Infants’ Detection of Audiovisual Synchrony in Language Development.

George Hollich
Purdue University

“Are you synching what I’m synching?”

Synchro...y Matters

“Neurons that fire together wire together.”
- Hebb

Synchronous firing is critical to binding disparate areas of the neocortex.
- Edelman et al.

“Rhythm is gonna get you”
- Gloria Estefan.

A talk in two parts

Part I: Infants use synchrony to segment words from speech.

Part II: Infants use synchrony to learn the meaning of words.

“Synchrony is foundational to early language development.”

Part I

Audiovisual Synchrony & Speech Segmentation

In a world of many complicated signals.

Synchrony gets attention

If something visual is moving simultaneous with a sound... this can, literally, help you hear better.
Visual Hypothesis

Infants should be able to use the visual synchronization between the face and the speech stream to segment words from that stream in a noisy/blended stimulus.

In collaboration with Rochelle Newman & Peter Jusczyk

Design

Familiarization

Test

Three types of video

Synchronized Display - Video was synchronized with the target audio.

Unsynchronized Display - Video was the opposite of the target audio.

Static Display - Video was a single static frame presented throughout.

Sample Trial

Combined Results

Interim Conclusions

- Infants successfully segmented the speech stream at 0dB signal-to-noise ratio!
- +15dB over previous work without faces.
- Infants can use what they see to hear better.
- Synchronized visual information aids in stream separation and subsequent segmentation.
What caused these results?

- Face-specific/viseme-specific information?
- Labs of Massaro and Werker have shown infants to be poor at phoneme-specific integration.
- Perhaps ANY synchronized visual would help.

Oscilloscope Results (n = 26)

- Infants showed evidence of segmentation even when it was a correlated oscilloscope pattern.
- Infant domain general sensitivity to any form of synchronized visual information allows them to segment the speech stream.

Results (n = 26 * 4)

Part II

Audiovisual Synchrony & Word Learning
Gogate & Colleagues

- In an experimental task, infants only learn word meaning if object is moved synchronous with word!
- Observational data indicates mothers who use AV synchrony in labeling have children with higher vocabularies.

Emergentist Model

In Hollich, Hirsh-Pasek, and Golinkoff (2000), we proposed a model of an active word learner which has the following properties:

- **Multiple Cues** - Attentional, Social, Linguistic
- **Differential Weighting** over time
- **Emergent** properties

Familiar Phase

25mo Results (n = 20)

<table>
<thead>
<tr>
<th></th>
<th>Synch</th>
<th>Non-Synch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>Original</td>
<td>80</td>
<td>70</td>
</tr>
<tr>
<td>Diff Label</td>
<td>70</td>
<td>60</td>
</tr>
</tbody>
</table>

15mo Results (n = 20)

<table>
<thead>
<tr>
<th></th>
<th>Synch</th>
<th>Non-Synch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>80</td>
<td>70</td>
</tr>
<tr>
<td>Original</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td>Diff Label</td>
<td>60</td>
<td>50</td>
</tr>
</tbody>
</table>

Results

- Older infants use synchrony to reliably attach a label.
- Not so for younger infants.
- They likely are conservative in their labeling strategy.
Nonetheless

Synchrony is very helpful, and is one way whereby infants gain a toehold onto the process of segmenting words from speech and learning their meaning.

How?

Kinds of synchrony

- Onset
- Duration
- Amplitude

In Phase
Out of Phase

Need for Microgenetic Analysis

- All of this suggests that even something as simple as detection of audiovisual synchrony is more than an all or nothing process.
- To better understand the mechanism we need to know what is happening moment-by-moment, and have principled predictions about what infant behavior SHOULD look like, IF they are using a particular algorithm.

Sensory-oriented models

Prince & Hollich (2005)

Frame-by-Frame Coding

Prince & Hollich (2005)
Consider a preferential looking task with two faces -- only one of which is synchronized with the audio. (Pickens et al., 1994)

Not perfect but close

- Biggest effects around offsets and onsets.
- Children definitely become bored, leading to a switch in preference, suggesting should model habituation.
- At times, either visual or audiovisual models account for a significant portion of the data.
- Likely individual differences in integration ability, like to model that as well.

Synchrony Matters

- More modalities/neural assemblies in synch, the more stable the representation.
- Thus, synchrony helps highlight important aspects over external and internal background noise.