Evaluating the learnability of vowel categories from Infant-Directed Speech

Jahnavi Narkar¹, Ekaterina A. Khlystova¹, Connor J. Mayer², Ann Aly³, Ji Young Kim⁴, Megha Sundara¹
¹Department of Linguistics, UCLA; ²Department of Language Science, UC Irvine; ³Tech Flow, Cape Coral, FL, USA; ⁴Department of Spanish and Portuguese, UCLA

BACKGROUND

• Hyper-articulation – increased distance between centroids of vowels – in infant-directed speech (IDS) is thought to facilitate acquisition (e.g., Trainor & Desjardins, 2002; Liu et al., 2005).
• But vowels in IDS are also more variable (Cristia & Seidl, 2014; Martin et al., 2015; Ludusan et al. 2021)

ALTERNATIVE APPROACH

➢ Evaluate distributional overlap
➢ By combining category distance and variability
➢ Measures used extensively in socio-phonetics and machine learning (e.g., Hay, Warren & Drager, 2006; Kelly & Tucker, 2020)
➢ Independently test learnability via previously implemented Gaussian learner (Feldman et al., 2013)

METHODS

• Four connected speech corpora analyzed:
  - English IDS: Providence Corpus (Demuth et al. 2007; ~20K tokens)
  - English ADS: Buckeye Corpus (Pitt et al. 2007; ~20K tokens)
  - Spanish IDS: adult-child dyads recorded in lab (Sundara et al. 2020; ~5K tokens)
  - Spanish ADS: adult Spanish speakers (Kim & Repiso-Pujol, 2021; ~5K tokens)
• Extracted F1, F2, F3 & duration in Voicespace (Shue et al., 2011)
• Indexing overlap between categories:
  (a) Pillai scores (0 = complete overlap; 1 = no overlap e.g., Hays et al. 2006)
  (b) KL divergence - machine learning statistic to measure the difference between 2 distributions (0 = complete overlap; larger number = less overlap)
• Extracting vowel categories: Bayesian model of distributional learning (Feldman et al., 2013)

RESULTS

Do vowel categories in IDS have less overlap than in ADS?

• Pillai scores to generate dissimilarity metric for vowel pairs in IDS and in ADS
• 2-D Multi-Dimensional Scaling (MDS) solution to visualize dissimilarity space

In both Spanish and English, some evidence that IDS vowels have less overlap

Extracting vowel categories via a Gaussian learner

• Trained a distributional model (Feldman et al. 2013) on F1, F2, F3, duration

Spanish IDS

<table>
<thead>
<tr>
<th>F1 (Barks)</th>
<th>F2 (Barks)</th>
<th>F3 (Barks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>1.5</td>
<td>1.8</td>
</tr>
<tr>
<td>2.0</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>3.0</td>
<td>3.5</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Spanish ADS

<table>
<thead>
<tr>
<th>F1 (Barks)</th>
<th>F2 (Barks)</th>
<th>F3 (Barks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>1.5</td>
<td>1.8</td>
</tr>
<tr>
<td>2.0</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>3.0</td>
<td>3.5</td>
<td>4.0</td>
</tr>
</tbody>
</table>

English IDS

<table>
<thead>
<tr>
<th>F1 (Barks)</th>
<th>F2 (Barks)</th>
<th>F3 (Barks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>1.5</td>
<td>1.8</td>
</tr>
<tr>
<td>2.0</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>3.0</td>
<td>3.5</td>
<td>4.0</td>
</tr>
</tbody>
</table>

English ADS

<table>
<thead>
<tr>
<th>F1 (Barks)</th>
<th>F2 (Barks)</th>
<th>F3 (Barks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>1.5</td>
<td>1.8</td>
</tr>
<tr>
<td>2.0</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td>3.0</td>
<td>3.5</td>
<td>4.0</td>
</tr>
</tbody>
</table>

• Pillai scores to generate dissimilarity metric for vowel pairs in IDS and in ADS
• 2-D Multi-Dimensional Scaling (MDS) solution to visualize dissimilarity space

KL divergence

• Calculated (symmetric) KL divergence for vowel pairs in IDS and ADS
• Greater absolute value of divergence (less overlap) in ADS
• But relatively more pairs in IDS with greater divergence (less overlap)

CONCLUSIONS

• Mixed findings in IDS
  • Pillai score for the vowel system somewhat more dispersed
  • Relatively more vowel pairs in IDS have greater KL divergence
• However, Bayesian distributional learner has lot of difficulty with connected speech
  • Worst on English 9-vowel system, though better in ADS
  • In some conditions it extracts 5 vowels, but only in Spanish IDS
• Overall, no clear evidence for facilitation in IDS

FUTURE DIRECTIONS

• Improvement needed in distributional learners to handle variation in naturalistic speech
• Perhaps IDS plays a different role in category learning
• Could the greater spread in IDS be helpful to identify relevant acoustic cues for vowel categories?

ACKNOWLEDGMENTS

This work was supported in part by NSF BCS-2028034 to MS. We also thank members of the UCLA Phonetics Lab for their feedback on this work.

REFERENCES


182nd Meeting of the Acoustical Society of America, Denver, CO
25 May 2022