

Automatic Speech Parser for ALS Data

A MATLAB Software developed by Saara Raja

PURPOSE:

- This is a software program that is used to automatically parse audio and motion data obtained from (Amyotrophic Lateral Sclerosis) ALS patients by the University of Texas at Dallas.
- It was developed for use of the Speech Disorders and Technology lab run by Dr. Jun Wang.
- Its function is to optimize the data processing stage by automatizing functions that must otherwise be done manually.
- The data parsed by the Automatic Speech Parser will be extracted and formatted for future statistical analysis.

RESEARCH DESIGN:

- Speech data is collected from ALS patients by attaching sensors to specific positions on the patients' tongue and lips.
- An electromagnetic articulograph (NDI Wave system) is used to detect the sensors and thereby record 3-dimensional tongue position and movement in real time (Figure 1).
- For speech detection, patients repeat 20 commonly used phrases, for which both the audio and tongue motion are recorded
- All 20 phrases are recorded in a single sound bite during data collection and must be manually parsed into individual phrases for future analysis.

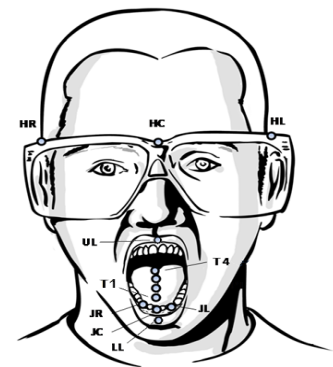


Figure 1

SOFTWARE DESIGN:

- Once the sound and motion data files are loaded into the automatic parser, both data types are displayed visually (Figure 2).
- The audio data is displayed first in audiowave form; the waveforms for all 20 phrases will be displayed.
- The speech motion data is displayed below the audio, and is segmented by Sensor Type (Tongue Tip, Tongue Back, Upper Lip, Lower Lip, Jaw Left, Jaw Right) and Dimension (x, y, z)
- The user has the option to listen to the audio by clicking 'Play Audio' or to begin the automatic parsing process by clicking 'Start Tagging'

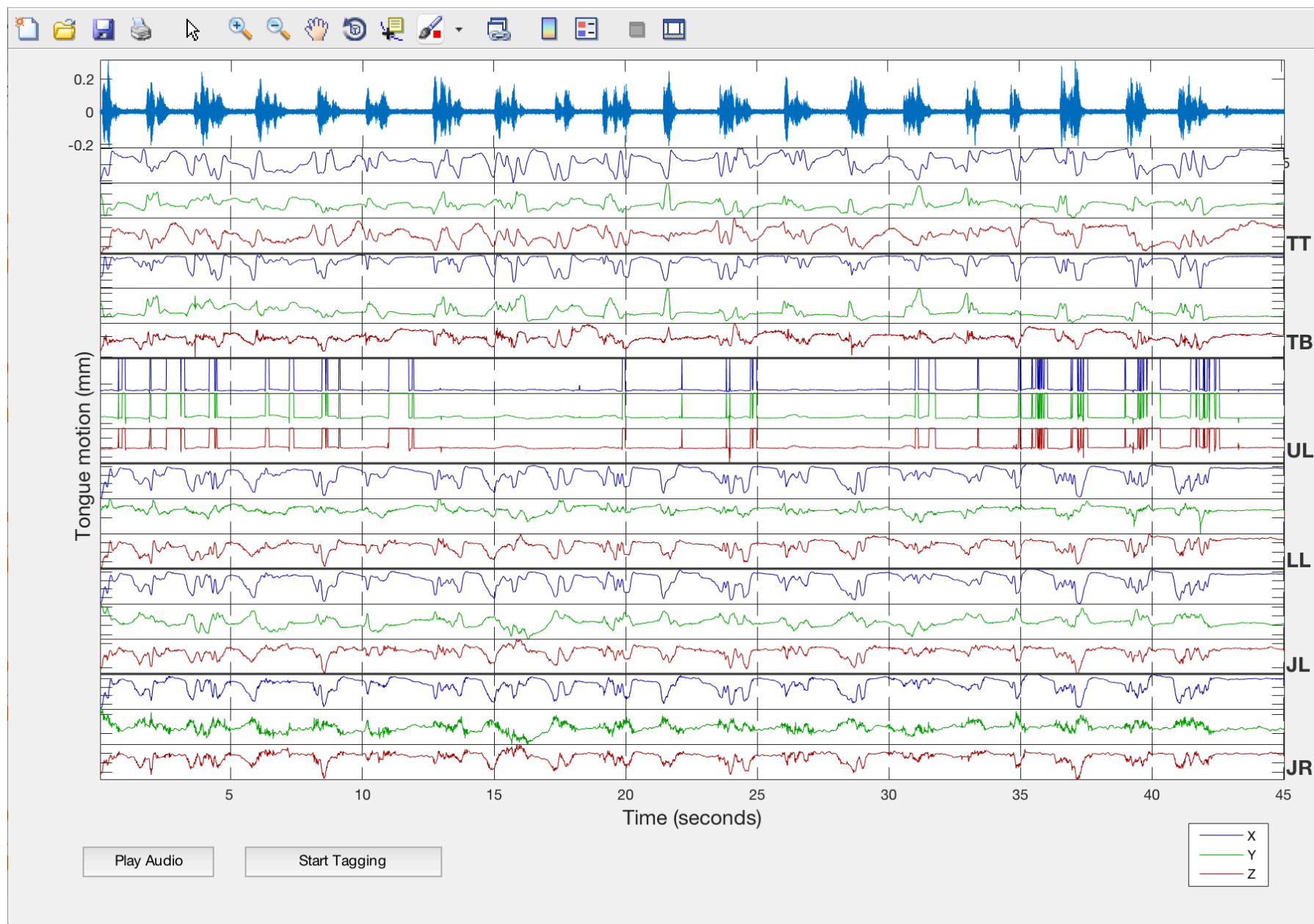


Figure 2

SOFTWARE ALGORITHM:

- The program will successfully separate each phrase from background noise automatically by detecting a predetermined threshold in audio amplitude (Figure 3).
- When the waveform amplitude's absolute value exceeds the threshold, the software recognizes that this is the beginning of speech.
- Likewise, from that point onward, when the waveform absolute value goes below the threshold, the software recognizes that this is the end of the phrase.
- Within each phrase there will be pauses between sentences in which the amplitude will go below the threshold (such as the pauses between the words 'How are you?').
- To avoid incorrectly splitting each word into a phrase, the parser will only tag the end of phrases if the dip below the threshold exceeds 0.3 seconds.
- The program will loop through the entire sound bite until all 20 phrases are tagged.

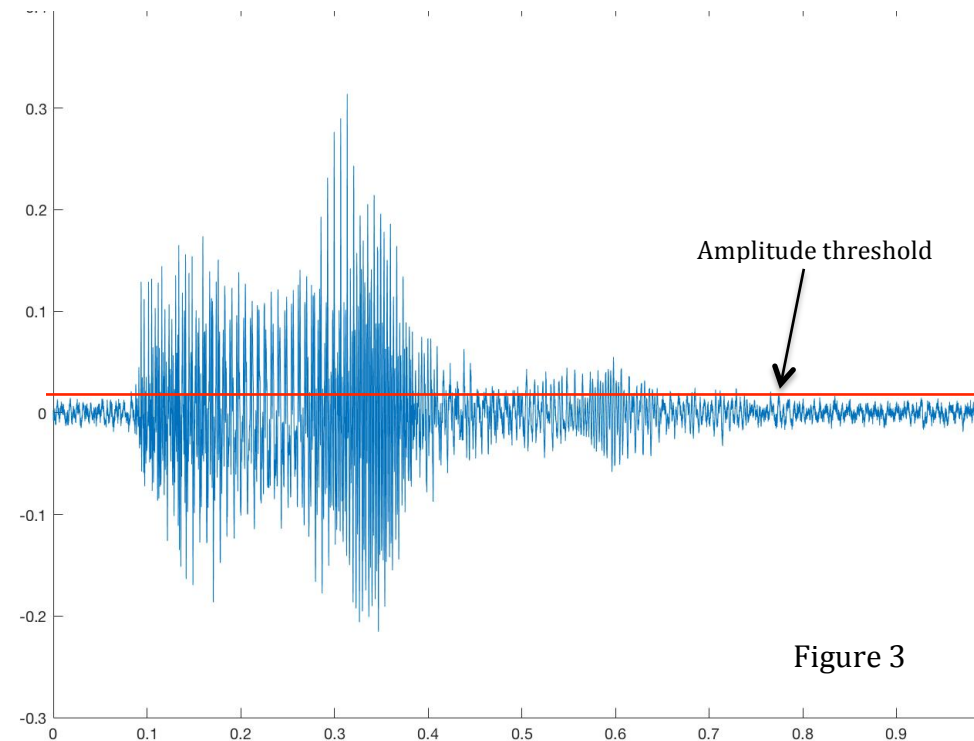


Figure 3

CORRECTING ERRORS:

- The software has several in-built features to prevent errors in parsing (Figure 4).
- The graphic user interface automatically zooms in on what it thinks is an individual phrase; if this viewing window is incorrect, the user can enter time bounds and click 'Adjust' in order to adjust the viewing window.
- The user can also manually change the tags that mark the beginning and end of the phrase by clicking on either 'Tag 1' or 'Tag 2' respectively.
- When the user is satisfied with the tag positions, he/she can click 'Accept Tags' to save the changes.

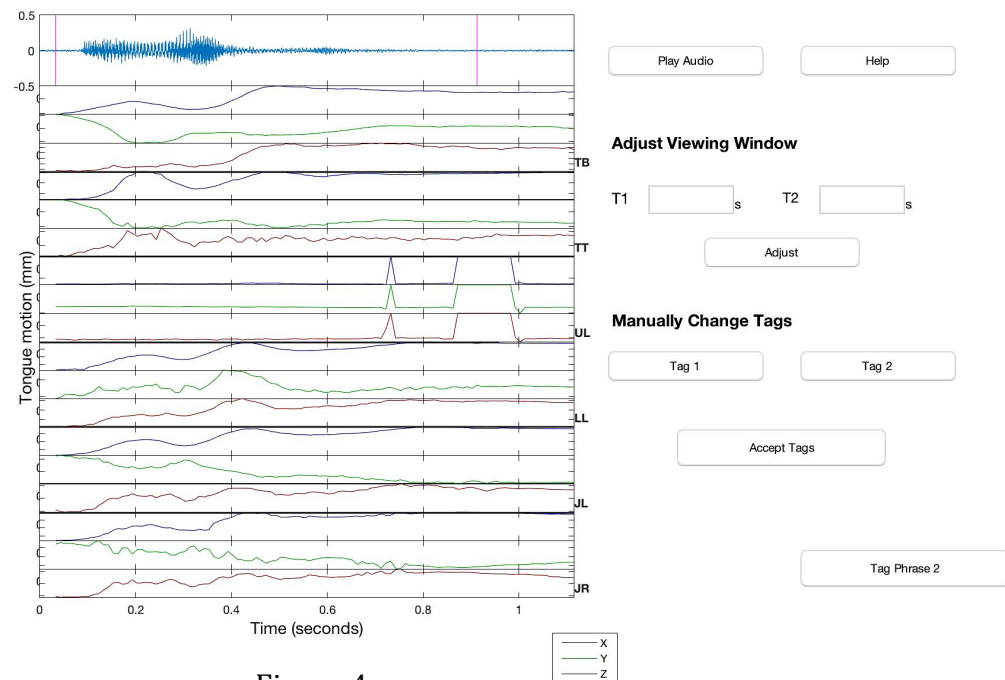


Figure 4

DEMONSTRATION:

- You can view the code for the Automatic Speech Parser at <https://github.com/saararaja/Automatic-Sound-and-Motion-Data-Parsing>
- To run a demo, simply download the code and sample data and run on MATLAB 2015 or above