How babies learn words and developing environments

Linda B. Smith
Indiana University
What is the role of the environment?

Why we are the way we are
What is essential
The malleable points in development
The mechanisms of change
Co-developing systems – the baby and the baby’s environment

This brain extends into the world


What we can do and thus the environments for development and learning change with development and learning.
Talk Overview

Developing Environments

over developmental time - an ordered set of lessons

in real time – the moment to moment experiences that drive change and learning

developmental niches

a pathways approach
Egocentric vision

3 objects at time
High resolution, tightly controlled in the lab

30 objects on the floor
Free-flowing, cluttered in the lab

At home, everyday experience

Yoshida & Smith (2008) Infancy
Yu et al (2009) IEEE Transactions on Autonomous Mental Development
Pereira, James, Jones & Smith (2010) Journal of Vision
Smith, Yu, Pereira, (2011) Developmental Science
Street, James, Jones & Smith (2011) Child Development
Yu & Smith (2012) Cognition
Yurovsky, Smith & Yu (2013) Developmental Science
James, Swain, Jones & Smith (2014) JCD
Pereira, Smith & Yu (2014), Psychological Bulletin & Review
Fausey, Jayaraman & Smith, (2016) Cognition
Yu & Smith (2016) Current Biology
Clerkin, Hart, Rehg, Yu & Smith (2016) Royal Society

Collaborations with
Chen Yu
Karin James
David Crandall
Egocentric vision
Gaze distribution toddlers in various contexts

Early studies 70° diagonal cameras
Middle studies 90° diagonal cameras
Current studies 120° diagonal cameras
At home 90° diagonal cameras

Bambach, Smith Crandall & Yu (2016)
Talk Overview

Developing Environments

over developmental time - an ordered set of lessons

in real time – the moment to moment experiences that drive change and learning

Developmental niches

a pathways approach
Home-view project  
(with the support of the National Science Foundation)

Building a corpus of developmentally indexed egocentric scenes

75 infants
3 weeks – 24 months
4 to 6 hours of head-camera video in the home
(no experimenters present, camera on hats)
records at 30Hz
over 54 million extracted images
Streams of 15 seconds of continuous recording sampled at 1/5 Hz
(once every 5 sec)
$R^2 = .42, F(1, 20) = 16.11. \ p < .001$

Mean frames sampled and analyzed per infant = 3310

Three individuals account for most of the faces

$R^2 = .14, F(1, 20) = 4.516. p < .05$. $R^2 = .16, F(1, 20) = 5.24. p < .05$. $R^2 = .23, F(1, 20) = 7.413. p < .05$
Proportion faces within 2 feet of head camera

$R^2 = .37, F(1, 20) = 13.61, p < .05$
It’s faces that decline with age, not people in view

Jayaraman, Fausey & Smith (2017) *Developmental Psychology*
It’s faces that decline with age, not people in view

Jayaraman, Fausey & Smith (2017) *Developmental Psychology*
(A) Faces

(B) Hands

(C) Faces minus Hands

Fausey, Jayaraman & Smith, (2016) Cognition
At all ages more than 70% of the hands in view are touching an object
Cumulative waking hours and face hours

Age in months

Cumulative waking hours
Cumulative face hours

0 500 1000 1500 2000 2500 3000 3500

0 2 4 6 8 10 12

4/5/17
Bloomington Indiana

Mumbai India

No age-related decline in faces in view

More different people

Same age-related decline in proximity of faces
The learning environment

Developmental time
The learning environment

Developmental time
Talk Overview

Developing Environments

over developmental time — an ordered set of lessons

in real time — the moment to moment experiences that drive change and learning

Developmental niches

a pathways approach
Multi-sensory project
(NSF, NICHD, Eye-Institute, AFOSR)
Chen Yu lead

Dual head-cameras, or head-mounted eye trackers
Motion sensors (hands, heads, eyes)
Audio
Multiple room cameras
Parent-infant play with multiple toys
Nearly 200 infants (longitudinal & cross-sectional)
9 months to 36 months
Birth &
12 months &
24 months

Object names

18 month olds

First words
Name explosion

4/5/17

Object names

speech
transcription
call response?

child's view

parent's view

bird eye view

child head motion

parent head motion

4/5/17
The Special Role of Hands

Objects come in and out of view, often one is larger (closer) than other in play, hands select these objects.

Yoshida & Smith (2008) *Infancy*
Yu et al (2009) *IEEE Transactions on Autonomous Mental Development*
Pereira, James, Jones & Smith (2010) *Journal of Vision*
Smith, Yu, Pereira, (2011) *Developmental Science*
Not just in the lab, not just in uncluttered contexts

One child, frames sampled 1/5 hz-- image size of baby doll

12X18 (216 grid squares)

This pattern characterizes visual experience and object naming at home and in more cluttered contexts – not just in the lab

Tay et al (in progress)
Sumarga et al (in progress)
Objects are often large and close in the infant view. These are associated with handling (by infant and to lesser degree by parent) and with being named by the parent.

Optimal moments for learning object names?
Novel Words & Objects

"zeebee"  "tema"  "dodi"

Parent-Toddler Free play

Word Learning Test

"habble"  "wawa"  "mapoo"

How do the naming moments associated with learned names differ from the naming moments associated not learned names?

Novel Words & Objects

"zeebee"  "tema"  "dodi"

"habble"  "wawa"  "mapoo"

Parent-Toddler Free play

Word Learning Test


Learn when the referent is visually salient (and more salient than competitors) for an *enduring* period of time *before and after naming*
a direct consequence of the baby’s body of how their sensory-motor system works, and how they interact with the world
The baby’s body


Head stability

Holding behavior

Head and eye alignment

Gaze distribution head camera image

4/5/17
in the toddler, visual attention and learning involve the whole sensory-motor system

real time coupling and self organization of eyes, head, and hands

*Learning object names is about the coordinated focus of head, eyes, and hand, stabilized visual attention, and the reduction of visual competition*
Developmental niches

This is a particular moment in developmental time when the baby’s body works this way and it may be a relevant factor in helping them rapidly learning words.
Talk Overview

Developing Environments

over developmental time - an ordered set of lessons

in real time – the moment to moment experiences that drive change and learning

Developmental niches

a pathways approach
8 to 10 month olds

Working on sitting stably object manipulation standing

Ahead of time of “traditional” view of start of word learning
Homeview Project

Building a corpus of developmentally indexed egocentric scenes in the home

- 75 infants (to date)
- 3 weeks – 24 months
- 4 to 6 hours of head-camera video in the home (no experimenters present)
  (no experimenters present, camera on hats)
- records at 30Hz
- over 5 million extracted images
The **visual frequency of objects** in scenes at 8-10 months – but not the frequency of their names -- predicts **age of acquisition of object names**

8 infants 8 to 10 months of age
**147 “mealtimes”**
Images sampled every 5 sec
Sample of 6000 images

Clerkin, Smith & Yu (in progress)
20 infants
**300 “mealtimes”**
Images sampled every 5 sec
Sample of 32,000 images
Coding –the visual objects in the scenes

About 500 naïve adults (Amazon Mechanical Turk)
Task: name the five most obvious objects in the scene, using basic level nouns
Coders saw images in sequential sets of 20
Each scene was coded by 4 individuals
Highly cluttered scenes
Lots of uncertainty
745 unique object categories
Object name categories

Of the 745 unique object labels provided by Turkers

**First Nouns** (133) – normatively learned by 16 months (Communicative Developmental Inventory - CDI)

**Early Nouns** (59) – normatively learned between 16 and 30 months (Toddler CDI)

**Later Nouns** (553) – other nouns named by coders that were not on either inventory

- AoA: 6.15 years (SD=1.56)
Cumulative waking hours and face hours

Unique basic level categories (ordered by frequency)

First – under 16
Early – 16 to 30 months
Later – average AoA 5.5 years

4/5/17
<table>
<thead>
<tr>
<th>First Nouns</th>
<th>Early Nouns</th>
<th>Later Nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table</td>
<td>Tray</td>
<td>Shelf</td>
</tr>
<tr>
<td>Chair</td>
<td>Washing Machine</td>
<td>Bag</td>
</tr>
<tr>
<td>Shirt</td>
<td>Napkin</td>
<td>Container</td>
</tr>
<tr>
<td>Bowl</td>
<td>Knife</td>
<td>Cabinet</td>
</tr>
<tr>
<td>Bottle</td>
<td>Tissue</td>
<td>Lid</td>
</tr>
<tr>
<td>Spoon</td>
<td>Basket</td>
<td>Counter</td>
</tr>
<tr>
<td>Window</td>
<td>Sofa</td>
<td>Fireplace</td>
</tr>
<tr>
<td>Cup</td>
<td>Dryer</td>
<td>Bin</td>
</tr>
<tr>
<td>Plate</td>
<td>Bench</td>
<td>Tablecloth</td>
</tr>
<tr>
<td>Glasses</td>
<td>Can</td>
<td>Straw</td>
</tr>
<tr>
<td>Food</td>
<td>Yogurt</td>
<td>Handle</td>
</tr>
<tr>
<td>Door</td>
<td>Bucket</td>
<td>Seat</td>
</tr>
<tr>
<td>Telephone</td>
<td>Sauce</td>
<td>Wood</td>
</tr>
<tr>
<td>Couch</td>
<td>Belt</td>
<td>Outlet</td>
</tr>
<tr>
<td>Picture</td>
<td>Walker</td>
<td></td>
</tr>
</tbody>
</table>

**Type Occurrence by Group**

- Mean
- Median

4/5/17
Cumulative waking hours and face hours

Unique basic level categories (ordered by frequency)

First – under 16
Early – 16 to 30 months
Later – average AoA 5.5 years

Number of Occurrences

Unique basic level categories (ordered by frequency)
Objects in scenes
Parent names (nouns)

Frequency

Rank order by object frequency
The visual frequency prior to systematic parent naming predicts normative age of acquisition
The pathway to the first mapping of heard names to seen things begins prior to word learning, in the statistics of visual objects in the baby’s visual scenes.
Developmental Niches

10 month olds

A few high frequency objects
Many rarer objects
Clutter
Few heard names

18 month olds

Objects handled
One dominant object in scene
Many more and different objects
Names for visually dominant objects

Incremental statistical learning about visual objects?

In the moment visual selection and mapping names to things?
Birth & 12 months & 24 months

First words
Name explosion

Object names

There's your ball
Changing statistics in collaboration with Jason Gold, Rowan Candy, Christina Deserio
Learning: Brains, Machines and Children
Emerging Area of Research Award Indiana University

Olaf Sporns    Michael Ryoo
Karin James    Martha White
Chen Yu        Sriraam Natarajan
Linda Smith    David Crandall
Franco Pestilli Michael Jones
Rob Goldstone  David Landy
(not pictured)  

4/5/17
A deep learning network
Winner!
One Punch Knockout
(Initial slow learning)
Rapid one trial Learning
Immediate correct generalization to new instances
Good in clutter
Good in suboptimal viewing conditions
Recognition from many views

Vs.

Only recognizes what trained on
A Unified Theory of Learning for Brains, Machines and babies?

How training sets are selected by the learner’s own activities in time
How training experiences are ordered (real time, developmental time)
How learning across different environments builds itself

How later learning reuses earlier learning

How learning environments vary and the consequences of these differences

Creating hidden competencies and deficits with long term consequence for the learner and learning to learn
**Home-view**

- **faculty**
  - Chen Yu
  - Rowan Candy
  - Jason Gold

- **post-docs**
  - Swapnaa Jayaraman
  - Sven Bambach
  - Drew Abney

- **graduate students**
  - Elizabeth Clerkin
  - Hadar Karmazyn Raz
  - Christina DeSerio
  - Charlene Tay

- **technicians**
  - **Charlotte Wozniak**
    - Tom Dedek

**Lots of wonderful undergraduates**

- Financial Support: NSF, NIH, AFOSR, Indiana University

---

**Multisensory**

- **Chen Yu**
- David Crandall

- **Sumarga Suanda**
- Lauren Slone
- Drew Abney
- Sven Bambach
- Lei Yuan

- **Catalina Suarez-Rivera**
- Yayun Zhang
- Linger Xu

- **Seth Foster**
- Charlene Tay

- **Charlotte Wozniak**
  - Anting Chen