# Measuring Fricatives

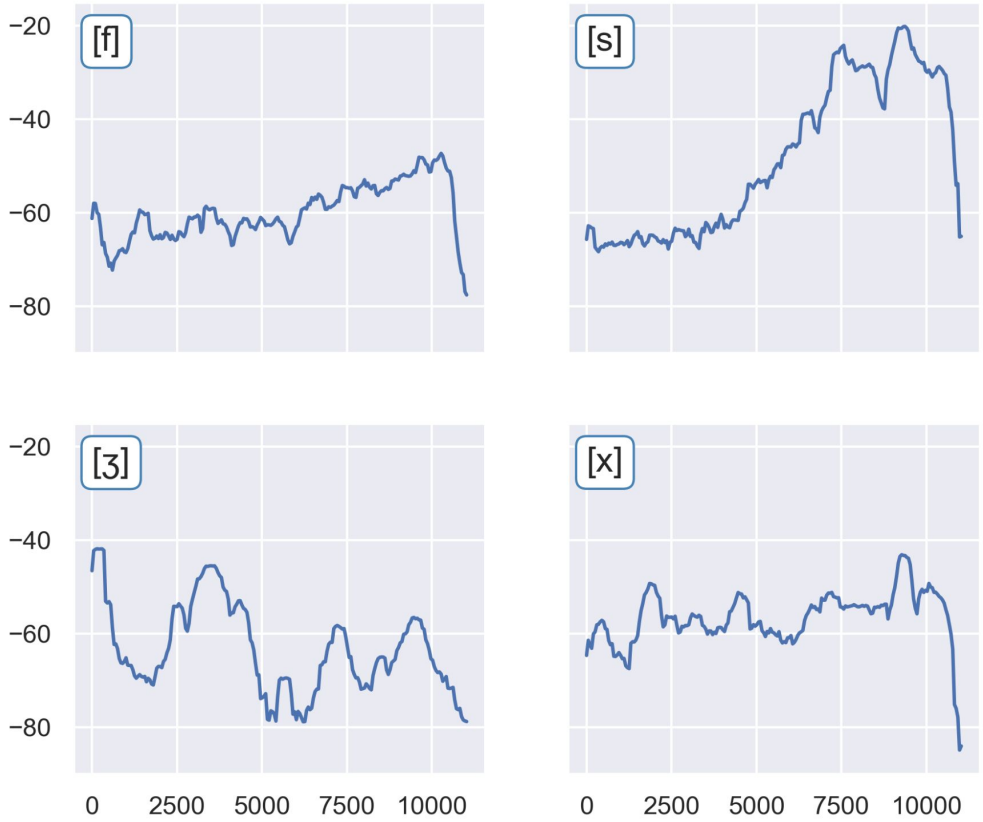
Keith Johnson
phonetics lab meeting
Feb 22, 2024

# Mexican Spanish

Fricatives differ in terms of the shape of their spectra.

Examples here are spectra from the midpoint of fricatives produced by one speaker of Mexican Spanish.

What are some things you notice about them?



# Moments Analysis

Emily Grabowski’s data on analysis preferences in phonetics indicates that moments analysis is the dominant method used to measure fricative acoustics.

What is ‘moments analysis’?

# Moments Analysis

1. Introduced by Forrest et al. 1990
2. Treat the spectrum as a statistical distribution and then take the first four statistical moments (mean, standard deviation, skew, kurtosis)
3. Calculation
    a. Take a spectrum (FFT, Tandem, Multi-taper) - two arrays, frequencies (f) and amplitudes (a)
    b. Take a frequency band (300-7000Hz)
    c. Normalize the amplitude values so they sum to 1 (e.g. Norm[i] = a[i]/sum(a) )
    d. Mean = COG = sum(f * Norm)
    e. SD = sum((f-COG)$^2$ * Norm)
    f. Skew = sum((f-COG)$^3$ * Norm)
    g. Kurtosis = sum((f-COG)$^4$ * Norm)
4. COG and SD seem to be the most useful of these spectral shape parameters

# Shadle's critique of moments analysis

- Unstable spectral estimation (recommends using multi-taper spectrum)
  - what is multi-taper?
- Normalization of amplitude removes a key acoustic property of fricatives
  - for example sibilants are louder than the other fricatives
- Gaussian assumption is inappropriate (especially when there is more than one peak)
- Parameters are not articulatorily interpretable in source-filter theory of speech production.
  - compare this to the interpretation we have for formants as vocal tract resonances

# Major Peak Analysis (MPA)

- Inspired by Shadle's suggested analysis
- Assume that for many fricatives the first vocal tract resonance is a major peak in the spectrum, and is key to interpreting the place of articulation of the fricative.
- Goal, find the F1 of fricatives.
- Why is it hard to do this?
  - Length of the resonant cavity is different for each fricative (the front cavity)
  - LPC analysis (the main tool for finding resonances in vowels) doesn't work well with voiceless speech
- Answer: use heuristic peak-picking.

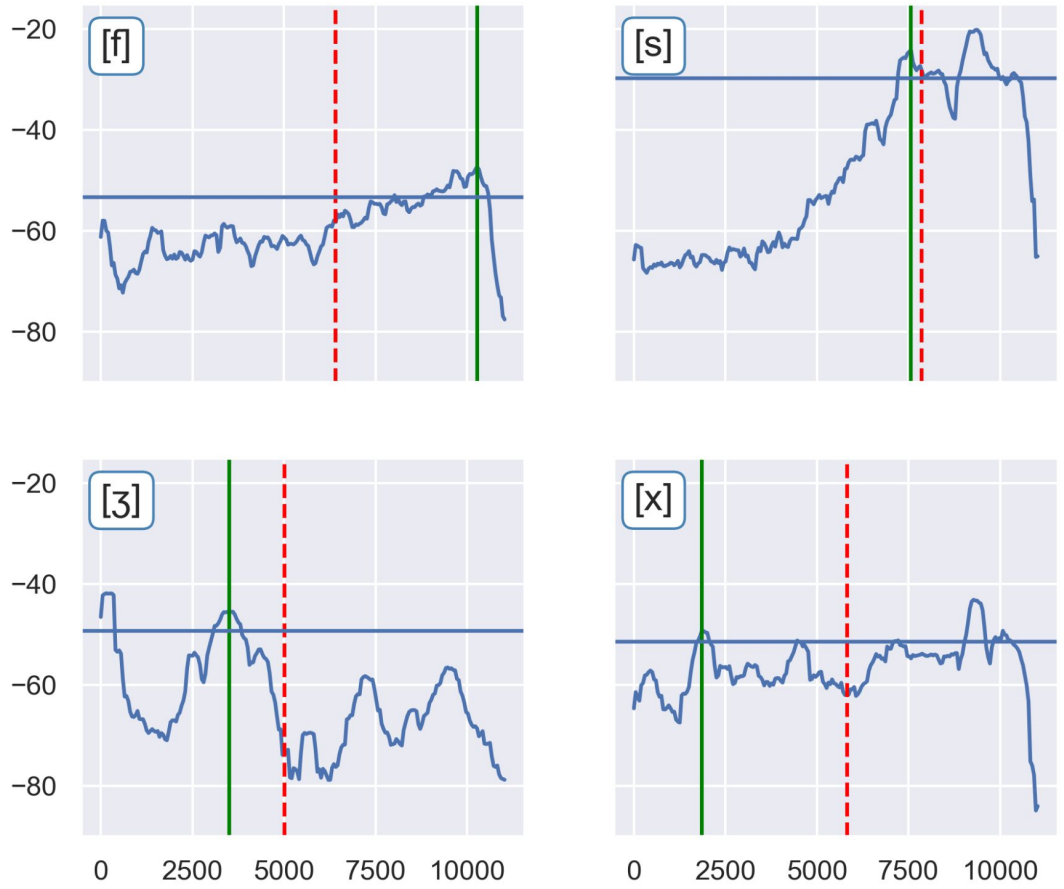
# Major Peak Analysis (MPA)

- Two measurements
  - Fm - the frequency of the lowest major peak in the fricative spectrum
  - Am - the amplitude of the lowest major peak in the fricative spectrum
- Algorithm
  - Use a multi-taper spectrum, using the **`nitime.multi_taper_psd()`** function
  - Return the first spectral peak that meets these criteria
    - η - minimum frequency (500 Hz)
    - δ - minimum distance between peaks (500 Hz)
    - α - minimum amplitude (80% of the amplitude range of the fricative spectrum)
    - φ - minimum prominence (6 dB up from the closest valley)
  - Relax α and φ, if no peak is found
  - This was implemented with the **`scipy.find_peaks()`** function

# Mexican Spanish

COG - dashed red line
Fm - solid green line

horizontal line is 80% of the spectral amplitude range, for 500 Hz to 11000 Hz.

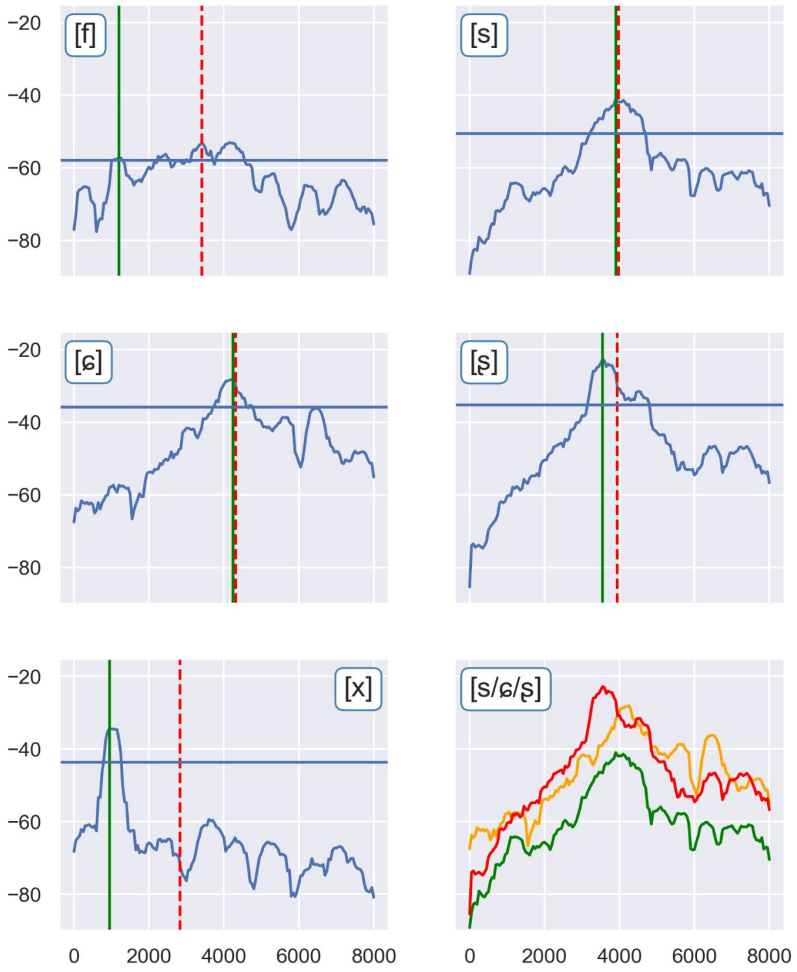


# Beijing Mandarin

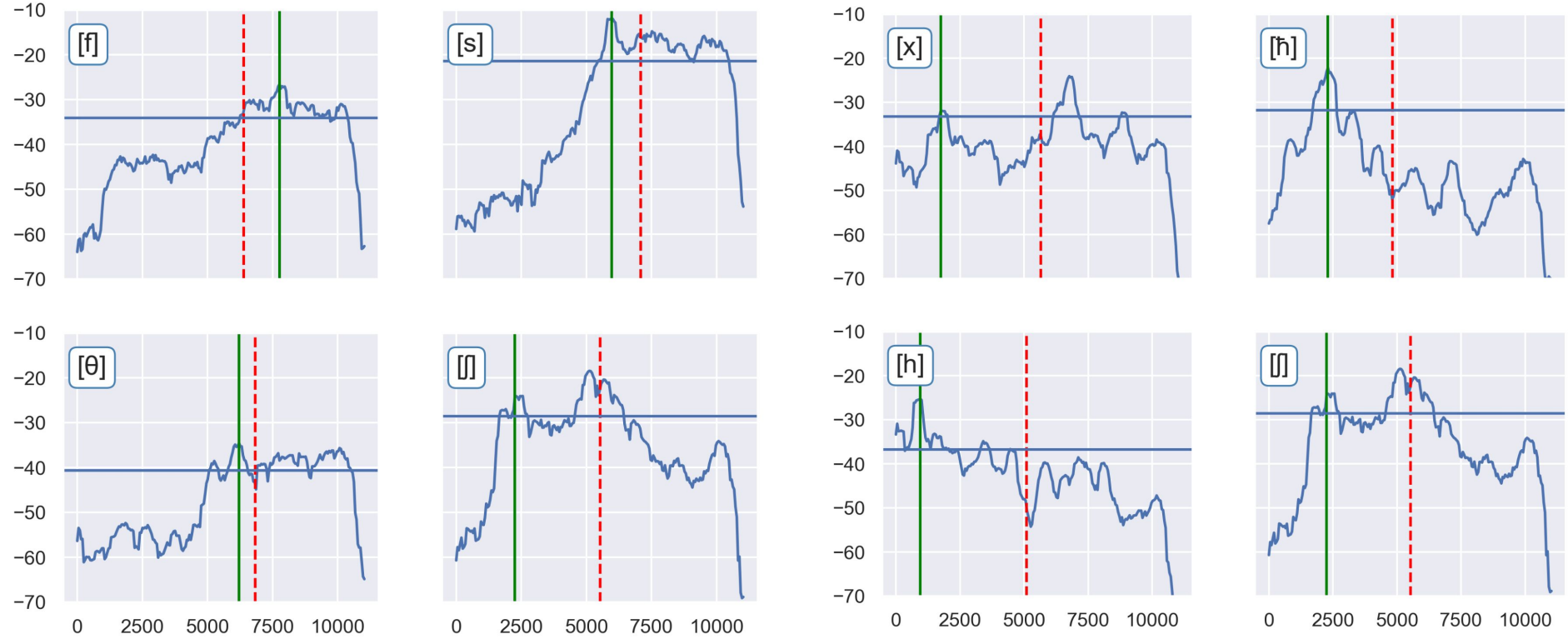
Example fricative spectra from a speaker of Beijing Mandarin.

- Coronals are pretty Gaussian, as assumed in Moments analysis.

What else do you notice?



# Syrian Arabic – further examples of fricative spectra

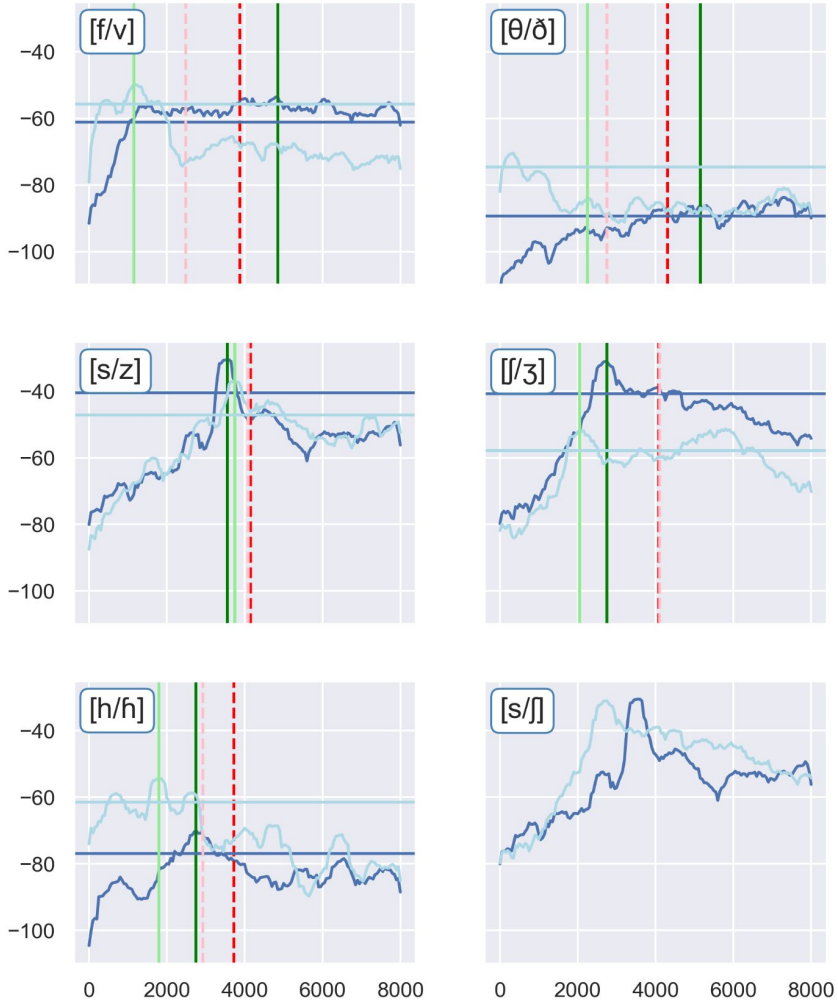


# American English

Example spectra of the voiced (fainter lines) and voiceless (darker lines) fricatives of American English.

- are [s] and [ʃ] actually voiced in these examples?

The major peak is marked with the solid green lines, and the COG is marked with dashed red/pink lines.



# Fricative Spaces

Syrian Arabic (Halbi, 2016)
- one male talker, news reading
- fricative inventory [f θ s ʃ x ħ h]
- 9,580 fricative tokens (3401 [s] to 556 [θ])

Mexican Spanish (DIMEx, Pineda et al. 2004)
- 100 talkers (49 m, 51 f), sentence reading, 60 sentences each.
- fricative inventory [f s x ʒ]
- 2,959 fricative tokens (2479 [s] to 77 [ʒ])

Beijing Mandarin (ASC, Li Aijun, et al. 2000)
- 10 (5 m, 5 f) talkers, reading discourses (30 minutes/talker)
- fricative inventory [f s ɕ ʂ x]
- 19,043 fricative tokens (7412 [ʂ] to 1376 [s])

American English (TIMIT, DARPA 1990)
- 628 talkers(70% m), sentence reading, ~11 sentences each.
- fricative inventory [f v θ ð s z ʃ ʒ h ɦ]
- 28,169 fricative tokens (10115 [s] to 217 [ʒ])

# Fricative classification scores

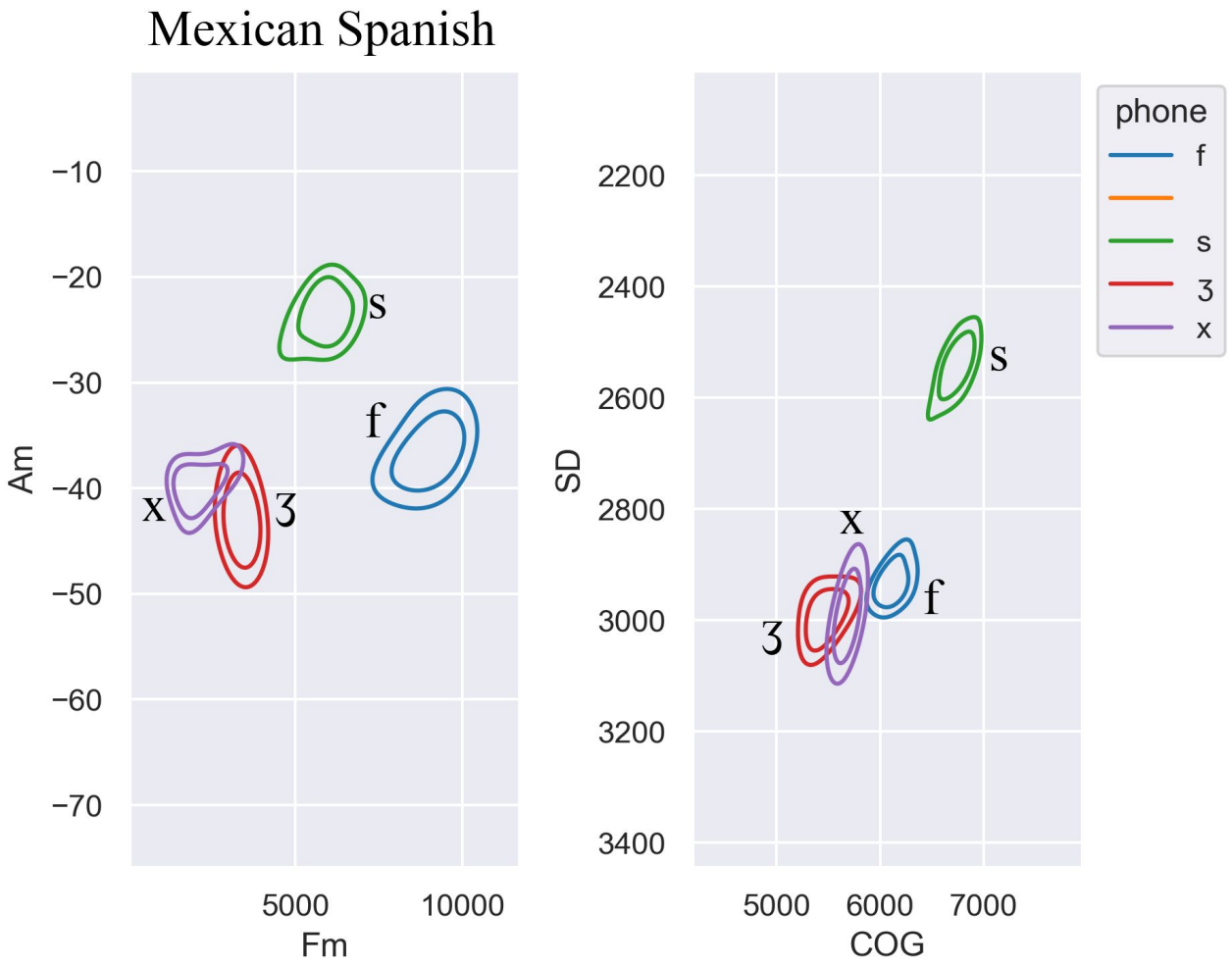
Linear Discriminant Analysis - % correct fricative classification, sample an equal number of tokens of each fricative in each language. 5-fold cross-validation.

| | | Arabic | Spanish | Mandarin | English | Average |
|---|---|---|---|---|---|---|
| | chance | 14% | 25% | 20% | 10% | 17% |
| Major Peak Analysis | Fm | 44 | 51 | 50 | 38 | 46 |
| | Fm + Am | 59 | 68 | 64 | 58 | 62 |
| | Fm + Am + dur | 69 | 71 | 65 | 59 | 66 |
| Moments Analysis | COG | 53 | 57 | 53 | 44 | 52 |
| | COG + SD | 67 | 59 | 66 | 55 | 62 |
| | COG + SD + dur | 73 | 65 | 67 | 55 | 65 |

# Spanish

MPA predicts that [f] and [ʒ] are distinct, Moments sees them as quite similar.

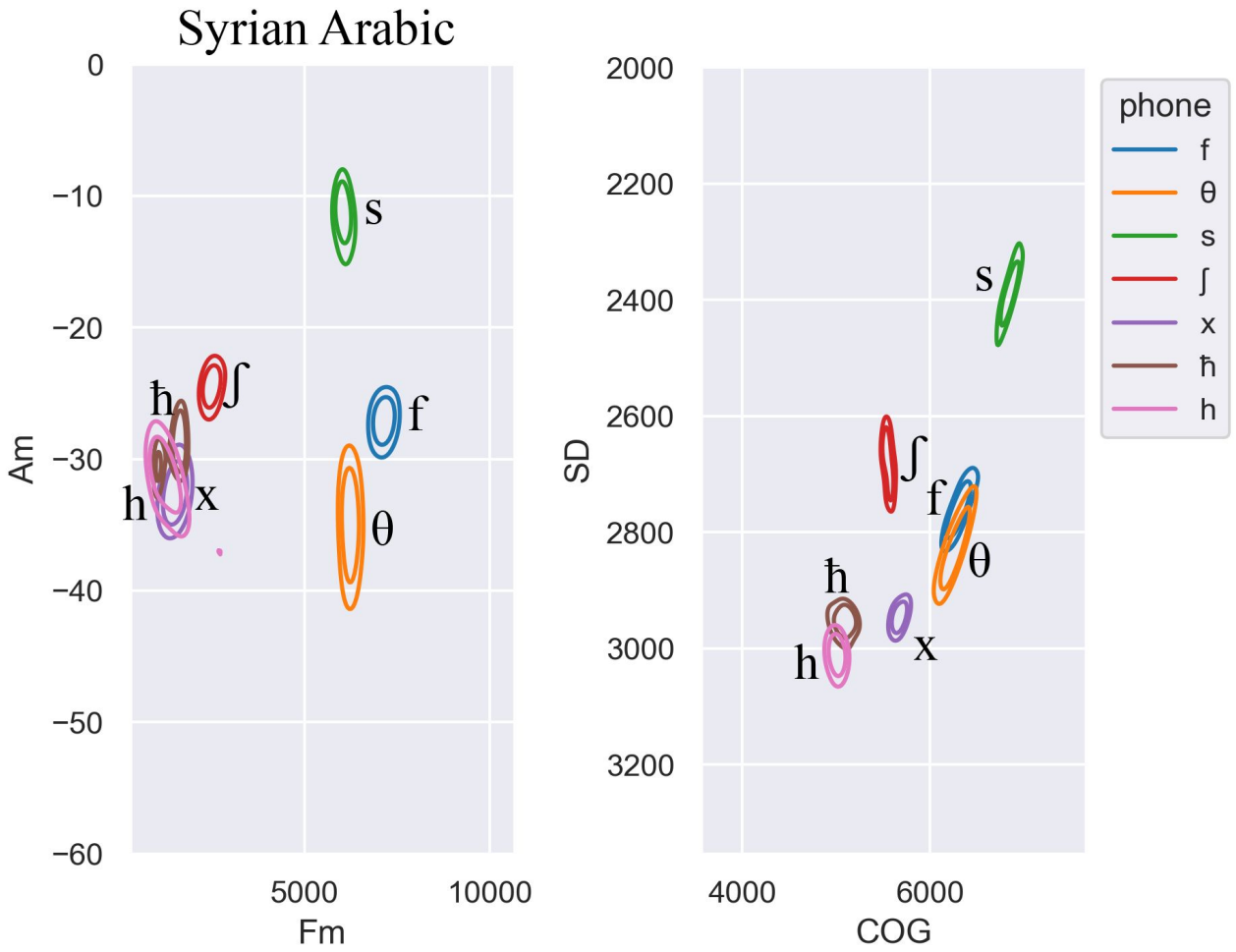
[x] and [f] are distinct in the MPA space but close in Moments analysis.



# Arabic

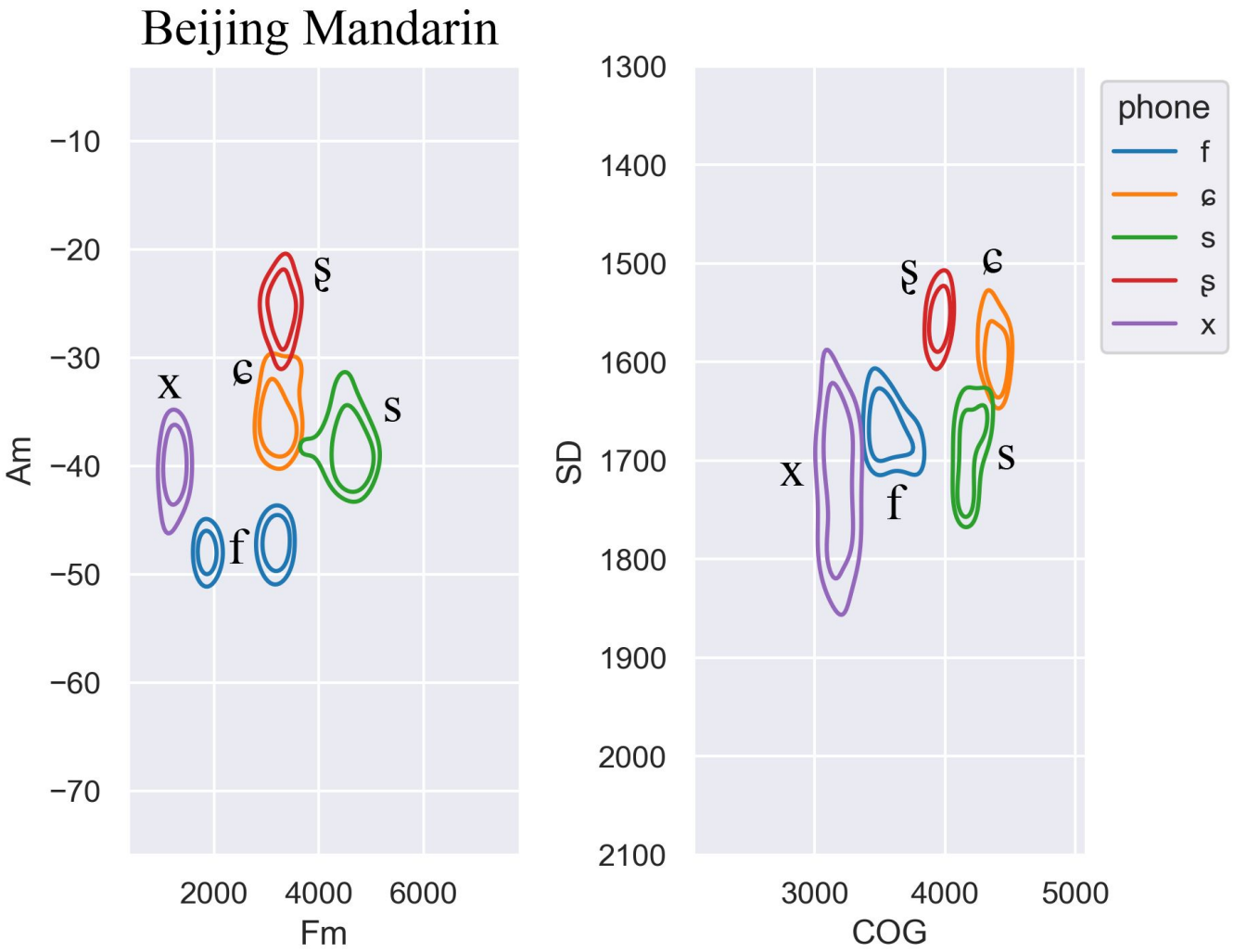
MPA predicts that [ʃ] is more similar to the back fricatives.

Moments predicts that [x ħ] and [h] are distinct. MPA predicts that they are confusable.



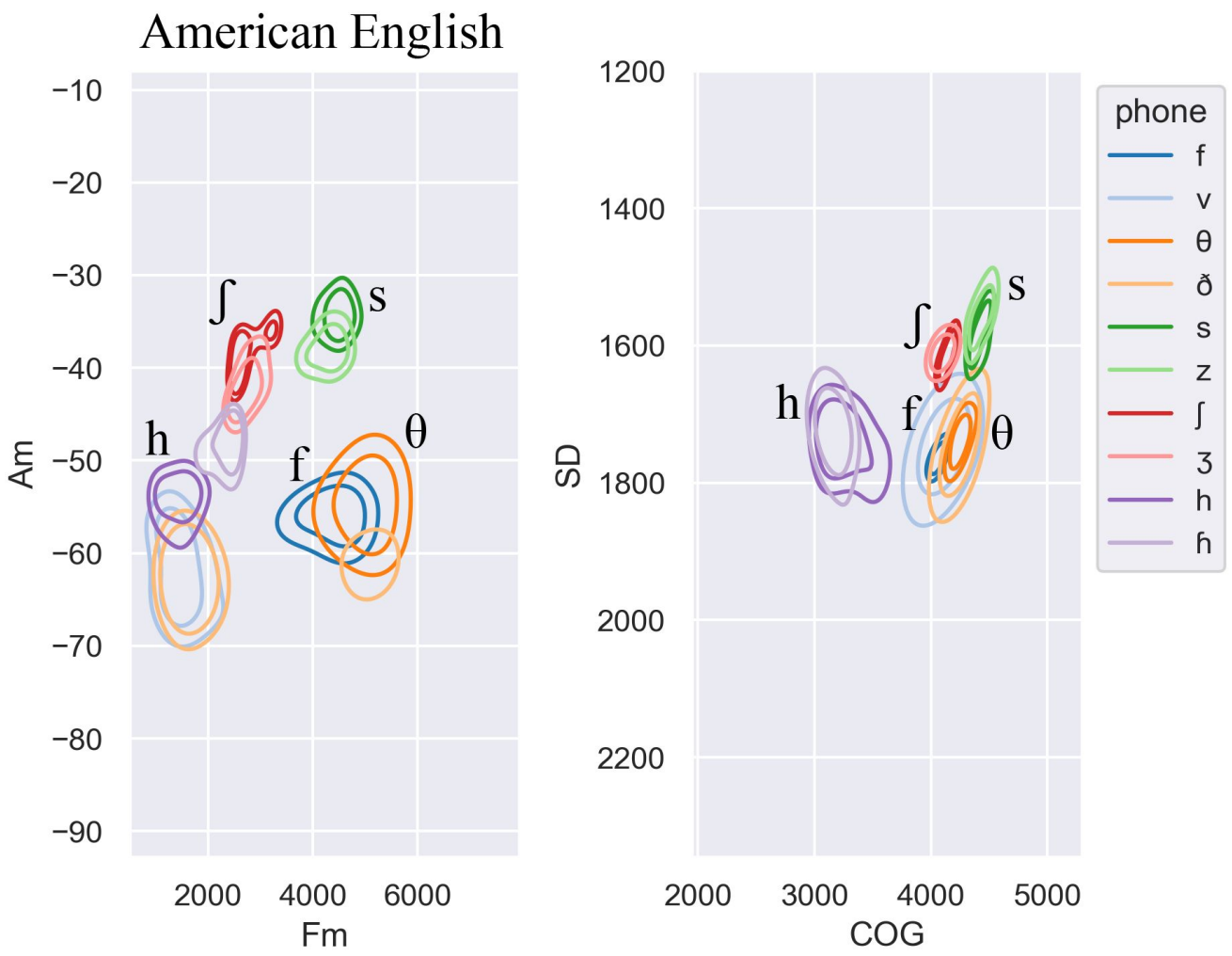
# Mandarin

MPA has [s] distinct from [ɕ] by frequency of the major peak. Moments has them distinct primarily by SD

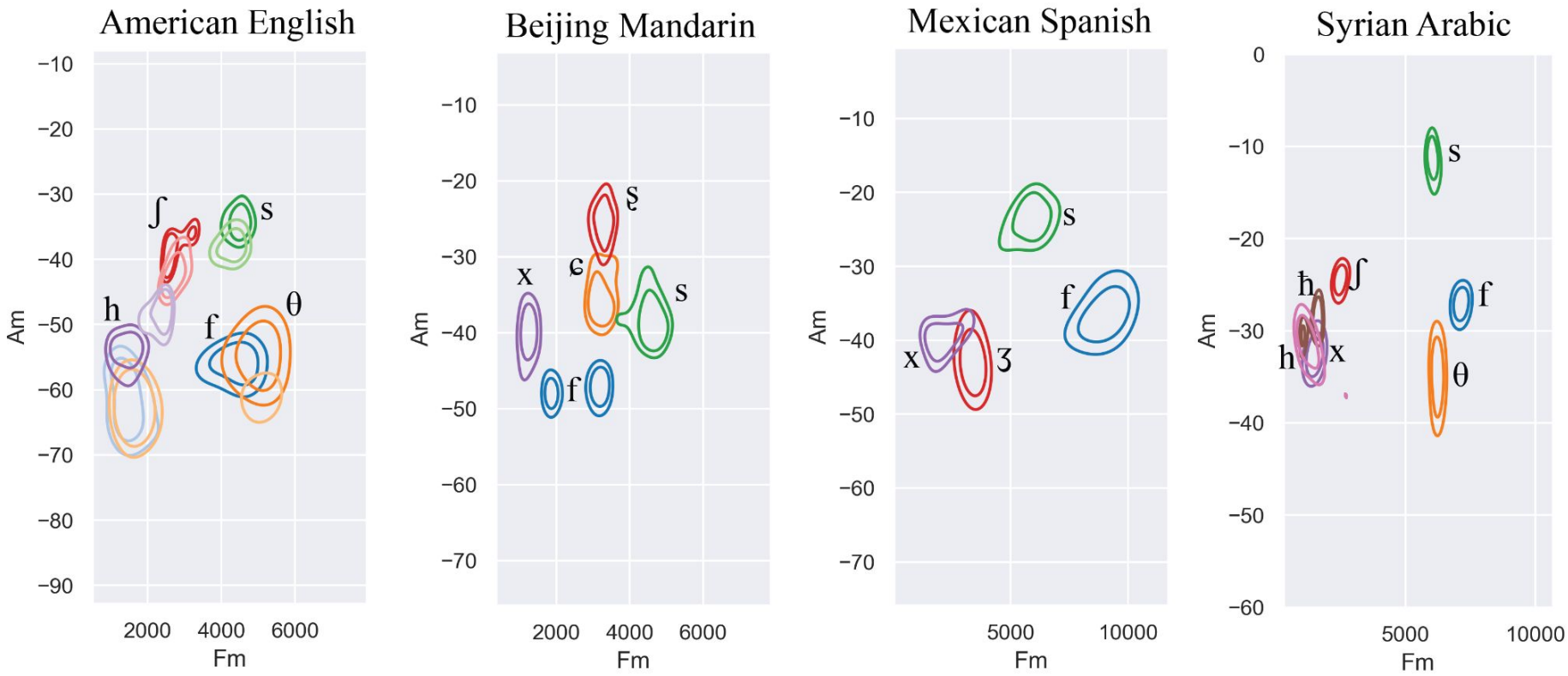


# English

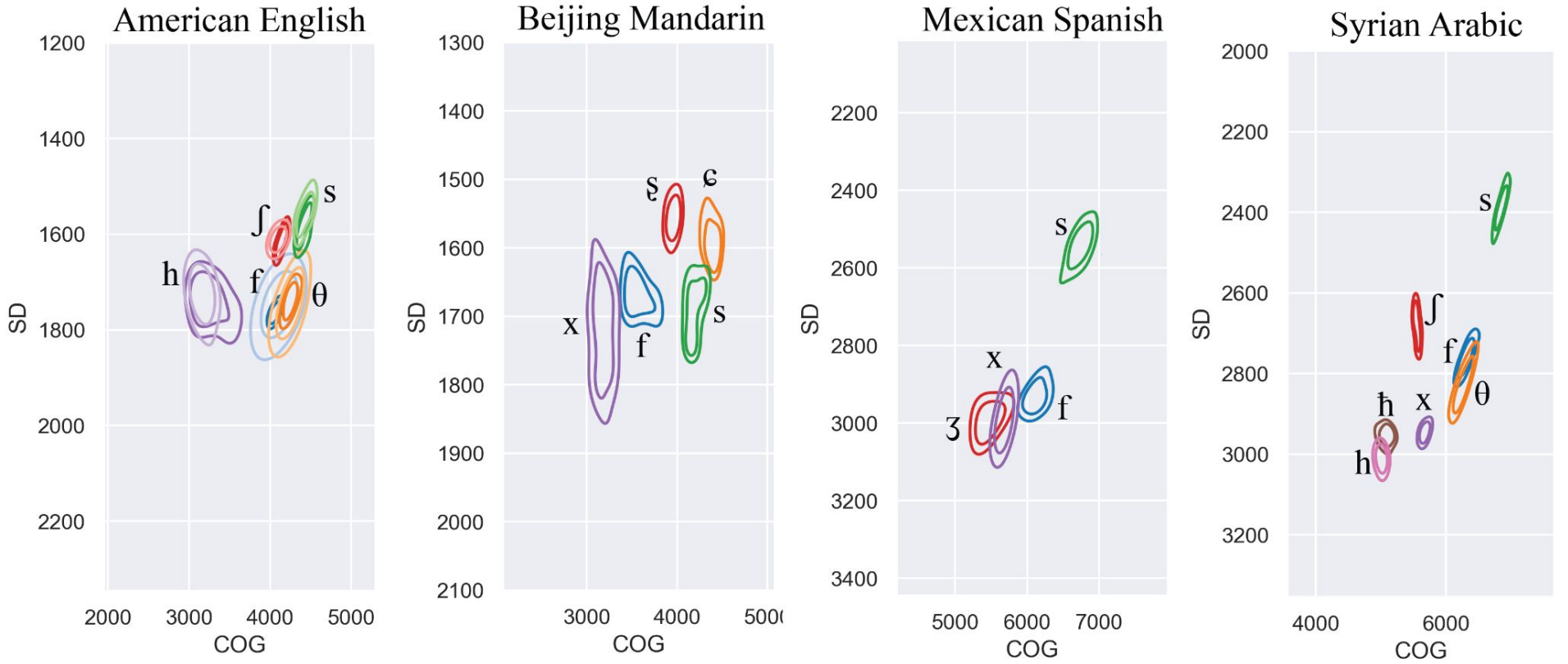
Similar spaces. Voiced fricatives show different spectral measures than voiceless in MPA but not in Moments.



# MPA spaces, cross-linguistic comparisons



# Moments spaces - cross linguistic comparisons



# Future Directions

- Put analysis script onto github
- Compare two corpora of the same language (next up the Buckeye Corpus)
- Test perceptual similarity predictions (see Melguy and Johnson, 2023)
- Explore additional parameters (e.g. F2 of back fricatives)