Tutorial for Manual Relative Fundamental Frequency (RFF) Estimation

Elizabeth S. Heller Murray & Jennifer M. Vojtech

Contents

A. Introduction

This tutorial was developed to assist individuals in using Praat for acoustical analysis. Praat is a freeware program developed by Boersma and Weenink (2013) for the purpose of phonetic research. Below are instructions for (i) downloading and navigating Praat, and (ii) using Praat to estimate fundamental frequency (f_o) and relative fundamental frequency (RFF) from an acoustic signal.

A.1. Downloading Praat

1. Navigate to http://www.fon.hum.uva.nl/praat/
2. Under the heading "**Download Praat**," select your operating system
   Note: This tutorial demonstrates using Praat on a Windows system
3. On the following page, select the bit edition of your operating system
   Note: On Windows, this is achieved by pressing the Windows key → search "system information" → System type
4. Open the zip folder and drag the Praat icon onto the desktop.
5. You are now equipped with the Praat program

A.2. Loading a File

- Upon opening Praat, two windows will appear: Praat Objects and Praat Picture
- We will mainly be using Praat Objects, so Praat Picture can be closed
- To **load a file**:
    1. Click **Open** from on top of the window menu
    2. Select **Read From File**
       *Note: If you choose **Open long sound file** to open your file, you will be unable to excise pieces of data in your sound clip*
    3. Select desired file and click **View and Edit**

A.3. Viewing a File

- After selecting **View & Edit**, a new window will appear with multiple sections
- If there are two channels to the audio signal, the upper two boxes will correspond to those channels (i.e., top box is channel 1, middle box is channel 2)
- If there is one channel to the audio signal, the upper box will correspond to this channel
- The bottom box contains the spectrogram for the audio signal

A.4. Using Praat

To select the portion of the sample that you would like to work with:

1. In the top box, click at the desired location to begin sectioning, and drag until the desired endpoint
2. The sample should now be highlighted in pink between two red bars showing either endpoint
3. Below is a table with a variety of commands that can be used to manipulate your selection:

Command	Feature
Tab	Start/stop playing the sample
Ctrl + n	Zoom to selection
Ctrl + i	Zoom in
Ctrl + o	Zoom out
Ctrl + x	Delete highlighted section
Ctrl + b	Zoom back
Ctrl + a	Show all
Esc	Interrupt playing

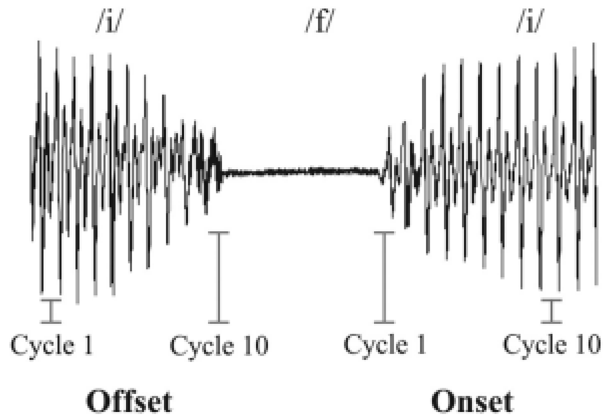
A.5 Online tutorials

Below is a link to an informative tutorial online, which can serve as a fine introduction to Praat: http://www.stanford.edu/dept/linguistics/corpora/material/PRAAT_workshop_manual_v421.pdf

B. Instructions for accurate f_0 detection and RFF estimation

B.1. Onset and Offset Cycles

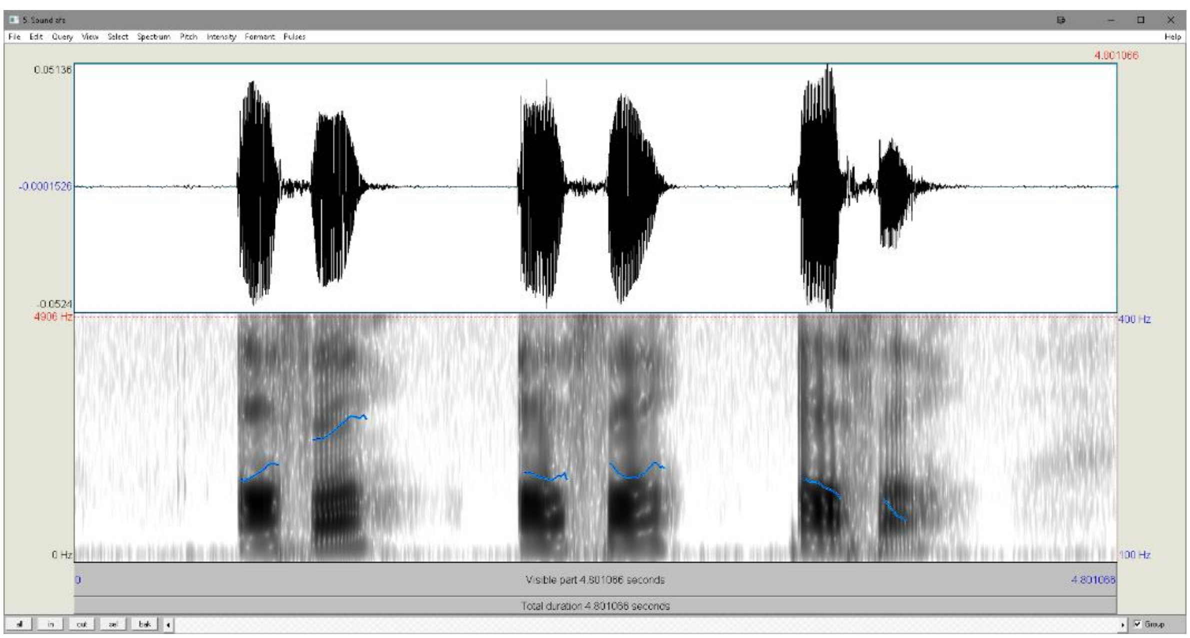
RFF is calculated using the **10 cycles prior to the voiceless consonant** ("offset cycles") and the **10 cycles following the voiceless consonant** ("onset cycles"). The figure below shows an example of RFF offset and RFF onset in the nonsense word /ifi/ (Heller Murray, Hands, Calabrese, & Stepp, 2015).



B.2. Manual Calculation of RFF

In this section, an acoustic sample containing three repetitions of the nonsense word /afa/ are examined.

1. Examine the acoustic waveform and its respective pitch contour, as in the figure below. Ensure that the sample follows suit of a voiced sonorant—voiceless consonant—voiced sonorant. Occasionally, there may be coarticulation, glottalization, or whispering in the sample, which could lead to inaccurate reporting of RFF values.

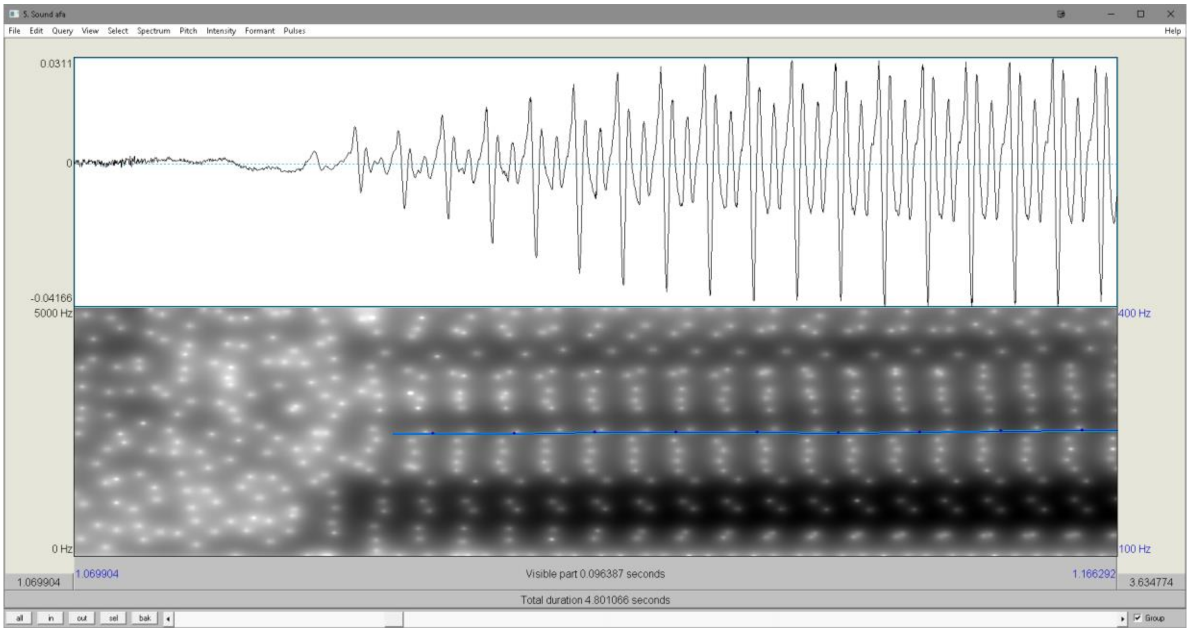


2. Listen to the file and determine if the voice is male or female; change the pitch range accordingly by selecting **Pitch → Pitch settings…** and entering the following starting range:
   - Female: 90–500Hz
   - Male: 60–300Hz

   Note: RFF calculation may be more accurate when this range is further optimized to the individual's voice

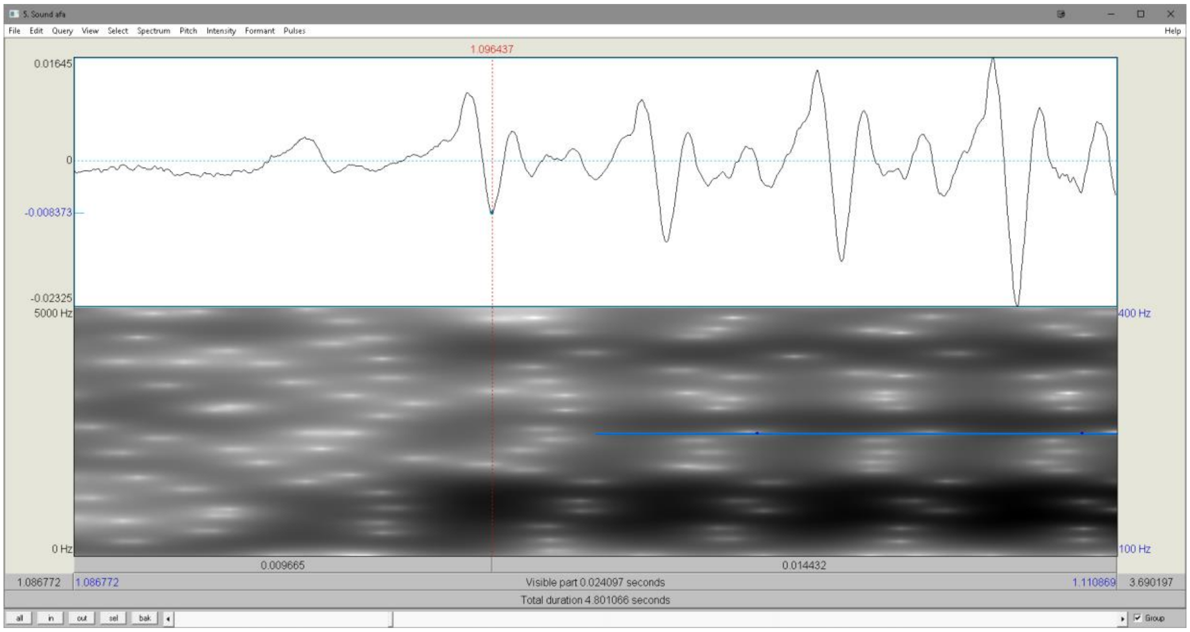
3. Navigate to **Pitch → Advanced Pitch settings…** and set to **Standards**

4. Zoom in on the section you want to analyze; it should look similar to the figure below
   Note: This is the vowel onset of the nonsense word /afa/.

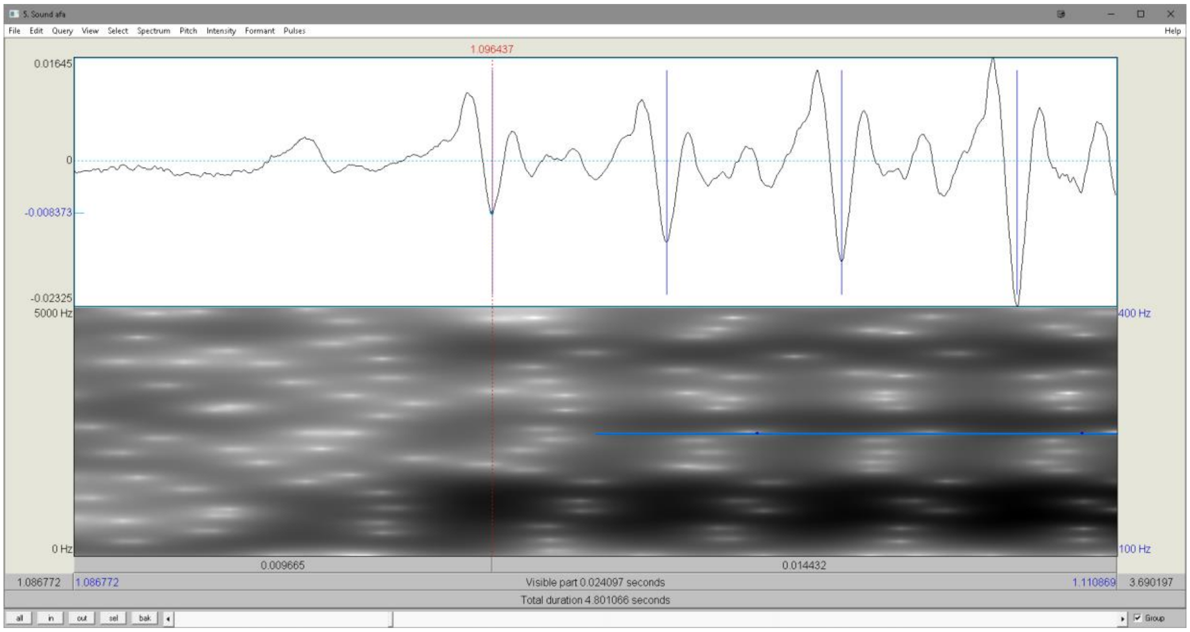


5. Ensure that the pitch contour is fairly smooth and regular—be wary of **large jumps in pitch**. While some large jumps in pitch make sense with the context of the sample (e.g., irregular vibrations of the sound /th/), Praat may produce a pitch contour that you do not agree with. Be sure to investigate any pitch jumps that occur in the acoustic sample prior to proceeding.

6. Once you agree with the pitch contour, zoom into the cycles closest to the consonant and place the cursor on the **peak or trough** of what you believe is the first cycle. Below is an example of selecting the first cycle of the vowel onset at a trough.



7. Once the cursor is marking the time point of the first cycle, leave it where it is and select **Pulses → Show Pulses** from the menu; pulse markers should appear. The pulses may or may not be aligned with the cursor (i.e., first cycle). The figure below shows a lucky example where the pulses are initially lined up with the cursor.



8. Calculation of RFF is dependent on the pulse timings of the 10 vocal cycles closest to the voiceless consonant (in this example, the /f/). Therefore, the next steps are to: (i) get the pulses to line up with the cursor, and (ii) achieve 10 visible vocal cycles (or 11 pulses) on the screen.

   - Leaving the cursor where it is, zoom out until 11 pulses are displayed on the screen, starting at where the cursor is.
   - Use trial and error to get the pulse to line up with your cursor. This can include moving forward or backward in time (with the bar at the bottom), zooming in or out, changing the advanced settings, or modifying the pitch range.

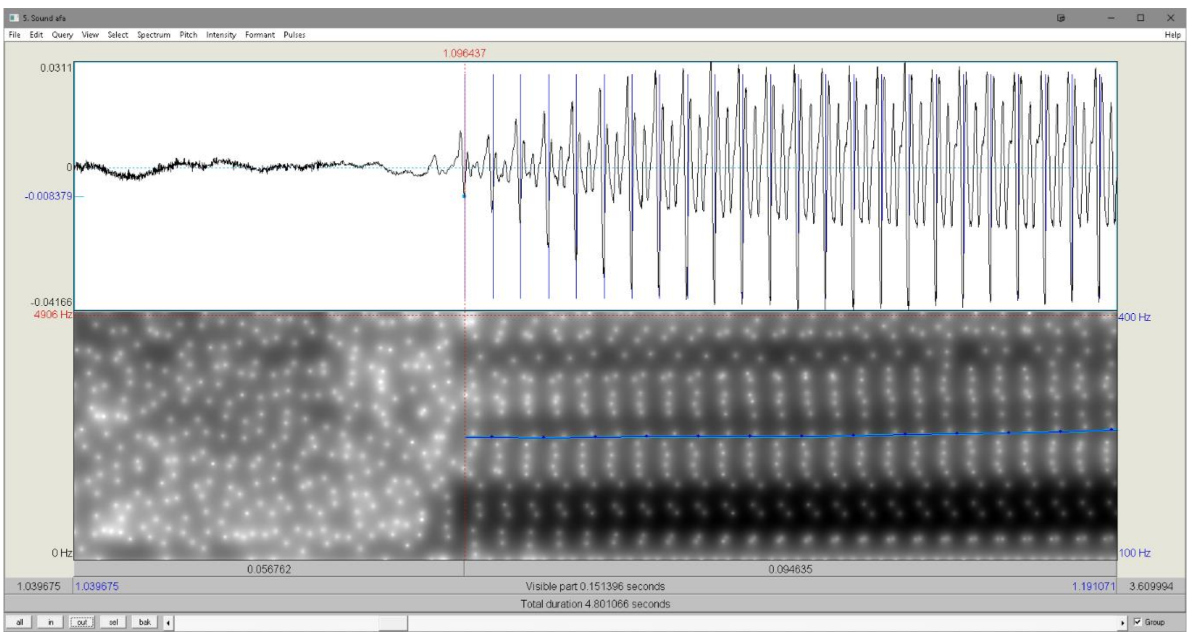
   It is suggested that you change just one advanced setting at a time to see what it does the pitch detection. The most common advanced settings that are helpful to change are the “silence threshold” and the “voicing threshold”.

   Here are descriptions of the functions contained in advanced settings:

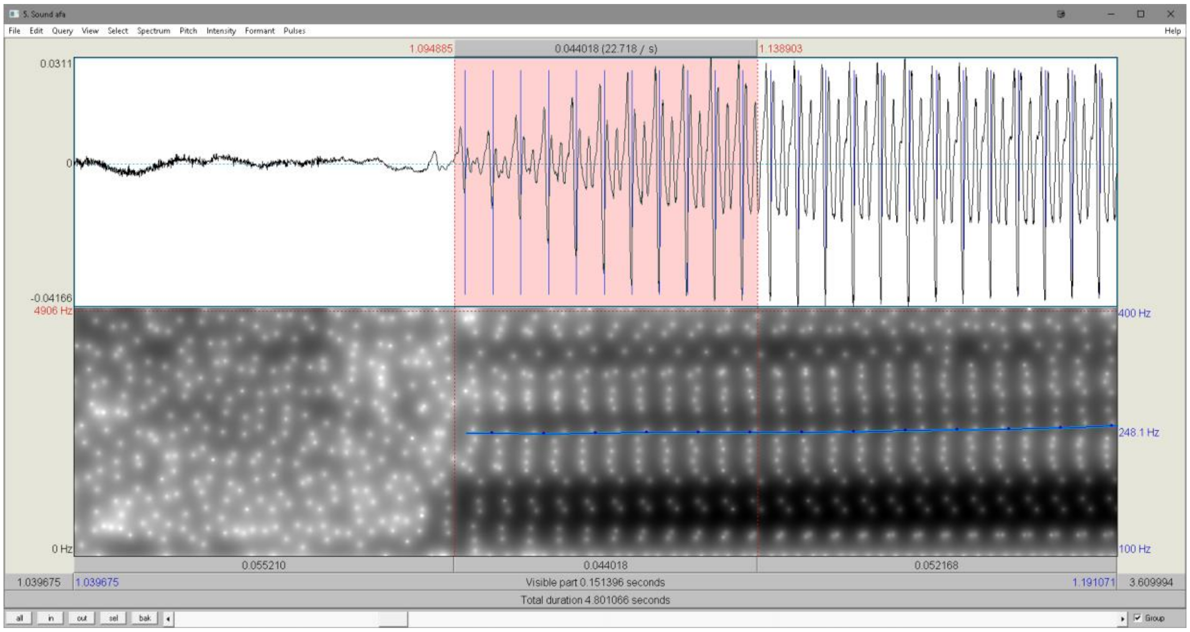
Function	Standard Value	Description
Pitch ceiling	600 Hz	Candidates above this frequency will be ignored
Silence threshold	0.03	Frames that do not contain amplitudes above this threshold (relative to the global maximum amplitude), are probably silent
Voicing threshold	0.45	The strength of the unvoiced candidate, relative to the maximum possible autocorrelation. To increase the number of unvoiced decisions, increase this value

Octave cost	0.01/octave	Degree of favoring of high-frequency candidates, relative to the maximum possible autocorrelation. To more strongly favor recruitment of high-frequency candidates, increase this value
Octave-jump cost	0.35	Degree of disfavoring of pitch changes, relative to the maximum possible autocorrelation. To decrease the number of large frequency jumps, increase this value
Voiced/unvoiced cost	0.14	Degree of disfavoring of voiced/unvoiced transitions, relative to the maximum possible autocorrelation. To decrease the number of voiced/unvoiced transitions, increase this value

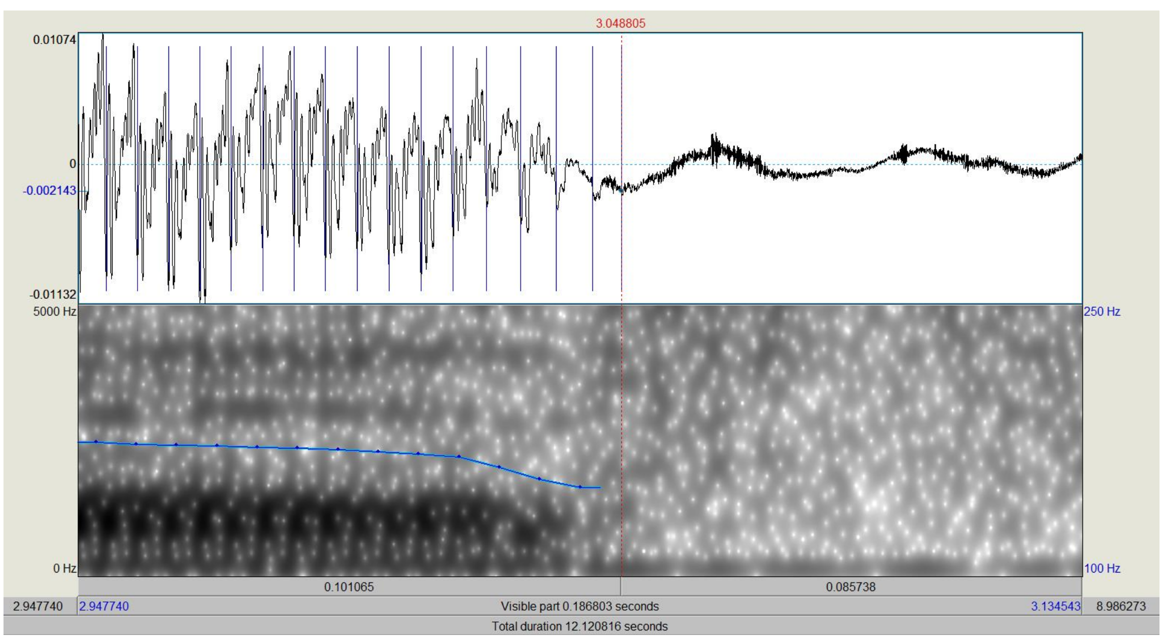
9. In this particular example, the window was zoomed out quite far in order to get the pulses to line up with the cursor (see figure below).
   Note: The cursor has not moved since deciding upon the time point of the first cycle. Additionally, this is the part that takes the longest. Remember, ***you are smarter than Praat****! If you feel the cycle starts in a certain place, don't waver on that – make Praat register to fit what you believe.*

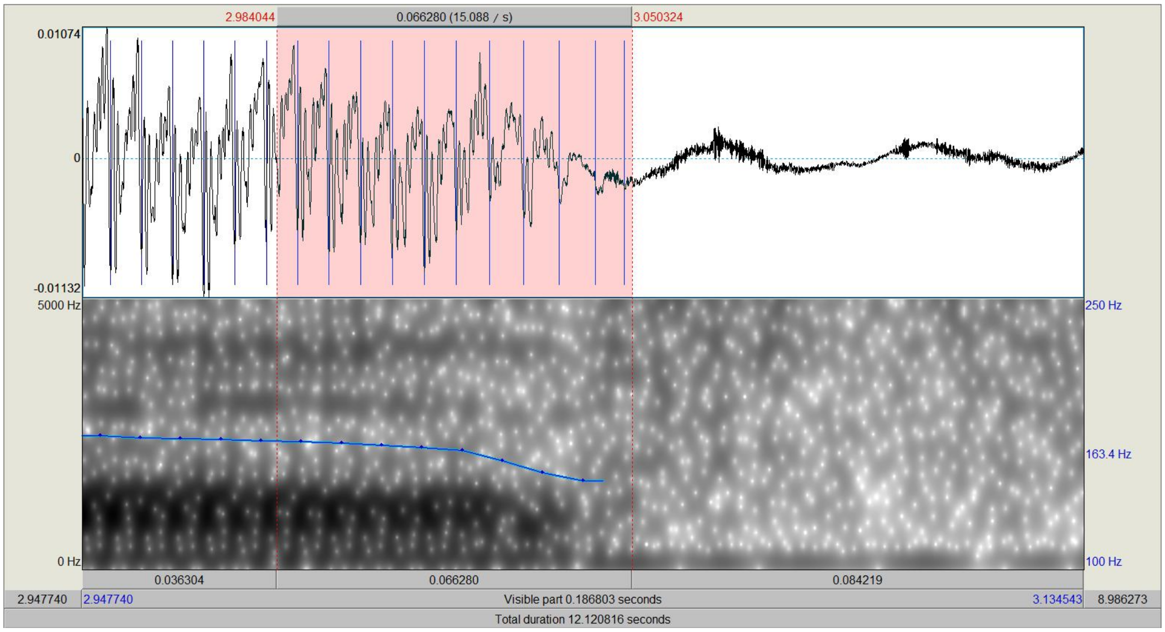


10. Once the pulses are lined up with the cursor and there are 10 vocal cycles (11 pulses) on the screen (starting with the pulse marked by the cursor), highlight the 11 pulses.



- For onset cycles (pictured above), the cursor is the beginning of onset cycle 1, so onset cycles 1–10 are selected, moving closer to the middle of the vowel.
- For offset cycles (pictured below), the cursor will be marking the end of offset cycle 10. Start with the pulse marked by the cursor and move backward 10 cycles, selecting offset cycles 10–1, moving closer to the middle of the vowel.





11. Select **Pulses → Pulse listing**; a window will pop up with 11 numbers, which correspond to the times of each of the 11 pulses.

Time (s)
1.096439158098
1.100457909330
1.104507586287
1.108555051804
1.112607138287
1.116640033262
1.120664949114
1.124685897184
1.128706828183
1.132722211761
1.136741280145

12. Now you have the start and end times for each cycles, find the difference between the start and end times to determine each period (10 cycles).

Period (s)
0.004289166
0.004523697
0.004559858
0.00454876
0.004554918
0.004559856
0.004544646
0.00454365
0.004554293
0.004518185

13. Next, convert each period into instantaneous frequency (1/period).

Frequency (Hz)
233.1455529
221.0581509
219.3050737
219.8401324
219.5429295
219.3051771
220.0391631
220.087398
219.5730434
221.3278289

14. Finally, convert the individual cycle frequencies into RFF values (in semitones) using the following formula:

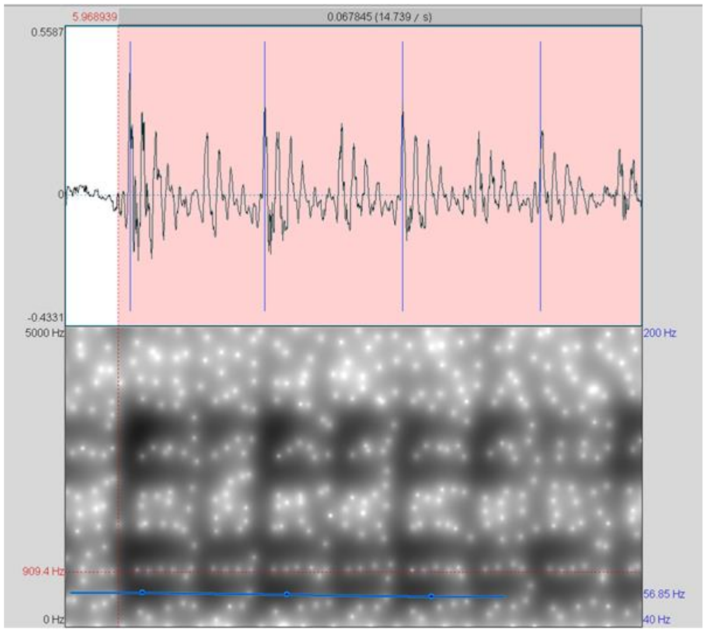
$$\text{RFF (ST)} = 39.86 * \log_{10}\left(\frac{f}{f_{ref}}\right)$$

The reference cycle, f_{ref}, in the example of vowel onset is the 10th onset cycle since it is further away from the fricative and closest to the center of the vowel. For offset cycles, the reference cycle would be the 1st offset cycle since it is furthest away from the fricative and closest to the center of the vowel. Remember, the RFF of the reference will always be 0, since it is comparing the reference frequency to the reference frequency.

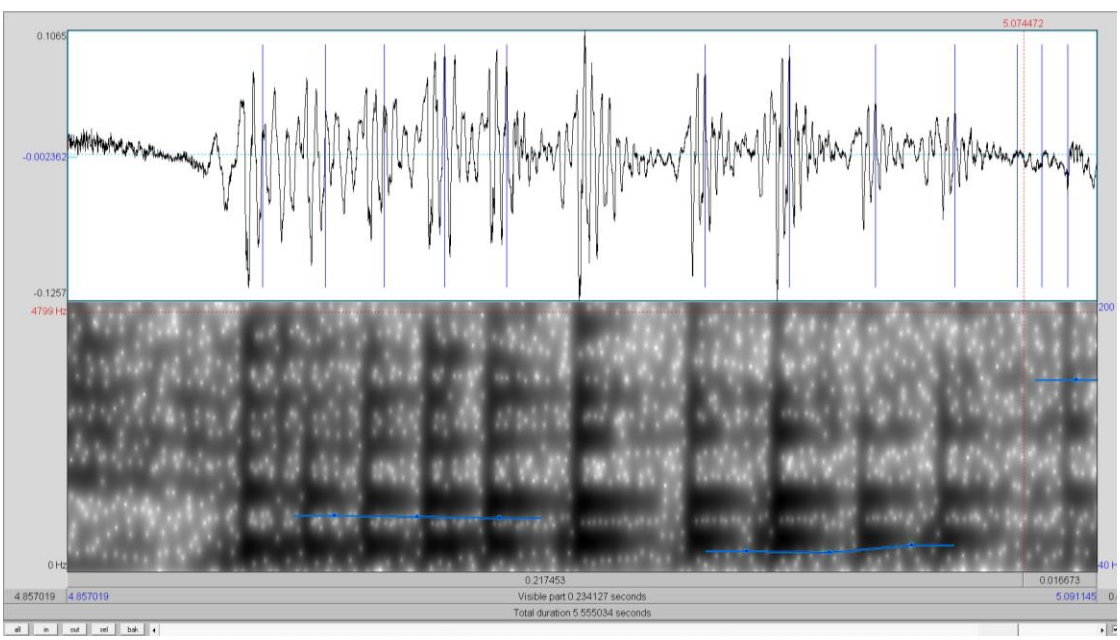
RFF (ST)
0.900552598
-0.021107174
-0.158948061
-0.116760998
-0.140181472
-0.158939902
-0.101094487
-0.097299858
-0.137806964
0

B.3. Troubleshooting and Rejecting RFF Values

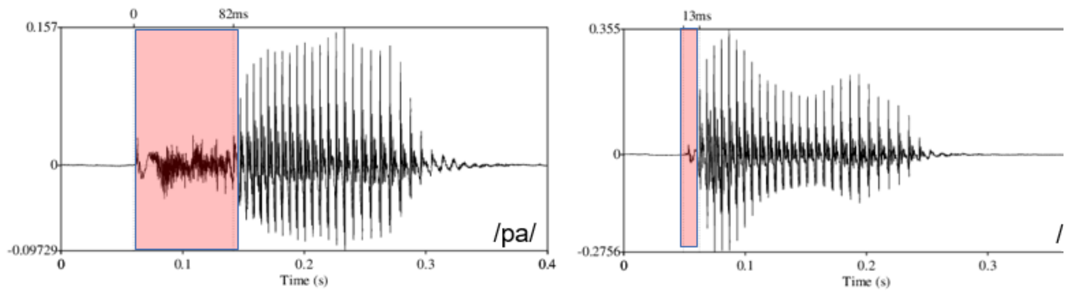
1. Praat may track the pulses of a sample in a way you do not agree with. Luckily, there are a few different options to change Praat's pitch detection method. For example, in the figure below, Praat seems to only detect every other cycle. The first thing you want to do is change your pitch settings. Listen to the voice and determine approximately what the pitch is, then reduce your pitch range. If that doesn't help, change some of the advanced pitch settings (see above for explanation).



2. Occasionally, an acoustic signal may be extremely irregular or glottalized. It is highly unlikely that an accurate RFF value can be estimated from such a signal. Here is an example where there is clear irregularity in the signal. The second half of the cycles are significantly longer than the rest of the sample. This is the end of the sentence and highly glottalized. You can tell because it looks like it skips a whole cycle, and the frequency changes drastically.



In general, it is best to reject RFF values based on the following criteria:

Glottalization	This can be determined according to the shape and periodicity of the waveform, how the speaker sounds, or the resulting RFF values. If the sample is glottalized, the resulting RFF values will make large, unexpected jumps from one cycle to the next. For instance, if the RFF values of three consecutive cycles were 0.08, 3.0, 0.5, then the corresponding acoustic signal is likely irregular or aperiodic in nature.
Voicing of the voiceless stop consonant	This can be determined by measuring the voice onset time (VOT) for voiceless stop consonants (i.e., /p/, /t/, /k/). VOT should be between 40–110 ms for voiceless stops, and -20–20 ms for voiced stops (i.e., /b/, /d/, /g/). See figure below for an example of /pa/ (left) vs. /ba/ (right). The VOT for /pa/ is 82 ms, while the VOT for /ba/ is 13 ms. Resulting RFF values should be checked for jumps of 1–2 ST between consecutive cycles, which can be indicative of stop consonant voicing. 
Failure to reach steady state	Calculation of RFF across each cycle depends on a reference frequency (f_{ref}) to compare the instantaneous frequency of the current cycle to. Failure of 10th offset or onset vocal cycle to reach steady state will produce erroneous RFF values. This can be observed as: (i) the 2nd value for offset RFF has a magnitude greater than 0.8 ST or (ii) the 9th value for onset RFF has a magnitude greater than 0.8 ST.
RFF value(s) are greater than 10 ST in magnitude	This represents a large, irregular jump in frequency between the reference cycle and the cycle of interest. RFF value(s) greater than 10 ST in magnitude are not physiologically possible.
Dichotic RFF values (e.g., -2, 1, -2, 1, …)	Quasiperiodic RFF values are indicative of glottal fry.
Less than 10 cycles	RFF must be calculated using the 10 transitionary cycles between the voiced sonorant and voiceless consonant (or vice versa).

C. References

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer (Version 5.3.51). Retrieved from http://www.praat.org

Heller Murray, E. S., Hands, G. L., Calabrese, C. R., & Stepp, C. E. (2015). Effects of Adventitious Acute Vocal Trauma: Relative Fundamental Frequency and Listener Perception. *Journal of Voice*.

Manual last updated on January 11th, 2018 by Jennifer M. Vojtech