Quantifying Diphthongs

A statistical technique for distinguishing formant contours

Adam Baker

University of Arizona

November 12, 2006

Outline

1. Importance of formant contours.
2. Ways of analyzing contours.
3. Smoothing Spline ANOVA
4. Examples
   - mostly about short a in English
5. Open issues (and possible solutions)

- There's code for everything (URL at the end of the presentation)
- Paul de Decker & Jennifer Nycz

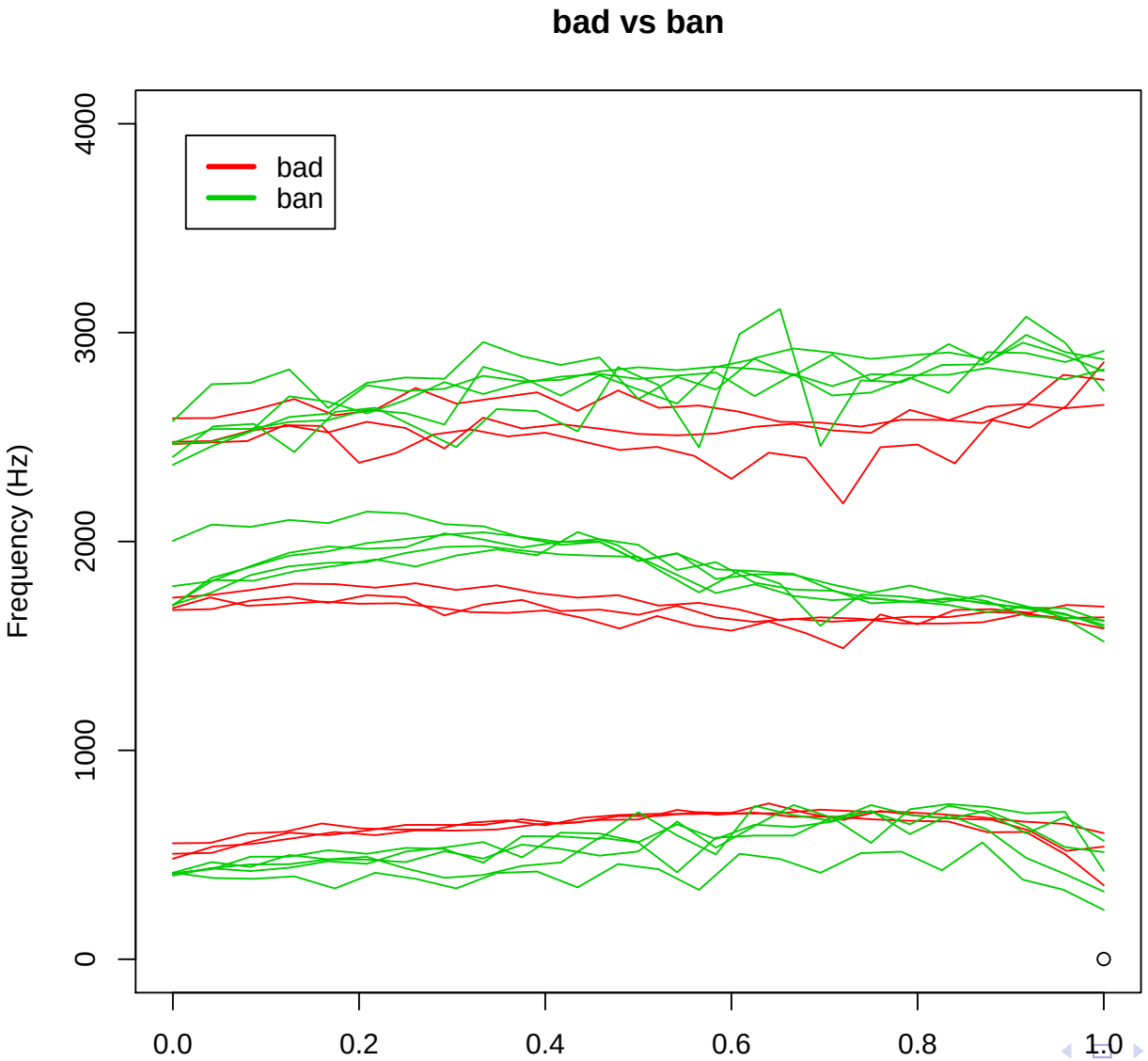
Role of formant transitions

- Diphthongs
  - Social significance: "tense æ" — transcribed [eːə] for Philadelphia English
  - Other pronunciations in other dialects, NOT [æ̃]
- We know that vowel formant transitions are important for perception
  - Silent-center vowel experiments (Strange et al. 1983, *et seq.*) show that the formant transitions of a vowel are sufficient for identification.
  - Parametric synthesis — meeting a single acoustic target isn't everything.
  - Gated perception experiments: the fastest-changing parts of a speech signal contain the most lexical information (Furui 1986, Warner 1998).

Previous representations of contours

- Conventionally vowels are characterized by point measurements.
  - mid-point
  - formant trajectory extrema
- This is a huge loss of data—we lose the time dimension.
- It's possible to investigate formant contours by comparing overlay displays, but this does not have statistical rigor.

ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



Previous representations of contours

- Vowel formant contours can be represented with a discrete cosine transform coefficients.
  - breaks the curve into three (Zahorian & Jagharghi 1993) or six (Watson & Harrington 1999) parameters.
- This keeps dynamic information, but the numbers are not readily interpretable.
- Handy for machine learning, but not really helpful for description.

Role of curves in general

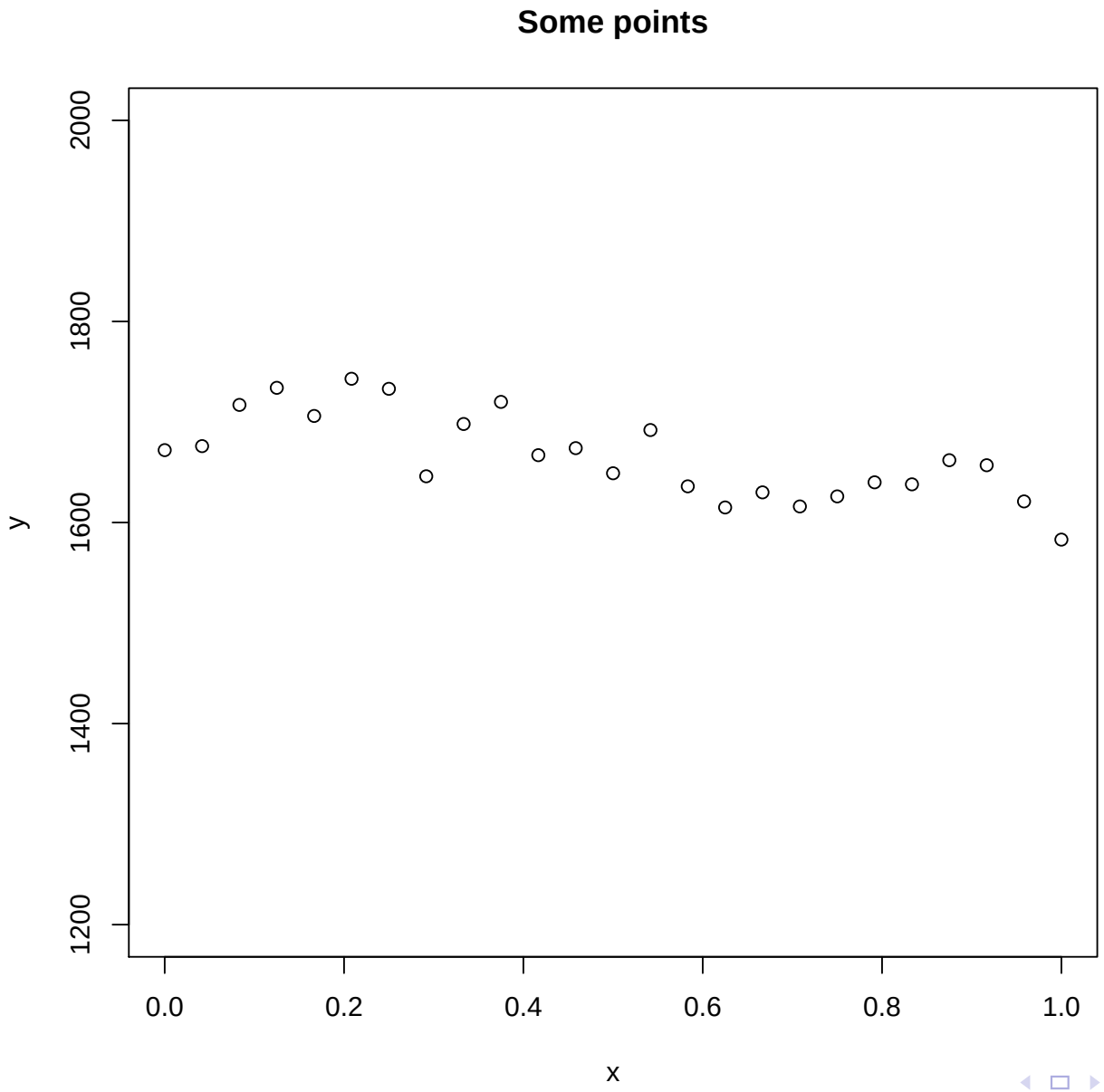
- Curves are often important in linguistics.
  - Tongue shapes
  - Formant contours outside of diphthongs
    - *oily*/*oiler*, *rural*/*whirl*
    - Mark Karan: *ferry*/*furry*
  - Pitch contours
- Not being able to compare curves is something of a hindrance.

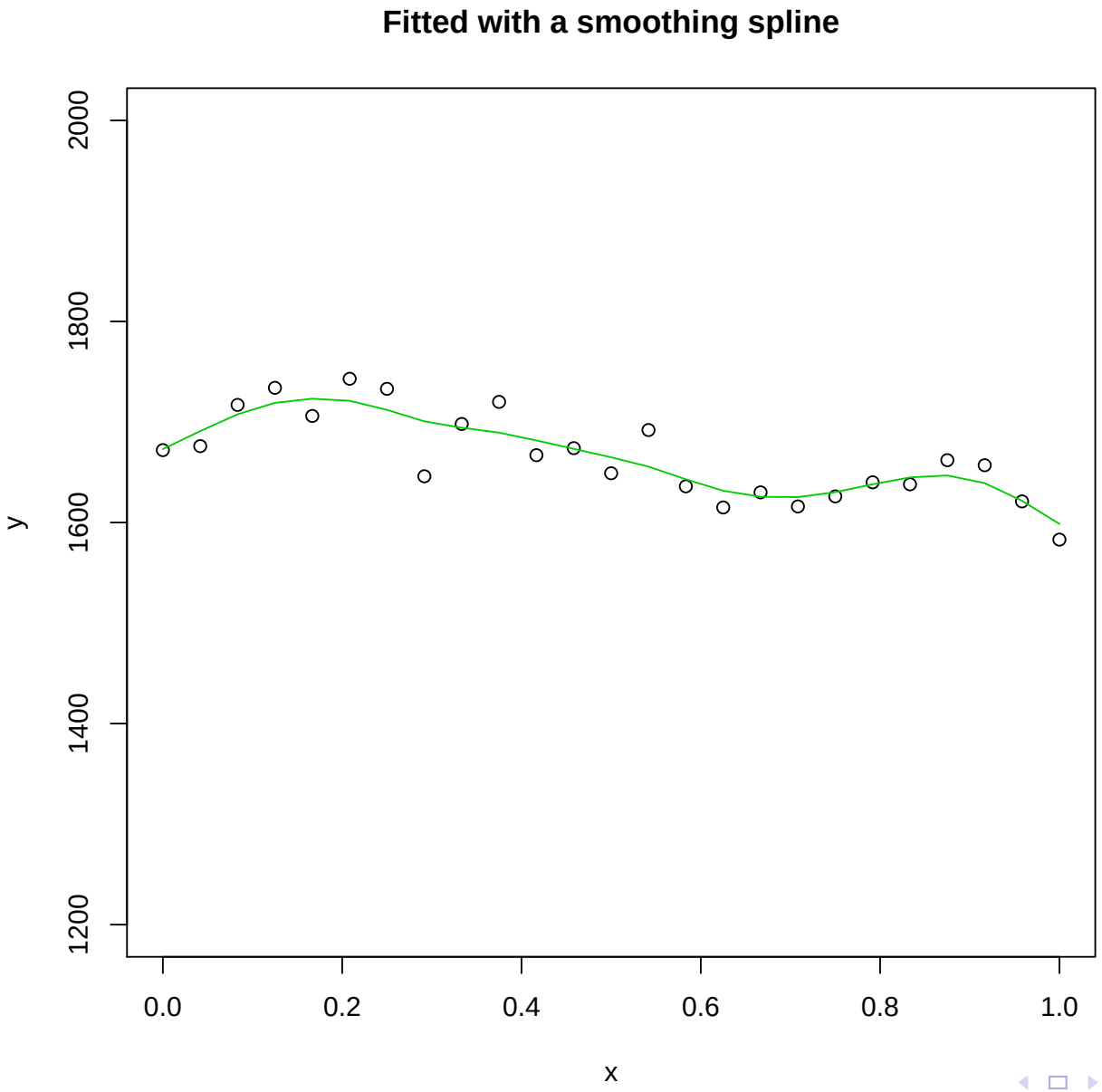
Smoothing Spline ANOVA

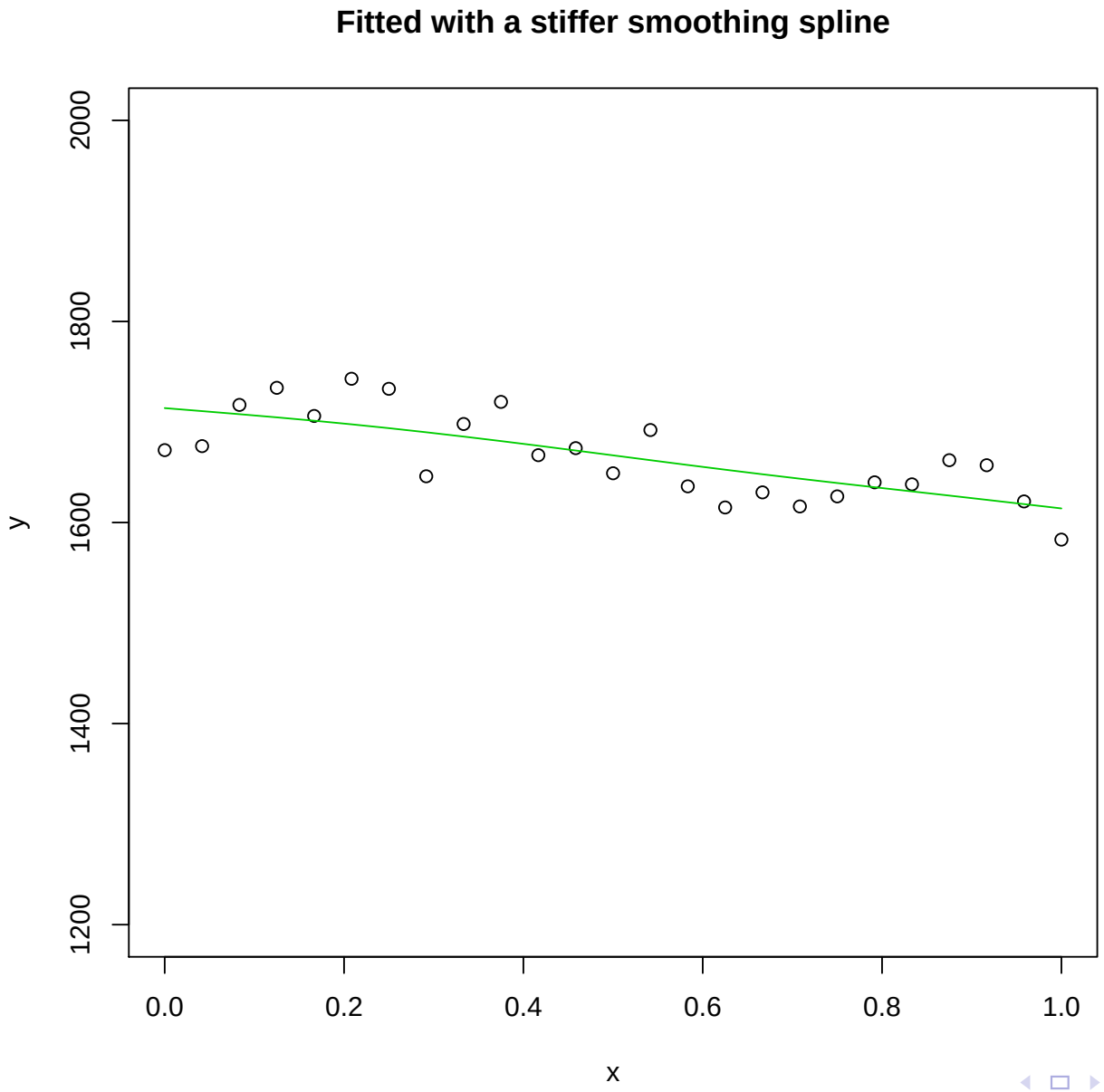
- The Smoothing Spline Analysis of Variance (SSANOVA) is a domain-general statistical tool for comparing curves.
  - it's been applied to, e.g., circadian rhythms, epidemiological studies
- Davidson (2006) brought SSANOVA to linguists by using to compare tongue shapes.
  - HUGE help to ultrasound research
- SSANOVA allows us to ask the "same-or-different" question for curves.
- It's used here for the comparison of formant contours.

Conceptual overview

- The "Smoothing Spline" part of a Smoothing Spline ANOVA.
- A smoothing spline is a function used to approximate data.
- You determine the smoothness of the function with a smoothing parameter.
  - Can be done automatically


Some points
y
x
2000
1800
1600
1400
1200
0.0
0.2
0.4
0.6
0.8
1.0


Fitted with a smoothing spline
y
x
2000
1800
1600
1400
1200
0.0
0.2
0.4
0.6
0.8
1.0


Fitted with a stiffer smoothing spline
y
2000
1800
1600
1400
1200
0.0
0.2
0.4
0.6
0.8
1.0
x

Conceptual overview

- Some splines will approximate a data set better than others.
- The "badness" of the spline's fit in the SSANOVA is similar to the idea of error in a regular ANOVA.

- Regular ANOVA: compare the error that results from including a factor or not including it.
- SSANOVA: compare the badness-of-fit when two curves are used for a data set instead of one.

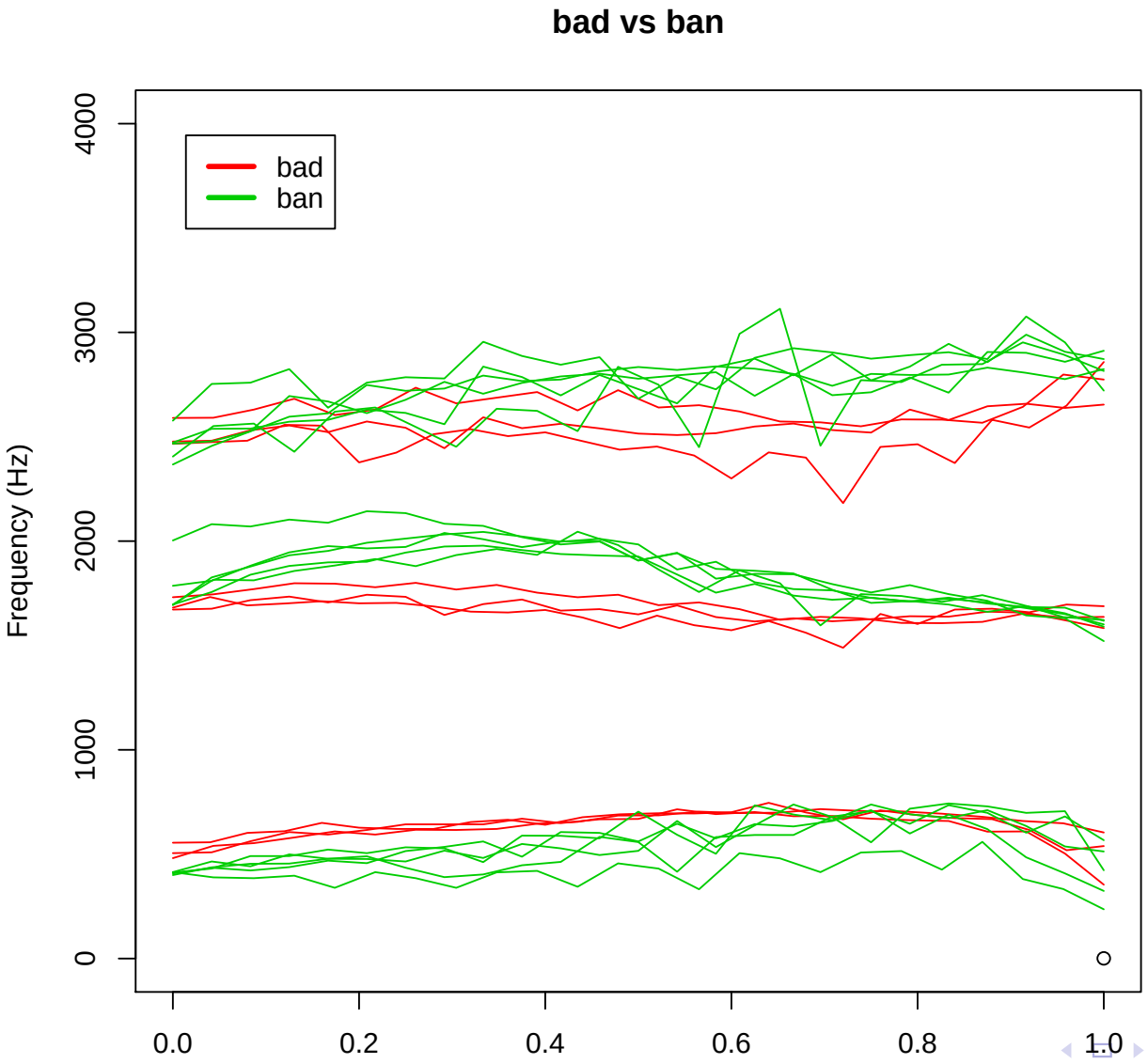
Conceptual overview

- What does "statistical significance" look like for curves?
- Two curves can be different along only a portion of their length.
- So the SSANOVA compares curves along their length.

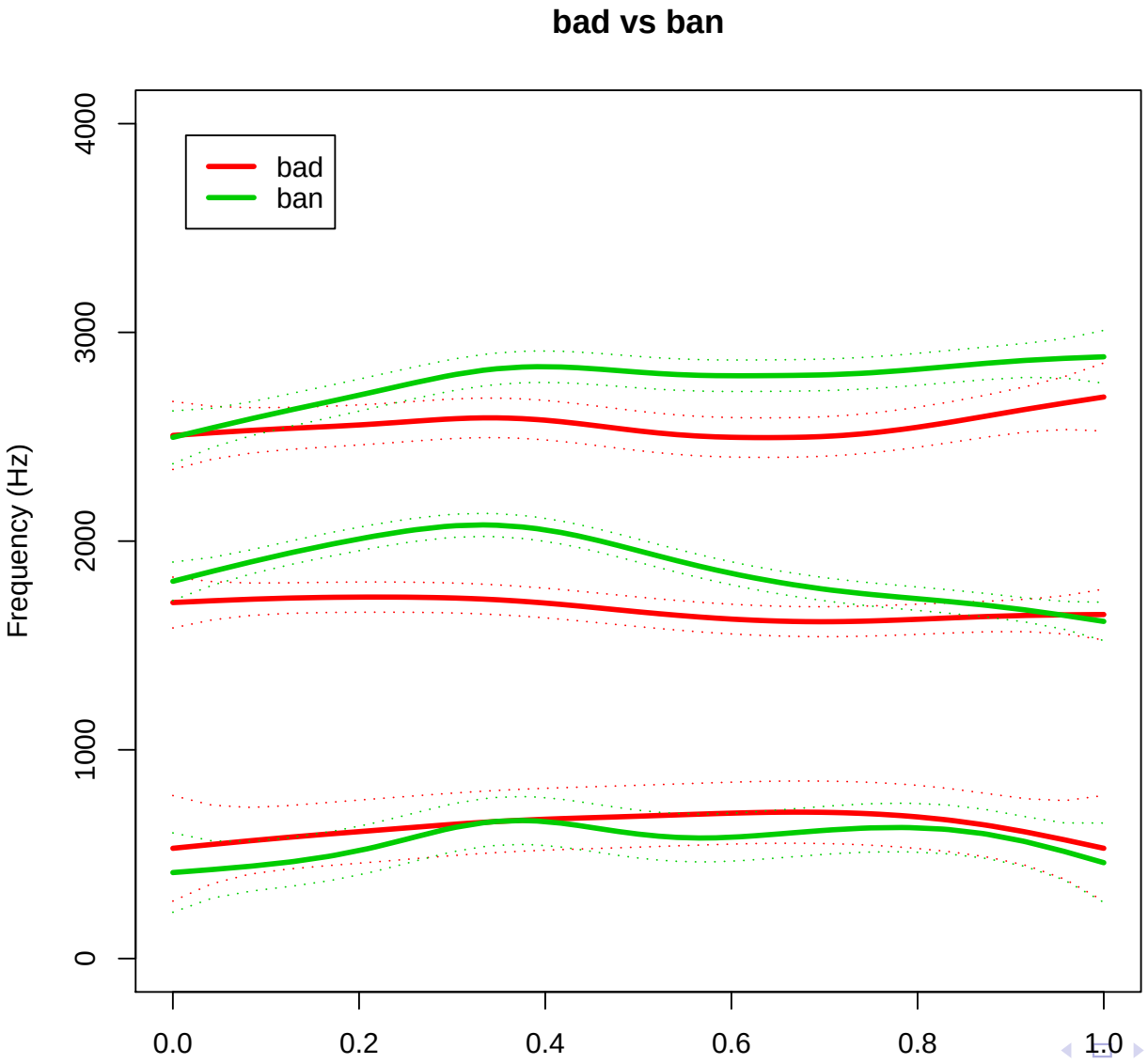
Conceptual overview

- Two lines are plotted representing the spline approximation of the data.
- 95% confidence intervals are plotted for both curve.
- We can be 95% confident that the mean of the curve lies within this interval.
- If the intervals of two curves do not overlap, we can be confident that they are distinct.
- This is conceptually similar to a p-value, but technically different.

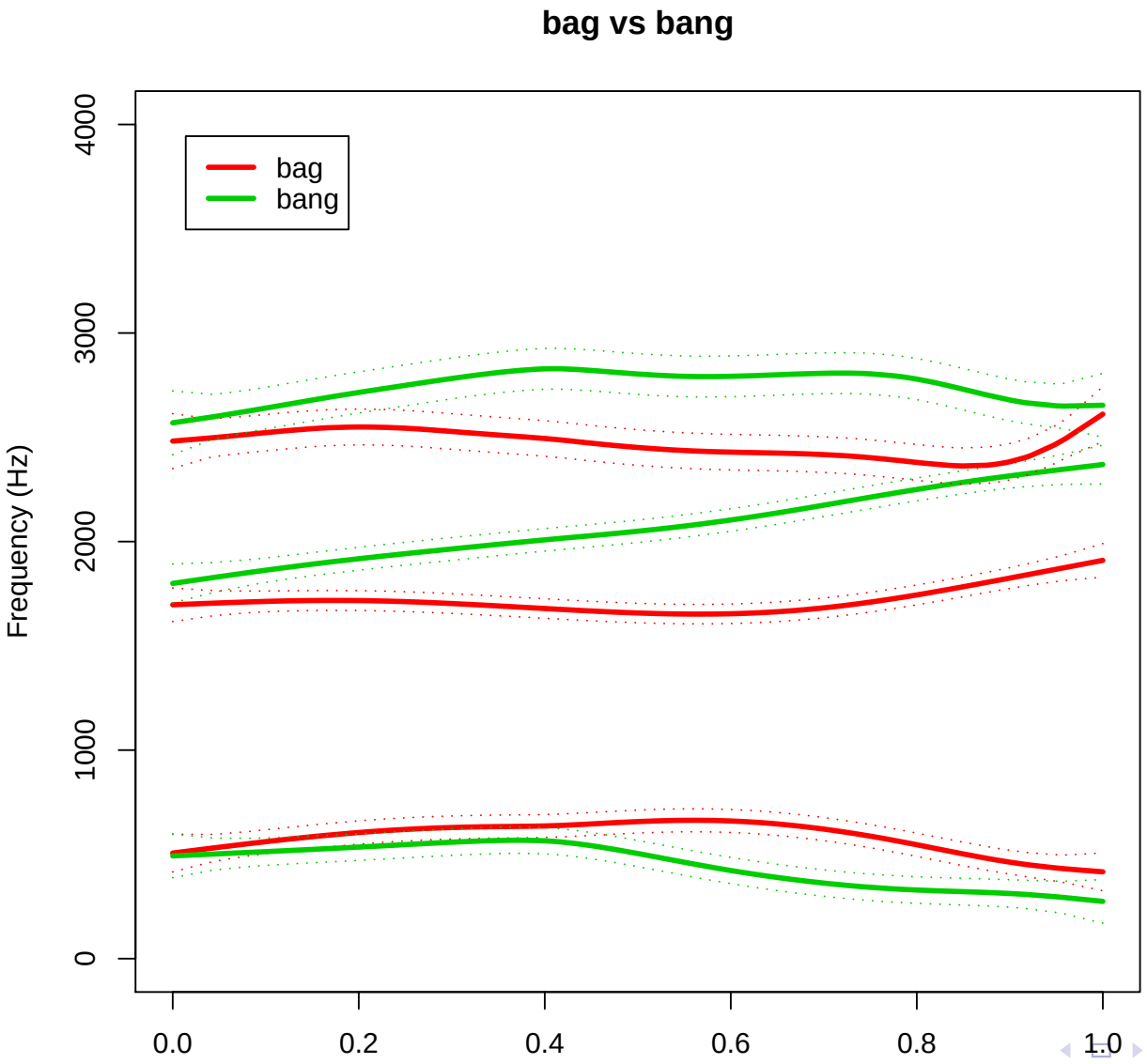
ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



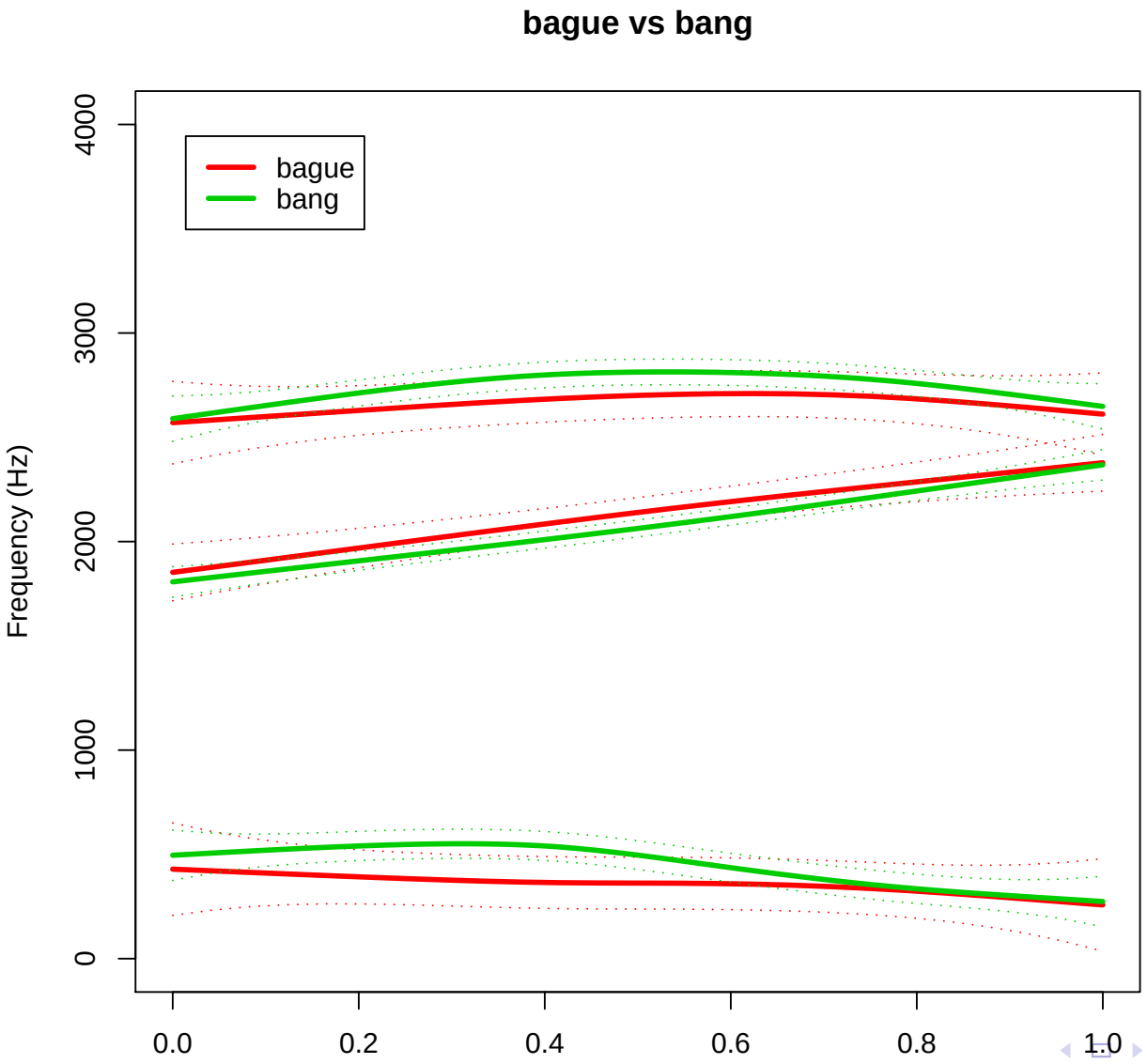
ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



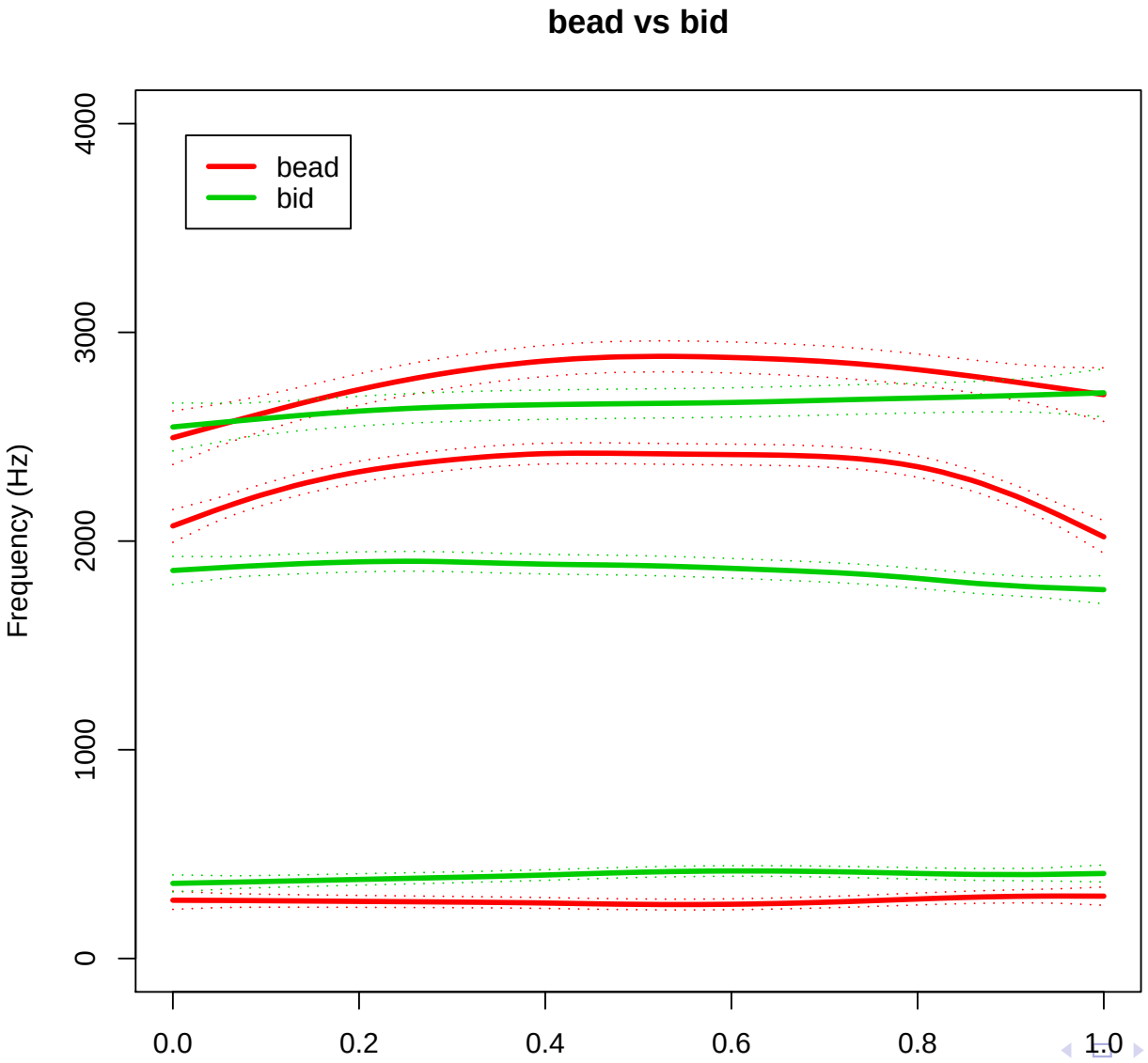
Open issues

1. Taking vowel duration into consideration.
2. Speaker normalization.

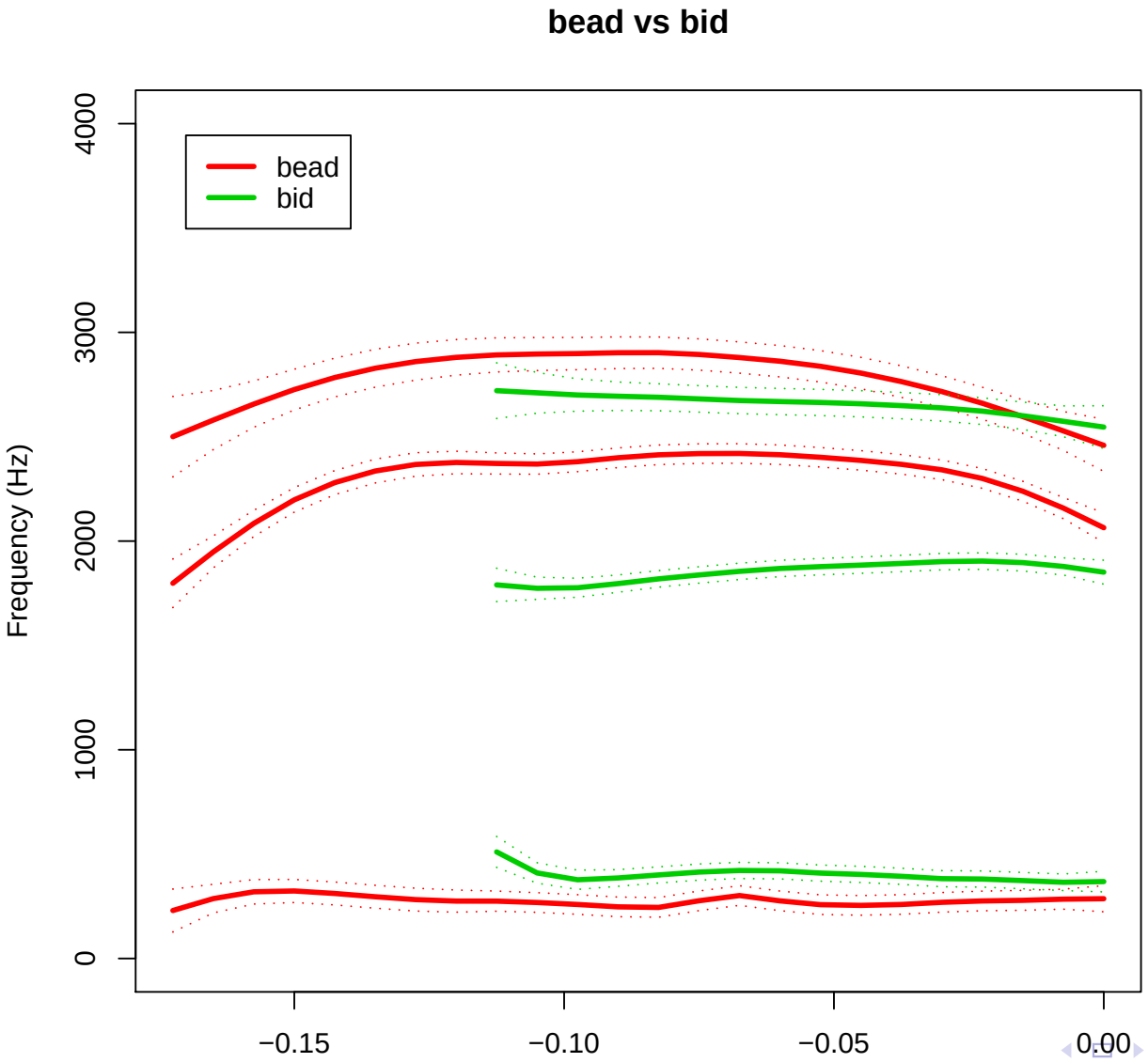
Vowel duration

- Not all productions of the vowel are of equal duration
  1. inter-token variability
  2. inherent duration differences (e.g., [i] v. [ɪ])
- The approach in the comparisons so far has been to normalize durations by stretching them linearly.
- One can also imagine...
  - Aligning the centers of vowels.
  - Aligning the right edges.
  - Aligning the left edges.
  - Performing some non-linear stretch.
- Aside from the non-linear stretch, these ideas have been implemented in the code on my web site.

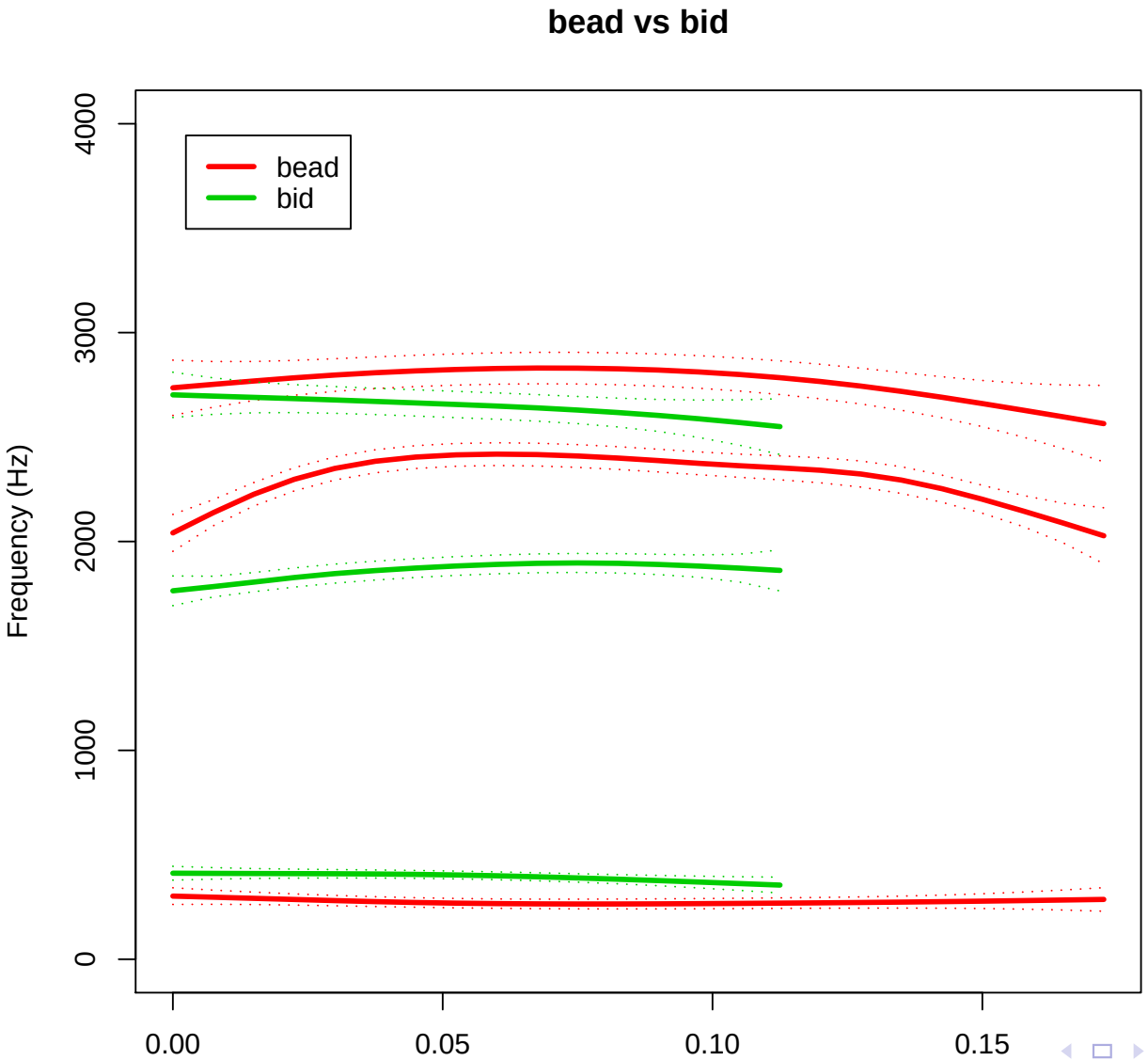
ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



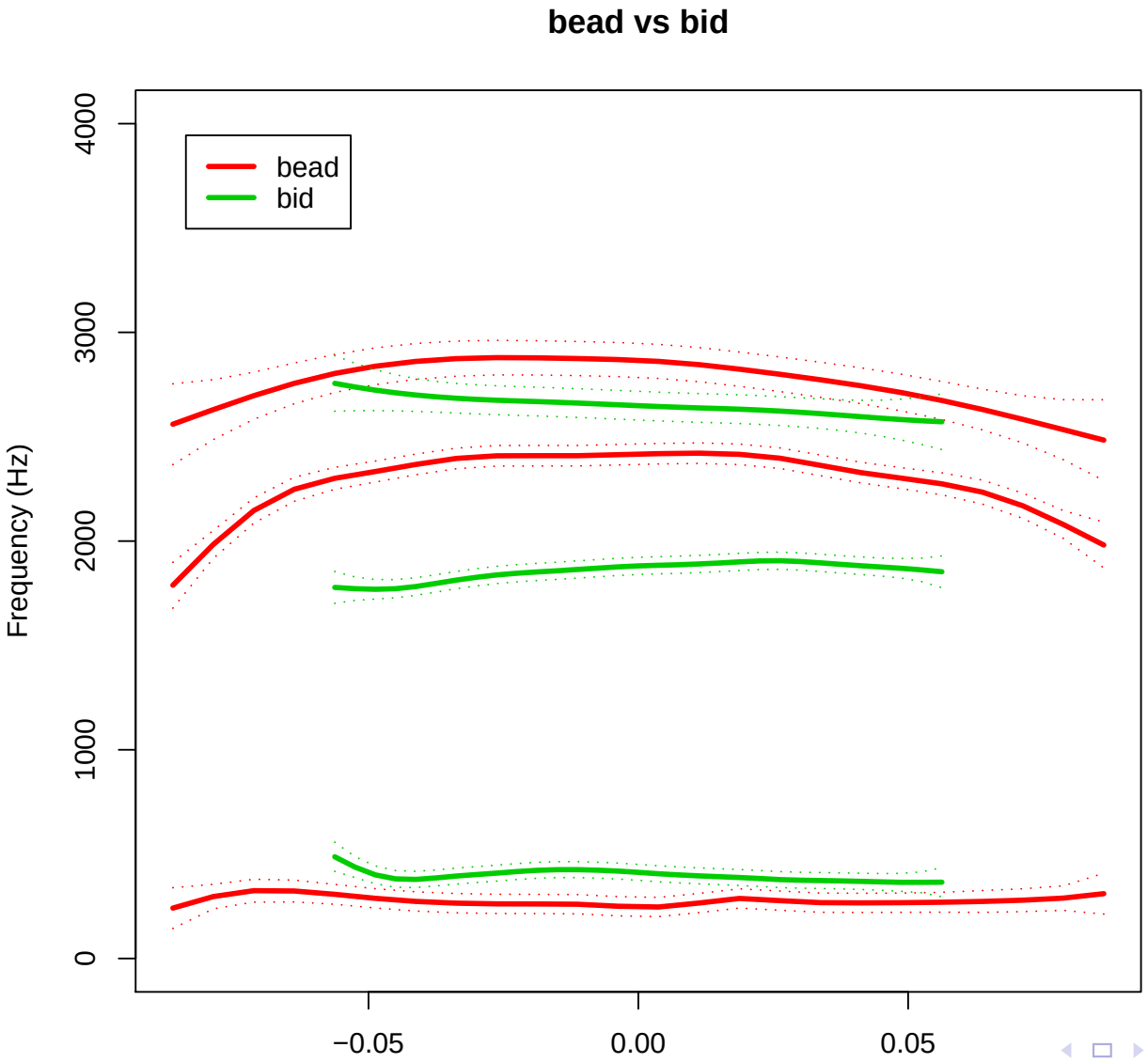
ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



ang2: 19 year-old male, has lived in Phoenix, AZ and Tucson, AZ



Vowel duration

- The commitment to a length normalization procedure is a theoretical one.
- Nothing is advocated or recommended here.

- But perhaps the SSANOVA can help us choose.
- "The curves are significantly different under length normalization regime X, but not under regime Y."

Speaker normalization

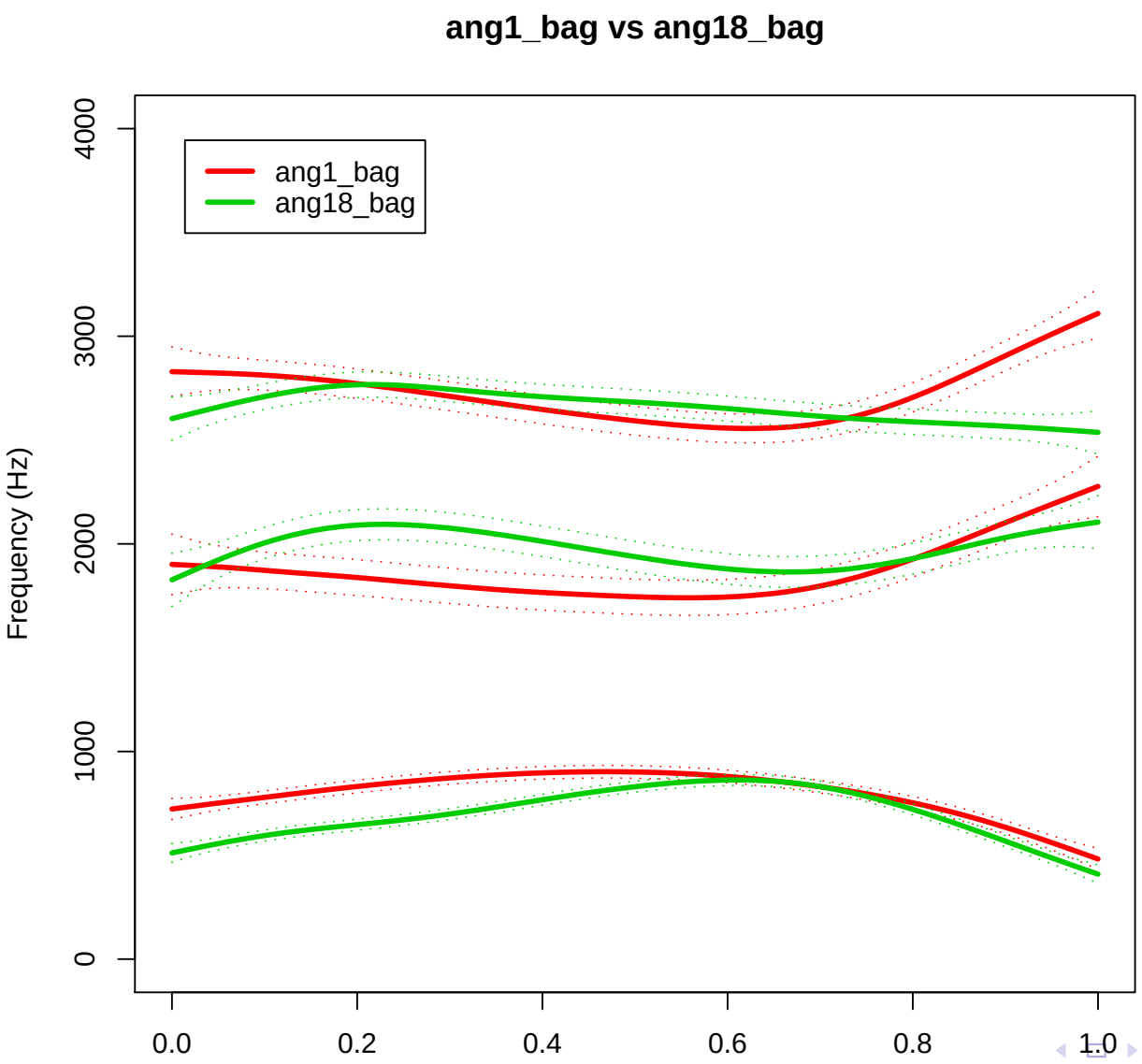
- Formant frequencies are often not directly comparable between subjects.
- The normalization of Nearey (1977) should be applicable even to contour data.
- Possibly need to rethink how to calculate mean vowel formants (do we measure at the midpoint? average of all formant measurements?)

Speaker normalization

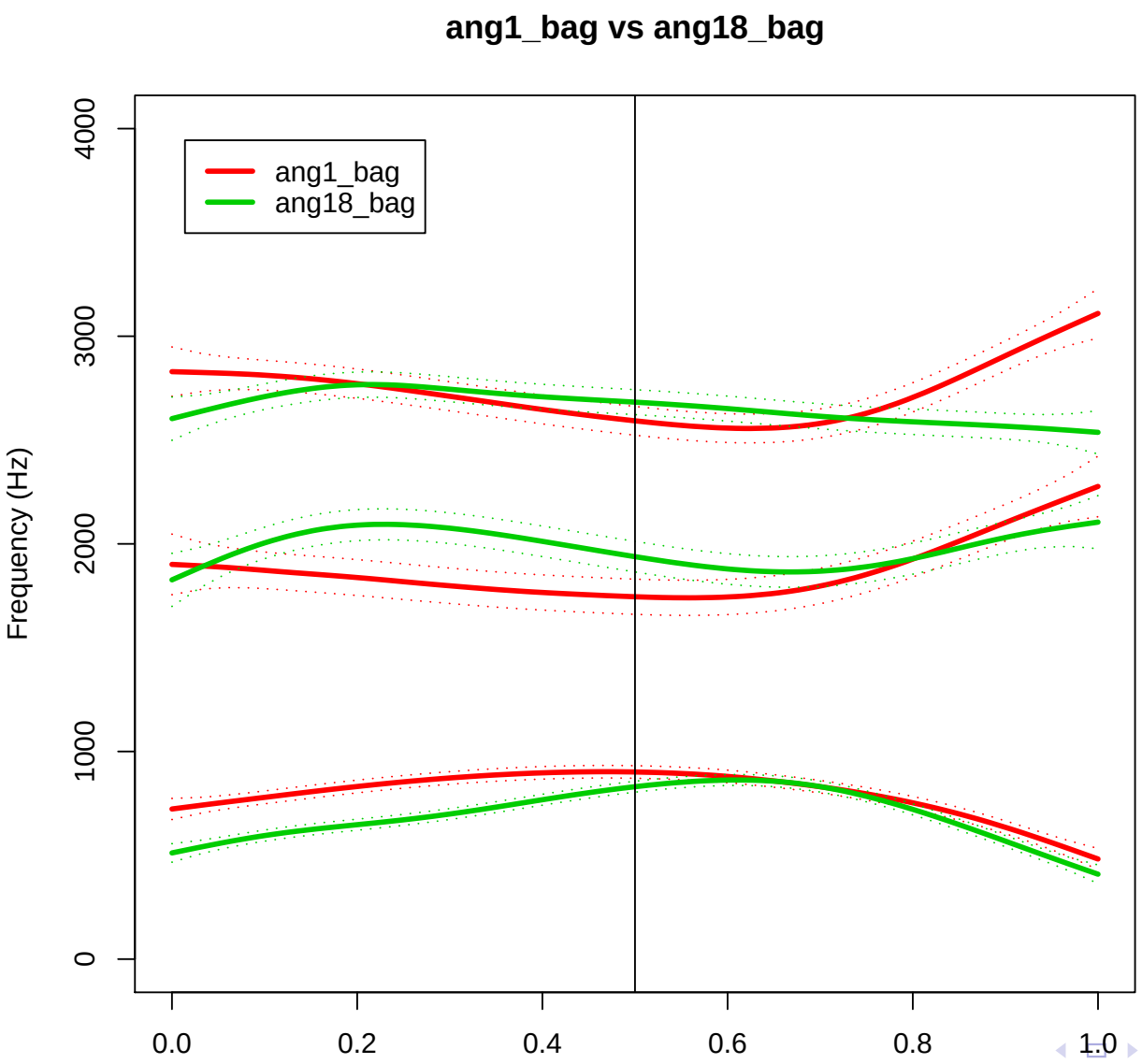
- We didn't get a full vowel set for the speakers shown here.

- Two 19-year-old females
- ang1: White, has lived in Phoenix and Tucson
- ang18: African American, has lived in Colorado, Utah, Tucson
- We'll compare the words "bag" and "dang"

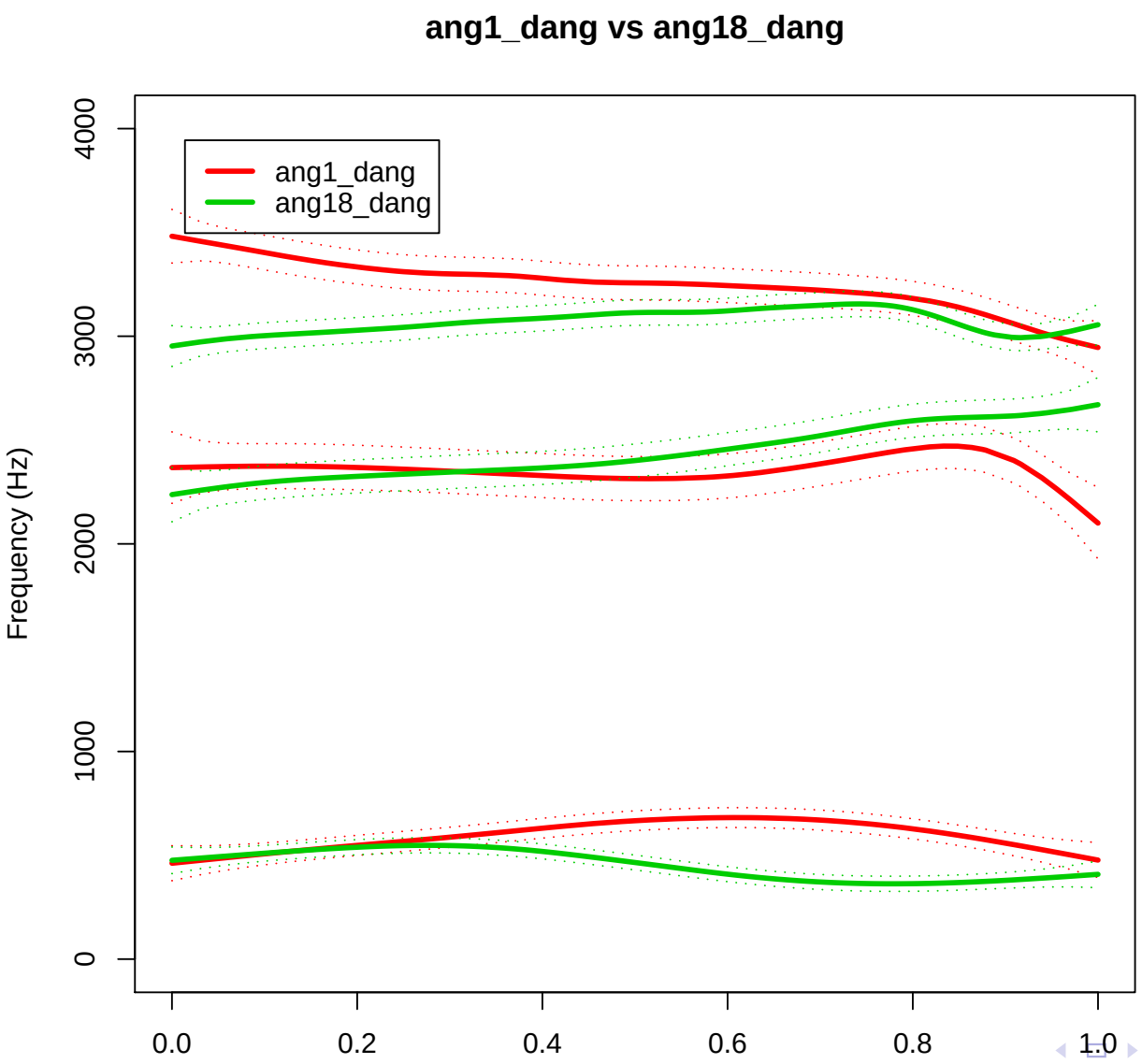
ang1 ang18


ang1_bag vs ang18_bag
ang1_bag
ang18_bag
Frequency (Hz)
0
1000
2000
3000
4000
0.0
0.2
0.4
0.6
0.8
1.0

ang1 ang18 Note the absence of a difference at the midpoint.



ang1 ang18


ang1_dang vs ang18_dang
ang1_dang
ang18_dang
Frequency (Hz)
0
1000
2000
3000
4000
0.0
0.2
0.4
0.6
0.8
1.0

Availability

- SSANOVA tests can be performed with the R or S statistical packages
  - Functions written by Wang & Ke (2002)
  - Helpful interface provided by Kyung Sin
  - Adapted for formants by Baker (this presentation)
- R is an open-source, cross-platform statistical environment
  - difficult to use, but (*ahem*) free
- Instructions for making it work at
  `http://www.u.arizona.edu/~tabaker/`

Conclusion

- Smoothing Spline ANOVA tests allow us to compare curves.
- This is particularly nice for vowel formants.
- There are a few issues to be worked out; nothing too serious.

- Instructions for making it work at `http://www.u.arizona.edu/~tabaker/`