Lingual implementations of non-lingual features: cross-linguistic data from Turkish, Japanese, and French

Adam Baker, Alina Twist, Yuko Watanabe, Diana Archangeli

University of Arizona

Ultrafest IV
New York University
September 29, 2007

What are distinctive features ?

1. Units of psychological computation
2. Descriptors of sound patterns
3. Phonetic units

- What is the proper phonetic interpretation of features ?
  - Articulatory
  - Acoustic
  - Perceptual
- The labels we use (high, back, ATR, round, etc.) are *articulatory*.

  "The articulator-bound features specify which articulators are active in producing specific contrasts, and prescribe how these articulators are to be configured." (Stevens 1998 : 249).

Looking at articulatory features

- The complex structure of the tongue makes any articulatory interpretation of [high], [low], [back], or [ATR] difficult.
- But there are "easier" articulatory features : [round], [nasal], [voice].
- These should unambiguously mean (for positive values) "round the lips," "lower the velum," and "adduct the vocal folds."
- Note that the mechanical linkages between the articulators in question and the tongue are negligible for the tongue.

- Simple question : are these actually non-lingual features ?

Three languages

- We're looking at the phonetic implementation of these features in three languages in which they are phonologically active.
- [round] in Turkish
  - (Classic) height-dependent roundness harmony in suffixes
- [voice] in Japanese
  - Approximately : devoicing of {i,ɯ} between voiceless consonants, and between a voiceless consonant and a final word boundary.
- [nasal] in French
  - Alternation between oral and nasal vowels in feminine and masculine adjectives (examples from Tranel 1987)
  - divine [divin] / divin [divɛ̃]
  - brune [bryn] / brun [brœ̃]
  - certaine [sɛrtɛn] / certain [sɛrtɛ̃]

Common protocols

- Subjects sat in an ophthalmic exam chair in a quiet room.
- Palatron head tracking system was used.
- Subjects read prompts from a computer screen.
- Used a mouse to click through the prompts at their own pace.
- Items presented in randomized blocks.
- Each prompt appeared five times.

Turkish

- Frame : Lütfen "_____" de.
- Five minimal pairs for [i]/[y], [ɯ]/[u], and [e]/[ø].
  - Almost all monosyllabic
  - Frame did not elicit a suffix
  - Example : is/ys

  Note : there is not an [e]/[ø] alternation in Turkish
- Two subjects

Japanese

- Frame : Mouichido _____ to ittekudasai.
  Gamen ni _____ ga mieta.
  (presented in katakana)
- $C_1V_1C_2V_2C_3V_3$
  - Real words
  - V_2 is {i,ɯ}
  - One or both of C_2 and C_3 voiceless.
  - Pitch-accent conducive to devoicing of V_2
  - Example : mosɯto/mozɯto
- One speaker of the Tokyo dialect

French

- Frame : Dites « _____ » s'il vous plaît.
- Minimal pairs to elicit various contrasts including :
  - ɛ/ɛ̃
  - o/ɔ̃
  - ɑ/ɑ̃
- Example : mɛs/mɛ̃s
- One subject

Common protocols

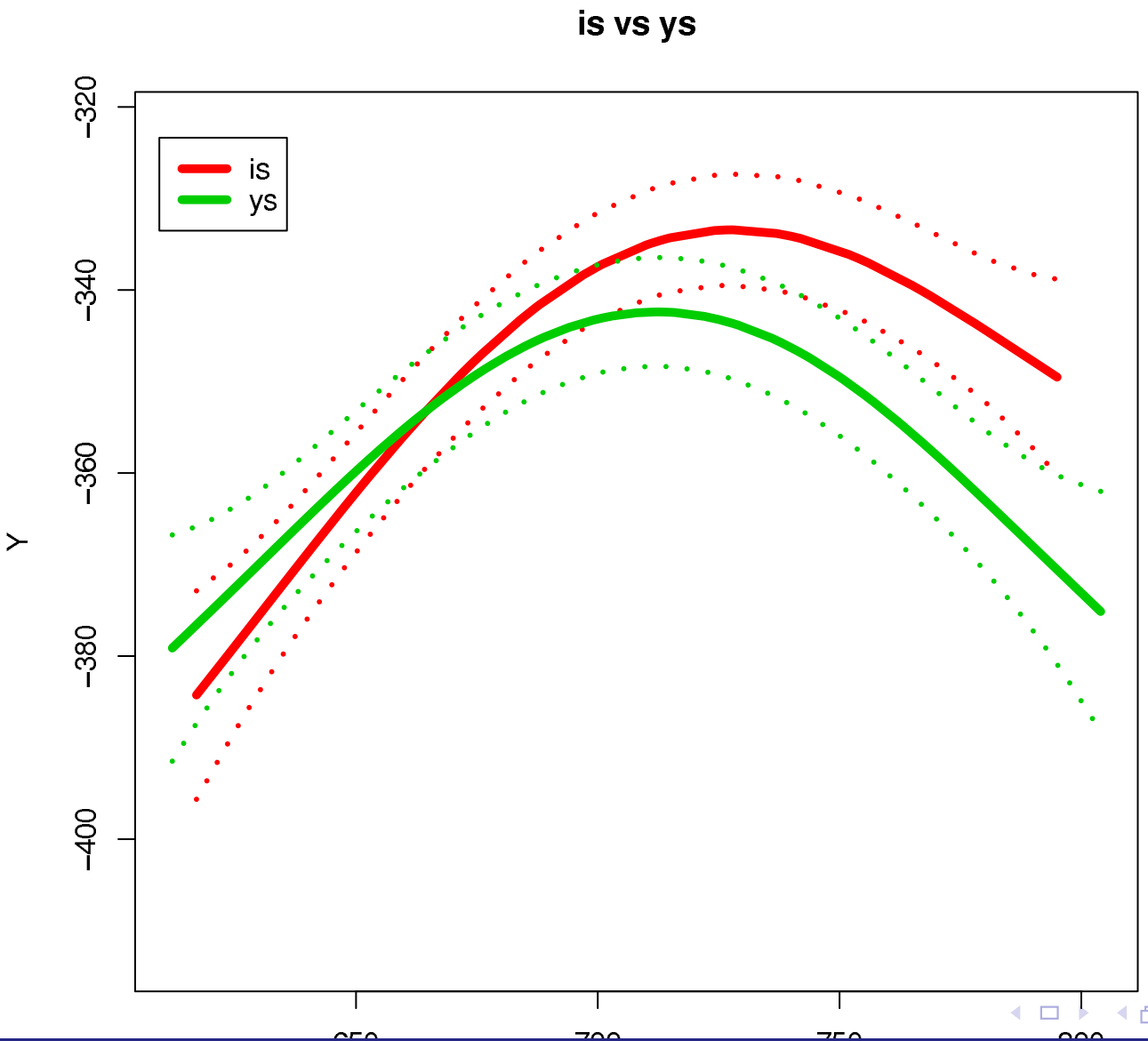
- One frame per vowel selected with articulatory criteria.
  - extreme of the lingual gesture
- Tongue traced with Palatoglossatron (Baker 2006)
- Adjustments for head and transducer movement made with Palatron algorithm (Mielke et al. 2005).
- For more detail, Google "APIL wiki."

- SSANOVA performed on minimal pairs to test for significant differences.

Turkish Results

- Speaker 1
  - 5/6 significant for [i]/[y]
  - 4/6 significant for [ɯ]/[u]
  - 5/5 significant for [e]/[ø]
- Speaker 2
  - 4/5 significant for [i]/[y]
  - 4/4 significant for [ɯ]/[u]
  - 6/6 significant for [e]/[ø]

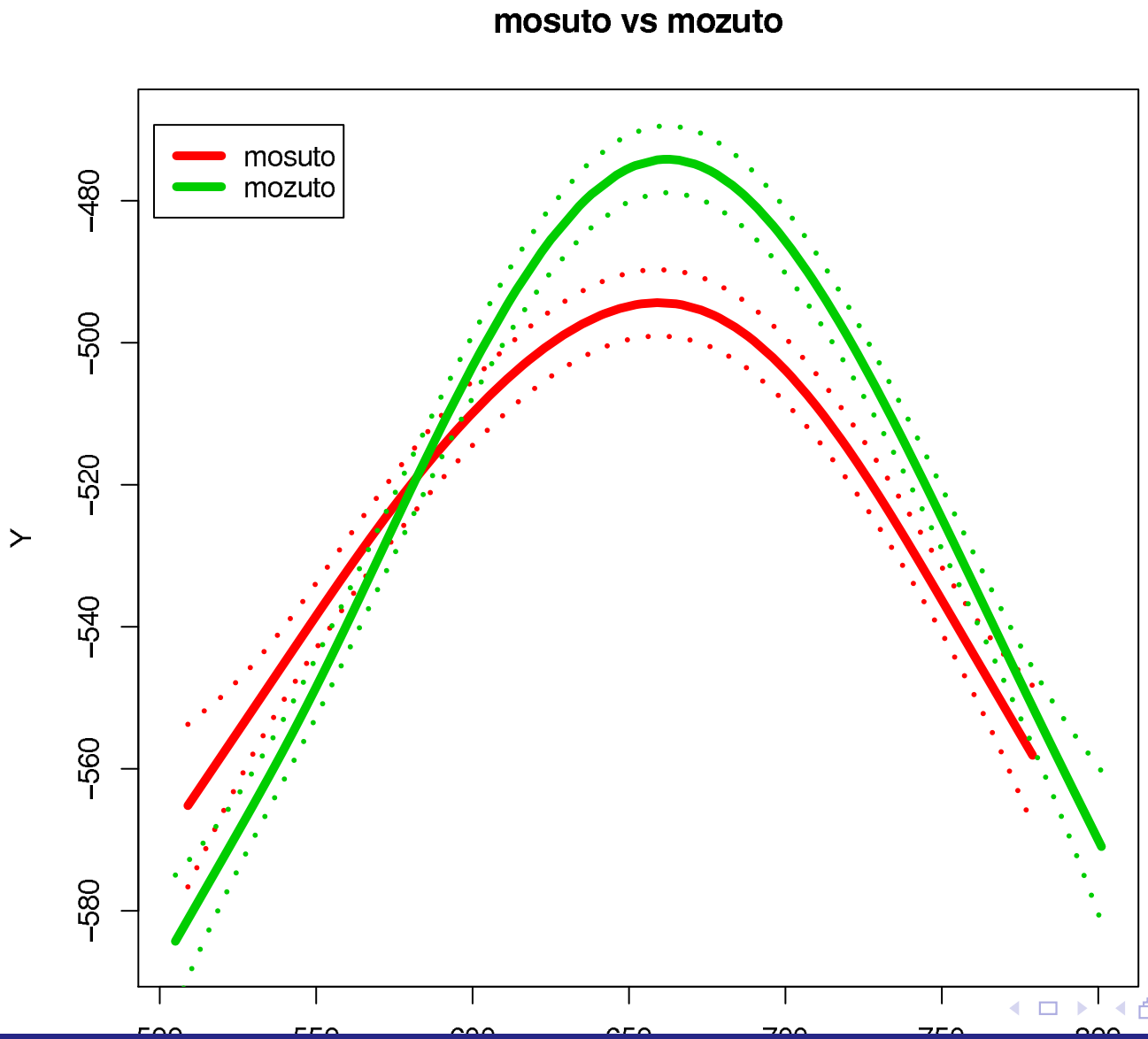
Turkish Results



Japanese Results

- Speaker 1
  - 6/6 significant for [i]/[i̟]
  - 5/6 significant for [ɯ]/[ɯ̟]

Japanese Results



French Results

- Speaker 1
  - 6/6 significant for [o]/[õ]
  - 2/4 significant for [ɑ]/[ɑ̃]
  - 0/2 significant for [ɛ]/[ɛ̃]
- Low confidence in these results

  (Note only 2/4 sig. for [ɑ]/[ɑ̃])

Important qualifications

- Small number of subjects (to be addressed)
- Overly simple statistical model
  - Equivalent of 20 t-tests ?
  - Better to have phonetic context as a factor in the statistical model
- But, Robinson Crusoe only needed the one footprint...

Conclusions

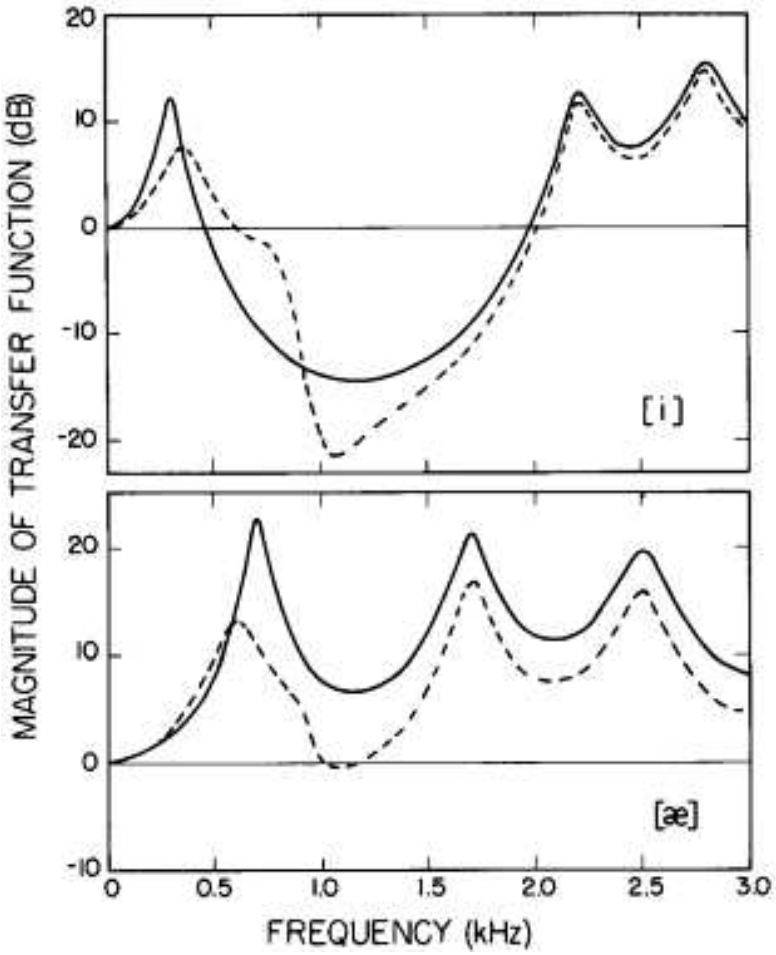
- [+*round*] doesn't just mean "round the lips"
- [−*voice*] doesn't just mean "abduct the vocal folds"
- There are different lingual postures as well.
- No straightforward articulatory interpretation

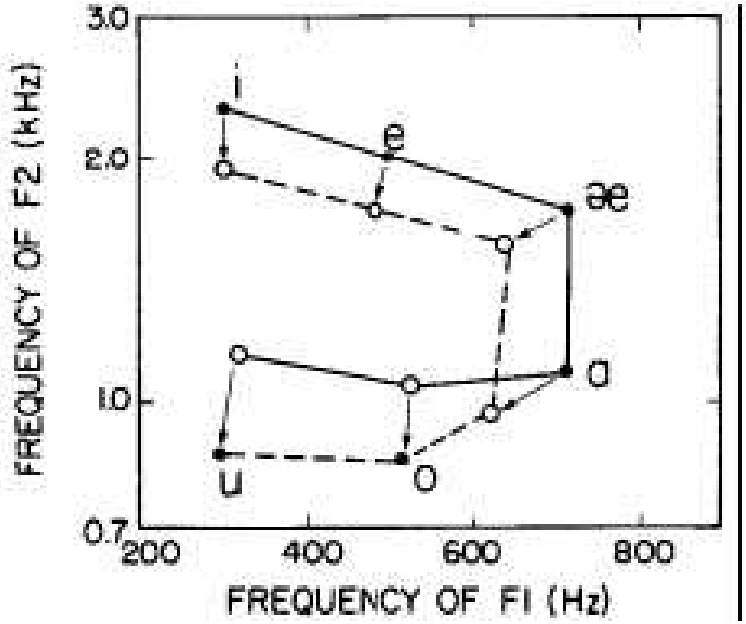
- Two possible theoretical consequences
  1. Features lose simple articulatory interpretation
  2. Bundles of features, rather than individual features, receive phonetic interpretation.

Conclusions

1. Features are not simply articulatory.
   - Might just call [*high*], [*back*], and [*round*] as [*a*], [*b*], and [*c*].
   - Or, look for acoustic interpretation
     - But we know that the acoustic implementation of distinctive features is not simple either.
     - Voiceless vowels have been shown by Ogasawara (2006) to be quite distinct from voiced vowels — no formants present.
     - The effect of nasalization varies by vowel.
     - The effect of a lip constriction varies by vowel.
     - The following graphs are from Stevens (1998).
   - Or, look for a perceptual interpretation.

Effects of nasalization (left) and rounding (right)

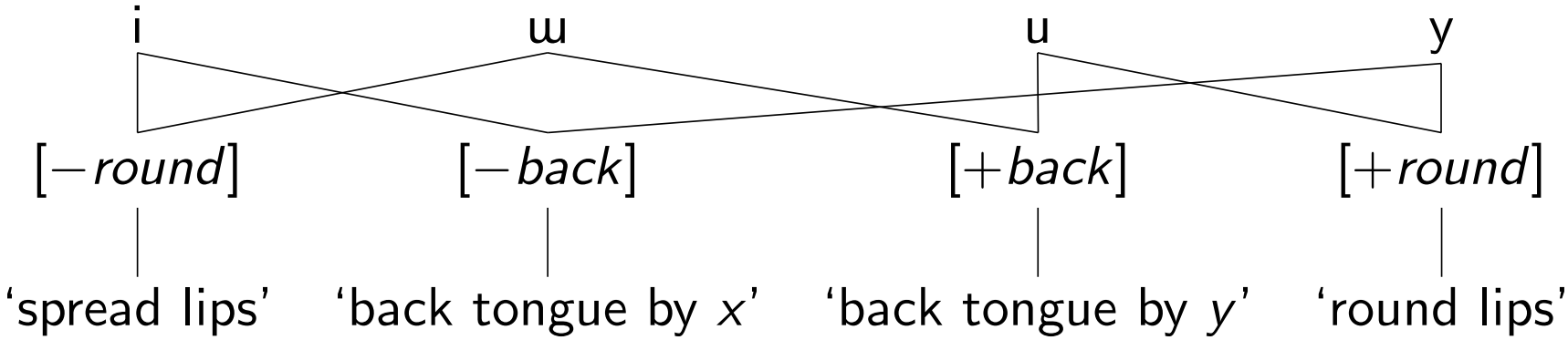




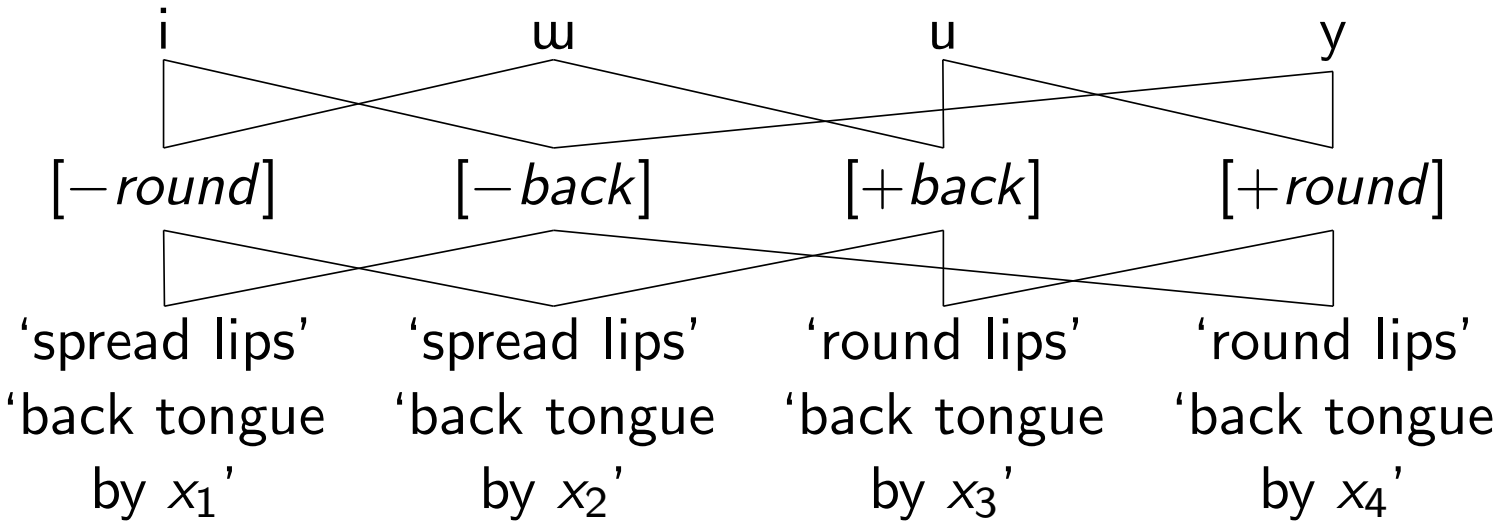
Conclusions

2. We need to interpret bundles of features
   - To abuse some terms already in use...
     - Context-free : each feature is independently interpreted.
     - Context-sensitive : a feature is interpreted based on the values of other features as well.
   - Context-free implementation of features was the vision of Jakobson, Fant and Halle (1957).
   - Seems also to be the idea behind Chomsky & Halle (1968).
   - We argue that, *insofar as distinctive feature theory is to have a bearing on articulatory implementation*, it's the only worthwhile system.

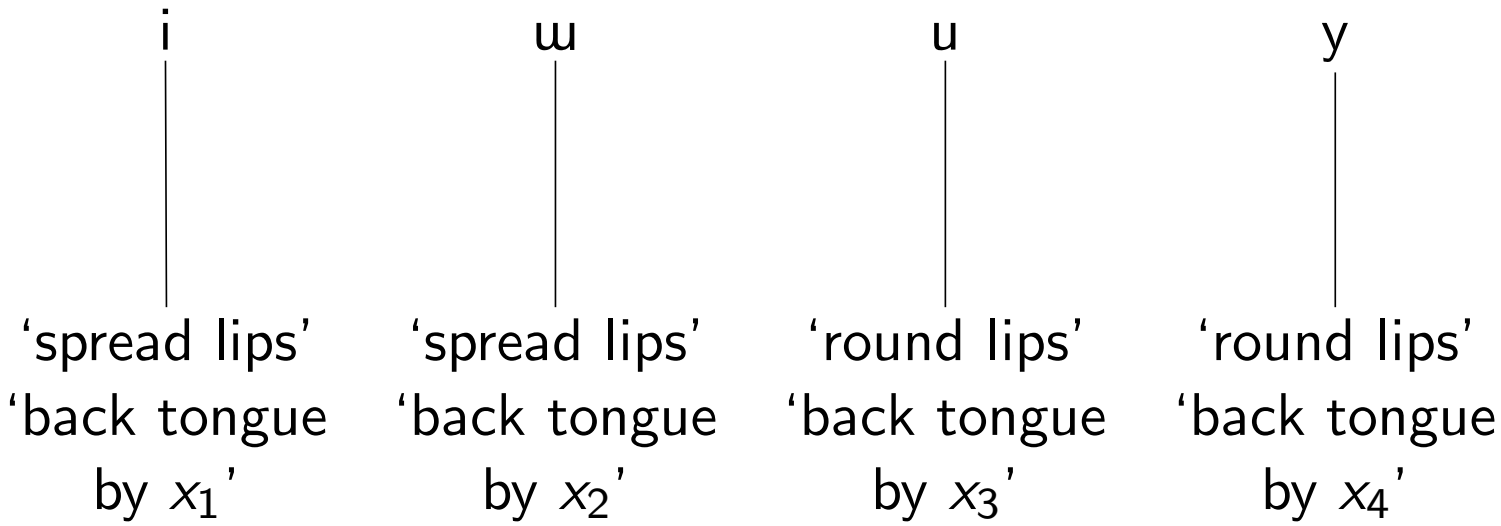
Context-free interpretation



Context-sensitive interpretation



A simpler representation



Possible objection—The theory is right, the data are wrong.

- That is, the phonetic analysis (e.g., Japanese [i̟] should be [ç], Turkish [y] should be [ʏ].)
- The featural analysis should be more refined.

- Allowance for more complex analyses will inevitably ameliorate the situation.
- This begs the question of whether the theory can ever be judged inadequate.
- It's a matter of personal judgment when this ceases to be worthwhile.
- I can describe any language in the world with 16 features... sampled at 22050 Hz!

Are innate features psychologically necessary ?

- Could a speaker with phonetic knowledge but no abstract feature system group the Turkish high vowels into {i,ɯ} and {y,u} ?
- Could a speaker pick [i] and [ɯ] from {i,ɯ,e,o,a}, and associated them with [i̟] and [ɯ̟] ?

- Yes of course : speakers can distinguish sounds based on their grossly similar phonetic properties.
- Sounds don't have to differ *just* by the relevant phonetic parameter.
- It just has to be salient in order to be helpful.

- In Emergent Feature Theory (Mielke 2004), speakers create sound patterns based on phonetic similarity.

Conclusion

1. Units of psychological computation
2. Descriptors of sound patterns
3. Phonetic units

- Turkish [*round*] and Japanese [*voice*] lack a simple phonetic interpretation.
- This requires a more complex interpretation of features.
- One might conclude that articulatory interpretations of features is not worthwhile.
- Beyond that, one may chose to look at perceptual interpretations of features, or just at segments.

Thank you !