

Quantifying Phonetic Informativity: An Information Theoretic Approach

James Whang

jamesw@snu.ac.kr



SEOUL
NATIONAL
UNIVERSITY

SAMSUNG

1. Selective Perception

- Speech perception is selective, weighting cues/features informative for categorical contrast higher than non-informative ones.
- However, multiple cues signal same category, making some redundant (Clements 2009).
 - VOT and F0 in Korean lenis/aspirated stops:
 $/t/$ vs. $/t^h/$
 - Peripherality and length in English high vowels:
 $/i/$ vs. $/i:/$

Research questions

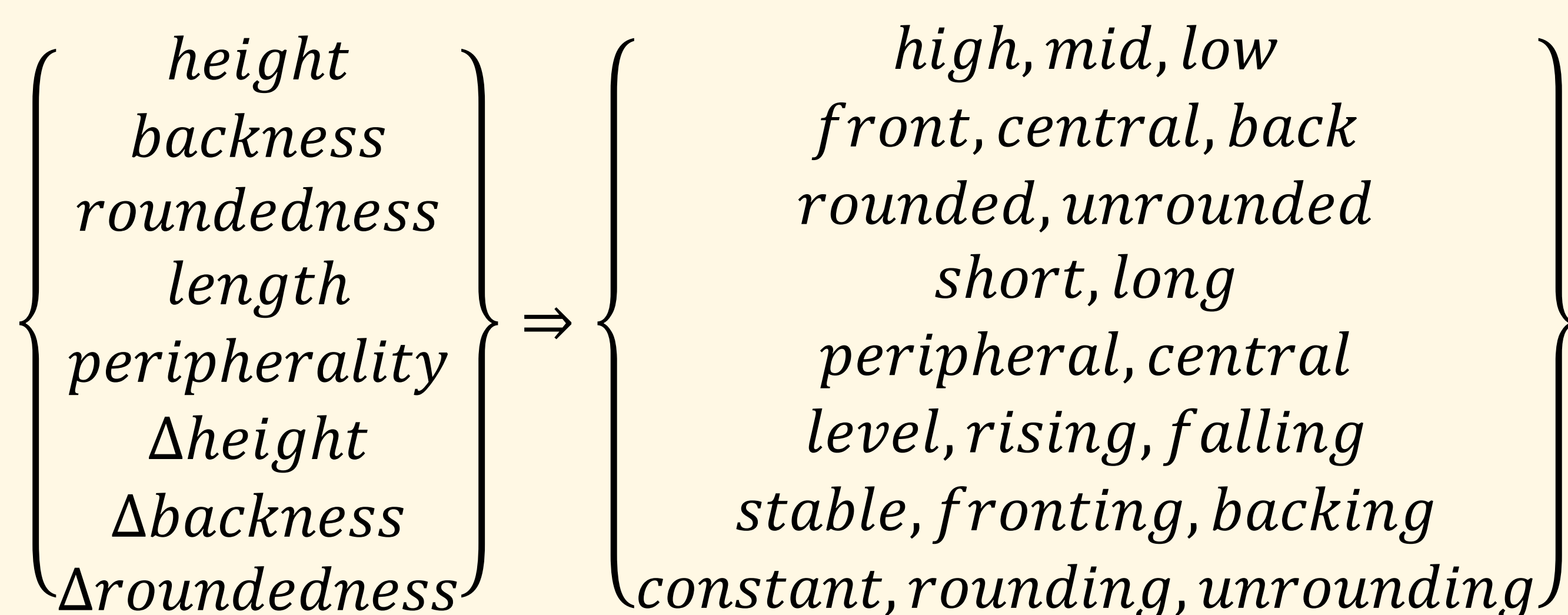
- Which cues are redundant vs. informative?
- How do we quantify redundancy/informativity?

3. Current Study

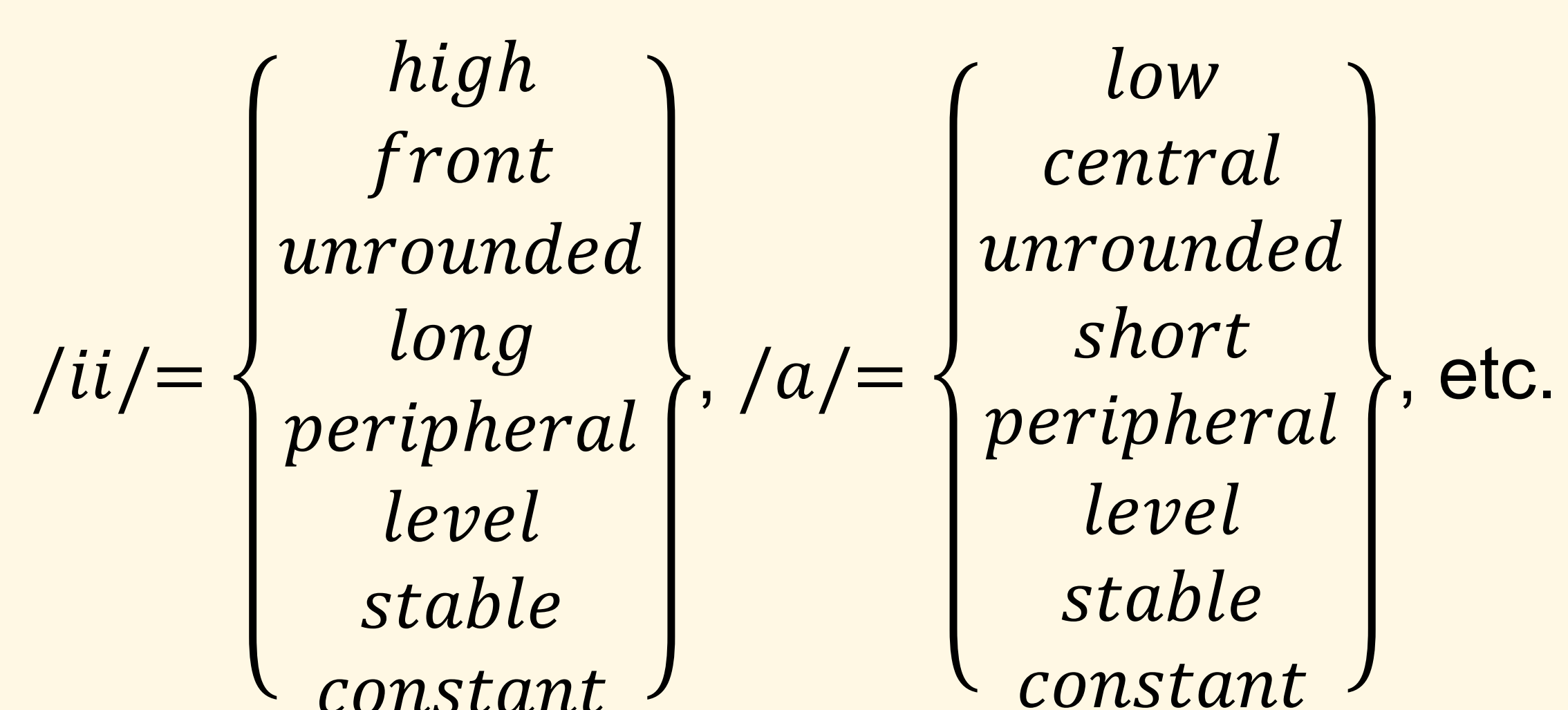
- Japanese vowels** as test case using CSJ-RDB (500K-word subset of Corpus of Spontaneous Japanese; Maekawa & Kikuchi 2005).
 - 679,123 vowels total

Procedure:

1. Define feature set



2. Convert vowels



3. Calculate surprisal (redundancy) for each feature

- Feature freq = \sum vowels containing feature
- Contexts = Given set A , all subsets of $A \setminus x$, where x is target feature

4. Calculate entropy for each feature (overall informativity) based on surprisal values

2. Quantifying Informativity

- Use Information Theory (Shannon & Weaver 1949) to both identify redundant cues and quantify the information held by each phonetic feature (or cue).

Surprisal: How informative is x in given context?

0 surprisal = completely redundant

$$-\log_2 \Pr(x|\text{Context})$$

Entropy (H): How informative is x overall?

0 entropy = does no work in language

$$\sum \Pr(x|\text{Context}) * -\log_2 \Pr(x|\text{Context})$$

4. Results

- Most informative cues
 - high* (H = 97.80)
 - vs. *mid* (H = 75.85), *low* (H = 42.49)
 - long* (H = 88.58)
 - vs. *short* (H = 38.44)
- Redundant cues (0 surprisal and entropy)
 - All Δ features: no diphthongs in Japanese (CSJ)
 - peripheral*: no lax vowels in Japanese
 - High informativity of *high* predicts heightened sensitivity to high vowels (Whang 2019)
 - High informativity of *long* and low informativity of *peripheral* predicts reliance on longness (not shortness!) to distinguish English tense/lax distinctions (Strange et al. 2001, 2011)
 - Low informativity of Δ features predict difficulty perceiving vowel-intrinsic spectral movement

5. Future Directions

- Redundant cues susceptible in language change?
- Compare informativity in different languages.
 - How much L1-L2 transfer?
 - How much training/data necessary?
- Using feature vectors instead of sets to quantify segment-internal timing relations.
 - Phonetic cues can be simultaneous or sequential.
 - Quantify within- and across-vector relations.