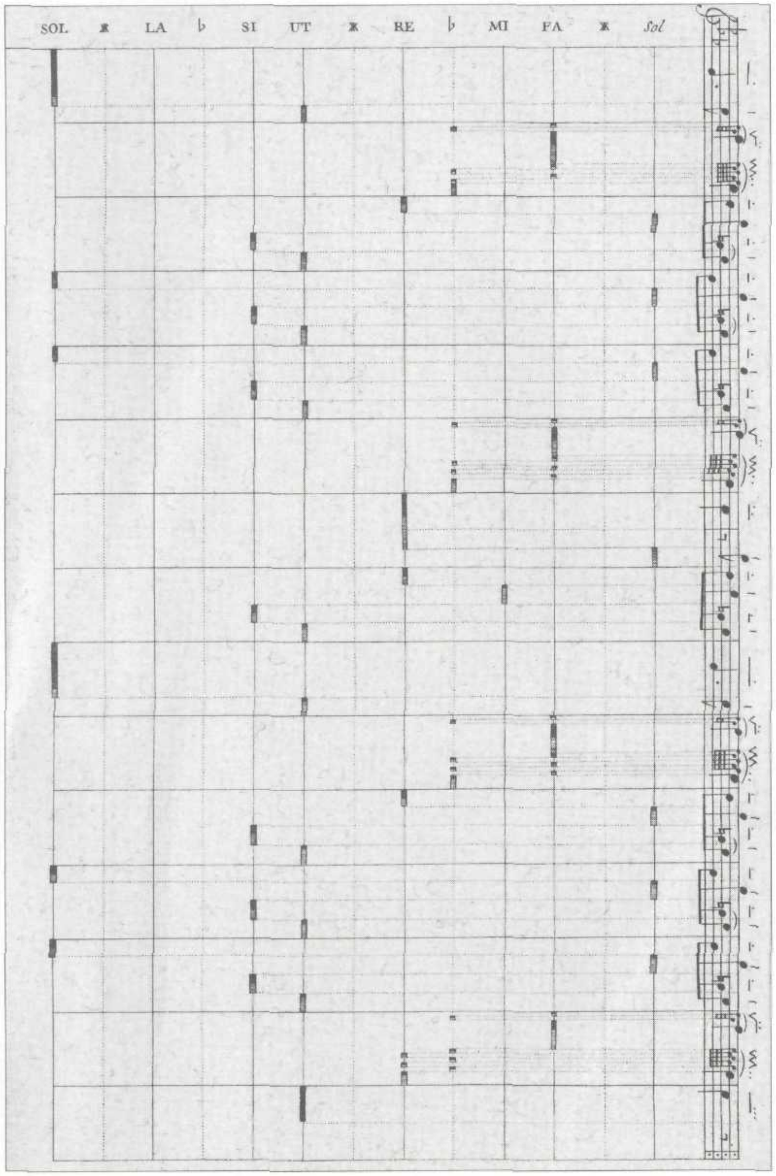


L'art du Facteur d'Orgues, Vol. 4, Plates 114 and 115.





L'art du Facteur d'Orgues,
Vol. 4, Plates 116 and 117.



L'art du Facteur d'Orgues, Vol. 4, Plate 118.

> the trill would become so rushed that they would make nothing but a sort of repulsive chirping, instead of a flattering and agreeable modulation.[77]

As an illustration of Engramelle's argument, I've educed his *notage* of the "Barcelonnette" twice in track twenty-eight: first with all the marks corresponding to notes left in, and then with all the marks corresponding to silent parts erased. (Stereo channels are separated by one octave throughout.) Which of the two versions do you prefer? Does the first version "repulse" you, as Engramelle predicted?

The rendition of Balbastre's "Romance" presented in track one was based on images from which all parts of notes marked as silent were erased. I assume this is how Engramelle would have wanted it. In rare instances, the visual transition from the sounded part to the unsounded part was a little ambiguous in the copies of the plates I had at my disposal, but I think the margin of error would be only about a millisecond. Virtually all the rhythmic irregularities you hear represent actual irregularities of spacing in the original engravings. As David Fuller states, inscriptions like these are "real recordings from the past, and they are completely free of any ideas about performance conceived since the 18th century."[78] And if I erased slightly too much or too little of any of the "silences," I certainly wasn't being guided by any preconceptions about how an ideal musical performance ought to sound.

(((☼)))

Source: Dom François Bedos de Celles, *L'art du Facteur d'Orgues*, Vol. 4 (Paris: L. F. Delatour, 1778), Plates 114-118.

(((((((ENDNOTES)))))))

[1] Thomas H. Maugh, "Sound Recording is Older than You Thought," *Los Angeles Times,* 29 March 2008.

[2] IstvanN1961, http://www.youtube.com/watch?v=J4-Bm6tAM00, accessed 12 May 2012.

[3] Dom François Bedos de Celles, *L'art du Facteur d'Orgues,* Vol. 4 (Paris: L. F. Delatour, 1778), 596.

[4] Alexandre Tolhausen, *Dictionnaire technologique dans les langues française, anglaise et allemande* (Leipzig: Bernhard Tauchnitz, 1883), 575.

[5] See e.g., François Bédos de Celles, *The Organ-Builder,* translated by Charles Ferguson (Raleigh: The Sunbury, 1977), 1:326ff.

[6] Helmut Kowar, "Sound Recordings as a Tool for Musicological Research into Musical Automata," *ARSC Journal* 38 (Fall 2007): 205.

[7] Bedos, *Facteur d'Orgues*, 4:596-7.

[8] Bedos, *Facteur d'Orgues*, 4:620.

[9] Bedos, *Facteur d'Orgues*, 4:632.

[10] David Fuller, *Mechanical Musical Instruments as a Source for the Study of* Notes Inégales (Cleveland Heights, Ohio: Divisions, 1979), 15.

[11] Walter Mendelson, "Preface," in Fuller, *Mechanical Musical Instruments*, 4.

[12] Ralph K. Potter, George A. Kopp, and Harriet C. Green, *Visible Speech* (New York: D. Van Nostrand, 1947).

[13] *On the Theory and Practice of Voice Identification* (Washington DC: National Academy of Sciences, 1979), 6.

[14] *Voice Identification,* 81.

[15] *Voice Identification*, 6.

[16] "AudioPaint," http://www.nicolasfournel.com/audiopaint.htm, accessed 12 May 2012.

[17] *Voice Identification*, 3, 75.

[18] Jared Verner and Robert N. Lehman, *Identifying Individual Bald Eagles with Voiceprints: a feasibility study*, United States Department of Agriculture Research Note PSW-359 (December 1982), [1].

[19] E. W. Scripture, *Researches in Experimental Phonetics: The Study of Speech Curves* (Washington DC: Carnegie Institution of Washington, 1906), 39.

[20] The source disc is identified in Edward Wheeler Scripture, *The Elements of Experimental Phonetics* (New York: Charles Scribner's Sons, 1902), 575.

[21] "ImageToSound," http://www.stopmotion-software.com/freeware.html, accessed Fall 2008 but not available at the time of writing in April and May 2012.

[22] Potter, Kopp, and Green, *Visible Speech*, 315.

[23] Milton Metfessel, *Phonophotography in Folk Music: American Negro Songs in New Notation* (Chapel Hill: University of North Carolina Press, 1928), 11.

[24] Metfessel, *Phonophotography,* 18.

[25] Metfessel, *Phonophotography,* 20-22.

[26] Metfessel, *Phonophotography,* 47-48.

[27] Metfessel, *Phonophotography,* 78, 80.

[28] Metfessel, *Phonophotography,* 29, n.

[29] Eugenie Lineff, *The Peasant Songs of Great Russia as they are in the folk's harmonization* (St. Petersburg, David Nutt, 1905), XIII, italics in original.

[30] Lineff, *Peasant Songs,* XXIX.

[31] Lineff, *Peasant Songs,* XV.

[32] Lineff, *Peasant Songs,* XVIII.

[33] Lineff, *Peasant Songs,* XX.

[34] Lineff, *Peasant Songs,* XXI.

[35] Lineff, *Peasant Songs,* XXII.

[36] J. A. Simpson and E.S.C. Weiner, eds. *The Oxford English Dictionary,* 2nd ed. Oxford: Clarendon Press, 1989 (5:75). The word *eduction* as I use it is synonymous with the technical term *output transduction*.

[37] "WIPO Performances and Phonograms Treaty," http://www.wipo.int/export/sites/www/treaties/en/ip/wppt/pdf/trtdocs_wo034.pdf; "Basic Proposal for the Substantive Provisions of the Treaty for the Protection of the Rights of Performers and Producers of Phonograms to Be Considered by the Diplomatic Conference," http://www.wipo.int/edocs/mdocs/diplconf/en/crnr_dc/crnr_dc_5-chapter1.html, accessed 11 May 2012.

[38] For a more detailed analysis of this recording, see Patrick Feaster and Stephan Puille, "E. Berliner's Grammophon: The Beginnings of the German Talking Machine Industry: Part Two," *The Sound Box* 28:3 (Sept. 2010): 3-7.

[39] *Prometheus* 1:14 (1890), 213.

[40] "The Improved Gramophone," *Electrical World* 12 (18 August 1888), 80.

[41] For more on the history of this practice, see Patrick Feaster, "'A Compass of Extraordinary Range': The Forgotten Origins of Phonomanipulation," *ARSC Journal* 42:2 (Fall 2011): 163-203.

[42] *Almanach Musical, 1776* (Paris: Ruault, [n.d.]), 32-34.

[43] "The National Recording Registry 2002," http://www.loc.gov/rr/record/nrpb/registry/nrpb-2002reg.html

[44] Patrick Feaster and Stephan Puille, "E. Berliner's Grammophon: The Beginnings of the German Talking Machine Industry: Part Two," *The Sound Box* 28:3 (Sept. 2010): 3-7, at page 7.

[45] [Friedrich] Schiller, *The Song of the Bell,* transl. Edwin Ellis Griffin (London: Hutchings and Crowsley, 1886), 11.

[46] "Vereinsversammlung am 26. November 1889," *Elektrotechnische Zeitschrift* 10 (December 1889), 554.

[47] Patrick Feaster, ed. and transl., *The Phonautographic Manuscripts of Édouard-Léon Scott de Martinville* (Bloomington, Indiana: FirstSounds.org, 2009), 5.

[48] Feaster, *Phonautographic Manuscripts,* 44.

[49] Scripture, *Researches,* 13-14.

[50] The piece that kicked off the furor was Jody Rosen, "Researchers Play Tune Recorded Before Edison," *New York Times,* 27 March 2008.

[51] On Scott's important relationship with this organization, see Serge Benoit, Daniel Blouin, Jean-Yves Dupont, Gérard Emptoz, "Chronique d'une invention: le *phonautographe* d'Édouard-Léon Scott de Martinville (1817-1879) et les cercles parisiens de la science et de la technique," *Documents pour l'histoire des techniques* 17 (June 2009): 69-89.

[52] For further information about the items featured in tracks 14-16, as well as a listing of all fifty known Scott phonautograms, see Patrick Feaster, "Édouard-Léon Scott de Martinville: An Annotated Discography," *ARSC Journal* 41:1 (Spring 2010): 43-82.

[53] Paula Findlen, ed., *Athanasius Kircher: The Last Man Who Knew Everything* (New York: Routledge, 2004).

[54] Athanasius Kircher, *Musurgia universalis sive ars magna consoni et dissoni in X. libros digesta* (Roma: Grignani Corbelletti, 1650), 313.

[55] In Kircher, *Musurgia,* this phrase appears at the top of alternate pages beginning at 337, replacing "magia phonocamptica" as the title of Book IX.

[56] Kircher, *Musurgia,* 323.

[57] Kircher, *Musurgia,* 328.

[58] [Victor] Hensen, "Ueber die Schrift von Schallbewegungen," *Zeitschrift für Biologie* 23 (1887): 291-302, at 292.

[59] Paul Wendeler, "Ein Versuch, die Schallbewegung einiger Consonanten und anderer Geräusche mit dem *Hensen*'schen Sprachzeichner graphisch darzustellen," *Zeitschrift für Biologie* 23 (1887), 303-320, at 309-310.

[60] Rudolph König, "Die manometrischen Flammen," *Annalen der Physik und Chemie* 146:2 (1872), 161-199, at 176-177.

[61] Edward L. Nichols and Ernest Merritt, "The Photography of Manometric Flames," *Physical Review* 7 (1898), 93-101, at 93-95.

[62] Nichols and Merritt, "Photography," 100.

[63] M. Pouillet, *Notions générales de physique et de météorologie à l'usage de la jeunesse* (Paris: Béchet Jeune, 1850), 386-89.

[64] Thomas Young, *A Course of Lectures on Natural Philosophy and the Mechanical Arts* (London: William Savage, 1807), 1:369.

[65] Young, *Course*, 1:389-91; plate XXV, fig. 353. The plate, like others in the same book, bears the inscription "Pub. by J. Johnson, London 1 July 1806. Joseph Skelton sculp."

[66] Leonhard Euler, *Tentamen novae theoriae musicae* (Petropolis: Academia Scientiarum, 1739), 36-7; Tab. I facing page 36.

[67] [Leonhard Euler], *Letters of Euler on Different Subjects in Physics and Philosophy Addressed to a German Princess,* transl. by Henry Hunter (London: printed for the translator, and for H. Murray, 1795), 1: 14-15.

[68] See Jamie Croy Kassler, *The Beginnings of the Modern Philosophy of Music in England* (Aldershot, Hampshire, England and Burlington, Vermont: Ashgate, 2004), 137-8.

[69] Louis Figuier, "Sciences. Méthode pour la fixation graphique du son, par M. Léon Scott," *La Presse* (Paris), Nov. 28, 1857, pp. 2-3.

[70] Joshua Steele, *An Essay Towards Establishing the Melody and Measure of Speech to be Expressed and Perpetuated by Peculiar Symbols* (London: J. Almon, 1775), 14.

[71] Steele, *Essay,* 39.

[72] James Rush, *The Philosophy of the Human Voice* (Philadelphia, J. Maxwell, 1827), 48-49.

[73] Rush, *Philosophy,* 134-135.

[74] J. Cummings Vail, ed., *Early History of the Electro-Magnetic Telegraph* (New York: Hine Brothers, 1914), facing pages 16 and 17.

[75] "Professor Morse's Account of the Origin and Progress of his Telegraph System," *American Telegraph Magazine* 1:3 (15 December 1852), 97-114, at 114.

[76] Bedos, *Facteur d'Orgues*, 4:610

[77] Bedos, *Facteur d'Orgues*, 4:621

[78] Fuller, *Mechanical Musical Instruments,* 14.

PRODUCED BY
Patrick Feaster and Steven Lance Ledbetter

COMPILED, EDITED, AND EDUCED BY
Patrick Feaster

MASTERED BY
Michael Graves, Osiris Studio

ANNOTATIONS AND ESSAYS
Patrick Feaster

ART DIRECTION AND DESIGN
Barbara Bersche

SPECIAL THANKS TO
Zach Cincotta, David Giovannoni, April G. Ledbetter, Stephan Puille, Ronda L. Sewald

More material is surfacing all the time. For breaking developments
visit www.phonozoic.net and www.firstsounds.org

DUST-TO-DIGITAL
Atlanta, Georgia

Dust-to-Digital
PO Box 54743
Atlanta, GA 30308
www.dust-digital.com

flagi

as tum ri

bet is be

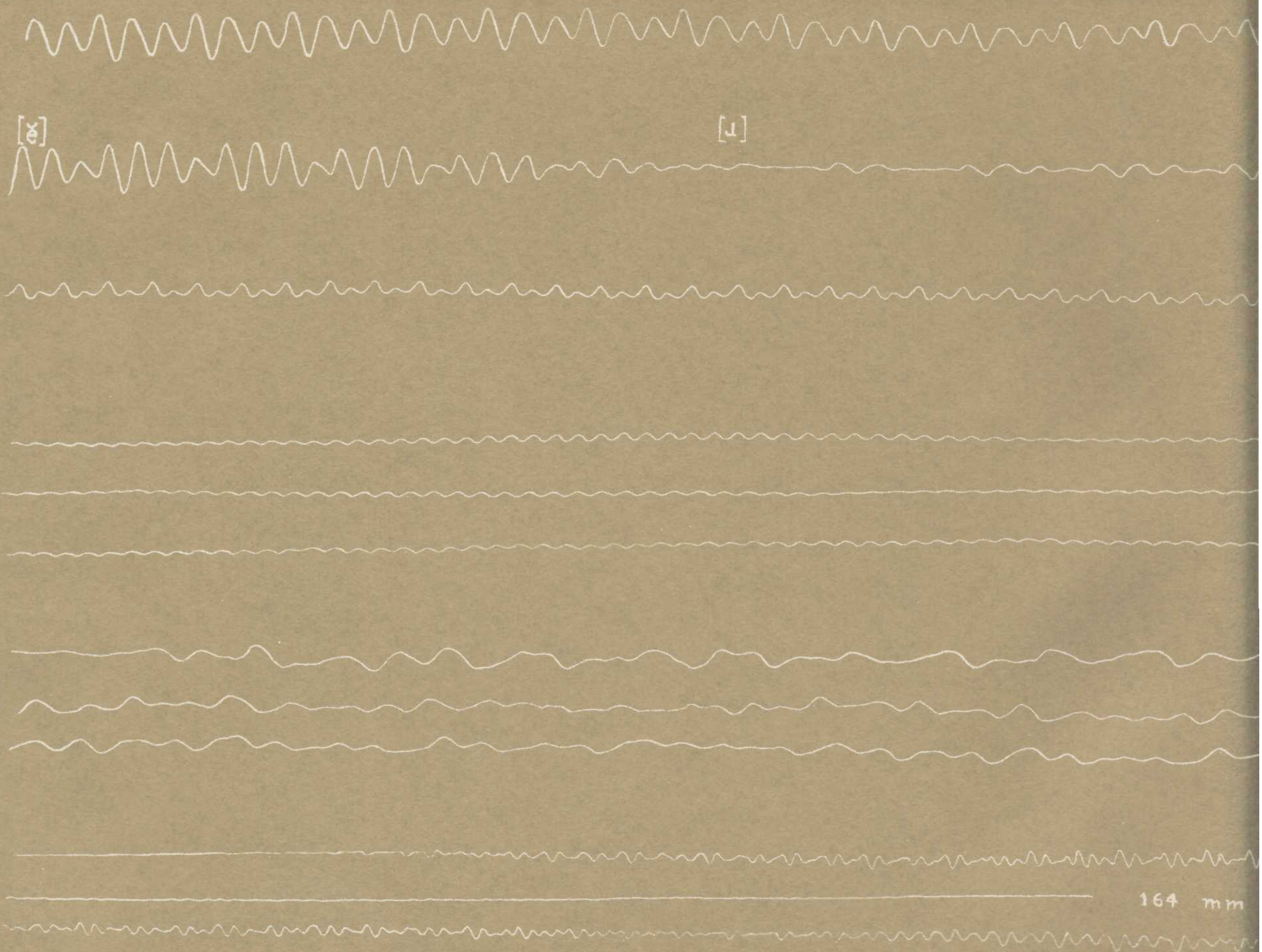
Sem beas flagitant ua li ra

re

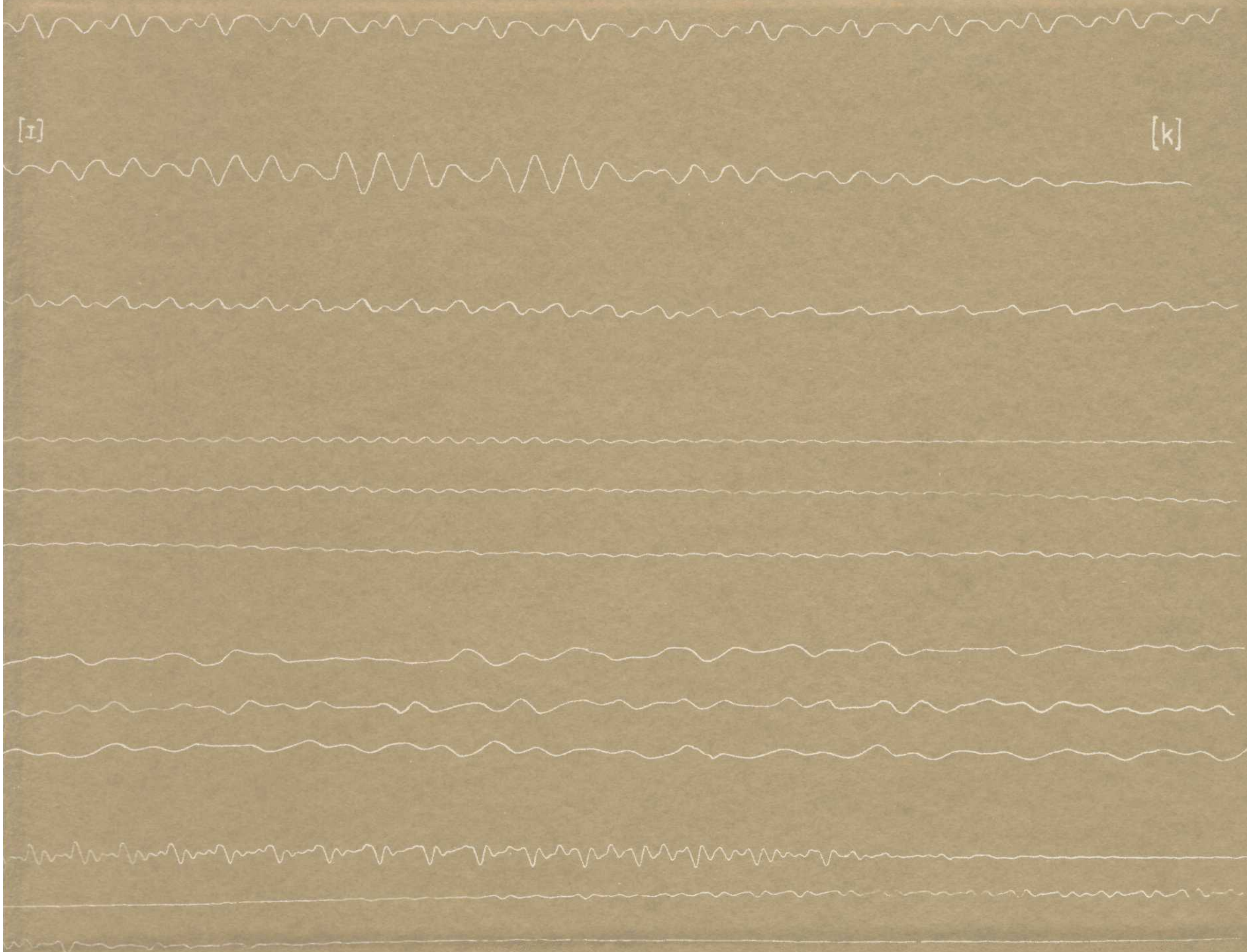
mis be lis

li ra ma

re



[ě]
[ɹ]
164 mm


[I]
[k]