6.S077 Lecture 2

• For animals: Biology+Experience = what you know
• A general research strategy: How “poverty of stimulus” argument for bees extends to baby talk
• You can’t learn something unless...
• How babies learning to talk is like/unlike honeybee learning to navigate
• How to combine statistics with biology

• Infant = in+fant = ”no language”
• By end of the hour, you’ll learn part of how infants learn (part of) language – the language sound system and something about how infants learn (part of) the sound system of their language
• Like honeybees, and unlike honeybees: combine methods from statistics, and a universal, apparently pre-existing innate constraint
• In general, babies are like songbirds, at least for learning a sound system
Learning from Small (or almost no) Data Method

Honeybee (Apis), solar ephemeris function

Azimuth (direction of sun in sky), degrees
(North=0, East=90)

Local solar time (hour of day)

Shaded= space of all physically possible
Solar ephemeris functions
The honeybee learning picture

Experience from flight $\rightarrow$ + what’s initially in brain =

Final state = Biology + Experience

Key q: How much biology? How much experience?

How is this any different?

Any Learning System
1. Initial State
2. Final State
3. External experience
4. Learning algorithm

Biology + Experience = Final State suggests a research strategy...

(Pharoah Psamtik I did the experiment....) [without IRB approval...]

(sounds)
If no experience, revert to “ground genomic state” or “wild type”; this tells us what’s there at the start

Reversion to “Wild type” without experiential input:
Zebra finch
Initial state, human babies

Typical French cry

Typical German cry

Source: Mampe, Friederici, Christophe & Wermke, Current Biology, 2009

Human stages sound system

1. Reflex to Cooing: ‘vegetative’ sounds, crying (in rhythmic pattern of mother’s language), ...
2. Precanonical: Syllables ma, ...
3. Canonical: Repetitive babbling: na-na
4. First words

Let’s illustrate.
Compression of conversational speech—Poverty of the lexical (word) stimulus

Play first

Play second

Play third

Play fourth

Courtesy Warner and Tucker
Compression of conversational speech—Poverty of the lexical (word) stimulus

Gonna hafta = Going to have to

Chillinn’the = Chilling in the

Notice that this small “four word” excerpt becomes totally clear from the speech that follows it

Yet our conscious percept is that we analyze the initial fragment, as we hear it

This rehabilitates the notion of “psychological moment”, in which unconscious processes move back and forth, but the percept is linear
An animal model for human learning?

What’s the same?

- “Critical period” for learning from external experience
- Babbling (subsong), practice & self-practice
- Plasticity frozen at puberty (by hormonal change – testosterone)
- Left-lateralization for system
- Brain circuitry control
- Beads on a string structure
Songbirds – Zebra finch “critical period” learning

Subsong

Plastic song

Adult song
Human (and animal?) sound system like a “tweaked template” system – 8 “dials”

The Articulatory System

1. Lips
2. Teeth
3. Alveolar ridge (tooth ridge/gum ridge)
4. Hard palate (roof of the mouth)
5. Soft palate (velum)
6. Nasal passage
7. Tongue
8. Jaw
9. Vocal cords
All English sounds

Table 1. Distinctive Features of American English Consonants

|     | p | b | m | f | v | θ | d | n | s | z | l | r | ñ | s | j | f | d3 | j1 | k | g | n | w | ? | h |
| Back|  + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + |
| High|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Coronal|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Anterior|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Labial|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Continuant|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Lateral|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Nasal|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Sonorant|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Strident|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Voiced|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

Table 2. Distinctive Features of American English Vowels

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>e</th>
<th>æ</th>
<th>u</th>
<th>o</th>
<th>o3</th>
<th>a</th>
<th>ð</th>
<th>að</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>low</td>
<td>back</td>
<td>rounded</td>
<td>ATR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The simplest linear patterns = regular

\[ \text{ba:d} \rightarrow \text{bat}; \text{de:g} \rightarrow \text{dek} \] (Heinz, 2007)
\[ \text{fi:t}3, \ast \text{fi:te:z} \]
(Chandlee & Jardine, 2013)

Birdsong & human sound systems: what’s the same?

Bengalese finch
\((\text{Lonchura striata domestica})\)
Source: K. Okanoya, 2003
Sound system components: birds & people

“Beads on a string” model:
1. Beads – chunks or “states” that are categorical classes (remember: “s-sh”)
2. Linear sequence – one state follows another, in constrained way
   (e.g., “slo” starts a possible English word, but “rdz” does not)
   = A finite-state automaton

Categorical production and perception

Today: Let’s address just one part of that: how do we find the “chunks” in the input at all?

twasbrilligandtheslithytovesdidgyreandgimbleinthewabe
ekalaloakapaiiaokekahikolamuölelokanakamalokoonänüpepeczarnelubrdzawegontowestrzechyarkikryjącewsobiezakopcone

twas brillig and the slithy toves did gyre and gimble in the wabe
e kala loa ka pai ia o kekahi kolamu ölelo kanaka ma loko o nää nüpepaczarne lub rdzawe, gontowe strzechy i arki kryjące w sobie zakopcone
Challenge: segmentation
twasbrilligandtheslithytovesdidgyre

\{\text{pabiku, tibudo, golatu, daropi}\}

Challenge: Combining Inference with Cognitive Constraints
(How real people solve real problems can help real computers)

Problem: \textit{twasbrilligandtheslithytovesdidgyreandgimble}

“Standard” solution: \textit{prettybaby pre-ty-ba-by}
Graph of transition probabilities (tp): $\Pr(x_{i+1} \mid x_i)$ & look for local minima or threshold
“Standard” claim: works great; “stats is all you need” (\textit{Science}, 1996)

\begin{align*}
\Pr(\text{bi} \mid \text{pa}) &= 1.0; \quad \Pr(\text{ku} \mid \text{bi}) = 1.0; \quad \Pr(\text{ti} \mid \text{ku}) = 0.3, \\
\Pr(\text{bu} \mid \text{ti}) &= 1; \quad \Pr(\text{do} \mid \text{bu}) = 1.0; \quad \Pr(\text{da} \mid \text{do}) = 0.3 \\
\Pr(\text{ro} \mid \text{da}) &= 1; \quad \Pr(\text{pi} \mid \text{ro}) = 1.0; \quad \Pr(\text{go} \mid \text{pi}) = 0.3 \\
\Pr(\text{la} \mid \text{go}) &= 1.0; \quad \Pr(\text{tu} \mid \text{la}) = 1.0 …
\end{align*}

Works great? NO!!!

\begin{itemize}
\item \text{pabiku}
\item \text{pigola} \rightarrow \text{pi golatu}
\end{itemize}
Phonetically transcribed mother’s speech to “Adam” fed to segmentation program

“CMU pronunciation dict”: http://www.speech.cs.cmu.edu/cgi-bin/cmudict

1= primary stress
U= utterance end

Big drum.
Horse.
Those are checkers
bih1gdrah1mU
hhao1rsU
dhow1zaa1rcheh1kerzU
Actual results on actual speech to children? works lousy! What’s the answer? But, add a **ONE** universal constraint about human language and it works GREAT!

![Precision and Recall, Pure Stat Interference vs. Stat Inference + UG, 250,000 child-directed examples](chart)

**Using the universal constraint**

<table>
<thead>
<tr>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI</td>
<td>41.6</td>
</tr>
<tr>
<td>SI-UG</td>
<td>95.9</td>
</tr>
</tbody>
</table>

**transitional statistics**

What IS this ONE universal constraint???? HINT: you all **know** it!

---

**Summary for today**

- **Small data for small minds:** Statistics can be profitably combined with pre-existing (“innate”) constraints, even in humans
- **Sound system for language looks like a (bit more complex version) of “tweaked template” as in honeybee navigation**
- **Correct representation key to learning success**