

# *What We're Talking About When We're Talking About Sound*

Tanya Clement, Assistant Professor  
iSchool  
University of Texas at Austin

# HiPSTAS

1. Introducing High Performance Sound Technologies in Access and Scholarship
2. A brief look at current sound/text technologies
3. A brief history of sound technologies and information science
4. Four use cases

# HIGH PERFORMANCE SOUND TECHNOLOGIES FOR ACCESS AND SCHOLARSHIP



[About](#) [Resources](#) [Call for Participation](#) [Organizers](#) [Readings](#) [Meetings](#)

## Welcome to HiPSTAS

By TANYA CLEMENT | Published: NOVEMBER 14, 2012 | [Edit](#)

**The HiPSTAS application is now available!**  
**DEADLINE EXTENDED to February 1, 2013**  
Please [apply](#).

Welcome to HiPSTAS (High Performance Sound Technologies for Access and Scholarship).

### CONTACT

[hipstasinfo\[at\]utlists.utexas.edu](mailto:hipstasinfo[at]utlists.utexas.edu)

### SEARCH

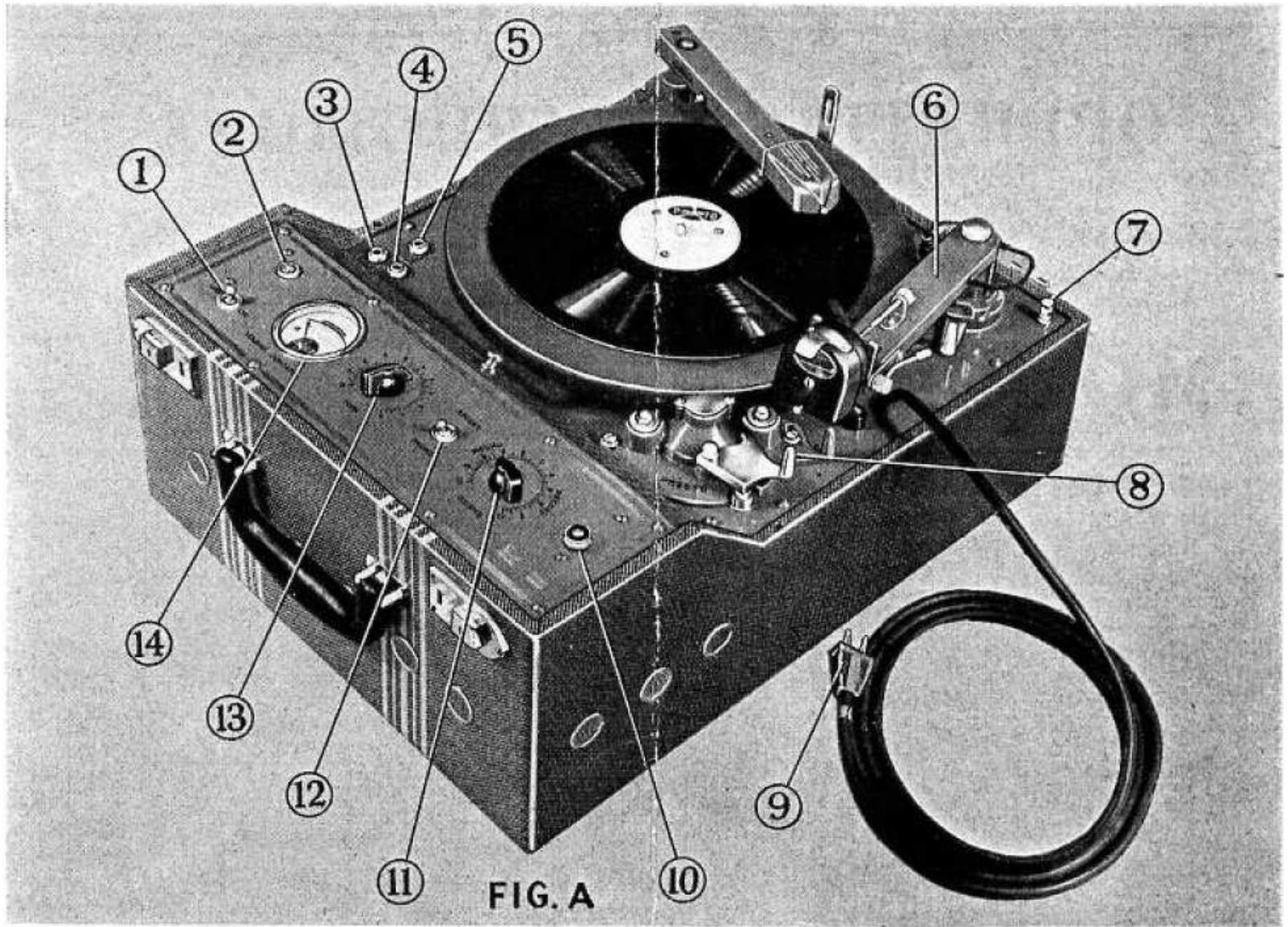
### RECENT POSTS

- [Welcome to HiPSTAS](#)

### META



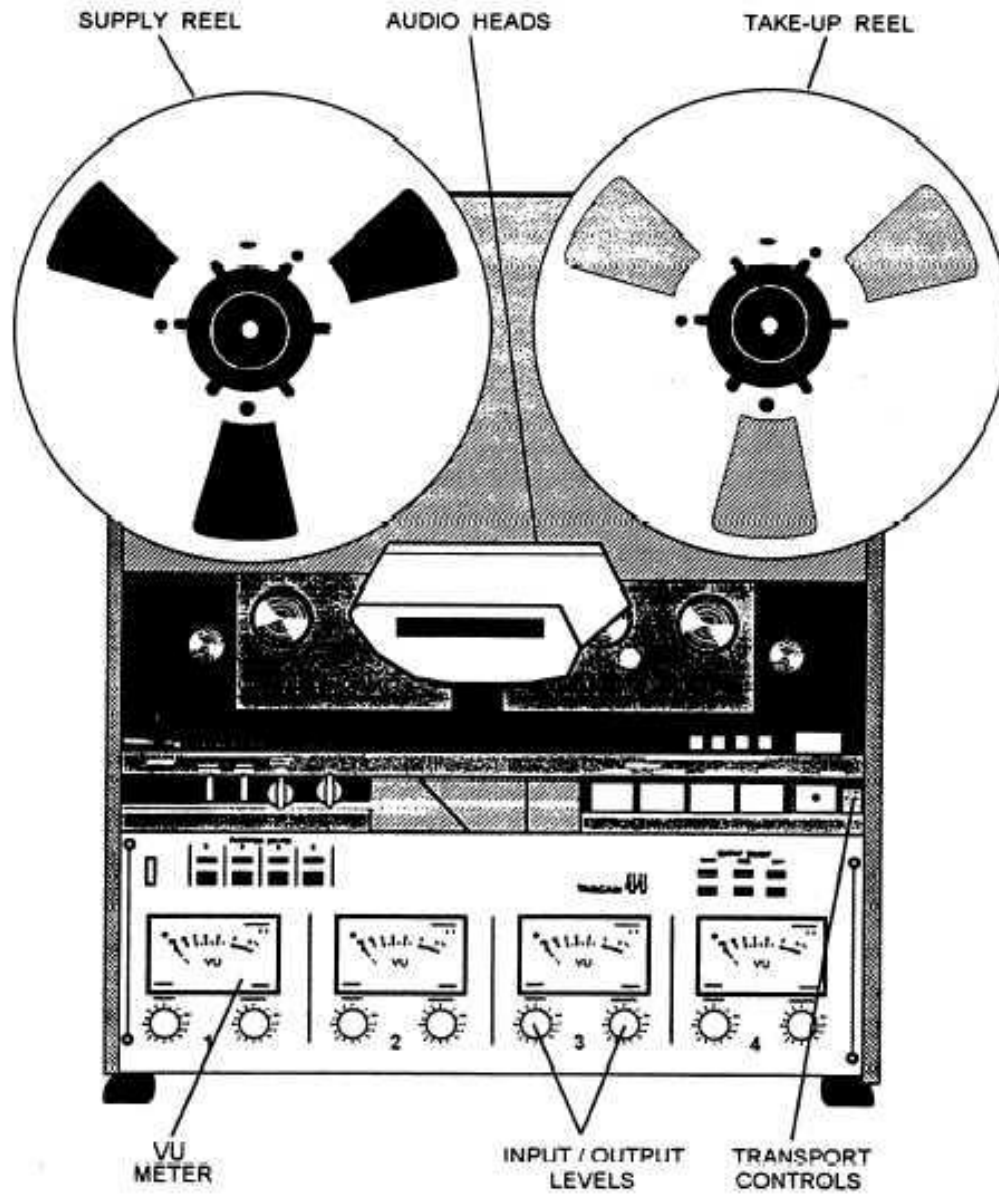
**Recording equipment in the back of John Lomax's car, late 1930s. (American Roots Music 60)**



The 1930s PRESTO Model "K" Recorder



te and aluminum PRESTO disc used by Alan Lomax



**Reel-to-reel tape recorder/reproducer.**

# The State of Recorded Sound Preservation in the United States: A National Legacy at Risk in the Digital Age

T. A. EDISON.  
Phonograph or Speaking Machine  
No. 200,521. Patented Feb. 19, 1878.



*Witnesses*  
*Chas. A. Smith,*  
*Rowland O. Silliman*  
*Inventor*  
*Thomas A. Edison*  
*per Leonard W. Searle*  
1878

August 2010

Commissioned for and sponsored by the  
 **National Recording  
Preservation Board**  
OF THE LIBRARY OF CONGRESS

COUNCIL ON LIBRARY AND INFORMATION RESOURCES  
AND THE LIBRARY OF CONGRESS

<http://www.clir.org/pubs/abstract/reports/pub148>







SCHOOL OF  
information

Stampede at the Texas Advanced Computing Center

# HiPSTAS collections

- poetry from PennSound at the University of Pennsylvania
- folklore at the Dolph the Briscoe Center for American History at UT Austin
- speeches at the Lyndon B. Johnson Library and Presidential Museum (LBJ Library) in Austin
- storytelling traditions at the Native American Projects (NAP) at the American Philosophical Society in Philadelphia

HiPSTAS is NOT about  
searching or analyzing sound  
by searching or indexing text

# Closed Captioning Search

## Query-Free News Search

Monika Henzinger  
Google Inc.  
2400 Bayshore Parkway  
Mountain View, CA 94043  
USA  
monika@google.com

Brian Milch  
UC Berkeley  
Computer Science Division  
Berkeley, CA 94720-1776  
USA  
milch@cs.berkeley.edu

Bay-Wei Chang  
Google Inc.  
2400 Bayshore Parkway  
Mountain View, CA 94043  
USA  
bay@google.com

Sergey Brin  
Google Inc.  
2400 Bayshore Parkway  
Mountain View, CA 94043  
USA  
sergey@google.com

### ABSTRACT

Many daily activities present information in the form of a stream of text, and often people can benefit from additional information on the topic discussed. TV broadcast news can be treated as one such stream of text; in this paper we discuss finding news articles on the web that are relevant to news currently being broadcast.

We evaluated a variety of algorithms for this problem, looking at the impact of inverse document frequency, stemming, compounds, history, and query length on the relevance and coverage of news articles returned in real time during a broadcast. We also evaluated several postprocessing techniques for improving the precision, including reranking using additional terms, reranking by document similarity, and filtering on document similarity. For the best algorithm, 84%-91% of the articles found were relevant, with at least

closed captioning company) and WebTV (now a division of Microsoft), broadcasts URLs in an alternative data channel interleaved with closed caption data [17, 2]. When a WebTV box detects one of these URLs, it displays an icon on the screen; if the user chooses to view the page, the WebTV box fetches it over the Internet.

For both of these systems the producer of a program (or commercial) chooses relevant documents by hand. In fact, the producer often creates new documents specifically to be accessed by TV viewers. To our knowledge, there has been no previous work on *automatically* selecting web pages that a user might want to see while watching a TV program.

In this paper we study the problem of finding news articles on

12th International World Wide Web Conference, 2003

## LBJ ORDERS SOME NEW HAGGAR PANTS

August 9, 1964 - 1:17 p.m.

.....

In this telephone call, which has become one of the more famous exchanges LBJ recorded, the President asked a leader of the Haggard clothing company for some custom-made pants, providing specific (and sometimes graphic) instructions on how they should be customized for him.

**LISTEN >>**

Original tape: **WH6408-16 PNO 4**

**DRAFT TRANSCRIPT**

Presidential Recordings Program, Miller Center of Public Affairs, University of Virginia



**President Johnson:** Joe, is your father the one that makes clothes?

**Haggar:** Yes, sir. **We're all together.**

**President Johnson:** Uh-huh. You-all made me some real lightweight slacks that he just made up on his own, sent to me three or four months ago. It's a kind of a light brown and a light green, rather soft green and soft brown.



**Navigation**

- [Main Page](#)
- [Log in](#)
- [Log out](#)
- [All Pages](#)
- [Recent changes](#)
- [Random page](#)
- [Help](#)

Category: LBJ

## President Johnson's Television Report Following Renewed Aggression in the Gulf of Tonkin

Layout: 



03:12 / 06:48

**Mark Video Segment:**  00:00:00    00:00:00

**& Share**  [Mail Video](#)  [Copy](#)  [Bookmark](#) [more...](#)

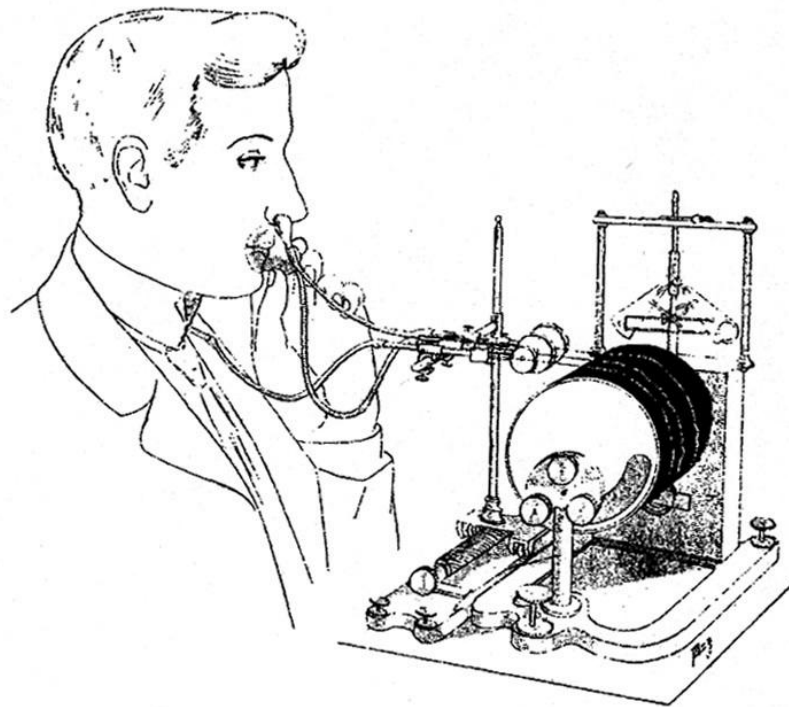
|  |                                  |
|--|----------------------------------|
| <b>Search</b>  | <input type="button" value="🔍"/> |
| <a href="#">Congressional Leadership Meeting Notes</a>   | <input type="button" value="🔍"/> |
| <a href="#">President Johnson's Memoir, Gulf of Tonkin Excerpt</a>   | <input type="button" value="🔍"/> |
| <a href="#">Photo 1: President Johnson's Midnight Address</a>  | <input type="button" value="🔍"/> |
| <a href="#">Lady Bird Johnson's Diary, August 4, 1964</a>  | <input type="button" value="🔍"/> |
| <a href="#">Photo 2: President Johnson's Midnight Address</a>  | <input type="button" value="🔍"/> |
| <a href="#">On this Day in History</a>   | <input type="button" value="🔍"/> |
| <a href="#">Table of Contents</a>  | <input type="button" value="🔍"/> |
| <a href="#">Additional Resources</a>   | <input type="button" value="🔍"/> |
| <a href="#">August 4: Recorded Telephone Conversations</a>   | <input type="button" value="🔍"/> |
| <a href="#">Map</a>  | <input type="button" value="🔍"/> |
| <a href="#">Lady Bird Johnson's Diary transcript</a>   | <input type="button" value="🔍"/> |
| <b>Transcript</b>  | <input type="button" value="🔍"/> |
| <i>There were no U.S. losses. The performance of commanders and crews in this engagement is in the highest tradition of the United States Navy.</i>  |                                  |
| <i>But repeated acts of violence against the Armed Forces of the United States must be met not only with alert defense, but with positive reply.</i> |                                  |

[http://solstice.ischool.utexas.edu/projects/index.php/LBJ\\_Gulf\\_of\\_Tonkin\\_Speech](http://solstice.ischool.utexas.edu/projects/index.php/LBJ_Gulf_of_Tonkin_Speech)



# Based on a history of sound machines, psychoacoustics and information theory

1. The **phonoscope**: reproducing the speaker's voice
2. The **ear phonoautograph**: the middle ear and tympanic instruments
3. The **audiometer**: the inner ear and psychoacoustics



Inscription de la parole.

FIGURE 2. Recording speech, circa 1897. From Abbé Pierre-Jean Rousselot and Fauste Laclotte, *Précis de prononciation* (Paris, 1902), 14.

sensitive instruments capable of capturing entire speech acts—instruments that

**Jean-Pierre Rousselot's "phonoscope" (1880s).** From Richard Sieburth's "The Sound of Pound: A Listener's Guide" on PennSound

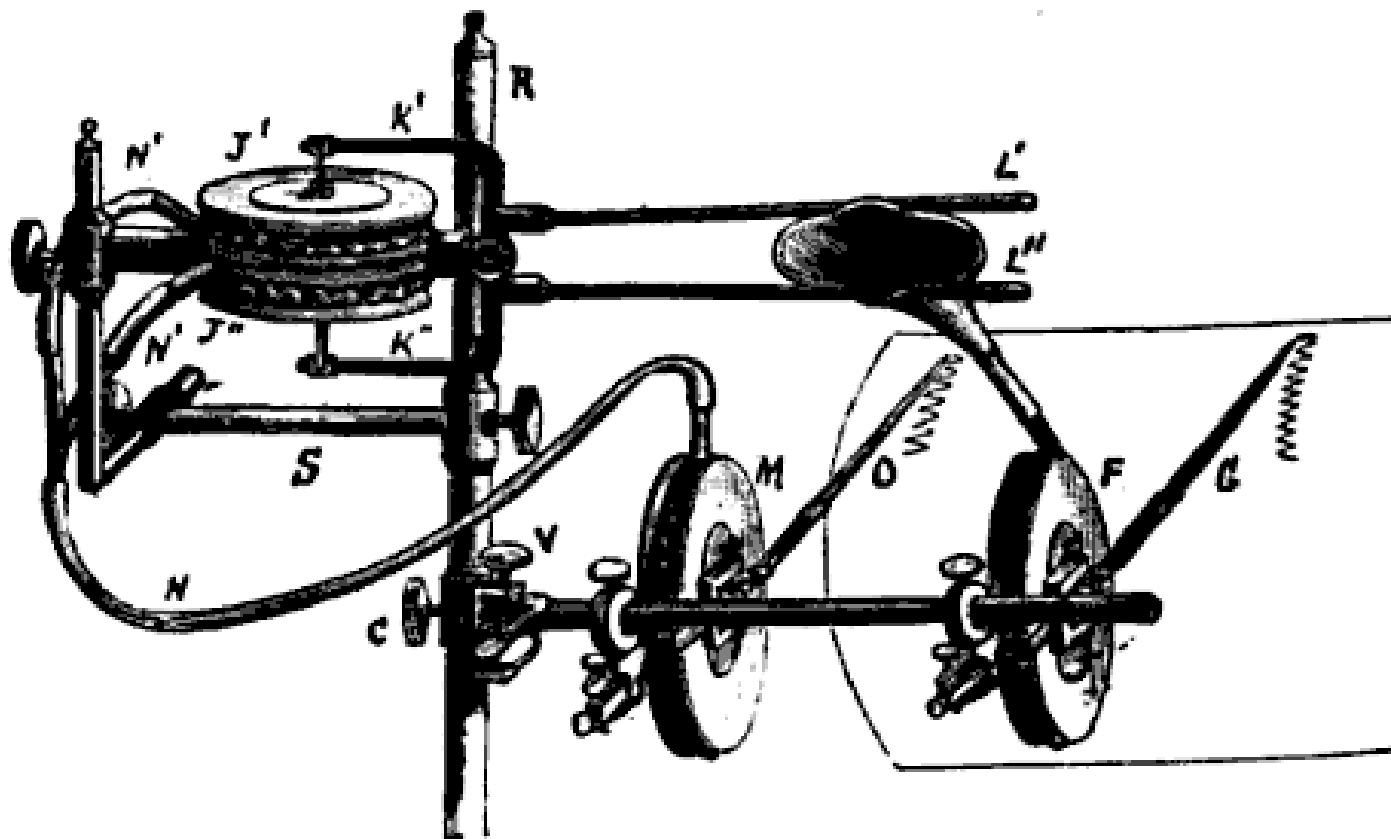


FIG. 278.

Scripture, Edward Wheeler. *The Elements of Experimental Phonetics*. C. Scribner's Sons, 1902: 354.



Fig. 35.

Explorateur du voile du palais.

Rousselot, Pierre Jean. Principes de phonétique expérimentale. H. Welter, 1901: 94

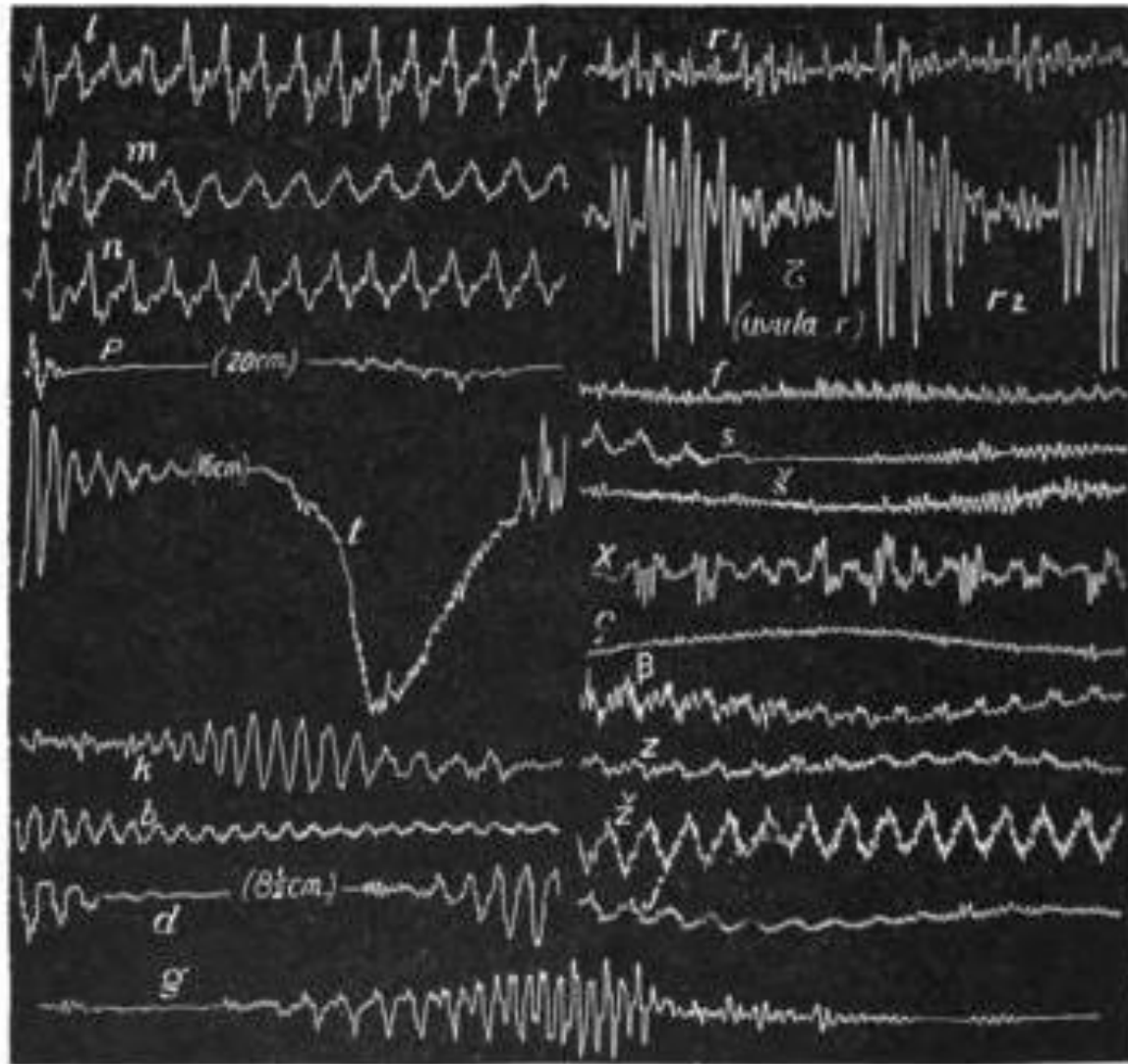
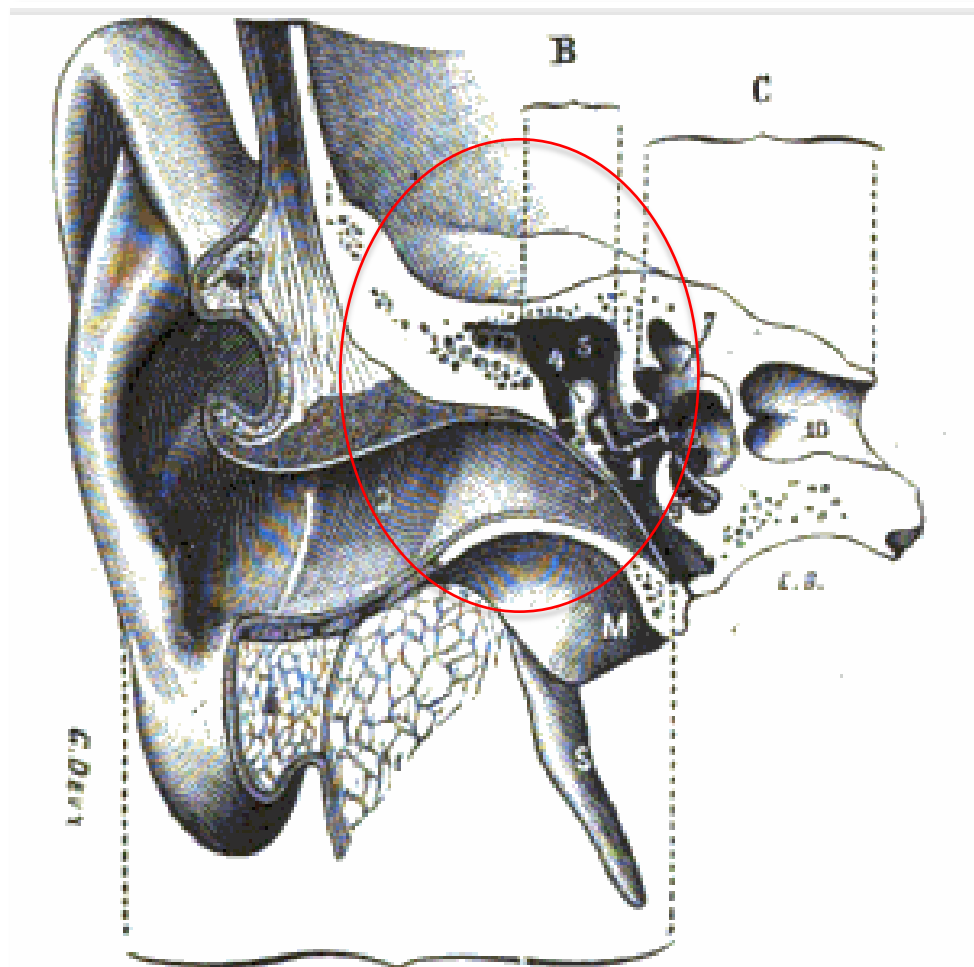


FIG. 33.

Scripture, Edward Wheeler. *The Elements of Experimental Phonetics*. C. Scribner's Sons, 1902: 42.



Alexander Bell and Clarence Blake, ear phonograph, 1874



A (TESTUT, *Anatomie.*)

Fig. 7.

Coupe vertico-transversale de l'appareil auditif.

Oreille externe. — B. Oreille moyenne. — C. Oreille interne. —



Detail of a phonautogram made in 1859 made from Scott's early phonaugraph



# Writing tempo and volume

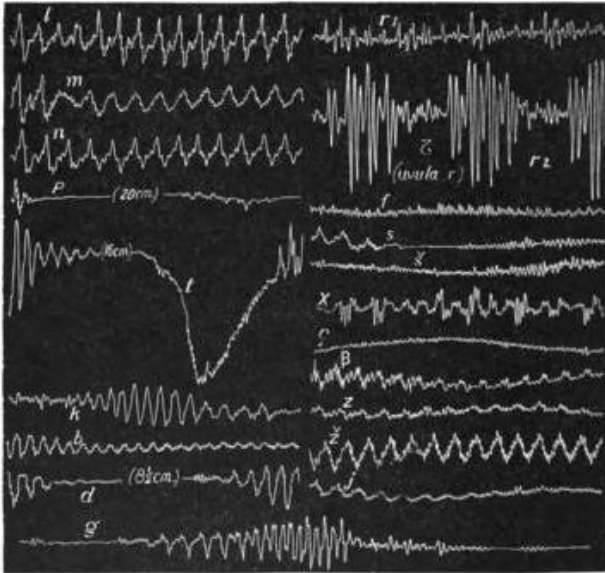
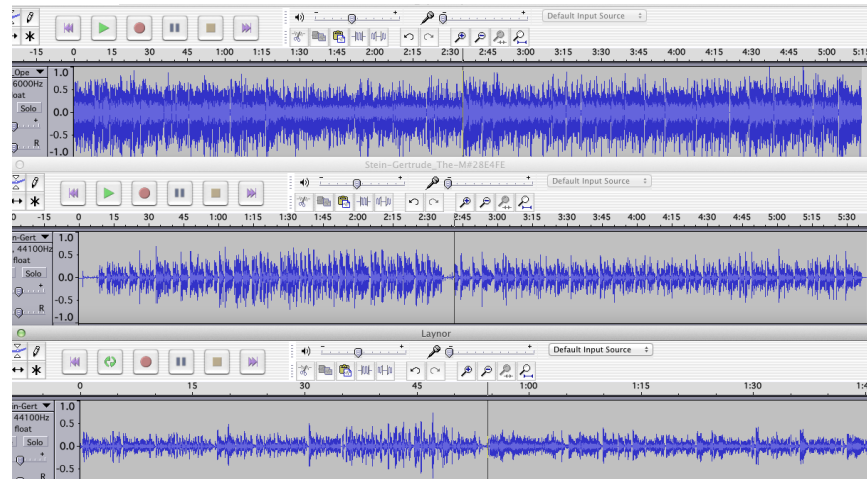
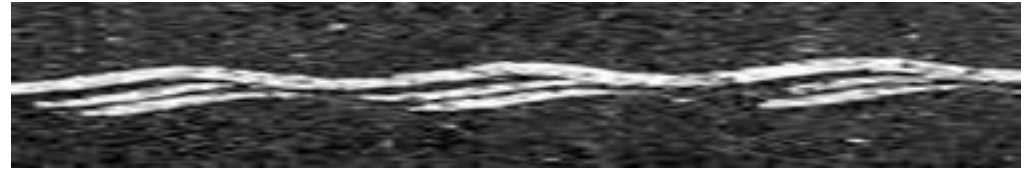
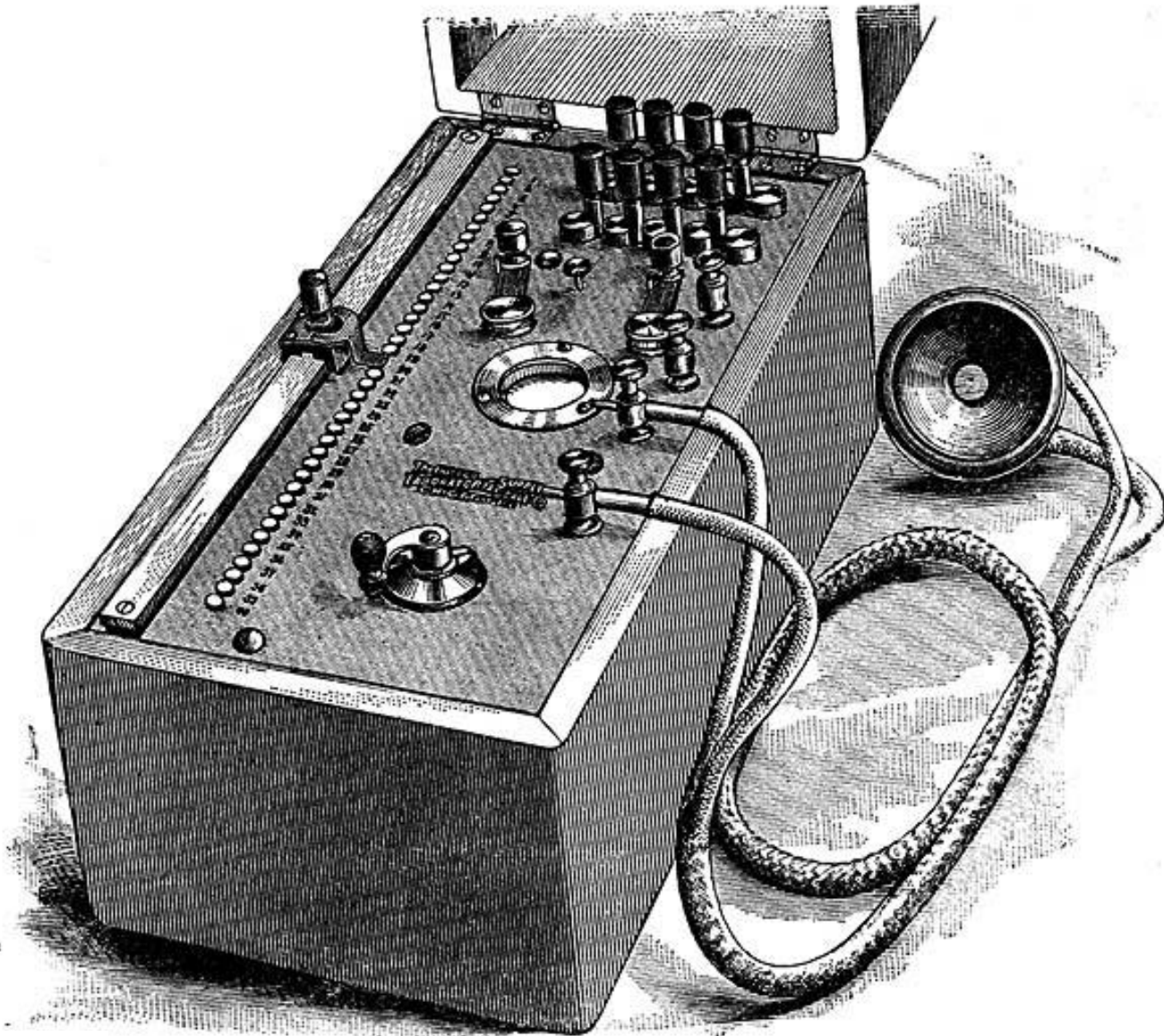
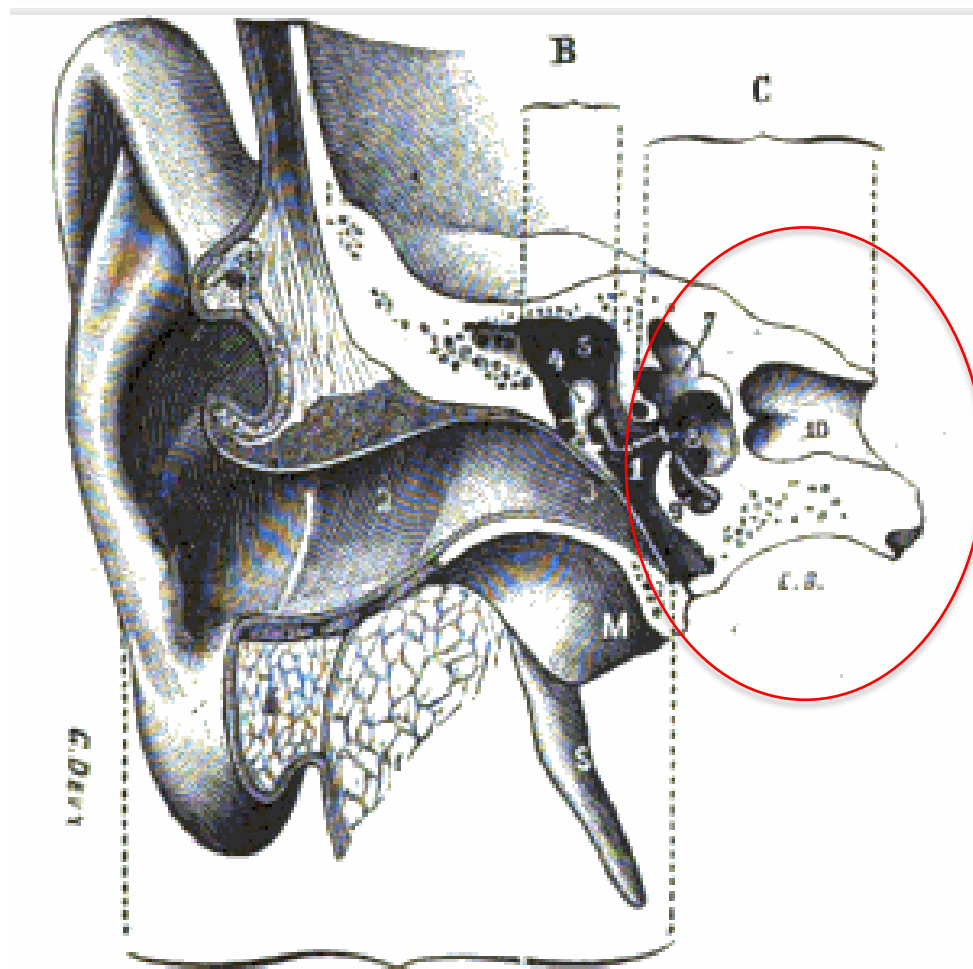


FIG. 33.





Carl Seashore, audiometer, 1899

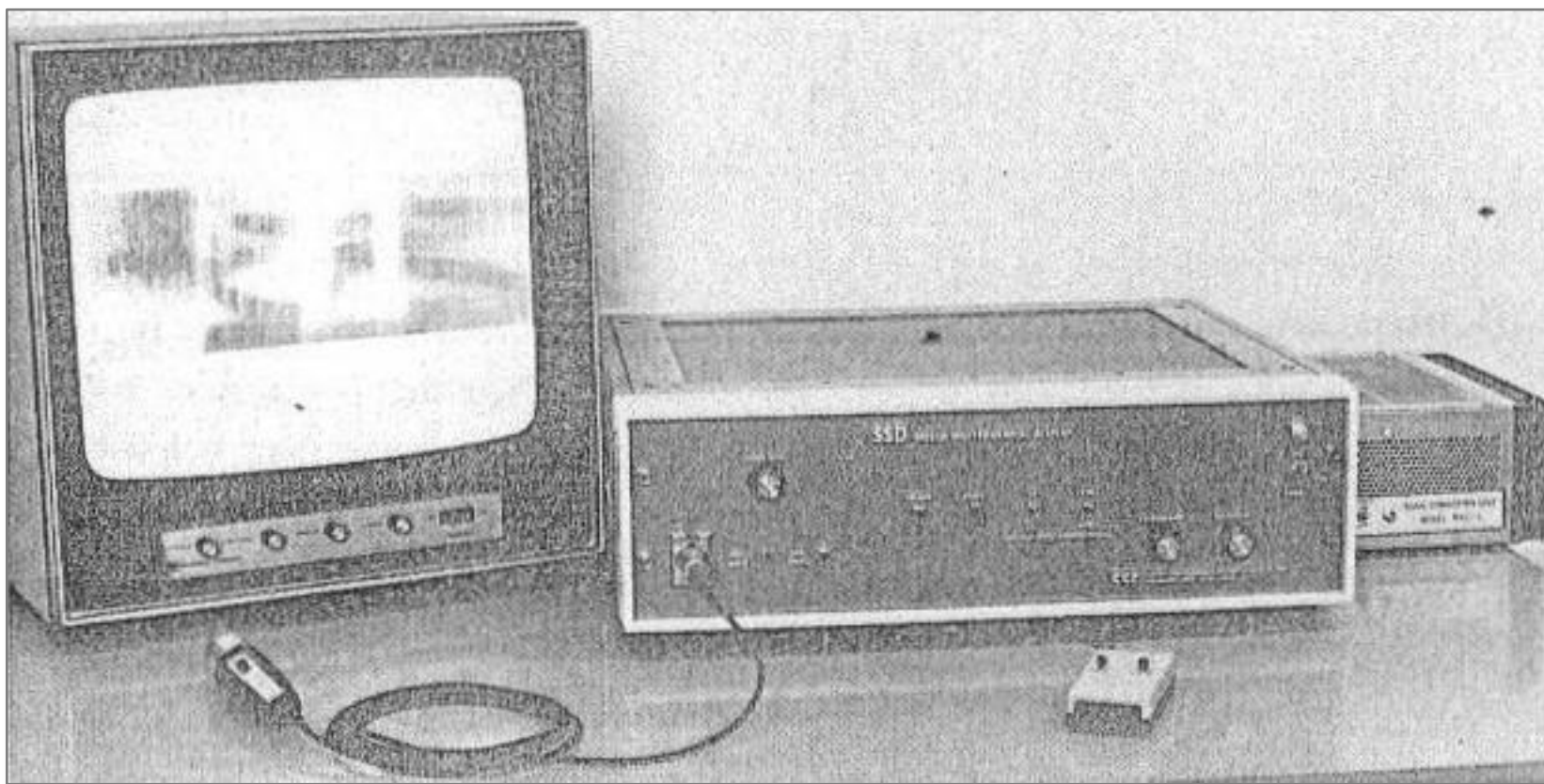


A (TESTUT, *Anatomie.*)

Fig. 7.

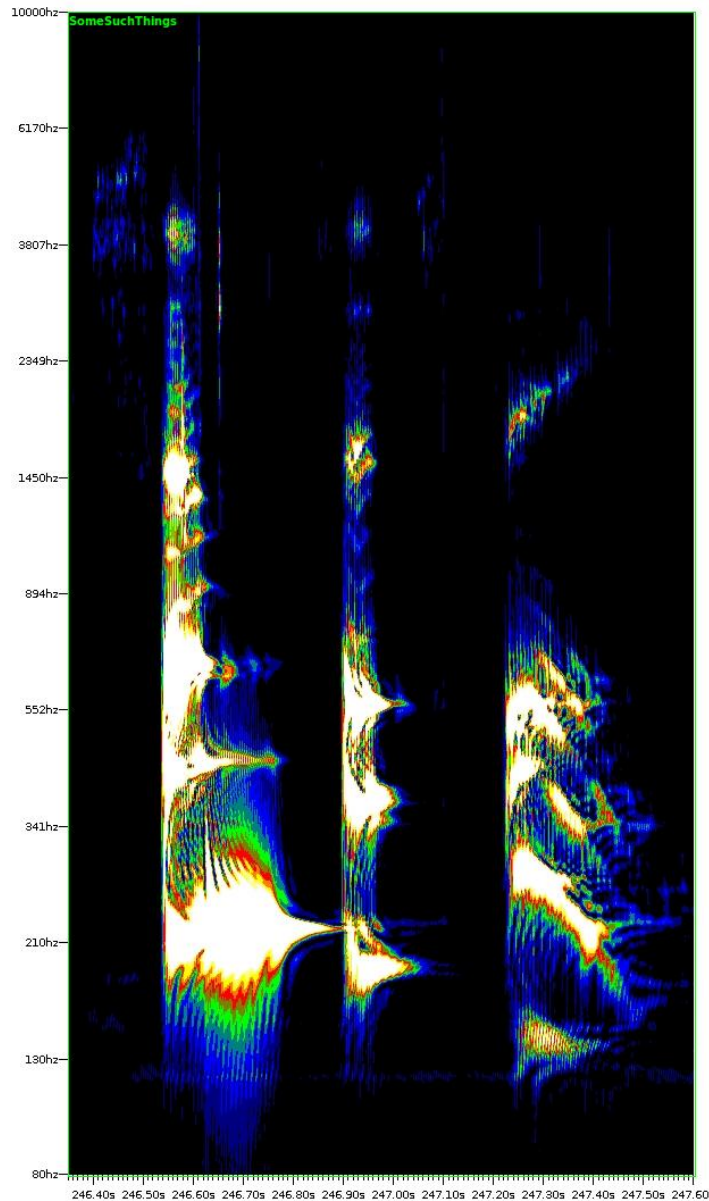
Coupe vertico-transversale de l'appareil auditif.

Oreille externe. — B. Oreille moyenne. — C. Oreille interne. —



In the 1940s and '50s, the spectrographic display at Bell Telephone Laboratories  
<http://www.spectrogramsforspeech.com/>

HZ, a unit  
of  
frequency



Energy represented by  
a heat based color scheme.

White – hottest, most intense

Yellow

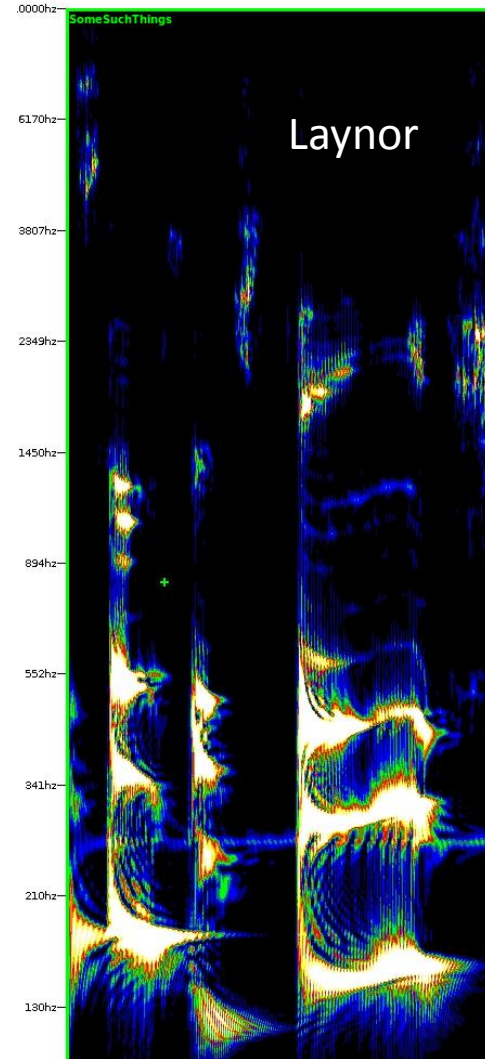
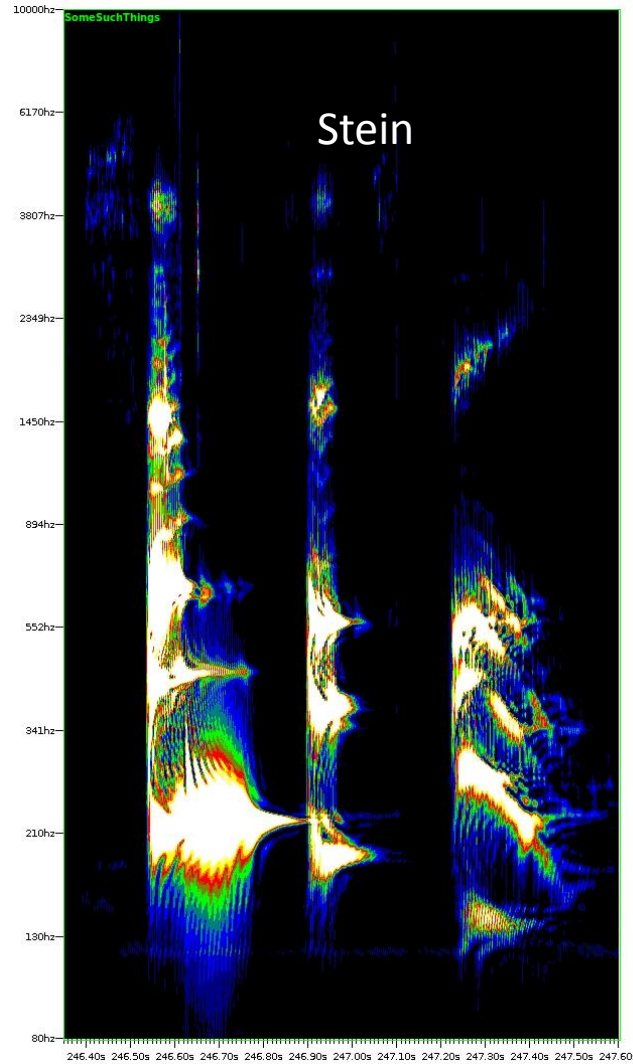
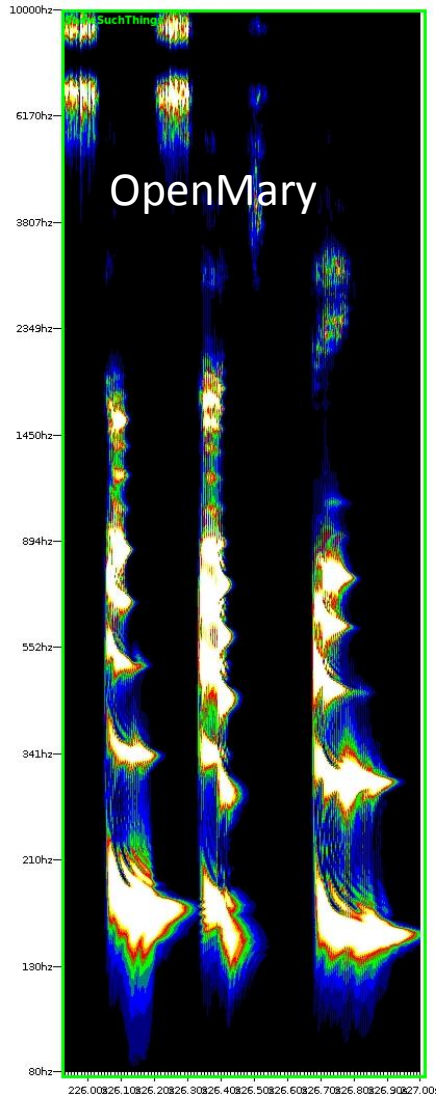
Red

Green

Blue

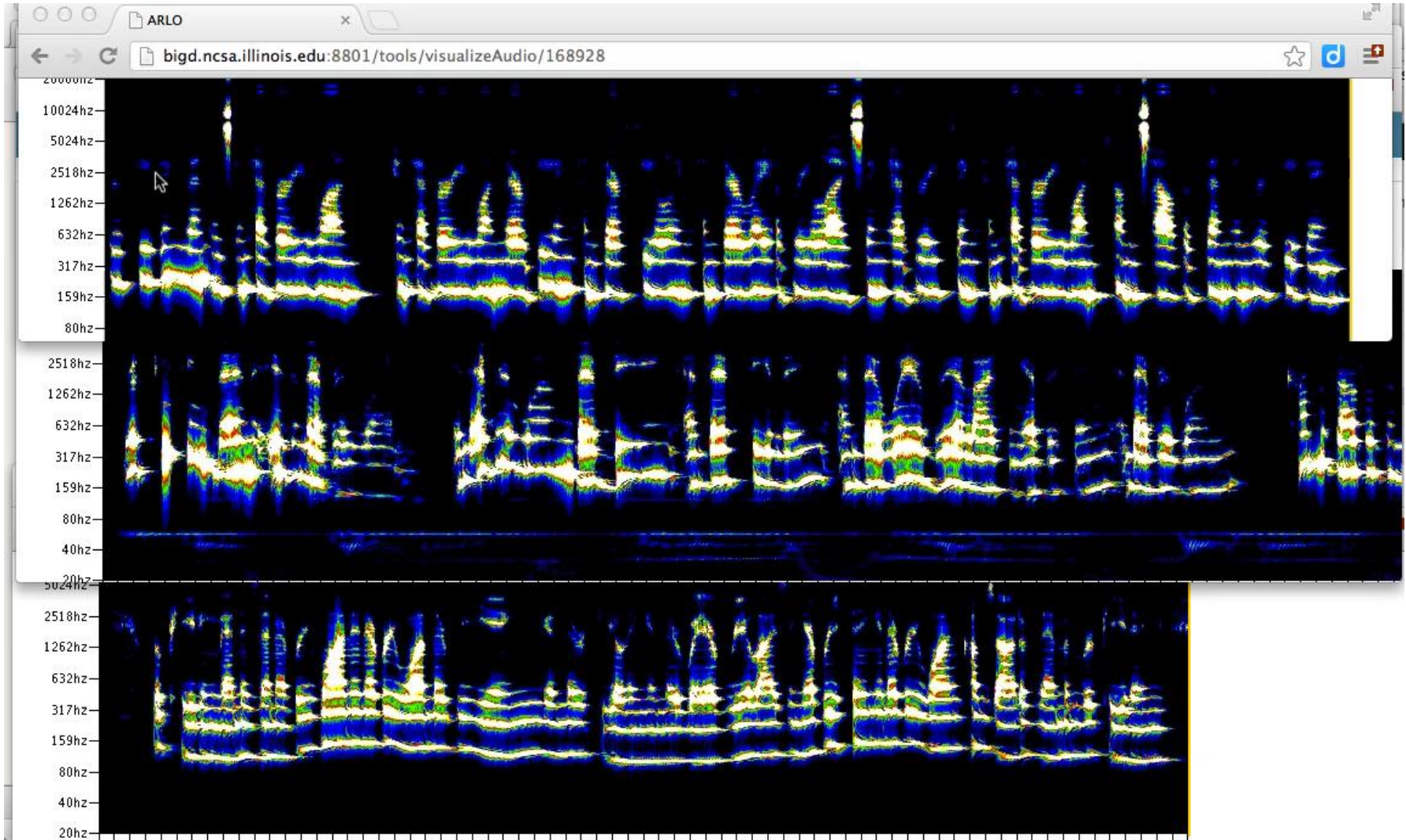
Black – coolest, least intense

# Use case #1: same content, different voices

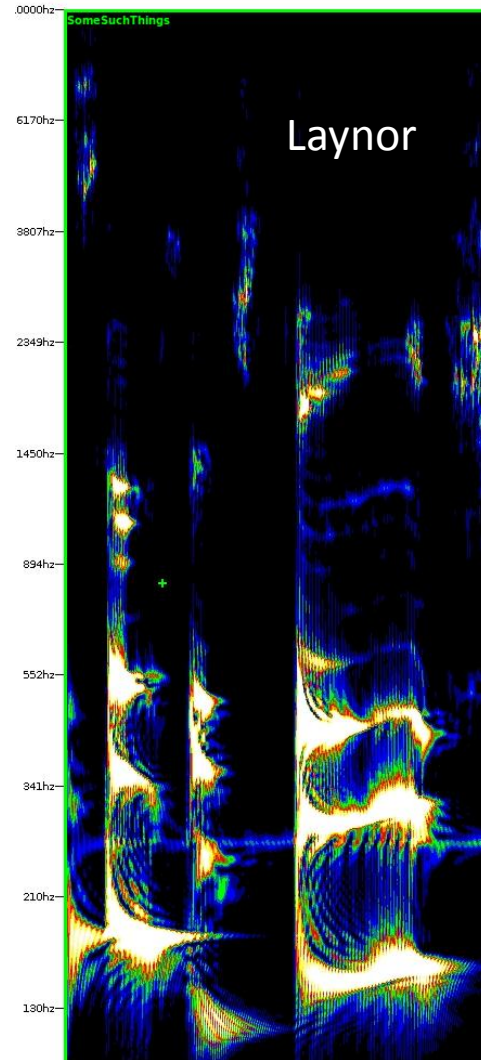
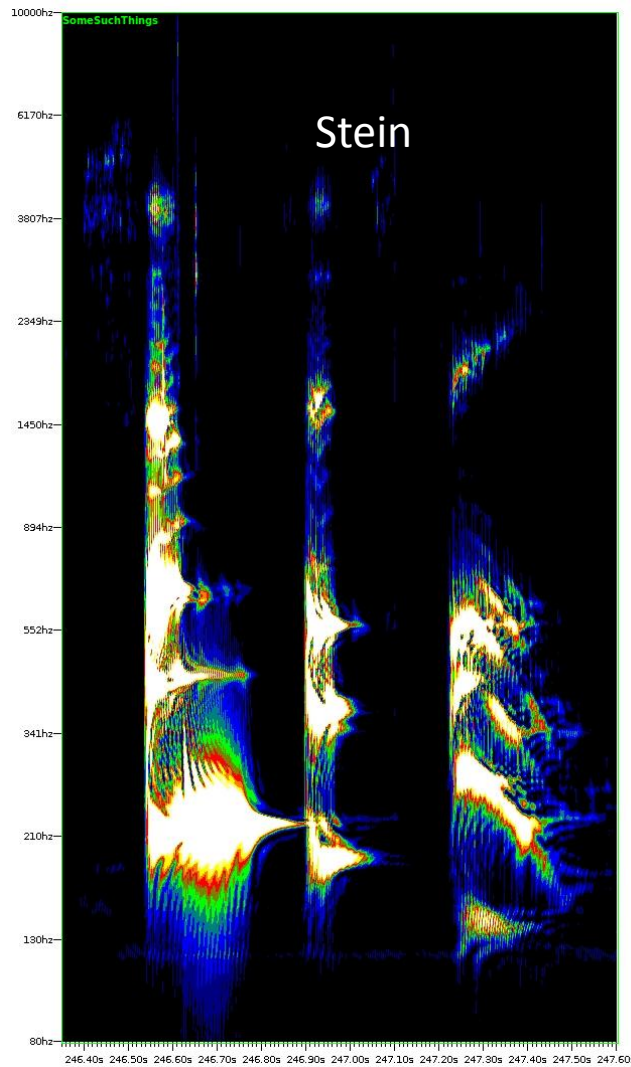
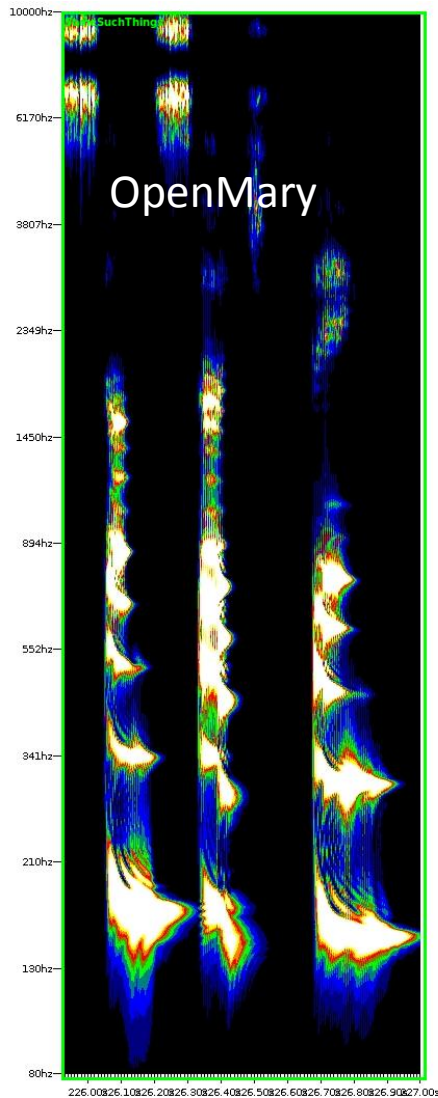


Spectrogram created with ARLO of the line “. . . some such thing.” (by OpenMary, Gertrude Stein, and Gregory Laynor) from Gertrude Stein's *The Making of Americans*

# Use case #1: same content, different voices



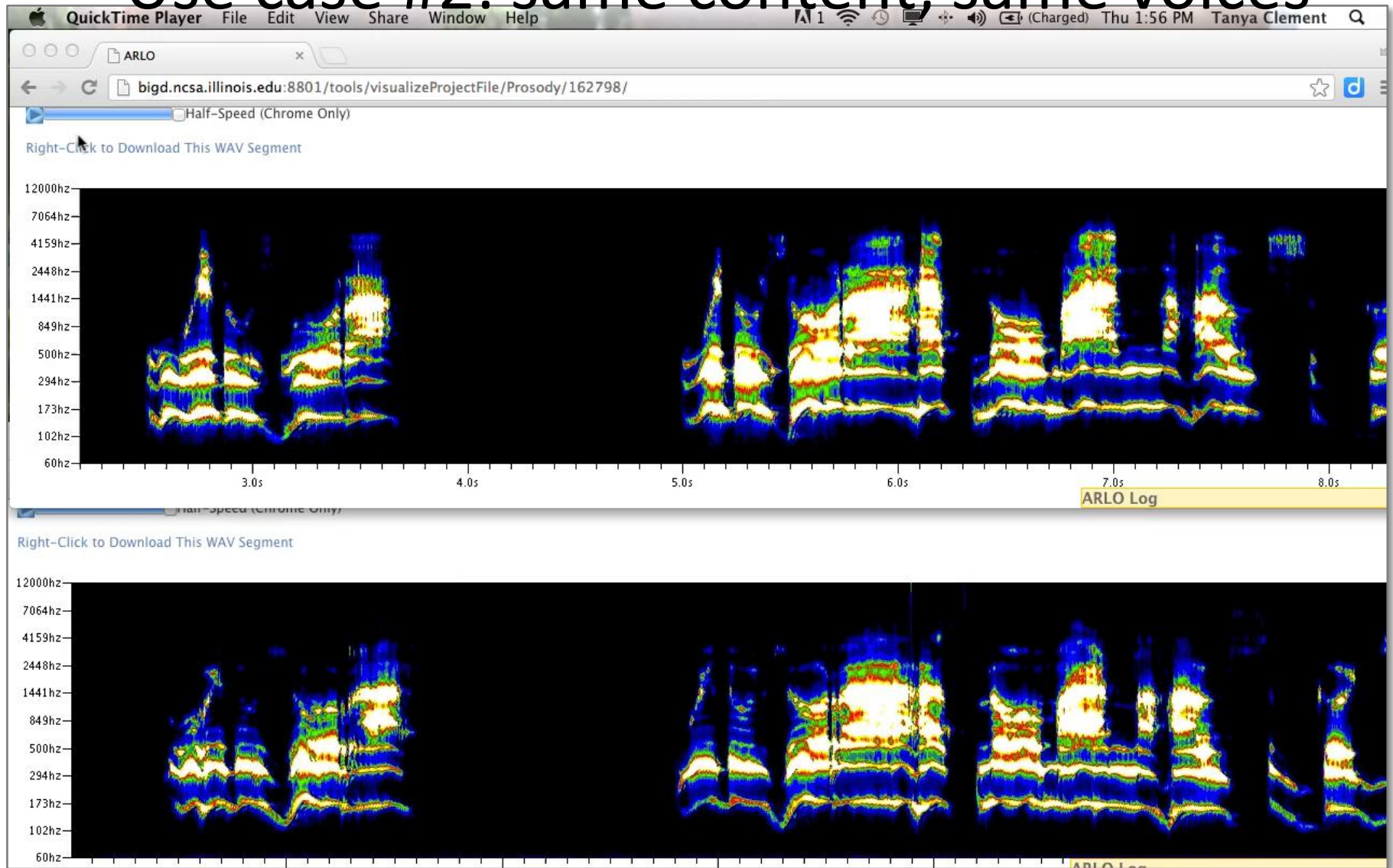
Spectrogram created with ARLO of the line “. . . some such thing.” (by OpenMary, Gertrude Stein, and Gregory Laynor) from Gertrude Stein's *The Making of Americans* generated by ARLO (Adaptive Recognition with Layered Optimization)



Spectrogram created with ARLO of the line “. . . some such thing.” (by OpenMary, Gertrude Stein, and Gregory Laynor) from Gertrude Stein's *The Making of Americans*

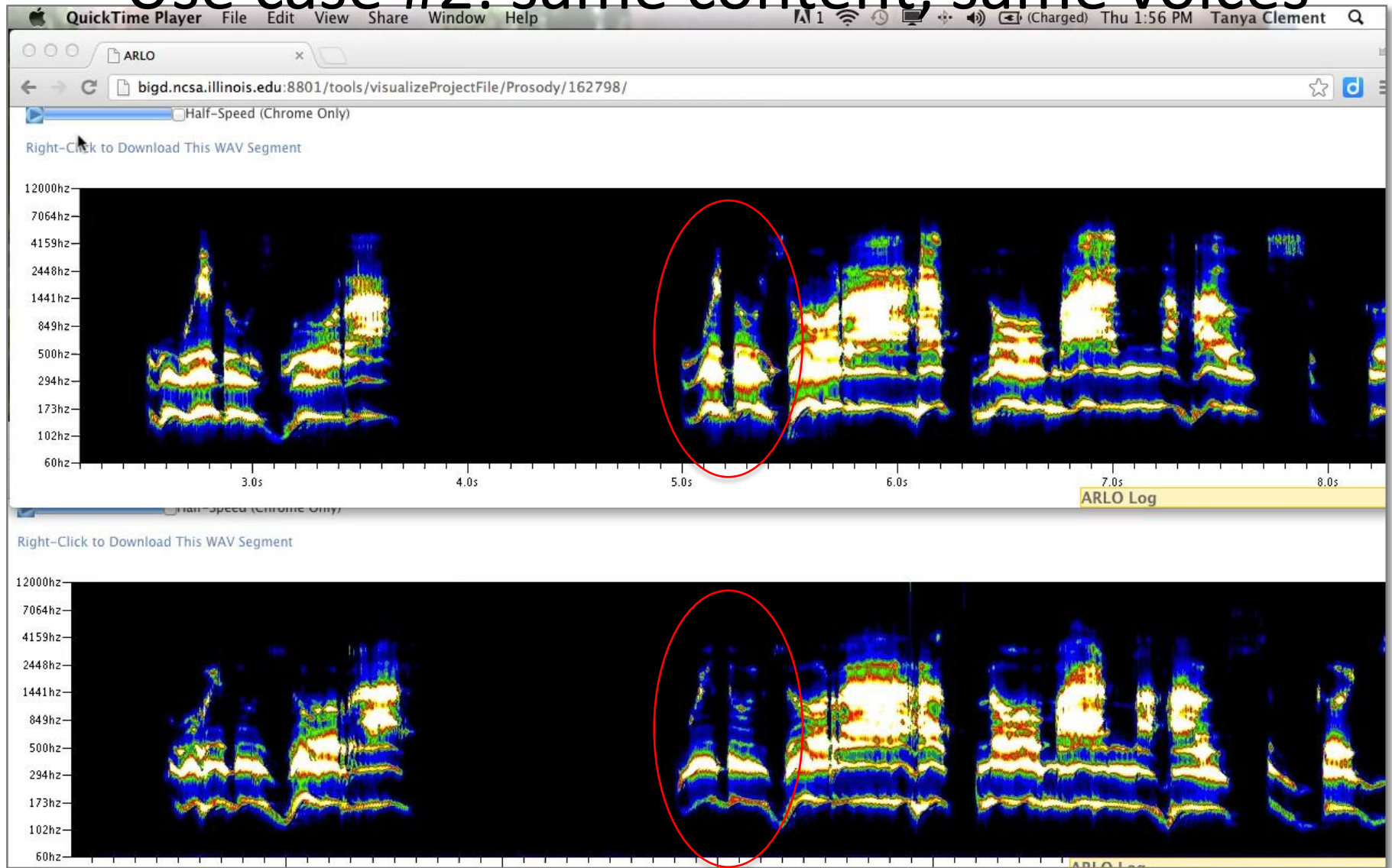


# Use case #2: same content, same voices



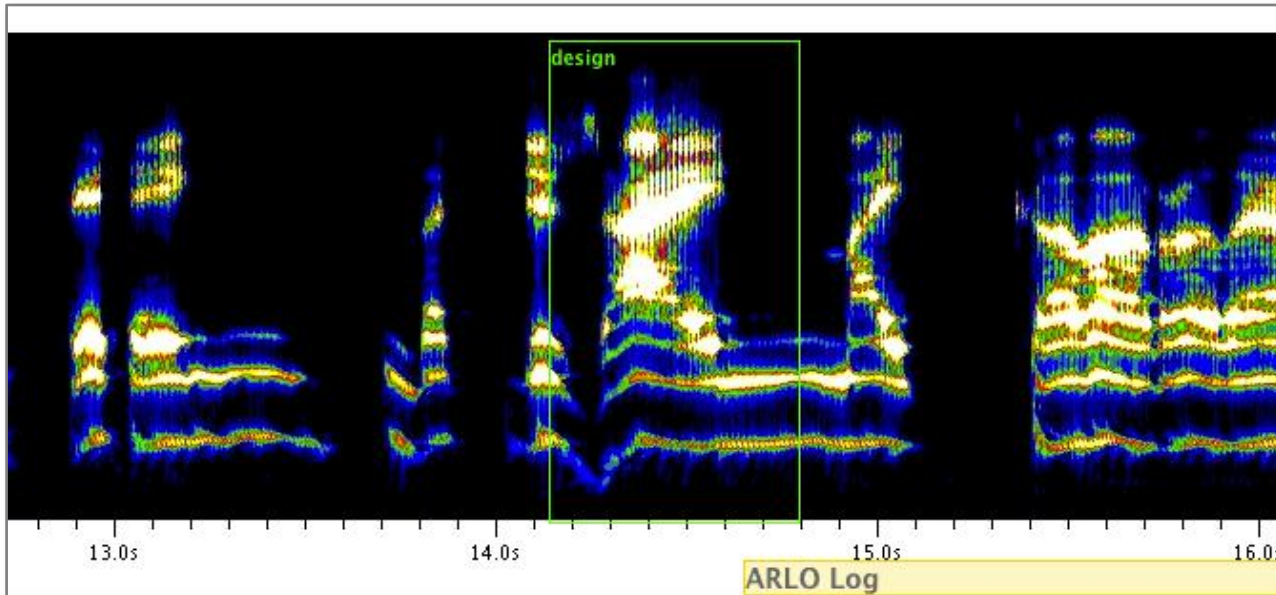
Pound's Canto XLV recorded in Harvard, 1939 and DC, 1958

# Use case #2: same content, same voices

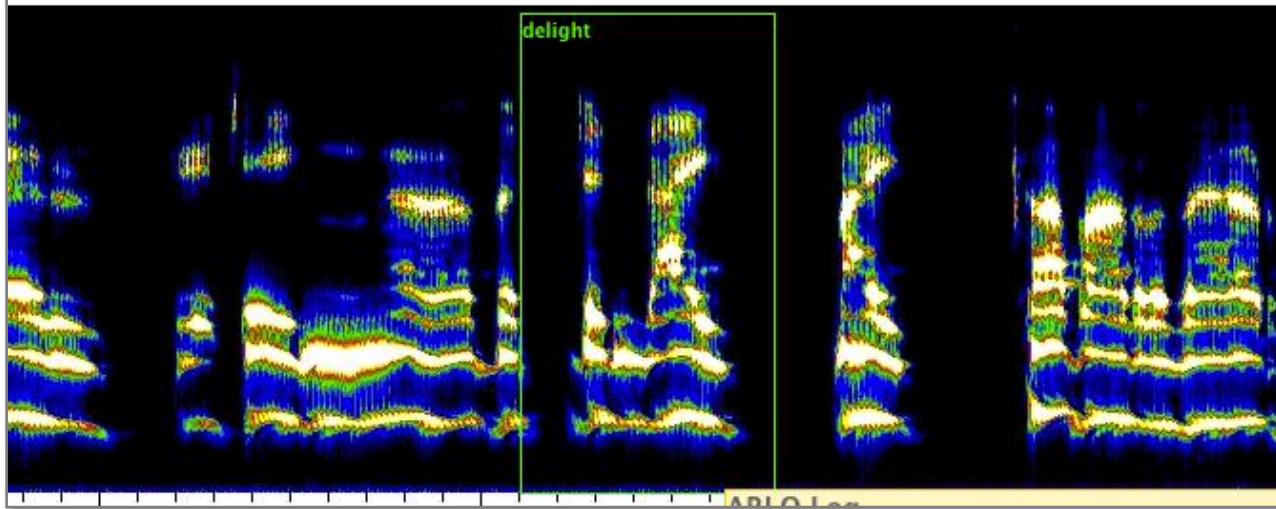


Pound's Canto XLV recorded in DC, 1958 and Harvard, 1939

# Use case #2: same content? Same voices



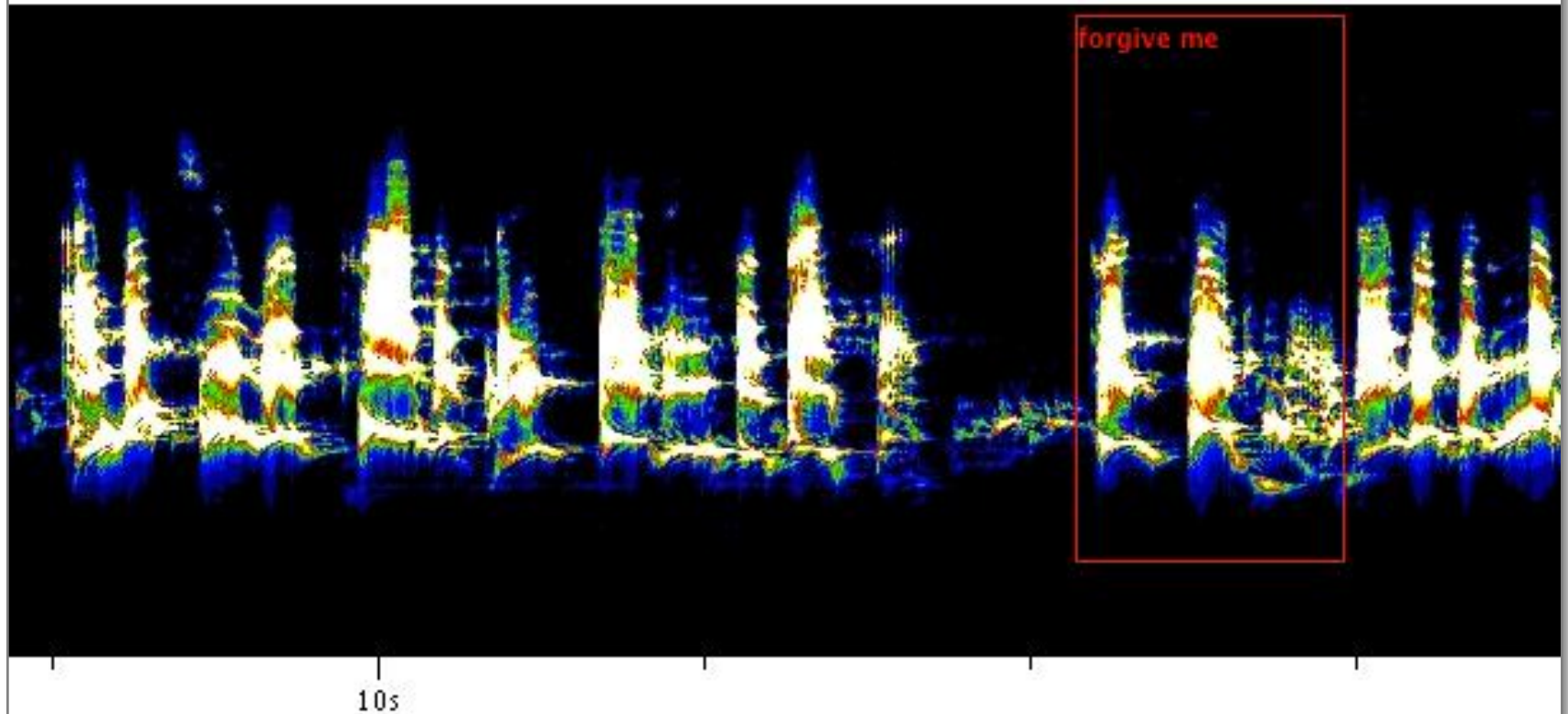
DC, 1958



Harvard, 1939

“design” vs. “delight” in Pound’s Canto XLV recorded in DC, 1958 and Harvard, 1939

# Use case #3: finding similar patterns



“Forgive me” in William Carlos Williams, “This is Just to Say”, Dec., 1951, Harvard

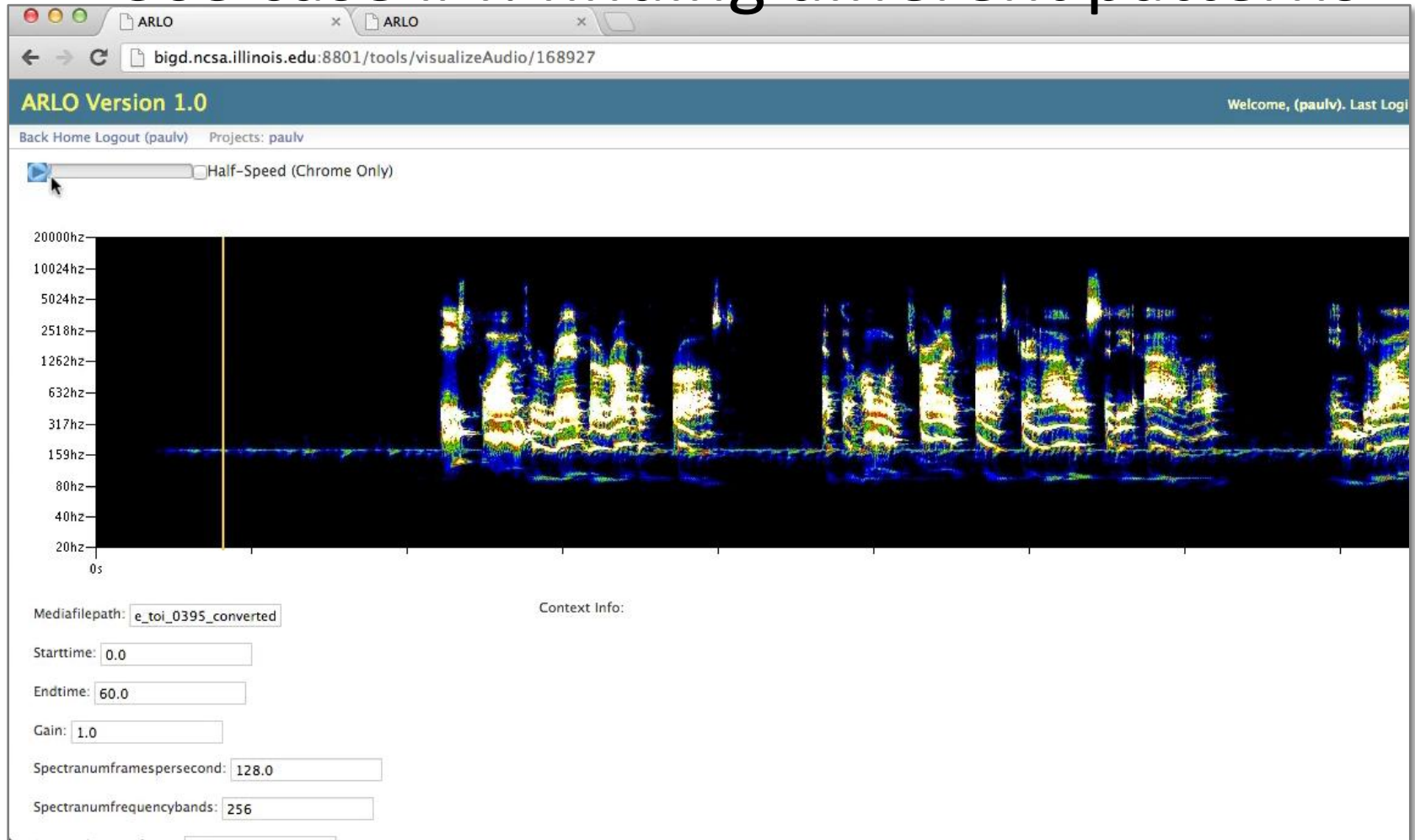
# Use case #3: finding similar patterns

The screenshot shows the ARLO web interface. At the top, there are two browser tabs labeled 'ARLO'. The address bar shows the URL 'bigd.ncsa.illinois.edu:8801/tools/projectFiles/paulv'. The page has a navigation menu on the left with sections for 'Tag Set' and 'Options'. The 'Tag Set' section includes links like 'Catalog', 'Create Transcript', and 'Export Tag Set'. The 'Options' section includes 'Tag Analysis', 'Discover', 'Performance Testing', 'Supervised Tag Discovery', 'Unsupervised Tag Discovery', 'Manage Jobs', and 'Multi-View'. Below the navigation is the 'Audio Files' section, which has a toggle for 'Enable All Audio Files' and a table of audio files. The table has columns for 'Active', 'Alias', and 'Duration (s)'. The 'Active' column contains 'Active' and a link to 'Visualize/Tag | Edit Meta Data'. The 'Alias' column contains the filename. The 'Duration (s)' column contains the duration in seconds. The 'Start Transcript' link is visible at the end of each row. At the bottom right of the interface, there is a yellow bar labeled 'ARLO Log'.

| Active | Alias  | Duration (s)       |
|--------|--|--------------------|
| Active | e_toi_0395_converted.wav   | 502.4426666666664  |
| Active | Ginsberg-Allen_01_Howl_Big-Table-Chicago-Reading_1959                                | 0.0                |
| Active | Owens_Frost  | 502.4426666666664  |
| Active | LBJ_Make_The_Pockets   | 2.986666666666667  |
| Active | Williams-WC-This_is_Just_to_Say-November_16_1950-Van_Nuys_CA.wav                     | 22.987755102040815 |
| Active | Williams-WC-This_is_Just_to_Say-December_4_1951-Harvard.wav                          | 74.71020408163265  |
| Active | Williams-WC-This_is_Just_to_Say-August_1950-Rutherford_NJ.wav                        | 17.110204081632652 |
| Active | Ginsberg-Allen_04_Howl_SFSU_10-25-56.wav   | 1107.4873469387755 |
| Active | Ginsberg-Allen_04_Ah-Sun-Flower_New-York_12-15-69.wav                                | 74.26612244897959  |
| Active | Ginsberg-Allen_03_The-Sick-Rose_New-York_12-15-69.wav                                | 100.02285714285715 |
| Active | Ginsberg-Allen_01_Howl-I-III_NY_5-4-95.wav   | 1357.7665306122449 |
| Active | Brown-Lee-Ann_02_Blake-Ah-Sunflower_Segue_NY_5-26-01.wav                             | 106.37061224489796 |
| Active | Bernstein-Charles_The-Sick-Rose_Poems-for-the-New-Millennium_KWH-UPenn_10-7-2009.wav | 55.30122448979592  |

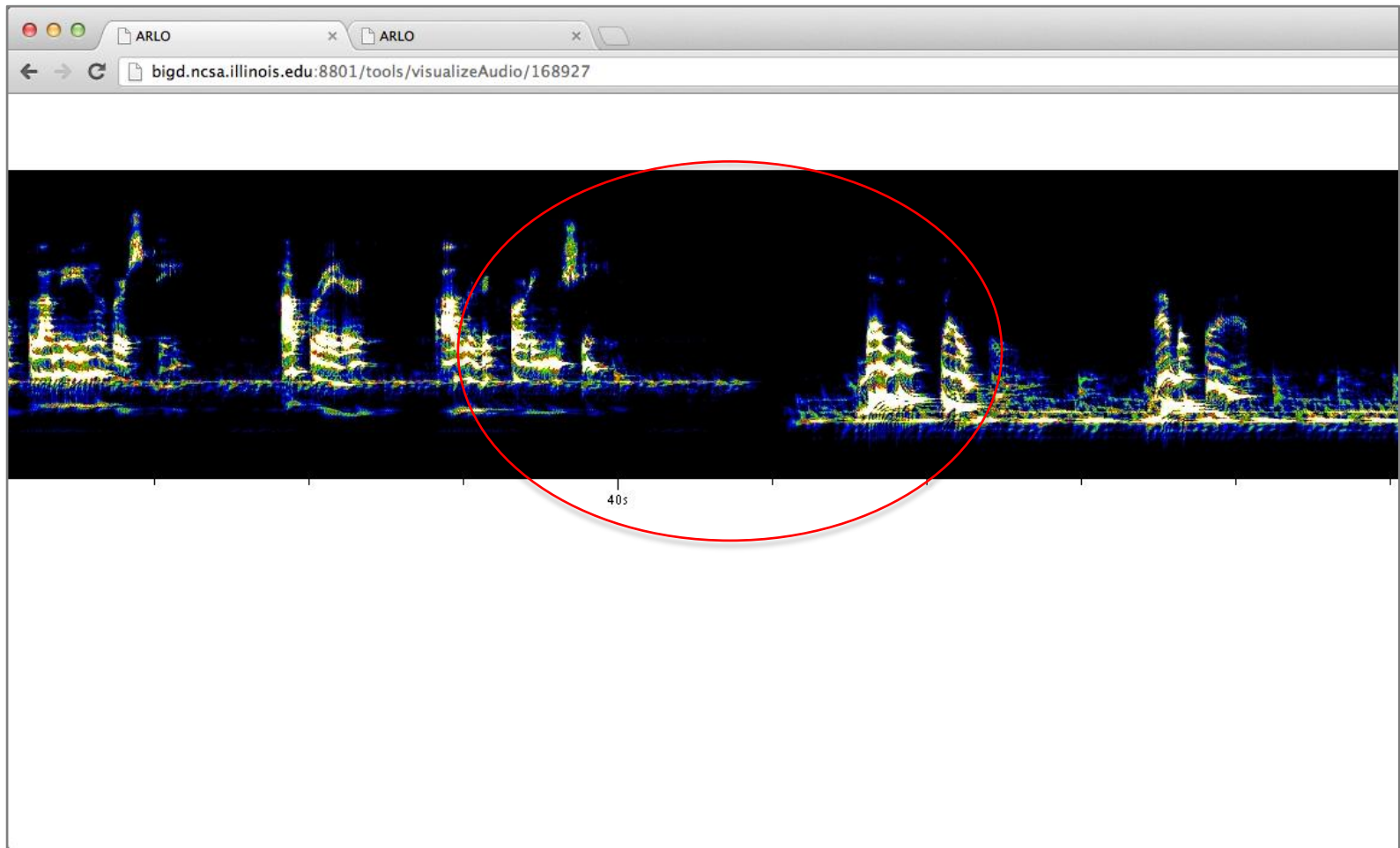
Finding “Forgive Me” in ARLO

# Use case #4: finding different patterns



Frost reading poems on Side B. of William A. Owens' folklore recordings

# Use case #4: finding different patterns



Frost reading poems on Side B. of William A. Owens' folklore recordings

# Use case #4: finding different patterns

Ojibwe:

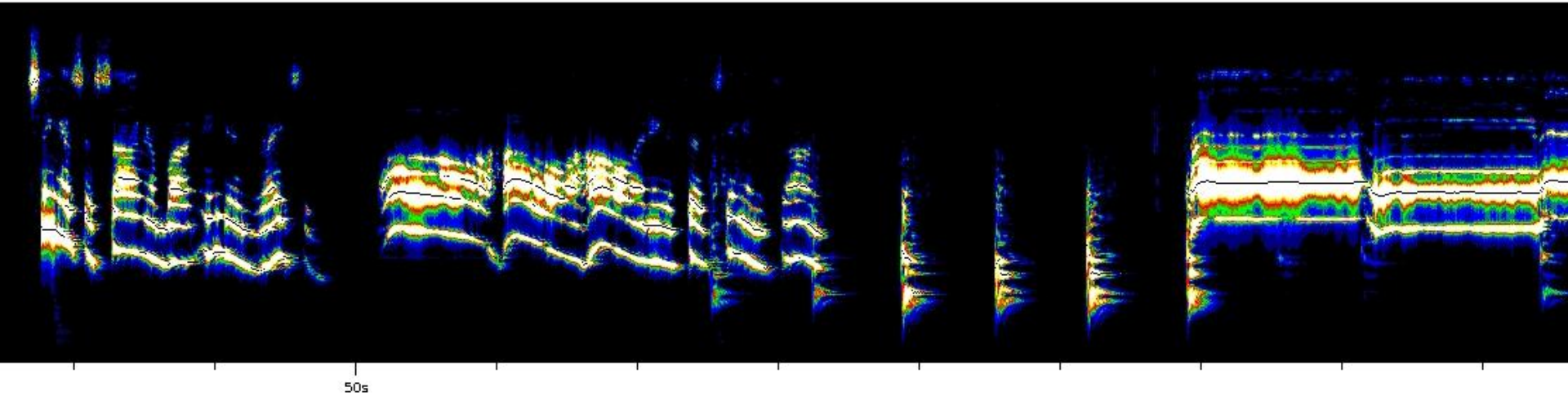
The Language of Teaching,  
The Language of Prayer



Larry Aitken, Tribal Historian, Leech Lake Band of Ojibwe  
Courtesy of Penn Museum



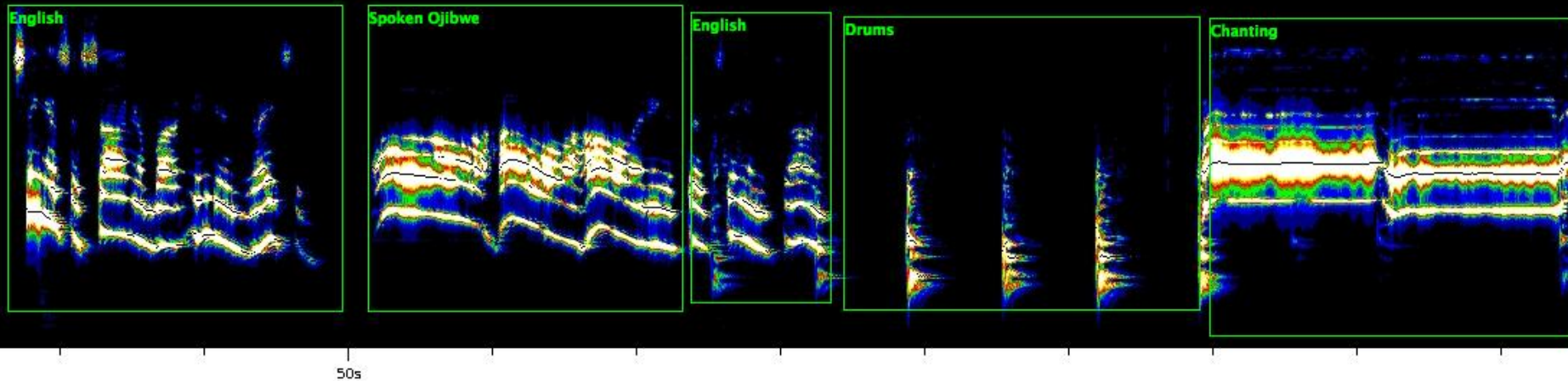
# Use case #4: finding different patterns



50s

**Excerpt from a 2007 interview with Larry Aitken, the tribal historian from the Leech Lake Band of Ojibwe conducted by Dr. Tim Powell when he was director of the Center for Native American studies at the University of Pennsylvania. The second image has been tagged by a user to show the different ways that English, spoken Ojibwe, drum beats, and chanting are visualized.**

# Use case #4: finding different patterns



**Excerpt from a 2007 interview with Larry Aitken, the tribal historian from the Leech Lake Band of Ojibwe conducted by Dr. Tim Powell when he was director of the Center for Native American studies at the University of Pennsylvania. The second image has been tagged by a user to show the different ways that English, spoken Ojibwe, drum beats, and chanting are visualized.**

# What are we talking about when we talk about sound?

- Tempo, pitch, tone/timbre, dynamics
- Damping ratios, gain, frequencies, spectra, energy, and pitch energy
- What do humanists and archivists tag?  
Laughter, silence, emotions, applause, pauses, feedback, changing speakers, changing tracks, changing genres

*Thank you!*

Questions?

[tclement@ischool.utexas.edu](mailto:tclement@ischool.utexas.edu)

*Special thanks to Tim Powell, my co-PI's David Tcheng and Loretta Auvil, and the National Endowment for the Humanities.*