Algorithms for learning the raising/control distinction from semantic information

W. Garrett Mitchener Misha Becker

College of Charleston & UNC Chapel Hill

August 1, 2007

Table of Contents

A Type of Structural Ambiguity

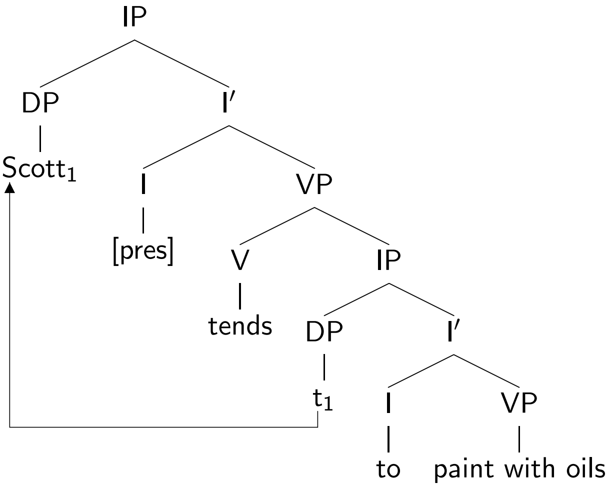
(1) Scott ______ to paint with oils.

a. Raising: Scott tends to paint with oils.

b. Control: Scott likes to paint with oils.

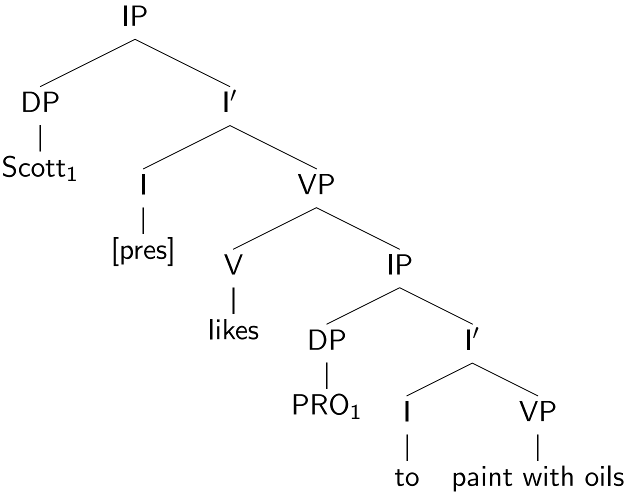
Raising

Scott tends to paint with oils.



Control

Scott likes to paint with oils.

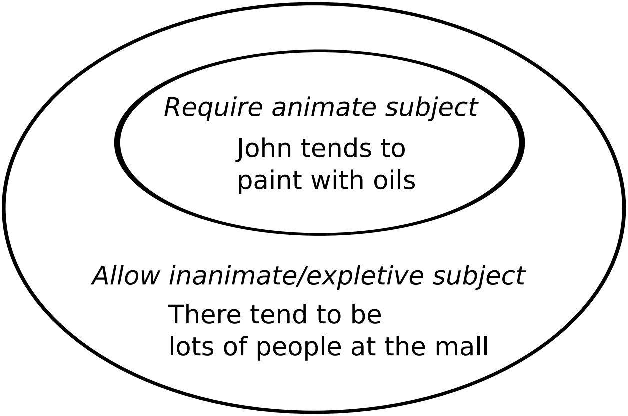


Raising & control comparison

- Raising verbs
  - No semantic relationship to subject
  - Raise subject of lower verb for syntactic subject
  - or use expletive subject (*it*, *there*)
- Control verbs
  - Semantic relationship to subject
  - Subject usually animate
  - Cannot occur with expletive subjects

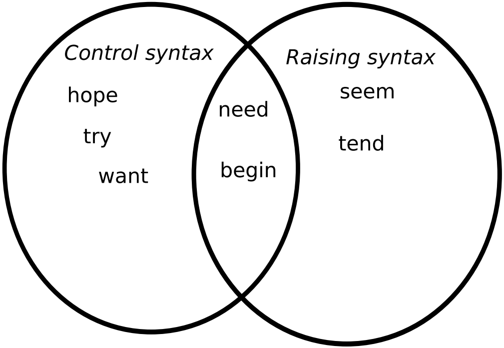
(2) a. It seems to be raining.
b. * It tried to be raining.
c. There tend to be arguments at poker games.
d. * There want to be arguments at poker games.

The subset principle



Assume a more restrictive grammar until you receive positive evidence that forces you to expand it.

The problem with subset-based learning



(3) a. It began to rain. (Raising)
b. Rodney began to beat up Zoe. (Control)

Child grammaticality test

Type	Item	3-year-olds		4-year-olds		5-year-olds	
		OK	silly	OK	silly	OK	silly
control/ compat.	The flower wants to be pink						**
control/ incompat	The flower wants to fly away		*		**		**
raising/ compat.	The hay seems to be on the ground		*		**		*
raising/ incompat	The hay seems to be excited		*		**		**

$*p \leq 0.05$, $**p \leq 0.01$

M. Becker, *There began to be a learnability puzzle*

Potential Semantic Cues from Input

$NP_{subject}$ ______ to Predicate

animate subject, eventive predicate → control verb

inanimate subject, stative predicate → raising verb

Results from adult experiment

Table: Responses by Subject Animacy and Predicate Eventivity

Sentence Type	Control	Raising	Ambig.	Other
Animate+Eventive	65%	5%	15%	15%
(*The driver . . . to hit the car*)				
Animate+Stative	40%	32.5%	20%	7.5%
(*These students . . . to belong*)				
Inanimate+Eventive	32.5%	17.5%	25%	25%
(*The boulder. . . to hit the car*)				
Inanimate+Stative	2.5%	70%	22.5%	5%
(*These shapes . . . to belong*)				

from Becker (2005)

CHILDES data

Verb	A+E	A+S	I+E	I+S
try	86	0	0	0
want	354	53	2	0
need	34	4	0	4
seem	0	4	4	5
gonna	1065	132	31	27

Switchboard data

Verb	A+E	A+S	I+E	I+S
try	149	12	1	0
want	342	123	3	0
need	208	62	3	23
seem	24	57	23	71
gonna	44	11	3	6
tend	36	37	10	9

Goals

- Graded classification of verbs
- Psychologically reasonable
  - Start permissive, become restrictive
  - Parallel processing
  - Simple operations

Raw proportions

For comparison: Fraction of animate subject + eventive predicate

CHILDES		
try	C	1.000
want	C	0.866
gonna	R	0.849
need	A	0.810
seem	R	0.000

Switchboard		
try	C	0.920
want	C	0.731
need	A	0.703
gonna	R	0.688
tend	R	0.391
seem	R	0.137

Switchboard data

Verb	A+E	A+S	I+E	I+S
try	149	12	1	0
want	342	123	3	0
need	208	62	3	23
seem	24	57	23	71
gonna	44	11	3	6
gonna-st	48	11	3	6
tend	36	37	10	9

gonna-st: A sensitivity test, same as *gonna* but with 4 more A+E.

Raw proportions

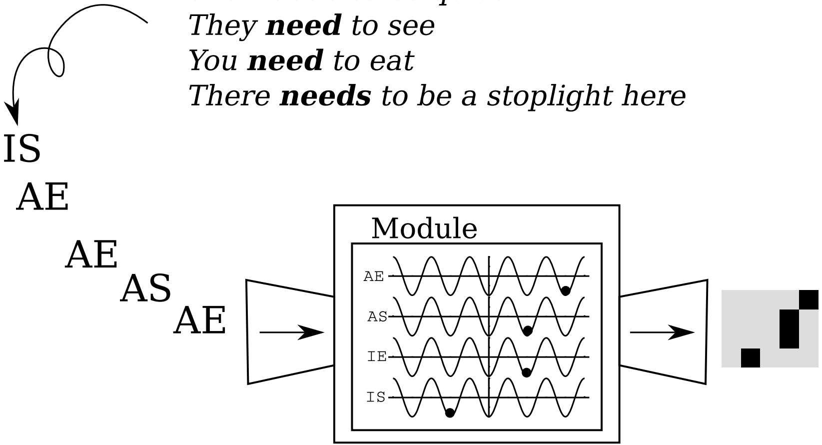
For comparison: Fraction of animate subject + eventive predicate

CHILDES		
try	C	1.000
want	C	0.866
gonna	R	0.849
need	A	0.810
seem	R	0.000

Switchboard		
try	C	0.920
want	C	0.731
gonna-st	R	0.706
need	A	0.703
gonna	R	0.688
tend	R	0.391
seem	R	0.137

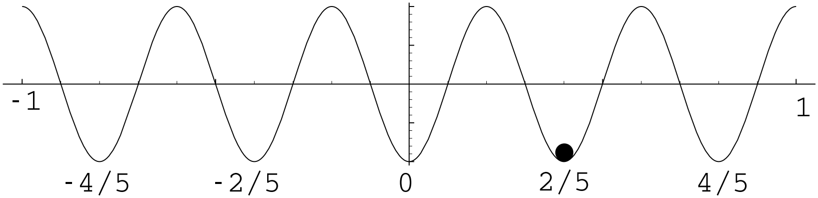
Particle-based learning module

*I **need** to leave*
*She **needs** to be quiet*
*They **need** to see*
*You **need** to eat*
*There **needs** to be a stoplight here*



Basic potential

- One particle for each of the four sentence types

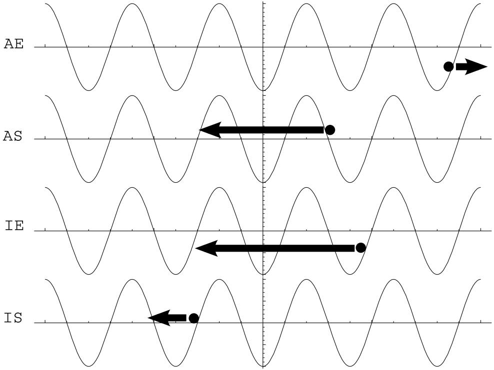


forbidden .... required

$$v = \cos 5\pi x$$

Force from a sentence

$$\mathbf{x}(t) = \begin{pmatrix} x_{\mathrm{AE}}(t) \\ x_{\mathrm{AS}}(t) \\ x_{\mathrm{IE}}(t) \\ x_{\mathrm{IS}}(t) \end{pmatrix}$$



Force from a sentence

If a sentence of type T arrives at time t_0, then it creates a force on the particles with potential

$$L(\mathbf{x}, t, T, t_0) = \begin{cases} 0 & \text{if } t < t_0, \\ \frac{1}{2} e^{-(t-t_0)/\sigma} \left\| \mathbf{x} - \mathbf{p}_T \right\|^2 & \text{if } t \geq t_0. \end{cases}$$

$$\mathbf{p}_{\mathrm{AE}} = \begin{pmatrix} 1 \\ -1/4 \\ -1/4 \\ -1/2 \end{pmatrix} \quad \mathbf{p}_{\mathrm{AS}} = \begin{pmatrix} -1/4 \\ 1 \\ -1/2 \\ -1/4 \end{pmatrix}$$

$$\mathbf{p}_{\mathrm{IE}} = \begin{pmatrix} -1/4 \\ -1/2 \\ 1 \\ -1/4 \end{pmatrix} \quad \mathbf{p}_{\mathrm{IS}} = \begin{pmatrix} -1/2 \\ -1/4 \\ -1/4 \\ 1 \end{pmatrix}$$

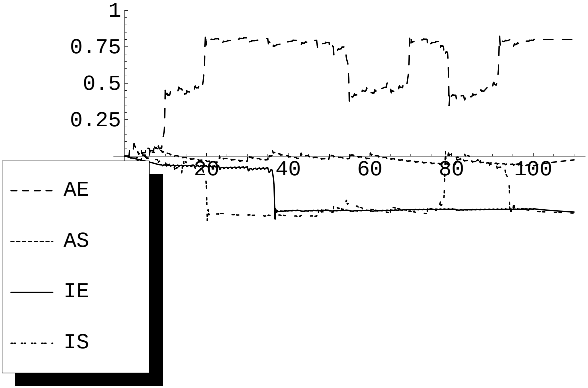
Complete potential

$$\frac{d^2\mathbf{x}(t)}{dt^2} = \mathbf{F}_{\text{field}} + \sum_{\text{inputs}} \mathbf{F}_{\text{input}} + \mathbf{F}_{\text{friction}}$$

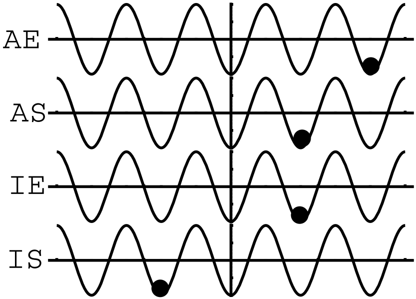
$$= -\operatorname{grad} V(\mathbf{x}(t)) - \gamma \left(\sum_i \operatorname{grad} L(\mathbf{x}(t), t, T_i, t_i) \right) - \beta \frac{d\mathbf{x}(t)}{dt}$$

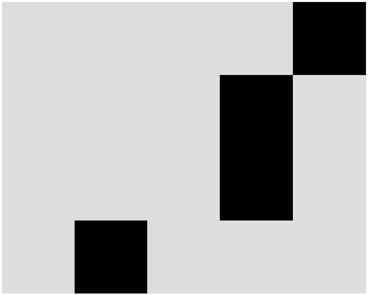
- $\gamma = 8$: relative magnitude of sentence force
- $\sigma = 10$: duration of sentence force
- $\beta = 8$: relative magnitude of drag force
- $t_i = i$: sentences arrive one every time unit

Sample run: learning *need* from CHILDES data

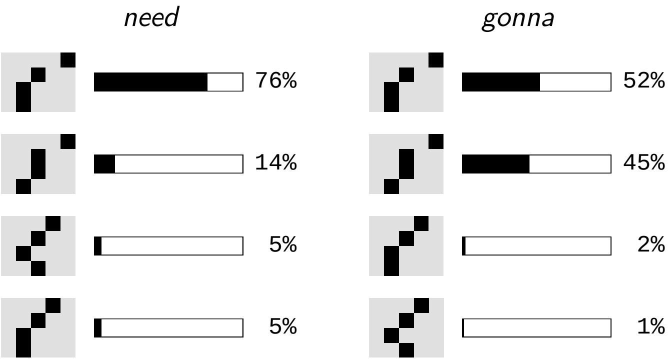


Results of many runs: Histogram key



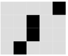
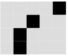


Results from 100 runs



Histograms of the learning algorithm using the sentence type proportions from CHILDES

Summary of most common patterns

	try		*want*		*need*		*gonna*	
	C	SB	C	SB	C	SB	C	SB
A 	100	47	5	0	14	0	45	23
B 	0	53	93	71	76	59	52	21
C 	0	0	1	16	5	27	2	21
other	0	0	1	13	5	14	1	35

Frequent patterns and the percentage of trial runs in which they occur. C means CHILDES, SB means Switchboard.

- *A* and *B* $\Longrightarrow$ Control tendency
- pattern variety $\Longrightarrow$ Raising tendency

Better index

Order by $A + B + \max(A, B, C)$

CHILDES		
try	C	200
want	C	191
$\Downarrow$ *raising*		*178.5*
need	A	166
$\Uparrow$ *control*		*157.5*
gonna	R	149
seem	R	0

Switchboard		
try	C	153
want	C	142
$\Downarrow$ *raising*		*130*
need	A	118
$\Uparrow$ *control*		*105*
gonna-st	R	92
gonna	R	67
tend	R	2
seem	R	0

Conclusion

- On-line learning
- Improved margins and robustness over A+E ratio
- Mimics child learning: permissive $\longrightarrow$ more restrictive
- Biologically reasonable

Implications

The learner assumes:

- isomorphic strings need not = isomorphic structures
- semantic relationships need not be local
- inanimate/expletive subjects are unlikely agents
    - $\rightarrow$ verbs that take inanimate/expletive subjects are unlikely to assign a subject θ-role

These assumptions should be made in general for verb learning (not just this case)

Expletive subjects

CHILDES

	there	it
gonna	1*	0
seem	0	6
others	0	0

Switchboard

	there	it
seem	8	?
used	4	?
gonna	1	?
tend	0	?
appear	0	?
need	16	?
others	0	?

*Sarah103: "I think there's gonna be a war tomorrow when it's time for us to leave"

A Continuum?

unambiguously raising	unambiguously control

tend, used-to, seem, need, begin, want, try, claim

Acknowledgements

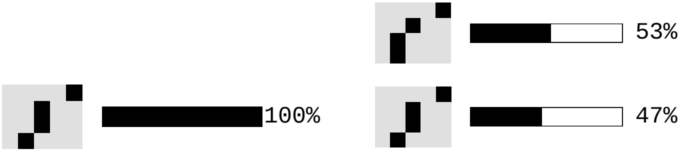
We would like to thank Joan Bresnan and her lab for making the annotated Switchboard corpus available to M. Becker for this research.

Also thanks to Susannah Kirby, Louise Lam and Seema Patidar for coding assistance.

Thanks to Charles Yang for many suggestions.

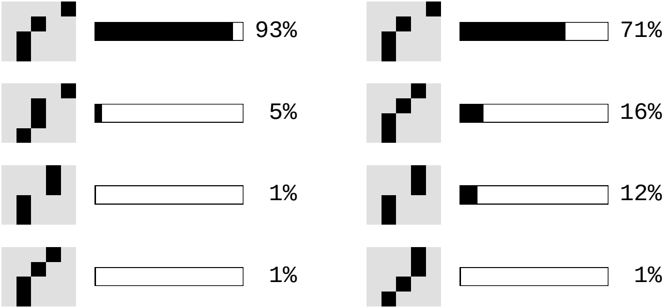
This research was supported by NSF grant 0734783 to Garrett Mitchener and a Junior Faculty Development Grant from UNC (2006) to Misha Becker.

Results from many runs: Learning *try* (control)



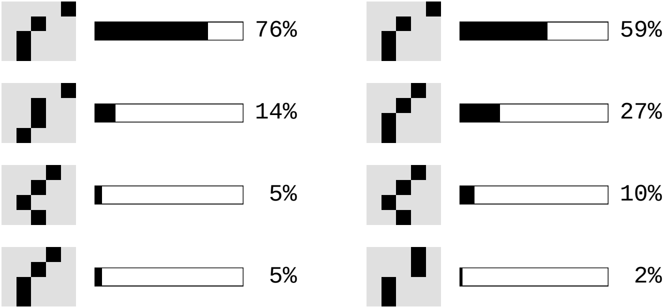
Histograms of the learning algorithm for *try* using the sentence type proportions from CHILDES (left) and Switchboard (right).

Results from many runs: Learning *want* (control)



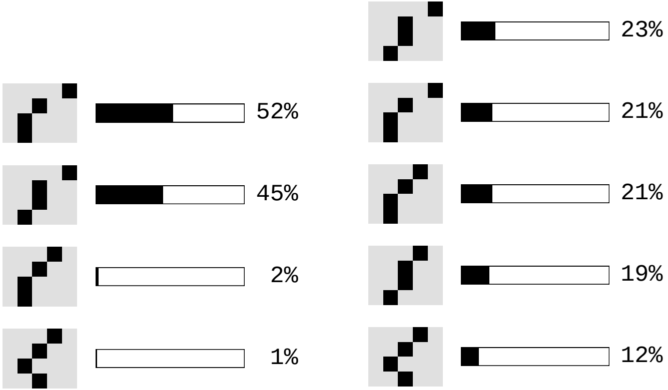
Histograms of the learning algorithm for *want* using the sentence type proportions from CHILDES (left) and Switchboard (right).

Results from many runs: Learning *need* (ambiguous)



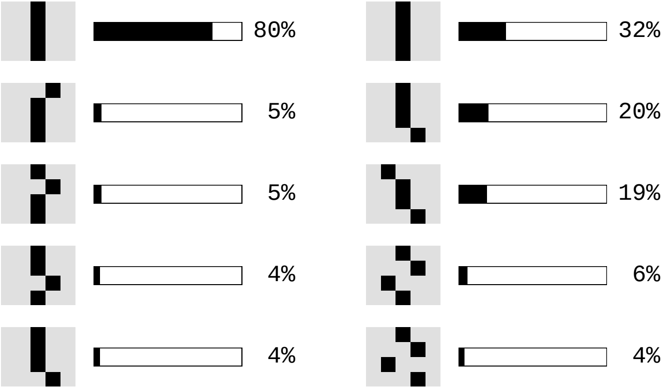
Histograms of the learning algorithm for *need* using the sentence type proportions from CHILDES (left) and Switchboard (right). Only the most common patterns are shown.

Results from many runs: Learning *gonna* (raising)



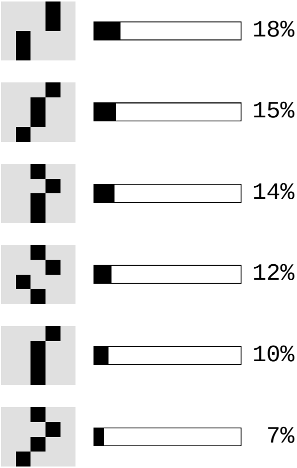
Histograms of the learning algorithm for *gonna* using the sentence type proportions from CHILDES (left) and Switchboard (right). Only the most common patterns are shown.

Results from many runs: Learning *seem* (raising)



Histograms of the learning algorithm for *seem* using equal sentence type proportions (left) and the sentence type proportions from Switchboard (right). Only the five most common patterns are shown.

Results from many runs: Learning *tend* (raising)



Histograms of the learning algorithm for *tend* using the sentence type proportions from Switchboard. Only the six most common