Foot structure and Rhythm

- Organization of syllables into feet

\[
\text{F} \quad \text{F}
\]
\[
\text{S} \quad \text{W} \quad \text{S} \quad \text{W}
\]

- Language differences in rhythm
  (‘stress-timing’ vs. ‘syllable-timing’)

- Isochrony of feet in stress-timed languages? No.

- Characteristics of languages that lead to impression:
  - $\%V$
    in syllable-timed lgs. a higher proportion of the speech signal is vocalic (voiced).
  - $\Delta C$
    in stress-timed lgs, there is greater variability of the consonant interval durations.
Classifying Lg Rhythm

Figure 1: Results from [6] for cv-variability represented by $\Delta C$ and $%V$ respectively: Stress-timed languages (EN: English, DU: Dutch, PO: Polish) can be distinguished from syllable-timed languages (SP: Spanish, IT: Italian, FR: French, CA: Catalan).

Figure 3: Cross plotted mean values for rate normalized variability of unvoiced intervals (varcoUV) and percentage over which speech is voiced (VO).
Causes for classification

- **Phonotactics**
  - Possibility of consonant clusters of various lengths limit the likelihood of hearing “syllable-timing.”

- **Foot-structure**
  - Temporal consequences of on rhythm (observed in stress-timing languages):
    - Greater length of stressed vs. unstressed syllables (vowel reduction)
    - Shortening of syllable durations in polysyllabic feet
Polysyllabic Shortening

- Port (1981)
- Kim & Cole (2005)
- Radio News corpus
- 12 minutes of speech

2: Mean syllable duration against the number of syllables in a foot. ▼-within-ip, ○-across-ip, Δ-across-IP.
Phonotactics vs. Foot structure

- Foot-structure effects expected in English, but not Spanish.
- English vs. Spanish (Nava et al, 2008)
  - “The North Wind and the Sun”
    - 30 English speakers read story in English
    - 45 Spanish (Mexican) speakers read story in Spanish
    - 20 less proficient L1Sp/L2Eng read in English

- Measurement technique
  - VR (Dellwo et al, 2007)
    - Total voiceless duration / total voiced duration (within pauses)
  - easy to automatically extract from large dataset
  - correlates with impressionistic judgements of syllable-timing (low VR) vs. stress-timing (high VR)
Results and TaDA modeling

- Spanish has lower VR
- foot effects
- phonotactic differences

- TaDA
  - models English syllable structure and phonotactics
  - no foot effects currently

- TaDA as Control
  - automatically generate “North Wind and Sun”:
    - English TaDA
    - Spanish TaDA
      (Parrell, Nava, Proctor et al)

- Results show effect of English phonotactics

- TaDA Models the behavior of L2 speakers with low prosodic proficiency (no foot effects).
Coupling Graph Model of Foot Effects

- Adding hierarchical coupled oscillators to the coupling graph

- English: foot-to-syllable coupling ($\lambda_{FS}$) >> syllable-to-foot ($\lambda_{SF}$) where $V$ is also $S$

```
  F
 /\  \
/   \  \\
V     V  V
C   C   C  C  C   C   C  C  C
  \\
 b  a  k  i  n  g

  F
 /\  \
/   \  \\
V     V  V
C   C   C  C  C   C   C  C  C
  \\
 c a s s e r o l e s
```
Polysyllabic shortening Simulation

\( \lambda_{FS} = 5 \)
\( \lambda_{SF} = 1 \)
\( \omega_{0F} = 1 \)
\( \omega_{0S} = 2 \)

2 syllables per foot

3 syllables per foot

Foot oscillator
Syllable oscillator

Time (s)
Stress Asymmetry

- How can the differential durations of stressed and unstressed syllables be modeled?

- Hypothesize clock slowing gesture ($\mu_T$) that is active at phases of Foot oscillator corresponding to stressed syllables.

- $\mu_T$ slows clock of Foot and Syllable oscillators (and all constriction gesture) in proportion to its activation level ($a_\mu$).

- Maximum strength of $\mu_T$ gesture will determine the degree or temporal asymmetry between stressed and unstressed syllables.
Stress Asymmetry Simulation

2 syllables per foot

3 syllables per foot

Foot oscillator

Syllable oscillator

Time

\[ a_\mu \]