Computational Cognitive Science

Lecture 14: Language Acquisition

Frank Keller

5 November 2026

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

Overview

Background on Word Learning
Modeling Word Learning
- Bayesian Formulation
- Generative Model
- Evaluation

Results
- Lexicon and Referent Accuracy
- Mutual Exclusivity
- Object Individuation

Reading: Frank et al. (2009).

Background on Word Learning

Word Learning

We have talked about how people learn categories, concepts, and causal relations.

Now we will turn to how children learn language. In particular, we will look at word learning (lexicon acquisition).

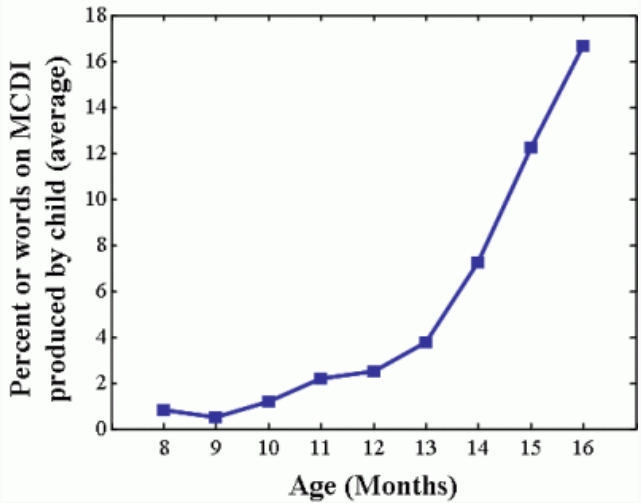
Psychological studies show that:

- between birth and adulthood, children learn about 60,000 words (8–10 words per day on average);
- during the second postnatal year, word learning accelerates dramatically: vocabulary explosion;
- often children learn a new word based on a single example of its use: one-trial learning.

This is an amazing developmental feat. How is it possible?

Word Learning

Growth curve from the MacArthur-Bates Communicative Development Inventory (MCDI):



Referential Ambiguity

Philosophers think this should be impossible (Quine 1960):

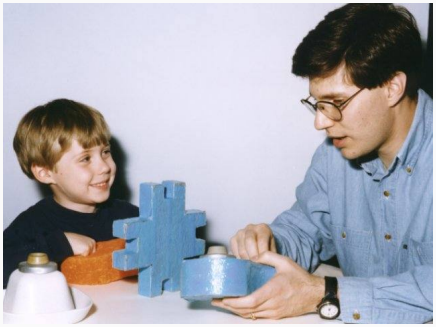


Developmental psychologists think:

- referential ambiguity is a serious problem, but:
- concrete nouns are learned first, then verbs, adjectives, abstract nouns follow;
- coincides with the development of syntax (two-word stage);
- one-trial learning happens when a new word is encountered in a context in which the other objects are familiar.

Social Learning

Crucially, children rely on social context to learn words. They infer other peoples' intentions based on:



- eye gaze;
- body position;
- pragmatics/salience.

Evidence for this comes from:

- apparent lack of learning from video;
- tracking of others' gaze at six months;
- learning new words using gaze at 18 months.

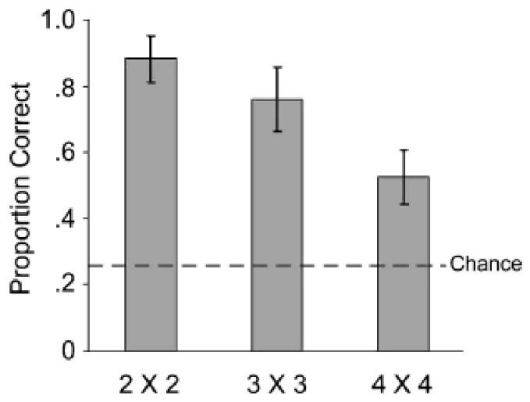
Cross-situational Learning

Cross-situational learning resolves referential ambiguity over time:



Cross-situational Learning

Evidence: both adults and infants can learn word meanings from cross-situational experience in the lab (Yu and Smith 2007):



$n \times n$ condition: trial presented n words and n possible referents.

Modeling Word Learning

Bayesian Formulation

Frank et al. (2009) propose a Bayesian model of word learning:

- combines cross-situational learning and inferring speaker's intention;
- learns a lexicon L from a video corpus C annotated with objects present and words spoken.

Goal: infer a lexicon given a corpus:

$$P(L|C) \propto P(C|L)P(L)$$

Here, L is a set of word-object pairs. $P(C|L)$ is defined using a generative model, and $P(L)$ favors smaller L:

$$P(L) \propto e^{-\alpha|L|}$$

(Remember the size prior in the model of concept learning in lecture 9!)

Word Learning as Joint Inference

Word learning creates a "chicken-and-egg" problem:

Lexicon $\Longleftrightarrow$ Speaker's intention

- If we know some word meanings, we can infer what the speaker is talking about.
- If we know what the speaker is talking about, we can infer what new words mean.

Frank et al.'s key idea is to solve these two problems simultaneously:

$$P(L, I|O, W)$$

The lexicon helps infer intentions, and intentions help infer the lexicon.

Time for a short quiz on WooClap!

Go to wooclap.com

Enter the event code in the top banner

Event code

HZAQQPM



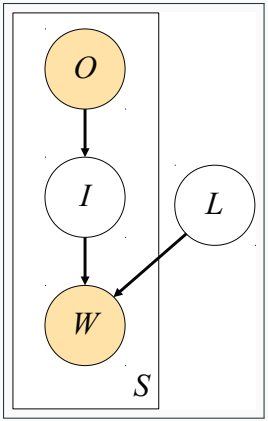
Generative Model

The corpus consists of independent situations. For each situation $s \in C$:

- O is the set of objects present in the situation;
- chose $I \subseteq O$, the set of intended referents (objects the speakers wants to talk about); note that I can be empty;
- generate the words W the speaker utters, conditional on I and L (L includes referential and non-referential words).

In practice, situation = utterance.

Graphical model notation: shaded circle: observed random variable; empty circle: hidden random variable; arrow: conditional dependence; plate: replicated S times.



Generative Model

$$
\begin{aligned}
P(C|L) &= \prod_{s \in C} P(O_s, W_s|L) \\
&= \prod_{s \in C} \sum_{I_s \subseteq O_s} P(O_s, I_s, W_s|L) \\
&= \prod_{s \in C} \sum_{I_s \subseteq O_s} P(O_s) P(I_s|O_s) P(W_s|I_s, L) \\
&\propto \prod_{s \in C} \sum_{I_s \subseteq O_s} P(I_s|O_s) P(W_s|I_s, L)
\end{aligned}
$$

Because intentions are unobserved, we sum over all possible intentions I_s.

Generative Model

Generate intentions from objects: uniform distribution:

$$P(I_s|O_s) \propto 1$$

Generate words from intentions and lexicon: words are independent. For each word w in W_s:

- choose referential ($p = \gamma$) or non-referential ($p = 1 - \gamma$):

$$P(W_s|I_s, L) = \prod_{w \in W_s} \left[\gamma \sum_{o \in I_s} \frac{1}{|I_s|} P_R(w|o, L) + (1 - \gamma) P_{NR}(w|L) \right]$$

- $P_R(w|o, L)$: choose uniformly from lexical items that refer to correct object;
- $P_{NR}(w|L)$: choose uniformly from all words in corpus.

Data

Small hand-annotated corpus:

- two 10-minute videos of mothers playing with preverbal infants using a small number of toys;
- speech is transcribed;
- each line of transcription is annotated with all visible mid-size objects, and true intention (for evaluation only).

words	“do bunnies go jumping through the forest?”
objects	BOOK, BIRD, RATTLE, MIRROR, BUNNY
intention	BUNNY

Evaluation

Evaluate model predictions against:

- gold-standard lexicon (word-object pairings);
- gold-standard intentions for each utterance (coded manually).

Compute precision (proportion of pairings that were correct) and recall (proportion of total correct pairings that were found).

Compare against related models:

- simple statistics (co-occurrence frequency, conditional probability, mutual information);
- translation model: baseline that learns word-object associations across situations, without representing speaker intentions.

Results

Results: Lexicon Accuracy

Model	Precision	Recall	F score
Association frequency	.06	.26	.10
Conditional probability (object|word)	.07	.21	.10
Conditional probability (word|object)	.07	.32	.11
Mutual information	.06	.47	.11
Translation model (object|word)	.07	.32	.12
Translation model (word|object)	.15	.38	.22
Intentional model	**.67**	**.47**	**.55**
Intentional model (one parameter)	.57	.38	.46

Take-home: jointly modeling intentions reduces spurious word-object mappings.

Results: Lexicon Accuracy

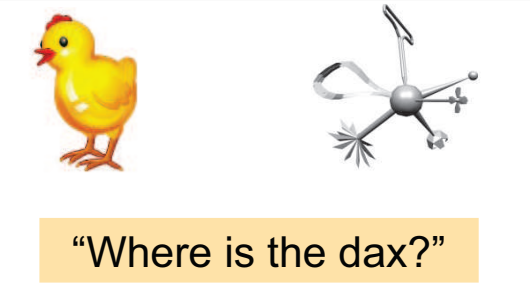
Word	Object
bear	**bear**
bigbird	**bird**
bird	**duck**
birdie	**duck**
book	**book**
bottle	bear
bunnies	**bunny**
bunnyrabbit	**bunny**
hand	**hand**
hat	**hat**
hiphop	mirror
kittycat	**kitty**
lamb	**lamb**
laugh	cow
meow	baby
mhmm	hand
mirror	**mirror**
moocow	**cow**
oink	pig
on	ring
pig	**pig**

Results: Referent Accuracy

Model	Precision	Recall	F score
Association frequency	.27	**.81**	.40
Conditional probability (object|word)	.59	.36	.45
Conditional probability (word|object)	.32	.79	.46
Mutual information	.36	.37	.37
Translation model (object|word)	.57	.41	.48
Translation model (word|object)	.40	.57	.47
Intentional model	**.83**	.45	**.58**
Intentional model (one parameter)	.77	.36	.50

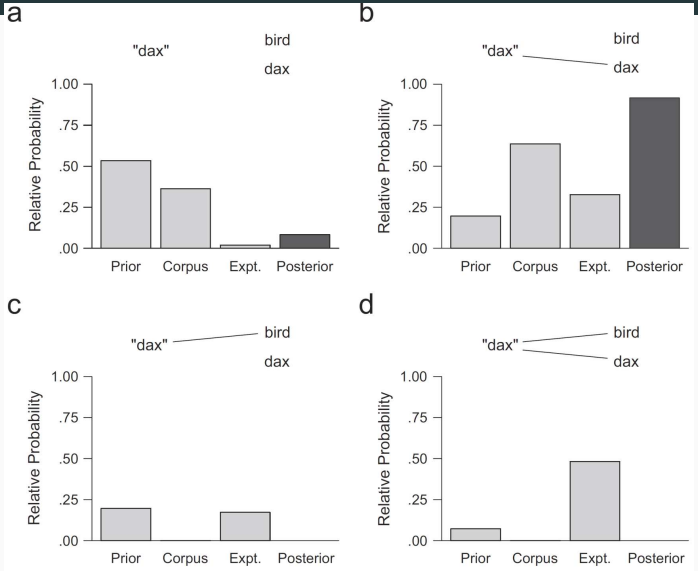
Results: Mutual Exclusivity

Children as young as 16 months map novel words to novel objects:



- some researchers have postulated a principle of mutual exclusivity to account for this;
- but it could also be general pragmatic principles at work;
- is mutual exclusivity learned or innate?

Results: Mutual Exclusivity



Corpus: likelihood of the corpus data; Expt: likelihood of experimental situation.

Results: Mutual Exclusivity

The model prefers to map the novel word *dax* to the novel object.
Why is mapping *dax* to BIRD unlikely?

- The lexicon already contains a well-supported mapping: *bird* $\rightarrow$ BIRD.
- If both *bird* and *dax* refer to BIRD, then every earlier use of *bird* becomes less probable: $P(\textit{bird}|\text{BIRD}, L)$ is reduced because another label could have been used.
- This lowers the likelihood of the previously observed corpus.

The prior also favors sparse lexicons, so unnecessary additional word-object mappings are penalized.

Mutual exclusivity emerges from Bayesian inference; it does not need to be built in as a separate principle.

This also gives the model a form of one-trial learning.

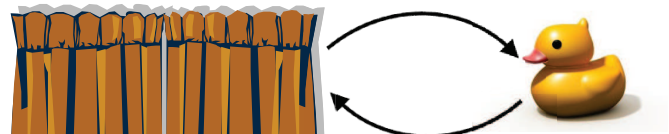
Results: Object Individuation

Infants as young as 12 months can use words to individuate objects.
Experiment by Xu (2002):

1. an object emerges behind a screen, they hear a word, the object disappears again;
2. a different object emerges and disappears, they either hear the same word or a different word;
3. the screen disappears revealing either one or two objects.

Children show longer looking times when the number of objects in step 3 mismatches the number of words in steps 1 and 2.

Results: Object Individuation

1. 

"Look, a toy!"

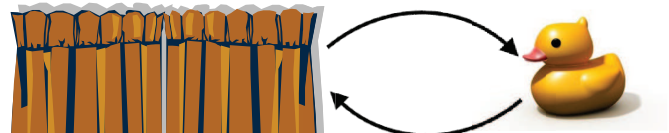
2.

"Look, a toy!"

3. 

vs. 

Results: Object Individuation

1. 

"Look, a duck!"

2.

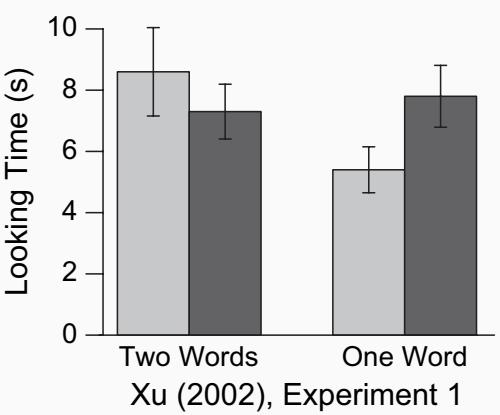
"Look, a ball!"

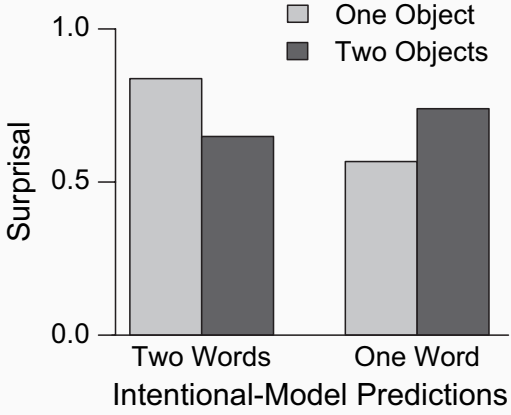
3. 

vs. 

Results: Object Individuation

Captured by the model in terms of surprisal (more on surprisal in lecture 16):





Discussion

Strengths:

- model combines cross-situational and social/intentional learning;
- corpus evaluation: does the model recover word-object mappings and intentions?
- qualitative/quantitative behavioral predictions: mutual exclusivity, one-trial learning, object individuation, intention reading, etc.

Weaknesses:

- only tested on very small corpus;
- only deals with concrete nouns;
- no model of syntax.

Summary

- Infants learn words based on multiple cues (eye gaze, body position, pragmatics/salience).
- Encountering a word in multiple situations is crucial for word learning.
- The model of Frank et al. (2009) combines observed words and objects with latent speaker intentions to infer a lexicon.
- It aggregates information across situations.
- It outperforms a simple translation model and captures mutual exclusivity and object individuation.
- Next week: how to model lexicon acquisition with realistic video data and modern computer vision/machine learning.

References

Frank, Michael C., Noah D. Goodman, and Joshua B. Tenenbaum. 2009. Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science* 20(5):578–585.

Quine, Willard. 1960. *Word and Object*. MIT Press, Cambridge, MA.

Xu, F. 2002. The role of language in acquiring object concepts in infancy. *Cognition* 85:223–250.

Yu, Chen and Linda B. Smith. 2007. Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science* 18(5):414–420.