Cross-linguistic prosodic differences and their effects on the interpretation of error patterns in child speech

Timothy Arbisi-Kelm
*University of Wisconsin, Madison*

Mary E. Beckman
*Ohio State University*

Special thanks to the following...

Principal Investigator: Jan Edwards (University of Wisconsin-Madison)

Lab members: Hyunju Chung, Junko Davis, Eunjong Kong, Fangfang Li, Sarah Schellinger, Asimina Syrika, Peggy Wong

Funding sources: NIH traineeship to first author, and NIDCD grant 02932 to Jan Edwards (PI)

Cross-linguistic phonological differences: inventory of autosegmental content specifications

• Focus is often on differences in phoneme inventories:
  – e.g., English has /s/ and /ʃ/, while Greek has only /s/.
• Or on differences in phonotactic constraints:
  – */tu/ in Cantonese
  – */kjɔ/ in English
  – */si/ in Japanese

Cross-linguistic prosodic differences (1)

But languages also differ dramatically in terms of prosodic structure:

• Phrasal shapes
  – English, Greek: pitch shapes (“accents”) from intonation linked to metrically prominent lexical stresses
  – Japanese: phrasal pitch shapes where “accents” are lexical tones; phrasing and pitch range manipulation, but no “stress”
  – Cantonese: intonation-phrase-final tones appended after lexical tone on last syllable

Cross-linguistic prosodic differences (2)

• Word shape
  – English: primarily 1-2 syllables, trochaic bias (Hayes 1980; Halle & Vergnaud 1987)
  – Greek: trisyllabic stress window aligned to end of word; iambs as common as trochees (Joseph & Philippaki-Warburton, 1987)
  – Japanese: predominantly 2-3 syllables, no stress, contrastive vowel and consonant length
  – Cantonese: predominantly monosyllabic, with each syllable equally prominent, specified for tone

Cross-linguistic prosodic differences (2)

• Syllable reduction:
  – English /æ/ and Greek /i, u/ can delete in metrically weak syllables that cannot align to pitch accents — e.g. *potato, qi’monas/ ‘winter’, but not *ç’ilja/ “lips”
  – Japanese: high vowel devoicing or deletion constrained by vowel length rather than by metrical strength
  – Cantonese: “syllable fusion”= consonant and vowel lenition and even deletion, but with preservation of lexical tone (Wong 2004, 2006)
Prosody in language acquisition

• Native language prosody is one of the earliest aspects of language that children learn:
  – Distinguish correct vs. incorrect pause placement in clauses at 6 mos. (Jusczyk, Hirsch-Pasek et al., 1992)
  – English-speaking babies show preference for trochaic vs. iambic words at 9 mos. (Jusczyk et al., 1999)

English-acquiring babies (13-20 mos) produced only recognizable trochees, while the French-speaking babies produced recognizable iambns Vihman, DePaolis, & Davis (1998).

Prosodic evidence in speech errors and disfluencies

• English speech errors (e.g., Shattuck-Hufnagel 1987; Fromkin 1973; Dell 1985) often exchange consonant gestures at foot beginnings.
  ‘deep lake’ → ‘leep dake’

Alignment evidence in speech errors and disfluencies:
Metrical structures are language specific

• Japanese errors (Kubozono 1989) suggest an affinity between analogous mora positions in different syllables, regardless of the gestural content licensed by the position.
  pansento → pansento

The παιδολογος project — cross-linguistic research on phonological acquisition

• Comparing word-initial lingual obstruents in real words and nonwords across Cantonese, English, Greek, and Japanese.
• Languages chosen because all have a rich inventory of lingual obstruents, as well as salient prosodic differences.
• Participants (20 2-5 yr-old children for each language) completed a word-repetition task, presented with both audio and visual stimuli.

The influence of ambient language on late babbling

English-acquiring babies (13-20 mos) produced only recognizable trochees, while the French-speaking babies produced recognizable iambns Vihman, DePaolis, & Davis (1998).
Analysis — transcription by native-speaker phonetician

- Initial consonants transcribed as:
  - 1 (correct and fluent)
  - 0 (incorrect in place and/or manner)
  - V (correct except for voicing/aspiration)
  - E (“Effortful”; reserved for fricatives and affricates)
- Also coded:
  - “Split CV” (pause/resyllabification between consonant and vowel)
  - Devoicing

**“E” for “effortful”**

- Analogous to adult stuttering — a disfluency arising from the demands of coordinating a difficult consonant constriction (e.g., tongue tip for [s]) with:
  - respiratory system for utterance initiation
  - laryngeal posture and subglottal pressure for initial syllable
  - tongue body coordination for “following” vowel

**Metrical conditions for “E” in Greek**

- Cases of “E” for fricatives most frequent in four-syllable nonwords.
- Greek has a “three-syllable window” for stress; therefore four-syllable forms necessarily have word-initial unstressed syllables.
- Child seems to focus more effort on getting non-initial stressed syllable right.

**“Split CV”**

- “split CV”: disfluency after a plosive release, especially when stop glottal gesture repeated.
- “split CV” suggests struggle with coordinating a precise lingual gesture with the following vowel.
- Percept is insertion of epenthetic vowel.

**Autosegmental/Metrical conditions for “split CV”**

- In English, such “split CV” cases are often seen when child is attempting the particularly complex gestural configuration of the “palatalized velar” (or /k/ before /yu/ diphthong).
- Often there is also a stereotypical /u/ for /k/ substitution.

**Different metrical resolution for /k/ in Japanese**

In Japanese, the more common resolution of the difficulty of this gestural configuration is to substitute an alveopalatal (cf. also Tsurutani, 2004).
Metrical conditions for syllable deletion

- Greek word-initial unstressed syllables in 4-syllable forms were often deleted, so that the word began with the stressed second syllable.
- Compare this with the "trochaic bias" of English, where initial unstressed syllables are deleted also in disyllabic and trisyllabic forms (e.g., /ˈnæ.nə/ for banana).

Vowel devoicing in Cantonese equals syllable fusion

- Cantonese syllable fusion: vowel deletion can occur without tone loss.
- Devoicing of vowel targets first part of syllable nucleus, leaving space to realize tone.

Devoicing in English, too!

- Devoicing in English should only occur in unstressed syllables, which were not used in our elicitation protocol.
- But we did see cases of devoiced syllables, as another kind of "E" with difficult sibilant fricatives

Or is it all in the ear of the adult perceiver?

- So what do we do when speech errors “break the rules”?
  - Interpretation 1 (TAK, listening to “stress”): “Residual cues to stress (e.g., high intensity of consonant burst and alternation with weak following vowel) preserves the syllable count.”
  - Interpretation 2 (MEB listening to tone pattern): “This is English, so the syllable is deleted and the stress shifts to the following syllable.”

Conclusion and Future Directions

- As children acquire the ambient spoken language, they must learn the metrical structures as well as the inventory of autosegmental content specifications.
- Children’s speech errors are therefore highly constrained by the prosodic structure of the ambient language.
- Future work:
  - statistical analysis of all error patterns
  - perception tests to address variation in prosodic structure licensing in complex errors

Thank you!