Explain me this: Learning what not to say

Adele Goldberg

adele@princeton.edu

- Learners want to understand messages given forms (comprehension).

and

- They need to choose forms to express the messages they want to convey (production).

Need to categorize

Need to categorize

form ~ function pairings: *constructions*

Constructions at varying levels of complexity and abstraction	
Word	eg., *Bosnia, break, notwithstanding*
Word (partially filled)	eg., *pre-N*, V-*ing*

Constructions at varying levels of complexity and abstraction	
Word	eg., *Bosnia, break, notwithstanding*
Word (partially filled)	eg., *pre-N*, V-*ing*
Idiom (filled)	eg., *Got milk?, give the Devil his due*
Idiom (partially filled)	eg., *Jog* <someone's> *memory,* *send* < someone> *to the cleaners*

Constructions at varying levels of complexity and abstraction	
Word	eg., *kast, breake, notwithstanding*
Word (partially filled)	eg., *pre-N*, V-*ing*
Idiom (filled)	eg., *Got milk?, give the Devil his due*
Idiom (partially filled)	eg., *Jog* <someone's> *memory, send* < someone> *to the cleaners*
Unusual constructions (partially filled)	The Xer the Yer (eg., *The more you think about it, the less you understand*) Sarcasm construction (*e.g., What am I, f**king Jiminy Cricket?*)
(unfilled) Ditransitive construction: Subj V Obj_1 Obj_2	eg., *He gave her a fish taco; He baked her a muffin.*
Passive: Subj aux VPpp (PPby)	eg., *The armadillo was hit by a car.*

Our quarry: how do speakers learn when they can generalize a construction and when they cannot?

How do we learn what *not* to say?

Certain things, one just shouldn't say

I haven't seen you for a long time–are you pregnant?
I only care about my grade in this course.
You look so old!
Who circumcised YOU?!

Positive evidence: types of expressions that are witnessed.

Assume you can say these sentences:

Scrape-*nu* the *vip* the *hap*.
Load-*nu* the *yib* the *vork*.
Flip-*nu* the *loof* the *rolm*.

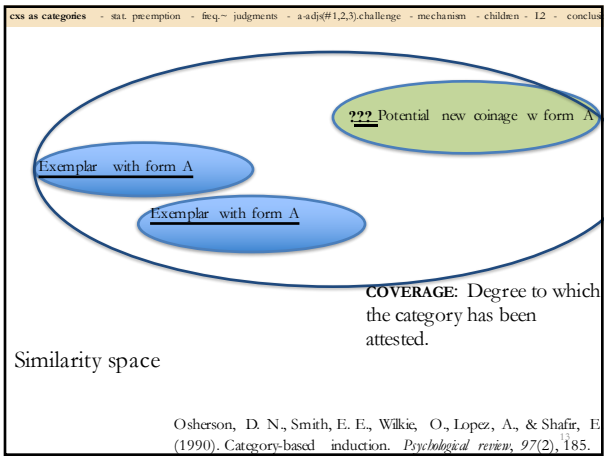
How likely is it that you can also say:
Rumple-*nu* the *pheb* the *jirm*.

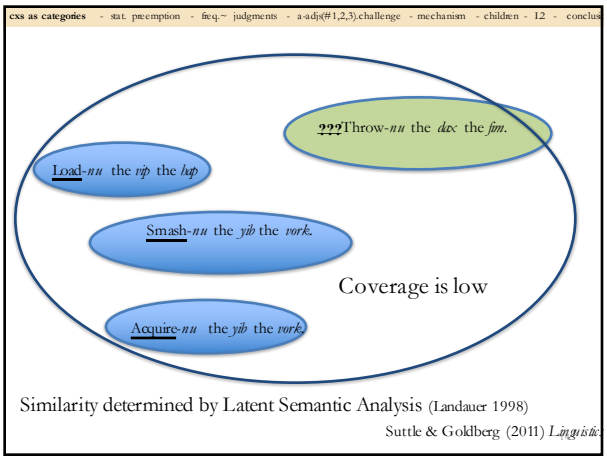
Subjects gave likelihood estimate from 0-100%

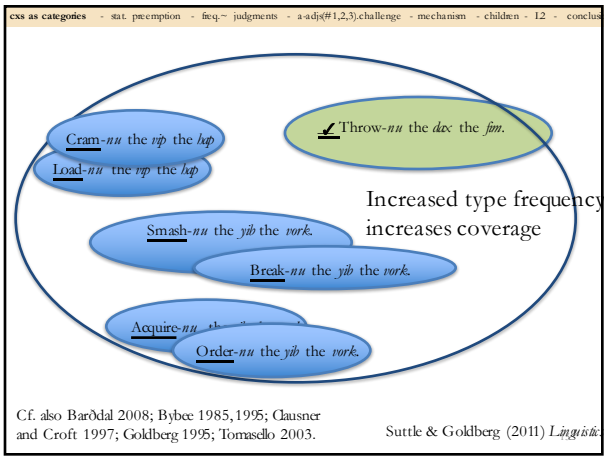
4 Studies (*N*= 55 distinct MT participants for each)
Results analyzed using mixed linear models, subjects & items as random effects.

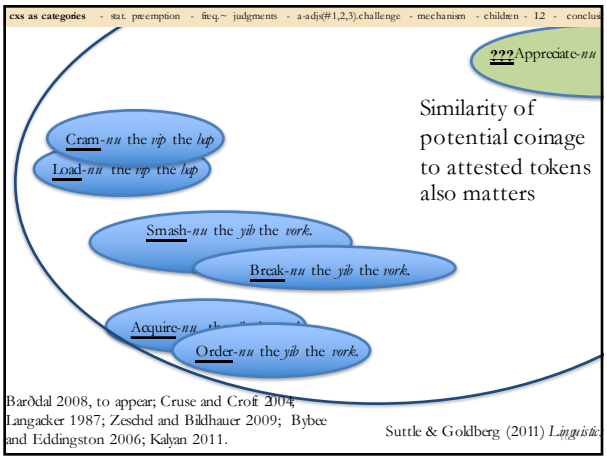
Suttle & Goldberg (2011) *Linguistic*

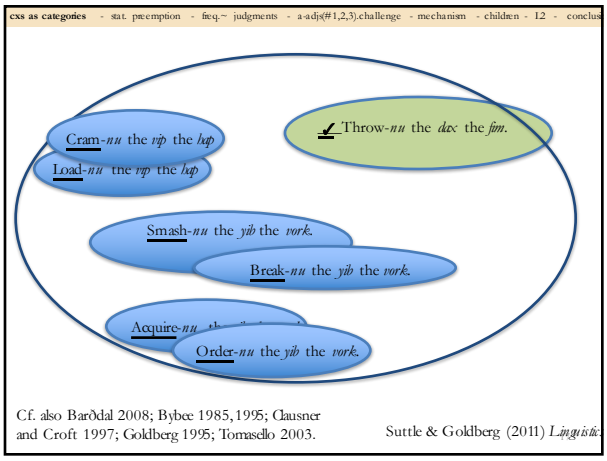
- **Varied type frequency**
- **Semantic variability**
- **Semantic similarity** of target sentence to closest attested instance

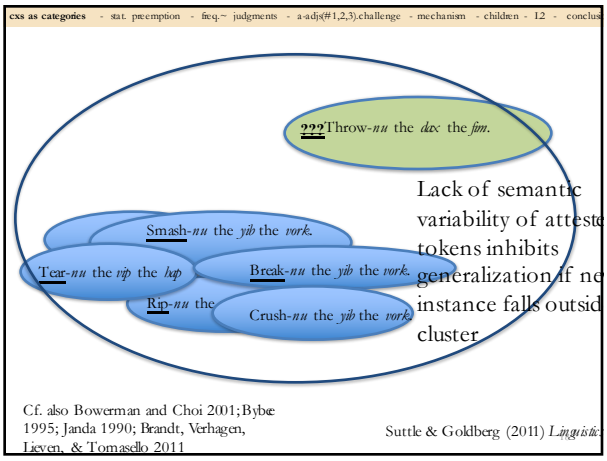

cxs as categories - stat. preemption - freq.~ judgments - a-adjs(#1,2,3).challenge - mechanism - children - 12 - conclus
???Potential new coinage w form A
Exemplar with form A
Exemplar with form A
COVERAGE: Degree to which the category has been attested.
Similarity space
Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E (1990). Category-based induction. Psychological review, 97(2), 185.


cxs as categories - stat. preemption - freq.~ judgments - a-adjs(#1,2,3).challenge - mechanism - children - 12 - conclus
???Throw-nu the dax the fim.
Load-nu the rip the hup
Smash-nu the yib the vork.
Acquire-nu the yib the vork.
Coverage is low
Similarity determined by Latent Semantic Analysis (Landauer 1998)
Suttle & Goldberg (2011) Linguistic


cxs as categories - stat. preemption - freq.~ judgments - a-adjs(#1,2,3).challenge - mechanism - children - 12 - conclus
Cram-nu the rip the hup
Load-nu the rip the hup
Throw-nu the dax the fim.
Smash-nu the yib the vork.
Break-nu the yib the vork.
Order-nu the yib the vork.
Increased type frequency increases coverage
Cf. also Barðdal 2008; Bybee 1985, 1995; Clausner and Croft 1997; Goldberg 1995; Tomasello 2003.
Suttle & Goldberg (2011) Linguistic


cxs as categories - stat. preemption - freq.~ judgments - a-adjs(#1,2,3).challenge - mechanism - children - 12 - conclus
???Appreciate-nu
Cram-nu the rip the hup
Load-nu the rip the hup
Smash-nu the yib the vork.
Break-nu the yib the vork.
Order-nu the yib the vork.
Similarity of potential coinage to attested tokens also matters
Barðdal 2008, to appear; Cruse and Croft 2004; Langacker 1987; Zeschel and Bildhauer 2009; Bybee and Eddington 2006; Kalyan 2011.
Suttle & Goldberg (2011) Linguistic


cxs as categories - stat. preemption - freq.~ judgments - a-adjs(#1,2,3).challenge - mechanism - children - 12 - conclus
Cram-nu the rip the hup
Load-nu the rip the hup
Throw-nu the dax the fim.
Smash-nu the yib the vork.
Break-nu the yib the vork.
Order-nu the yib the vork.
Cf. also Barðdal 2008; Bybee 1985, 1995; Clausner and Croft 1997; Goldberg 1995; Tomasello 2003.
Suttle & Goldberg (2011) Linguistic


cxs as categories - stat. preemption - freq.~ judgments - a-adjs(#1,2,3).challenge - mechanism - children - 12 - conclus
???Throw-nu the dax the fim.
Smash-nu the yib the vork.
Break-nu the yib the vork.
Crush-nu the yib the vork.
Lack of semantic variability of attested tokens inhibits generalization if ne instance falls outsid cluster
Cf. also Bowerman and Choi 2001; Bybee 1995; Janda 1990; Brandt, Verhagen, Lieven, & Tomasello 2011
Suttle & Goldberg (2011) Linguistic

Semantic (and phonological) categorization goes a long way to explaining which new instances sound acceptable

Still, there remain **odd gaps** in what we can say

(eg., Braine 1970; Lakoff 1970; Baker 1979; Bowerman 1988; Pinker 1989; Goldberg 1995; Ambridge et al. 2008; 2009; 2011; 2012):

??She explained him the story.
She told him the story.

??He vanished the rabbit.
He hid/banished the rabbit.

??She considered to go.
She wanted/hoped/planned to go.

?? The asleep boy
The astute/sleeping boy

Reliable direct feedback is not available

Me loves you, Mommy.

I have just completed a colorful mural on my bedroom wall with indelible markers.

(eg., Brown & Hanlon 1970; Hirsh-Pasek et al. 198

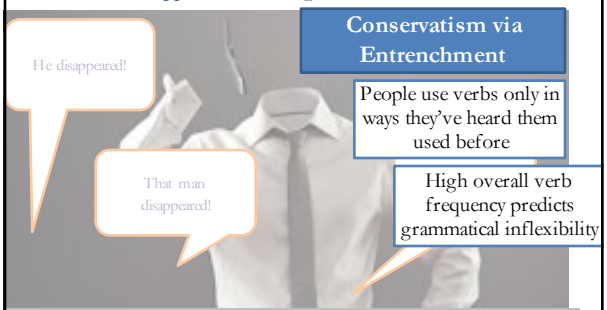
Is this how speakers learn to avoid saying:

?*She disappeared something.* ?

Braine & Brooks, 1995; Brooks et al. 1999; Theakston, 2004; Ambridge et al. 2008, 2011; 2012
Cf. also Hsu & Griffiths 2009; Stefanowitsch 2008; Perfors et al. 2010; Berwick 2012

Is this how speakers learn to avoid saying:

?*She disappeared something.* ?



Braine & Brooks, 1995; Brooks et al. 1999; Theakston, 2004; Ambridge et al. 2008, 2011; 2012
Cf. also Hsu & Griffiths 2009; Stefanowitsch 2008; Perfors et al. 2010; Berwick 2012

Verb Frequency Effect

?The magician vanished the rabbit.

>

??The magician disappeared the rabbit.

Brooks & Tomasello (1999)
Theakston (2004)
Ambridge et al. (2008)
Robenalt & Goldberg (2015, *Cognitive Linguistics;* to appear *Language Learning*)

BUT SPEAKERS CAN BE CREATIVE

"I grabbed the guy by the collar of his flowered shirt, **popped him a punch**" (COCA corpus)

"A 15-year-old **googled his way** to revolutionizing cancer detection…" (http://pixelsandpills.com)

""Mike couldn't believe she had managed to **flirt his wallet open** once again…" (COCA corpus)

HOW DO SPEAKERS LEARN WHAT *NOT* TO SAY?

Conservatism via Entrenchment:
People only repeat what they've heard

Statistical Preemption:
People learn to avoid certain novel formulations by systematically witnessing a competing alternative

Ambridge et al. (2014)
Boyd & Goldberg (2011)
Boyd, Ackerman & Kutas (2013)
Brooks & Tomasello (1999)
J. Claus (2014)
Goldberg (1995; 2006; 2011)

How do speakers learn what *not* to say?

Verb Frequency Effect

?The magician vanished the rabbit.
>
??The magician disappeared the rabbit.

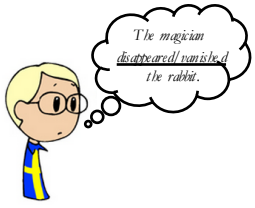
Brooks & Tomasello (1999)
Theakston (2004)
Ambridge et al. (2008)
Robenalt & Goldberg (to appear, *Cognitive Linguistics*; to appear *Language Learning*)

Statistical Preemption

People can be creative.
They can extend verbs to novel uses *as long as there isn't* a readily available alternative

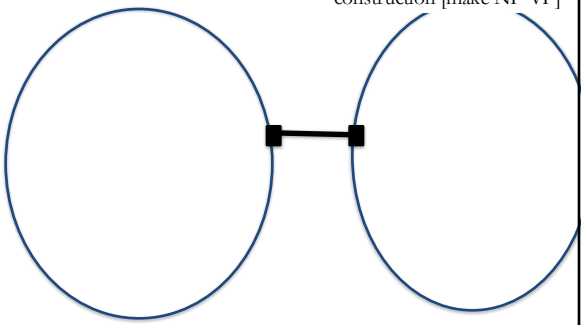
Disappear NP
sounds bad because **every time it would have been appropriate,** ***made NP disappear*** **occurred in its place.**

Make NP vanish has occurred less often.



Causative construction NP V NP

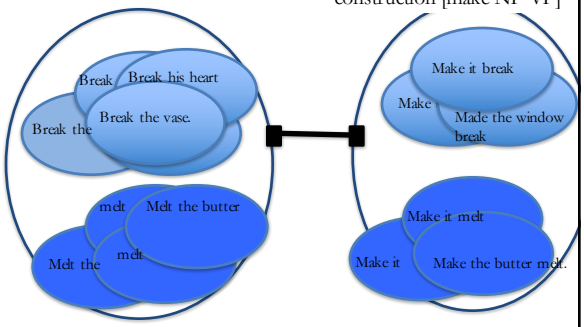
periphrastic causative construction [make NP VP]



Many verbs can appear in either, with corresponding difference in function

Causative construction NP V NP

periphrastic causative construction [make NP VP]

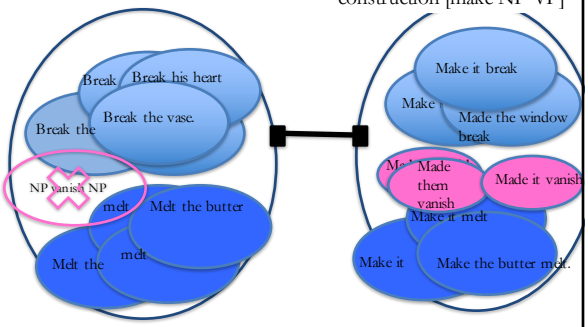


Many verbs can appear in either, with corresponding difference in function

31

Causative construction NP V NP

periphrastic causative construction [make NP VP]

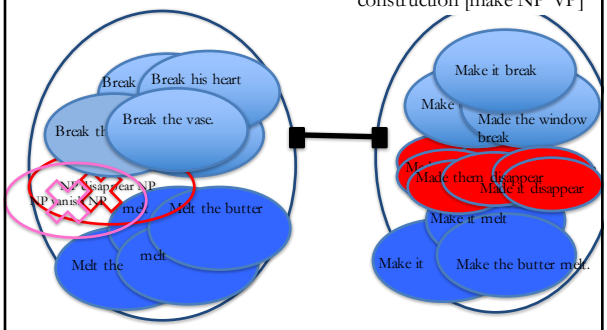


Provides evidence that *vanish* cannot occur in the causative construction

32

Causative construction NP V NP

periphrastic causative construction [make NP VP]

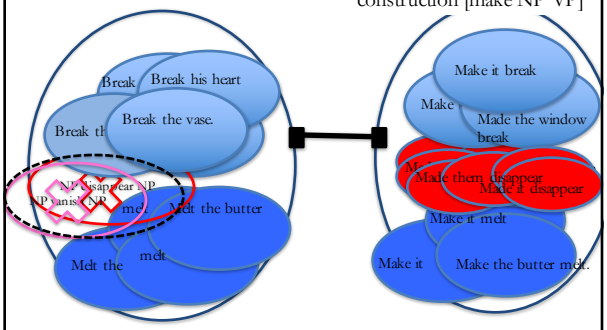


Increased type frequency provides better evidence

33

Causative construction NP V NP

periphrastic causative construction [make NP VP]



Possible to categorize non-occurrences as well? (stay tuned

34

Each construction has its own function (e.g., Bolinger 1971; Clark 1987; Langacker 1987; Kemmer & Verhagen 1994; Goldberg 1995).

Is that *a problem* for statistical preemption (Bowerman 1988; Pinker 1989)?

35

Is an advantage:

There will be contexts in which CxA should be preferred over CxB for verb$_i$.

If CxB is consistently witnessed instead… can learn that CxA is not possible for verb$_i$.

(Goldberg 1995)

36

Are higher frequency verbs *always* less flexible?

Conservativism via entrenchment: yes

Statistical preemption: only if another formulation is used in the same general context.

Study: Generate novel sentences with high and low frequency verbs (novelty confirmed via COCA)

37

Robenalt & Goldberg 2015: *Cognitive Linguistics*; 2016: *Language Learning*

Binned sentences via norming study

The editorial embarrassed the poor man out of town. — No-Competing Alternative (No-CA)
The editorial mortified the poor man out of town.

The chef coated ranch dressing over the salad. — Has-competing Alternative (Has-CA)
The chef doused ranch dressing over the salad.

38

Novel sentences without an agreed upon competing alternative phrasing (Novel-noCA)

The sound rattled/ reverberated the bats out of their hiding place.
The lifeguard swam/ paddled a pool toy to the kids.
The teacher frowned/ glowered a warning to the back of the class.
The crowd cheered/ hollered the reluctant candidate to the podium.
The editorial embarrassed/ mortified the poor man out of town.
The lion roared/ snarled the veterinarian out of the enclosure.
The editor smiled/ grinned the new reporter into his office.
The woman screamed/ shrieked the children out of the ice cream store.
The magician fascinated/ enthralled the toddlers into a trance.
Andrew insulted/ derided the potential member out of the club.

No-Competing Alternative (No-CA)

Novel sentences with an agreed upon competing alternative phrasing (Novel-hasCA)

The scientist infected/ corrupted bacteria into the sample.
The children soiled/ splotched mud onto the carpet.
The designer decorated/ embellished lace onto the invitation.
The dictator flooded/ inundated propaganda into the city.
The chef coated/ doused ranch dressing over the salad.
The housekeeper soaked/ drenched bleach into the towel.
The nurse bound/ bandaged cotton over the wound.
Natalie smacked/ swatted a newspaper onto the mosquito.
The landscaper surrounded/ bordered rocks around the garden.
The camper blocked/ obstructed a heavy backpack into the entrance.

Has-competing Alternative (Has-CA)

39

Acceptability Ratings

N=108 from MT

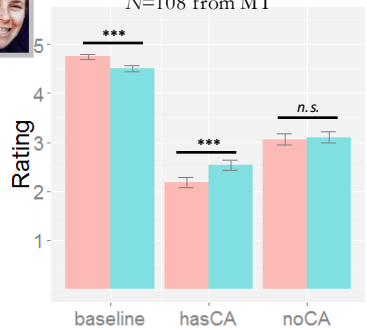
1	2	3	4	5
Completely unacceptable	Bad but not terrible	In between	Good but not perfect	Completely acceptable

The chef coated ranch dressing over the salad. _____
The magician vanished the rabbit. _____
Ashley was terribly mortified. ←baseline _____
……

40

Acceptability Ratings

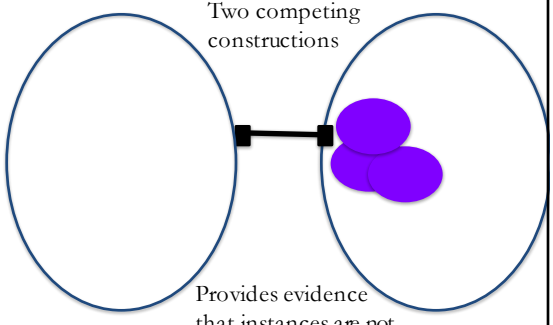
N=108 from MT



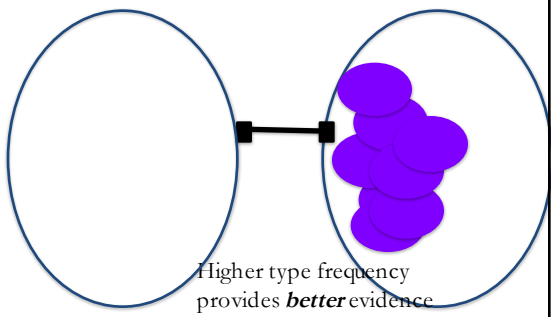
Statistical Preemption: Verb frequency only makes a difference if the sentence has a preemptive alternative.

Mixed Linear Model; subjects, items as random effects; length, semantic features of the verbs, plausibility as fixed effects

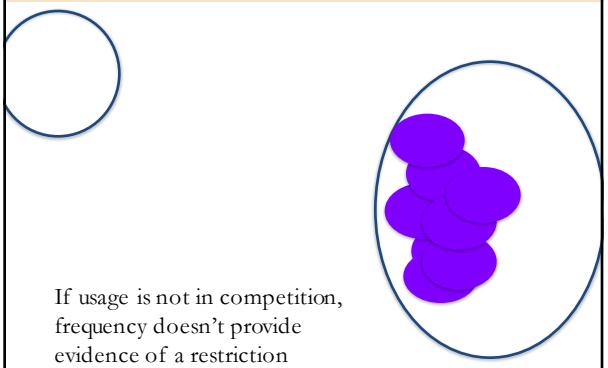
Robenalt & Goldberg 2015, *Cognitive Linguistics*



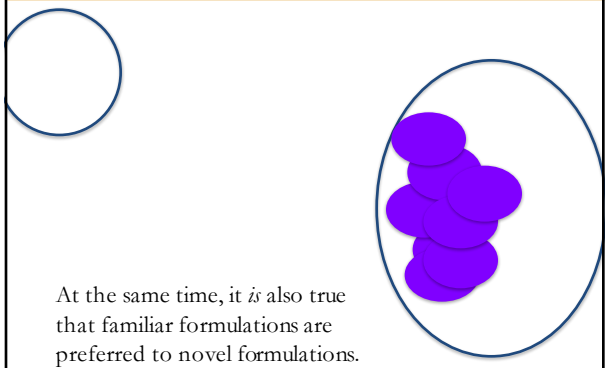
42



Higher type frequency provides ***better*** evidence that instances are not in second category



If usage is not in competition, frequency doesn't provide evidence of a restriction



At the same time, it *is* also true that familiar formulations are preferred to novel formulations.

Interim summary

People prefer to use familiar formulations.

But we <u>are</u> willing to use verbs creatively.

Frequency of verb *in competing constructions* is relevant to unacceptability.

Quantifying evidence for statistical preemption

Probability of CxB statistically preempting CxA for *verbi*:
P(CxB| a discourse context in which the learner might expect to hear CxA[*ver*

Confidence of statistical preemption for verbi:
ln frequency (CxB & [CxA would be at least as appropriate])

Goldberg, 2011, *Cognitive Linguistics*

A *pronominal recipient* and *lexical theme* favors the ditransitive construction over the dative.

She gave me the ball. >
She gave the ball to me.

Thompson 1990; 1995; Arnold et al. 2000; Bresnan et al. 2007; Collins 1995; Dryer 1986; Erteschik-Shir 1979; Givón 1979; 1984; Goldberg 1995; Goldberg 2006; Green 1974; Hovav Rappaport and Levin 2008; Oehrle 1975; Wasow 2002.

"Alternating" verbs: Probability of a DATIVE given a context appropriate for a DITRANSITIVE. COCA Corpus

Alternating verbs:	A Dative: [4] [v] [lexical NP] to [definite pronoun]	Ditransitive: [v] [definite pronoun] [lexical NP]	P (dative | dative with relevant restrictions + ditransitive)
Tell	36	3713	< .01
Give	111	7982	.01
Show	35	932	.04
Send	146	1098	.12
Sell	40	152	.21
Bring	111	415	.21
Read	81	275	.23
Lend	7	176	.04
Total	567	14743	.04 (Average)

P(CxB | a discourse context in which the learner might expect to hear CxA[*verbi*])
P (dative | discourse context in which one might expect to hear ditransitive) = .04

"NON-alternating" verbs: Probability of a DATIVE given context appropriate for a DITRANSITIVE. COCA Corpus

Non-alternating verbs:	**Dative:** [v] [lexical NP] to [definite pronoun]	Ditransitive: [v] [definite pronoun] [lexical NP]	**P (dative | dative with relevant restrictions + ditransitive)**
Explain	120	1	.99
Whisper	16	1	.94
Transfer	20	0	1.0
Return	74	11	.88
Entrust	13	0	1.0
Deliver	33	18	.65
Present	43	37	.53
Repeat	26	0	1.0
Total	345	69	.83 (Average)

P(CxB | a discourse context in which the learner might expect to hear CxA[*verbi*])
P (dative | discourse context in which one might expect to hear ditransitive) = .83

CONFIDENCE of statistical preemption for verbi:
ln frequency (CxB & CxA would be at least as appropriate)

Non-alternating verbs:	P (dative | context appropriate to ditrans)	Confidence of statistical preemption
Explain	.99	ln(120) = 4.78
Whisper	.94	ln(16) = 2.77
Transfer	1.0	ln(20) = 2.99
Return	.88	ln(74) = 4.30
Entrust	1	ln(13) = 2.56
deliver	.65	ln(33) = 3.49
present	.53	ln(43) = 3.76
repeat	1.0	ln(26) = 3.26

STRENGTH of statistical preemption: a function of PROBABILITY and CONFIDENCE of the preempting construction.

Relevant data is available to learners/analysts in large corpora.

Another subtle restriction : "a-adjectives":

??the/an **asleep** child
??the/an **afloat** ship
??the/an **alive** monster
??the/an **aghast** audience
??the/an **ablaze** building
??the/an **afraid** child

Semantic near-synonyms
a. the/a sleeping child
b. the/a scared man
c. the/a living monster
d. the/a burning building

Phonologically related NON-a-adjs
a. the/an adult male
b. the/an acute sickness
c. the/an astute comment
d. the/an aloof woman

Explanation for the distribution:

Historical "persistence":

	Old English
asleep	< *in sleep*
abloom	< *in bloom*
adrift	< *on drift*
afloat	< *on float*
ablaze	< *on blaze*

As PPs, **the on drugs man*

afraid <> "on fraid"

Instead *afraid* < p. ppl of *affray (v.): "to startle"*

Speakers have assimilated *afraid* to category of a-adjectives.

Synchronically: Requires usage-based model: speakers are aware of which adjectives they've heard in which constructions.

Constructionist (usage-based) view

- Learners record statistics about particular items' distribution.
- Constructional generalizations emerge from learners categorizing over the input.

A-adjective experiment

- Undergraduate native speakers of English (n = 32)
- Production task

Will novel a-adjectives get assimilated to the category? (exp. #1)

Do speakers make use of statistical preemption? (exp. #2) Wisely? (exp. #3)

(Boyd and Goldberg, 2011, *Language*)

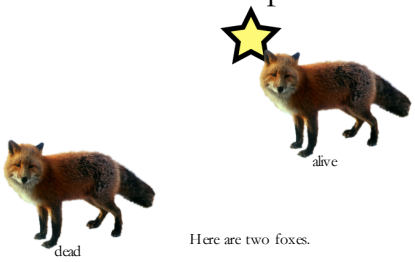
An example trial

dead

Here are two foxes.

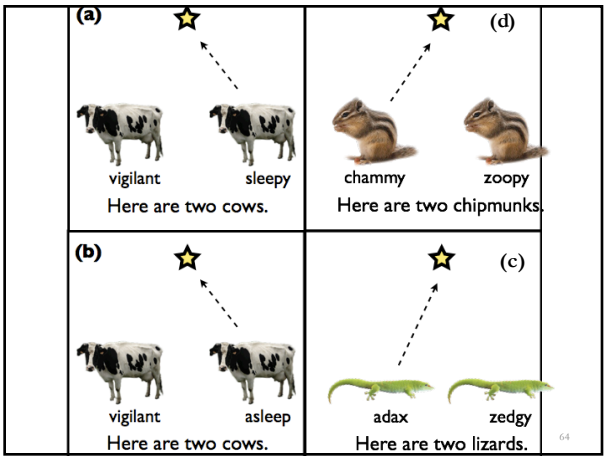
alive

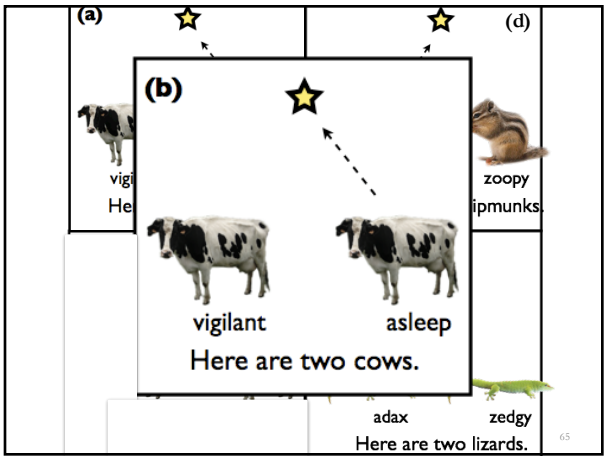
An example trial

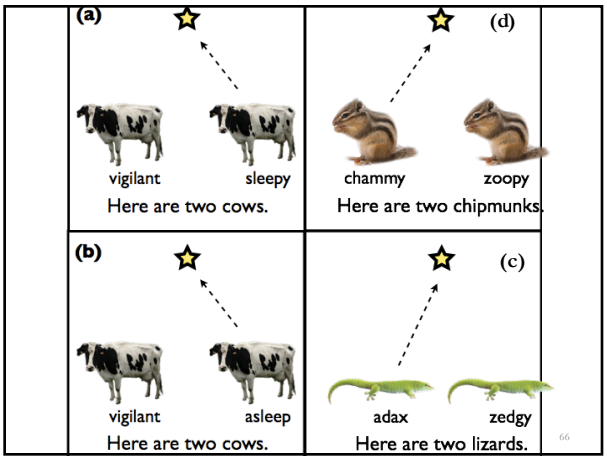


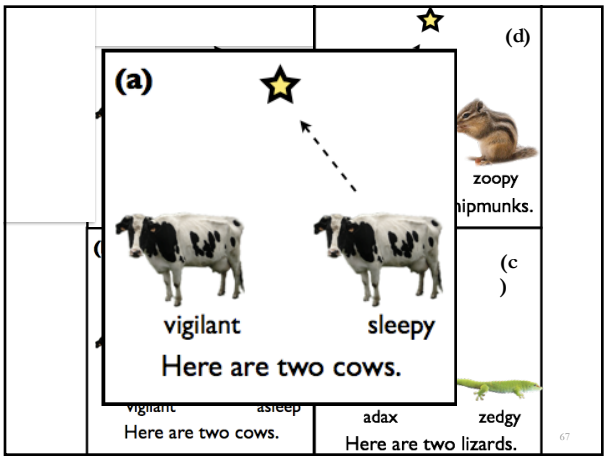
Here are two foxes.

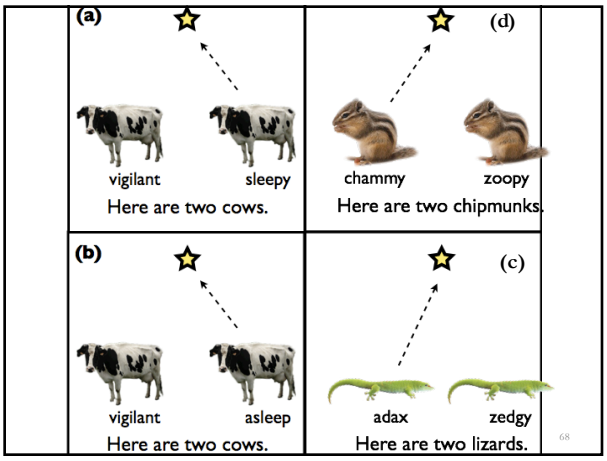
Which fox moved to the star?

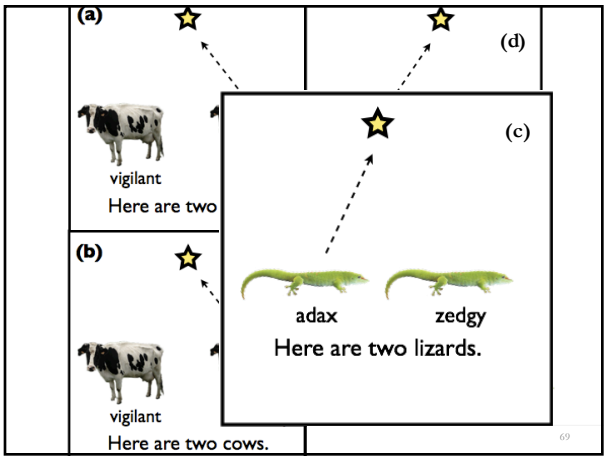


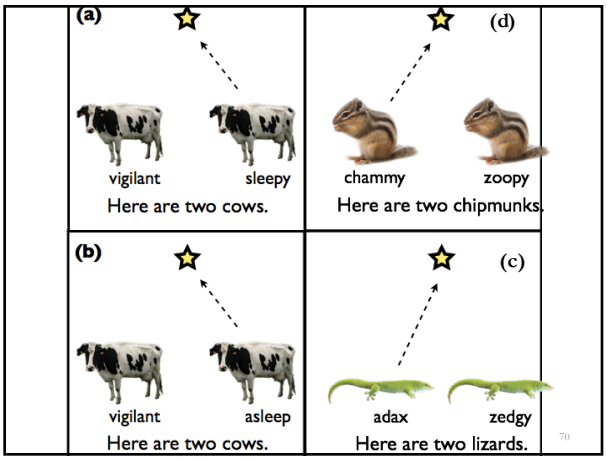


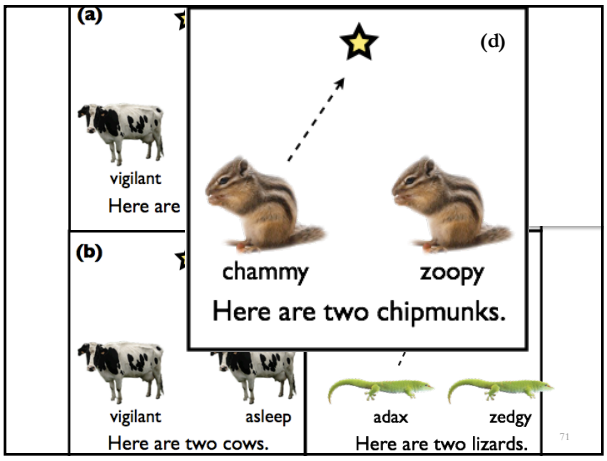












cxs as categories - stat. preemption - freq.~ judgments - **a-adjs (#1,2,3)**.challenge - mechanism - children - 12 - conclusion

24 counterbalanced fillers:

Prototypical adjectives: likely to elicit prenominal use: *the bad dog*

bad, good, smart, dumb, fast, slow, old, young, rich, poor, strong, weak

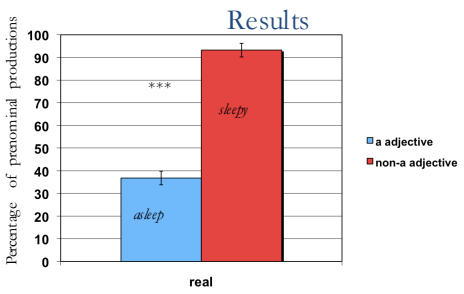
Present tense verbs: likely to elicit predicative use: the dog that bites

bites, camps, cooks, cries, gambles, reads, runs, smokes, snowboards, travels, votes, wr

Fillers elicited the intended structure 99% of the time.

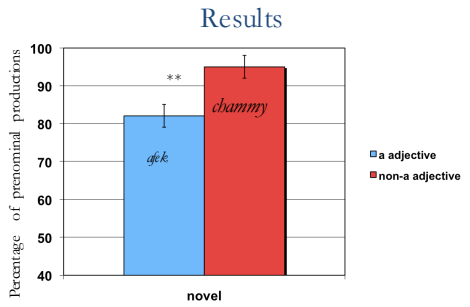
72

- Exposure Block
  - **6 "practice trials"**: 3 attributive and 3 predicative examples using non-experimental words.
- Test Block
  - 16 critical trials interleaved with 16 filler trials



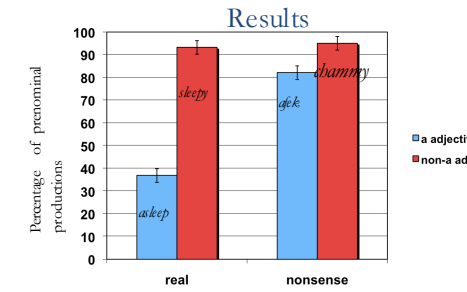
Speakers do in fact avoid using real a-adjs (*afraid, asleep*) prenominally

(Boyd and Goldberg, 2011, *Language*)



Speakers also avoid using nonsense a-adjs (*ablim, afek*) prenominally

(Boyd and Goldberg, 2011, *Language*)



Restriction on real adjectives is stronger than restriction on nonse adjectives.

(Boyd and Goldberg, 2011, *Language*)

Study #2: introduce preemptive context for novel a-adjective

An example trial

afek

zeky

Here are two foxes.

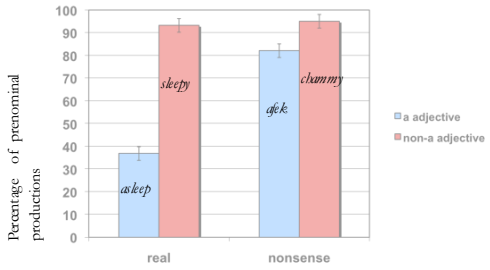
"The fox that's afek moved to the star"

Design

- Exposure Block
  - 6 "practice trials": participants witness 2 novel a-adjectives in Relative Clause: preemptive context.
- Test Block: same as before
  - 16 critical trials interleaved with 16 filler trials.
    - 2 novel a-adjectives were seen during exposure, 2 new novel adjectives.
  - To encourage response variability, fillers were strongly biased towards either an attributive or RC response.

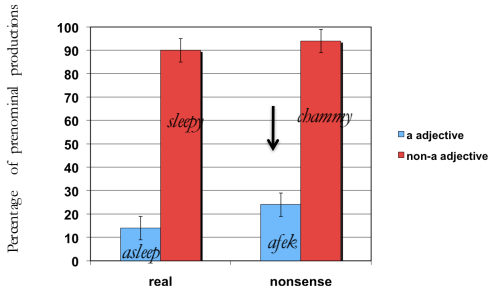
Recall: experiment **#1** results

Recall: experiment **#1** results



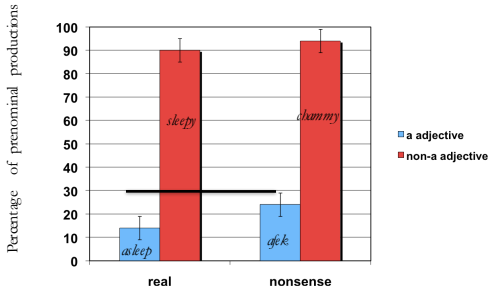
(Boyd and Goldberg, 2011, *Language*)

Results of **Exp. #2: with preemptive context**: (n=20)



(Boyd and Goldberg, 2011, *Language*)

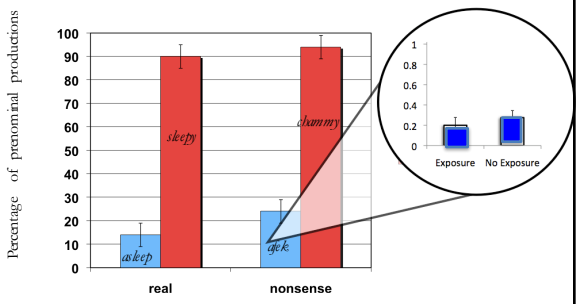
Results of **Exp. #2: with preemptive context**: (n=20)



No difference between real a-adjectives (14%) and novel (24%), $t(19) = 1.25$, $p = .23$.

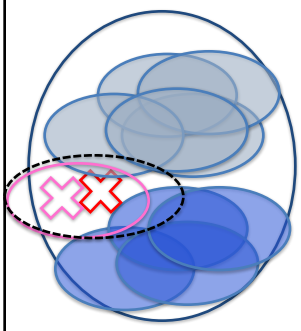
(Boyd and Goldberg, 2011, *Language*)

Results of **Exp. #2: with preemptive context**: (n=20)



No difference between novel a-adjectives seen during exposure (20%) and those that were (27.5%), $t(19) = -1$, $p = .33$.

(Boyd and Goldberg, 2011, *Language*)



It is possible to categorize non-occurrences!

Exp. 3: pseudo-preemptive context (conservatism?)

Are people savvy enough to know when a context is truly preemptive?

Exposure to pseudo-preemptive context:

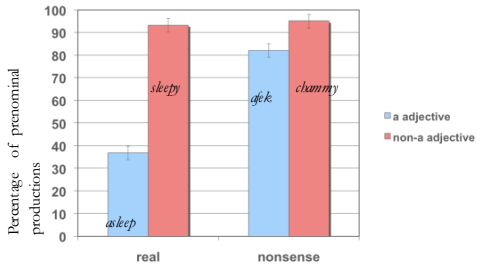
The fox that's adax and proud of itself…

Notice that prenominal attributive construction is unavailable:

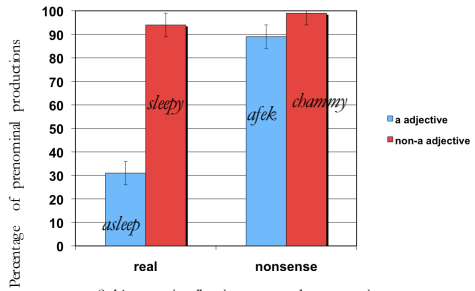
**The proud of itself fox…*

**The afek and proud of itself fox….*

Recall: experiment #1 results



Results of Exp. #3: with pseudo-preemptive context: (n=20)



Subjects rationally ignore pseudo-preemptive context (eg., *The fox that was ablim and proud of itself.*)

(Boyd and Goldberg, 2011, *Language*)

Interim summary

- Learners record the statistics of their language.
- Learners categorize their input into patterns.
- **Statistical preemption**: learners learn to avoid certain formulations because an alternative formulation is systematically used in the appropriate context.
- Learners are smart about what counts as a preemptive context.

Recent challenge to statistical preemption account of ***??an afraid boy***

Yang 2015 (*Language*), Bruening 2011:

- Children do not witness enough evidence for statistical preemption before they are 3.
- Learners don't need indirect negative evidence because a-adjectives pattern like particles, *on, up, out*.
- The suggested evidence: a-adjectives are claimed to occur with adverbs *right, straight, well, far*, which also apply to particles.

1. TYPICAL ADJECTIVES	2. *A*-ADJECTIVES	3. PREPOSITIONAL PHRASES:	4. LOCATIVE PARTICLES
red	*asleep*	*on the table*	*up*
sleepy	*afloat*	*into the room*	*down*
floating	*afraid*	*to the house*	*on*
full	*alone*	*at two o'clock*	*in*
huddled	*ablaze*	*in the mind*	*inside*
pinkish	*abloom*	*inside the box*	*around*
absurd	*alike*	*out of the city*	*away*
acute	*alive*	*around the ring*	*across*
aloof	*awake*	*away from them*	*out*
	aware		

BUT, many a-adjectives *never* occur with *any* of the key adverbs in all of COCA corpus:

**far asleep, *straight asleep, *well asleep*

**far awake, *straight awake, *well awake*

(and the ones that do occur only do so extremely rarely (*?well alive; ?far alon*

REGULAR adjectives occur *much* more commonly:

e.g.,

far greater (1300)

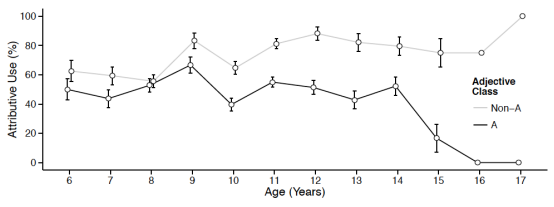
well pleased (70) (Yang classifies *across, away, around, ahead* as a-*adjectives*)

straight winning (14)

In any case, particles *can* occur in attributive position:

the near future
the past year
the outside world
the inside track

Finally, children do not learn the restriction until they are 10 years old:



Jessica Hao (2015) Princeton University senior thesis

Summary

Yang 2015 (*Language*), Bruening 2011:

- Children do not witness enough evidence for statistical preemption before they are 3.
  And they don't actually show the restriction until they are roughly 10 years old.
- Learners don't need indirect negative evidence because a-adjectives pattern like particles, *off, on, up, out*.
  This is a suggested *off button*. also *near future, past year, inside track*
- The suggested evidence: a-adjectives are claimed to occur with adverbs *right, straight, well, far*, which also apply to particles.
  Actually these adverbs almost *never* occur with a-adjectives, while they do occur with regular adjectives.

Evidence that a-adjectives are adjectives

- Semantics
- They appear after bare *seem* (Lakoff 1970; Jackendoff 1972)
  - *Pat seemed afraid/alone/alive.*
- They are conjoinable with other adjectives
  - *Chris was afraid but happy.*
  - *She was alive and healthy.*

(see corpus analysis in Goldberg & Boyd, to appear, *Language)*

97

→ Statistical preemption is consistent with available evidence, while the suggested positive evidence is not (pace Yang 2015; Bruening 2011)

98

Error-driven learning

We anticipate (predict) what others will say as they speak.

(Kutas & Hillyard 1984; Tanenhaus et al. 1995; McRae, Spivey-Knowlton, Tanenhaus 1998; Pickering and Garrod 2007, 2012, Johnson, Turk-Browne & Goldberg to appear, *LCN*)

99

Competition-dependent learning

Partial activation of competing form leads to learned dissociation (Anderson & Spellman 1995; Anderson, Green & McCulloch 2000; Detre 2010; Newman and Norman 2010).

100

Anderson & Spellman 1996:
Subjects learned paired associations, e.g.,
Fruit-Apple
Fruit-Pear
Fruit-Kiwi
Furniture-Table
Furniture-Phone
……
Then cued with a subset of these pairs such as:
Fruit-Pe___.

Note "Pear" is only partially cued, therefore subjects partially activate othe *prototypical* associates, e.g., Apple.
- memory for Fruit-**Apple** was weakened
- memory for Fruit-**Kiwi** was not weakened.
- memory for other uncued pairs, e.g., Furniture-Table was not weakened either.

101

If speakers anticipate (1), it is *partially activated.* If the competing form (2) is witnessed instead, subsequently, (1) is harder to retrieve (forgotten/suppressed).

1. ?? She explained him something.
2. She explained something to him.

(2) preempts (1).

102

Relationship to preemption

A construction that is in competition is weakened whenever another form "wins" (is used).

If a competing construction is *not* partially activated, there is no suppression.

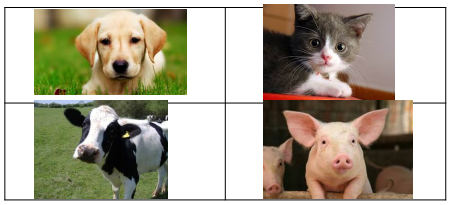
When do children generalize?
As soon as the tokens are recognized to form a pattern.
Different ages for different types of patterns.

e.g., Elena Lieven et al. (1997)
Mike Tomasello (2000; 2003)
Berko, J. (1958).

Much to do here: which dimensions of similarity are relevant to which constructions and how do children determine this? (Perek & Goldberg, to appear, *JML*)

Why do children overgeneralize?
They do not have the alternative readily available (in general, or at the moment of speaking).

Huttenlocker 1979; Benedict 1979; Gruendel 1977, Gershkoff-Stowe 2001



How do children eventually recover from overgeneralizations?
The more conventional alternative becomes more readily available (through more exposure & greater fluency.

Why might **L2** learners find it harder to learn arbitrary restrictions?

- ?? "could you recommend me some [place] to apply?"
- ?? "maybe it's better to explain me first"
- ?? "have you ever considered to go climbing in Ecuador? We would love for you to come and discover it!"
- ?? "the afraid boy hiding behind the board"

"The RAGE hypothesis"
(Grüter & Rohde; Lew-Williams & Fernald 2010)

- Non-native speakers have reduced ability to generate expectations during language processing,
- particularly based on grammatically encoded distinctions.

Encuentra la pelota

(la pelota)

(el zapato)

 Lew-Williams & Fernald (2007; 2010)

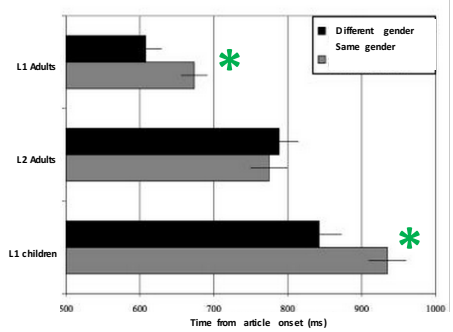
Encuentra la pelota

(la galleta)

(la pelota)

 Lew-Williams & Fernald (2007; 2010)

Referent Identification Time

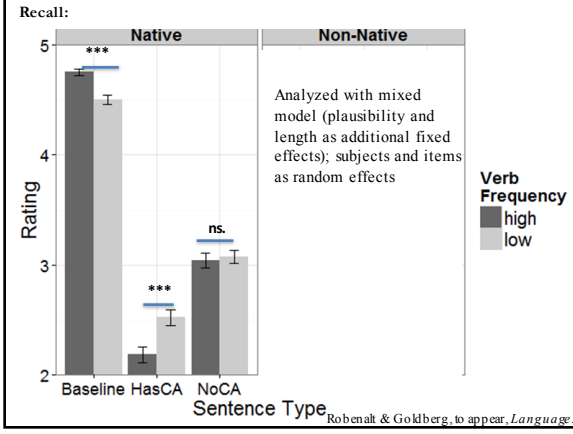


 Lew-Williams & Fernald (2007; 2010)

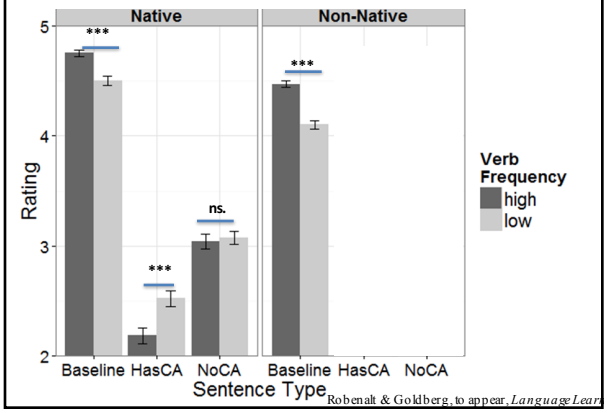
- L2 processing is a cost on WM
- Individual differences, degree of proficiency, and task can play a role in the extent to which processing is affected.

Grüter et al. (2014); Havik et al(2009); Lew-Williams & Fernald (2007; 2010); Linck et al. 2009; Martin (2013)

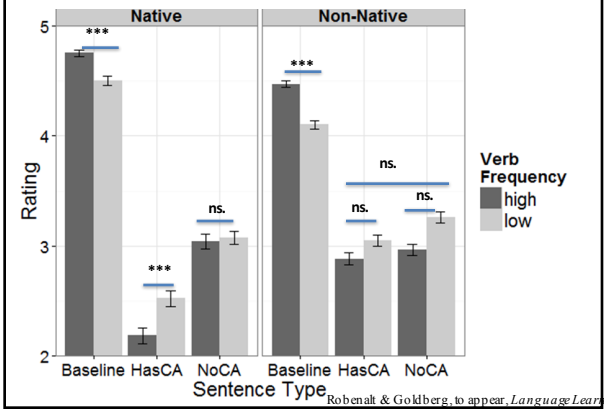
If non-native speakers don't anticipate upcoming utterances to the same extent as native speakers, they will have less opportunity for competition-dependent learning.



Robenalt & Goldberg, to appear, *Language Learn*



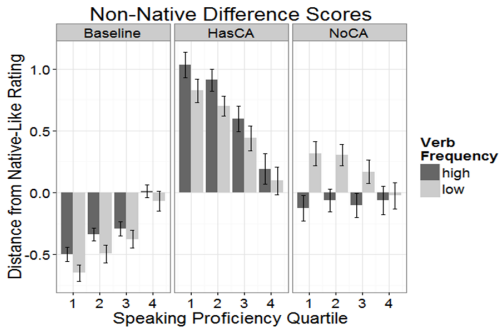
Robenalt & Goldberg, to appear, *Language Learn*



Robenalt & Goldberg, to appear, *Language Learn*

L2 learners tended to treat novel sentences as novel, without taking into account whether a readily available competing alternative existed or not.

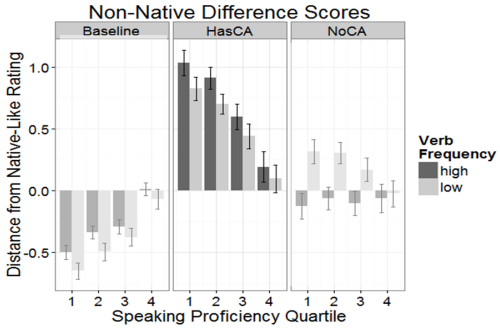
117



118

Robenalt & Goldberg, to appear, *Language Learn*

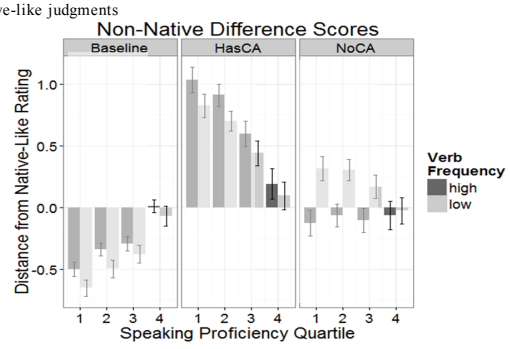
Not clear that L2 learners make use of statistical preemption to the sam extent as L1.



119

Robenalt & Goldberg, to appear, *Language Learn*

Speakers at the highest quartile of self-reported speaking proficiency showed native-like judgments



120

Robenalt & Goldberg, to appear, *Language Learn*

Conclusion: short version

It's categorization.

Conclusion: more specifically

- Learners record statistics of their language.
- We actively categorize the input, recognizing patterns of form and function.
- Productivity is generally determined by general principles of induction (coverage & similarity)
- Learners learn to avoid certain formulations because an alternative formulation is systematically used instead: **statistical preemption. (i.e competition driven learning)**
- Adults are smart about what counts as a preemptive context.

- Statistical preemption is a slow, gradual process; some arbitrary restrictions are learned very late.
- L2 learners may take less able to take full advantage of statistical preemption.

- Thanks to:

Labex, Pollet Samvelian!

Jeremy Boyd

Laura Suttle

Clarice Robenalt

Andrew Conway
Sam Glucksberg
Matt Johnson
Phil Johnson-Laird
Harvey Keselman

Florent Perek

Ben Ambridge

Thank you!

EINSTEIN Stiftung Berlin
Einstein Foundation Berlin

Freie Universität Berlin

PRINCETON UNIVERSITY