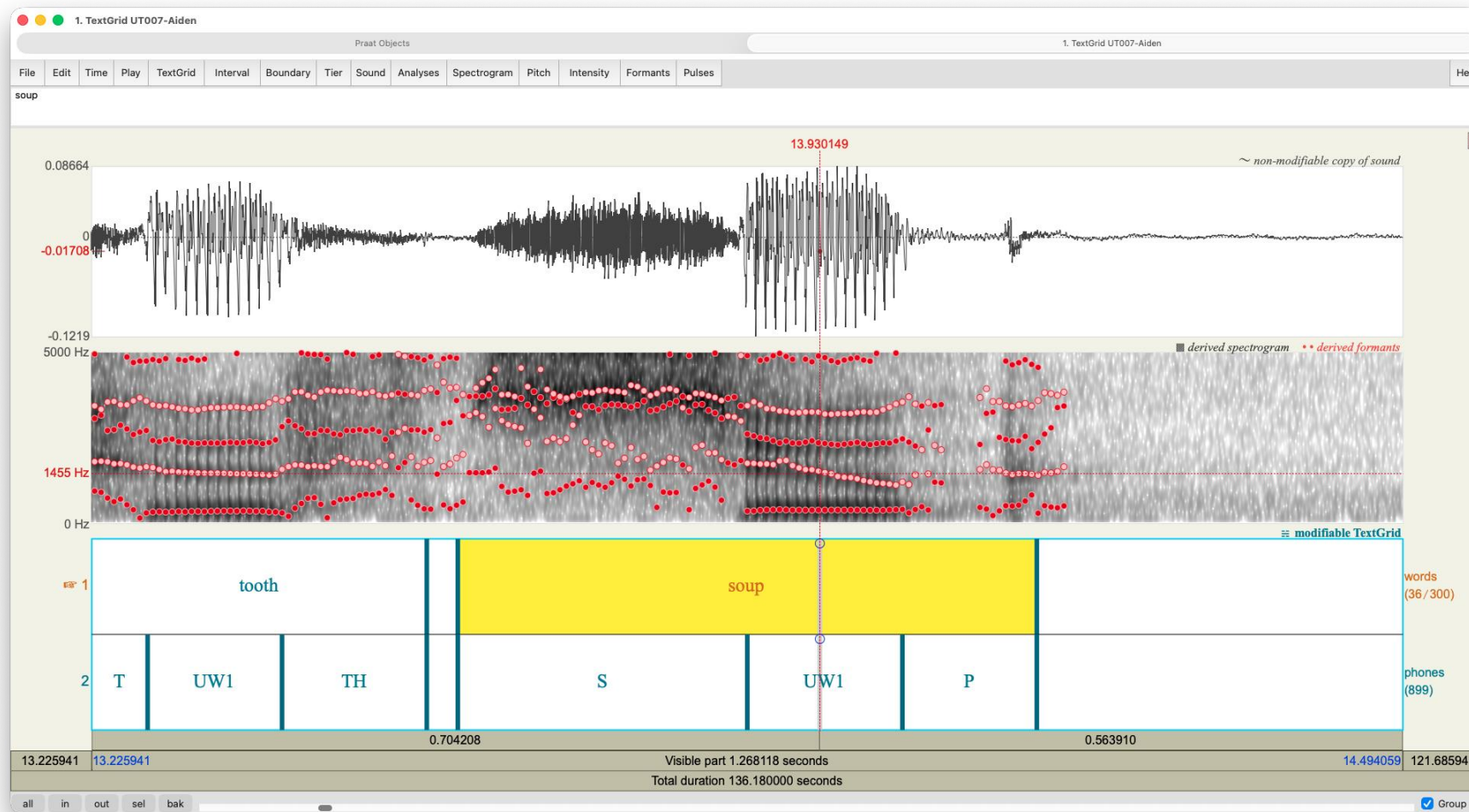


Recent Developments in Automatic Sociophonetic Processing

Joey Stanley
Brigham Young University

New Ways of Analyzing Variation 54
October 23, 2026
Montreal, Quebec

Transcription → Alignment → Formants



Audio → Spreadsheet

Transcription

- CLOx (Wassink, Squizzero, Fellin, Nichols 2018)
- Bed Word (Ma, Glass, & Stanford 2024)
- Whisper (Radford, Kim, Xu, Brockman, McLeavey, Sutskever 2022)
- etc.

Forced Alignment

- WebMAUS (Schiel 1999, Kisler, Reichel, & Schiel 2017)
- SPPAS (Bigi 2015)
- Gentle (Ochshorn & Hawkins 2015)
- Prosody-Lab (Gorman, Howell, & Wagner 2011)
- Montreal Forced Aligner (McAuliffe, Socolof, Mihuc, Wagner, & Sonderegger 2017)
- etc.

Formant Extraction

- FAVE-Extract (Rosenfelder, Fruehwald, Evanini, Yuan 2011)
- Fast Track (Barreda 2021a)
- new-fave (Fruehwald 2025a)
- etc.

Not Secure — darla.dartmouth.edu/uploadbound

DARLA: Automated Vowel Extraction

DARLA

Semi-Automated

Completely Automated

FAQs

Cite

Dartmouth Linguistics

Neukom Institute

Georgia Tech Linguistics

Automated Vowel Extraction with Manual Transcriptions and Utterance Boundaries

Upload your audio files and manual transcriptions (with sentence boundaries annotated [as described here](#)) to run forced alignment followed by formant extraction.

[Click here at any time to cancel or start over.](#)

Upload a .wav or .mp3 file: no file selected
Longer recordings (of at least 2 minutes) are recommended. Your uploaded files are stored temporarily on the Dartmouth servers in order to process your job, and deleted after.

Manual transcription as a .TextGrid file: no file selected
Textgrid should contain a tier named "sentence" with sentence/breath group intervals.

Filter out stop-words? ☒ Yes ☐ No
[This is the list](#) of stop-words we remove. (Link opens in a new tab.)

Filter out unstressed vowels? ☒ Yes ☐ No

Filter out vowels whose F1 or F2 bandwidths are over 300? ☒ Yes ☐ No

Your e-mail address:

We will not store or distribute your address.

VoxHumana is still in beta-testing mode. This means it is frequently getting updated and tweaked, which may occasionally cause things to break unexpectedly. Please be patient as I work through these bugs and please avoid overusing it (such as submitting many long files for transcription) so that others can try it out. I am happy to help fix issues, clarify how to use it, take suggestions for improvement, or just hear that you're using it and finding it helpful. Please reach out at voxhumana.ling@gmail.com.

UPLOAD



Drop audio here or click to browse

WAV · MP3 · FLAC · M4A · Max 1 GB

☒ TRANSCRIBE

Whisper

Language

English — english_us_arpa

Whisper model

Small (default, fastest)

▸ Advanced options

☒ ALIGN

MFA

Acoustic model

english_us_arpa

Dictionary

english_us_arpa

▸ Advanced options

☒ EXTRACT

new-fave

Extraction method ☒ new-fave (default) ☐ FAVE-extract (legacy, DARLA-style)

Formant ceiling (Hz) (optional)

Default: 10000

Number of formants (optional)

Default: 3

▸ Advanced options

Class code (optional)

Only if your instructor gave you one

SUBMIT

There are no jobs currently in the queue, so your job will start right away.



VoxHumana live demo

stanley.byu.edu/VoxHumana

Conclusions

- I hope this serves as a useful replacement for DARLA.
- I especially want students to get data quickly without the hassle of coding.
- This is still in beta-testing (especially for non-English jobs)
 - Feedback, bugs, questions
 - Suggestions for new features, languages, varieties, etc.
 - A quick note saying it worked well

Recent Developments in Automatic Sociophonetic Processing

Joey Stanley
Brigham Young University

Find VoxHumana as well as my other new tool, Pipeline, at

stanley.byu.edu