4pSC3
Modeling prosodic rhythm: Evidence from L2 speech

Emily Nava
University of Southern California

Louis Goldstein
USC, Haskins Laboratories

Elliot Saltzman
Haskins Laboratories

Hosung Nam
Haskins Laboratories

Michael Proctor
Yale University, Haskins Laboratories

Supported by NSF IIS 07-03624
Cross-linguistic rhythmic classification

- ‘Stress-timed’ vs. ‘syllable-timed’ languages (Pike 1945, Abercrombie 1967)
  - Original distinction cast in terms of ‘isochrony’; no evidence found for this basis
- Dauer 1983: Continuum of +/- ‘syllable’ or ‘stress’ timed
  - Vowel reduction: English √, Spanish ∅
  - Syllable structure inventory:
    - Open syllables: English 44%, Spanish 70%;
    - CV syllables: English 34%, Spanish 60%
  - Correlates of word-level stress:
    - Vowels in stressed syllables 50% longer than unstressed in English, Spanish only 10%
Measurement techniques

Measurement techniques for rhythmic classification:

- ΔV, ΔC: standard deviation of vocalic, consonantal intervals (Ramus et al. 1999)

- nPVI-V, rPVI-C: pairwise variability index is a measure of unit-to-unit variation in speech (Low & Grabe 2002)

- Rhythm Ratio: average of the ratio of adjacent syllables (Gibbon and Gut 2001)
Measurement techniques

- **Voicing ratio (VR)** (Dellwo et al. 2007)
  \[
  VR = \frac{\text{voiceless intervals}}{\text{voiced sequences}}
  \]
- **V%**: total percentage of speech that is voiced
- **VCL**: standard deviation of voiceless intervals
- Advantage: values generated automatically over large data set
Rhythm in L2 Speech

• L1 rhythmic influence on L2 rhythm: values between L1 and L2 (White 2007)
  – L1 Chinese, L2 English
    • nPVI-V (Low et al. 2000)
    • Δ C, %V (Lin & Wang 2005)

– L1 Mexican Spanish, L2 English bilinguals:
  • nPVI-V values between L1 and L2 related to much lower incidence of vowel reduction in their L2English (Carter 2005)
Rhythm in L2 Speech

• English: reduced, deletes vowel more common than German
  – Reduced vowels in German occur in final syllables, inflectional morphemes
  – Results: L1 English reduced/deleted vowels in L2 German at a higher rate than L1 German controls, L1 Italian/Romanian at a lower rate (Gut 2003)

• Acquisition cue: function words generally unstressed in English; unstressed vowel as underlying property of English “stress-timed” rhythm
Rhythm and L2 acquisition

• **Current study**: Acquisition of prosodic proficiency in English: the rhythm connection
  
  – Difference between stressed and unstressed syllables is greater in English than in Spanish  

• **L1Spanish/L2English speakers**:  
  What do native Spanish speakers learn about the foot when they acquire native-like competence in English?
Experiment 1: Rhythmic proficiency

• Stimuli: “The North Wind and the Sun” (English, Spanish)

• Participants:
  – 30 English Controls
  – 45 L1Spanish/L2 English
  – 20 monolingual Spanish

• Voicing ratio (VR = voiceless/voiced)
  – value automatically detected for each separate sentence of recorded passage

• Percentage voiced, SD of voiceless
  – V%: total percentage voiced speech
  – SD VCL: standard deviation of voiceless intervals
Experiment 1: Results V%, SD VCL

- +/- phrasal prosody: Separate experiment determined degree of English native-like prosodic proficiency
Experiment 1: Results discussion

- **English**: lower V%, higher SD VCL
- **Spanish**: higher V%, lower SD VCL
- **L2 results**: grant insight into acquisition process, don’t reveal whether acquisition of syllable level of foot level in English

- Which aspects of English-like rhythm have been acquired by L2 speakers?
- **Possible answer**: If learning is primarily at the syllable level, TaDA would reproduce those results
Modeling prosodic rhythm: TaDA

- **TaDA**: Task-dynamics model of speech production generates gestural patterning in time and the resulting acoustic output (Nam et al, 2005).
  - **Current model**: Incorporates knowledge about coordination for English at the syllable level (including effects of complex onsets and coda)
    - BUT **aprosodic**: does not incorporate temporal effects due to foot structure, so …
    - Model behaves like L2 speaker who has acquired the temporal structure of English at the syllable level but not the foot level
Information flow through TaDA

- Syllable-structure based Gesture coupling
- English Text or ARPABET
- Intergestural Coupling Graph
- Gestural planning oscillator variables
- Activation variables (Gestural Score)
- Tract/Constriction variables
- Model articulator variables

INTER-GESTURAL COORDINATION

INTER-ARTICULATOR COORDINATION

output speech
Coupling graph & gestural score: “spot”

- Labial closure
- Pharyngeal narrow
- Alveolar critic
- Alveolar closure

In-phase

Anti-phase

Time: 50 ms
TaDA English:

• Given English text input
  gestures taken from dictionary
  coupling graph generated

• Resulting construction has knowledge of English syllable structure, but not prosody
  – No vowel reduction in unstressed syllables

• Example: “cop top” vs. “copped stop”
  – TaDA appropriately accommodates English syllable structure: increased voiceless interval
Coupling graphs

“cop top”

TB (CLO)  →  LIP (CLO)  →  TT (CLO)  →  LIP (CLO)
   |                |                |
TB (wide)  |                |                |
            TB (wide)  |                |
                   |                |
                   TB (wide)

“copped stop”

TB (CLO)  →  LIP (CLO)  →  TT (CLO)  →  TT (CRIT)  →  TT (CLO)  →  LIP (CLO)
   |                |                |                |                |
TB (wide)  |                |                |                |                |
            TB (wide)  |                |                |
                   |                |                |
                   TB (wide)

In-phase  Green
Anti-phase  Red
TaDA English

“cop top” VC#CV

“copped stop” VCC#CCV

Time 680 ms
English Speaker

“cop top” VC#CV

“copped stop” VCC#CCV

Time 680 ms
TaDA English

“divide”

English Speaker

“divide”

No vowel reduction!

Time 600 ms
Spanish TaDA

• Preliminary version of Spanish TaDA
  – Includes dictionary entries and appropriate syllable structure
• Adjustments to coupling graphs
• But NO prosody!

• Prediction for Spanish TaDA: if foot is more ‘symmetric’ in Spanish than in English, then Spanish TaDA should more faithfully reproduce Spanish speech
TaDA Spanish

“david”

Spanish Speaker

“david”

Time 600 ms
Modeling prosodic rhythm: TaDA computational experiment

- **Synthesis**: “The North Wind and the Sun”
  - text input to TaDA
    - English
    - Spanish

- **Analysis**:
  - Output acoustics analyzed using same algorithms for natural English, Spanish
    - V%: total percentage voiced speech
    - SD VCL: std dev of voiceless intervals
Modeling prosodic rhythm: Results

[Graph showing various data points labeled with English, L2Eng+, L2Eng-, TaDA English, TaDA Spanish, Monolingual Spanish controls.]
Results and Implications

• TaDA English and L2 English speech without native-like prosody show similar rhythm measurement
  – TaDA takes into account syllable structure but does not account for the durational differences characteristic of English foot structure

• TaDA Spanish yields results close to native Spanish
  – Appropriate syllable structure is enough; durational difference in Spanish foot not as great as in English

• Acquisition of ‘asymmetric’ foot in English essential component to native-like prosody
  – L2Eng + group has acquired foot in English, while L2Eng - has not
References


• Carter, P. M. 2005. Quantifying rhythmic differences between Spanish, English, and Hispanic English. In R. S. Gess, & E. J. Rubin (Eds.), *Theoretical and experimental approaches to romance linguistics: Selected papers from the 34th linguistic symposium on romance languages* (Current issues in linguistic theory 272) (pp. 63–75). Amsterdam, Philadelphia: John Benjamins.

• Dellwo, V., Fourcin, A. and E. Abberton. 2007. Rhythmical classification based on voice parameters. International Conference of Phonetic Sciences (ICPhS)


References