



<https://webtrans.ldcresearch.org>
<<< try the demo

High quality recordings and transcriptions of speech via remote platforms

Jonathan Wright, Robert Parker, Jeremy Zehr, Neville Ryant, Mark Liberman, Christopher Cieri, James Fiumara
Linguistic Data Consortium, University of Pennsylvania
Corresponding Author: Jonathan Wright jdwright@ldc.upenn.edu <http://www.ldc.upenn.edu>



<https://webtrans.ldcresearch.org>
try the demo >>>



Problem

- Numerous research domains leverage speech transcriptions; clinical diagnostic research in particular is an active area for LDC and others
- Commercial Speech to Text is now quite accurate but its deficits still leave room for using human transcription
- Notable deficits include omitting disfluencies (see bottom right of poster) and poor recognition in new vocabulary domains and unseen audio conditions
- While a commercial product may best target the intended utterance, much research requires accurate representation of the speech, including disfluencies
- However, the software landscape for recording and transcribing speech leaves much to be desired; no current platform to our knowledge combines all the features described here

Solution

- Custom mobile apps or web apps provide portable, remote, lossless recording options
- Web app allows remote and collaborative creation of transcripts
- Transcription interface has rich display and extensive keybindings
- Management interface allows coordination of large projects
- Transcription guidelines provide rich annotation and standard for training
- Cloud services provide extensible functionality
- LDC Speech Activity Detection provides pre-segmentation; humans correct boundaries
- Commercial Speech To Text inserted and then corrected saves time
- Validation of transcript in real time and second pass workflow for accuracy

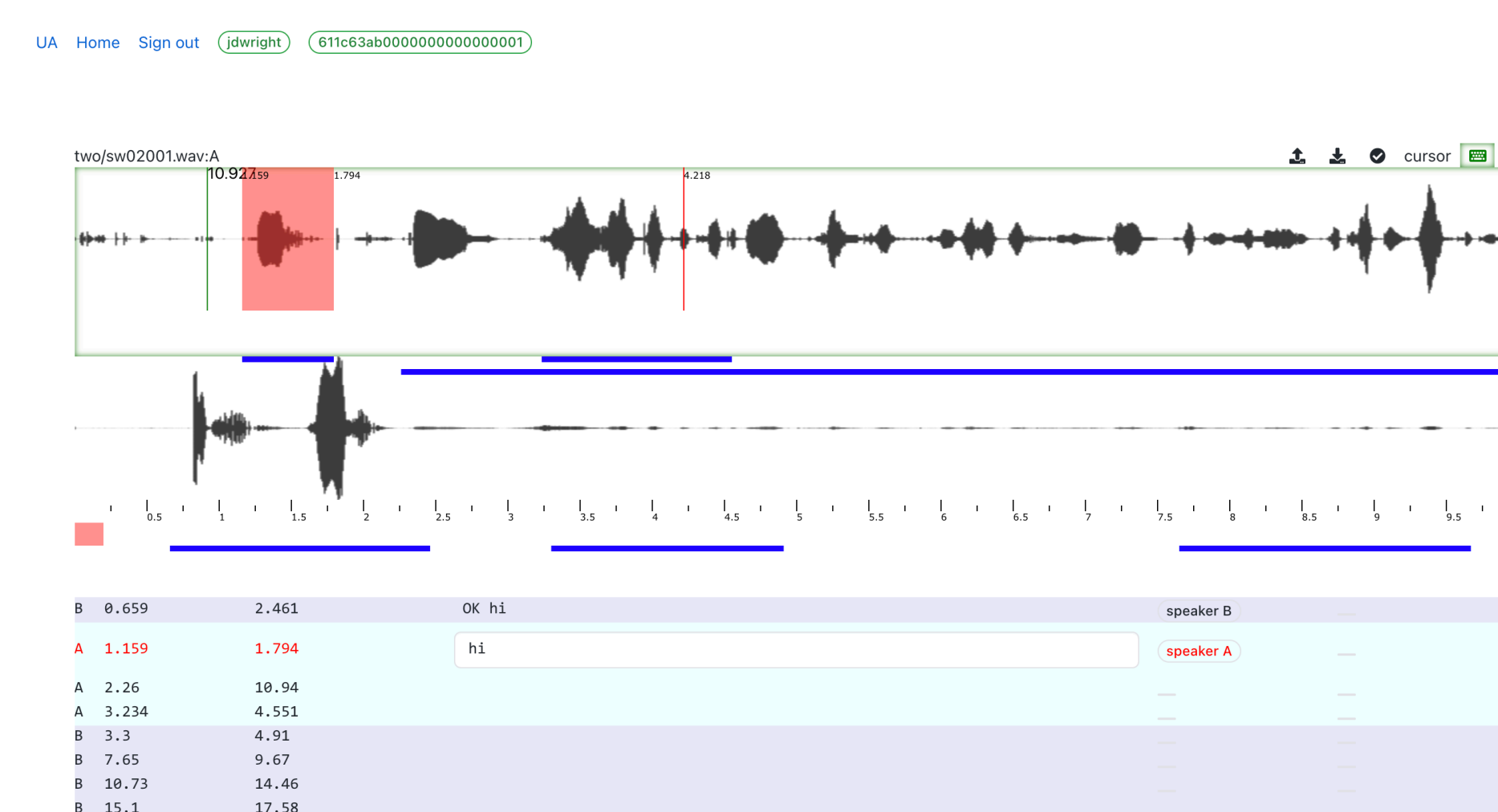
Distributed Infrastructure

- Rather than a single tool, we use a distributed infrastructure of many parts
- GUI is a traditional web app with Ruby on Rails backend
- Web app runs within docker for portability, including local installation
- LDC platform is a combination of separate services on AWS
- web app, S3 storage, Speech Activity Detection, Speech to Text, transcript validation
- Web standards are increasingly powerful and browsers are increasingly standards compliant
- Javascript is widely used and has little danger of obsolescence



Transcription Interface

- Central to fast transcription speeds is a simple and intuitive interface
- At top is the left to right oriented waveform, at bottom is the top to bottom oriented transcript
- The tabular transcript is made up of timestamps, transcript, speaker label, and section label
- Blue lines below the waveform indicate the segments present in the transcript
- Red in the waveform and transcript calls out the selected segment
- A green waveform line and green segment background indicate playback position
- Navigation, playback, and labeling can be performed solely with the keyboard
- Numerous zoom, scroll, and playback options



Channels, Speakers, Sections

- Channel separate from speaker
- Two channel mode automatic based on audio
- Speaker labels customizable and unlimited
- Labels could indicate some other dimension
- Section labels operate differently; they auto fill based on first and last segment

UA

Home

Sign out

jdwright

jump to readonly

Work

Projects

Browse

Kit Types

Tools

Data Sets

Invites

Scripts

Buckets

Features

Services

Workflows

Groups

Permissions

Help

Project Info

Tasks

Project Members

Speech Biomarkers

Return to project list

Task Info

Kits

Duals

Task Members

Output Tables

Mixer 6 Speech 9

Return to task list

All Kits

Create Kits

Bulk Reassign

301 rows

State anyanyUser anyany

301 matched

filter

ID	UID	State	Source	Continue	User	Broken Comment	Updated At
5437	5f0fba48000000000000000003	done	20090714_092751_L DC_120257_CH05_IN TV.wav	n/a	xiaonic	n/a	2020-07-21T21:42:35.472+00:00
7946	5f89e3a9000000000000000019	done	20090804_160044_L LDC_120318_CH05_I NTV.wav	n/a	neilsimp	n/a	2020-10-27T21:10:53.385+00:00
5435	5f0fba48000000000000000001	assigned	20090713_153843_L DC_120309_CH05_IN TV.wav	n/a	jdwright77	n/a	2020-07-16T02:24:08.447+00:00
4705	5f0f339e000000000000000031	done	20090713_144125_L DC_111268_CH05_IN TV.wav	n/a	jdwright	n/a	2020-08-28T18:43:32.530+00:00

Management Interface

- A single instance can accommodate any number of users and tasks
- All users have accounts, both managers and transcriptionists
- Managers create tasks, add users to tasks, and create assignments
- Assignments can be followed from unassigned to assigned to done
- Percentage of dual assignments can be set
- First and second pass tasks can be connected; as assignments are completed in the first pass task, assignments in the second pass task are automatically created with transcript copies
- Transcripts can be downloaded as plain text, tab delimited tables

Special Transcript Elements

- filled pauses
- repetitions
- partial words
- speech errors
- non-speech noises
- unintelligible regions

Download Transcript

Filename (.tsv will be appended)

☐ Include Headers

☒ Include Speaker

☒ Include Section

Create a transcript file, which you can preview below and then Download

Preview

Kit	Audio	Rep	End	Text	Speaker	Section
6197a560000000000000000000000000	CarrieFisher.wav	8.496	2.532	That's Carrie Fisher as Princess Leia	Terry	Li
6197a560000000000000000000000000	CarrieFisher.wav	2.75	4.56	and Harrison Ford as Han Solo	Terry	Introdu
6197a560000000000000000000000000	CarrieFisher.wav	4.78	6.35	In the First Star Wars Film	Introdu	
6197a560000000000000000000000000	CarrieFisher.wav	6.81	9.47	Fisher has just written a memoir about making the		
6197a560000000000000000000000000	CarrieFisher.wav	9.7	13.62	Included in the book is the diary she wrote while		
6197a560000000000000000000000000	CarrieFisher.wav	13.29	16.45	when she was having a secret affair with Harrison		
6197a560000000000000000000000000	CarrieFisher.wav	16.53	18.83	the memoir is called the Princess Diarist		
6197a560000000000000000000000000	CarrieFisher.wav	19.47	21.35	Carrie Fisher welcome back to Fresh Air	Terry	Li
6197a560000000000000000000000000	CarrieFisher.wav	21.58	23.92	so how did you find your Star Wars diary?	T	