Speech Compression: Autotune Documentation
Summer 2019 Computational Science REU
Colleen Olson
Dr. Bill Robertson, Austin Wassenberg, Alex Kaszynski
July 15th, 2019

How to Use FinalAutotune.py

This program takes in an audio file and categorizes the frequencies into specific chosen frequencies. It can produce either a chord (hardcoded to specific frequencies) or it produces a changing tone. That changing tone is determined by what the user inputs.

Give the program a minute or two to boot up. You will see a black screen.

Inputting a .wav file:

Every input that the program takes needs to be in a specific form. It's case-sensitive, spaces will throw it off, etc.

Your .wav file needs to be recorded in mono, NOT stereo. It needs to have a sampling rate associated with it. Recording in Audacity works if you change it to mono. The sampling rate is usually 8,000 or 44,100 samples/second.

The .wav file needs to be in the same directory (folder) that the .exe is in.

If your .wav file is named Example.wav, type in 'Example.wav'.

FrameTime:

The program works by breaking the sound file up into small windows and analyzing those windows independently. A good frameTime is 0.01 seconds. Too long, such as 0.1 seconds, will produce choppy output. Too short, such as 0.005, and speech sounds will disappear.

Simply type '0.01'.

Naming your file:

The edited sound file will be saved to the same directory that the other files are saved to, and will be a .wav file also. Enter a name, such as 'name', and the program will automatically add the .wav suffix.

Chord vs. List:

This choice determines what frequencies are played and how they are played. The chord option is hardcoded, so it will only play a predetermined chord.

If you choose List, you will be asked how long you want the list to be. Then you will be asked what notes you want in your list.

List works by breaking the sound file into sections; each section corresponds to an element in the list. For example, if you have a list of length - [c,d,e,f] - the sound file will be broken into 4 even sections. For the first section, the frequencies will play as C's. For the second section, they will play as D's.

When entering which notes you want in your list, type a single note (lowercase), and hit enter. Then type the next note, and hit enter.

Note choices ('sh' means sharp):

a
ash
b
c
csh
d
dsh
e
f
fsh
g
gsh

Check your directory for the new file.

Tips on choosing Lists:

If you make your list too long with too many changes, they will go by very fast and the recognizable speech will decrease. A good number of notes would be about 2 per second.

If there are periods of silence in your audio file, especially at the beginning and end, the program will still divide up the sections per note equally. This means that if you want to play a tune, the first and last notes may not sound or will be very quiet. The volume is dependent on the original volume of the audio file, after all.