

Transcribing the Digital Archive of Southern Speech: Methods and Preliminary Analysis

Rachel Miller Olsen, Michael L. Olsen,
Joseph A. Stanley & Margaret E.L. Renwick

The University of Georgia

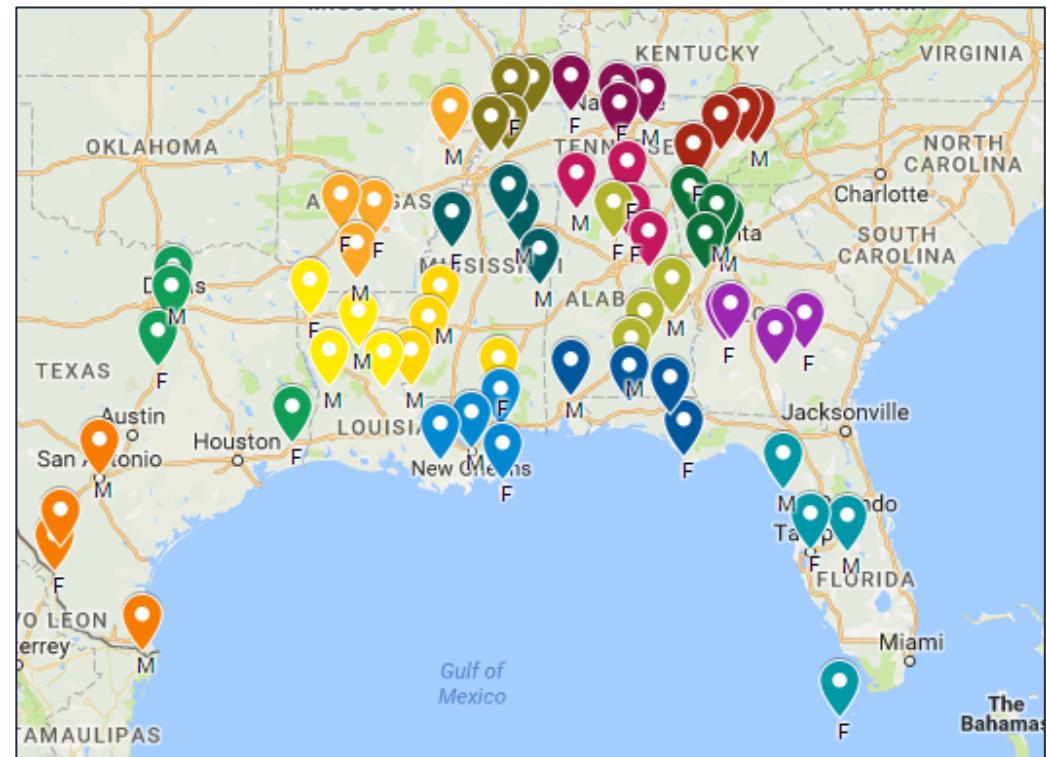
SECOL 84

Introduction

- ▶ Large-scale transcribed audio corpora are available
 - ▶ Buckeye Corpus, Santa Barbara Corpus, etc.
- ▶ How do these come to be? What's the on-the-ground process of building such a corpus?
- ▶ Here we discuss:
 - ▶ Methods for large-scale transcription
 - ▶ Early data & analysis resulting from transcription

Digital Archive of Southern Speech (DASS)

- ▶ 64 interviews
- ▶ 2.5-10hrs, $\mu=5.75$
- ▶ 372 hours of audio



LAGS Protocols

MUX 47 3B		7A	UG ATLANTA T 053.07
1			< 1A.4 - 6 >
2	bəx̥-θ	k'ɪ-tʃən	House where he grew up. [oʊ-ɛf / vɪz-k-tə-ʒən] həʊ-s]
3	bəd rɪ-θm	də-ɪnɪŋ ↗ rɪ-thum ↘	Long central hall. Porch had large [k'a-ɪ-ləmz] Two-story house.

- ▶ Pilot Study:
 - ▶ 1031 words/spkr x 10 = 10,310 words →
 - ▶ Searchable time-aligned corpus of 132,000 words

Transcribing DASS

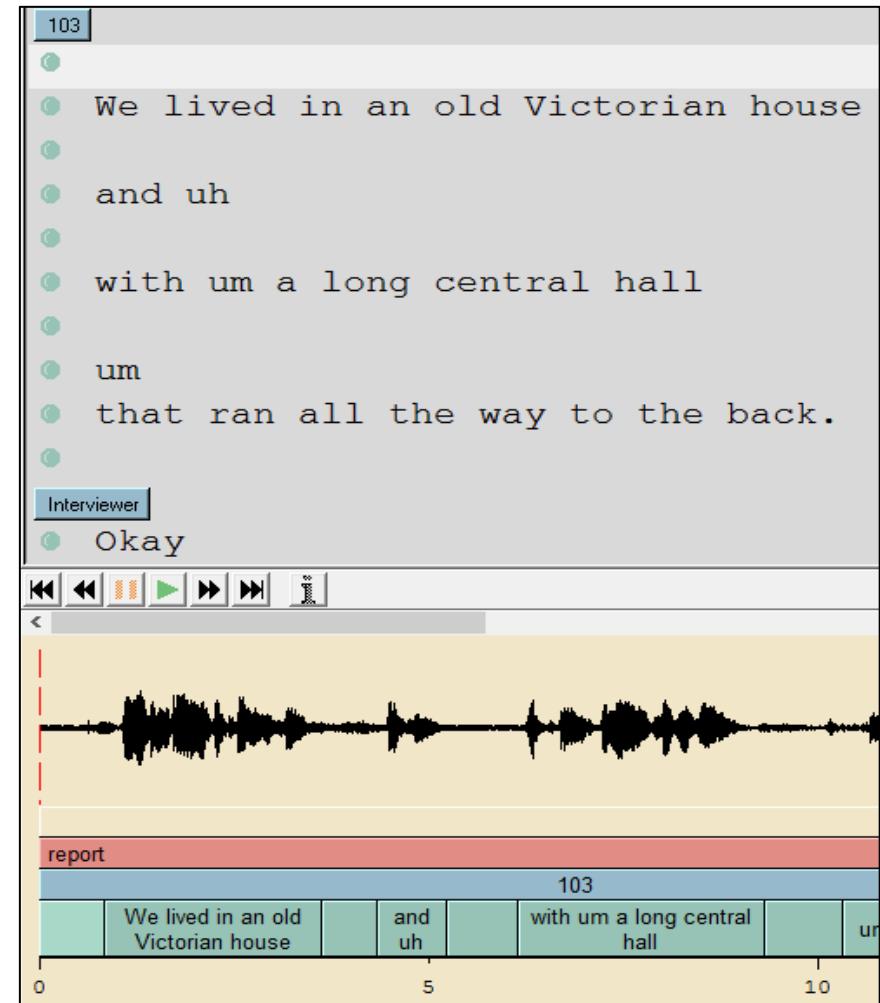
- ▶ 35 undergraduate student workers
- ▶ Each student worker is assigned one interview
- ▶ One reel at a time
- ▶ 408 reels/files, $\mu=54\text{mins}$

Transcriber

(Boudahmane et al. 1998–2008)



- ▶ Create & edit time-aligned orthographic transcriptions
- ▶ Easy-to-use graphical user interface
- ▶ .trs (native .xml)
- ▶ trans.sourceforge.net



Guidelines

- ▶ Transcriber protocols (~25 pages)
- ▶ Phrase Dictionary
- ▶ Two-phase listening
- ▶ Daily files + Multiple backups

Codes	Meaning
{D: }	Doubt
{X}	Unintelligible
{C: }	Comment
{NW}	Non-word (e.g. laugh, cough)
{NS}	Non-speech (e.g. dog barking)

Workflow

Transcription
(i.e. 2 listens)
complete

Spot-checked
for consistency

File conversion via
LaBB-CAT scripts
(Fromont & Hay 2012)

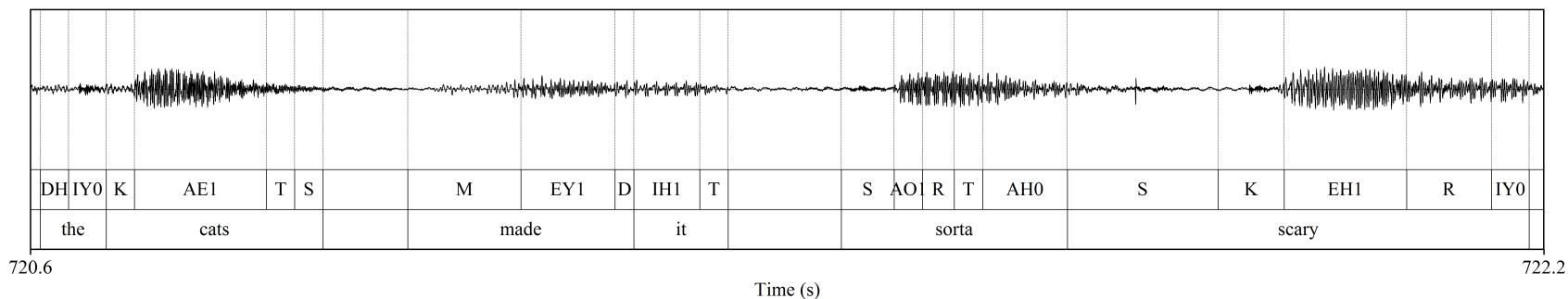
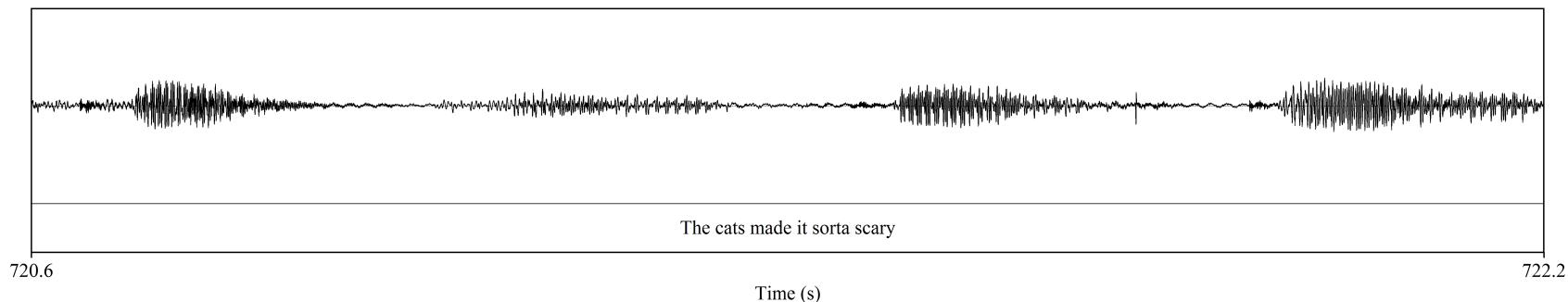
labbcat.sourceforge.net

.trs (.xml) → .txt
.trs → .TextGrid

Automatic
phonetic
analysis!

Forced Alignment

- ▶ Forced-aligned with DARLA (Reddy & Stanford 2015)

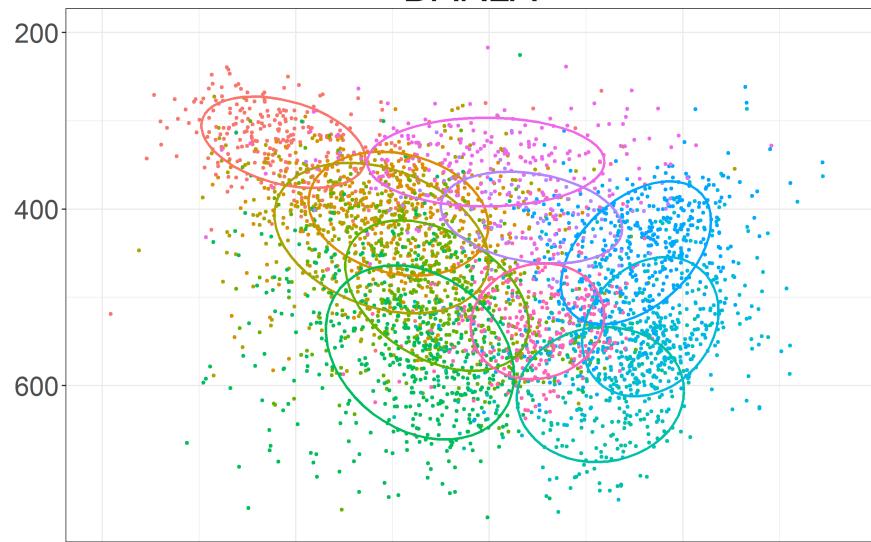


Phonetic Analysis

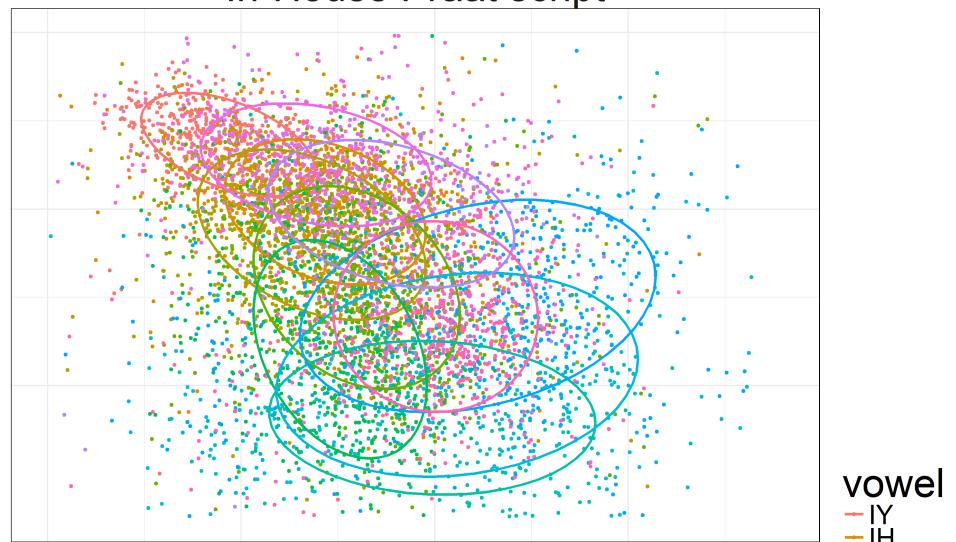
- ▶ Formant extraction: four different methods
 - ▶ In-house Praat script (Boersma & Weenink 2016)
 - ▶ DARLA (Reddy & Stanford 2015)
 - ▶ out-of-the-box FAVE (Rosenfelder et al. 2011)
 - ▶ based on ANAE means
 - ▶ modified FAVE (Rosenfelder et al. 2011)
 - ▶ based on Southern means

Comparision of formant extractors

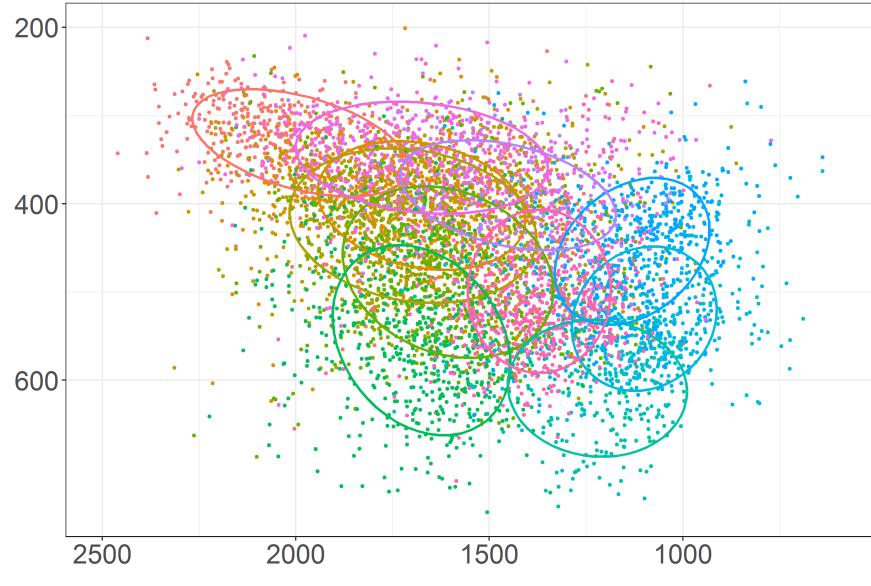
DARLA



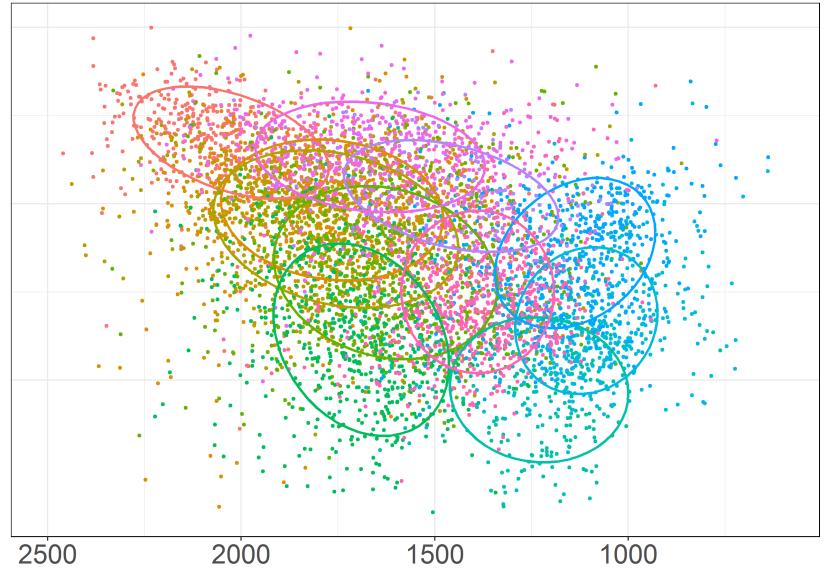
In-House Praat script



Out-of-the-box FAVE



Modified FAVE

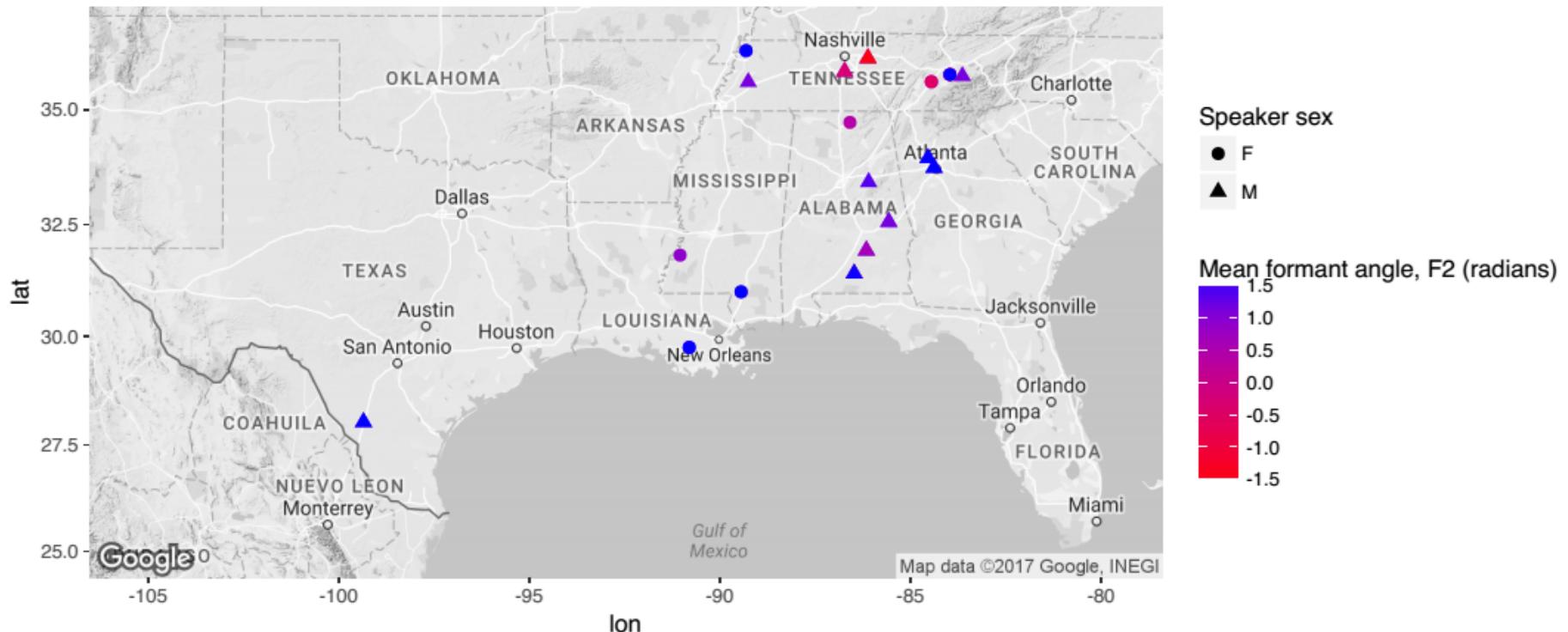


vowel

- IY
- IH
- EY
- EH
- AE
- AA
- AO
- OW
- UH
- UW
- AH

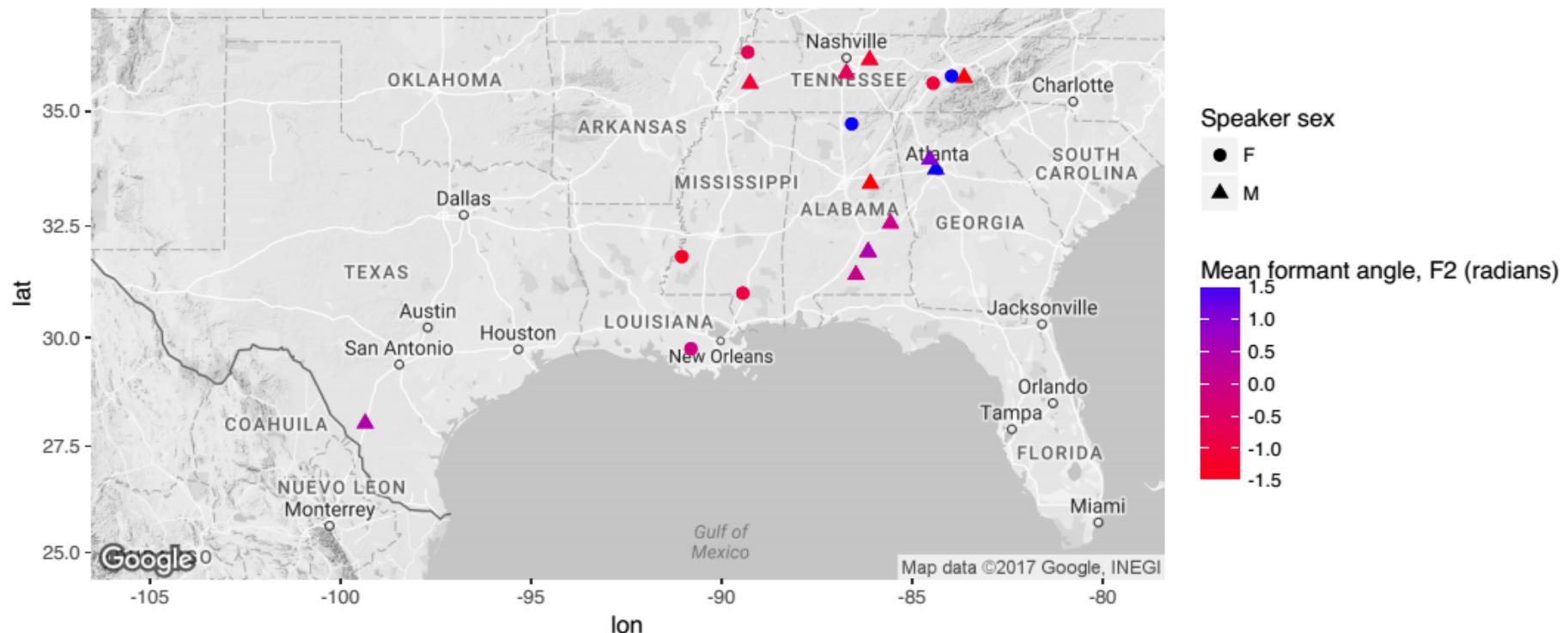
Preliminary Findings: Glide weakening

Mean formant angle, F2, before voiceless consonants: higher value = more diphthongal



Glide weakening (cont.)

Mean formant angle, F2, before voiced consonants: higher value = more diphthongal



Observations

- ▶ Large-scale transcription
 - ▶ Time to transcribe
 - ▶ Estimated: 10:1; Reality:13:1
- ▶ Phonetic Analysis
 - ▶ Comparison of formant measurements
 - ▶ In-house Praat script no good
 - ▶ DARLA filtered out 53%
 - ▶ Too early to tell if FAVE modifications were better

References

- Boersma, Paul & David Weenink. 2016. *Praat: Doing phonetics by computer [Computer program]*, Version 5.4.08. Retrieved from <http://www.praat.org>.
- Boudahmane, Karim, Mathieu Manta, Fabien Antoine, Sylvian Galliano & Claude Barras. 1998. *Transcriber* v. 1.5.2. <http://trans.sourceforge.net/>.
- Fromont, Robert & Jen Hay. 2012. LaBB-CAT. *Proceedings of the Australasian Language Technology Workshop*, vol. 10, 113–117. Dunedin, New Zealand.
- Gorman, Kyle, Jonathan Howell & Michael Wagner. 2011. Prosodylab-Aligner: A Tool for Forced Alignment of Laboratory Speech. *Canadian Acoustics* 39(3). 192–193.
- Kretzschmar, William A. 2011. *Linguistic Atlas Project*. www.lap.uga.edu.
- Labov, William, Ingrid Rosenfelder & Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language* 89(1). 30–65.
- Pederson, Lee, Susan L. McDaniel, & Carol M. Adams, eds. 1986-93. *Linguistic Atlas of the Gulf States*. 7 vols. Athens, GA: University of Georgia Press.
- Reddy, Sravana & James N. Stanford. 2015. Toward completely automated vowel extraction: Introducing DARLA. *Linguistics Vanguard* 1(1). 15–28. doi:10.1515/lingvan-2015-0002.
- Renwick, Margaret E.L. and Rachel M. Olsen. 2016. Voices of coastal Georgia. *Proceedings of Meetings on Acoustics*, 25, 60004. doi:10.1121/2.0000176.
- Rosenfelder, Ingrid, Joe Fruehwald, Keelan Evanini & Jiahong Yuan. 2011. *FAVE (Forced Alignment and Vowel Extraction) Program Suite*. <http://fave.ling.upenn.edu>.

Thank you!

This work is supported by NSF grant #1625680

Automated Large-Scale Phonetic Analysis: DASS Pilot

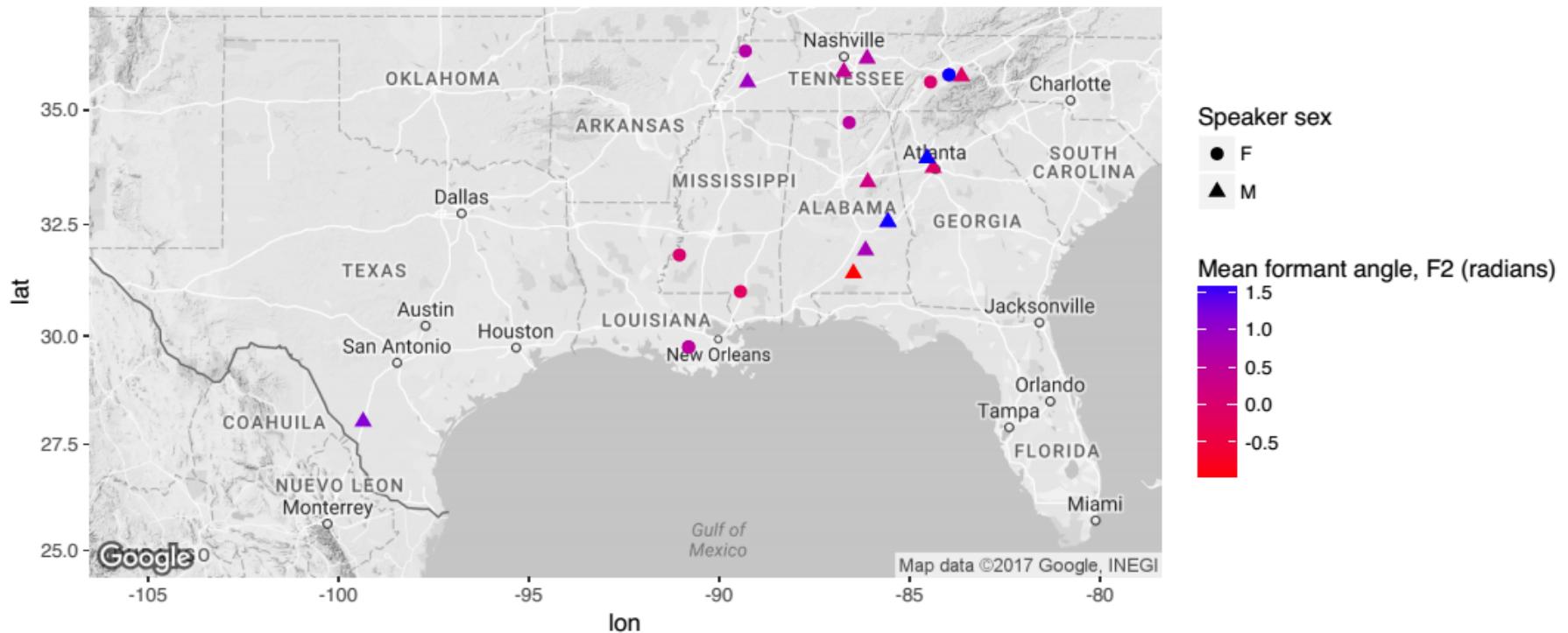
PIs: Drs. William Kretzschmar & Margaret Renwick.

Discussion

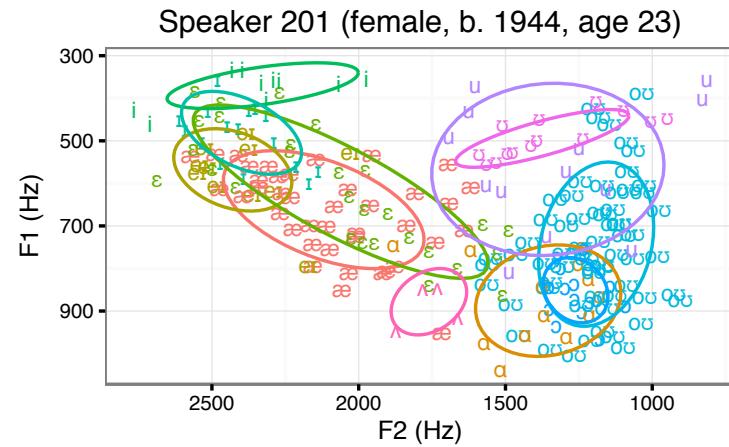
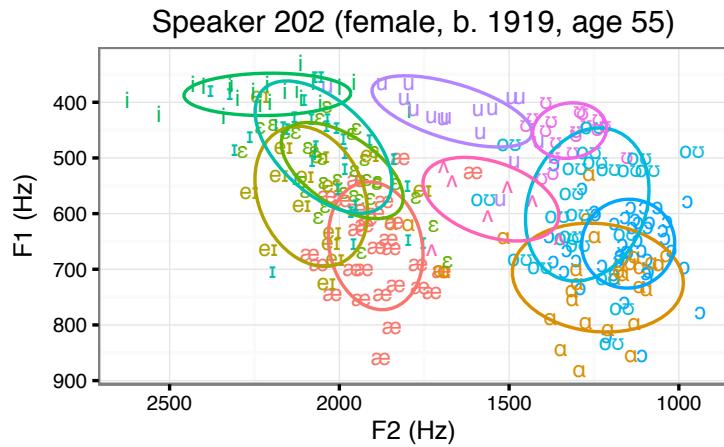
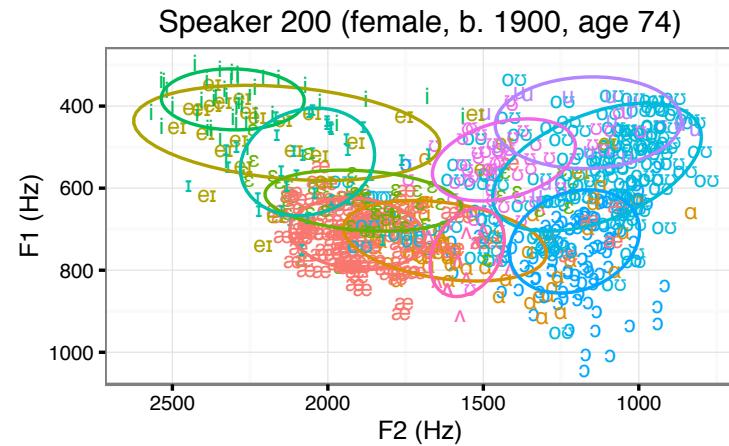
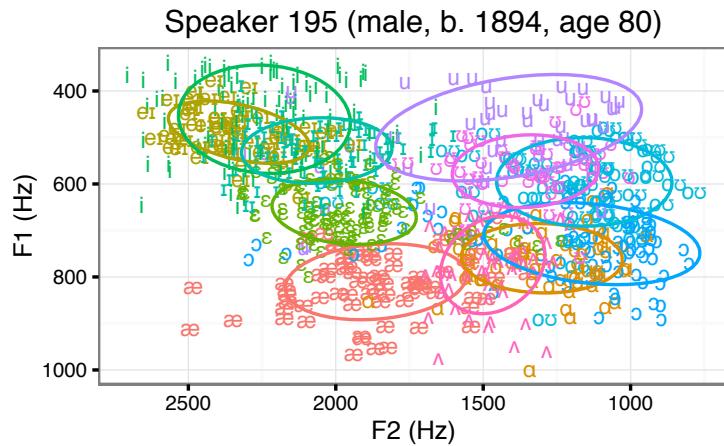
- ▶ Great free software available.
- ▶ Easy to use, even for novices.
- ▶ Linguistic Atlas data has much to offer!
- ▶ Large audio corpora can/should be built & can be analyzed.

Glide weakening

Mean formant angle, F2, final position: higher value = more diphthongal



Example Vowel Spaces



LAGS Speaker Area AK

► LAGS Protocols:

MMY	80	1B	2	LG ST. MARYS
1			1..	AK 109.02
c) mɛ̃nde	(2A.1)	c) sərtʃdɛ	(2A.1)	
	c) sən̩dɛ			1B.10 (+text 27.7.)
c) fr̩a..	dɛ			2A.5
2				
c) mɔ̃nɛ	n			1B.8

- 1031 tokens/spkr x 10 spkrs = 10,310 tokens
- Full transcription of interviews:
 - Searchable time-aligned corpus of 132,000 words

Linguistic Atlas of the Gulf States (LAGS)

