Some Novel Visualizations of Prosodic Variation

Prosody Visualization Challenge I: Speech Prosody 2018

Mark Liberman – University of Pennsylvania

Abstract

There are many possible ways to visualize prosody-related acoustic measurements. This poster explores two simple examples:

1. The joint distribution of delta f0 and delta amplitude.
2. A dipole plot of f0 differences as a function of time differences.

Because amplitude contours correspond approximately to syllabic "sonority", the relationship between f0 changes and amplitude changes tells us something about the phase relationships between f0 movements and syllable positions -- rises, falls, rise/falls, fall/rises, etc.

We can calculate the f0 differences versus time differences at various time scales, e.g. the scale of syllables and the scale of phrases. And the results show us the balance of f0 changes of different directions at different time scales.

In all plots, we use log measurements (e.g. semitones for f0 and dB for amplitude).

Methods

The input to the process is just a time function of f0 and amplitude measurements, sampled at 200 Hz (= 5 msec frame step). I've used get_f0a
http://languagelog.ldc.upenn.edu/myl/get_f0a.tgz
but any decent pitch tracker will be fine.

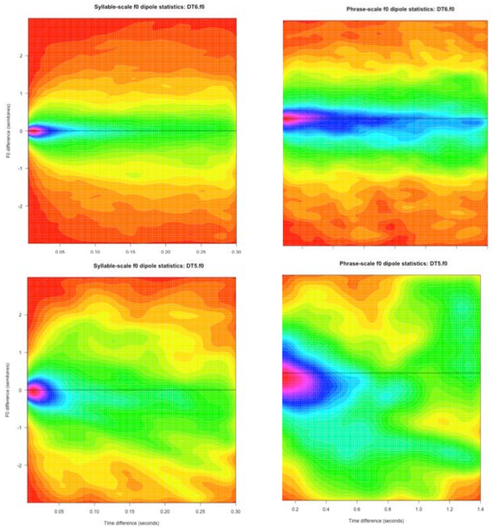
Then a couple of simple R scripts, one for each type of plot, do the rest of the work. Copies of these scripts can be found in
http://languagelog.ldc.upenn.edu/myl/DeltaF0DeltaAmp
and
http://languagelog.ldc.upenn.edu/myl/F0Dipole

These scripts assume an input file consisting of text lines, one per analysis frame, where the first field is the f0 estimate, the second field is indicates voicing (1 or 0), and the third field is RMS amplitude.

BELOW:

Top line: Syllable-scale and phrase-scale dipole plots for a Donald Trump rally, 10/2016

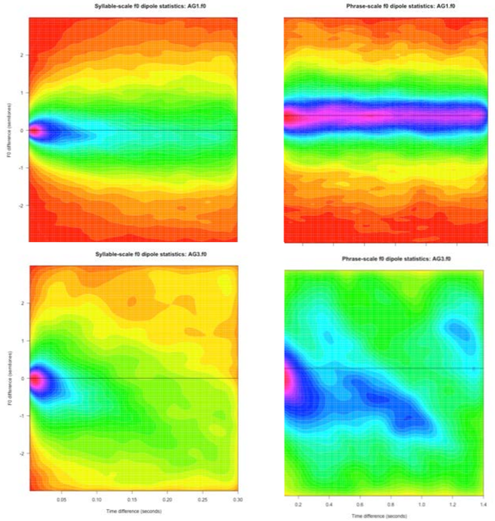
Bottom line: Syllable-scale and phrase-scale dipole plots for a Donald Trump interview, 12/2016



BELOW:

Top line: Syllable-scale and phrase-scale dipole plots for Allen Ginsberg reading *Howl*, 1956

Bottom line: Syllable-scale and phrase-scale dipole plots for Allen Ginsberg discussing *Howl*, 1956



Discussion:

In some but not all cases, these images evoke visually the acoustic impressions of the associated audio, and may thereby help us to understand the linguistic, stylistic, cultural or individual differences involved.

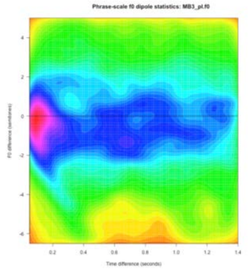
There are of course problems:

1. Pitch tracking often fails – and indeed the construct of "fundamental frequency" is almost as problematic as "formant".
2. Amplified, reverberant, and processed audio (e.g. studio-added AGC or other dynamic range compression) will show up in the delta-amplitude signal (as it does in acoustic perception).
3. There are many other features whose joint distributions are also relevant to our perceptions of prosody – various linear or nonlinear dimensionality reduction might yield more insightful pictures.

Some additional directions to explore:

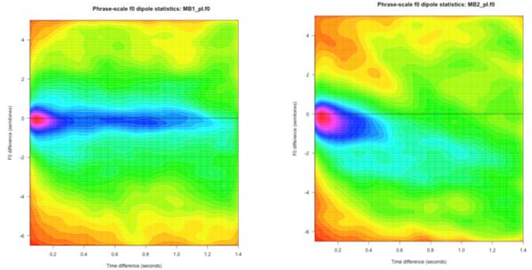
1. Animating the plots by moving a window through the input.
2. Attempting statistical analysis/classification based on such features.
3. Plotting reduced-dimensionality projection from larger feature set.

RIGHT: Miron Białoszewski commentary on "The Tempest"



BELOW:
Miron Białoszewski reading from "Aniela w miasteczku Foligno" (LEFT)

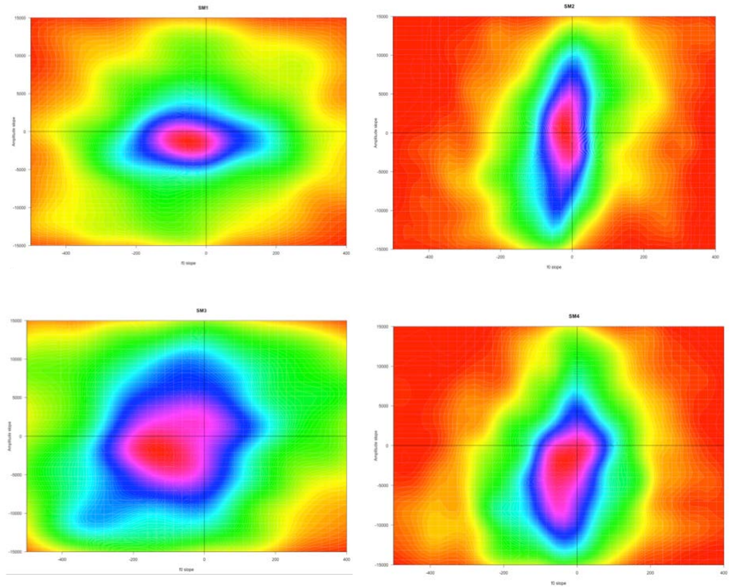
Miron Białoszewski reading from "Ach gdyby gdyby nawet piec zabrali" (RIGHT)



BELOW:

Top line: DeltaF0-DeltaAmp plot for M.L. King sermon (left)
DeltaF0-DeltaAmp plot for M.L. King interview (right)

Bottom line: DeltaF0-DeltaAmp plot for T.D. Jakes Sermon (left)
DeltaF0-DeltaAmp plot for T.D.Jakes interview (right)



References

"More on pitch and time intervals in speech", 10/15/2006, http://itre.cis.upenn.edu/~myl/languagelog/archives/003677.html
"Political sound and silence", 2/8/2016, http://languagelog.ldc.upenn.edu/?p=28950
"Poetic sound and silence", 2/12/2016, http://languagelog.ldc.upenn.edu/nll/?p=2404
"Some visualizations of prosody", 10/23/2016, http://languagelog.ldc.upenn.edu/nll/?p=28974
"Overall F0 trends at syllable and phrase scale", Spring 2016, http://www.ling.upenn.edu/courses/Spring_2016/ling521/F02.html
"Tunes, political and geographical", 2/2/2017, http://languagelog.ldc.upenn.edu/?p=30614