Haskins Laboratories

WORKSHOP
Jul-24, 2020

Formant Analysis with Big Data

Wei-rong Chen

Research scientist
Haskins Laboratories

Scope

- Best practice in formant analysis with large corpus.
- Use MATLAB to run PRAAT script for formant analysis.

Requirements:

- Computer software: MATLAB (R2015b or newer) , PRAAT (v6.xx or newer)
- Basic knowledge of formant, LPC, MATLAB & PRAAT

Formant analysis: LPC

- LPC (linear predictive coding):
- Most widely used algorithm for formant.

Predict the next sample by N past samples.

All-pole model of speech coding.

LPC in PRAAT: Burg method (Burg, 1967)

Burg, J.P. Maximum Entropy Spectral Analysis.
37th Annual International Meeting, SEG, Oklahoma, USA, Oct-31, 1967.

Limitations of LPC

- Since LPC is all-pole model, it cannot measure anti-resonances (zeros) in nasals or liquids.
- LPC measured formants are biased by F0. (Klatt 1986; Shadle et al 2016)

  The higher the F0, the more biased.

When F0 > 200 Hz:

- In **BEST** condition, the F0-biased formant error is at least 40 Hz. (Chen et al, 2019)
- **Formant variability** measured by LPC can be **completed replaced by F0 variability**.

Chen, W.-R., Whalen, D. H., & Shadle, C. H. (2019). F0-induced formant measurement errors result in biased variabilities. *JASA, 145*(5), EL360-EL366.
Klatt, D. H. (1986). Representation of the first formant in speech recognition and LF models of the auditory periphery. 12th ICA.
Shadle, C. H., Nam, H., & Whalen, D. H. (2016). Comparing measurement errors for formants in synthetic and natural vowels. JASA, 139(2), 713-727.

Data organization

- Two kinds of acoustic data:

1. Large sound file with multiple tokens.
2. Small sound file with single token.

If you have large sound files with multiple tokens, you may want to break them down to small sound files with single token.

Data organization - naming

- **GOOD practice:** use "**unique**" indexing letters and numbers

  e.g., "S01_B01_R0001.wav" (subject S01, block B01, token R0001)

  "RS0001.wav"

And you keep a spread sheet of **wordlist** storing the meaning of these indices.

- **BAD practice:** use the stimulus as file name

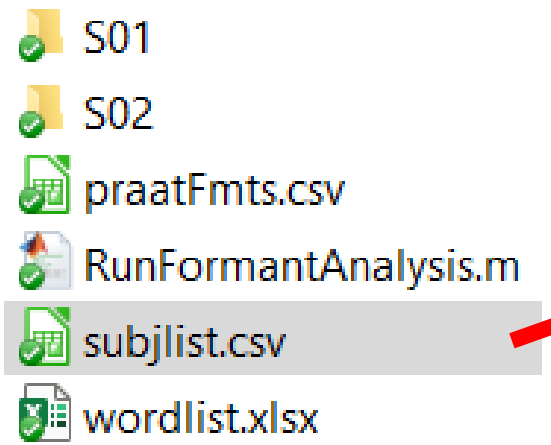
  e.g., "flower.wav"

  "Token accented.wav"

NEVER use space or any special characters in file & folder names!!!
(Only use alphabetic **letters** & **numbers** for naming!!!)

Data organization - structure

Organize data by subjects (speakers):
- Data of each subject are stored in each separate folder.
- Create a text (.csv) file listing all subjects and optionally the gender.



	A	B
1	Subject	Gender
2	S01	M
3	S02	M
4		

Preprocess – convert stereo to mono

- If your acoustic files were recorded in stereo format (2 channels), you need to convert them all to mono.
- PRAAT script: “mono_converter.praat” by Dr. Shigeto Kawahara

http://user.keio.ac.jp/~kawahara/scripts/mono_converter.praat

- Place the script in the folder containing .wav files
- Click and run the script
- The original .wav files will be copied to a newly created ‘./original/’ folder
- The converted mono .wav files will be stored in the folder.

Formant analysis in PRAAT

- Two steps:

1. **LPC Formant estimation:**

Estimates formants for each short time window (e.g., 20~50ms) stepping through the whole vowel duration. (formants jumps up and down)

2. **Formant tracking:**

Track continuous formant trajectories (penalize formant jumps)

LPC parameters in PRAAT

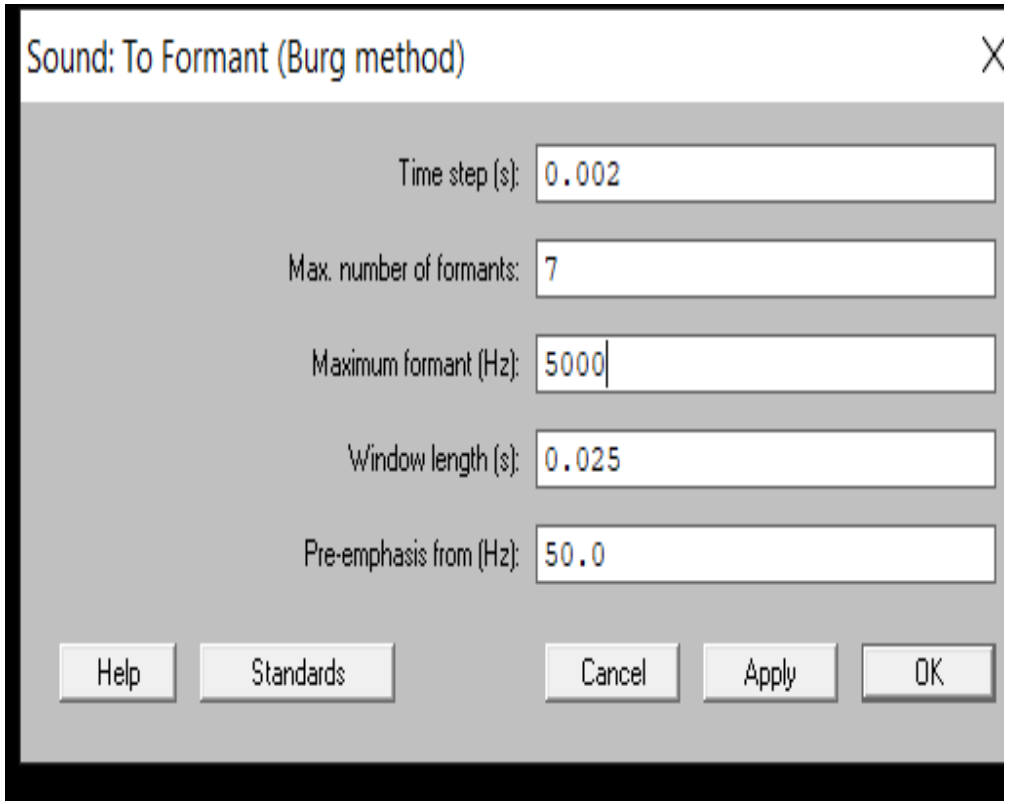
- Time step (s):
  Time interval between two consecutive windows.
  I usually use 0.002s (2ms). Shorter step takes longer time in calculation.

- Max. number of formants:
  = LPC orders /2
  We usually add 2 more than theoretical maximal number of formants within the cut-off frequency. (=7 if we cut-off at 5000 Hz)

- Maximum formant (Hz):
  = cut-off frequency (Nyquist frequency)

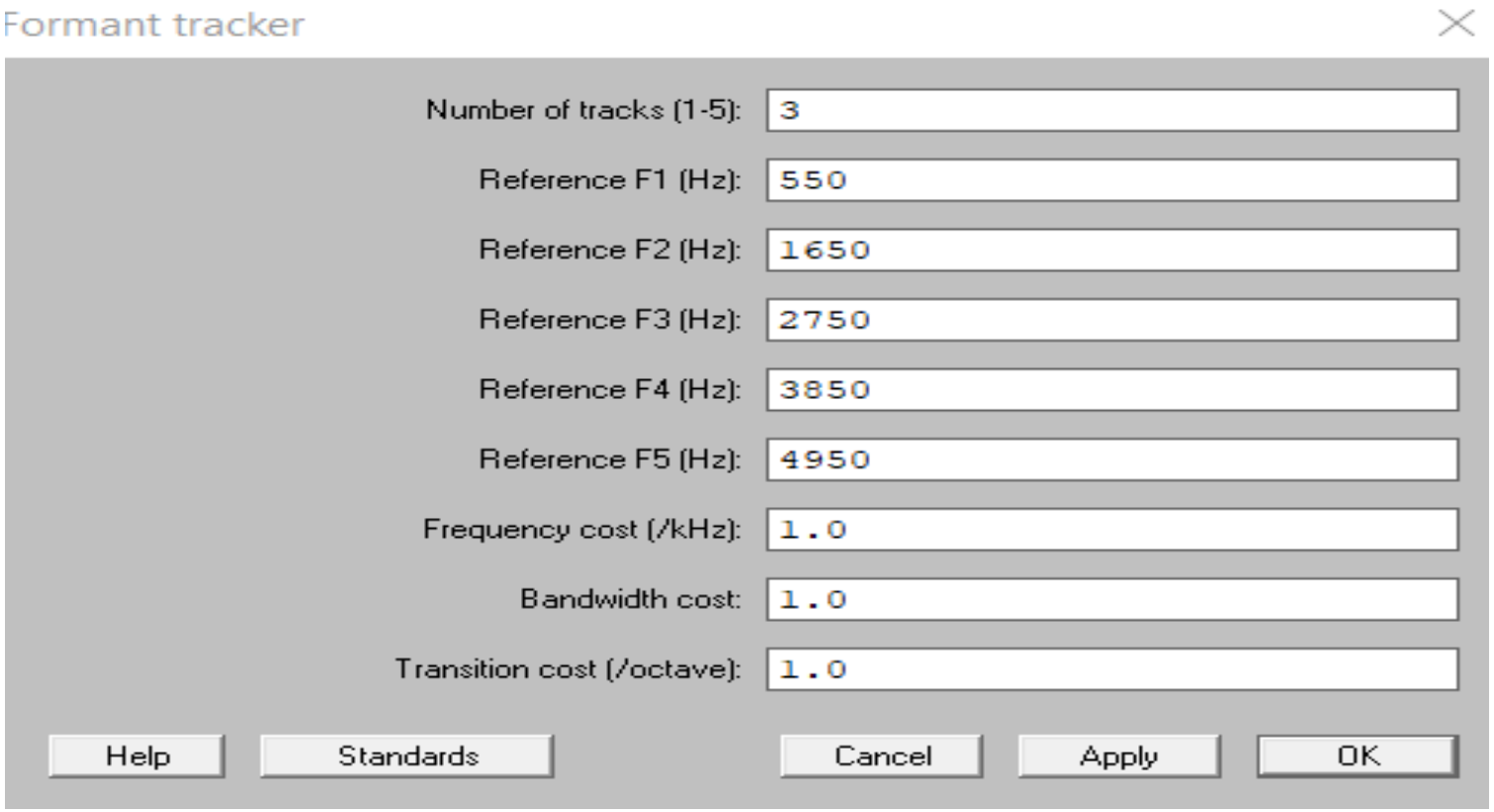
- Window length (s):
  Actually this is HALF window length.
  If you specify 0.025, the window size is 50ms.

- Pre-emphasis (Hz);
  Reducing the amplitude of ambient noise.



Formant tracking in PRAAT

Viterbi algorithm penalizing formant track drift away from formant references.
- NOT ideal for some vowels!



Seeding method of formant analysis

- Seeding method:

  Specify different formant references (seeds) for different vowels in tracking according to known information. (Chen, Whalen & Shadle, 2019).

- For example, we can use the published formant references reported in Peterson & Barney (1952) for each individual vowel.

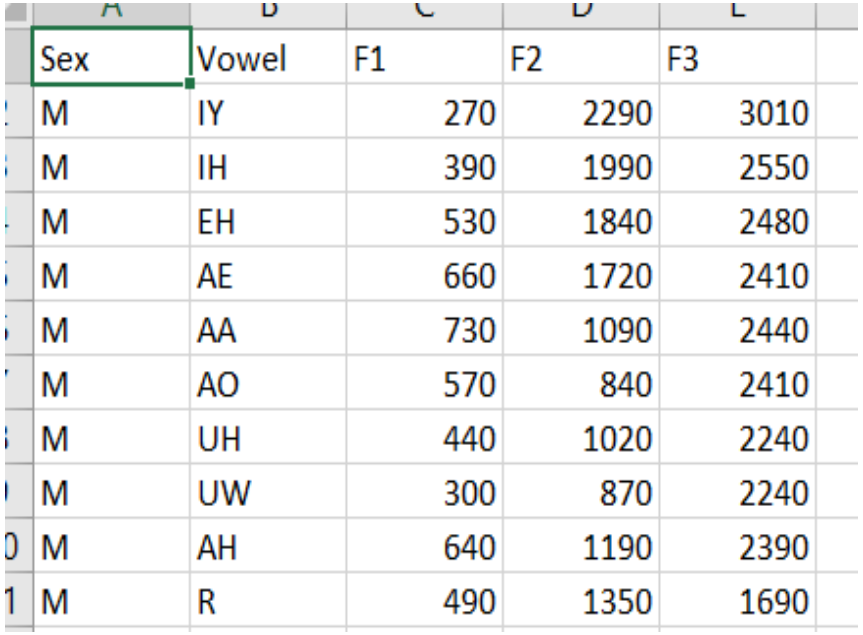
Chen, W.-R., Whalen, D. H., & Shadle, C. H. (2019).
F0-induced formant measurement errors result in biased variabilities.
JASA, 145(5), EL360-EL366.

How to use seeding method

1. Construct a text (.csv) file of formant references.

2. Create a skip-list for labels you don't want to analyze.

3. Run the provided codes.



	Sex	Vowel	F1	F2	F3
	M	IY	270	2290	3010
	M	IH	390	1990	2550
	M	EH	530	1840	2480
	M	AE	660	1720	2410
	M	AA	730	1090	2440
	M	AO	570	840	2410
	M	UH	440	1020	2240
	M	UW	300	870	2240
0	M	AH	640	1190	2390
1	M	R	490	1350	1690