

Table of Contents

Introduction

TABLE OF CONTENCT!!!!!!!!!!

**Introduction**

Researches show that there is a strong connection between different characteristics in speech and the ability to persuade others. This ability depends a lot on the level of confidence that a speaker shows, which can be measured using different speech characteristics.

In modern times in general, and after the outbreak of the corona pandemic in particular, the digital platform is gaining more momentum. You can see a lot of different fields that have been transformed into online. Communication through various online platforms is also in common use in workplaces. Today, the world of social media occupies a large place in daily communication, which evokes the need to improve the performance of verbal communication.

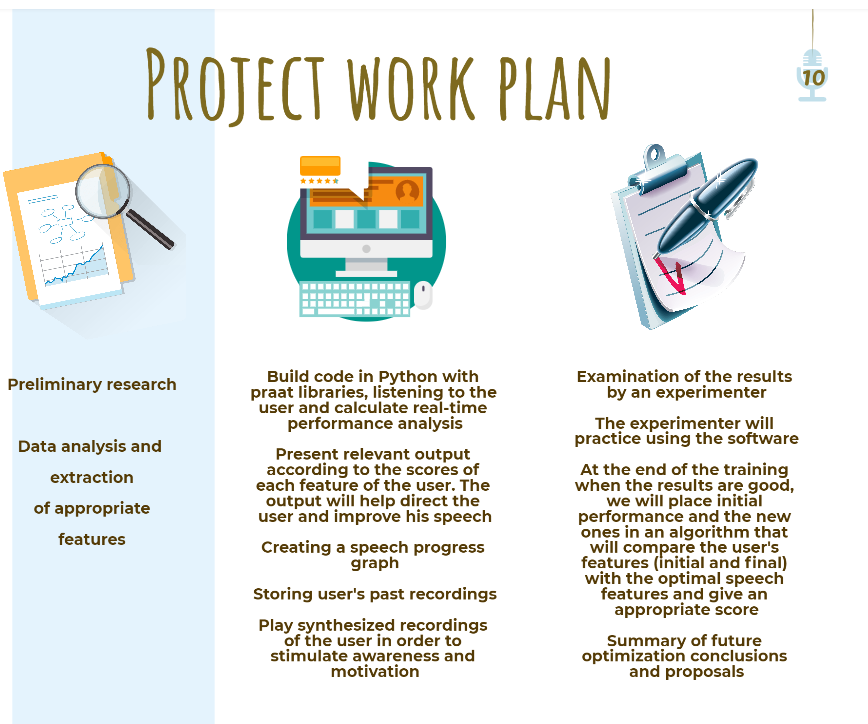
In developing speech recognition and analysis based training, we hope we can improve the speech ability of a person who has difficulty speaking confidently, or of any person who wants to strengthen abilities such as - persuasion, ability to stimulate motivation, trust and interest among listeners / audiences / interviews / dates / employers and more.

This software can suit people in a wide range of fields: sales agents, teachers, people who want to give a speech / lecture, interviewees, dating, people whose work platform is digital (Facebook, Instagram, YouTube, Zoom etc.) or anyone who wants Improve his vocal communication and persuasive abilities.

**Project Assumption**

In this project, we relied on a strong connection between different speech characteristics, and the level of confidence and persuasiveness that the speaker demonstrates.

We assumed that by improving the various metrics and characteristics of speech, it would be possible to provide impressive results in the speaker's performance and thereby enabling him to increase his persuasiveness and confidence.

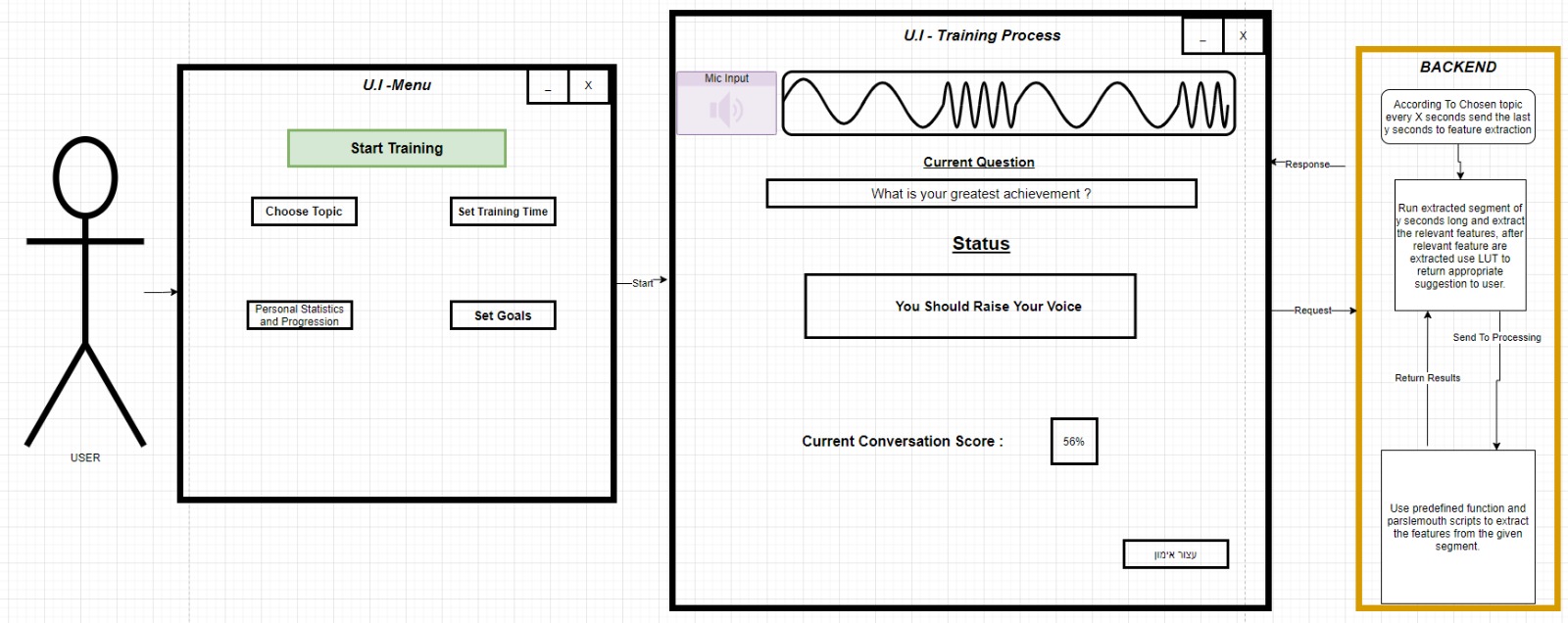


**The program**

We took the relevant measurement features from past research in the field.   
In order to obtain a linear graph indicating the relationship between the change in each trait we selected, and the level of confidence the interviewee exhibited, we used a dataset of recordings from a job interview experiment conducted at M.I.T University [1]. We examined which characteristics mostly affect the grade received by each interviewee in his or her job interview. At the grading stage, the scorers did not test knowledge or experience in a particular field, but mostly relied on the interviewer’s communication with the interviewee and his or her vocal characteristics. We chose to focus on several key features: Pitch, intensity-Max, intensity-Min and word repetitions.

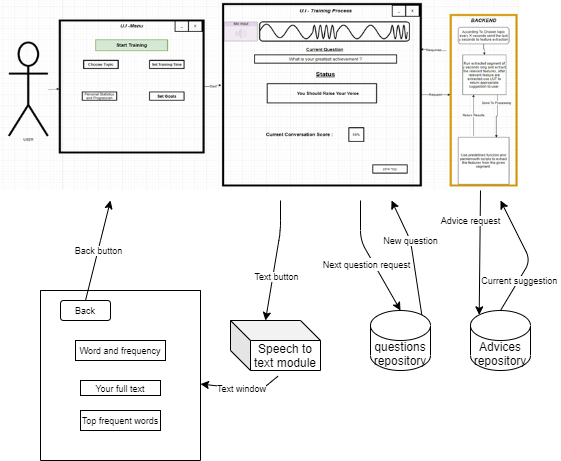
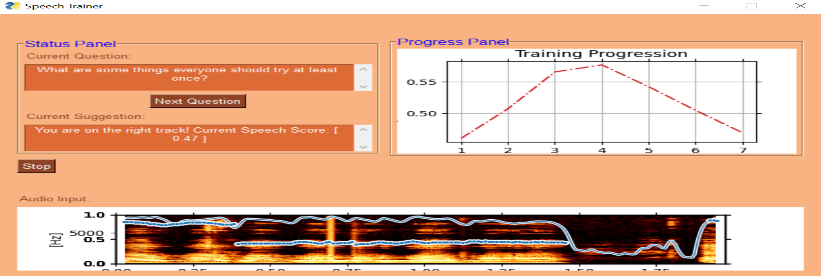
We developed the system in Python using special speech processing libraries.

**The Model Structure**



Next Question

Calculate text



1

2

4

5

6

7

8

3

A

B

Note: the numbered arrows are related to the scheme

**Backend**

1

1. User recording: From the moment 'start training' has been clicked, the user is recorded. The record was made by using the PyAudio library.

2. Calculation of speech characteristics: Using parallel programming, the software calculates behind the scenes the speech characteristics in real-time and also presents the user some improvement-suggestions or compliments among with his or her current score. We calculated the speech characteristics using the Praat library - ParselMouth.

2

3. Selection of speech characteristics that we calculated and chose to improve: (Based on past research):

3

* Pitch - We found that the Pitch feature is a very influential factor. We found that it can be improved by using descending intonation at the end of a sentence - which makes the user sound more confident [5] . We chose to explain to the user that as long as he is not asking a question, the tone should be lowered at the end of a sentence.

• Intensity Max - We found that speech volume is very important. Very weak speech can be a result of big distance/technical malfunction in the microphone, as well as from insecure and hesitant speech.

• Intensity Min - We found that intensity that is too high also detracts from the user's performance. This can be a result of speaking with a strong emotion (like anger, aggression).

• Word repetitions - We found that repetitiveness of words detracts from user performance. While using rich vocabulary creates better persuasive ability. That is why we chose to emphasize the speaker's text. If the speaker repeats many words, the system shows the word and its number of repetitions, in addition to printing all the spoken text. To calculate this feature, we used the Speech\_Recognition library.

4

4. Displays a spectrogram that includes the Intensity and the Pitch metrics in real-time speech.

5

5. Displays a progress graph based on the calculated user scores, in real time.

6. Presentation of questions according to the topic - We have prepared a wide database of questions -a total of 2,258 questions. This database includes questions from job interviews, dating questions and conversations. If the user asks for a new question, the system pulls out another question depending on the topic chosen and in a random form.

A

6

7. Displaying comments and directions - Depending on the feature to be improved, the system displays the appropriate message and guidance. If there are no comments and the metrics look fine, the system randomly displays a "compliment".

B

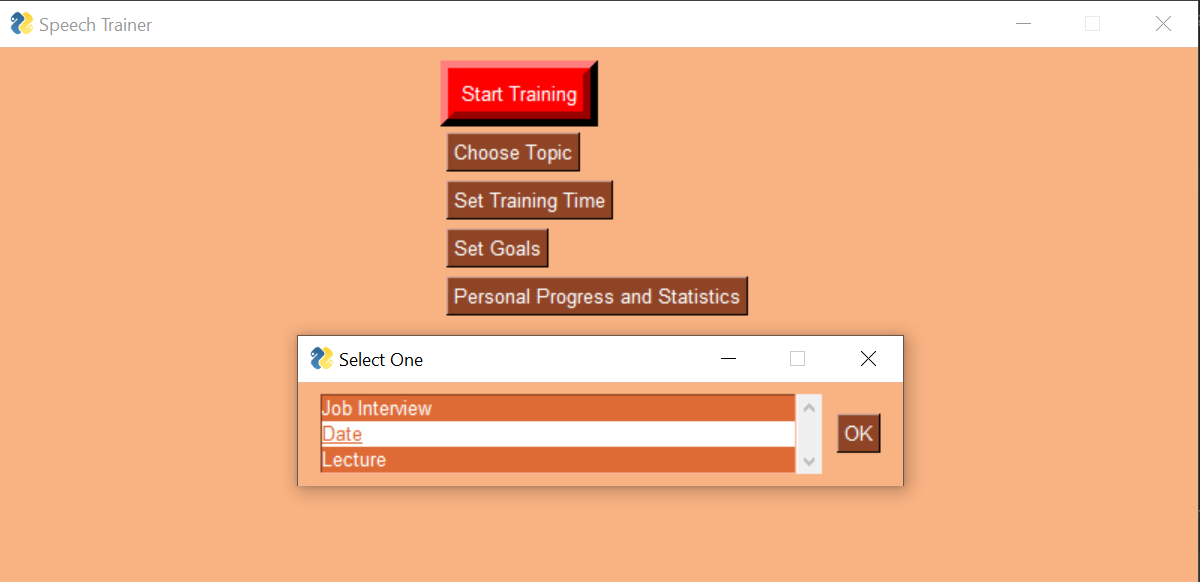
7

8

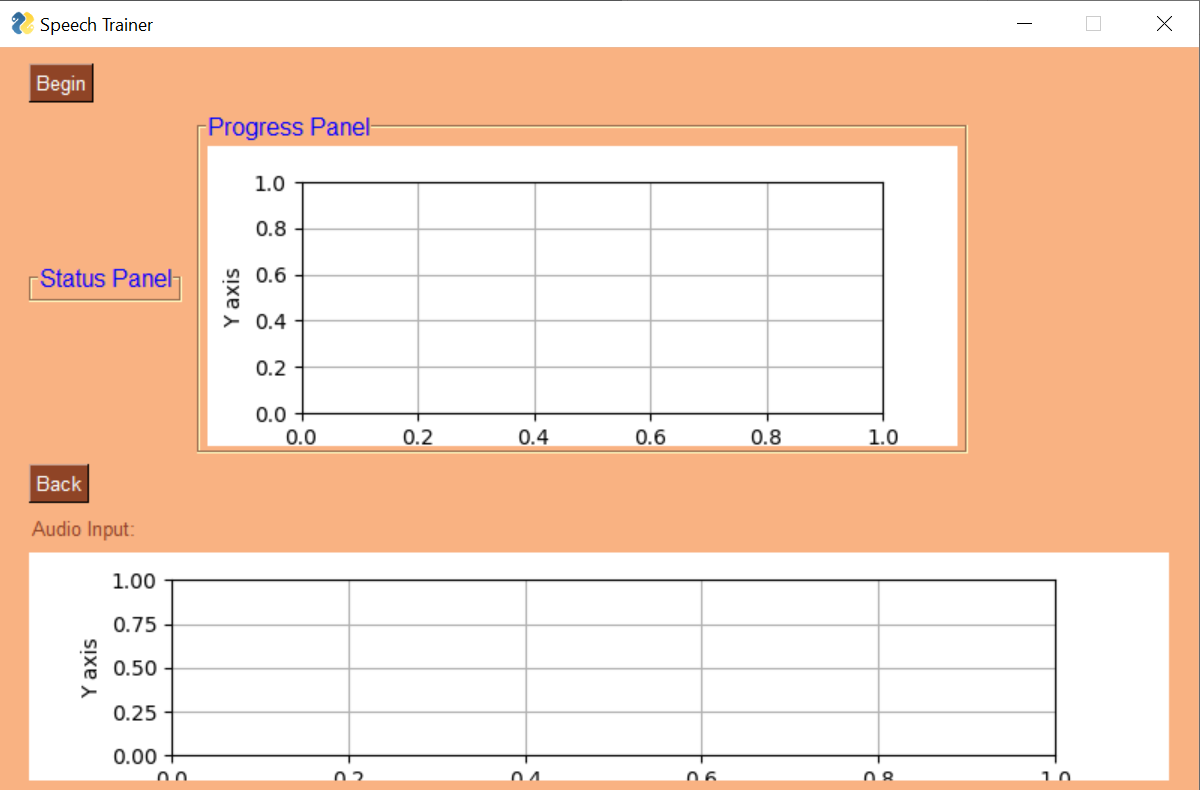
1. הצגת ציון נוכחי – את הציון חישבנו ע"י ................................... באמצעות חישוב הקשרים לינארים בין הציון הסופי למאפיין המתאים. זאת לפי DataSet ראיונות העבודה של אוניברסיטת MIT .
   * 0.6314 **+** 0.0168 \* **intensityMax = Score**

**UI**

**Choose a topic**



**Start training**



**During training**

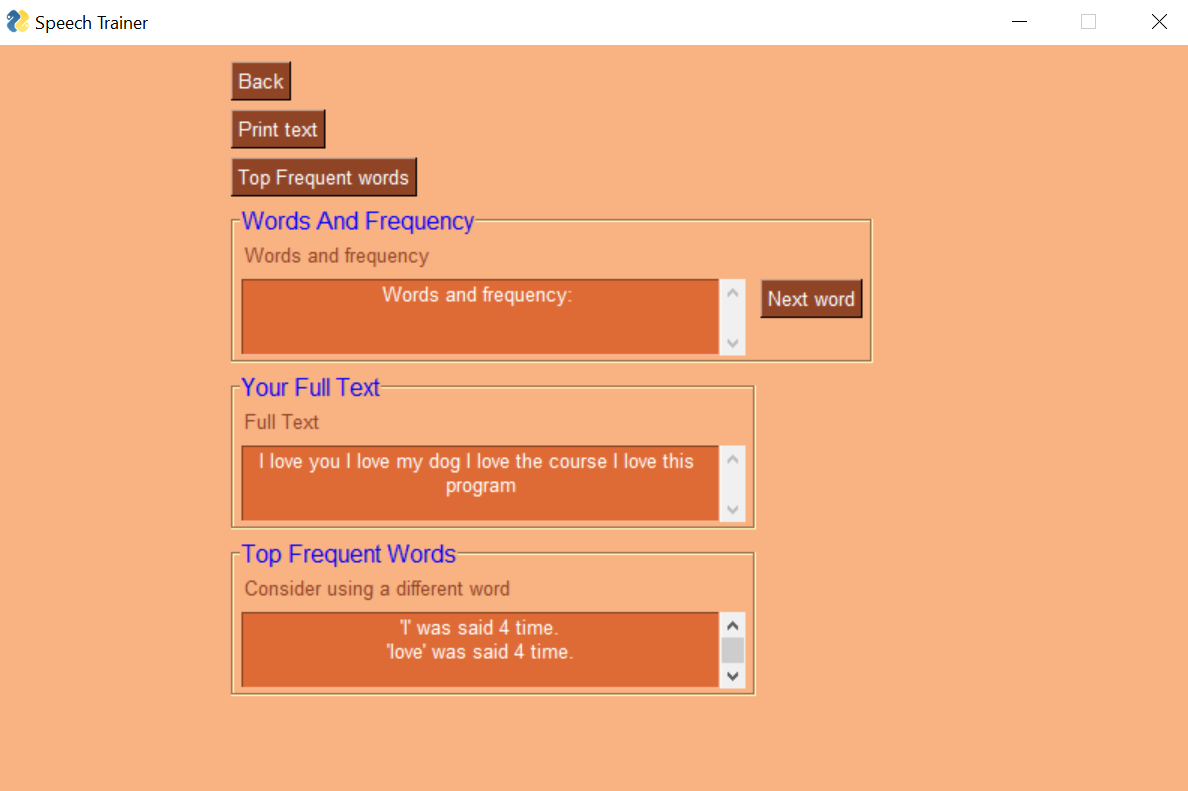


**End of training:**

The results are saved and displayed to the user.



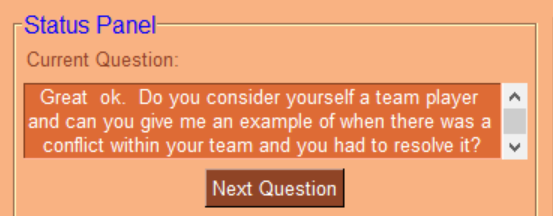
**Speech recognition – speech to text window**



**מודולים**

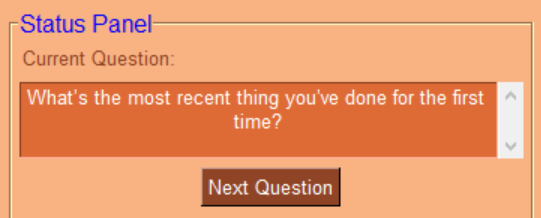
לשפר ולהסביר על הממודוליםםםםם

**Questions (by chosen topic) module**



Database of questions from job interviews:

We relied on transcripts of job interviews from a research experiment conducted by MIT University [1]. In total, we collected from the transcript of all the job interviews about 2,098 questions.



Database of dating questions:

We collected this database of questions while crawling a website: ...... a total of about 160 questions.

**Praat figure Module**



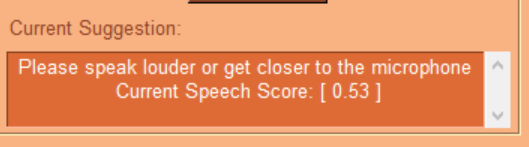
Shows the pitch and the intensity spectrograph in real-time.

**Graphical Progress Module**



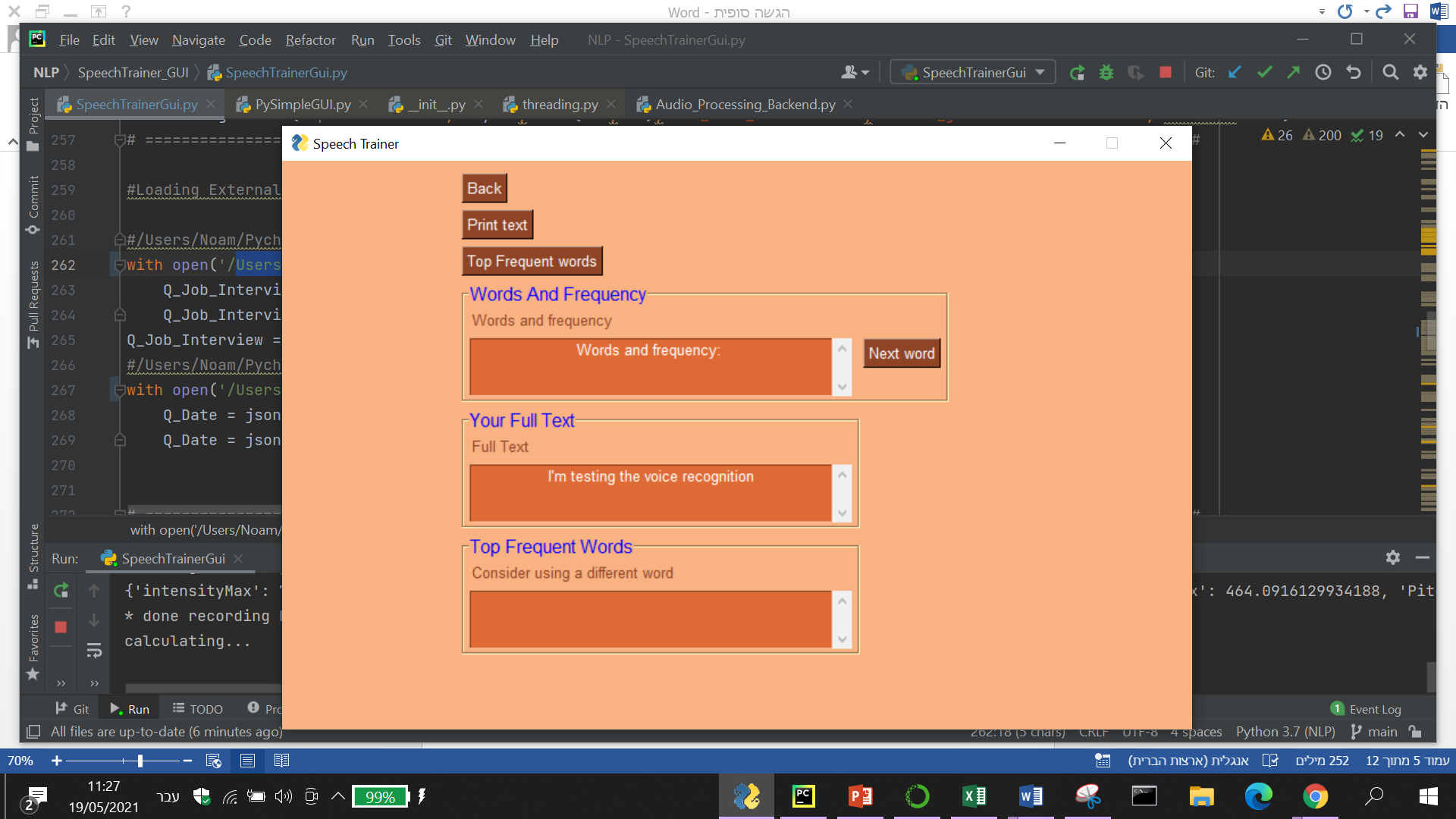
Shows the user's progress over time, by calculating the average scores.

**Advices or compliments Module**

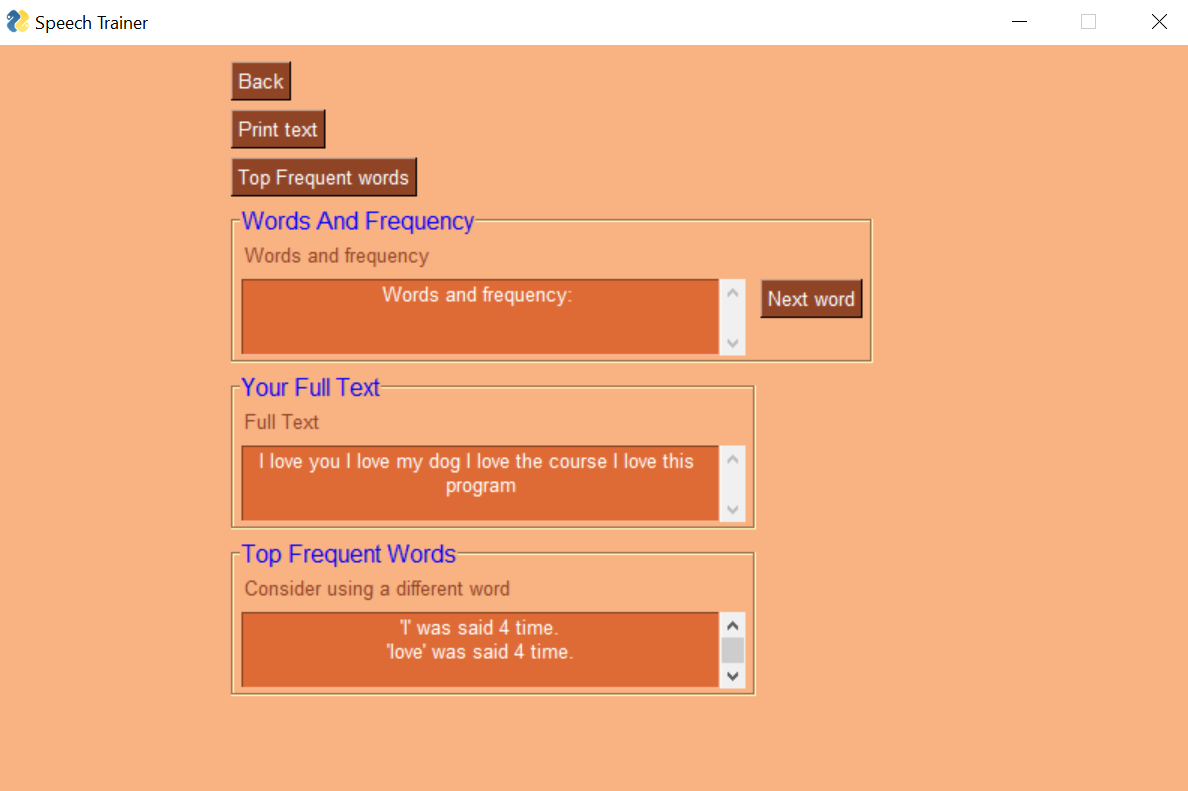


Informs if the user needs to improve some of the features. Otherwise, the program encourages him by generating a random compliment

**Speech recognition Module**

Prints the script spoken. If there are any word-repetitions, it presents them along with the amount of repetitions.

Case 1: Short sentence with no repetitions.



Case 2: Presents the repetitions.

**Scores**

**Functionality**

כאן צריכים לשרטט דיאגרמות של סימולציות -----> המשתמש מדבר בכעס המערכת מחשבת מדד מסוים -> המשתמש מעל -> מחזיר "תרגעעעעע גבר"

**NLP Tools**

כלים מעולם ה NLP

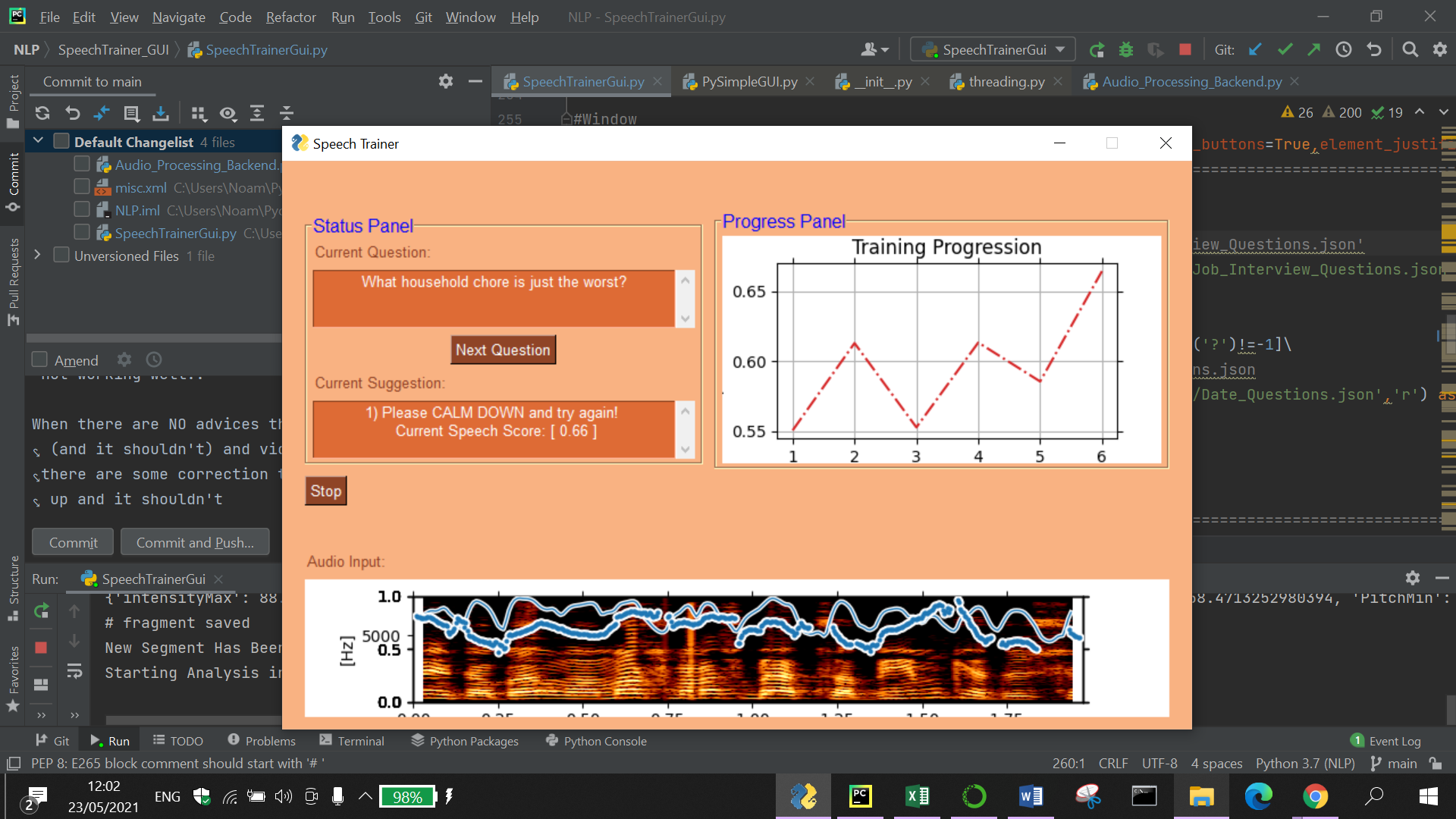
Speech Recognition

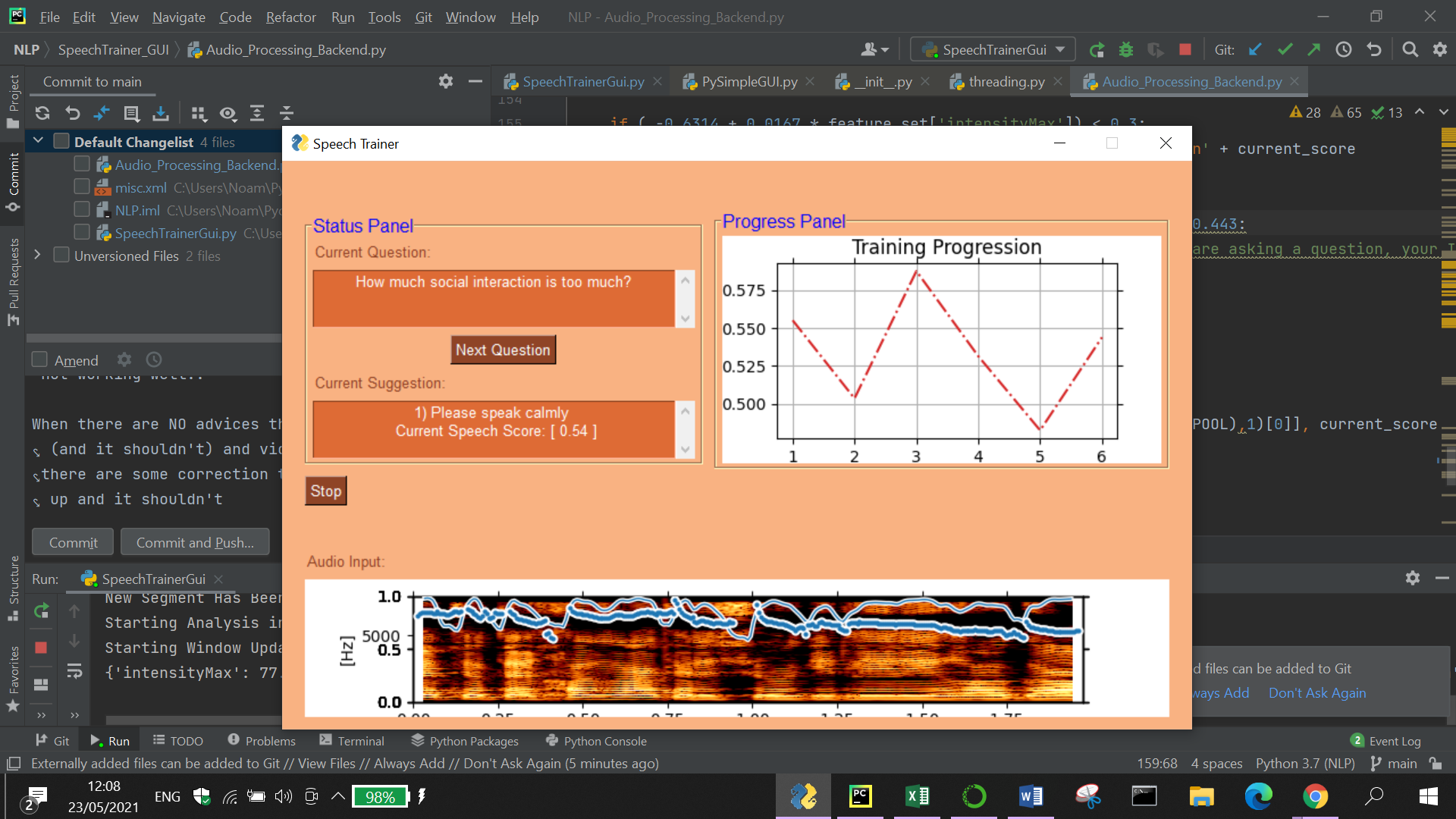
Speech to text

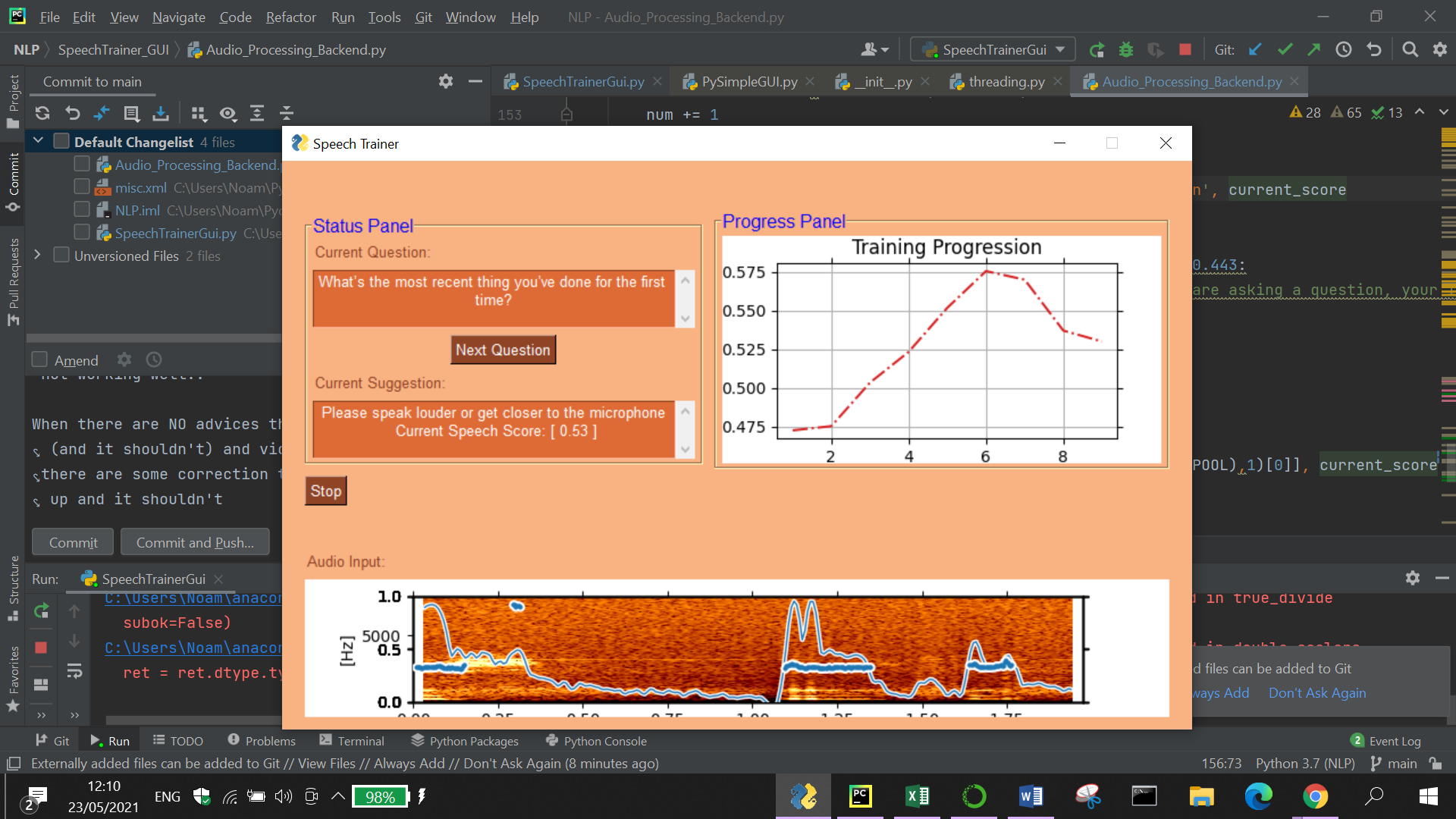
Praat FEATURES EXTRACTING + ANALYZING VIA PRAAT-PARSELMOUTH

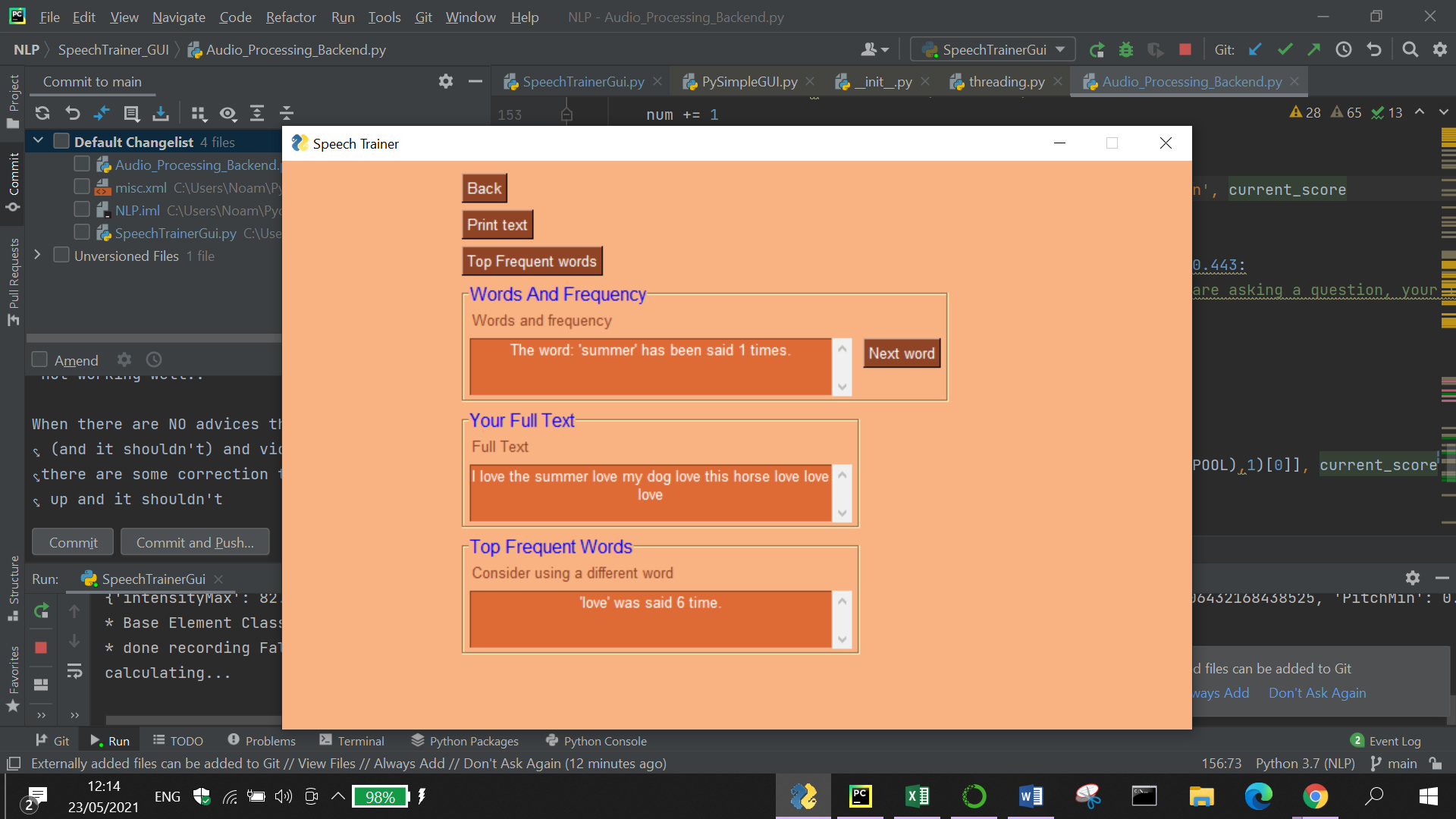
**Simulations**

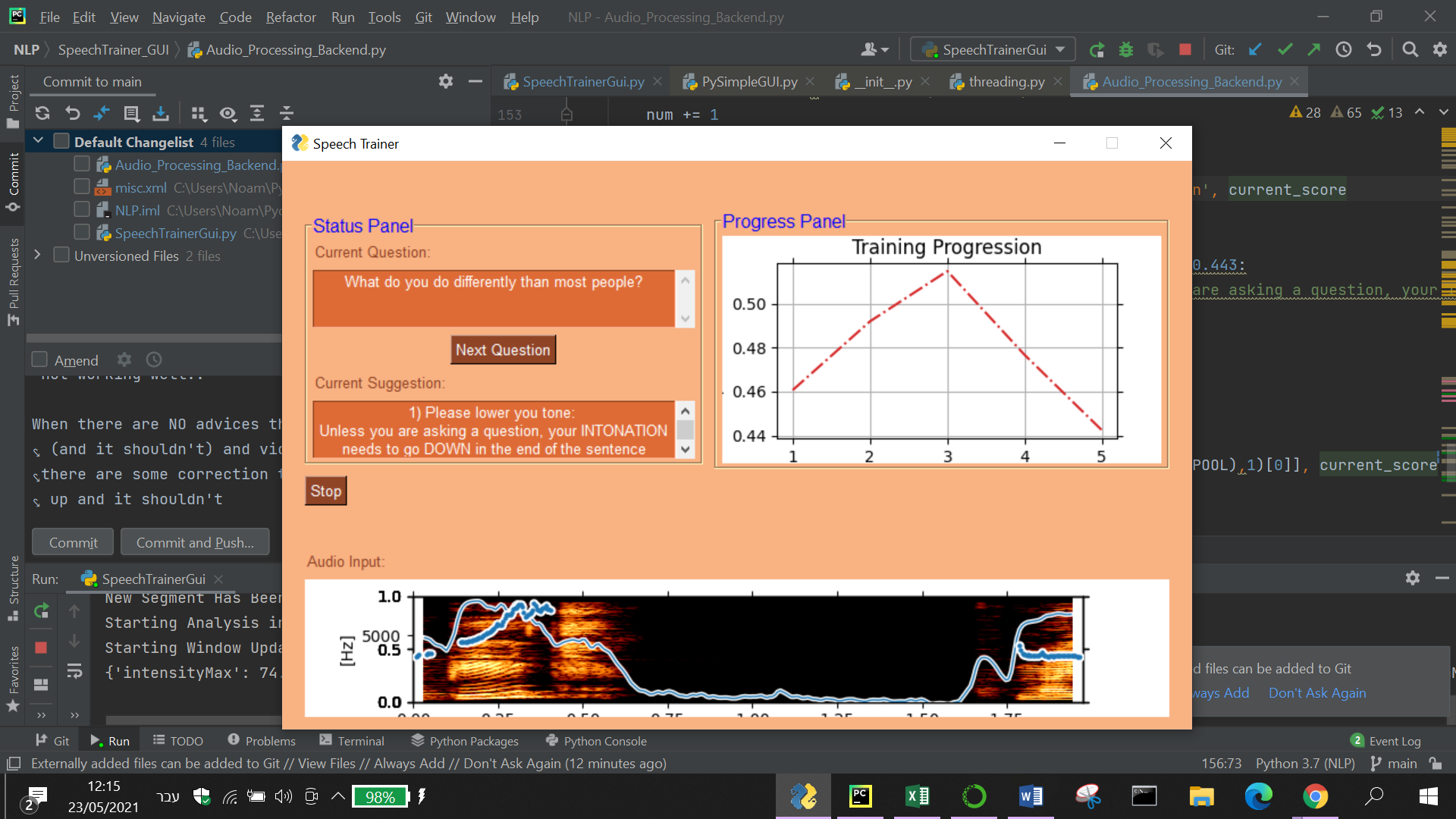
Very nervous speech



Restless speech:

Talking too quietly:

Repetitive words:

Rising intonation at the end of a sentence:

**Results and summary**

* **Did you answer the research question?**
* Indicate again your project topic and exemplify how you answered it.
* Summary and conclusions

**Bibliography and other sources**

[1]:I. Naim, M. I. Tanveer, D. Gildea and M. E. Hoque, "Automated prediction and analysis of job interview performance: The role of what you say and how you say it," 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2015, pp. 1-6, doi: 10.1109/FG.2015.7163127.

[2] H. S. Shim, S. Park, M. Chatterjee, S. Scherer, K. Sagae and L. Morency, "Acoustic and para-verbal indicators of persuasiveness in social multimedia," 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 2239-2243, doi: 10.1109/ICASSP.2015.7178369.

[3] Joshua J. Guyer1, Leandre R. Fabrigar1 and Thomas I. Vaughan-Johnston "Speech Rate, Intonation, and Pitch: Investigating the Bias and Cue Effects of Vocal Confidence on Persuasion" Personality and Social Psychology Bulletin · August 2018 DOI: 10.1177/0146167218787805

**הכוונות בנושא שיפור המדדים:**

[4] Magdin, M., T. Sulka, Júlia Tomanová and M. Vozár. “Voice Analysis Using PRAAT Software and Classification of User Emotional State.” Int. J. Interact. Multim. Artif. Intell. 5 (2019): 33-42.

<https://www.researchgate.net/publication/331881418_Voice_Analysis_Using_PRAAT_Software_and_Classification_of_User_Emotional_State>

[5] <https://lumen.instructure.com/courses/218897/pages/linkedtext54274>

**Appendix**