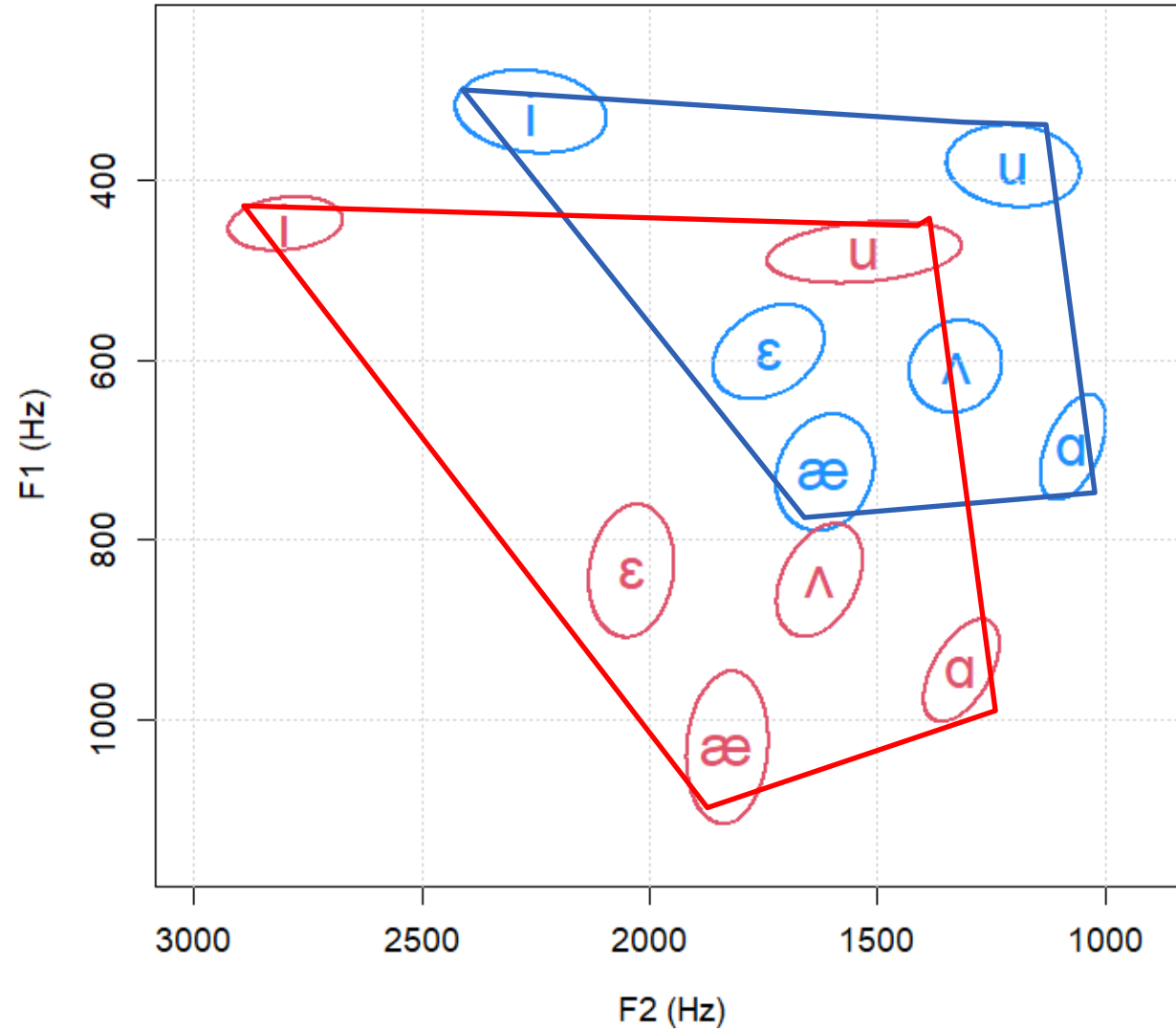# There is no female vocal tract: Abandoning essentialist ideology in phonetics

Santiago Barreda
University of California, Davis
sbarreda@ucdavis.edu

Michael Stuart
National Yang Ming Chiao Tung University

Vowel Formant Data

Male
Female

Fant (1966) - A note on vocal tract size factors and non-uniform F-pattern scalings

"The female to male scale factor is of the order of 18% averaged over the whole vowel system. The typical deviations from this rule are:

[…]

(c) Very open front or back vowels display a first formant "sex factor" k_1 which is substantially higher than the average.

[…]

These findings conform with anatomical constraints of the average female vocal tract. The particular scaling from male to female tract reduces the pharynx length more than the length of the mouth."
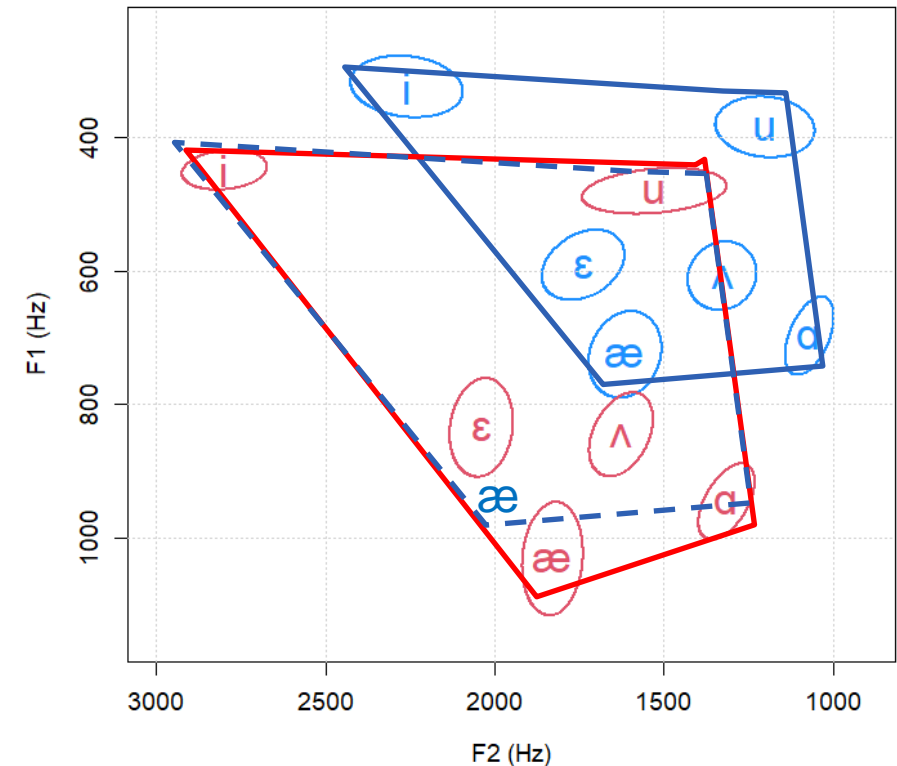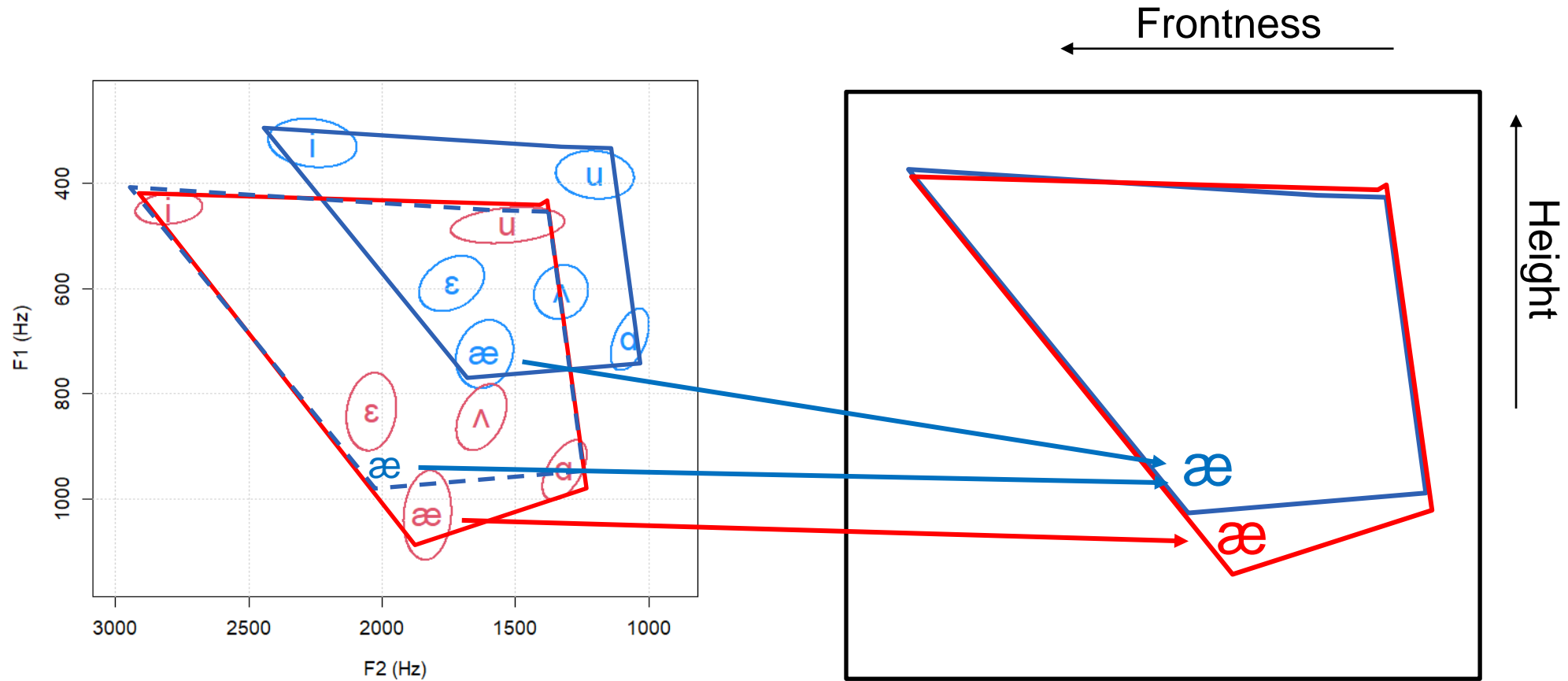
Fant (1966) - A note on vocal tract size factors and non-uniform F-pattern scalings

- Q: Why is the female vowel space a different shape than the male vowel spaces?

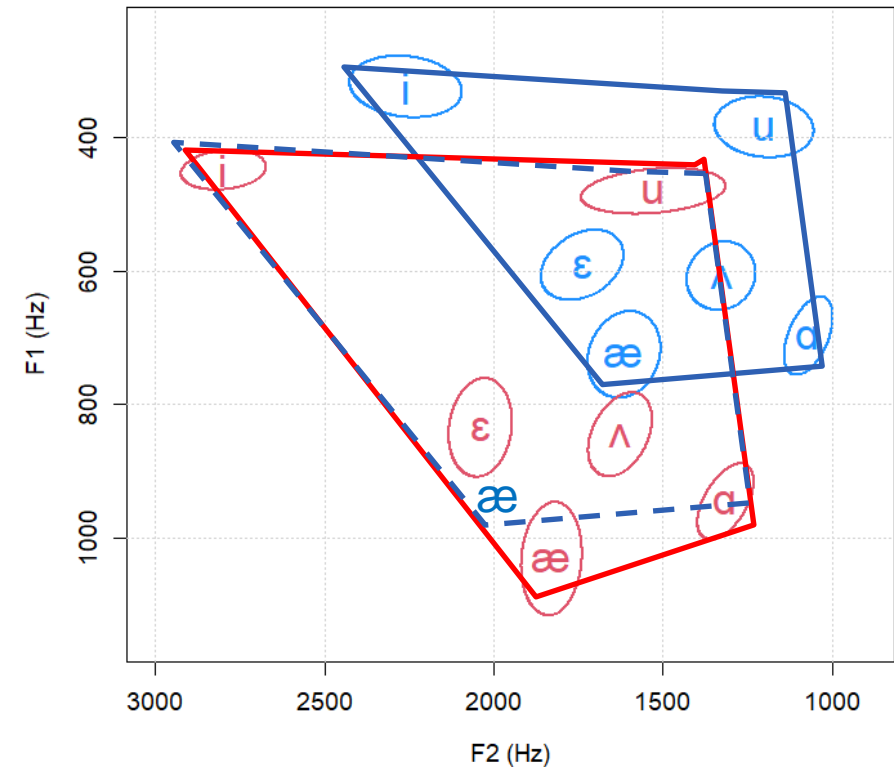- A: Because of the constraints of the female body.
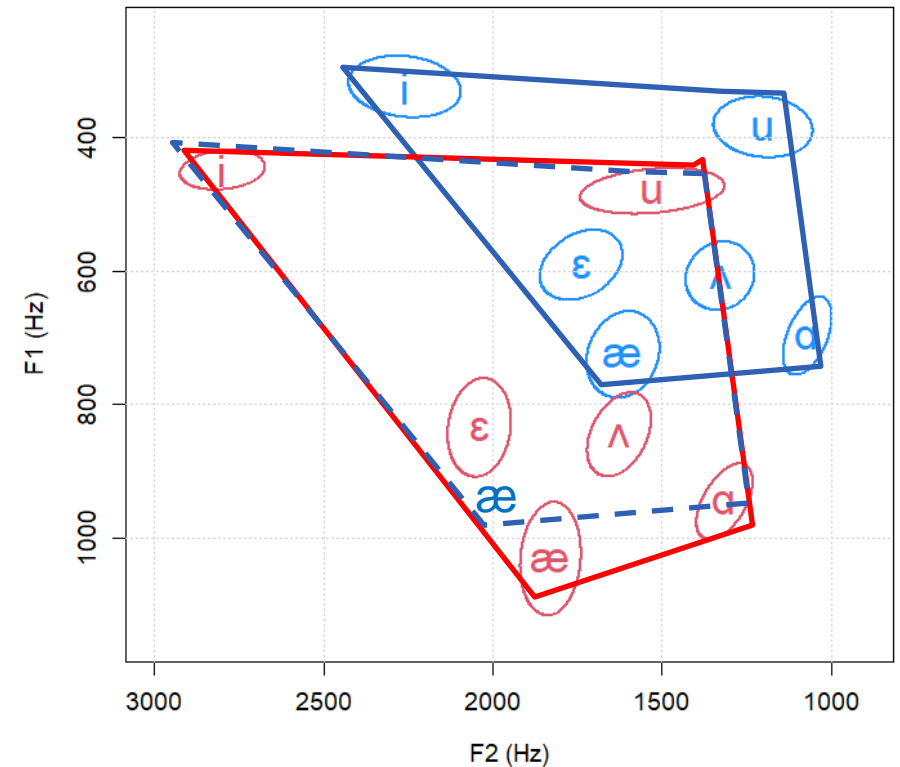
Non-uniform scaling = Phonetically salient

Fant (1966) - A note on vocal tract size factors and non-uniform F-pattern scalings

- Q: Why is the female vowel system *phonetically* different from the male system?

- A: Women produce these *linguistic* differences because of the constraints of the female body.

Sex and Gender

- Sex: Categories based on anatomy/physiology.

- Gender: Categories based on cultural/social practices.

- Both are socially constructed and are historically contingent.

  - e.g., conceptions of 'biological' sex 100 years ago did not include hormones (Sans 2017).

Opposing View on Sex and Gender

1. Essentialist:

   ○ Direct link between anatomy/physiology gender (sex == gender).
   ○ "Women speak in a feminine manner because it is inherent to their female bodies".
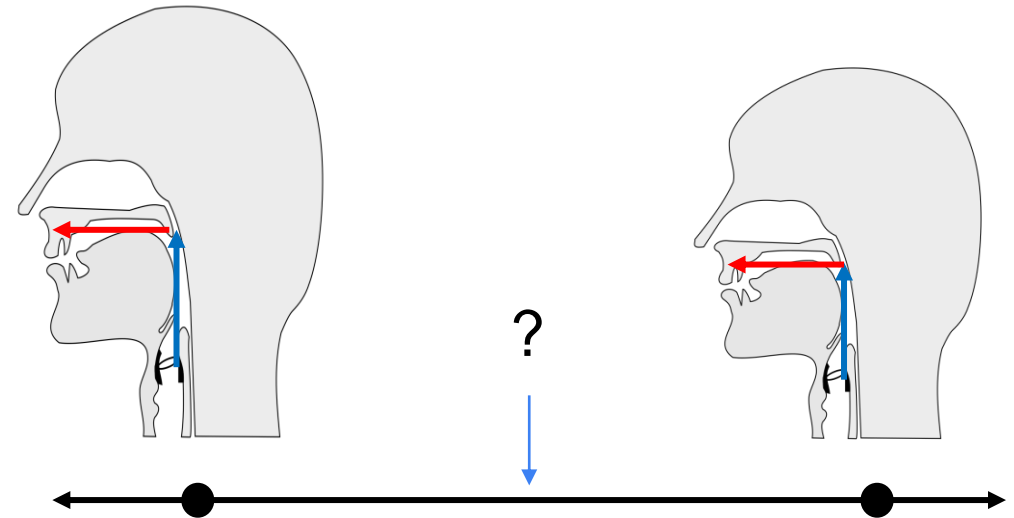
2. Constructivist:

   ○ No deterministic link between gendered speech and female bodies (sex != gender).
   ○ "To the extent that women produce 'feminine' speech, it is due to socialization and culture".

Fant (1966): Essentialist ideology

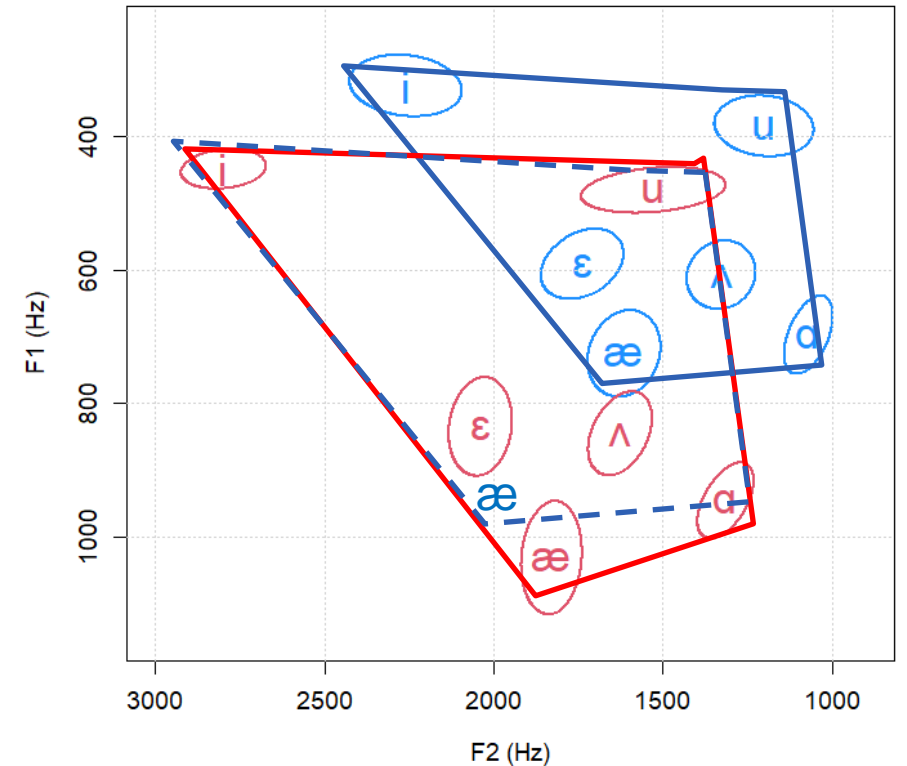The framing and conclusions conveys
the following perspective:

1. Discrete, easily separable male and
   female anatomical differences.

Fant (1966): Essentialist ideology

The framing and conclusions conveys
the following perspective:

1. Discrete, easily separable male and
   female anatomical differences.

2. Males set the standard; women try
   to copy it.

Fant (1966): Essentialist ideology

The framing and conclusions conveys the following perspective:

1. Discrete, easily separable male and female anatomical differences.

2. Males set the standard; women try to copy it.

3. The female copy fails because of 'constrained' female body.
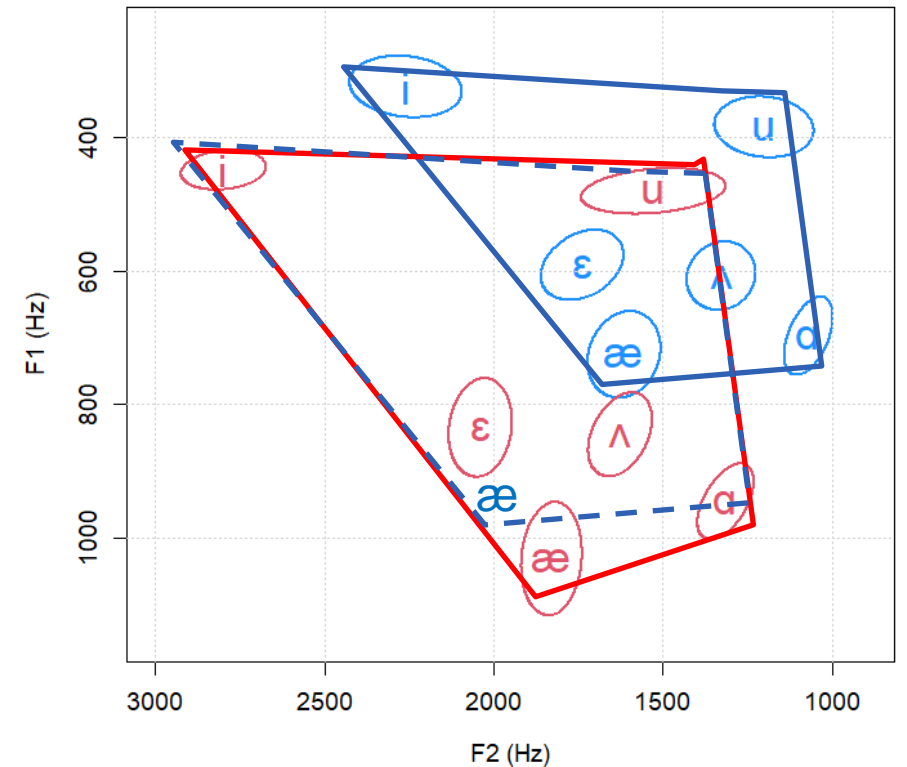
Main Arguments

1) There are <u>no</u> necessary differences in vocal tract anatomy between male and female speakers.

2) The anatomical differences between male and female speakers are <u>not</u> causally related to the linguistic properties of male and female speech.

3) Let's update how we think about normalization, from "removing anatomical variation" to "representing phonetic variation".

Part 1: There are <u>no</u> necessary differences in vocal tract anatomy between male and female speakers.

Fant (1966) - A note on vocal tract size factors and non-uniform F-pattern scalings

- "The main physiological determinant of the specific deviations from the average rule is that the ratio of pharynx length to mouth cavity length is greater for males than for females and that the laryngeal cavities are more developed in males." (p. 22)
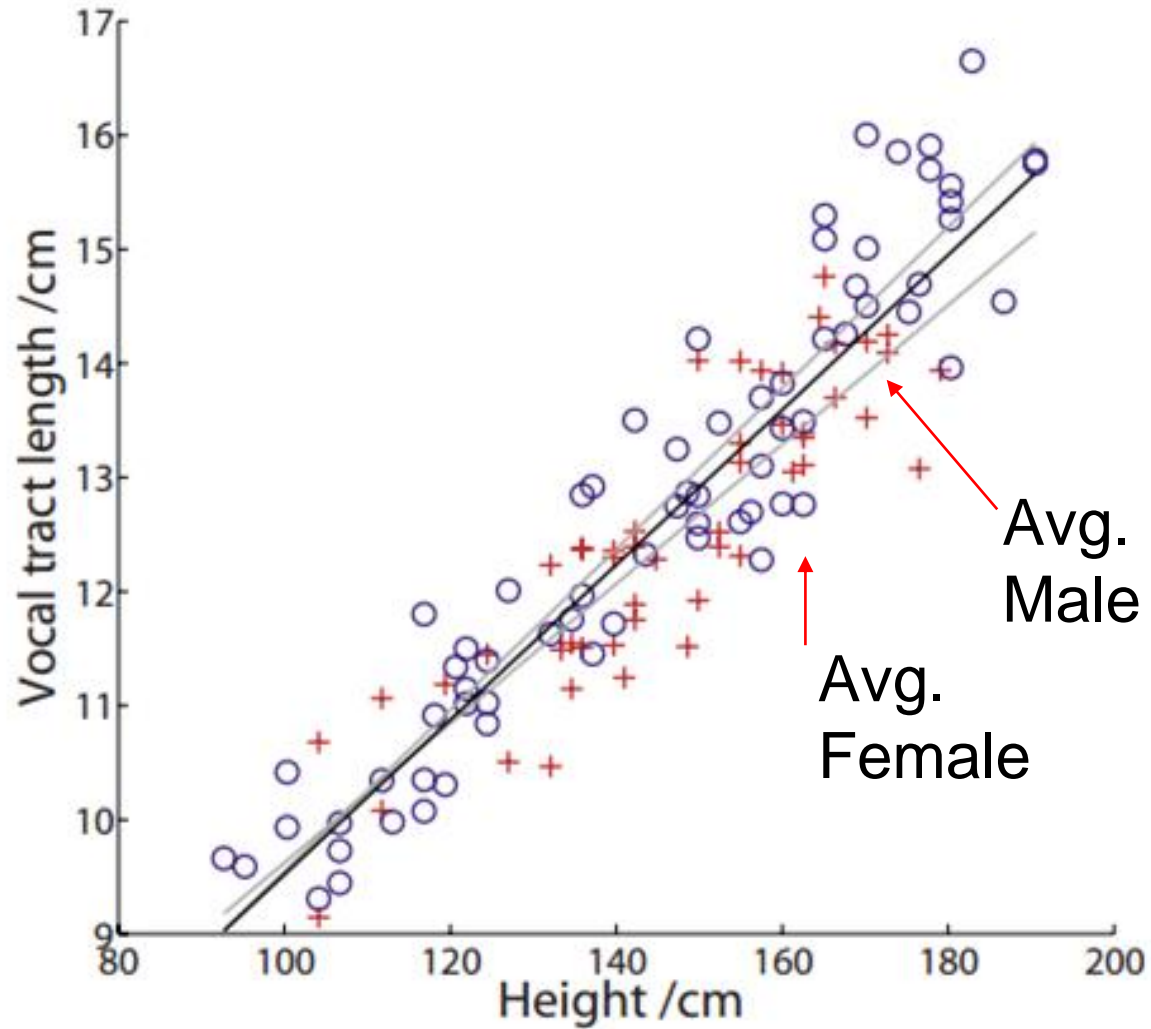
$$\frac{\text{Male Pharyngeal Cavity Length}}{\text{Male Oral Cavity Length}} > \frac{\text{Female Pharyngeal Cavity Length}}{\text{Female Oral Cavity Length}}$$

Male P/O > Female P/O
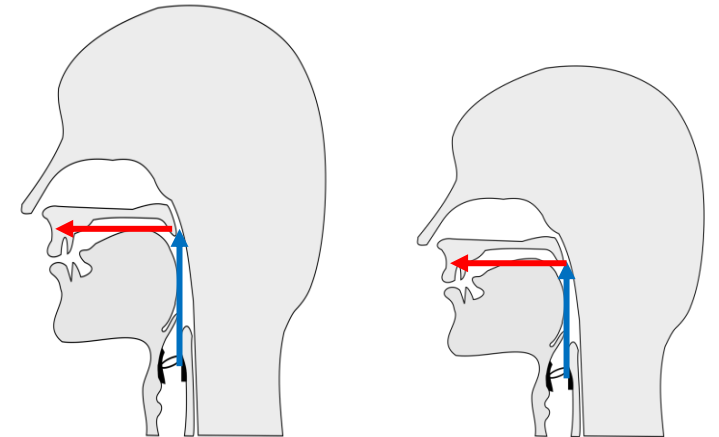
Simplified: Men have proportionally smaller heads (or longer necks) than women.

VTL ~ height

Turner et al. (2009)
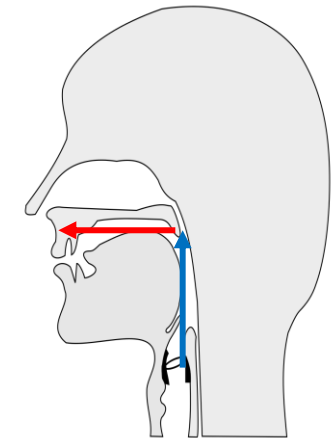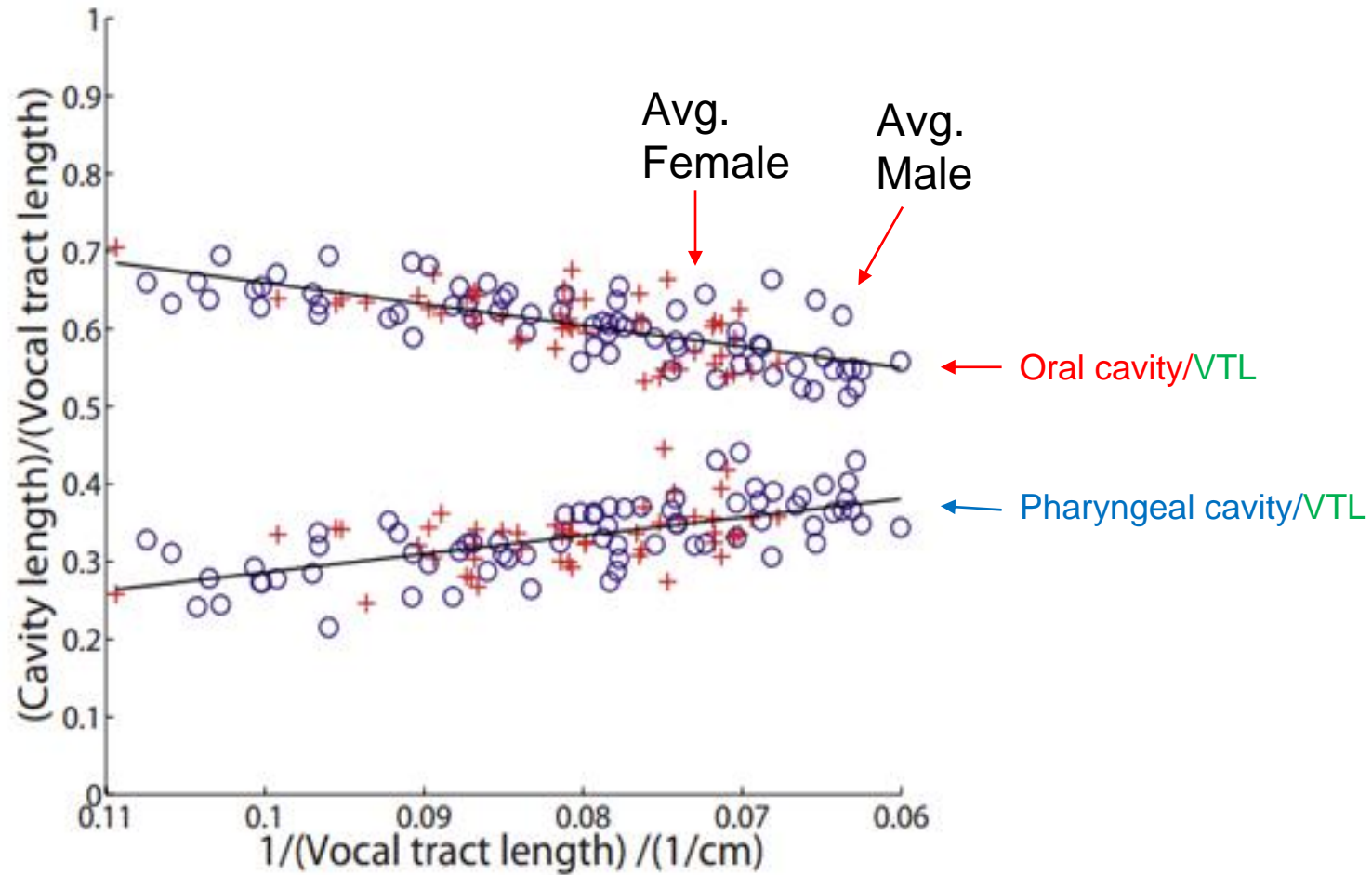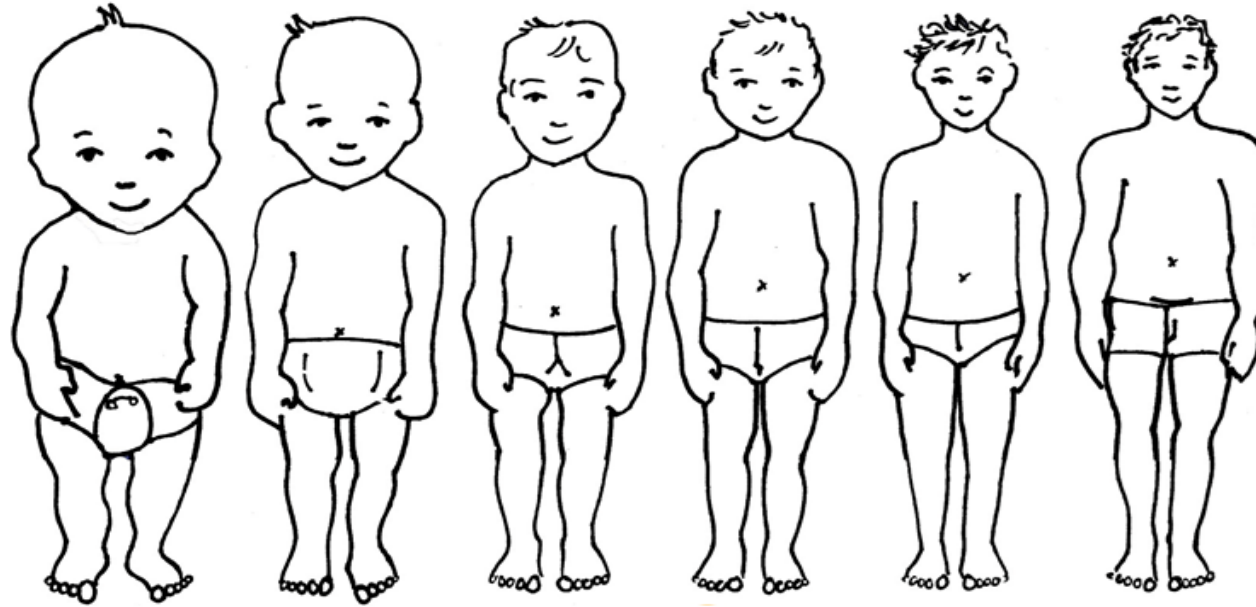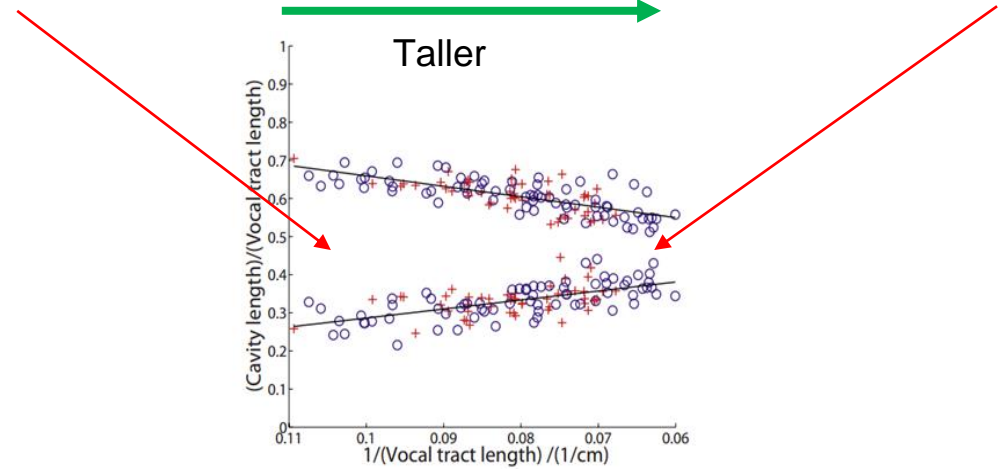Fitch and Geidd (1999)

VT Cavities ~ height
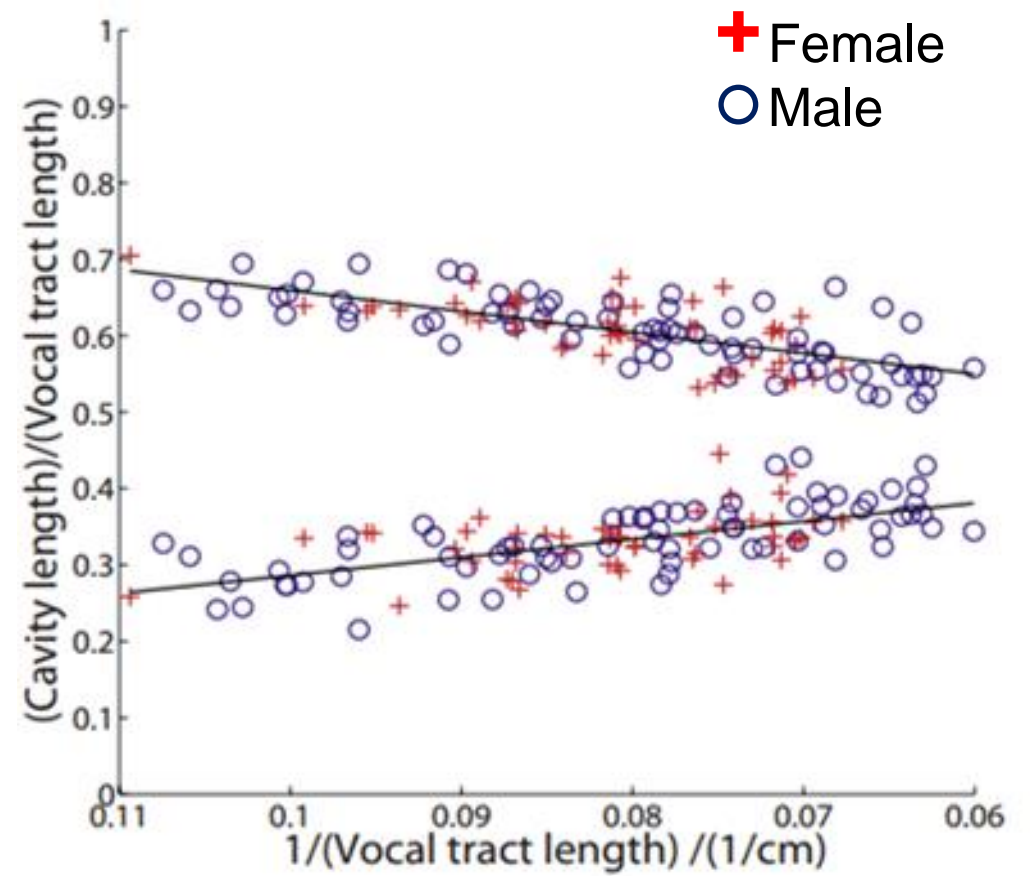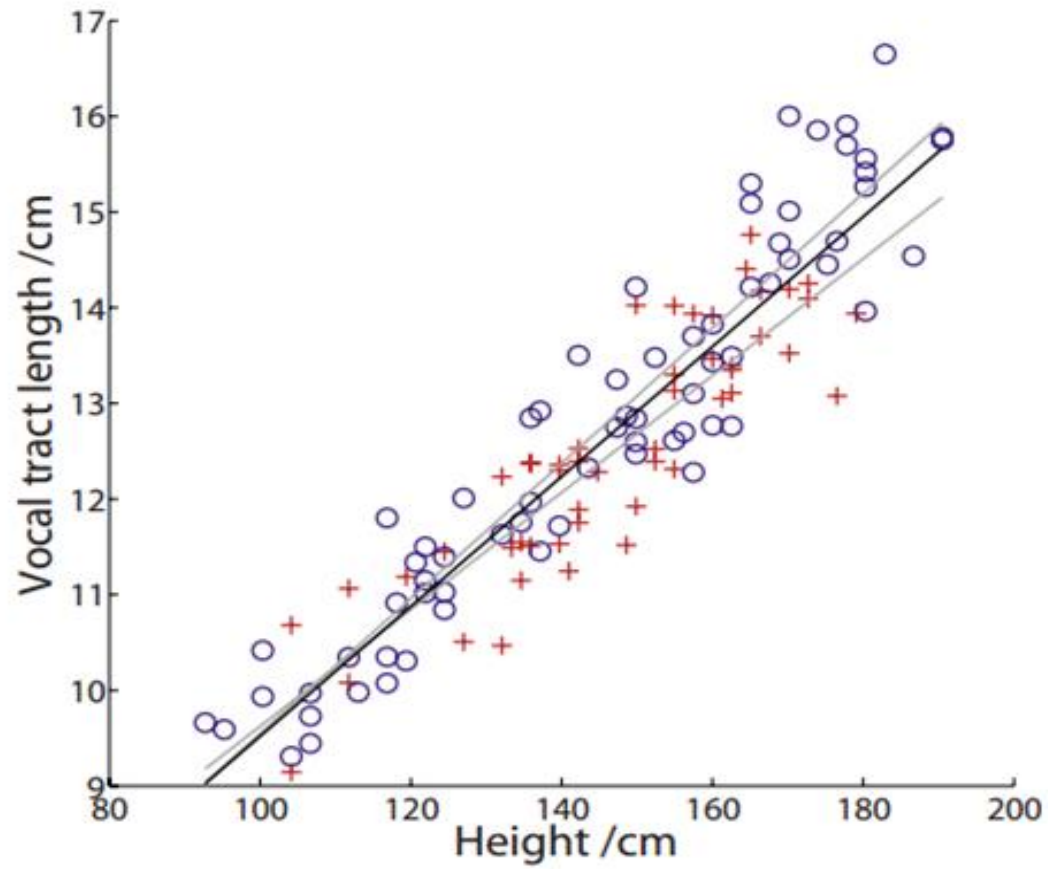

Turner et al. 2009
Fitch and Geidd 1999

Head ~ height

Taller

Height, not sex!

Part 2: Anatomical differences between females and males are <u>not</u> causally related to the phonetic properties of gendered speech.

Gendered Differences = Phonetically salient
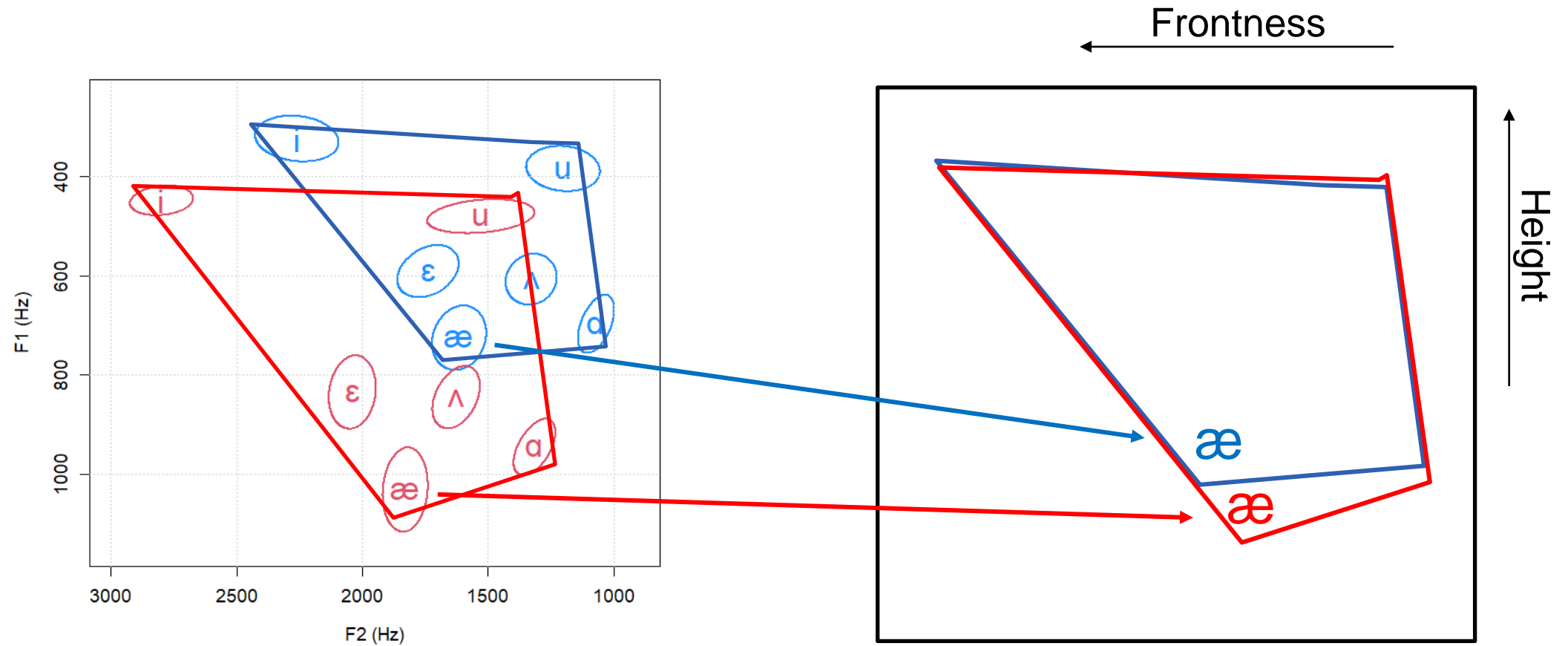
Gendered Dialect

"male/female vowel formant differences are variable across languages and dialects of the same language. […] these data suggest that gender differences are not solely due to vocal tract anatomical differences between men and women"

"the overall conclusion is clear. People (perhaps especially men) perform gender."

Johnson (2006), p. 487
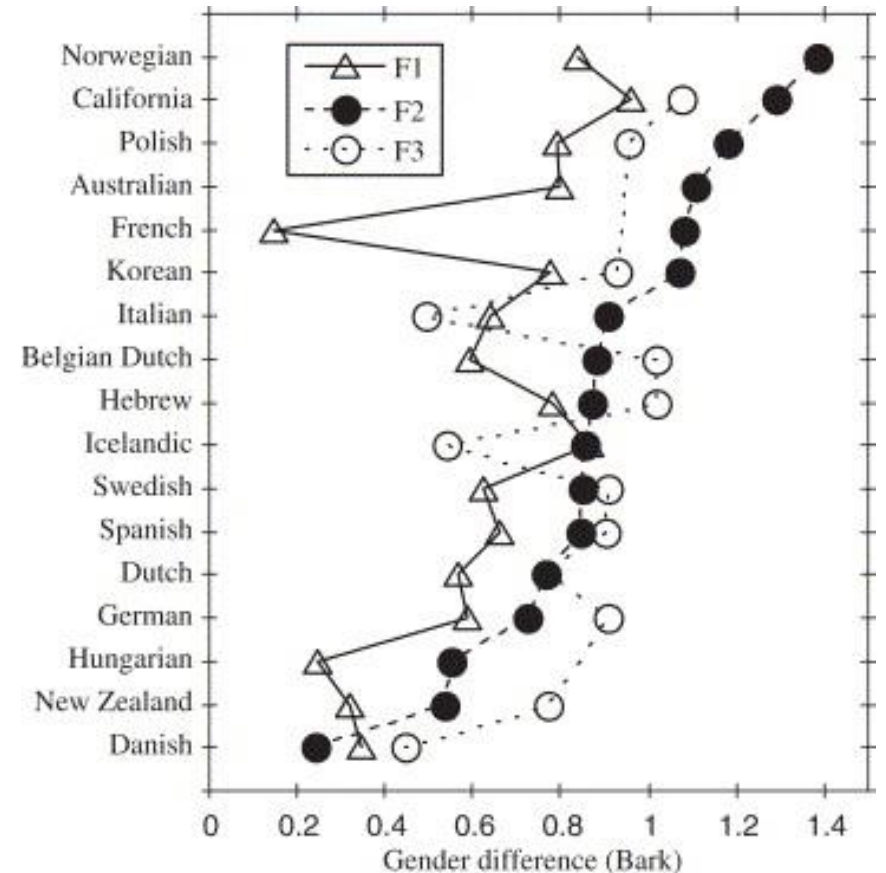
Gendered Dialect

- <u>Not</u> denying the existence of gendered linguistic differences.

- Denying the necessary connection of linguistic variation to anatomical differences between men and women.

Frontness

Height

æ

æ
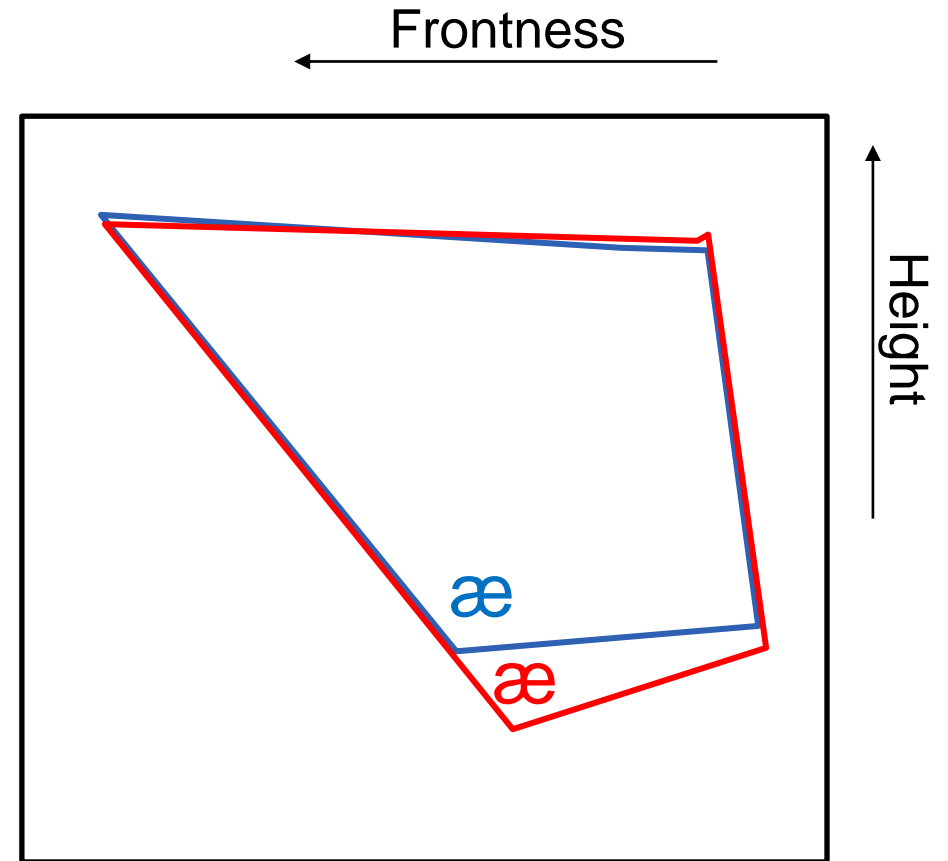
Inviolable principle: Linguistic Universality

- Any human that can reach full competence in language X could've reached full competence in language Y.

- Alternate version: The language faculty is not language-specific.

Linguistic Universality: Gendered Variation

Since any speaker can potentially reach full competence in any L1…..

- Could women learn a language that is just like the male targets (i.e., with a slightly higher /æ/)?

  - No: Women can't be native speakers of at least some languages.

  - Yes: Gendered *phonetic* variation should not be dismissed as 'anatomical'.

Male ("target")

Frontness

Female
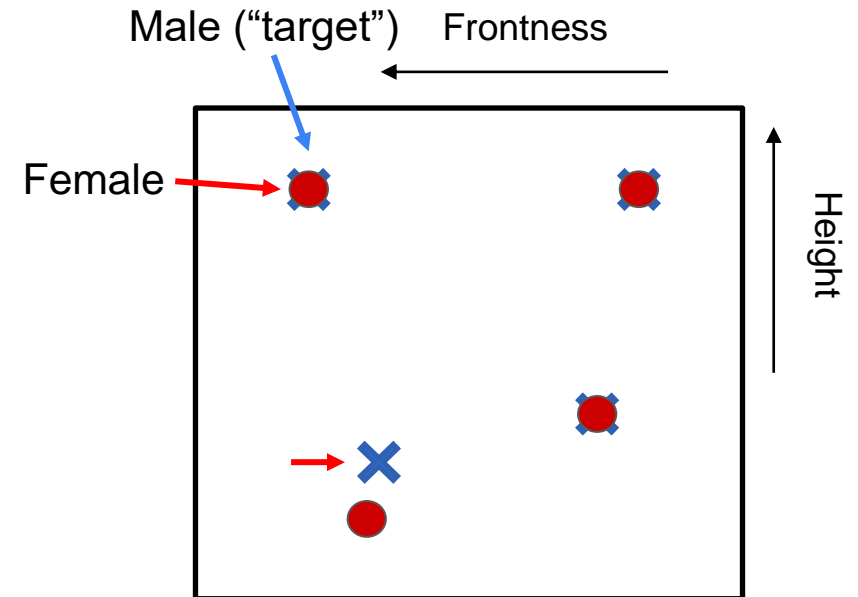
Height

Phonetic vs. Acoustic

- Vowel sounds have a dual reality:

  1. Physical acoustic characteristics (measured in dB and Hz).
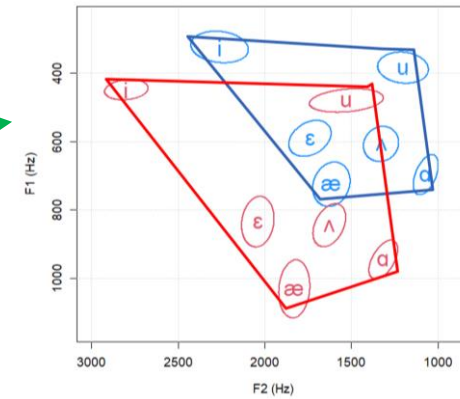
  (1)

  2. Linguistic/phonetic properties, there consistent of vowel quality, height and frontness.
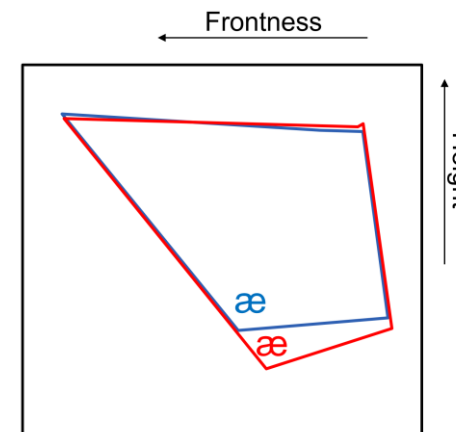
  (2)

- Human bodies can limit (1) but <u>not</u> (2).

Linguistic Universality: Gendered Variation

- Can these female speakers produce the dashed blue system?

- The answer is obviously <u>yes</u>.

- The gendered difference should not be attributed to "anatomy".

"Anatomical" Variation: A red herring

- "[A]s is well known, an infinite number of shapes of different lengths are realizable for a given set of formant frequencies and bandwidths" (Wakita 1977, p. 190).
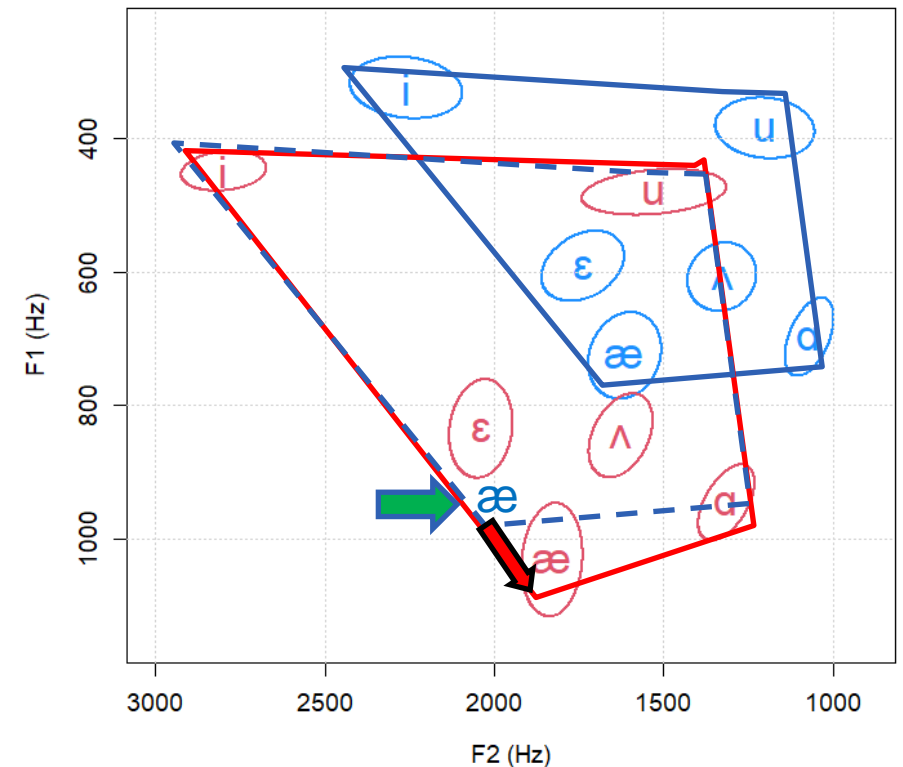
- The bodies of identical twins can produce substantially different formant patterns for the 'same' vowels (Zuo & Mok 2015).

- Conclusion: A single human body can produce arbitrarily different formant patterns, and an infinite number of anatomies and gestures can be combined to yield any given formant pattern.

Part 3: Let's update how we think about normalization, from "removing anatomical variation" to "representing phonetic variation".

Vowel Normalization Algorithms

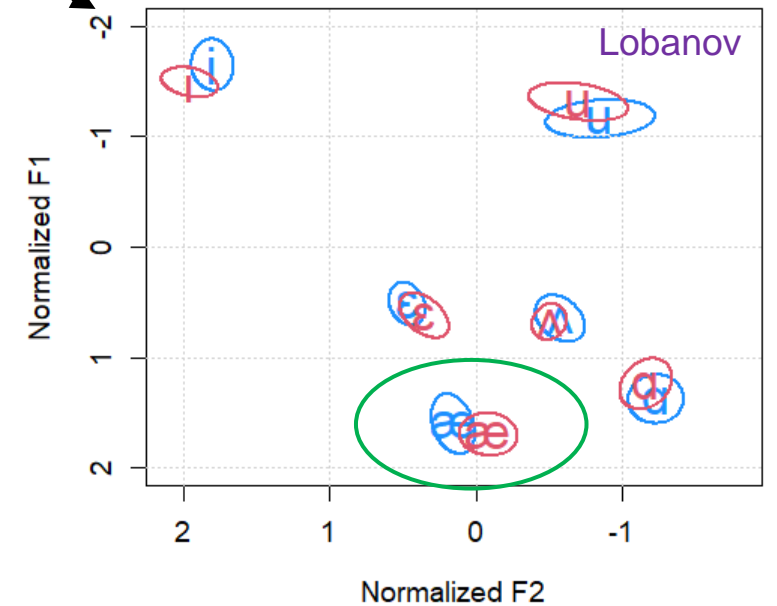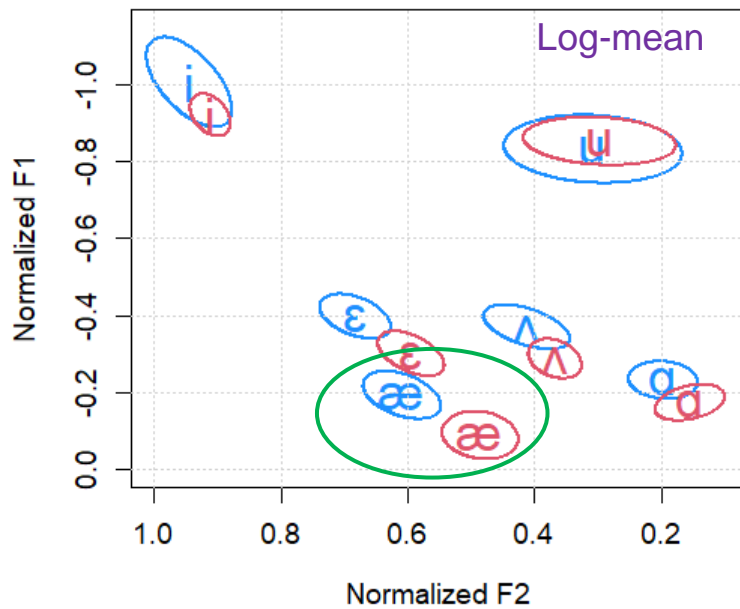- Map from an acoustic representation to a phonetic/linguistic representation.

$$f([F_1, F_2, F_3]) = [N_1, N_2, N_3]$$
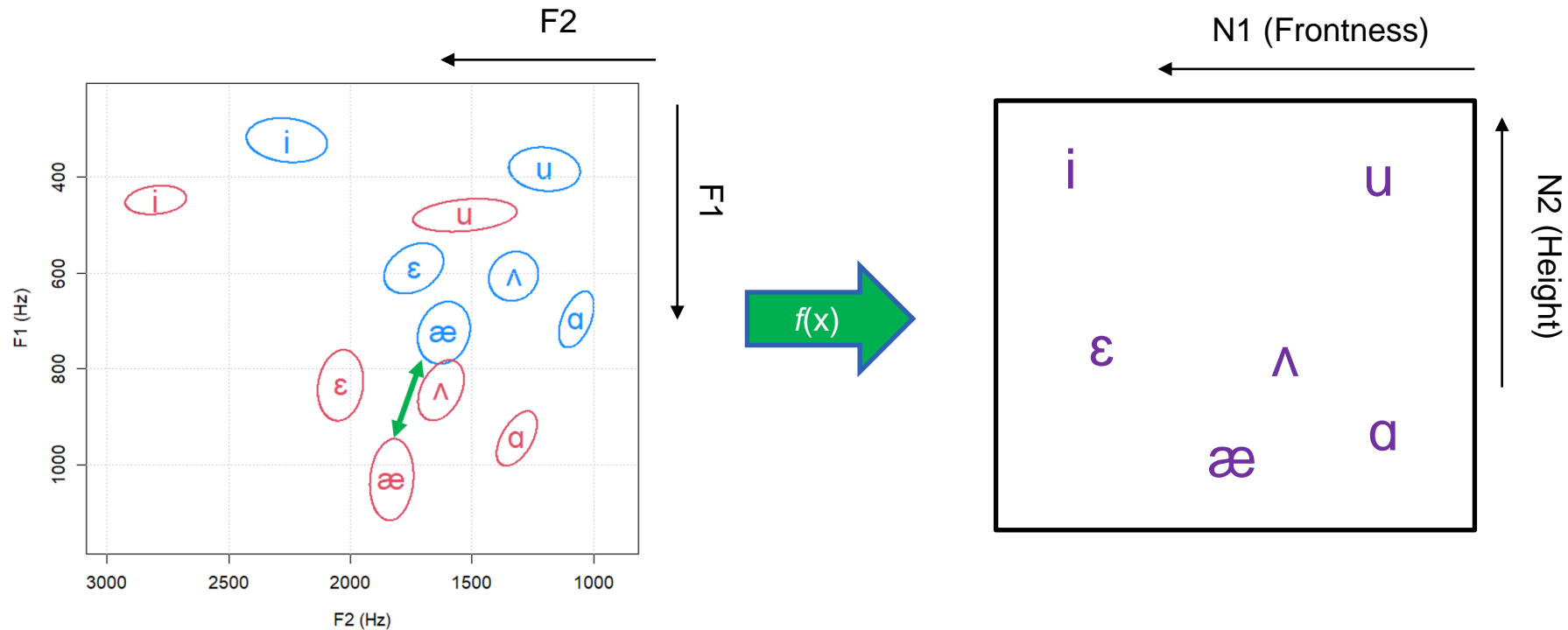
Choosing Normalization Algorithms
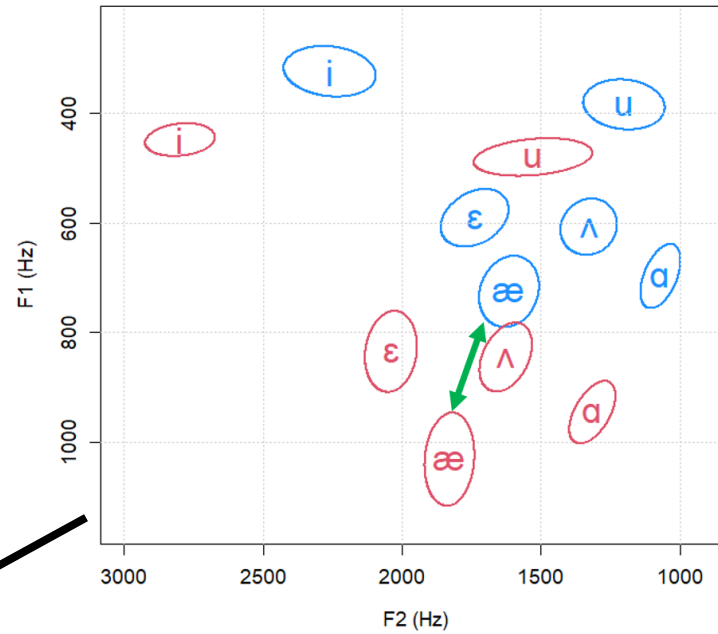
Variation in F1

What causes an observed F1 frequency?

Essentialist:

F1 = Sex + Phoneme   (where sex==gender)

Constructivist:

F1 = Sex + Phoneme   (where sex!=gender)

F1 = (Body + Gender) + Phoneme
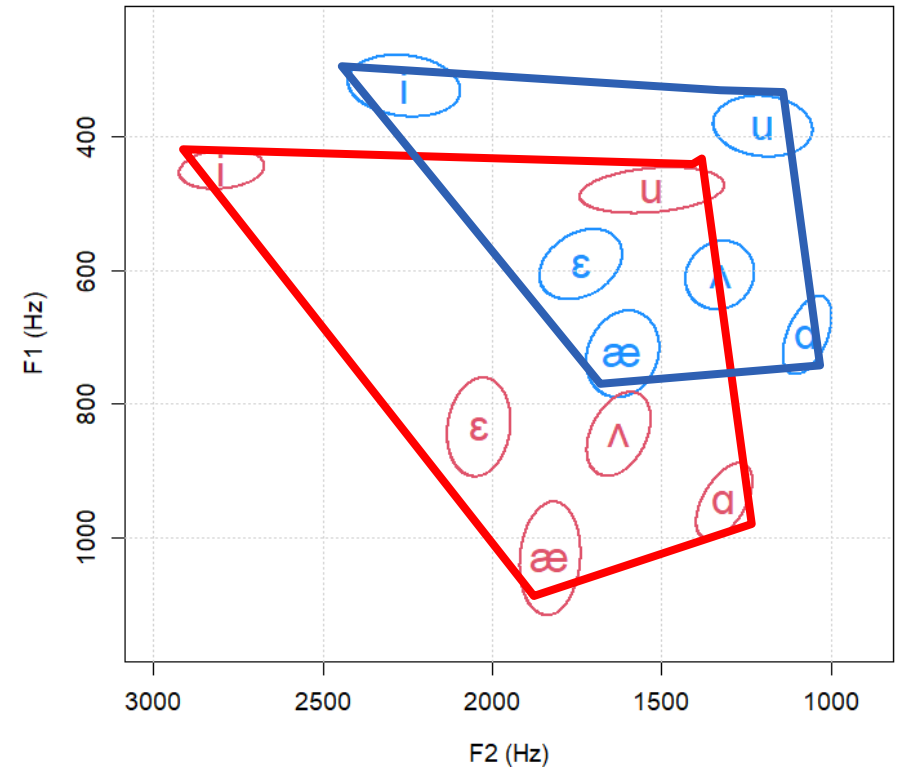
Normalizing variation in F1

What if we can successfully erase (non-phonetic) body differences?

- Essentialist view: Only phoneme information remains.

F1 = Sex + Phoneme   (where sex==gender)
N1 = ~~Sex~~ + Phoneme

- Constructivist view: Phoneme and gender information remains.

F1 = Body + Gender + Phoneme
N1 = ~~Body~~ + Gender + Phoneme

Adank et al., (2004): Erasing gender

- Evaluated normalization methods (in part) on their ability to <u>erase</u> gender information.

- "a procedure is successful at eliminating gender-related anatomical/physiological variation when [the LDAs are] performing at chance level" (p. 3103).

Gender ~ NF1

Essentialist expectation: <u>Impossible</u> to predict gender.
NF1 = ~~Sex~~ + Phoneme
Sex/Gender ~ ~~Sex~~ + Phoneme

Constructivist expectation: <u>Possible</u> to predict gender.
NF1 = ~~Body +~~ Gender + Phoneme
Gender ~ ~~Body~~ + Gender + Phoneme

Voeten et al. (2022) : Erasing individual variation

- "There are two goals to speaker normalization: retaining (socio)linguistic information (here, vowel identity and regional origin) and normalizing anatomical differences (here, speaker sex and individual differences between speakers matched on sex, age, and region)."

- "We consider a normalization method to perform well if it maintains differences between vowels and regions, normalizes differences between the two speaker sexes, and similarly normalizes individual speaker differences." Voeten et al. (2022)

F1 = Phoneme + Dialect + Sex + Individual differences
NF1 = Phoneme + Dialect + ~~Sex + Individual differences~~
NF1 = Phoneme + Dialect + ~~Sex + (Gender + SF1 + SF2 + … + SFN)~~

How did we get here?

1. Early on, a focus was placed on anatomical variation between men and women.

2. 'Anatomy' was identified as the source of differences between men and women.

3. 'Anatomy' was extended to generally explain much between-speaker variation.

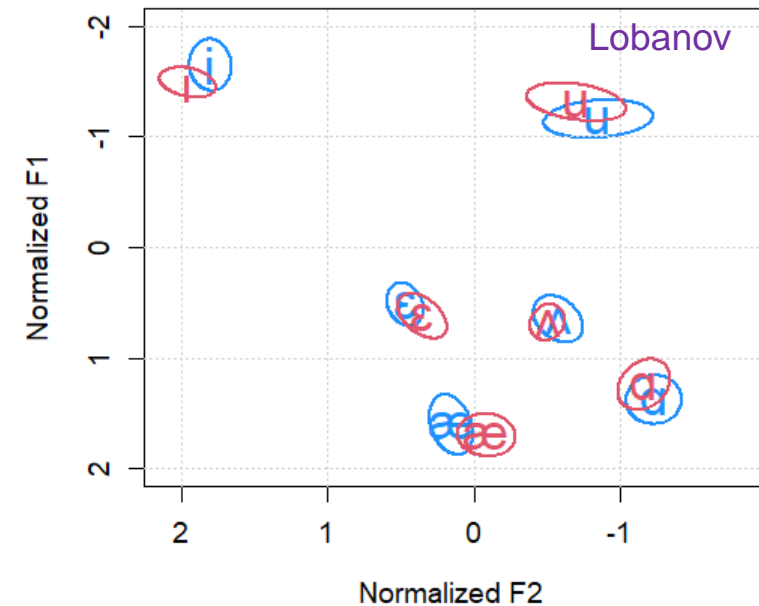4. This has led to wide-scale <u>overnormalization</u> of linguistic data.

Overnormalization

- Overnormalization: The erasure of legitimate phonetic and social information from our data.
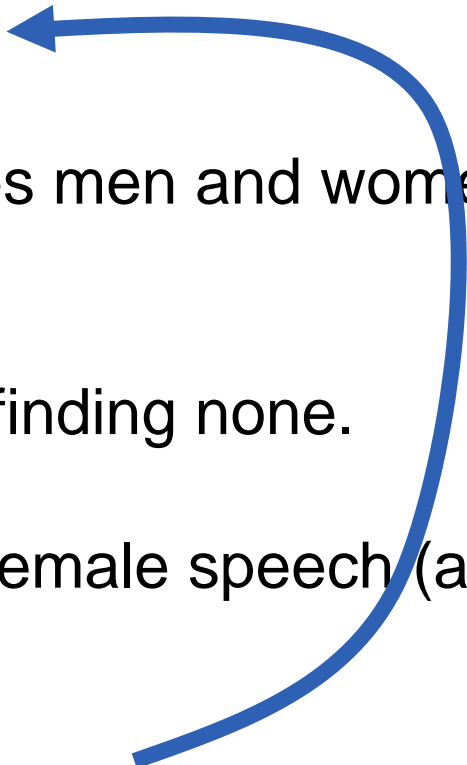
NF1 = Phoneme + ~~Body~~ + Gender + SF1 + SF2 + … + SFN
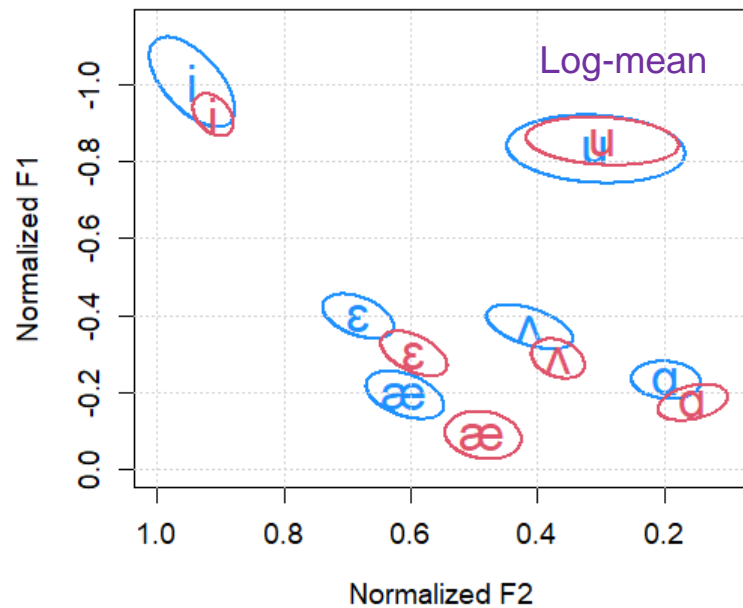
NF1 = Phoneme + ~~Body + Gender + SF1 + SF2 + … + SFN~~

Begging the question: Gender variation

- Begging the question of a lack of gender differences:

1. We collect data to see if men and women are different

2. We normalize, using a method selected because it makes men and women as indistinguishable as possible.

3. We then test for a difference between men and women, finding none.

4. We report that there is no difference between male and female speech (apart from the 'anatomical' differences we erased of course).

5. Future researchers continue with the assumption that there are no male/female differences apart from those related to anatomy.

Reproducing ideologies

- Eckert (2014): "the gender binary is maintained by the continued 'doing' of gender – every time a person uses a single-sex restroom, he or she is reproducing that binary" (p. 529).

- When we normalize with the intention of removing gender differences in speech because they are 'anatomical', we reproduce essentialist gender ideology unintentionally.
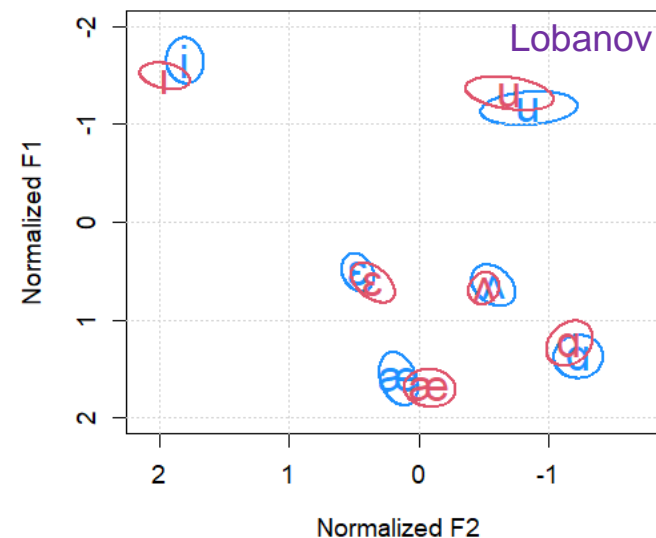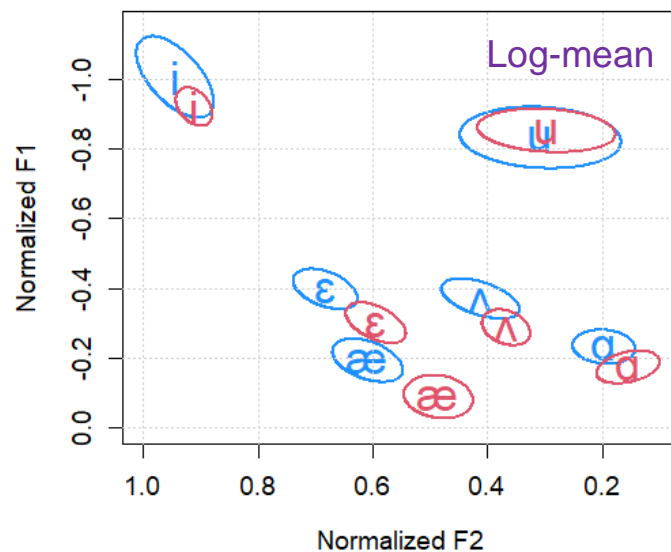
The way forward: Focus on phonetics

- Overnormalization is akin to overfitting: Removing 'too much' variation.

  - Without a phonetic (not phonemic) 'ground' truth, this is bound to happen.

- The only way to avoid overnormalization: Focus on preserving phonetic information, erase only non-phonetic acoustic variation.

- Forget trying to infer the *cause* of variation:

  - Focus on erasing non-phonetic variation, regardless of the source.

  - Focus on representing phonetic variation, regardless of the source.

Thank you!!

References

Eckert, Penelope. "The problem with binaries: Coding for gender and sexuality." *Language and Linguistics Compass* 8.11 (2014): 529-535.

Fant, G. (1966). A note on vocal tract size factors and non-uniform F-pattern scalings. *Speech Transmission Laboratory Quarterly Progress and Status Report*, *1*, 22-30.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *106*(3), 1511-1522.

Turner, R. E., Walters, T. C., Monaghan, J. J., & Patterson, R. D. (2009). A statistical, formant-pattern model for segregating vowel type and vocal-tract length in developmental formant data. *The Journal of the Acoustical Society of America*, *125*(4), 2374-2386.