Psych 156A/ Ling 150: Acquisition of Language II

Lecture 11

Grammatical Categorization II

Announcements

HW3 available (due 5/27/14) — update made to question #2 on 5/8/14. Please download the updated version if you haven't already. (If you downloaded it after 5/8/14, you have the latest version.)

Computational problem

Determine that there are grammatical categories like Noun and Verb that behave similarly with respect to their morphosyntax (the way they go together with other words).

Noun = {penguin, goblin, glitter, cheese}

Morphosyntax: Nouns can take determiners like "the" and "a"

{the penguin, a goblin, the glitter, a king}

Computational problem

Determine that there are grammatical categories like Noun and Verb that behave similarly with respect to their morphosyntax (the way they go together with other words).

Verb = {swim, dance, flutter, smell}

Morphosyntax: Verbs can take adverbs that modify them, like "really"

{really swim, really dance, really flutter, really smell}

Assessing child knowledge

How do we know when children achieve adult-like knowledge?

Charles Yang, 2010

"Language use is the composite of linguistic, cognitive and perceptual factors many of which, in the child's case, are still in development and maturation. It is therefore difficult to draw inferences about the learner's linguistic knowledge from his linguistic behavior."

But
"…child language [can] be interpreted in terms of adult-like grammatical devices…"

Example adult-like grammatical device: Categories like Noun and Verb

Morphosyntax of nouns

Nouns can combine with certain types of words in the input to make larger units (ex: Noun Phrases).

Determiner + Noun ("the kitty")
[NP → Det + N]

Adjective + Noun ("cute penguins")
[NP → Adj + N]

Morphosyntax of nouns

This is thought to involve knowledge of the category *Noun*

- Impact: Rules for combining nouns together with other words to generate utterances involve this symbol (along with other symbols), rather than individual lexical items.

NP → Det Noun **NP → Adj Noun**

rather than

VP → the kitty **VP → cute penguins**

Morphosyntax of verbs

Verbs can also combine with certain types of words in the input to make larger units (ex: Verb Phrases [VP]).

Verb + Object ("hug the kitty")
[VP → Verb + Det + N]

Auxiliary + Verb ("can hug")
[VP → Aux + Verb]

Morphosyntax of verbs

This is thought to involve knowledge of the category *Verb*

– Impact: Rules for combining verbs together with other words to generate utterances involve this symbol (along with other symbols), rather than individual lexical items.

VP → Verb Det N | **VP → Aux Verb Det N**

rather than

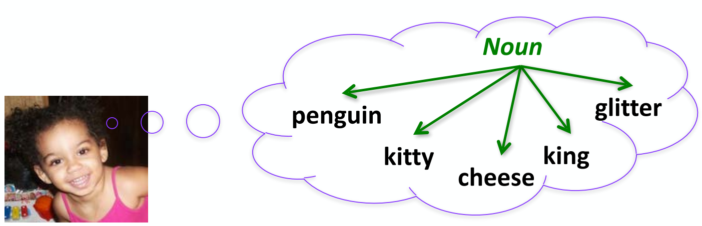
VP → hug the kitty | **VP → can hug the kitty**

Combinatorial power

- Why do we believe this is what (adult) rules look like?
  - Expressive power: The ability to generate novel utterances, composed of recognizable pieces (words). Since the utterances haven't been heard before, they must be generated based on rules whose primitives are more abstract than individual lexical items.
  - This kind of combinatorial diversity is sometimes called productivity.

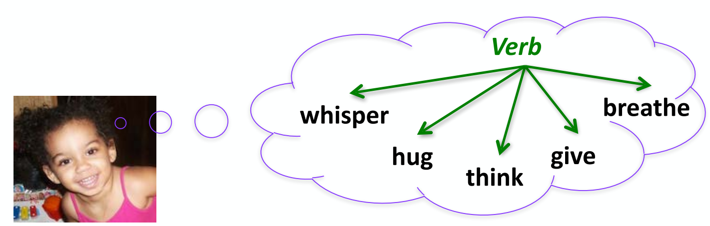
Assessing knowledge of grammatical categories

When do children first develop knowledge of the abstract category of *Noun*?



Assessing knowledge of grammatical categories

When do children first develop knowledge of the abstract category of *Verb*?



Development of grammatical categories

Some studies suggest that grammatical category knowledge may already be in place around the age of two

- Determiners, Nouns: Valian 1986, Valian, Solt, & Stewart 2008
- Auxiliary verbs: Stromswold 1989, Rispoli, Hadley, & Holt 2009, Rissman, Legendre, & Landau 2013
- Verbs: Kowalski & Yang 2012

Development of grammatical categories

Other studies suggest that it may appear significantly later:

- Determiners, Nouns: Pine & Lieven 1997
- Auxiliary verbs: Wilson 2003, Theakston & Lieven 2005, Theakston, Lieven, Pine, & Rowland 2005, Theakston & Lieven 2008, Theakston & Rowland 2009
- Verbs: Tomasello 1992, Tomasello 2006

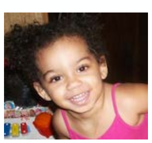
Assessing knowledge of grammatical categories

How can we tell?

One indicator:

Knowledge about how one word (noun or verb) combines with other words is transferred within the category.

…could think…



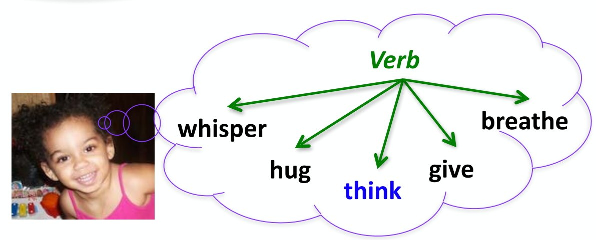
Assessing knowledge of grammatical categories

How can we tell?

One indicator:

Knowledge about how one word (noun or verb) combines with other words is transferred within the category.

…could think…

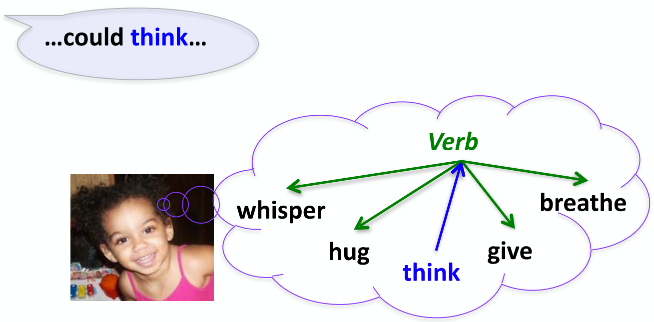


Assessing knowledge of grammatical categories

How can we tell?

One indicator:

Knowledge about how one word (noun or verb) combines with other words is transferred within the category.

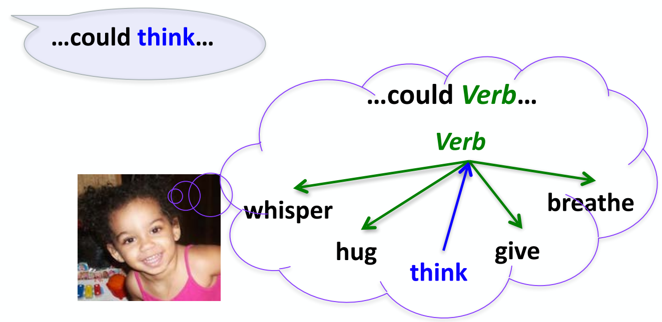


Assessing knowledge of grammatical categories

How can we tell?

One indicator:

Knowledge about how one word (noun or verb) combines with other words is transferred within the category.

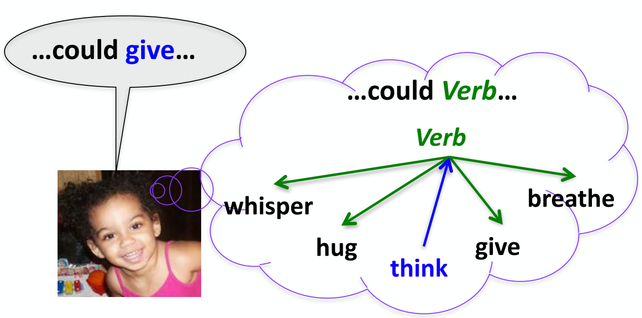


Assessing knowledge of grammatical categories

How can we tell?

One indicator:

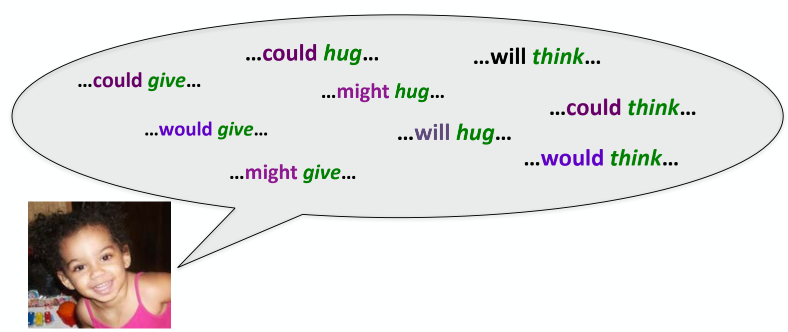
Knowledge about how one word (noun or verb) combines with other words is transferred within the category.



Assessing knowledge of grammatical categories

How can we tell?

This causes the child to combine words of the same category with similar words, so that there's overlap in usage within a category.



Assessing knowledge of grammatical categories

How can we tell?

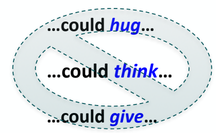
This overlap (which is combinatorial diversity) is something we can quantitatively assess to gauge productivity with respect to categories. In particular, we can look at the lexical substitution for a category (Tomasello 1992, Pine & Lieven 1997, Naigles, Hoff, & Vear 2009, Yang 2010, 2011, 2013).

...could *give*...
...could *hug*...
...will *think*...
...might *hug*...
...could *think*...
...would *give*...
...will *hug*...
...would *think*...
...might *give*...

Assessing knowledge of grammatical categories

Premise: If children's noun or verb usage shows enough combinatorial diversity (productivity), as measured by the lexical substitution (overlap), this suggests they have rules that are based on the more abstract symbols *Noun* and *Verb*, rather than rules that are lexically-based.

...the *Noun*...



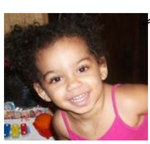
...could *Verb*...

Expected productivity

How much overlap do we expect to see if a child knows the category *Noun* or *Verb*? Let's think this through for *Verb*.

For example, should we expect every verb to combine with every auxiliary?

...could *give*...
...could *hug*...
...will *think*...
...might *hug*...
...could *think*...
...would *give*...
...will *hug*...
...would *think*...
...might *give*...

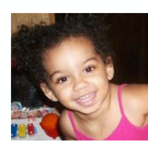


Expected productivity

How much overlap do we expect to see if a child knows the category *Noun* or *Verb*? Let's think this through for *Verb*.

For example, should we expect every verb to combine with every auxiliary?

...could *give*...
...could *hug*...
...will *think*...
...might *hug*...
...could *think*...
...would *give*...
...will *hug*...
...would *think*...
...might *give*...

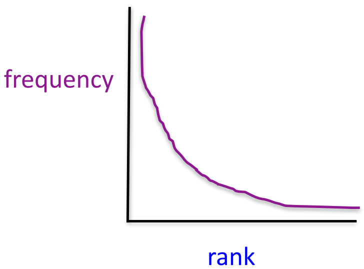


Probably not if these are naturalistic outputs. We don't say everything we know when we speak – we say things to communicate our intended meaning at the time.

Expected productivity

How much overlap do we expect to see if a child knows the category *Noun* or *Verb*? Let's think this through for *Verb*.

In fact, it turns out naturalistic linguistic output shows power law behavior (a Zipfian distribution)...

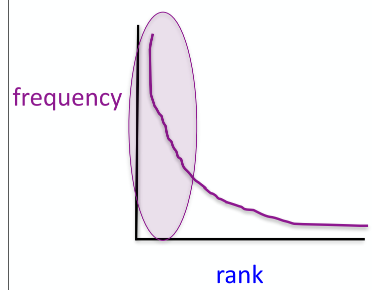


verb	freq	rank
get	101	1
go	100	2
…		
feel	8	58
…		
dream	1	251
…		

Expected productivity

How much overlap do we expect to see if a child knows the category *Noun* or *Verb*? Let's think this through for *Verb*.

In fact, it turns out naturalistic linguistic output shows power law behavior (a Zipfian distribution), where a few things are said very frequently...

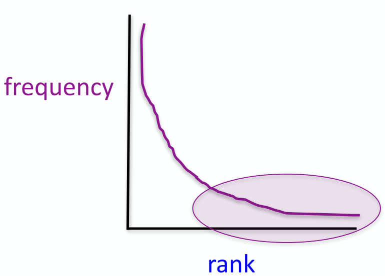


verb	freq	rank
get	101	1
go	100	2
…		
feel	8	58
…		
dream	1	251
…		

Expected productivity

How much overlap do we expect to see if a child knows the category *Noun* or *Verb*? Let's think this through for *Verb*.

In fact, it turns out naturalistic linguistic output shows power law behavior (a Zipfian distribution), where a few things are said very frequently and most things are said very infrequently.

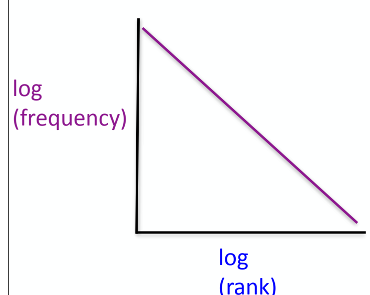


verb	freq	rank
get	101	1
go	100	2
…		
feel	8	58
…		
dream	1	251
…		

Expected productivity

How much overlap do we expect to see if a child knows the category *Noun* or *Verb*? Let's think this through for *Verb*.

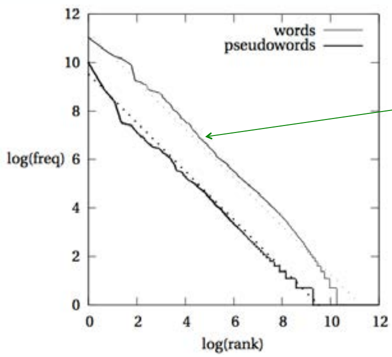
In fact, it turns out naturalistic linguistic output shows power law behavior (a Zipfian distribution), where a few things are said very frequently and most things are said very infrequently.



This shows up as a linear relationship in logarithmic space.

Naturalistic linguistic output

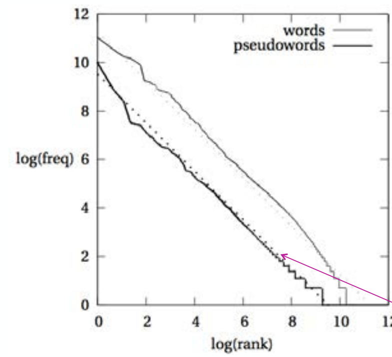
Checking the Zipfian distribution on the Brown corpus
(adult-generated speech and text)



Words are generated with a Zipfian distribution. (Yang 2010)

Naturalistic linguistic output

Checking the Zipfian distribution on the Brown corpus
(adult-generated speech and text)



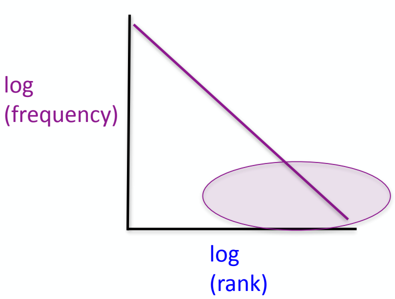
"The lower line is plotted by taking 'words' to be any sequence of letters between *e's* (Chomsky 1958). The two straight dotted lines are linear functions with the slope -1, which illustrate the goodness of the Zipfian fit." (Yang 2010)

Even "unnatural" linguistic output has a Zipfian distribution.

Expected productivity

How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

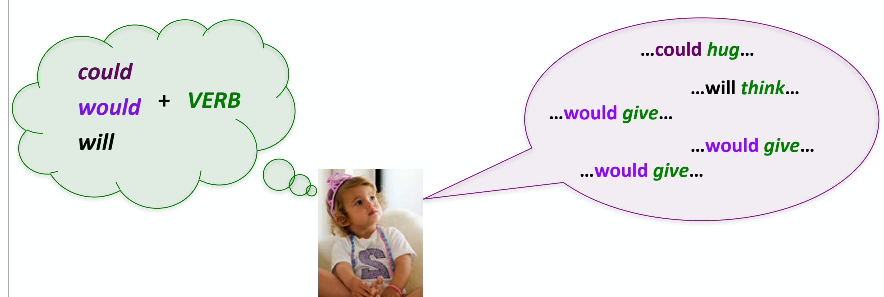
One implication: We can't expect much overlap in combinatorial usage for nouns or verbs that only are used a few times (and certainly not for those that are only used once).



Expected productivity

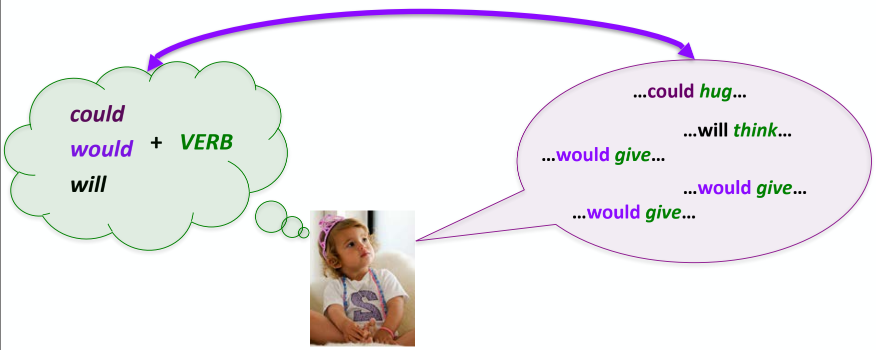
How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

We need to somehow factor in that a child may know that combinatorial usage transfers to other nouns or verbs, but just doesn't choose to say those other nouns or verbs with other words.



Expected productivity

Yang (2010, 2011, 2013): A formal definition of how many lexical items in a category (like *Noun* or *Verb*) are expected to have overlap in a productive system that uses category-based rules.



A formal definition of productivity

Yang 2010, 2011, 2013

Expected overlap: How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

$$E(L,V,S)=\frac{1}{V}\sum_{r=1}^{V}E(r,V,L,S)$$

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

{hug, give, take, read, want, think, …} = ???

A formal definition of productivity

Yang 2010, 2011, 2013

Expected overlap: How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

$$E(L,V,S)=\frac{1}{V}\sum_{r=1}^{V}E(r,V,L,S)$$

The probability that a particular word with rank *r* will have overlap…

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

verb	freq	rank
get	111	1
go	99	2
…		
keep	8	58
…		
use	1	247
…		

A formal definition of productivity

Yang 2010, 2011, 2013

Expected overlap: How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

$$E(L,V,S)=\frac{1}{V}\sum_{r=1}^{V}E(r,V,L,S)$$

The probability that a particular word with rank *r* will have overlap…

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

where this is equivalent to this

$$E(r,V,L,S)=1+(L-1)(1-p_r)^S-\sum_{i=1}^{L}[(l_i p_r+1-p_r)^S]$$

with

$$p_r=\frac{1}{r*H_V},\quad H_V=\sum_{i=1}^{V}\frac{1}{i}$$

A formal definition of productivity

Yang 2010, 2011, 2013

Expected overlap: How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

$$E(L,V,S)=\frac{1}{V}\sum_{r=1}^{V}E(r,V,L,S)$$

…given the total number of vocabulary items V combining with the lexical items from the class of interest…

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

verb (used with auxiliary) total = 58
like
go
…
cook
…
make
…

A formal definition of productivity

Yang 2010, 2011, 2013

Expected overlap: How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

$$E(L,V,S)=\frac{1}{V}\sum_{r=1}^{V}E(r,V,L,S)$$

…and the number of combinatorial vocabulary items from the lexical class of interest *L*…

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

auxiliary total = 23
don't
can't
…
do
…
are
…

A formal definition of productivity

Yang 2010, 2011, 2013

Expected overlap: How much overlap do we expect to see if a child knows the category *Noun* or *Verb*?

$$E(L,V,S)=\frac{1}{V}\sum_{r=1}^{V}E(r,V,L,S)$$

…and the number of word usages observed with the combinatorial vocabulary items of interest *S*.

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

auxiliary + verb usage total = 222
don't close the door
can't reach it
…
do not come
…
are you cooking that?
…

Comparison with child data

Yang 2010, 2011, 2013

Once we know this expected overlap, we can look at the overlap we actually observe in the empirical data and see if they match.

$$E(L,V,S)=\frac{1}{V}\sum_{r=1}^{V}E(r,V,L,S)$$

Calculating observed overlap:
If word is used with more than one lexical item within the lexical class (ex: auxiliaries), overlap for that word = 1.

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

…could give…
…would give…
…could give…
…could give…

overlap for *give* = 1

Comparison with child data

Yang 2010, 2011, 2013

Once we know this expected overlap, we can look at the overlap we actually observe in the empirical data and see if they match.

$$E(L,V,S) = \frac{1}{V}\sum_{r=1}^{V} E(r,V,L,S)$$

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

Calculating observed overlap:
If word is used with more than one lexical item within the lexical class (ex: auxiliaries), overlap for that word = 1.
Otherwise, overlap for word = 0.

…could give…
…could give…
…could give…
…could give…

overlap for *give* = 0

Comparison with child data

Yang 2010, 2011, 2013

Once we know this expected overlap, we can look at the overlap we actually observe in the empirical data and see if they match.

$$E(L,V,S) = \frac{1}{V}\sum_{r=1}^{V} E(r,V,L,S)$$

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

Calculating observed overlap:
If word is used with more than one lexical item within the lexical class (ex: auxiliaries), overlap for that word = 1.
Otherwise, overlap for word = 0.

$$\text{Observed overlap} = \frac{\text{total overlap from all words}}{\text{total number of words}}$$

Comparison with child data

Yang 2010, 2011, 2013

Once we know this expected overlap, we can look at the overlap we actually observe in the empirical data and see if they match.

$$E(L,V,S) = \frac{1}{V}\sum_{r=1}^{V} E(r,V,L,S)$$

The proportion of observed lexical items in the vocabulary that are expected to have some overlap: 0.0 - 1.0

Calculating observed overlap:
If word is used with more than one lexical item within the lexical class (ex: auxiliaries), overlap for that word = 1.
Otherwise, overlap for word = 0.

$$\text{Observed overlap} = \frac{\text{total overlap from all words}}{\text{total number of words}}$$

Example:
Suppose we have 4 words (give, eat, tell, hug)
Suppose the overlap scores are (1,0,1,1).
Observed overlap = (1+0+1+1)/4 = ¾ = 0.75

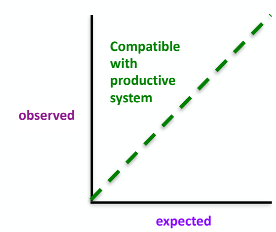
Comparison with child data

Yang 2010, 2011, 2013

Once we know this expected overlap, we can look at the overlap we actually observe in the empirical data and see if they match.

If observed overlap = expected overlap, this child's output is compatible with knowing the grammatical category (*Noun* or *Verb*).

Individual words are used as if they were part of the same category.



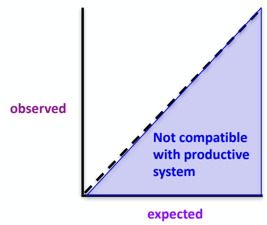
Comparison with child data

Yang 2010, 2011, 2013

Once we know this expected overlap, we can look at the overlap we actually observe in the empirical data and see if they match.

If observed overlap < expected overlap, this child's output is not compatible with knowing the grammatical category (*Noun* or *Verb*).

Individual words are used as if they were not part of the same category.



When do children know *Noun*?

Yang 2010, 2011

One aspect of *Noun* knowledge =

Nouns can be used with determiners *a* and *the*

NP → Det Noun

Data: Child-produced utterances from the six American English corpora of the CHILDES database (age range 1;1 to 5;1).

First 100, 300, and 500 productions from all children to capture earliest stage of language production which should (presumably) be the least productive.

When do children know *Noun*?

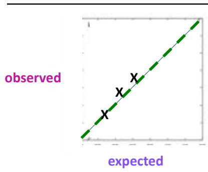
Yang 2010, 2011

One aspect of *Noun* knowledge =

Nouns can be used with determiners *a* and *the*

NP → Det Noun

	Observed Overlap	Expected Overlap
First 100 utterances	21.8	22.4
First 300 utterances	29.1	29.1
First 500 utterances	34.2	33.9



Observed overlap = Expected overlap

When do children know *Noun*?

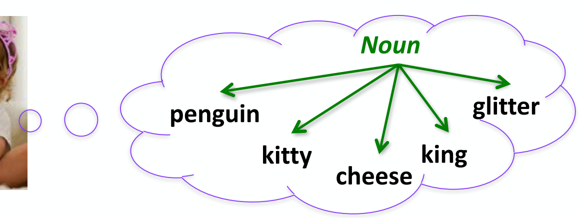
Yang 2010, 2011

One aspect of *Noun* knowledge =

Nouns can be used with determiners *a* and *the*

NP → Det Noun

Implication: Very early child language consistent with knowing *Noun* (at least with respect to combining these Determiners and Nouns)



When do children know *Verb*?

Pearl & Braunwald in prep.

Aspects of *Verb* knowledge =

Verbs can be used with subjects

Sentence → Subject Verb

"I hug"

When do children know *Verb*?

Pearl & Braunwald in prep.

Aspects of *Verb* knowledge =

Verbs can be used with objects

VP → Verb Object

"hug birdie"

When do children know *Verb*?

Pearl & Braunwald in prep.

Aspects of *Verb* knowledge =

Verbs can be used with non-object phrases, like indirect objects, predicate adjectives, prepositional phrases, locatives

VP → Verb Non-Object

"give birdie a hug"

"be happy"

"give a hug to birdie"

"go home"

When do children know *Verb*?

Pearl & Braunwald in prep.

Aspects of *Verb* knowledge =

Verbs can be used with negations

VP → Negation Verb

"don't cry"

When do children know *Verb*?

Pearl & Braunwald in prep.

Aspects of *Verb* knowledge =

Verbs can be used with auxiliary verbs

VP → Auxiliary Verb

"will hug"

When do children know *Verb*?

Pearl & Braunwald in prep.

Aspects of *Verb* knowledge =

Verbs can be used with wh-words when asking questions

Question → Wh-word… Verb

"what do you want?"

When do children know *Verb*?

Pearl & Braunwald in prep.

Aspects of *Verb* knowledge =

Verbs can be used with embedded clauses

VP → Verb Embedded-Clause

"look what I did"

When do children know *Verb*?

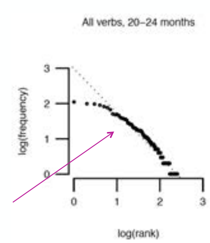
Pearl & Braunwald in prep.

Data:

Longitudinal data from a typically developing child (Laura) in naturalistic contexts

Focusing on 20 to 24 months of age.

Total verb vocabulary items (types): 254
Total verb usages (tokens): 2157



Laura's verb usage shows a Zipfian distribution

When do children know *Verb*?

Pearl & Braunwald in prep.

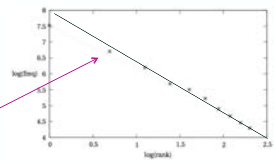
Data:

Longitudinal data from a typically developing child (Laura) in naturalistic contexts

Focusing on 20 to 24 months of age.

Total verb vocabulary items (types): 254
Total verb usages (tokens): 2157

This is similar to estimates of children's transitive verb usage that are based on larger corpora (1.1 million words from CHILDES) done by Yang (2010).

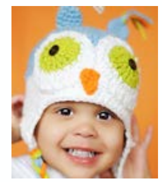


When do children know *Verb*?

Pearl & Braunwald in prep.

Data:

Laura's verb usage appears to be typical, compared against a group of 93 children between 20 and 24 months from the American English CHILDES database



When do children know *Verb*?

Pearl & Braunwald in prep.

20 to 24 months CHILDES

Total verb types: 322

Total verb tokens: 10432

20 most frequent verbs

be, *put*, get, go, want, see, have, open, take, sit, do, *close*, look, eat, *fix*, come, *read*, make, write, *fall*

Laura uses 15 of their top 20.

Laura's 20 most frequent verbs

get, go, have, is, see, do, want, *like*, come, open, look it, sit, take, made, write, eat, *hurt*, *carry*, make, *hear*

The other children use 16 of Laura's top 20.

When do children know *Verb*?

Pearl & Braunwald in prep.

Laura's word order also appears adult-like, with respect to Subjects, Verbs, and Objects.

Total verb uses with subjects, objects, or both: 1688

Verbs with Subjects only

SV	VS
499	9
98.2%	

Verbs with Objects only

VO	OV
477	6
98.8%	

Verbs with Subjects and Objects

SVO	SOV	VSO	VOS	OSV	OVS
689	2	1	2	1	2
98.9%					

When do children know *Verb*?

Pearl & Braunwald in prep.

If observed overlap = expected overlap, productions are compatible with a productive system at 20-24 months of age.

If observed overlap < expected overlap, they are not.

Observed overlap
Expected overlap
Obs* – *Exp

When do children know *Verb*?

Pearl & Braunwald in prep.

Observed – Expected = 0.0 → Could be productive, could know *Verb* at 20-24 months of age.

Observed – Expected < 0.0 → Not productive, does not know *Verb* at 20-24 months of age.

Observed overlap
Expected overlap
Obs* – *Exp

When do children know *Verb*?

Pearl & Braunwald in prep.

Observed – Expected = 0.0 → Could be productive, could know *Verb* at 20-24 months of age.

Observed – Expected < 0.0 → Not productive, does not know *Verb* at 20-24 months of age.

	subj	obj	non-obj	neg	aux	wh	emb cla
Observed overlap	0.50	0.50	0.54	0.36	0.48	0.13	0.60
Expected overlap	***0.72***	***0.76***	***0.64***	***0.45***	***0.52***	***0.56***	***0.72***
Obs* – *Exp	***-0.22***	***-0.26***	***-0.10***	***-0.09***	***-0.04***	***-0.43***	***-0.12***

When do children know *Verb*?

Pearl & Braunwald in prep.

Comparing observed to expected overlap for 7 classes, it doesn't seem like verbs combine freely with words from different lexical classes for this child at 20-24 months.

Lexical substitution knowledge is not transferring across different verbs.

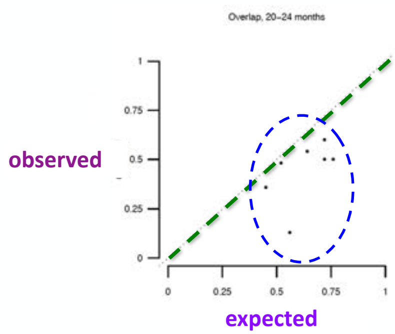
	subj	obj	non-obj	neg	aux	wh	emb cla
Observed overlap	0.50	0.50	0.54	0.36	0.48	0.13	0.60
Expected overlap	***0.72***	***0.76***	***0.64***	***0.45***	***0.52***	***0.56***	***0.72***
Obs* – *Exp	***-0.22***	***-0.26***	***-0.10***	***-0.09***	***-0.04***	***-0.43***	***-0.12***

When do children know *Verb*?

Pearl & Braunwald in prep.

Comparing observed to expected overlap for 7 classes, it doesn't seem like verbs combine freely with words from different lexical classes for this child at 20-24 months.

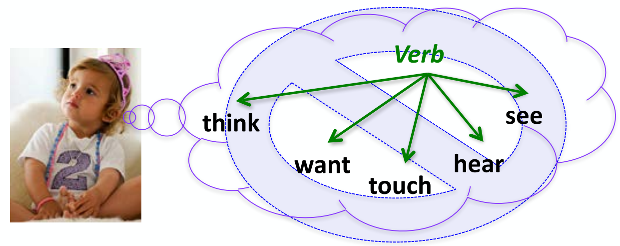
Lexical substitution knowledge is not transferring across different verbs.



When do children know *Verb*?

Pearl & Braunwald in prep.

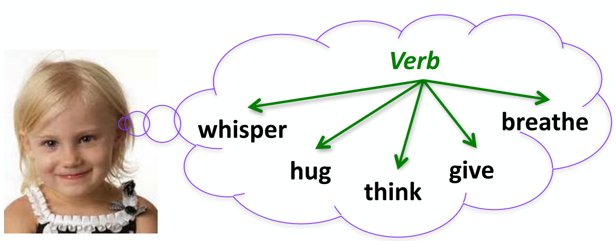
Despite the number of verbs this child is producing (254 verb vocabulary items) and the regularity of word order (>98% SVO, SV, or VO), development of the grammatical category knowledge of *Verb* does not occur until after 24 months for this child.



When do children know *Verb*?

Pearl & Braunwald in prep.

Maybe later...



Recap: Grammatical categorization

Productivity, as measured by the lexical overlap of words within a grammatical category, is one way to assess whether children seem to have knowledge of a particular grammatical category.

Natural language use seems to have a Zipfian distribution, where many combinations are rarely (or never) heard. This can make it hard to learn (more on this next time), but it can also make it hard to figure out what knowledge children have.

Yang (2010, 2011, 2013) offered a formal metric for figuring out exactly how much overlap words should have, given that language use has a Zipfian distribution.

Based on this metric, it seems like children may attain knowledge of *Noun* earlier than they attain knowledge of *Verb*.

Questions?

You should be able to do all the review questions for grammatical categorization and up through question 3 on HW3.