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 syllable “sonority”, the relationship between F0 changes and amplitude changes tells us something about the phase relationships between F0 movements and syllable positions – rises, falls, rises/falls, falls/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). For example, for F0:

http://www.langlog.upenn.edu/myl/get_f0a.tgz

This script assumes an input file consisting of text lines, one per analysis frame, where the first field is the F0 estimate, the second field is an index of 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

Below:

Mirror Białoszewski commenting on “The Tempest”

Below:

Mirror Białoszewski reading from “Ach gdyby gdyby nawet piec zabrali” (RIGHT)
Mirror Białoszewski reading from “Aniela w miasteczku Foligno” (LEFT)

Discussion:

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.

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

Mark Liberman – University of Pennsylvania

Some Novel Visualizations of Prosodic Variation
Prosody Visualization Challenge I: Speech Prosody 2018

PowerPoint® software.

www.genigraphics.com

Email: info@genigraphics.com

US and Canada: 1-800-790-4001

Next business day within the US and Canada.

Once your poster file is ready, visit

www.genigraphics.com

to order a high-quality, affordable poster print.