

Learning the LDC Web Based Transcription Tool

Jonathan Wright

February 2022

1 Introduction

In recent years, LDC has developed a web based infrastructure for the efficient creation of high quality speech recordings and transcripts for use in engineering, clinical, and linguistic research. There are several components to the infrastructure including the browser based transcription tool, cloud storage, Speech Activity Detection, Speech To Text, Forced Alignment, and recording apps. Unlike a typical application, this software has multiple possible configurations, from a local installation on a laptop, to a publicly available cloud service, or a similar installation on a private network.

The tutorial begins with the use of an LDC produced mobile recording app, and how to set up recordings locally or in the cloud. We then guide participants through installation of the transcription tool on their own laptops to transcribe local audio files. We then illustrate how cloud configurations work, and participants can use existing LDC instances. Next, the tutorial goes through the basic transcription features and how those features can be customized. This includes the use of external services like Speech Activity Detection or Speech To Text, and when using a cloud instance this includes user and task management. Finally, the tutorial will cover the options for downloading the transcripts.

2 Outline

2.1 Recording

We begin with demonstrating our recording software. Participants will be invited to download an app to their mobile device, as well as use a web based app from their laptops. These apps are very simple, essentially involving record/stop, and simple meta data selection, so the discussion will primarily be about customization. The demo apps will upload to LDC cloud storage, so we will illustrate how they can be connected to new storage accounts. Furthermore, a one size fits all approach to field recording is unrealistic, and so we will cover the modification of the app's meta data or appearance.

2.2 Transcription

Most of the tutorial will focus on our transcription software, informally known as *webtrans*.

2.2.1 Installation

We begin with installation and configuration. Participants will be invited to install the software on their own laptops. The application is a web app, so using it locally still involves using via a web browser. The first step is to install docker, a software system for running other applications in isolation, and then docker will handle the installation of webtrans. The process is simple, and allows the user to transcribe audio completely locally with no network connection. Next, the participants will connect their local instances to LDC cloud services, to illustrate how they might connect to their own cloud services (for example cloud storage of audio). Next, we discuss more distributed configurations of webtrans, like installing on an institutional server, or installing in the cloud, which means end users don't need any software beyond a web browser.

2.2.2 Transcription

Next in the tutorial we focus on the transcription tool itself. The tool has been designed with efficiency in mind. Users can optionally use ASR or Speech Activity Detection to start off the transcript, and there are a variety of keybindings that allow mouse free usage. The tool has a built in tutorial that highlights the most important features; we will go over this and other features as well. Participants will be able to experiment with their own local instance, or use one of LDC's cloud instances if desired. Before moving on, we'll demonstrate how to download the transcripts and look at their representation.

2.2.3 Task Management

While creating transcripts may be sufficient for some users, others will want user and task management functionality, which we examine next. In any server based installation, multiple users will be able to create accounts and log in at the same time. Using one of LDC's cloud based instances, we will demonstrate how different tasks can be created for different sets of users. Tasks can vary not only in the data covered, but the details of how assignments are made and what options are available in the transcription tool. For example, one available option is to create a second pass task, where the transcripts created by one set of users are made available to another set of users for careful review.

2.3 HLT services

While HLT services would have been touched upon in previous parts of the tutorial, for example, in the use of Speech Activity Detection to automatically create transcription segments, in this part of the tutorial we will examine the services more closely. We will look at the details of some services that connect to webtrans, and as well as how to connect new services that might be available to end users.