A Longitudinal Analysis of the Spectral Peaks of Vowels for a Japanese Infant

Kentaro Ishizuka, Ryoko Mugitani, Hiroko Kato, and Shigeaki Amano

NTT Communication Science Laboratories
NTT Corporation
Outline

• A longitudinal study of vowel development for a Japanese infant
• Speech data
  – From 4 to 60 months, over 18,000 vowels
• Change in spectral peaks with age
  – First and second spectral peaks
  – Vowel space expansion
Background

• Spectral peak analysis of children’s vowels
  – Peaks correspond to maturity of articulation
  – Many systematic studies after 3 years old

• Analysis of infants’ vowels (< 3 years old)
  – Articulatory organs mature rapidly before 3
  – Few studies (cover this age partially)

• Longitudinal studies [Buhr, 1980; Bond et al., 1982]
  – Few data
  – Qualitatively rather than quantitatively
Purpose

• Reveal early stages of vowel developmental process
  – Changes in spectral peaks with age
  – Covers 4 to 60 months
  – Longitudinal analysis
  – Quantitative analysis using many data
Acoustical analysis

- **Speech data**
  - Natural speech by a mother-child dyad
  - Japanese vowels: /a/, /e/, /i/, /o/, and /u/
  - Month age: 4, 6, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 55, and 60
  - 18,168 vowels in total

- **Spectral peaks extraction**
  - 12-order LPC analysis
  - Extract the two lower spectral peaks ($f_1$ and $f_2$)
  - Transformed into Bark scale
Mean $f_i$ values

- **Low vowel**
- **Middle vowel**
- **High vowel**

Diverges corresponding to the height of the articulation position

Decrease after 36 months
Mean $f_2$ values

Diverges corresponding to the horizontal articulation position

Fine-tuned after 30 months

Front vowel

Center vowel

Back vowel

Age (month)
$f_1$-$f_2$ plane with phoneme labels

- Vowel space expands with age
- Boundaries between vowels become clearer until 18 months
- Corresponds to anatomical findings
  [Vorperian et al., 2005]
Distance between vowels

Mahalanobis distance between vowels

The distance converges at around 18 months

The distance becomes large especially in early age
Quantitative analysis of vowel space expansion

Reference points
Mother’s “corner” vowel centroids

Categories
Π₁: vowel /a/: centroid \( c₁ \)
Π₂: vowel /i/: centroid \( c₂ \)
Π₃: vowel /o/: centroid \( c₃ \)

Discriminant function for vowel data \( d_j \)

\[
D_i^j = \left| d_j - c_i \right|^2 - \left| d_j - c_k \right|^2 \quad k = \mod(i / 3) + 1
\]
Histograms of $D_i$

Broadly distributed until 12 months

Biased to the two end of the histograms = Vowels are clustered around either of the two ref. points
Differences between histograms

- The differences become large until around 12 months, then decrease with age.
- The vowels clustered in the vowel space in terms of the physical acoustic phenomenon.
Conclusions

- Longitudinal vowel development analysis
- Longitudinal trends in vowel development correspond to the articulation positions
- The trends correspond to the rapid anatomic development of the vocal tract
Apendix A:
Number of analyzed vowels

<table>
<thead>
<tr>
<th>Age (month)</th>
<th># of vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-6</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>7-12</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>13-18</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>19-24</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>24-30</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>31-36</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>37-42</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>43-48</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>49-54</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
<tr>
<td>55-60</td>
<td>/a/ /e/ /i/ /o/ /u/</td>
</tr>
</tbody>
</table>
Apendix B:
Fitted polynomial curves

Polynomial curve for Mahalanobis distance

\[ \text{Dist} = a_0 + a_1 \cdot \left( \frac{\text{Age} + a_2}{a_3} + a_4 \right) \cdot \exp \left( a_5 \cdot \frac{\text{Age} + a_2}{a_3} \right) \]

Polynomial curve for differences between histograms

\[ \text{Diff} = a_0 + a_1 \cdot \left( \frac{\text{Age} + a_2}{a_3} \right) \cdot \exp \left( a_4 \cdot \frac{\text{Age} + a_2}{a_3} \right) \]