Bootstrapping in language acquisition:
A computational investigation of word learning and lexical category acquisition

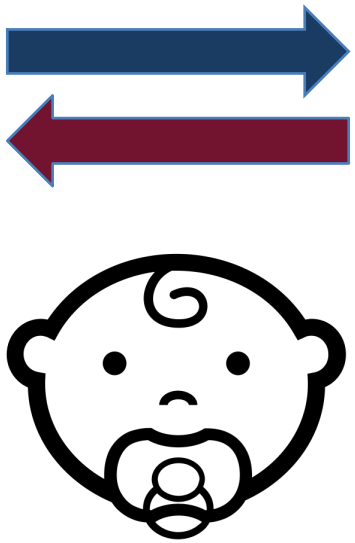
Giovanni Cassani & Robert Grimm

26th August 2016 – CLiPS meeting

CLiPS
Computational Linguistics & Psycholinguistics
University of Antwerp

Bootstrapping language learning

Learn words to infer phonemic contrasts



Learn phonemes to recognize words

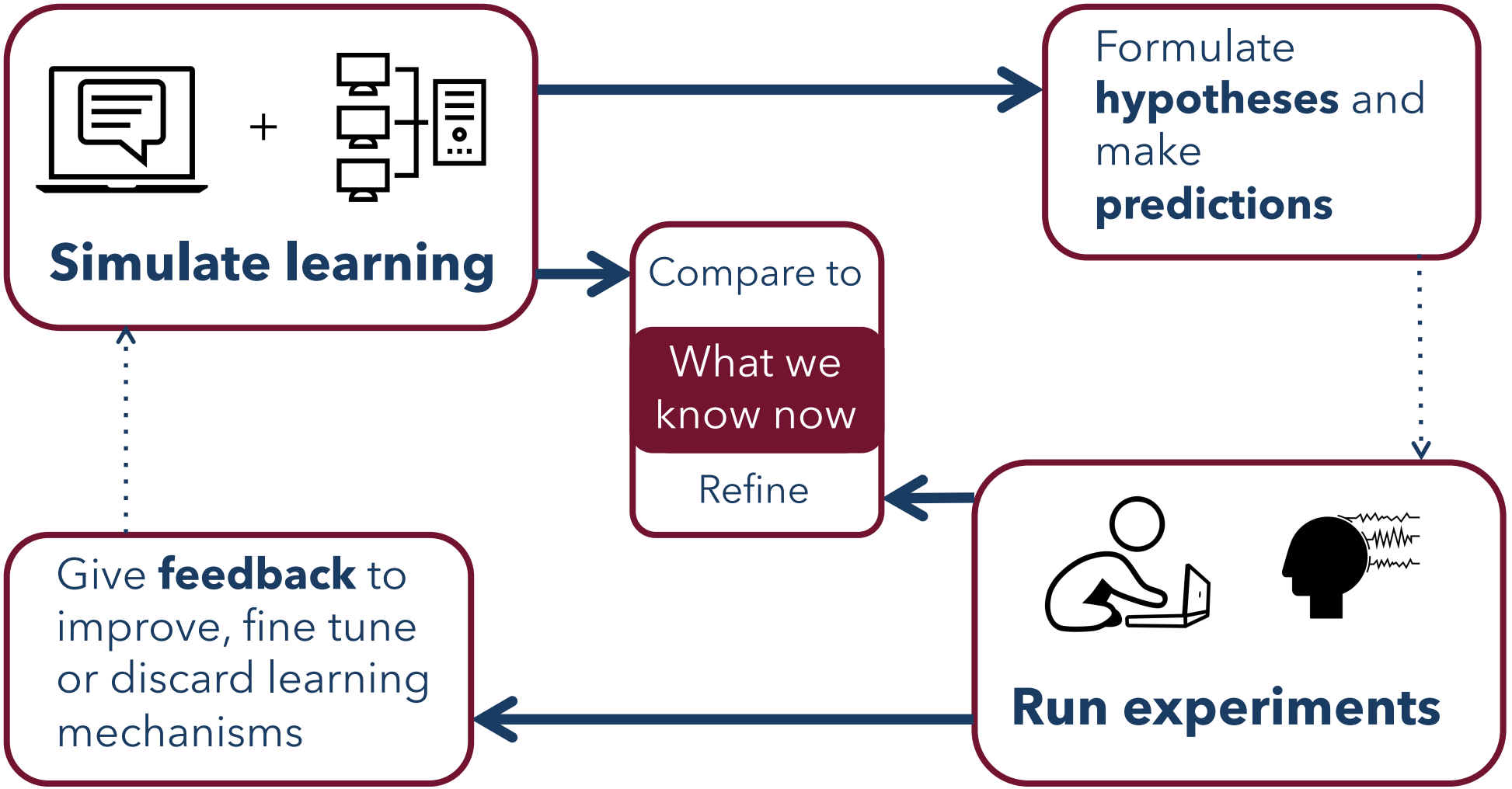
Segment speech to recognize linguistic units

Use linguistic units to segment speech

Overview

- Why modeling?
- Learning words
  - What's been done
  - What could we do better
  - What have we done so far
- Acquiring lexical categories
  - State-of-the-art
  - Missing links
  - Understanding as unraveling
  - How I am doing it

The bootstrap loop



The bootstrap loop

Simulate learning

Formulate **hypotheses** and make **predictions**

Compare to

What we know now

Refine

Run experiments

Give **feedback** to improve, fine tune or discard learning mechanisms

The simulation gain

Reverse-engineer the learning child

Run experiments considering the full input children receive (linguistic and extra-linguistic) and disentangle causal effects afterwards

Explore many more possibilities and generate grounded, motivated hypotheses

Overview

- Why modeling?
- Learning words
  - What's been done
  - What could we do better
  - What have we done so far
- Acquiring lexical categories
  - State-of-the-art
  - Missing links
  - Understanding as unraveling
  - How I am doing it

Models of Word Learning

- Model particular aspects of word learning
  - Map sound to meaning representation
  - Learn to abstract across speaker-specific word forms
  - Identify intended referent from set of possible referents

- Match model results to behavioral measures
  - Lexical confusion effects
  - More speaker variability = faster learning
  - Objects that appear in many different scenes are more easily linked to their labels

What could we do better

- Holistic rather than fragmented view of word learning
  - Combine sub-tasks to model entire process of word learning

- Improve evaluation of fit for model and human data
  - Increase number of data points
  - Quantify degree of fit

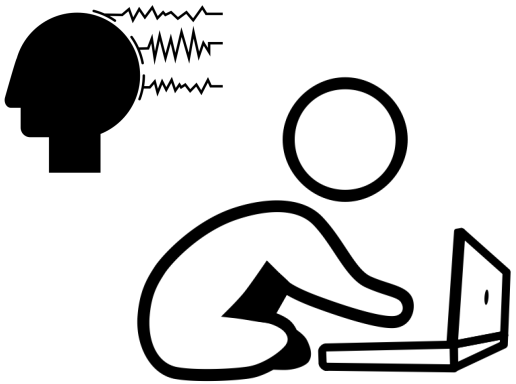
What have we done so far

- Two related aspects of word learning
  - Segment smaller units from incoming speech stream
  - Map segmented units to meaning representations
- What is the effect of:
  - Time of segmentation
  - Associative links between word forms
- Measure effect of model-derived statistics on Age of Acquisition

Overview

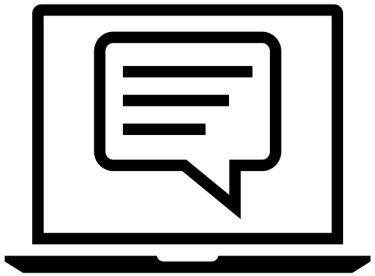
- Why modeling?
- Learning words
  - What's been done
  - What could we do better
  - What have we done so far
- Acquiring lexical categories
  - State-of-the-art
  - Missing links
  - Understanding as unraveling
  - How I am doing it

State-of-the-art



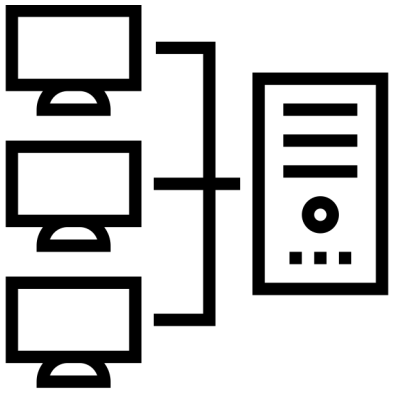
Experimental:

What does it take for children to form categories and what do they look like



Corpus studies:

What information is in the child-directed speech that is relevant to category learning



Computational:

What learning mechanisms succeed in learning categories

Missing links

“That doesn’t seem plausible…”

Little formal characterization of learning mechanisms

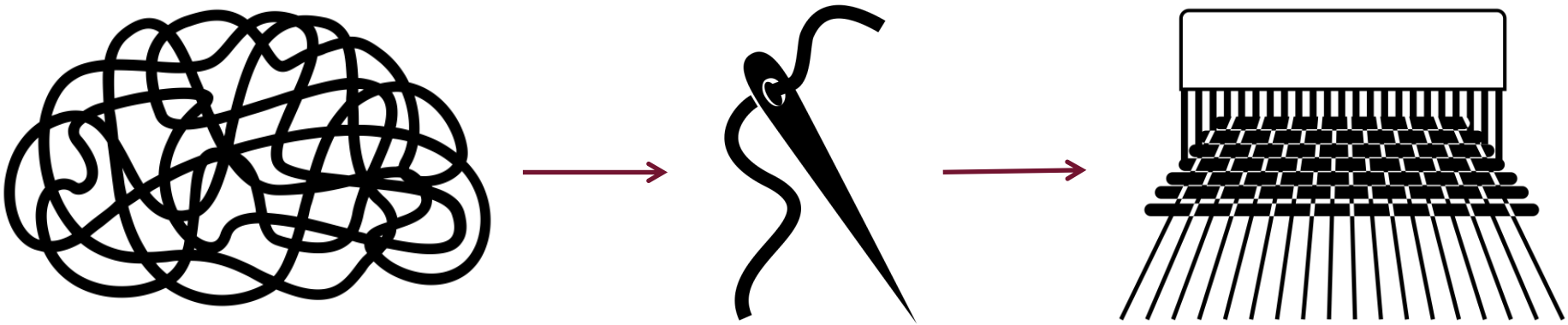
It is learned, but how?

Few connections between learning outcomes, what drives learning and how it unfolds over time

Scaling up

Few links between laboratory results, real scenarios and simulations

Understanding as unraveling



Recognize the full problem, isolate sub-problems, consider and place them in a broader picture.

How I am doing it

Restrict starting assumptions

Avoid pre-supposed, high-order linguistic units and representations

Harness computational models to **explore complex relations** between different parts of the input together

Test the same predictors on different tasks to place possible causes in a **broad, interconnected picture**

Link important information to the mechanisms that use it to solve a task **the way children appear to do**

Thank you!

Questions?

Credits

All icons come from **The Noun Project**.

 *Created by* Cindy Duong

 *Created by* Christopher Smith

 *Created by* Icon Fair

 *Created by* Aenne Brielmann

 *Created by* Alfredo Hernandez

 *Created by* Claire Jones

 *Created by* Mister Pixel

 *Created by* emilegraphics

Ghost track: Bestiary (or *can you* ***not****?*)

- Train and test a Neural Network on the same data (*Cognition,* 2011)
- Estimate a parameter from behavioral data; run a model that relies on the parameter using the same data used for estimation; use high correlation between model output and data to prove the model mimics humans (*Cognitive Psychology,* 2013)
- "Marginally significant", "Trend towards significance", … (*Everywhere,* any year)
- Use the most complex model possible