Ode to Communication, Emergentism, Cue validity, Competition

Adele E. Goldberg
Princeton University
1997
MacWhinney, Bates & Kliegl (1984): “the forms of natural languages are created, governed, constrained, acquired and used in the service of communicative functions.”

Understand messages, given forms (comprehension) &
Choose forms, given intended message (production) &
Conform to the conventions of their community
Need to learn

form \sim function pairings: \textit{constructions}
<table>
<thead>
<tr>
<th>CONSTRUCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word</strong></td>
</tr>
<tr>
<td><a href="#">Double object construction</a></td>
</tr>
<tr>
<td><a href="#">Passive</a></td>
</tr>
</tbody>
</table>
## CONSTRUCTIONS

<table>
<thead>
<tr>
<th>Word</th>
<th>e.g., Brian, CLAN, CHILDES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word (partially filled)</td>
<td>e.g., N-Bank, Name-fest, #-th</td>
</tr>
</tbody>
</table>
### CONSTRUCTIONS

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>e.g., Brian, CLAN, CHILDES</td>
</tr>
<tr>
<td>Word (partially filled)</td>
<td>e.g., N-Bank, Name-fest, #-th</td>
</tr>
<tr>
<td>Idiom (filled)</td>
<td>e.g., dead as a doornail; when pigs fly</td>
</tr>
</tbody>
</table>
## CONSTRUCTIONS

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>e.g., Brian, CLAN, CHILDES</td>
</tr>
<tr>
<td>Word (partially filled)</td>
<td>e.g., N-Bank, Name-fest, #-th</td>
</tr>
<tr>
<td>Idiom (filled)</td>
<td>e.g., dead as a doornail; when pigs fly</td>
</tr>
<tr>
<td>Idiom (partially filled)</td>
<td>e.g., Jog &lt;someone’s&gt; memory, I hope this &lt;N&gt; finds you well</td>
</tr>
</tbody>
</table>
## CONSTRUCTIONS

<table>
<thead>
<tr>
<th></th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>e.g., Brian, CLAN, CHILDES</td>
</tr>
<tr>
<td>Word (partially filled)</td>
<td>e.g., N-Bank, Name-fest, #-th</td>
</tr>
<tr>
<td>Idiom (filled)</td>
<td>e.g., dead as a doornail; when pigs fly</td>
</tr>
<tr>
<td>Idiom (partially filled)</td>
<td>e.g., Jog &lt;someone’s&gt; memory, I hope this &lt;N&gt; finds you well</td>
</tr>
<tr>
<td>The Xer the Yer Cx (partially filled)</td>
<td>(e.g., The more you think about it, the less you understand)</td>
</tr>
<tr>
<td>CONSTRUCTIONS</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td><strong>Word</strong></td>
<td>e.g., Brian, CLAN, CHILDES</td>
</tr>
<tr>
<td><strong>Word (partially filled)</strong></td>
<td>e.g., N-Bank, Name-fest, #-th</td>
</tr>
<tr>
<td><strong>Idiom (filled)</strong></td>
<td>e.g., dead as a doornail; when pigs fly</td>
</tr>
<tr>
<td><strong>Idiom (partially filled)</strong></td>
<td>e.g., Jog &lt;someone’s&gt; memory, I hope this &lt;N&gt; finds you well</td>
</tr>
<tr>
<td><strong>The Xer the Yer Cx</strong> (partially filled)</td>
<td>(e.g., The more you think about it, the less you understand)</td>
</tr>
<tr>
<td><strong>Double object construction:</strong> (unfilled) Subj V Obj1 Obj2</td>
<td>(e.g., He gave her a fish taco; He baked her a muffin.)</td>
</tr>
</tbody>
</table>
## CONSTRUCTIONS

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>e.g., Brian, CLAN, CHILDES</td>
</tr>
<tr>
<td>Word (partially filled)</td>
<td>e.g., N-Bank, Name-fest, #-th</td>
</tr>
<tr>
<td>Idiom (filled)</td>
<td>e.g., dead as a doornail; when pigs fly</td>
</tr>
<tr>
<td>Idiom (partially filled)</td>
<td>e.g., Jog &lt;someone’s&gt; memory, I hope this &lt;N&gt; finds you well</td>
</tr>
<tr>
<td>The Xer the Yer  Cx (partially filled)</td>
<td>(e.g., The more you think about it, the less you understand)</td>
</tr>
<tr>
<td>Double object construction:</td>
<td>(e.g., He gave her a fish taco; He baked her a muffin. )</td>
</tr>
<tr>
<td>Passive:</td>
<td>(e.g., The armadillo was hit by a car)</td>
</tr>
<tr>
<td>Subj aux VP_{pastpart}(PP_{by})</td>
<td></td>
</tr>
</tbody>
</table>
Long term memory: vast, associative, content-addressable

Constructions:
emergent clusters, conditioned by enormous range of factors

Failure to access an ideal construction to express intended message
- overextensions
- regularization
- conservativism
- polysemy
- CREATIVITY

co-activated constructions combine or COMPETE:

4th, 5th, 100th, gazillionth vs. 3rd

RYAN RULZ OK 🔄 @ryanqnorth · Jun 1
Strawberry jam: hi i'm strawberry jam
Blueberry jam: hi i'm blueberry jam
Raspberry jam: hi i'm raspberry jam
Orange jam: BoNjOuR, you may call me MARMALADE
POLYSEMY:
If best match is inaccessible, use a good-enough match.

baby

kitten
co-activated constructions combine or COMPETE

4th, 5th, 100th, gazillionth vs. ?3th
baby gorilla, baby aardvark vs. ?baby cat

3rd kitten
She sneezed.

Simple action

She sneezed her tooth across town.

+Caused motion

She sneezed a terrible sneeze.

+Creation

She sneezed herself awake.

+Change of state

She sneezed her way to the ER.

+Creation of a path & motion thru it
Nouns (nouns, names, and other things) as verbs
(Clark & Clark 1979)

- “I'm gonna go benadryl myself to sleep.”

- “Trump is trying to out-Nixon Nixon.”

- “few people want to be memorialized um-ing, you know-ing, and remember that time when we got drunk-ing their way into infamy.”
Cue reliability: $P(\text{meaning} \mid \text{construction})$
MacWhinney 1987; Bates & MacWhinney 1989; MacWhinney 1997; MacWhinney 2001; MacWhinney 2018

Constructions have high cue validity for function
Goldberg, Casenhiser & Sethuraman 2005; Ellis 2017; Perek & Goldberg 2017; Thorthathiri & Rattinger 2016
Verbs often provide a highly reliable cue:

She ___ him something. (DO construction)

\[ P(\text{give} \mid \text{DO}) \approx 0.50 \]

Goldberg, Casenhiser & Sethuraman 2004; Casenhiser & Goldberg 2005; Ellis 2006; Ellis & Ferreira-Junior 2009; Wulff 2013
Verbs often provide a highly reliable cue:

She ___ him something. (DO construction)

\[ P(\text{give} \mid \text{DO}) \approx 0.50 \]

She ___ it someplace. (Caused-motion construction)

\[ P(\text{put} \mid \text{NP Loc}) \approx 0.40 \]

Goldberg, Casenhisier & Sethuraman 2004; Casenhisier & Goldberg 2005; Ellis 2006; Ellis & Ferreira-Junior 2009; Wulff 2013
Verbs often provide a highly reliable cue:

She ___ him something.  (DO construction)

\[ P(\text{give} \mid \text{DO}) \approx 0.50 \]

She ___ it someplace  (Caused-motion construction)

\[ P(\text{put} \mid \text{NP Loc}) \approx 0.40 \]

She ___ someplace  (Intrans. motion construction)

\[ P(\text{go} \mid \text{Loc}) \approx 0.54 \]

Goldberg, Casenhiser & Sethuraman 2004;
Casenhiser & Goldberg 2005; Ellis 2006; Ellis & Ferreira-Junior 2009; Wulff 2013
Constructions as generalization over a structured set of item-based patterns

(It is) \( <\text{adj}> \) of \( \text{NP}_{\text{agent}} \) \( \text{VP}_{\text{to}} \)
e.g., It’s _____ of you to be here.

SPEAKER JUDGES
SOMEONE’s ACTION

It’s nice/good of you to be here. (67% AmEng: nice)
??It’s tall of you to reach the top shelf.
It’s big of you to reach the top shelf.
??It was good of the dishwasher to save water.

\[ P (nice \mid \text{Adj of NP cx}) \approx 0.16 \]
Odd gaps:
(e.g., Braine 1970; Lakoff 1970; Baker 1979; Bowerman 1988; Pinker 1989; Goldberg 1995; Ambridge et al. 2008; Yang 2015):

??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.

• https://video.twimg.com/ext_tw_video/1136946839826116609/pu/vid/406x720/9OMrIxf4fGTy6OSX.mp4?tag=10

(e.g., Brown & Hanlon 1970; Hirsh-Pasek et al. 1984; cf. Boh...)
• co-activated constructions combine or COMPETE

4th, 5th, 100th, gazillionth vs. 3rd
baby gorilla, baby aardvark vs. baby cat
tell/guarantee me something vs. explain me something

explain something to me
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)
Is an advantage!
Statistical Preemption:
People learn to avoid certain novel formulations by systematically witnessing a competing alternative.

Bohannon, MacWhinney & Snow 1990
Ambridge et al. (2014)
Boyd & Goldberg (2011, 2015)
Boyd, Ackerman & Kutas (2013)
Brooks & Tomasello (1999)
J. Claus (2014)
Robenalt & Goldberg (2015)
Perek & Goldberg (2017)
Bohannon, MacWhinney & Snow 1990
How do speakers learn what not to say?

The magician made the rabbit disappear!
Verb Frequency Effect

?The magician vanished the rabbit.

> 

??The magician disappeared the rabbit.

Brooks & Tomasello (1999)
Theakston (2004)
Ambridge et al. (2008)
People can be creative. They can extend verbs to novel uses *as long as there isn’t a readily available alternative*

*Make NP vanish* has occurred less often.

The magician disappeared/vanished the rabbit.
Are higher frequency verbs *always* less flexible?

Statistical preemption: only if another formulation competes

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

Please paraphrase:

*The chef coated ranch dressing over the salad.*

: Has-Competing Alternative

Robenalt & Goldberg, 2015, 2016, Tachihara & Goldberg, 2019
The editorial embarrassed the poor man out of town.

: Has no readily available, conventional competing alternative

The editorial embarrassed the man and he left town.

The editorial made the man so embarrassed that he left town.

The embarrassing editorial caused the man to leave town.

The man left town because he was embarrassed.

Rabenalt & Goldberg, 2015, 2016, Tachihara & Goldberg, 2019
Binned sentences via norming study

The editorial *embarrassed* the poor man out of town.
The editorial *mortified* the poor man out of town.

The chef *coated* ranch dressing over the salad.
The chef *doused* ranch dressing over the salad.

Robenalt & Goldberg, 2015, 2016, Tachihara & Goldberg, 2019
No-Competing Alternative (No-CA)

No agreed upon competing alternative phrasing

- 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.

Has-competing Alternative (Has-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.
The chef coated ranch dressing over the salad.  

The magician vanished the rabbit.  

Ashley was terribly mortified.  

......

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>Completely unacceptable</td>
<td>Bad but not terrible</td>
<td>In between</td>
<td>Good but not perfect</td>
<td>Completely acceptable</td>
</tr>
</tbody>
</table>

N=108 from MT
Mixed Linear Model; subjects, items as random effects; length, semantic features of the verbs, plausibility as fixed effects

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

$N=108$ from MT
Mixed Linear Model; subjects, items as random effects; length, semantic features of the verbs, plausibility as fixed effects

N=108 from MT
Mixed Linear Model; subjects, items as random effects; length, semantic features of the verbs, plausibility as fixed effects

$N=108$ from MT
*Mixed Linear Model; subjects, items as random effects; length, semantic features of the verbs, plausibility as fixed effects*

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

People prefer to use familiar formulations.

Frequency of verb *in competing constructions* is relevant to unacceptability.
The perils of ignoring competition
Threshold Proposal (w/o competition)

Yang (2016)

The Core

Past Tense -ed
Plural -s

The Periphery

Brought
Oxen
Threshold Proposal (w/o competition)

Yang (2016)

A construction “goes” productive when:

<table>
<thead>
<tr>
<th>Domain Size</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>
A construction “goes” productive when:

<table>
<thead>
<tr>
<th>Domain Size</th>
<th>Allowable # of exceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>( e = N/\ln N )</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>
Threshold Proposal (w/o competition)
Yang (2016)

A construction “goes” productive when:

<table>
<thead>
<tr>
<th>Domain Size</th>
<th>Minimum # of rule following cases</th>
<th>Allowable # of exceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>$M = N - \frac{N}{\ln N}$</td>
<td>$e = \frac{N}{\ln N}$</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
Threshold Proposal (w/o competition)  
Yang (2016)

A construction “goes” productive when:

<table>
<thead>
<tr>
<th>Domain Size</th>
<th>Minimum # of rule following cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>$M = N - N/\ln N$</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>
### Threshold Proposal (w/o competition)

Yang (2016)

<table>
<thead>
<tr>
<th>N</th>
<th>Minimum # of required rule-following cases: M (N - N/lnN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>50</td>
<td>37</td>
</tr>
<tr>
<td>116</td>
<td>92</td>
</tr>
<tr>
<td>118</td>
<td>93</td>
</tr>
<tr>
<td>200</td>
<td>162</td>
</tr>
<tr>
<td>17,000</td>
<td>12,440</td>
</tr>
</tbody>
</table>
Testing threshold proposal (w/o attention to competition)
Schuler, Yang and Newport (2016)

3M/1-1-1-1-1-1-1e Condition:

<table>
<thead>
<tr>
<th>Domain Size</th>
<th>Minimum # of rule following cases</th>
<th>Allowable # of exceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>$M = N - \frac{N}{\ln N}$</td>
<td>$e = \frac{N}{\ln N}$</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
Testing threshold proposal (w/o attention to competition)
Schuler, Yang and Newport (2016)

5M/1-1-1-1e Condition

<table>
<thead>
<tr>
<th>Domain Size N</th>
<th>Minimum # of rule following cases $M = N - N/\ln N$</th>
<th>Allowable # of exceptions $e = N/\ln N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
% productive use of morpheme (-\textit{ka})

Schuler, Yang Newport (2016)

N=15
Enter: competition
Counterbalanced:
Construction’s function: plural or classifier
Entities: animals or crayons
Condition order
Morpheme labels (po or fep)

32 children, 4;0-6;0 (M = 4;8)
within-subjects design 2 conditions

Hernandez, Floyd, & Goldberg, 2019
Recall:

Testing threshold proposal (w/o attention to competition)
Schuler, Yang and Newport (2016)

5M/1-1-1-1e Condition

<table>
<thead>
<tr>
<th>Domain Size N</th>
<th>Minimum # of rule following cases $M = N - N/\ln N$</th>
<th>Allowable # of exceptions $e = N/\ln N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
New 5M/4E condition

P(M) = .55
P(e) = .45

<table>
<thead>
<tr>
<th>Domain Size</th>
<th>Minimum # of rule following cases</th>
<th>Allowable # of exceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>M = N - N/lnN</td>
<td>e = N/lnN</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

Hernandez, Floyd, & Goldberg, 2019
Recall:

tested $M$ (min # of rule-following cases)
Schuler, Yang, & Newport (2016)

$P(M) = .33$
$P(e_1) = .11$
$P(e_2) = .11$
$P(e_3) = .11$
$P(e_4) = .11$
$P(e_5) = .11$
$P(e_6) = .11$

$M = 3$
$e = 6$

<table>
<thead>
<tr>
<th>Domain Size N</th>
<th>Minimum # of rule following cases $M = N - N/\ln N$</th>
<th>Allowable # of exceptions $e = N/\ln N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
New 3M/0e condition

\[ P(M) = .33 \]

No competitor

\[ M = 3 \]

<table>
<thead>
<tr>
<th>Domain Size</th>
<th>Minimum # of rule following cases</th>
<th>Allowable # of exceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>( M = N - N/\ln N )</td>
<td>( e = N/\ln N )</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

Hernandez, Floyd, & Goldberg, 2019
32 children, 4;0-6;0 ($M = 4;8$) within-subjects design 2 conditions

Counterbalanced:
Construction’s function: plural or classifier
Entities: animals or crayons
Condition order
morpheme label: po or fep
Pretest 1: count domain (N)

1 2 3 4 5 6 7 8 9

Hernandez, Floyd, & Goldberg, 2019
Pretest 2: Label Items

OR

bear
giraffe
elephant
goat
lion
camel
tiger
gorilla
zebra
orange
green
pink
yellow
brown
purple
red
gray
blue
Mr. Chicken
5/4 Condition

Animals & Classifier

P(dax | classifier) = 0.55
P(fep | classifier) = 0.45

Hernandez, Floyd, & Goldberg, 2019
5/4 Condition:
Crayons & Plurals

M = 5

Hernandez, Floyd, & Goldberg, 2019
3/0 Condition
Crayons & Plural

$P(po \mid \text{plural}) = 0.33$

No competitor!

$M = 3$
Recall

Schuler, Yang, and Newport (2016):

\[
P(M) = .55
\]

\[
P(e_1) = .11
\]

\[
P(e_2) = .11
\]

\[
P(e_3) = .11
\]

\[
P(e_4) = .11
\]

\[
P(e_5) = .11
\]

\[
P(e_6) = .11
\]
% productive use of dominant morpheme

$P(M) = .33$
$P(e) = .45$

Hernandez, Floyd, & Goldberg, 2019
Yang (2016) predictions:

\[
P(M) = 0.55
\]
\[
P(e) = 0.45
\]

\[
P(M) = 0.33
\]

% productive use of dominant morpheme

Hernandez, Floyd, & Goldberg, 2019
% productive use of dominant morpheme

P(M) = .33
P(e) = .45

Hernandez, Floyd, & Goldberg, 2019
% productive use of dominant morpheme

$\beta = -8.824, z = -2.918, p = .0035$

Hernandez, Floyd, & Goldberg, 2019
$P(M) = .33$

3M/0e
No competition

$P(M) = .55$

5M/4e
Serious competition

$P(e) = .45$

Hernandez, Floyd, & Goldberg, 2019
Long term memory: vast, associative, content-addressable

**Constructions:**
emergent clusters, conditioned by enormous range of factors

Speakers aim to express their intended message:

Relevant constructions are activated and Combine or COMPETE

Failure to access an ideal construction to express intended message leads to:
grammatical errors (*explain me this*)
overextensions: *dog*
regularization
conservativism

When there *is* no available better alternative:

novel polysemy
creativity
Thank you!