Remy "helps" to read "Syntactic categories in the speech of young children" (Valian 1986)

Language Acquisition

9.85 – Infant cognition

Kim Scott

With many thanks to and slides from Melissa Kline

Mini-lectures

- Do we teach children to speak?
- Do we help children to speak?

What's amazing about language?

Human language is...

- **Productive (potentially infinite)**
  - The Red Sox beat the Yankees.
  - Mario said that the Red Sox beat the Yankees.
  - Jill thought that Mario said that the Red Sox beat the Yankees...

  Used in everyday communication...
  - i hope you did not say i told you he thought your chances were bad

Human language is…

- Productive (potentially infinite)
- **Context-independent:** We can talk about things that aren't…
  - Present
  - Visible
  - Tangible
  - Real

Human language is...

- Productive (potentially infinite)
- Context-independent
- **Uniquely human**

Chimp vocalizations...

- Pant-hoot - food enjoyment, social excitement, and sociability feelings
- Pant-grunt - directed towards dominant individuals by submissive individuals
- Distance calls
  - To draw attention to danger
  - To draw attention to food sources
  - To establish location of other groups in the area
- Short bark – when hunting
- Tonal bark – "given in the presence of large snakes"

Pant-hoot

http://www.youtube.com/watch?v=p3b-MBaa3bg

Alex (Avian Learning/Language Experiment)

- "Which one Two?"

- "How many Red?"

- "How many Wood?"

http://www.youtube.com/watch?v=7yGOgs_UlEc

Alex parrot experiment photos removed due to copyright restrictions.
http://alexfoundation.org/?page_id=183&show=gallery

What's missing?

Human language is…

- Productive (potentially infinite)
- Context-independent
- Uniquely human
- **Richly structured**

(not prescriptive rules)

Don't say "ain't"

Don't end sentences in prepositions:

    That is something which I can't put up with.

    *That is something up with which I cannot put.

Don't split infinitives:

    To boldly go where no man has gone before.

    ? To go boldly where no man has gone before.

    ? Boldly to go where no man has gone before.
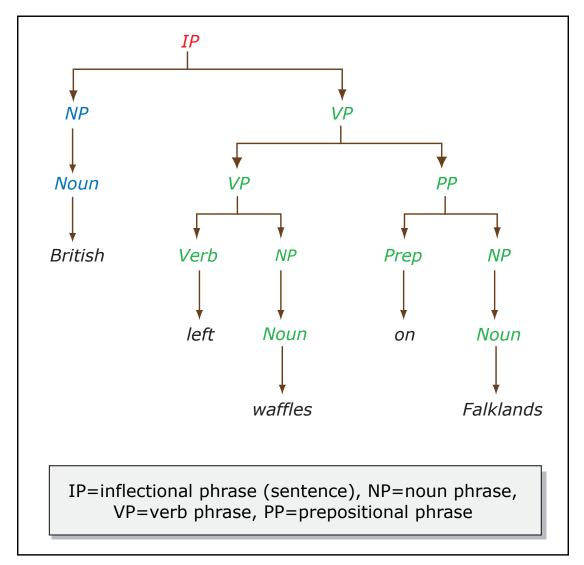
Human language is richly structured

Colorless green ideas sleep furiously

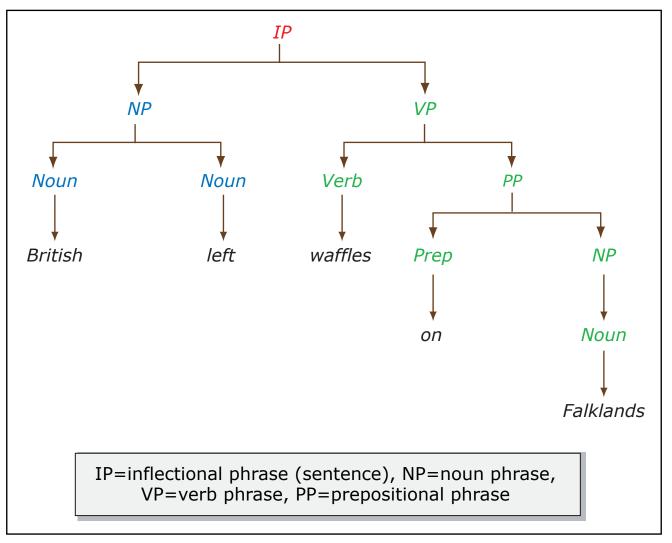* Furiously green sleep ideas colorlessly

The big grey cat

* The grey big cat

"British left waffles on Falklands"

IP=inflectional phrase (sentence), NP=noun phrase, VP=verb phrase, PP=prepositional phrase

Image by MIT OpenCourseWare.

"British left waffles on Falklands"

IP=inflectional phrase (sentence), NP=noun phrase,
VP=verb phrase, PP=prepositional phrase

Human language is…

- Productive (potentially infinite)
- Context-independent
- Uniquely human
- Richly structured
- **LEARNABLE**

In fact, everyone learns it…

"It is a very remarkable fact that there are none…that they cannot arrange different words together, forming of them a statement by which they make known their thoughts; while on the other hand, there is no other animal, however perfect and fortunately circumstanced it may be, which can do the same."

Renee Descartes

In fact, everyone learns it…

(barring very extreme situations)

...On roughly the same timescale

...In roughly the same order

...Making similar errors

...Across all languages studied

Progression of language acquisition

- Vocal play (16wks-6 months)
- Babbling (6-10 months)
- "Single-word stage" (10-18mo)
  - *Kitty, dada, up!, allgone, wassat?, bye-bye*
- "Two-word stage" (~18mo)
  - *Byebye plane, See baby, More hot, I shut*
  - Vocabulary growth increases rapidly
- Telegraphic speech (~2yo)
  - *Where wrench go?*
  - Grammatical elements start to appear, in a relatively fixed order
- Full, adult like sentences (~2.5yo)

Levels of structure

- Children must learn:
  - Phonetics & Phonology
  - Lexicon
  - Syntax & Morphology
  - Semantics
  - Pragmatics/Discourse

Levels of structure

- Children must learn:
  - Phonetics & Phonology
  - Lexicon
  - Syntax & Morphology
  - Semantics
  - Pragmatics/Discourse

Minimal units of meaning:
cat, dog, mouse, see, watch, -s, -ing

Levels of structure

- Children must learn:
  - Phonetics & Phonology
  - Lexicon
  - Syntax & Morphology
  - Semantics
  - Pragmatics/Discourse

Structure of words (combining lexical items)

Structure of sentences (combining words)

Syntax vs. Semantics

Syntactically correct / * Syntactically incorrect:

- Colorless green ideas sleep furiously.
- *Furiously sleep ideas green colorless.

- *The keys to the cabinet is on the counter

- I wiped the counter clean
- *I cleaned the counter shiny

- The fish swam/*I swam the fish
- The lamp broke/I broke the lamp

- Jabberwocky

Pragmatics/Discourse

Bill hid Dave's car keys -- he was drunk. It wasn't the first time that Dave had gotten that drunk.

*Bill hid Dave's car keys -- he hates spinach. It wasn't the first time that they had seen the movie.
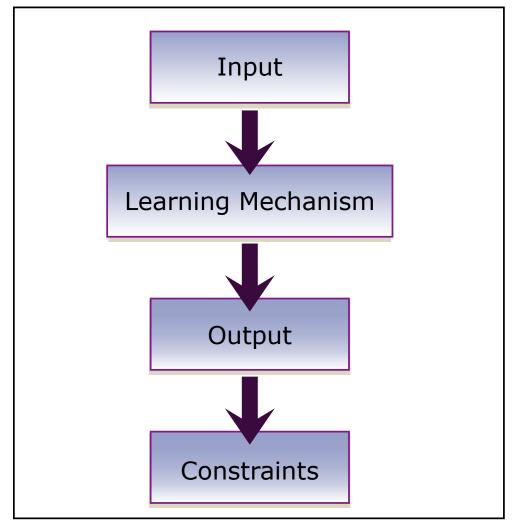
Plan for today

Input

↓

Learning Mechanism

↓

Output

↓

Constraints

Image by MIT OpenCourseWare.

The input

Infants get a wide variety of language experience.

- – Overheard adult conversations
- – **Motherese / child directed speech**
- – Interaction
- – Responses to own utterances
- – Corresponding events, contexts

What is it that they really need?

Motherese

- Child-directed speech is
  - Slower
  - More animated
  - Wider pitch variation
  - Wider vowel contrasts
  - Preferred by newborns (Cooper and Aslin 1990)
  - ~60% grammatical sentences, the rest well-formed clauses and stock phrases (Newport, Gleitman, & Gleitman 1977)
- Near-universal: Used by parents and non-parents, in almost all cultures

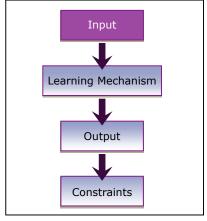# Is motherese a specialized teaching language?

Image by MIT OpenCourseWare.

- Grammatical features
    - Heavy on questions ("Where's your nose?) and commands ("Get your feet out of the laundry!"); only ~30% declarative sentences
    - Syntactic structures not introduced sequentially
    - Some tuning to what the child can understand (receptive language), but not what the child can produce
- Hard to find independent effects of quality/quantity of CDS on development
- Probably helpful in phonological and word learning, communicating affect and intention, **drawing attention to speech … but doesn't "teach" language.**
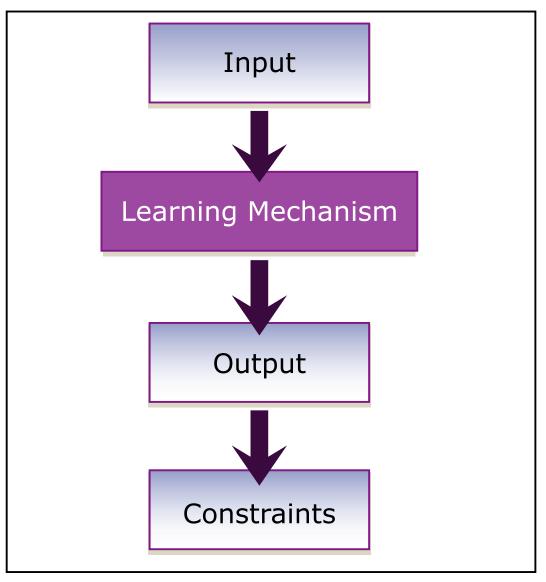
28

Input

Learning Mechanism

Output

Constraints

Image by MIT OpenCourseWare.

Can imitation and reinforcement explain language learning?

Behaviorism (Skinner):

– Language is a behavior

– Language learning is like any other learning… behavior shaping

What are some challenges for this approach?

Image by Ross Hawkes on Flickr.

Can imitation and reinforcement explain language learning?

Behaviorism (Skinner):

– Language is a behavior

– Language learning is like any other learning… behavior shaping

What are some challenges for this approach?

– Creativity and error

– Absence of (use of) appropriate corrections

– Enrichment of the input (deaf children of hearing parents, creoles)

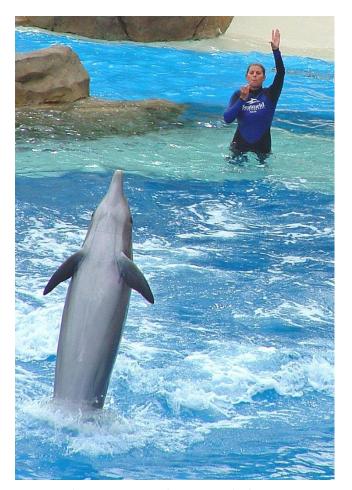**How dependent is language learning on the input?**

Image by Ross Hawkes on Flickr.

Creativity in language

- Who deaded my kitty cat?
- Don't giggle me!
- I undarked it.
- Could you write a small downercase E?
- I'd like to go laundrying.
- Remy: I want MAMA to put it in the fridge.
  Mama: Okay, fine.
  Remy: Mama's nice. Mama's a nice put-er in fridge.
- It is not a hanger. It is missing parts so it can't be a hanger. It is just a HANG.
- There's a wack in there!
- Oh, you mean there's wax?
- No, just one wack!

Creativity in language

Remy's full long name is Remington Malakai Scott. What is…

Uncle Idan's full long name?

- Uncley Uncley Uncley Don Scott

Mama's full long name?

- Mamaton Malakai Scott

Uncle Max's full long name?

- Maxican Maxican Scott

Kids say (meaningful) things they've never heard before

- When Kenzo gets here, I'm gonna go all the way to the door and I'm gonna ask Kenzo's papa why did I break the jellyfish necklace?
- It's important Grandma only has one head... not two, not three.
- The mama rabbit wants to hop while the baby rabbit holds on tightly to the mama rabbit's back!
- If you picked me up I wouldn't get squashed by you.
- I don't know why anybody loves me.
- I think the bee kissed me with his toes. He wasn't gentle enough with me.
- Going round and round in a circle... chasing the OTHER steam! (1;11)
- I'm working hard to get up the Mama. (1;10)

Can imitation and reinforcement explain language learning?

Behaviorism (Skinner):

– Language is a behavior

– Language learning is like any other learning… behavior shaping

What are some challenges for this approach?

– Creativity and error

– Absence of (use of) appropriate corrections

– Enrichment of the input (deaf children of hearing parents, creoles)

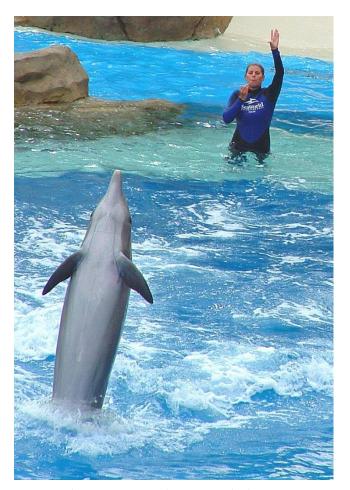**How dependent is language learning on the input?**

Image by Ross Hawkes on Flickr.

Adults generally correct *meaning*

- I need to save him. He needs to be saven!
- Mama smells like play-dough. Mama doesn't LOOK like play-dough. Yellow is play-dough.
- Are all liquids can be poured?
- [Holding up Meefo] Which of one animals would you like?
- Uncle Julian's water bottle has some eyes on HIS water bottle...
- What do I see blue for spraying out of?

- I think cheese is inside wood chips.
- I didn't like my clementine because it was turning into an orange.
- I wanted Mama because I screamed a little bit.
- Mama will grow up like Grandpa, and then Mama will have a scratchy face!
- I'm a Remy, but I'm not a human. Also I'm a toddler!
- You took a long shower! Sometimes people's computers break.

Adults generally correct *meaning*

Child: Mamma isn't a boy, he a girl.
Adult: That's right.
Child: Draw a boot paper.
Adult: That's right, draw a boot on paper.
Child: Her curl her hair.
Adult: Mm-hmm.

Child: There's the animal farmhouse.
Adult: No, that's a lighthouse.

Source: Brown, Roger and Camille Hanlon. "Derivational complexity and the order of acquisition in child speech." In Cognition and the Development of Language. Edited by John R. Hayes. Wiley, 1970, pp. 11-53.

Children's resilience against correction

- Those not Mama feet. Those Mama FOOTS. That pretty silly! (2;0)

- Remy (2;7): Do gooses have feet?
- Me: Uh, yes, geese have feet.
- Remy: But do gooses?

Children's resilience against correction

Adult: He's going out.

Child: He go out.

Adult: Adam, say what I say: Where can I put them?

Child: Where I can put them?

Source: Brown, Roger and Camille Hanlon. "Derivational complexity and the order of acquisition in child speech." In *Cognition and the Development of Language.* Edited by John R. Hayes. Wiley, 1970, pp. 11-53.

Children's resilience against correction

Child: Want other spoon, Daddy.

Father: You mean you want THE OTHER SPOON?

Child: Yes, I want other one spoon please Daddy.

Father: Can you say "the other spoon"?

Child: Other… one… spoon.

Father: Say… "other".

Child: Other.

Father: "Spoon."

Child: Spoon.

Father: "Other… spoon."

Child: Other… spoon. Now give me other one spoon?

Source: Braine, Martin. "The Acquisition of Language in Infant and Child." In The Learning of Language. Edited by C. E. Reed. Appleton-Century-Crofts, 1971, pp. 7-95.

Children's resilience against correction

- Child: A teatop, teatop, teatop, teatop
- Mother: It's a teapot
- Child: Teatop
- Mother: Teapot, not teatop. Teapot
- Child: Sugar

Source: Howe, Christine J. "The Case for Innate Knowledge." In *Language Learning: A Special Case for Developmental Psychology?* Psychology Press, 1993, pp. 19.

…despite usual sensitivity to correction!

Remy: Why not is Mama happy about that? (2;11)

→ If reinforcement is what's going on, how come it (a) generally doesn't happen and (b) doesn't seem to work?

Can imitation and reinforcement explain language learning?

Behaviorism (Skinner):

– Language is a behavior

– Language learning is like any other learning... behavior shaping

What are some challenges for this approach?

– Creativity and error

– Absence of (use of) appropriate corrections

– Enrichment of the input (deaf children of hearing parents, creoles)

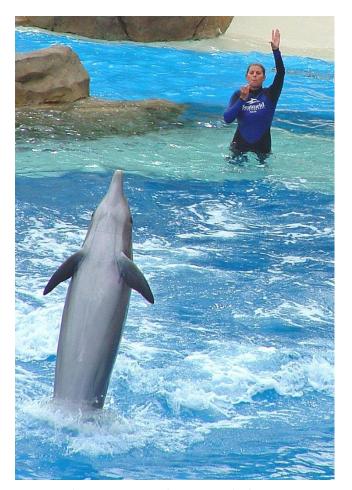**How dependent is language learning on the input?**

Image by Ross Hawkes on Flickr.

The Forbidden Experiment

- Consider situations in which children lack access to some critical piece of information for language learning...

Herodotus (484-425 BCE)

"Now the Egyptians, before the reign of their king Psammetichus, believed themselves to be the most ancient of mankind. Since Psammetichus, however, made an attempt to discover who were actually the primitive race, they have been of opinion that while they surpass all other nations, the Phrygians surpass them in antiquity. This king, finding it impossible to make out by dint of inquiry what men were the most ancient, contrived the following method of discovery:- He took two children of the common sort, and gave them over to a herdsman to bring up at his folds, strictly charging him to let no one utter a word in their presence…

Herodotus (484-425 BCE)

"His object herein was to know, after the indistinct babblings of infancy were over, what word they would first articulate. It happened as he had anticipated. The herdsman obeyed his orders for two years, and at the end of that time, on his one day opening the door of their room and going in, the children both ran up to him with outstretched arms, and distinctly said "Becos." ...He informed his lord, and by his command brought the children into his presence...."

Herodotus (484-425 BCE)

"Psammetichus then himself heard them say the word, upon which he proceeded to make inquiry what people there was who called anything "becos," and hereupon he learnt that "becos" was the Phrygian name for bread. In consideration of this circumstance the Egyptians yielded their claims, and admitted the greater antiquity of the Phrygians."

The Forbidden Experiment

- Consider situations in which children lack access to some critical piece of information for language learning…

Language and the deaf child

- Many deaf children have hearing, non-signing parents
- Oralist tradition
- Goldin-Meadow & colleagues studied 10 children in this situation from ages 1-4
  - Single manual gestures around 12 months
  - Two- and three-word sign sequences at 2 years
    - With syntactic organization!
  - No development of functional morphemes, tense, case, etc.

Language and the blind child

- Lack of access to word *reference*
- How could blind child learn the meanings of *mountain, bird, cloud?*
- Language acquisition by blind children is unexceptional: same timeline, same character
- Even for the verbs *look* and *see*
  - "touch the table but don't look at it"
    - Landau & Gleitman (1985)

The Forbidden Experiment

- "Isabelle" – removed from abusive home at age 6; age-appropriate language by 7
- Helen Keller – blind and deaf from age 2, started learning sign language at 7
- "Genie" – removed at age 13
  - Acquired some language: "Another house have dog"; "No more take wax"
  - But no progression past telegraphic stage:
  - "Where is may I have a penny?"
  - "I where is graham cracker on top shelf?"
- "Chelsea" – rediagnosed at age 31!
  - "Breakfast eating girl"; "Banana the eat"

Children Inventing Language

- Late ASL exposure affects acquisition (Newport 1990)
  - 50-year-olds who started learning from birth, age 4-6 or after age 12
  - Late learners are inconsistent with syntax
- But their children enrich the language!
  - "Simon" - ASL input came from 2 late-learners
  - Simon's own language developed structures *not in the input*

Pidgins and Creoles

- Pidgin – a lingua franca created for communication
  - Shares some features with early child speech
  - Single clauses, few if any function words
- The children of pidgin speakers create creoles
  - Multiclausal sentences
  - Grammaticization
  - Function morphemes

Nicaraguan Sign Language

- Schools opened in the 1970s & 1980s bring together populations of deaf children for the first time
- Transition from a shared pidgin to a much more complex sign language
- Syntactic complexity emerges
- Interesting cognitive effects of learning only the simplified language!

Can training account for language learning?

No: Evidence for the child's role...

- Generalizations that couldn't have been learned from direct experience
- Success across a wide range of inputs (some very impoverished)

Children go beyond the input, but not all on their own.
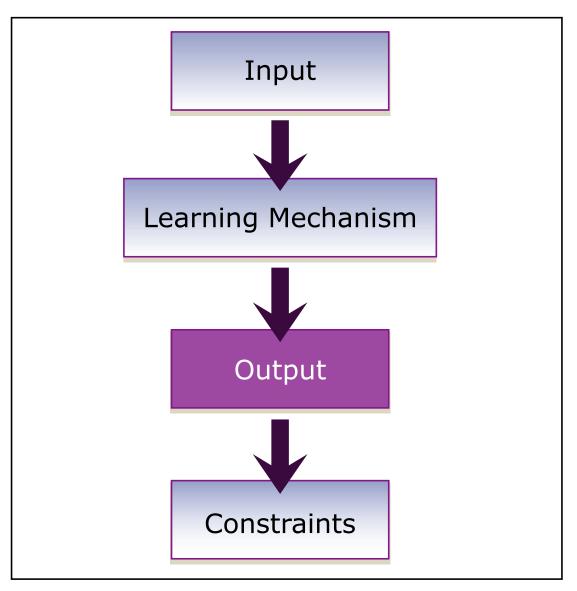

Image by Ross Hawkes on Flickr.

Image by MIT OpenCourseWare.

Output: what do kids figure out?

- Output ≠ production!
- Even very early on, have abstract notions of syntactic categories
- …not just the ability to interpret sentences based on current vocabulary, but expectations about new words!
- Preferences for grammatical sentences by ~16 months, although these are driven by function words
- Syntactic priming (3yos):

  Hearing "Give the lion the ball!" makes it easier to understand "Show the horse the book!" and harder to understand "Show the horn to the dog" (Thothathiri & Snedecker, 2008)

Some early abstract knowledge

Table removed due to copyright restrictions.
Table 1. Soderstrom, Melanie, Katherine S. White, et al. "Receptive Grammatical Knowledge of Familiar Content Words and Inflection in 16-Month-Olds." *Infancy* 12, no. 1 (2007): 1–29.

Gertner, Fisher, & Eisengart (2006)

Figure removed due to copyright restrictions. Figure 1 from Gertner, Y., Fisher, C., et al. "Learning Words and Rules: Abstract Knowledge of Word Order in Early Sentence Comprehension." *Psychological Science* 17, no. 8 (2006): 684-91. doi:10.1111/j.1467-9280.2006.01767.x.

"The duck is gorping the bunny!"

21-month-olds identify abstract participant roles (agent/patient)…

and use word order to interpret a novel transitive verb

…Even before they know what a word MEANS!

28-month-olds succeed!

Figure removed due to copyright restrictions. Figure 1 from Yuan, S., Fisher, C. ""Really? She Blicked the Baby?": Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening." *Psychological Science* 20, no. 5 (2009): 619-26.

Yuan & Fisher 2009

Not just transitive/intransitive…

Figure removed due to copyright restrictions. Figure 2 from Scott, R. M., Fisher, C. "Two-Year-Olds Use Distributional Cues to Interpret Transitivity-Alternating Verbs." *Language and Cognitive Processes* 24, no. 6 (2009): 777-803.
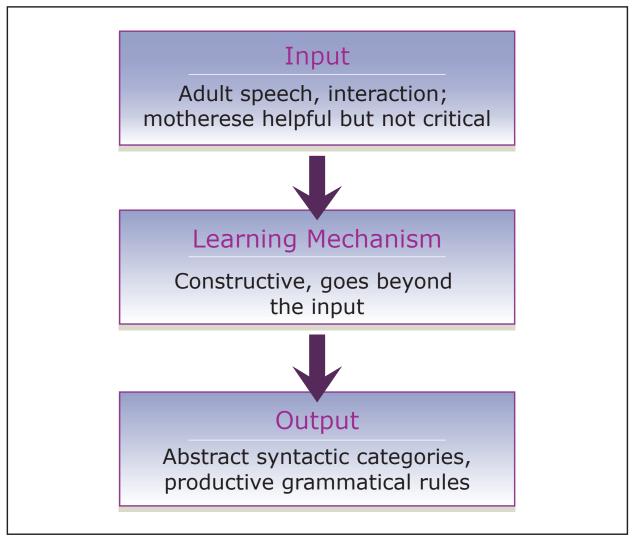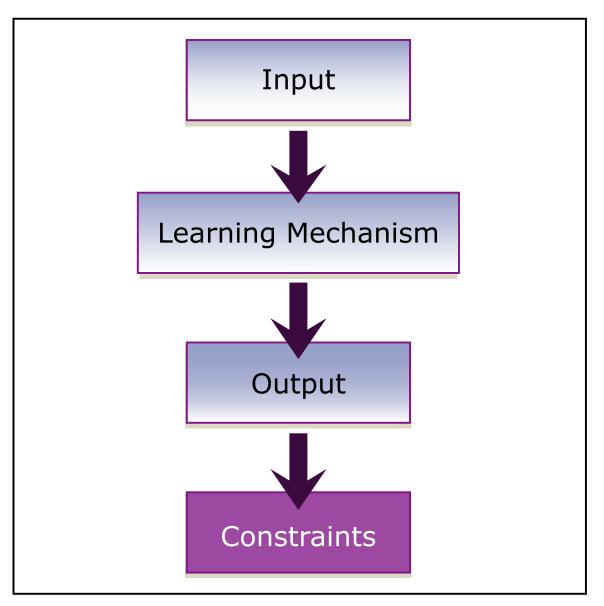
Scott & Fisher 2009

Last time….

Image by MIT OpenCourseWare.

Image by MIT OpenCourseWare.

Clearly children learn language from their parents…

"There is a massive correlation between being born in England and coming to speak English and being born in France and speaking French."

Source: Gleitman, Lila R. and Elissa L. Newport. (1995). "The invention of language by children: Environmental and biological influences on the acquisition of language." In Language: An Invitation to Cognitive Science . Edited by Lila R. Gleitman, and Mark Liberman. MIT Press, 1995.

…but this learning MUST be constrained.

The induction problem

- Children need to (and as we've seen, they do!) generalize from the input.

- But which generalizations? An analog of the "gavagai" problem:

  My cat **is** fuzzy.

  Is my cat fuzzy?

The induction problem

- Children need to (and as we've seen, they do!) generalize from the input.

- But which generalizations? An analog of the "gavagai" problem:

> My cat <u>is</u> fuzzy.

> <u>Is</u> my cat fuzzy?

- Take some more complex sentences…

> Buttons <u>is</u> a cat who <u>is</u> fuzzy.

> Buttons ,who <u>is</u> a cat, <u>is</u> fuzzy.

	Move first "is"	**Move second "is"**	**Move "is" from main clause**
Buttons is a cat who is fuzzy.	Is Buttons a cat who is fuzzy?	*Is Buttons is a cat who fuzzy?	Is Buttons a cat who is fuzzy?
Buttons , who is a cat, is fuzzy.	*Is Buttons, who a cat, is fuzzy?	Is Buttons, who is a cat, fuzzy?	Is Buttons, who is a cat, fuzzy?

Language is structure-dependent

- The only rule that works for both sentences refers to the <u>structure</u> of the sentence, not the order of words
  - Requires a representation of something like 'clause' and 'main clause'
  - Input for this kind of question is rare at best
- And all rules of language operate in this way
- Where does the structure bias come from?

INNATE

EARLY-ABSTRACTING

- **Universal grammar (Chomsky) and "Poverty of the stimulus" argument**
- **Bootstrapping approaches for categorizing words...**

- **Bayesian inference approaches to grammar**

DOMAIN-GENERAL ←———————————→ DOMAIN-SPECIFIC

LATE-ABSTRACTING

- **Verb island hypothesis, item-based learning (Tomasello)**
- **Distributional approaches**

LEARNED

Chomsky: Poverty of the stimulus

1. There are patterns in language that can't be learned just from positive evidence.
2. Children only GET positive evidence.
3. But they all learn these patterns.

…Hence, they must be using additional innate knowledge about language.

Universal grammar: an innate, language-specific set of cognitive structures that limits the possible languages a child can learn

Chomsky: Poverty of the stimulus

1. There are patterns in language that can't be learned just from positive evidence.

   – But some successes of statistical approaches

2. Children only GET positive evidence

   – Remy: Was the one [which] near my house was blinking?

   – Me: Probably, but you should say "Was the one [which was] near my house blinking?"

   – Disagreement about WHAT would be count as negative evidence and what would be "enough"

   – Failures of expectation and implicit negative evidence

3. But they all learn these patterns.

   – Children don't appear to consider the linear hypothesis… but we may not all learn exactly the same grammar, and maybe not right away.
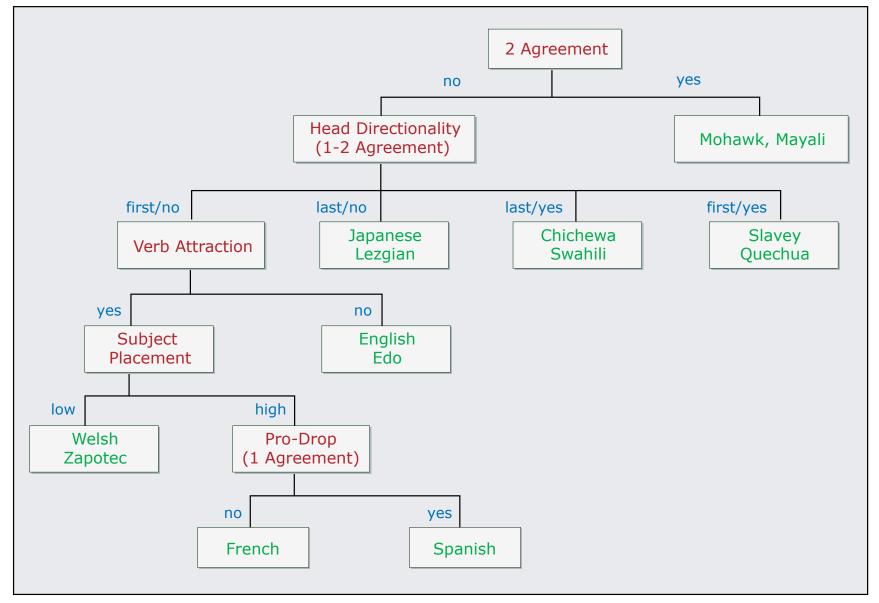
Principles and Parameters

2 Agreement
— no → Head Directionality (1-2 Agreement)
— yes → Mohawk, Mayali

Head Directionality (1-2 Agreement)
— first/no → Verb Attraction
— last/no → Japanese, Lezgian
— last/yes → Chichewa, Swahili
— first/yes → Slavey, Quechua

Verb Attraction
— yes → Subject Placement
— no → English, Edo

Subject Placement
— low → Welsh, Zapotec
— high → Pro-Drop (1 Agreement)

Pro-Drop (1 Agreement)
— no → French
— yes → Spanish

So what drives language acquisition?

Universal grammar (the principles) are **genetically determined.**

- Experience/exposure
  - Triggers correct settings of the parameters
- Maturation
  - …A-chain movement is like permanent teeth coming in – innate but late
- Evidence interpreted in adult-like terms!

How do children avoid overgeneralization without negative feedback?

But how do children categorize words?

Even if we innately expect language to have rules based on abstract parts of speech, we have to figure out which words are nouns, verbs, prepositions, etc.

$$S \rightarrow VP \ NP_{SUBJ}$$
$$NP \rightarrow N \ (det)$$
$$VP \rightarrow NP_{OBJ} \ V$$

the: N
boy: V
threw: N
rocks: det

(Pinker 1984)

But how do children categorize words?

Some proposals for linking universal grammar to input…

- Semantic bootstrapping/innate linking rules (Pinker, Grimshaw, Macnamara):

  - Make use of connections between **semantic** (e.g. object, action) and **syntactic** categories (e.g. noun, verb)

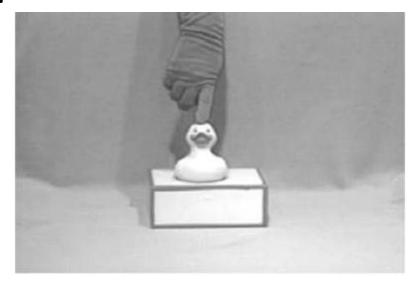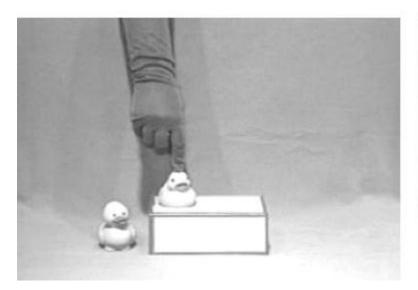| Grammatical element	Semantic inductive basis[12]
SYNTACTIC CATEGORIES	
Noun	Name of person or thing
Verb	Action or change of state
Adjective	Attribute
Preposition	Spatial relation, path, or direction
Sentence	Main proposition
GRAMMATICAL FUNCTIONS[13]	
Subject	Agent of action; cause of causal event; subject of an attribution of location, state, or circumstance; argument with "autonomous reference"
Object and Object2	Patient or theme
Oblique	Source, goal, location, instrument
Complement	Proposition serving as an argument within another proposition

(Pinker 1984)
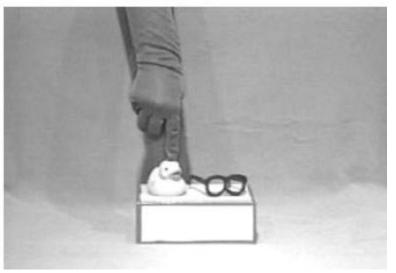
But how do children categorize words?

Some proposals…

- Fisher et al. (2010)—Verb learning by **structure mapping:**
  - Bias to expect the number of NOUNS in a sentence to equal the number of PARTICIPANT ROLES in its meaning
  - Not unique to verbs (learning transitive/intransitive): any predicate
    - I'm happy for you
    - It's under the table
  - Two-year-olds also use the number of nouns to learn new prepositions!
  - Appears to be **unlearned:** present in homesign, children speaking languages that have a much less reliable correspondence
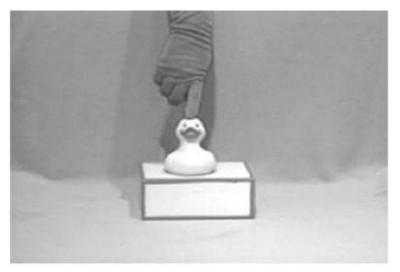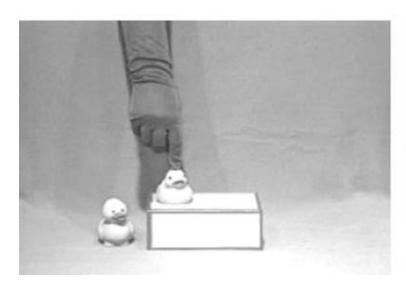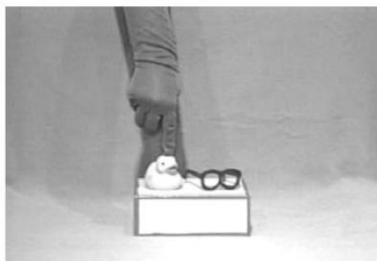
This is a corp!

Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.

This is a corp! What else is a corp?

This is acorp my box!

Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.

This is acorp my box! What else is acorp my box?

But how do children categorize words?

Some proposals…

- Fisher et al. (2010)—Verb learning by **structure mapping:**
  - Bias to expect the number of NOUNS in a sentence to equal the number of PARTICIPANT ROLES in its meaning
  - Not unique to verbs (learning transitive/intransitive): any predicate
    - I'm <u>happy</u> for you
    - It's <u>under</u> the table
  - Two-year-olds also use the number of nouns to learn new prepositions!
  - Appears to be **unlearned(!):**
    - present in homesign
    - Present in children speaking languages that have a much less reliable correspondence between # nouns and # participant-roles

Theory space

INNATE

DOMAIN-GENERAL

DOMAIN-SPECIFIC

LEARNED

- **Bayesian inference approaches to grammar**

EARLY-ABSTRACTING
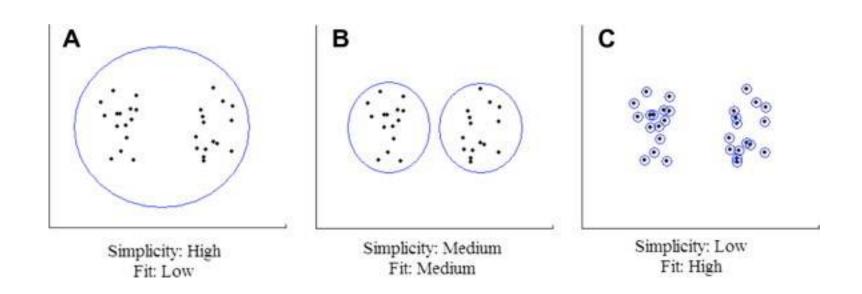
- **Universal grammar (Chomsky) and "Poverty of the stimulus" argument**
- **Bootstrapping approaches for categorizing words...**

LATE-ABSTRACTING

- **Verb island hypothesis, item-based learning (Tomasello)**
- **Distributional approaches**

A domain-general possibility

- Language is an induction problem: it may be strictly "unlearnable," but could a rational learner make good guesses from the input?

- Maybe what we have is **"poverty of the imagination"**
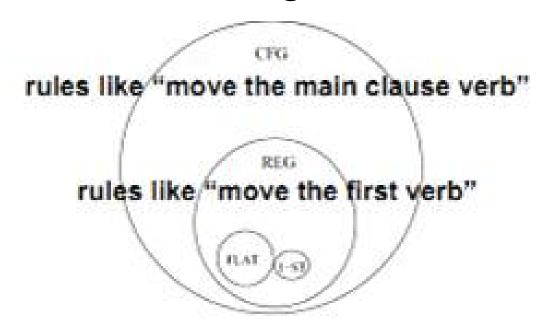
An early attempt: Bayesian inference of grammar

Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.

Perfors, Tenenbaum, & Regier (2006)

- Instead of learning how one rule words (e.g. is-movement)

- Learn how the whole grammar works…

Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.

Perfors, Tenenbaum, & Regier (2006)

Used corpus of sentences spoken by adults to children

- About 2300 unique sentence types

- Broke down into 6 levels based on frequency of sentence type

Three types of hypotheses:

- "Flat" grammar: all sentence types listed

- Regular grammar: rules for adding to the start of a sentence

- Context-free grammar

Table removed due to copyright restrictions. Table 1. Perfors, Amy, Joshua B. Tenenbaum, and Terry Regier. "Poverty of the Stimulus? A Rational Approach." *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (2006): 663-8.

Perfors et al. discussion

Goal: Maximize P(Grammar | Corpus)

Results:

- a context-free grammar becomes the "simplest option" at stage 3
- CFGs also generalize better to additional input

Maybe we don't have enough information to learn specific grammatical rules, but we could learn them as part of a bigger structure.

- The frequency of examples like "Is the dog which is in the corner hungry?" is irrelevant!

Some caveats:

- Computational-level approach: explains **why** the context-free grammar might be chosen, but not **how.**
- Agnostic about innateness

Theory space

INNATE

DOMAIN-GENERAL

DOMAIN-SPECIFIC

LEARNED

- **Bayesian inference approaches to grammar**

EARLY-ABSTRACTING

- **Universal grammar (Chomsky) and "Poverty of the stimulus" argument**
- **Bootstrapping approaches for categorizing words…**

LATE-ABSTRACTING

- **Verb island hypothesis, item-based learning (Tomasello)**
- **Distributional approaches**

How could we learn grammar without having syntactic categories first?

- Tomasello: Language learning depends on general cognitive mechanisms, and social abilities

- Adult language is too complex to be learned, right away – so children must be learning something else

- **"Poverty of the child grammar"**

Back to the two-word stage…

- Remember the two-word stage: children don't make word order errors
  - Two-word utterances as 'samples' from longer sentences
- Do young children organize their language in terms of categories of nouns, verbs, abstract syntactic relationships?

Maybe not…

- A diary study – Tomasello (1990) recorded his daughter's early language (<2 yo): her language was 'grammatical' but item-based
  - *Draw __for __; Draw __ on __; __ Draw on __*
  - *Cut __*
  - No transfer of structure: categories are "draw-er," "thing drawn on," "thing drawn with" rather than subject, object, instrument.
- Pine & Lieven (1997): 2-3 year olds start using "a" and "the" with distinct sets of nouns
- Young children in production studies tend to resist generalization (*Toy blicked -> I blicked the toy*)
- Remy says:
  - <u>I suspect </u>I wanted to wake up… <u>I bet </u>I want to look at some boats.
  - <u>It seems to me </u>I would like a strawberry.
  - No <u>is not a good word</u>. You don't want to <u>is not a good word</u>.

The Verb Island Hypothesis

- Early in development children use verb-specific representations to form sentences (kicker KICKS kick-ee)

- Gradually, representations merge and become abstract or general over different verbs (NP V NP)

- Classic overgeneralizations (*Don't giggle me!*) appear after this takes place, age 3 yo and later

Problems with the Verb Island Hypothesis

- Early syntactic awareness
  - Remember Gentner et al paper: even when children can't say these sentences, they still *interpret* them correctly
  - Children have at least broad syntax/semantics expectations by 18 months!
  - Gap between receptive and productive language: sensitivity to exactly the function words they skip
- Resisting production might be rational
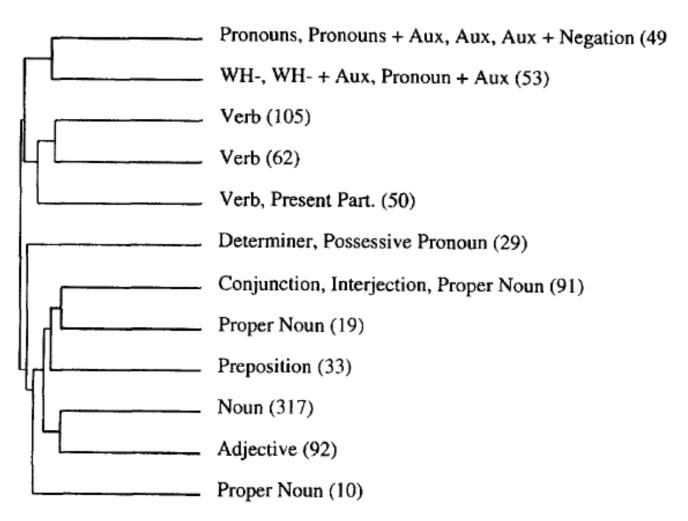- But the idea of usage-based grammar is important.

Other late-abstracting approaches

Distributional information: contexts where we find a particular word

- Maratsos & Chalkely, 1980: Classic distributional theory
  - Learning e.g. gender classes from repeated frames
- Redington et al., 1988: clustering based on bigrams is surprisingly informative…

The cow jumped over the moon.

Redington et al. 1998

Pronouns, Pronouns + Aux, Aux, Aux + Negation (49

WH-, WH- + Aux, Pronoun + Aux (53)

Verb (105)

Verb (62)

Verb, Present Part. (50)

Determiner, Possessive Pronoun (29)

Conjunction, Interjection, Proper Noun (91)

Proper Noun (19)

Preposition (33)

Noun (317)

Adjective (92)

Proper Noun (10)

Remy says: statistical information is at least **part** of the story…

- Me: No, I'm taking my shower alone today.
- Remy: It's MY shower-alone-today!
- Me: What type of animal is a spider?
- Remy: A spider is a tiny spider. Down came the rain and washed the spider out!
- Ear… earwax… candlewax… candles… HAPPY BIRTHDAY!
- Take me out to the ball game, take me out to the crowd… Buy me some peanuts and crackerjacks, I don't care if I never get back for it's root, root root for the home team, if they don't win it's a shame… for it's one, two, three, four, five, six, seven…

Conclusions

- Language is complex, richly structured, and LEARNABLE

- Children **construct** language from the input – and sometime surpass it

- Language is an induction problem – language learning must be constrained

  - Domain-general or domain-specific?

  - Initial representations?

  - What hypotheses will the child consider?

9.85 Infant and Early Childhood Cognition
Fall 2012