Technology Assisted Reading Acquisition (TARA): Children Acquiring Literacy Naturally

Dom Massaro
massaro@ucsc.edu

University of California, Santa Cruz
Psyentific Mind psyentificmind.com
Early Literacy (Current Belief)

• Speech is Special
  – Specialized Module
  – Is Acquired Unintentionally

• Reading is Artificial
  – Must be Taught
  – Wait Till Schooling Begins
Language and Speech are Special

independent language organ

independent course of development

not standard perception and learning

Noam Chomsky
Univ. of Maryland, Jan. 26, 2012
Value of the Face
Realistic Emotion and Affect
Baldi’s “visible speech” can be driven by either 1 a recorded voice or 2 text that is automatically converted into a synthesized voice. Either voice is automatically aligned with Baldi’s articulation.
Mandarin

Translation:
The new-age system for Taiwanese language involves word change.

Bao
Baldi Speaks

Aligned with Natural Speech
Baldi: upcoming virtual vocalist

Examples:
http://mambo.ucsc.edu/psl/primus.html

Primus
Examples:
http://mambo.ucsc.edu/psl/primus.html

verse 2
my notes on tap at my side

verse 3
a fluorescent tan covers my hide
Examples:
http://mambo.ucsc.edu/psl/primus.html

verse 2
my notes on tap at my side

verse 3
a fluorescent tan covers my hide
Realistic Speech
Facial control parameters.
1  jaw rotation
2  lower lip f-tuck
3  upper lip raising
4  lower lip roll
5  jaw thrust
6  cheek hollow
7  philtrum indent
8  lip zipping
9  lower lip raising
10 rounding
11 retraction
Speech "Visibility" Improves Comprehension

Percentage of Words Correct

Subject

Auditory Speech with Baldi
Auditory Speech Alone
Speech "Visibility" Improves Comprehension

- **Auditory Speech with Gary**
- **Auditory Speech Alone**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Percentage of Words Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10%</td>
</tr>
<tr>
<td>5</td>
<td>20%</td>
</tr>
<tr>
<td>9</td>
<td>30%</td>
</tr>
<tr>
<td>13</td>
<td>40%</td>
</tr>
<tr>
<td>17</td>
<td>50%</td>
</tr>
<tr>
<td>21</td>
<td>60%</td>
</tr>
<tr>
<td>25</td>
<td>70%</td>
</tr>
<tr>
<td>29</td>
<td>80%</td>
</tr>
<tr>
<td>33</td>
<td>90%</td>
</tr>
<tr>
<td>37</td>
<td>100%</td>
</tr>
<tr>
<td>41</td>
<td>100%</td>
</tr>
<tr>
<td>45</td>
<td>100%</td>
</tr>
<tr>
<td>49</td>
<td>100%</td>
</tr>
<tr>
<td>53</td>
<td>100%</td>
</tr>
<tr>
<td>57</td>
<td>100%</td>
</tr>
<tr>
<td>61</td>
<td>100%</td>
</tr>
<tr>
<td>65</td>
<td>100%</td>
</tr>
<tr>
<td>69</td>
<td>100%</td>
</tr>
</tbody>
</table>
Speech "Visibility" Improves Comprehension

27% Natural vs. 21% Synthetic
See What I’m Saying
Multi-Sensory Pattern Recognition
<table>
<thead>
<tr>
<th></th>
<th>BA</th>
<th>2</th>
<th>3</th>
<th>A</th>
<th>4</th>
<th>DA</th>
<th>none</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>none</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>