INTONATIONs Detection for English Accent and Audio Processing

Presented By
Ayush Jain
1309713026

Intonation

- It is the rise and fall of words pitch use for expressing emotions and attitudes, to identify grammatical structure, to show what information in the utterance is new and what is already known, to show how clauses and sentences go together in spoken discourse, to organize speech into units that are easy to perceive, memorize and perform, to act as a marker of personal or social identity etc.

Sentence Structure

•Every sentence is a combination of various parts of speech in different order. They contains different Noun Phrase, Verb, and Prepositional Phrase etc.

Example:

- S -> NP VP
- VP -> V NP
- Articles, Proper-Noun, Adjective etc.

Inputs

•Inputs include an audio raw file either with extension .raw or .wav with Sample Rate 16000HZ, Sample Format 16-bit Float PCM, and Channel MONO.

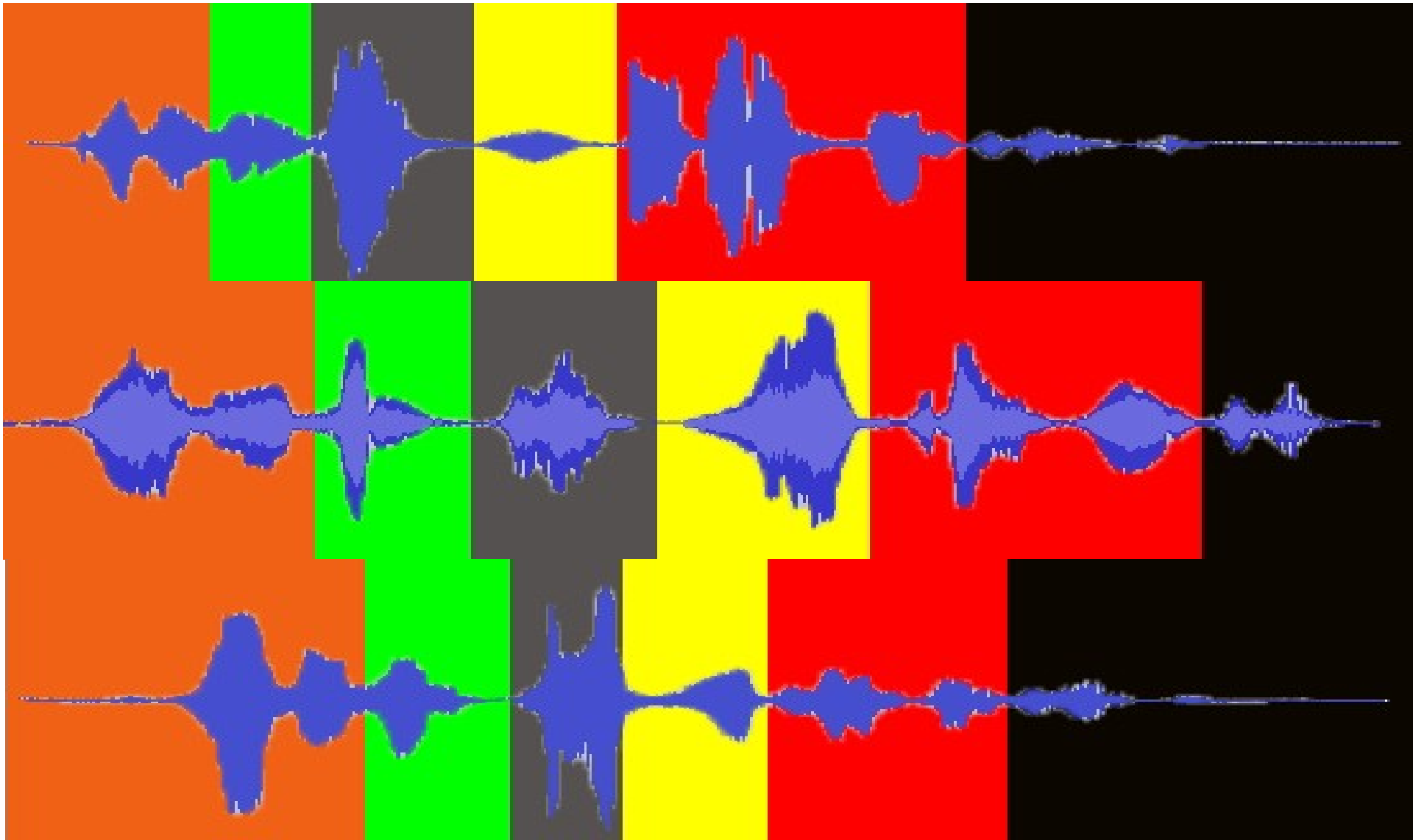
• Different tools are available online to convert files in this format.

• Audacity is used to synthesis the audio.

Tools

- Audacity –AN open source tool for manipulating audio files.
- Eclipse – An IDE for software Development.
- Database – To show and display datas.
- Android Stdio – An IDE for developing android mobile applications.

AUDIO IN AMPLITUDE FORM



Sentence recognize	Conclusion Drawn	Question
Ravi will drive me to the cinema tonight	Not somebody else but Ravi.	Who will drive you to the cinema tonight?
Ravi **will** drive me to the cinema tonight	It's definite and certain to happen.	Will Ravi drive you to the cinema tonight?
Ravi will**drive** me to the cinema tonight	How is it happening?	How will Ravi take me to the cinema tonight?
Ravi will drive **me** to the cinema tonight	Not somebody else but me.	Who Ravi will drive to the cinema tonight?
Ravi will drive me **to** the cinema tonight	Not from or back but to.	Where Ravi will drive me tonight?
Ravi will drive me to the **cinema** tonight	Where?	Where Ravi will drive me tonight?
Ravi will drive me to the cinema **tonight**	When?	When Ravi will drive me to the cinema?

Algorithm

- Highest Peaks
- Discrete Average
- Moving Average
- Area Covered
- Time Constraints‘ of words
- A number of top picks.

Algorithm

Peaks<Intensity, Duration>= Words Sentence, push (Intensity, Duration) from
Words Start Duration to End Duration.

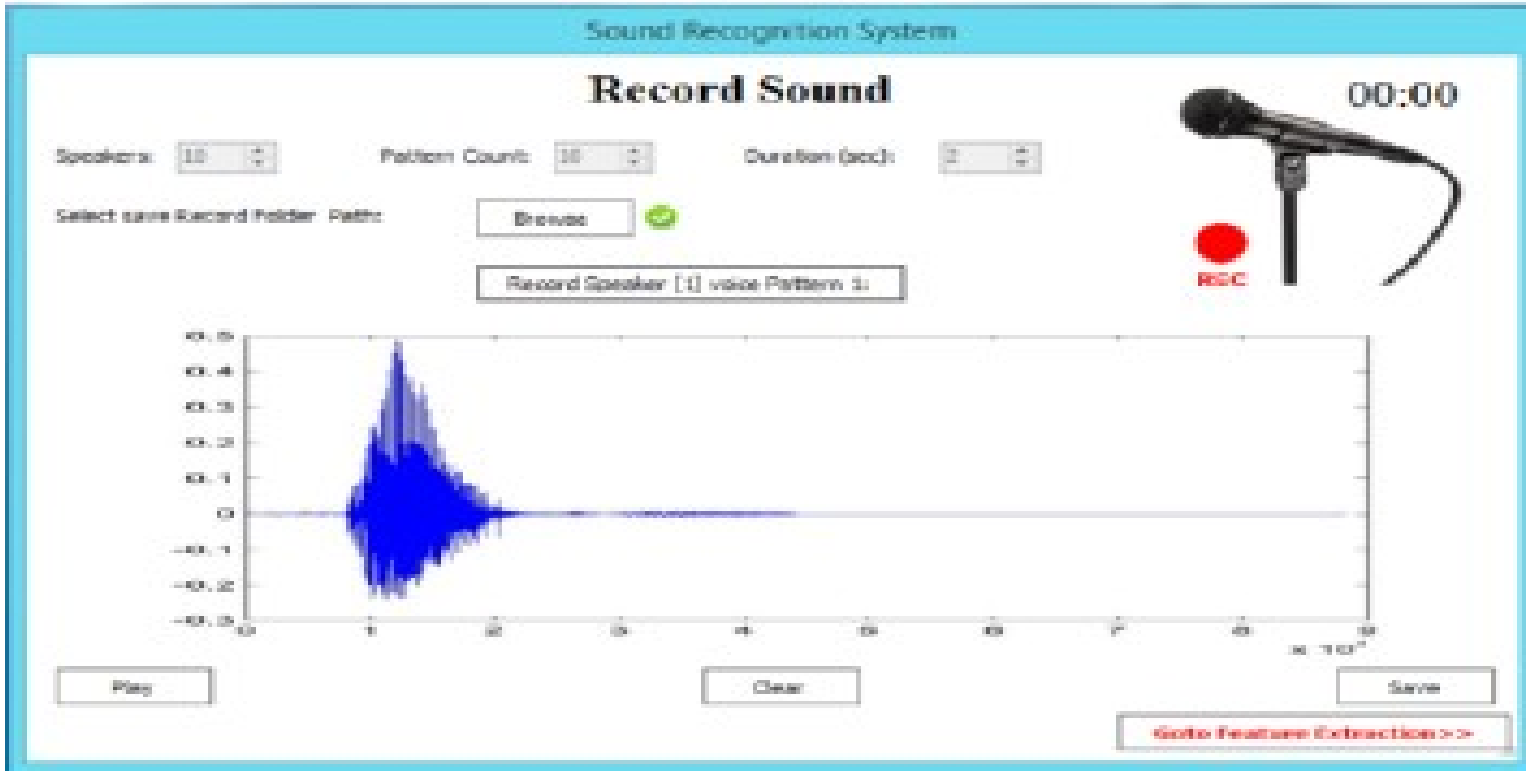
Top<Intensity, Duration>= Pick top 100 in Peaks on basis of Intensity.

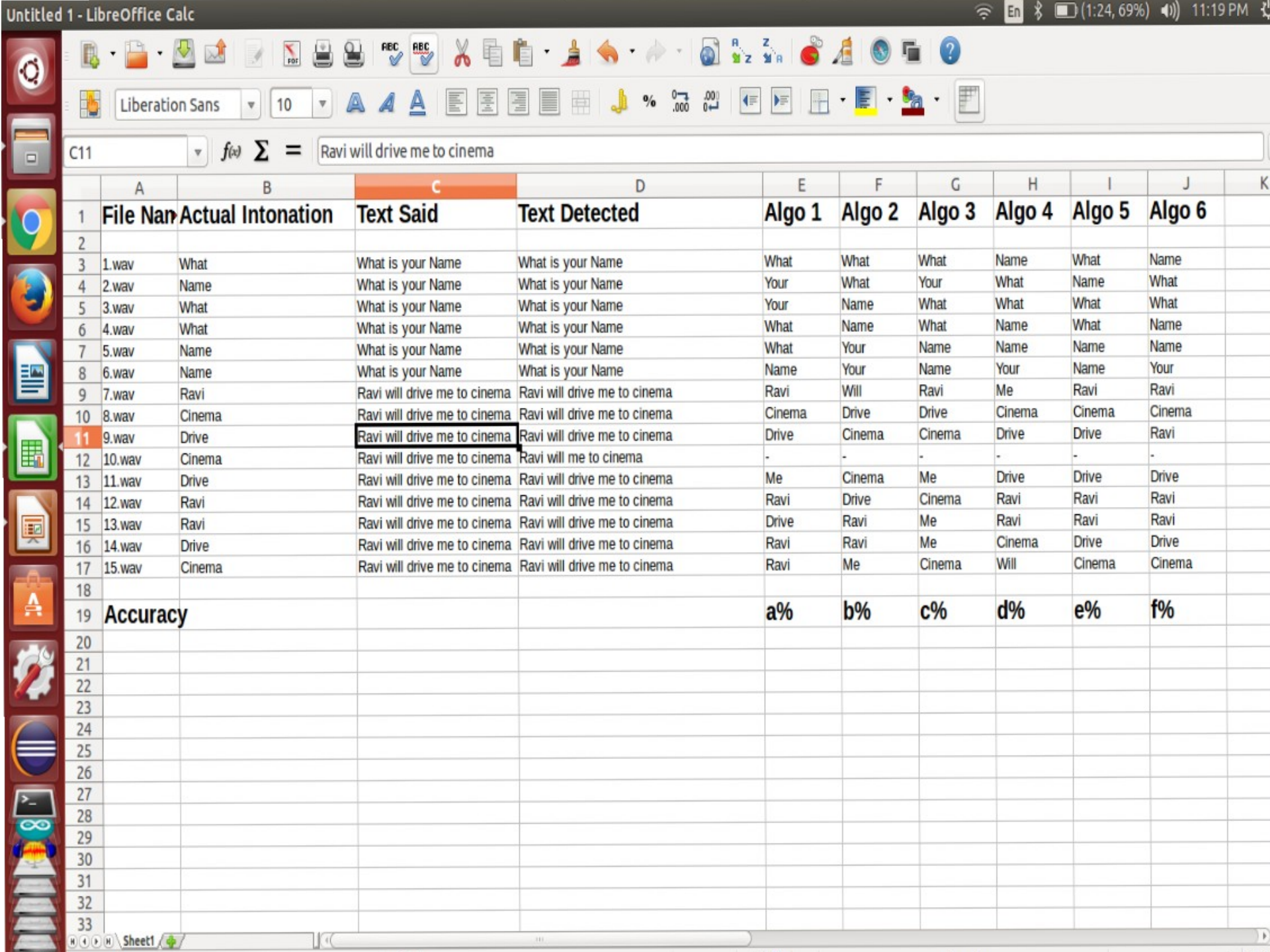
Total Intensity Sum of word =
If (top<Duration> (Word Start to Word End Duration))
Total Intensity sum =top<Intensity>

Score of word =Total Intensity Sum/ (Word End Duration- Word Start Duration)

GUI

It will consists of a basic GUI.





	A	B	C	D	E	F	G	H	I	J
1	File Nan	Actual Intonation	Text Said	Text Detected	Algo 1	Algo 2	Algo 3	Algo 4	Algo 5	Algo 6
2										
3	1.wav	What	What is your Name	What is your Name	What	What	What	Name	What	Name
4	2.wav	Name	What is your Name	What is your Name	Your	What	Your	What	Name	What
5	3.wav	What	What is your Name	What is your Name	Your	Name	What	What	What	What
6	4.wav	What	What is your Name	What is your Name	What	Name	What	Name	What	Name
7	5.wav	Name	What is your Name	What is your Name	What	Your	Name	Name	Name	Name
8	6.wav	Name	What is your Name	What is your Name	Name	Your	Name	Your	Name	Your
9	7.wav	Ravi	Ravi will drive me to cinema	Ravi will drive me to cinema	Ravi	Will	Ravi	Me	Ravi	Ravi
10	8.wav	Cinema	Ravi will drive me to cinema	Ravi will drive me to cinema	Cinema	Drive	Drive	Cinema	Cinema	Cinema
11	9.wav	Drive	Ravi will drive me to cinema	Ravi will drive me to cinema	Drive	Cinema	Cinema	Drive	Drive	Ravi
12	10.wav	Cinema	Ravi will drive me to cinema	Ravi will me to cinema	-	-	-	-	-	-
13	11.wav	Drive	Ravi will drive me to cinema	Ravi will drive me to cinema	Me	Cinema	Me	Drive	Drive	Drive
14	12.wav	Ravi	Ravi will drive me to cinema	Ravi will drive me to cinema	Ravi	Drive	Cinema	Ravi	Ravi	Ravi
15	13.wav	Ravi	Ravi will drive me to cinema	Ravi will drive me to cinema	Drive	Ravi	Me	Ravi	Ravi	Ravi
16	14.wav	Drive	Ravi will drive me to cinema	Ravi will drive me to cinema	Ravi	Ravi	Me	Cinema	Drive	Drive
17	15.wav	Cinema	Ravi will drive me to cinema	Ravi will drive me to cinema	Ravi	Me	Cinema	Will	Cinema	Cinema
18										
19	Accuracy				a%	b%	c%	d%	e%	f%

Output

- It helps to detect any kind of audio file effectively.
- It also helps to correctly detect the stressed part in the sentence of audio file.

USAGE

- In Field of Robotics.
- Machine Learning
- Telecommunication for Improving English
- Sarcasm Detection

References

- Automatic sentence stress feedback for non-native English learners.
- Wikipedia .
- Toby Model.

Conclusion

- The above algorithm will help to identify and detect stress and improve vocabulary. And the APIs can be used by many different telecommunications companies to improve their workers vocabulary.