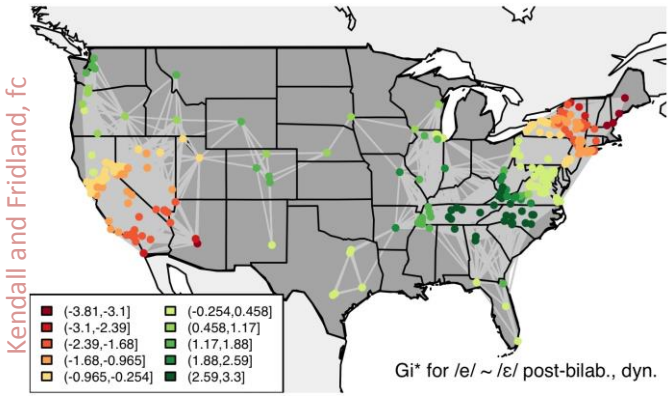
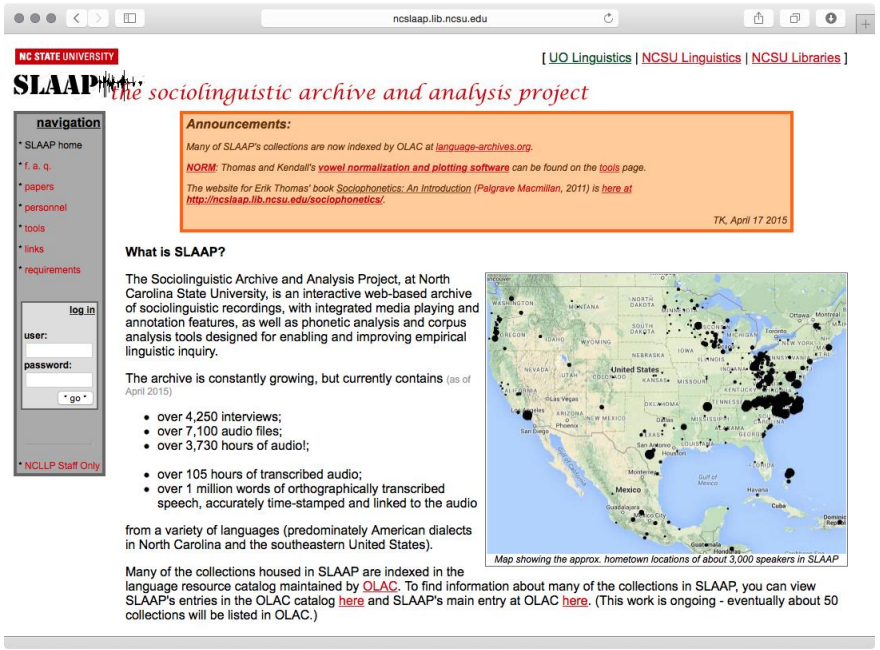
Me & the UO Language Variation & Computation Lab

- Sociophonetician and sociolinguist researching variation and change in regional and ethnic varieties of U.S. English
  - My dissertation (2009; and 2013 book) on "corpus sociophonetics" of speech rate and pause variation in U.S. English
  - Currently, developing a public corpus of spoken African American English
    - Funded by NSF (SBE-BCS-Linguistics)
  - Currently, with Valerie Fridland (UNR), pan-regional study of production and perception of vowels and vowel shifts
    - Funded by NSF (SBE-BCS-Linguistics)



- In terms of speech technology,
  - Develop and maintain *Speech Data Management Systems*
  - Main e.g. Sociolinguistic Archive and Analysis Project (SLAAP)
    - http://slaap.lib.ncsu.edu
  - Also, NORM/Vowels.R
    - Tools for plotting/transforming acoustic vowel data



SLAAP: Kendall 2007

How does my field impact speech technology?

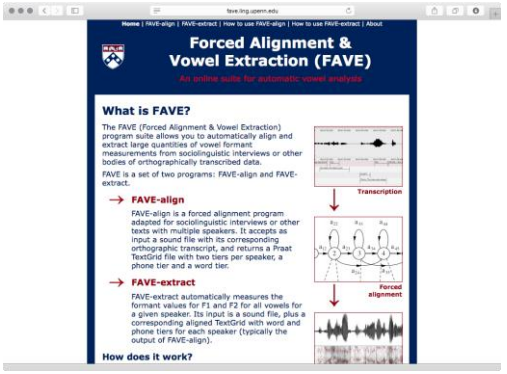
- Primary research questions:
  - How does language variation & change relate to social and cognitive factors?

- Primary questions for speech technology:
  - How can we discover/identify/analyze sound change in progress?
  - How do we differentiate important variation from unimportant variation (noise)?
  - How do we find/assess relevant data?
  - Existing tools and foci indicate that sociolinguists are looking for cheap/automatic time-aligned transcription and ability to acquire "analytic data" quickly/cheaply.

- Largely, sociolinguists are (avid?) users of speech technology but rarely creators
  EXCEPTIONS ➔
  - Most work uses Praat (Boersma & Weenink 2001-2015) for manual/semi-automatic analysis.

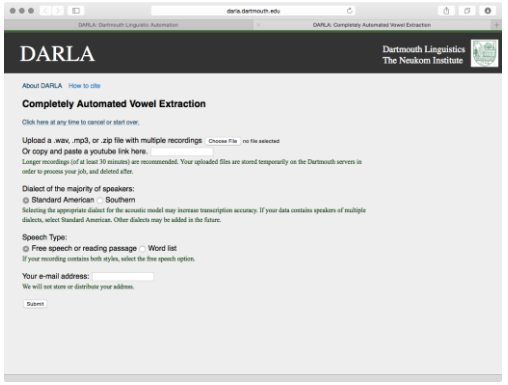
- Existing…
  - State of the art = forced-aligned and probabilistic formant extraction

FAVE: Rosenfelder et al. 2011)



  - Also, Prosodylab aligner (Gordon et al. 2011)
  - Frontier?? = completely automated vowel extraction

DARLA: Reddy & Stanford 2015



What challenges do we face to *impact* ST?

- Much sociolinguistic/variationist data are non-standard ("unconventional corpora" Beal et al. 2007)
- The features of interest are in flux and (can be) dialect dependent
  - E.g. Northern Cities shifted vowels, the low back merger in American English
- Preexisting speech models don't match varieties under examination
- Interested in speaker characteristics and not just speech

- Our solutions are somewhat overly specific (to question at hand) and may not apply to new datasets or new questions
  - E.g. FAVE is state of the art, but still has limitations
    - It uses a sample of American English (from ANAE) as its reference…

- Again, sociolinguists are generally (relatively naïve) *users* of speech technology

What challenges do we face to impact or *use* ST?

- Lots of diverse data
  - SLAAP contains > 4,000 interviews, > 3,700 hours of speech
  - But individual projects (≈ varieties) can be as small as ~6 interviews

- My bias is on the archive/data management side:
  - No uniform guidelines/standards for data/metadata
    - *NSF & other "data management" guidelines are improving things…*
  - No interoperability between "archives" and low discoverability
    - Most "archives" are researchers' desktop computers

- Conventional tools often have unknown error rates/types for non-standard speech
- Logistical challenges include:
  - Lack of technical expertise within sociolinguistics (some exceptions)
  - To use ST but also just to understand ST possibilities or to articulate questions
  - Low interest by speech technologists in sociolinguistic projects(??) or more likely a large disciplinary divide between sociolinguistics and speech technology

➔ *Can speech technologists educate this and other (potential?) user populations?*

A sociolinguistic/sociophonetic wish-list?

- What would ideal speech technologies look like from a sociolinguistic perspective?

- Again, bias on the archive side: searchable (by metadata and by content/feature) interoperable distributed archives
  - *Improved sociolinguistic archiving could represent a huge boon to speech technology, NLP, etc. in that it massively ramps up the amount and diversity of speech data available for R & D, representing a range of real-world speech types*

- Searchable = acoustic landmark detection for speech features
  - E.g.: "I want to find young Southern males with high rates of consonant cluster reduction" or "What rates of consonant cluster reduction do young Southern males exhibit?"
- Transcription "on the fly"(ish)
  - Requires flexible ASR/language models robust to disfluent, conversational speech
  - *Also could provide relatively cheap assessments of ST success rates*
    - *E.g. Researchers could approve/disapprove or hand-correct transcripts to improve speech technology systems as a part of their own research*