Revisiting the role of prosody in early language acquisition

Megha Sundara

UCLA Phonetics Lab
Outline

• Part I: Intonation has a role in language discrimination

• Part II: Do English-learning infants have a trochaic bias in early word segmentation?
PART I:
LANGUAGE DISCRIMINATION
Newborns discriminate languages

• Early research
  – Native vs. non-native languages (Bahrick & Pickens, 1988; Mehler et al., 1988)
    • Based on familiarity with and recognition of their native language

• More recently
  – Discriminate some language pairs even when unfamiliar with both (Nazzi et al., 1998)
    • Based on infants’ sensitivity to prosody, “specifically the rhythmic, timing differences” between languages
What is rhythm?

• Languages traditionally divided into 3 rhythm classes (Pike, 1945; Abercrombie, 1967)
  – Stress-timed (E.g. English, Dutch, German)
  – Syllable-timed (E.g. Spanish, French, Italian)
  – Mora-timed (E.g. Japanese, Telugu, Kannada)

• Languages within rhythm classes share syntactic (Mehler & Christophe, 2000), phonological (Dauer, 1983) and acoustic-phonetic characteristics (Ramus et al., 1999; Low et al., 2000)
Rhythm class: Acoustic-phonetic bases

Fig. 1. Distribution of languages over the (%V, ΔC) plane. Error bars represent ± 1 standard error.
Rhythm divorced from intonation

• ...the rhythm-based language discrimination hypothesis (R hypothesis), stems from evidence that newborns are sensitive to prosody, that is, the overall properties of utterances such as intonation and rhythm. The R hypothesis states that infants extract prosodic, and more specifically, rhythmic properties of sentences [italics added] and that they sort sentences into a small number of classes or sets based on rhythmic, timing properties [italics added]. (Nazzi et al., 1998, p. 757)
Part I: Language discrimination

• English learners’ discrimination of English vs. German (joint work with Chad Vicenik)
  – Intonation sufficient to distinguish between rhythmically similar languages.
  – Adult listeners attend to intonation to distinguish them.
  – Infants fail to discriminate when intonation is removed.
Method

• Stimuli
  – 8 female speakers each
  – 20 sentences per speaker
  – Adult-directed speech
  – Sentences based on the Nazi et al., 2000 stimuli
What’s in the input?

• Acoustic analysis
  – Rhythm measures
    • 11 measures
      – %S, ΔO, ΔS, VarcoS, VarcoO, Mean S, Mean O, rPVI S, nPVI S, rPVI O, nPVI O
  – Intonation measures
    • 6 measures
      – Min f0, max f0, mean f0, number of rises, average rise, average slope of f0
Classification using Logistic Regression

Rhythm & Pitch cues

Rhythm & Pitch cues

Rhythm cues

Pitch cues

Percent

Vicenik & Sundara, under review
Method

- **Stimuli**
  - 8 female speakers each
  - 20 sentences per speaker
  - Adult-directed speech
  - Sentences based on the Nazzi et al. stimuli

- **Adult listeners**
  - N =15 per condition

- **3 conditions**
  - Low-pass filtered
  - Rhythm-only
    - ?a?a?a
    - (Sasasa)
  - Intonation only
    - aaaaa
Low-pass filtered

• Cut off at 400 Hz, 50 Hz smoothing

• Sample
Intonation only

• Sample
Adult perception results

![Graph showing discrimination scores for Low-Pass Filtered, Rhythm Only, and Intonation Only conditions. The x-axis represents the conditions, and the y-axis represents the discrimination score. The graph includes data points and error bars.]

Vicenik & Sundara, under review
Infant listeners

- 5- and 7-month-olds
- Tested using Headturn Preference Procedure (HPP)
  - Identical to Nazzi et al.’s procedure
ILUSTRATION OF THE HEAD-TURN PREFERENCE PROCEDURE (HPP)

1. light
2. speaker
3. VCR
4. computer
5. observer
6. caretaker
7. baby
Design

• Two phases
  – Familiarization phase
    • Either English or German (counterbalanced)
    • 4 passages by 2 different speakers
    • Listen to each passage for at least 20 s (80s total)
  – Test phase
    • 8 trials
    • 4 new passages by 2 new speakers
    • Listening time to familiar and novel language averaged
Infant results

Vicenik & Sundara, under revision

Listening time (s)

- 5-mo Full cue
- 7-mo Full cue
- 7-mo Low-pass filtered
- 7-mo Flat intonation

Familiar (green) vs. Novel (gray)

* *
Part I: Language discrimination

• In language discrimination (joint work with Chad Vicenik)
  – Intonation sufficient to distinguish between rhythmically similar languages, English vs. German.
  – Adult listeners attend to intonation to distinguish them
  – Infants fail to discriminate when intonation is removed

• Cannot ignore the role of intonation in language discrimination
PART II: WORD SEGMENTATION
The segmentation problem
Timeline of word segmentation

• Studies by Jusczyk and colleagues

• English-learning babies can segment
  – Monosyllabic CVC (e.g., cup): 7.5 mo
  – Trochaic bisyllables (Sw, e.g., doctor): 7.5 mo
  – Iambic bisyllables (wS, e.g., guitar): 10.5 mo
  – Monosyllabic VC (e.g., eel): 16 mo
Trochaic bias

- Matches ambient language prosody
  - Seen in Dutch, and English (Houston et al., 2000), but not Canadian French 8-month-olds (Polka & Sundara, 2012) segment trochees

- Cannot be learned from the distribution of stress for words in isolation
  - Only a minority of 2 syllable utterances in English are trochaic

- "...infants acquire the trochaic parsing bias as a generalization over a “protolexicon” of word forms extracted on the basis of the forms relatively high conditional probability and frequency..."  
  - Not just frequency
Central idea in Swingley, 2005

• Statistical probability used to cluster bisyllables necessary for the emergence of a trochaic bias

• Support from artificial language learning studies
  – 6- to 7-month-olds weight transitional probability over prosody (Thiessen & Saffran, 2003)
  – 8-, and 11- month-olds weight prosody over transitional probabilities (Johnson & Jusczyk, 2001)
Our prediction

If infants rely on statistical clustering, **iambs** should **not** be difficult to segment
Test for sensitivity to iambs

- 6-month-old English-learning infants (n = 8)
- Using the Headturn Preference Procedure
- Tested on beret, device, guitar and surprise
  - Familiarized with 2 passages
    - Criteria - 60 s to each passage
  - Tested on all four isolated words
ILUSTRATION OF THE HEAD-TURN PREFERENCE PROCEDURE (HPP)

1. light
2. speaker
3. VCR
4. computer
5. observer
6. caretaker
7. baby
Methods

• Testing with HPP in two stages
  – Familiarization stage
    • E.g. Your device can do a lot. Her device only fixes things. My new red device makes ice cream. The pink device sews clothes......
    • E.g. The big red surprise is for you. The small pink surprise is for Dawn. Your surprise will be fantastic. I think Dawn got the old surprise.........
  – Test stage
    • 2 Familiar word lists
      – device.....device......device....doctor.....
      – surprise....surprise...surprise.....candle....
    • 2 Control / Novel word lists
      – beret.....beret......beret....beret.....
      – guitar....guitar...guitar.....guitar....
Results

Familiar > Novel: 7 out of 8 infants!
Modified timeline of word segmentation

• English-learning babies can segment
  – Iambic bisyllables (wS, e.g., guitar): 6 mo
  – Monosyllabic CVC (e.g., cup): 7.5 mo
  – Trochaic bisyllables (Sw, e.g., doctor): 7.5 mo
  – Iambic bisyllables (wS, e.g., guitar): 10.5 mo
  – Monosyllabic VC (e.g., eel): 16 mo

• Statistical clustering of bisyllables precedes the trochaic bias
Follow-up

• Do English-learning 6-month-olds rely exclusively on statistical probabilities to segment iambic patterns?

• Use prosodic distribution at utterance boundaries instead (Aslin et al., 1996)?
  – How do you get boundaries?
    • Prosodic: Intonation
    • Segmental: Preboundary lengthening and post boundary strengthening

However, only 9-month-old English-learning infants (and 10-month-olds French learning infants) are sensitive to phrasal boundaries (Jusczyk et al., 1992; Gout et al., 2004).
Stress at utterance boundaries

Corpus

Stage 1: Extract all input utterance from entire CHILDES corpus
- Utt $U$ is input to child $c$ if $U$ uttered when $c$ present, not by $c$
- For every child $c$ in CHILDES, extract English-only input

Stage 2: Lexical stress
- Custom software: lexical search-and-replace using CELEX
- The words are replaced by their stress patterns
  
  \[
  0 \quad 1 \quad 0 \quad 01 \quad 0 \quad 0 \quad 1 \quad 10
  \]

  Size: 340,000 utts– 2 million wds– 1.5 million syls
  Equivalent to about 2 months of language input

Stage 3: De-segment and analyze
- Remove wd bdry’s (e.g. $0 \ 1 \ 0 \ 01 \ 0 \ 0 \ 1 \ 10 \rightarrow 0100100110$)
- Statistical properties of resulting distribution

Daland, 2010
Stress at utterance boundaries

<table>
<thead>
<tr>
<th>Utterance-initial</th>
<th>Utterance-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Trochees</td>
<td>• Trochees</td>
</tr>
<tr>
<td>• Iambs</td>
<td>• Iambs</td>
</tr>
<tr>
<td>• (other)</td>
<td>• (other)</td>
</tr>
<tr>
<td>74340</td>
<td>88519</td>
</tr>
<tr>
<td>79529</td>
<td>98571</td>
</tr>
<tr>
<td>(186729)</td>
<td>(153508)</td>
</tr>
</tbody>
</table>

Daland, 2010
Follow-up

• Do English-learning 6-month-olds rely exclusively on conditional probabilities to segment iambs?

• Use prosodic distribution at utterance boundaries instead (Aslin et al., 1996)?
  – How do you get boundaries?
    • Prosodic: Intonation
    • Segmental: Preboundary lengthening and post boundary strengthening

• Use prosodic distribution at phrasal boundaries (Christophe et al., 2008; Daland, 2009)?
Summary & Conclusion

Part I:

• Intonation can be used for language discrimination, even for prosodically similar languages

• Listeners, adults and infants, attend to intonation while discriminating languages

• Need to rethink role of intonation and its interaction with rhythm
Summary & Conclusion

Part II:

• English learning infants can segment iambics at 6-mo

• Need to rethink the contents of protolexicon of infants and the interaction between statistical learning and prosody