Forum for Tools for Listening to Text in Performance

Please post here your questions and issues with Drift, Gentle and related analysis of your audio files. DO NOT delete any posts. Thanks!

Comments from Tanya:

1. Gentle is producing a transcript.txt file but nothing is in it.
2. The gentle interface doesn't show what file it is by name -- you have to guess from the text on the page. That's kind of frustrating when you have to leave the page to give it time to work and by the time you go back (if you're me) you've forgotten what you were working on.
3. It would be great if the files that we download from gentle can be automagically named according to the name of the file we upload. (I'm having to rename all the generically named downloaded files). The Drift downloads are named well.
4. It's taking drift2 a very long time to download the files. Sometimes the downloaded file just shows up on my system hours later but there is no way of knowing that drift2 is still working -- is there a way to get some sort of sense of the process as it's working? Kind of like how you can see when you are downloading a file how much of the download is completed. In other projects, we've asked the user for an email and then just emailed the files when it's done. I don't know if that's an option here.

When I click "Download csv" in drift2, I get an error message: "Failed -- Server problem." Thoughts? Other than that, drift and gentle are working GREAT with my super low-quality performance recordings from the seventies. (FYI, I'm working with a 34-second clip, which may be too long for the server to handle. Is that it?) Hi (it's Neil), will you email this issue to me and I will send it right to Robert? Not sure to whom I ought to direct the reply.

Comments from the PodcastRE crew

- Having a tough time with gentle with longer files. This is a known issue to you guys, I know, but just wanted to confirm that when we try to upload files that are longer than 10 minutes, we run into significant issues.
- We've tried to compare a number of files in drift, and the results are interesting. But like gentle, drift seems designed for short audio clips. We have some files (several minutes in length) that we have run through drift and would like to see the frequency plot across the entire file, but the zoom feature only lets us scale to about 30 seconds. It would be good to get a more macro view of the entire file.

Calculating Three Basic Prosodic Measures in Excel:
Pause Length, Average Pitch and Pitch Range

NOTE: Using the basic R scripts we already have, the Python code we're working on, or the existing Matlab approach, will generate these and other prosodic measures MUCH more efficiently and automatically, for more than one recording at a time. So if you have a lot of data to work with already, please don't waste your time calculating prosodic measures this way for all of your recordings. If you're eager to try to the other approaches, just let me (Marit) know.

However, it IS worth your while to get to know the data for 5-10 recordings intimately, so it can be useful to play around with calculating these three basic prosodic measures in Excel. And of course this is way to start generating interesting data if you aren't yet comfortable with Python, R/R Studio, or Matlab.

TO CALCULATE PAUSE LENGTH

1) Download a CSV from Gentle (which shows the start and end times of words in your recording).

2) Calculate pause length between the first two words by typing "=C2-D1" in cell E2.

3) Copy and paste that formula into E3.

4) Copy E2 and E3 as a block by selecting both cells and clicking on the lower right corner. Paste into E4. That should copy the formula for the whole column.

You can use the data about pause length to analyze when speakers make long pauses—for punctuation, for breath, for suspense, for line breaks, etc. For more detailed examples, see the 15 examples in "After Scansion: Visualizing, Deforming and Listening to Poetic Prosody" at https://arcade.stanford.edu/content/after-scansion-visualizing-deforming-and-listening-poetic-prosody.

TO CALCULATE AVERAGE PITCH

1) Download a CSV from Drift.

2) Select all of column B and click on "Sort and Filter" in the upper right corner, and select "Filter."

3) Select "0," "60," "120" and close the filter box. This will prevent zeros and erroneous values (60 and 120) from being included in your averages. Note: If you see any octave jumps in the pitch contour in Drift—a pitch value that is exactly double or exactly half the pitch value for the previous value—you would want to filter out that value as well. However, compared to most pitch-tracking software, Drift produces few octave jumps, since its pitch-tracking algorithm was trained on recordings that simulate low-quality radio. (Also note that it is unusual for a human

speaker to go much below 80-100 Hertz, though it does happen—a gravelly Robert Frost can get into the 70s and even 60s. Small children and biological females with especially high pitch voices can reach pitch values of 400 or more.)

4) Go to the end of column B. Type “=A” and “AVERAGE” should show up. Type in the appropriate cell range, beginning with the first row that includes a pitch value, e.g. “=AVERAGE(B4:B602).

This will give you a value, in linear Hertz, for the speaker’s average pitch in the recording. Keep in mind that linear hertz are not ideal for comparing male and female speakers, or for comparing speakers with a very low and very high average pitch. This is because the linear values for pitch range increase exponentially, e.g. a speaker whose pitch range in a recording
is 100-200 Hz is as wide as a speaker whose pitch range is 200-400 Hz.

TO CONVERT LINEAR HERTZ TO LOG HERTZ

1) Type “=LOG10(B4)” in D4, for example. This would convert a value of 403 Hz into 2.605… log Hz.

TO CALCULATE PITCH RANGE

1) Select column B.

2) Click “Sort and Filter” and click on “Filter” in the upper right corner.

4) Click “Smallest to Largest.” If it asks to “Expand the current selection” or “Continue with current selection,” choose “Continue with current selection.” Otherwise it will include all columns in the sort.

Your pitch range, assuming you have filtered out 0, 60 and 120, is simply the lowest value to the highest. For instance, if the first pitch value is in B4, and it is 146, and the highest value is 403, in the last cell in the B column, your speaker’s pitch range in this recording is 146-403 Hertz. Again, to compare pitch ranges between speakers with very different average pitch (e.g. very low and very high), it’s best to convert to log Hertz. You might also find that two speakers with very similar average pitch may have quite different pitch ranges, one narrower and one wider, for instance.

Hi, this is Gretchen. More of an Excel question, but I’d like to compare data of two speakers in one spreadsheet. Possible? Simple?

Prosodic Measures / Python Scripts

kwsherwood@gmail.com

I'm running Gentle/Drift installed on my Macbook.
My process was

1) Use Gentle to create transcription / timing (output -> align.csv)
2) Use Drift to execute pitch trace (output -> PoemName.wav-drift.csv)
3) Execute python script in form:
   `Python prosodic_measures.py align.csv PoemName.wav-drift.csv`

Intially I got the error: `No module named lempel_ziv_complexity`

Users will need to know to install this module:

```
$ python3 -m pip install lempel_ziv_complexity
```

Here is a successful execution:

```
$ python3 prosodic_measures.py align.csv CB-Empty.wav-drift.csv

Warning: numba.jit seems to not be available. Using a dummy decorator for numba.jit() ...
Warning: numba.jit seems to be disabled. Using a dummy decorator for numba.jit() ...

Twelve Prosodic Measures
------------------------
1. WPM: 108
2. Average pause length: 0.37 seconds
3. Average pause rate: 0.15 per second
4. Rhythmic complexity of pauses: 0.4790021041816561
5. Average pitch: 125.1194335937502 Hz
6. Pitch range: 1.3333672466603712 octaves
7. Pitch speed: -1.220756100208768 octaves per sec
8. Pitch acceleration: -1.8444939716438655 octaves per sec per sec
9. Pitch entropy: 3.4358787491500595
```

1. Question ? Is the "njubma.jit" error significant? Or can it be ignored? -
2. Excecution of the other Python scripts seems to fail, though it's not clear to me this is the same problem as in question 1

```
$ python3 prosodic_measures_remove_multiples.py align.csv
CB-Empty.wav-drift.csv
Warning: numba.jit seems to not be available. Using a dummy decorator
for numba.jit() ...
Warning: numba.jit seems to be disabled. Using a dummy decorator for
numba.jit() ...
Traceback (most recent call last):
  File "prosodic_measures_remove_multiples.py", line 29, in <module>
    for index, filename in enumerate(sorted(os.listdir(gentle_dir))):
NotADirectoryError: [Errno 20] Not a directory: 'align.csv'
$ python3 prosodic_measures_remove_integers.py align.csv
CB-Empty.wav-drift.csv
Warning: numba.jit seems to not be available. Using a dummy decorator
for numba.jit() ...
Warning: numba.jit seems to be disabled. Using a dummy decorator for
numba.jit() ...
Traceback (most recent call last):
  File "prosodic_measures_remove_integers.py", line 26, in <module>
    for index, filename in enumerate(sorted(os.listdir(gentle_dir))):
NotADirectoryError: [Errno 20] Not a directory: 'align.csv'
$ python3 prosodic_measures_remove_nothing.py align.csv
CB-Empty.wav-drift.csv
Warning: numba.jit seems to not be available. Using a dummy decorator
for numba.jit() ...
Warning: numba.jit seems to be disabled. Using a dummy decorator for
numba.jit() ...
Traceback (most recent call last):
  File "prosodic_measures_remove_nothing.py", line 26, in <module>
    for index, filename in enumerate(sorted(os.listdir(gentle_dir))):
NotADirectoryError: [Errno 20] Not a directory: 'align.csv'
```
