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. /tʰ/
  - Peripherality and length in English high vowels:

    /i/ vs. /ɪ/

Research questions

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

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

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

Entropy (H): How informative is x *overall*?

0 entropy = does no work in language

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

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

$$\begin{Bmatrix} height \\ backness \\ roundedness \\ length \\ peripherality \\ \Delta height \\ \Delta backness \\ \Delta roundedness \end{Bmatrix} \Rightarrow \begin{Bmatrix} high, mid, low \\ front, central, back \\ rounded, unrounded \\ short, long \\ peripheral, central \\ level, rising, falling \\ stable, fronting, backing \\ constant, rounding, unrounding \end{Bmatrix}$$

2. Convert vowels

$$/ii/ = \begin{Bmatrix} high \\ front \\ unrounded \\ long \\ peripheral \\ level \\ stable \\ constant \end{Bmatrix}, /a/ = \begin{Bmatrix} low \\ central \\ unrounded \\ short \\ peripheral \\ level \\ stable \\ constant \end{Bmatrix}, \text{etc.}$$

3. Calculate surprisal (*redundancy*) for each feature
   - Feature freq = ∑vowels containing feature
   - Contexts = Given set A, all subsets of $A \backslash \mathrm{x}$, where x is target feature
4. Calculate entropy for each feature (*overall informativity*) based on surprisal values

4. Results

- Most informative cues
  - $high$ $(\mathrm{H} = 97.80)$
    - vs. mid $(\mathrm{H} = 75.85)$, low $(\mathrm{H} = 42.49)$
  - $long$ $(\mathrm{H} = 88.58)$
    - vs. $short$ $(\mathrm{H} = 38.44)$

- 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.

Clements (2009) The role of features in phonological inventories. **Shannon & Weaver** (1949) *The mathematical theory of communication.* **Maekawa & Kikuchi** (2005) Corpus-based analysis of vowel devoicing in spontaneous Japanese: An interim report. **Whang** (2019) Effects of phonotactic predictability on sensitivity to phonetic detail. **Strange, Hisagi, Akahane-Yamada & Kubo** (2011) Cross-language perceptual similarity predicts categorial discrimination of American vowels by naïve Japanese listeners. **Cohen Priva** (2015) Informativity affects consonant duration and deletion rates.