The development of verb classes: A computational adventure with implications for linguistic theory

Lisa Pearl
University of California, Irvine

June 5, 2017: Language Science Colloquium
University of California, Irvine
Today’s plan

Verb classes

done-to
The ice melted.
The penguin climbed.
doer

Computational modeling

Results & implications
Today’s plan

Verb classes

*done-to*
The ice melted.

The penguin climbed.
*doer*
**Verb classes**

Verbs allow a variety of options for where their arguments appear ...

- **try**
  - She tried to melt the ice.
  - *It tried that she melted the ice.*

- **seem**
  - The penguin seemed to climb the hill.
  - It seemed that the penguin climbed the hill.
Verb classes

Verbs allow a variety of options for where their arguments appear and how they’re interpreted.

**try**

She tried to melt the ice.

*It tried that she melted the ice.*

**seem**

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.
**Verb classes**

*Verbs* allow a variety of options for where their arguments appear and how they’re interpreted.

*It tried that she melted the ice.*

*She tried to melt the ice.*

*The penguin seemed to climb the hill.*

*It seemed that the penguin climbed the hill.*

**Verb classes**

- **melt**
  - *doer_{melt}*
  - *done-to_{melt}*
  - She melted the ice.
  - The ice was melted.
  - The ice melted.

- **try**
  - *doer_{try}*
  - She tried to melt the ice.
  - *doer_{melt}*
  - *done-to_{melt}*
  - *seem**
  - *doer_{climb}*
  - The penguin seemed to climb the hill.
  - It seemed that the penguin climbed the hill.
  - The penguin climbed.
  - The hill was climbed.
  - The penguin climbed.
Verb classes

Each verb has **certain linguistic patterns of behavior**, which are shared with other verbs in the same **verb class**.

- **subject-control**
  - She tried to melt the ice.
  - *It tried that she melted the ice.*

- **subject-raising**
  - The penguin seemed to climb the hill.
  - It seemed that the penguin climbed the hill.

- **unaccusative**
  - The ice melted.

- **ungergative**
  - The penguin climbed.

- **done-to**
  - The ice melted.

- **unergative**
  - laugh  dance
Verb classes

How do we tell how a new verb will behave?

The water daxed to blick.

want  need
subject-control

doermelt
She tried to melt the ice.

*It tried that she melted the ice.

doerclimb
The penguin seemed to climb the hill.
It seemed that the penguin climbed the hill.

done-tomelted
The ice melted.

melt
unaccusative
break  fall

appear
subject-raising

doerclimb
The penguin climbed.

climb
unergative

laugh  dance
Verb classes
We can recognize that it belongs to a specific verb class, and use that knowledge to predict its behavior.

The ice melted.

It daxed that the water blicked.

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.

The ice melted.

The penguin climbed.

The water blicked.
Verb classes

This is what we think kids are doing, too.

want  need
subject-control

doer_{melt}
try
She tried to melt the ice.

*It tried that she melted the ice.

doer_{climb}
seem
The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.

done-to_{melted}
The ice melted.

blick  melt
unaccusative

break  fall

done-to_{blicked}
The water blicked.

It daxed that the water blicked.

appear  dax
subject-raising

doer_{climb}

The penguin climbed.

unergative

climb

laugh  dance
Verb classes

Important developmental step:
Grouping verbs into useful **classes**.

melt  
**unaccusative**  
break  fall  
want need  
**subject-control**  
try  
appear  
**subject-raising**  
seem  
climb  
**unergative**  
laugh dance

So how might children do this?

And how can we test different proposals about how they might do this?
Verb classes

The penguin climbed.

Computational modeling

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Subject > Object > Indirect Object

Results & implications
Today’s plan

Computational modeling of language acquisition

A brief overview
Language acquisition = Information processing task

Given the available input ...

The penguin tried to climb.

The ice seemed to melt.
Language acquisition = Information processing task

Given the available input, information processing done by human minds...

The penguin tried to climb.

The ice seemed to melt.
Language acquisition = Information processing task

Given the available input, information processing done by human minds to build a system of linguistic knowledge ...

The penguin tried to climb.

The ice seemed to melt.
Language acquisition = Information processing task

Given the available input, information processing done by human minds to build a system of linguistic knowledge whose output we observe.

The penguin tried to climb. The ice seemed to melt.

The penguin wanted to dance. It appeared that the ice broke.
Language acquisition = Information processing task

To understand how children solve the acquisition task, we need to make explicit the relevant components of the task.

The penguin tried to climb. The ice seemed to melt. It appeared that the ice broke.

The penguin wanted to dance.
Language acquisition = Information processing task

A framework that makes components of the acquisition task more explicit.
A framework that makes components of the acquisition task more explicit.

Distinguishes between things external to the child that we can observe (input signal, child’s behavior) vs. things internal to the child (everything else).
Perceptual encoding:
Turning the input signal into an internal linguistic representation = perceptual intake.
Perceptual encoding:
Involves current grammar
Perceptual encoding:
Involves current grammar being **deployed in real time to parse** the input
Perceptual encoding:
Involves current grammar being deployed in real time to parse the input, often drawing on extralinguistic systems
Generating observable behavior
Involves current linguistic representations being used by production systems.
Doing inference

Generalization happens
Theoretical & computational methods

Doing inference

Generalization happens by using existing learning biases, (some of which may be innate and language-specific)
Doing inference

Generalization happens by using existing learning biases, (some of which may be innate and language-specific) operating over the acquisitional intake — what’s perceived as relevant for acquisition
Doing inference

Generalization happens by using existing learning biases, (some of which may be innate and language-specific) operating over the acquisitional intake — what’s perceived as relevant for acquisition to produce the most up-to-date hypotheses about linguistic knowledge.
The current linguistic hypotheses are used in subsequent perceptual encoding.
This whole process happens over and over again throughout the learning period.
This is language acquisition

An informative computational model of language acquisition captures these important pieces in an empirically-grounded way.
Informative computational models = informative about the learning strategies children use
A successful learning strategy is an existence proof that linguistic knowledge is attainable using the knowledge, learning biases, and capabilities comprising that strategy.
Important learning strategy components include:

- **knowledge** (= theories of **representation**)

---

**Learning strategies children use**

**Perceptual encoding**
- Developing grammar
- Parsing procedures
- Extralinguistic systems (audition, pattern recognition, memory, theory of mind, etc.)

**Production systems**

**Inference engine**
- Acquisitional intake
- Universal grammar

**Perceptual intake** (linguistic representations)
Important learning strategy components include

- theories of representation
- biases & capabilities that must exist for that knowledge to be successfully deployed during acquisition (= theories of the learning process).
Learning strategies children use

theories of representation

theories of the learning process

This is what we want to evaluate with computational modeling.

The penguin tried to climb.

The ice seemed to melt.
Today’s plan

Verb classes

The penguin climbed.

done-to

The ice melted.

doer

Computational modeling

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Subject

Object

Indirect Object

Results & implications

Today’s plan

Verb classes

The penguin climbed.

done-to

The ice melted.

doer
Today’s plan

Computational modeling

Information available and how to use it
The penguin tried to climb.
The ice seemed to melt.

**Syntactic cues**

**Syntactic frame** Shallow “syntactic skeleton” (Gutman et al. 2015)

Children are very adept at using **syntactic bootstrapping** to learn useful generalizations about how verbs behave (e.g., Fisher et al. 2010, Gutman et al. 2015, Harrigan et al. 2016).
The penguin tried to climb.

Syntactic cues
Syntactic frame

NP ____ S_{nonfinite} -surface morphology
NP ____ +past S_{nonfinite} +surface morphology
The penguin tried to climb.

The ice seemed to melt.
Information available

Syntactic frame
NP ___ S_{nonfinite} -surfmorph
NP ___+past S_{nonfinite} +surfmorph

Conceptual cues
Animacy

The penguin tried to climb.
The ice seemed to melt.

It’s useful:
It can distinguish verb classes like raising vs. control verbs, and psych object-experiencer verbs.

The ice seemed to melt.

```
Syntactic frame
NP ___ S_{nonfinite} -surfmorph
NP ___+past S_{nonfinite} +surfmorph
```

```
Conceptual cues
Animacy

Becker 2009, Kirby 2009, Kirby 2010,
Becker 2014, Becker 2015, Hartshorne et al. 2015
```

**Children use it:**
Young children have been shown to use this cue in experimental studies.
The penguin tried to climb.
The ice seemed to melt.

Syntactic frame
NP ___ S_nonfinite -surfmorph
NP ___+past S_nonfinite +surfmorph

Animacy
+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Conceptual cues
Thematic roles

Children could use them:
Thematic roles that indicate event participant roles are salient to very young children.
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax

She melted the ice with a blow dryer.

Subject                Object

Indirect Object

Syntactic frame
NP ___  $_{\text{nonfinite}}$ -surfmorph
NP ___+past $_{\text{nonfinite}}$ +surfmorph

Animacy
+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

How do we get from here to here?

Thematic roles  Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax
She melted the ice with a blow dryer.
Subject Object Indirect Object

Syntactic frame
NP ____ S_{nonfinite} -surfmorph
NP ____+past S_{nonfinite} +surfmorph

Animacy
+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Thematic roles map to one of three categories.

Theories of prior knowledge

Intermediate representations

The Uniformity of Theta Assignment Hypothesis

Thematic roles Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax

She melted the ice with a blow dryer.

Subject
Object
Indirect Object

Syntactic frame
NP ___ S\textsubscript{nonfinite} -surf\textsubscript{morph}
NP ___+past S\textsubscript{nonfinite} +surf\textsubscript{morph}

Animacy
+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Mapping to Syntax

Intermediate representations

These categories map to syntactic positions.

Thematic roles map to one of three categories.

UTAH

Thematic roles
Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax

She melted the ice with a blow dryer.

Subject
Object
Indirect Object

Thematic roles are ordered with respect to each other.

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

The (relativized) UTAH
Larson 1988, Larson 1990

Animacy

+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Syntactic frame
NP ___ S\textsubscript{nonfinite} -surfmorph
NP ___+past S\textsubscript{nonfinite} +surfmorph

Intermediate representations

Theories of prior knowledge

UTAH

Thematic roles
Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax

She melted the ice with a blow dryer.

Subject Object Indirect Object

Syntax frame
NP ___ S_{nonfinite} -surf
NP ___+past S_{nonfinite} +surf

Animacy
+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Mapping to Syntax

UTAH

Whichever ones are present map in order to syntactic positions.

Thematic roles Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax

She melted the ice with a blow dryer.

Subject
Object
Indirect Object

Animacy

+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Syntactic frame

NP ___ S_{\text{nonfinite}} \text{-surfmorph}
NP ___+past S_{\text{nonfinite}} \text{+surfmorph}

UTAH
rUTAH

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Standard UTAH and rUTAH implementations typically assume the mapping is also known a priori

Thematic roles
Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...

Intermediate representations

Theories of prior knowledge

Mapping to Syntax
Information available

**Conceptual cues + Syntactic-semantic knowledge**
**Thematic roles and how to use them**

**Syntax**

She melted the ice with a blow dryer.

**Subject**

**Object**

**Indirect Object**

**But these are separate components**

**Syntactic frame**

NP ____ $S_{\text{nonfinite}}$ -surfmorph

NP ____+past $S_{\text{nonfinite}}$ +surfmorph

**Animacy**

+animate

The penguin tried to climb.

-animate

The ice seemed to melt.

**Intermediate representations**

UTAH

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

**rUTAH**

**Thematic roles**

Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax

She melted the ice with a blow dryer.
Subject Object Indirect Object

Whether children expect a mapping a priori impacts how they perceive the intake for acquisition

Animacy
+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Syntactic frame
NP ___ S_{nonfinite} -surfmorph
NP ___+past S_{nonfinite} +surfmorph

Thematic roles
Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...

Intermediate representations

UTAH

rUTAH

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Theories of prior knowledge

Mapping to Syntax
The ice was melted by the girl.

The thematic roles are Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...

Mapping to Syntax

Theories of prior knowledge

Intermediate representations

Syntax

Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...

UTAH

rUTAH

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

melt: +movement

Salient when mapping doesn’t hold: Interpreted as movement

+ expect a mapping
**Information available**

+ expect a mapping

Salient when mapping doesn’t hold:
Interpreted as movement

**Syntax**

The ice was **melted** by the girl.

**Thematic roles**  
Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...

**UTAH**

**rUTAH**

Agent > Experiencer >  
Theme > Patient >  
(Source, Goal, Instrument)

**Mapping to Syntax**

**Theories of prior knowledge**

**Intermediate representations**

**Mapping to Syntax**

**The ice was melted by the girl.**

**Subject**

**Indirect Object**

**2nd highest**

**Highest**

melt: +movement

*melted* is the highest when mapping holds.

melted is 2nd highest when mapping holds.

melted is highest when mapping holds.

The mapping of *melt* to *movement* is interpreted as movement.
Information available

- expect a mapping

Children track grammatical positions of intermediate representations

Syntax

The ice was melted by the girl.

Subject

Indirect Object

done-to
doer

Mapping to Syntax

Theories of prior knowledge

Intermediate representations

UTAH

rUTAH

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Thematic roles Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

- expect a mapping

Children track grammatical positions of intermediate representations

Syntax

The ice was \textit{melted} by the girl.

\begin{align*}
\text{Subject} & \quad \text{Indirect Object} \\
2\text{nd highest} & \quad \text{Highest}
\end{align*}

Intermediate representations

\textbf{UTAH}

Agent > Experiencer >
Theme > Patient >
(Source, Goal, Instrument)

\textbf{rUTAH}

Thematic roles

Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Conceptual cues + Syntactic-semantic knowledge
Thematic roles and how to use them

Syntax

The ice was melted by the girl.

Subject

Indirect Object

Mapping to Syntax

+expmap

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

-expmap

UTAH

rUTAH

Intermediate representations

Thematic roles  Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Information available

Syntactic frame

NP ___ S_{nonfinite} -surfmorph
NP ___+past S_{nonfinite} +surfmorph

Animacy

+animate
The penguin tried
to climb.

-animate
The ice seemed
to melt.

Thematic roles and how to use them

Agent > Experiencer > rUTAH  +expmap
Theme > Patient > rUTAH  +expmap

Subject  Object  Indirect Object

Subject  Object  Indirect Object

movement?
Potential learning strategies

Animacy

The penguin tried to climb.

The ice seemed to melt.

Syntactic frame

NP ___ S_{nonfinite} -surfmorph
NP ___+past S_{nonfinite} +surfmorph

Thematic roles and how to use them

Agent > Experiencer > rUTAH +expmap
Theme > Patient > (Source, Goal, Instrument)

Subject Object Indirect Object

Movement?
Potential learning strategies

Thematic roles and how to use them

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)
Potential learning strategies

Animacy

-animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Syntactic frame

NP ___ S\_{\text{nonfinite}} -\text{surfmorph}
NP ___ +\text{past} S\_{\text{nonfinite}} +\text{surfmorph}

Thematic roles and how to use them

Choice 1

Agent > Experiencer > r\text{UTAH}
Theme > Patient > (Source, Goal, Instrument)

Choice 2

-\text{expmap}

+\text{expmap}

Subject Object Indirect Object

movement?
Potential learning strategies

Animacy

+animate
The penguin tried to climb.

-animate
The ice seemed to melt.

Syntactic frame

Choice 1

NP ___ $S_{\text{nonfinite}}$ -surfmorph
NP ___+past $S_{\text{nonfinite}}$ +surfmorph

Thematic roles and how to use them

Choice 2

Agent > Experiencer > rUTAH
Theme > Patient > (Source, Goal, Instrument)

Choice 3

Subject Object Indirect Object -expmap
Subject Object Indirect Object +expmap

movement?

3 binary choices = 8 strategies
Potential learning strategies

3 binary choices = 8 strategies

All strategies require learner’s initial state to be sufficient
to extract this information from the input
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy

Syntactic frame
- surfmorph  + surfmorph

Thematic roles and how to use them

UTAH
Agent > Experiencer > Theme > Patient >
(Source, Goal, Instrument)

rUTAH
-expmap
+ expmap

(subject Object Indirect Object)

movement?

"it's falling off"

(from Brown-Eve corpus in CHILDES Treebank)
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy

Syntactic frame
- surfmorph + surfmorph

Thematic roles and how to use them

UTAH
Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

rUTAH - expmap + expmap

Animacy

Subject Object Indirect Object

movement?
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy

Syntactic frame
- surfmorph  + surfmorph

Thematic roles and how to use them

UTAH

rUTAH

Agent > Experiencer >
Theme > Patient >
(Source, Goal, Instrument)

-expmap
+ expmap

Subject  Object  Indirect Object

Subject  Object  Indirect Object

movement?

-FALL

-animate subject: 1

Possible perceptual intake

"it’s falling off"

Input

Subject Theme-V1 animate

Inference engine

Acquisitional intake

Universal grammar
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy

Syntactic frame
+surfmorph

Thematic roles and how to use them
rUTAH
Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)
+expmap

Possible perceptual intake

"it’s falling off"

FALL
-animate subject: 1
Done-to as subject: 1

UTAH
-expmap
-surfmorph
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy

Syntactic frame
+surfmorph

Thematic roles and how to use them

rUTAH
Agent > Experiencer >
Theme > Patient >
(Source, Goal, Instrument)

+expmap

Animacy

Possible perceptual intake

“it’s falling off”

Possible perceptual intake

Input

FALL

-animate subject: 1

Done-to as subject: 1

NP V PRT

UTAH

-expmap

-surfmorph
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy
Syntactic frame
-surfmorph
Thematic roles and how to use them

rUTAH
Agent > Experiencer >
Theme > Patient >
(Source, Goal, Instrument)
+expmap

Input
“it’s falling off”

Possible perceptual intake

Possible
perceptual intake

FALL
-animate subject: 1
Done-to as subject: 1

NP V_{+prog} PRT

UTAH
-expmap
+surfmorph

Inference engine
Acquisitional intake
Universal grammar

NP
PRP
Aux
V1
VBG
PRT

Subject
Object
Indirect Object

Subject
Object
Indirect Object

movement?
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

**Animacy**

**Syntactic frame**
- surfmorph

**Thematic roles and how to use them**

rUTAH
Agent > Experiencer > Theme > Patient >
(Source, Goal, Instrument)

- expmap

Subject Object Indirect Object

Subject Object Indirect Object

movement?

**FALL**
- animate subject: 1
+ movement: 1

Theme is expected to map to object, not subject. Indicator of movement.

NP V_{prog} PRT

UTAH
+expmap
+surfmorph

Possible perceptual intake

“it’s falling off”

“it’s falling off”

Input

Possible perceptual intake

FALL
- animate subject: 1
+ movement: 1

Theme is expected to map to object, not subject. Indicator of movement.

NP V_{prog} PRT

UTAH
+expmap
+surfmorph
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy
Syntactic frame

-surfmorph
Thematic roles and how to use them

-UTAH

Input
"it's falling off"

Possible perceptual intake

Theme is only role so is default highest. Expected mapping is to highest syntactic position (subject).

FALL
-animate subject: 1
+movement: 0

NP V_{prog} PRT

rUTAH
Agent > Experiencer >
Theme > Patient >
(Source, Goal, Instrument)
+expmap
+surf morph
Potential acquisitional intakes

3 binary choices = 8 strategies
Each strategy has a different impact on the acquisitional intake

Animacy

Syntactic frame
- surfmorph

Thematic roles and how to use them

UTAH

Possible perceptual intake

"it's falling off"

Input

FALL
- animate subject: 1

Highest role as subject: 1

NP V_{prog} PRT

rUTAH
Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

+ expmap

+ surfmorph
8 modeled learners and their acquisitional intakes

“it’s falling off”

Possible perceptual intake

FALL
8 modeled learners and their acquisitional intakes.

Input: “it’s falling off”

Possible perceptual intake

Animacy: -animate subject: 1 All 8 learners
8 modeled learners and their acquisitional intakes

Input

“it’s falling off”

Possible perceptual intake

FALL

Animacy

Animate subject: 1

All 8 learners

Syntactic frame

-surfmorph

4 learners

NP V PRT

+surfmorph

4 learners

NP V_{prog} PRT

Subject Theme-V1

Animate
8 modeled learners and their acquisitional intakes

Input

“it’s falling off”

Possible perceptual intake

Animacy

Animate subject: 1

All 8 learners

Syntactic frame

- surfmorph

4 learners

NP V PRT

Intermediate representation

UTAH
Done-to as subject
2 learners

rUTAH
Highest as subject
2 learners

UTAH
Done-to as subject
2 learners

rUTAH
Highest as subject
2 learners

+$surf morph

4 learners

NP V prog PRT
8 modeled learners and their acquisitional intakes

“it’s falling off”

Possible perceptual intake

Animacy
- animate subject: 1

All 8 learners

Syntactic frame
- surfmorph

+ surfmorph

4 learners

NP V PRT

Intermediate representation

Mapping to syntax

UTAH
Done-to as subject

2 learners

- expmap

+ mvmt: 1

1 learner

rUTAH
Highest as subject

2 learners

+ expmap

+ mvmt: 1

1 learner

UTAH
Done-to as subject

2 learners

- expmap

+ mvmt: 0

1 learner

rUTAH
Highest as subject

2 learners

- expmap

+ mvmt: 0

1 learner

UTAH
Highest as subject

2 learners

- expmap

+ mvmt: 1

1 learner

rUTAH
Highest as subject

2 learners

- expmap

+ mvmt: 1

1 learner
Today’s plan

Verb classes

The penguin climbed.

Computational modeling

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Subject Object Indirect Object

Results & implications

The ice melted.
Today’s plan

Computational modeling

How do we model this?
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

“it's falling off”
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

"it’s falling off"

<3 years old

Brown-Eve corpus (Brown 1973) and the Valian corpus (Valian 1991), with syntactic & thematic annotations provided by the CHILDES Treebank (Pearl & Sprouse 2013).

Speech directed at 22 children between 18 and 32 months.

~40,000 utterances (~197,000 word tokens, 555 verbs)

Focus on the 239 verbs occurring 5 or more times.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

“it’s falling off”

CHILDES Treebank

<3yrs
18 and 32 months
~40,000 utterances
239 verbs

<4 years old

<3yrs + Brown-Adam subsection (Brown 1973), with syntactic & thematic annotations provided by the CHILDES Treebank (Pearl & Sprouse 2013).

Speech directed at 23 children between 18 and 48 months.

~51,000 utterances (~254,000 word tokens, 617 verbs)

Focus on the 267 verbs occurring 5 or more times.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

"it’s falling off"

CHILDES Treebank

<3yrs

18 and 32 months
~40,000 utterances
239 verbs

<4yrs

18 and 48 months
~51,000 utterances
267 verbs

<5 years old

<4yrs + Brown-Adam subsection (Brown 1973), with syntactic & thematic annotations provided by the CHILDES Treebank (Pearl & Sprouse 2013).

Speech directed at 23 children between 18 and 58 months.

~56,500 utterances (~285,000 word tokens, 651 verbs)

Focus on the 284 verbs occurring 5 or more times.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old.

"it’s falling off"

CHILDES Treebank

<3yrs
18 and 32 months
~40,000 utterances
239 verbs

<4yrs
18 and 48 months
~51,000 utterances
267 verbs

<5yrs
18 and 58 months
~56,500 utterances
284 verbs
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

Basic question: Is it possible for the child to use the **acquisitional intake** to achieve the **target knowledge/behavior**?
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Basic question: Is it possible for the child to use the **acquisitional intake** to achieve the **target knowledge/behavior**?

This is the goal of **learnability approaches**

*(computational-level of analysis: Marr 1982)*

Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Basic question: Is it possible for the child to use the **acquisitional intake** to achieve the **target knowledge/behavior**?

**Ideal learner model:** not concerned with the cognitive limitations and incremental learning restrictions children have.

Concerned with what **assumptions** are **useful** for children to have.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Basic question: Is it possible for the child to use the acquisitional intake to achieve the target knowledge/behavior?

Ideal learner model: Also an excellent first step to see if this is the right conceptualization of the acquisition task.
Learners use a generative model of how the observable data for each verb are created.

Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs
Learners use a generative model of how the observable data for each verb are created.

Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

FALL
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

Each verb appears in a certain number of instances in the input.

“it’s falling off”
“she fell down”
“don’t fall!”
“is London Bridge falling down?”
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

Each instance is observed some number of times.

“it’s falling off” (3x) “it’s falling off” “it’s falling off”
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Each verb belongs to some class which determines its linguistic behavior.

The learner doesn’t know beforehand how many classes there are or which verbs belong to which.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old.

Each verb belongs to some class which determines its linguistic behavior.

Objective: Infer verb class

"it’s falling off" (3x)  "it’s falling off"  "it’s falling off"

...but has a bias for fewer classes.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old.

<3yrs <4yrs <5yrs

“it’s falling off” (3x) “it’s falling off” “it’s falling off”

Depending on the verb class, the observed usage will have certain characteristics.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

These characteristics include binary choices such as whether the subject is animate or not.

Each class has a probability of preferring each option.

- **unaccusatives**
  - +anim Subject: 0.3
  - -anim: 0.7

"it’s falling off"

(3x) “it’s falling off”

“it’s falling off”

FALL
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

Binary choices:

+/-animate subject
+/-animate object
+/-animate indirect object
+/-movement (when +exp-mapping)

+animate  -animate

Subject  Object  Indirect Object

movement?

unaccusatives

+anim  Subject  -anim

0.3  0.7

-it’s falling off

(3x)  “it’s falling off”  “it’s falling off”
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

The learner infers these probabilities, and begins with no bias towards either option per class.

Unaccusatives

-anim +anim
Subject 0.3 0.7

“it’s falling off”

(3x)  “it’s falling off”
“it’s falling off”

Binary choices:

+animate -animate

Subject Object Indirect Object
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

These characteristics include multinomial choices such as which syntactic frame a verb appears in.

Each class has a probability of preferring each option.

\[
\begin{align*}
NP & \ V \ PRT \ & 0.3 \\
NP & \ V \ & 0.25 \\
\ldots \\
NP & \ V \ S \ & 0
\end{align*}
\]
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3 yrs  <4 yrs  <5 yrs

“it’s falling off”

Multinominal choices:
which syntactic frame is used $NP \ V \ PRT$
(if -exp-mapping)

position of doer/Highest role
position of done-to/next-highest role
position of done-by/third-highest role

Agent > Experiencer >
Theme > Patient >
(Source, Goal, Instrument)

Subject    Object    Indirect Object

$\alpha_{\psi}$

$\theta_{\psi_{c_j}}$

$\pi_{\phi_{c_j}}$

$\beta_{\phi_{0}}$

$\beta_{\phi_{1}}$

$\gamma_{c}$

$\theta_{c}$

$\gamma_{c}$

$\gamma_{c}$

NP V PRT 0.25 NP V S 0

unaccusatives

+anim Subject -anim

0.3 0.7

FALL

“it’s falling off”

(3x) $NP \ V \ PRT$

“it’s falling off”

“it’s falling off”
Inference: The learner forms different classes because the characteristics are sufficiently different for each class.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

Using the observed instances of verb usage, Bayesian inference can be used to determine ...

NP V  PRT  0.25
NP V S  0

unaccusatives

+anim  Subject  -anim
0.3  0.7

NP V  PRT
“it’s falling off”
(3x)

“it’s falling off”

“it’s falling off”

FALL
Using the observed instances of verb usage, Bayesian inference can be used to determine
• how many classes there are
Using the observed instances of verb usage, Bayesian inference can be used to determine:

- how many classes there are
- which class each verb belongs to
Using the observed instances of verb usage, Bayesian inference can be used to determine
- how many classes there are
- which class each verb belongs to
- what the characteristics are of each class
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

Using the observed instances of verb usage, Bayesian inference can be used to determine

- how many classes there are
- which class each verb belongs to
- what the characteristics are of each class

Best answer: maximizes the probability of the observed data.
Using the observed instances of verb usage, Bayesian inference can be used to determine
• how many classes there are
• which class each verb belongs to
• what the characteristics are of each class

\[ p_{c_j} = P(c_j|c_{-j}, \gamma_c, F_{-j}, \lambda) = p_{cat_j} \times p_{binary_{c_j}} \times p_{multinomial_{c_j}} \]

+ Gibbs sampling (method guaranteed to find optimal answer, given sufficient time to search the hypothesis space)
Goal: Determine if the information provided in the acquisitional intake is sufficient to identify verb classes this way.

Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<table>
<thead>
<tr>
<th>&lt;3yrs</th>
<th>&lt;4yrs</th>
<th>&lt;5yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>FALL</td>
<td>FALL</td>
<td>FALL</td>
</tr>
</tbody>
</table>

“it’s falling off”

NP V PRT 0.25

NP V S 0

unaccusatives

+anim  Subject -anim 0.3 0.7

-surfmorph +surf morph

+expmap -expmap

UTAH rUTAH
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  <4yrs  <5yrs

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs

Yields 12 verb behaviors
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“It was ___-en.”

done-to
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“It was ___-en.”

(done-to)

3yrs

+= hit, see, ...

-= know, remember, ...

verb classes
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“It was ___-en.”

done-to

4yrs

+= hit, scare, see, ...
-= know, love, remember, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“*It was ___-en.*”

**done-to**

5yrs

+= hit, love, scare, see, …

-= know, remember, …
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“Jack ___ Lily the thing.”
Survey of 32 experimental studies on children’s production and comprehension of specific verbs


"Jack ___ Lily the thing."

3yrs

+= give, read, *say, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“Jack ___ Lily the thing.”

4 yrs

+= give, read, *say, teach, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

**Survey of 32 experimental studies on children’s production and comprehension of specific verbs**


“Jack ___ Lily the thing.”

**5yrs**

+= ask, give, read, *say, teach, …
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“`It ___.”
done-to
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“It ___.”
done-to

3yrs
+= break, fall, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs

Control object & Raising object: Kirby 2009a, Kirby 2009b, Kirby 2010, Becker 2014

Control object

“I ___him to leave.”
done-recipient (main)
doer (embedded)
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs

Control object & Raising object: Kirby 2009a, Kirby 2009b, Kirby 2010, Becker 2014

Control object
“I ___ him to leave.”
done-recipient (main)
doer (embedded)

4yrs 5yrs
+= ask, tell
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs

Control object & Raising object: Kirby 2009a, Kirby 2009b, Kirby 2010, Becker 2014

Raising object

“I ___ him to leave.”
doer (embedded)

4yrs

5yrs

+= need, want
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


Control subject
“i ___ to leave.”
\[ \text{doer (main)} \]
\[ \text{doer (embedded)} \]
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


Control subject
“I ___ to leave.”
doer (main)
doer (embedded)

4yrs
5yrs
+= try, want
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


Raising subject

“I ___ to leave.”

doer (embedded)

4yrs 5yrs

+= seem
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs

Subject-experiencer and Object-experiencer psych verbs: Hartshorne et al. 2015

Subject-experiencer
“Jack ___ Lily.”
Experiencer
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs

Subject-experiencer and Object-experiencer psych verbs: Hartshorne et al. 2015

Subject-experiencer
“Jack ___ Lily.”
Experiencer

4yrs
5yrs

+= like, love
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs

Subject-experiencer and Object-experiencer psych verbs: Hartshorne et al. 2015

Object-experiencer
“Jack ___ Lily.”

Experiencer

4yrs 5yrs

+= frighten, scare, surprise
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


Non-finite to
“Jack ___ to go.”
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


Non-finite to
“Jack ___ to go.”

3yrs
+= get, start, suppose, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“Jack ___ that he can go.”
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


that
“Jack ___ that he can go.”

3 yrs
+= hope, know, say, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


*that*

“Jack ___ that he can go.”

5yrs

+= guess, hope, know, pretend, say, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


“Jack ___ whether/if he can go.”
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

verb classes

Survey of 32 experimental studies on children’s production and comprehension of specific verbs


whether/if

“Jack ___ whether/if he can go.”

5yrs

+= ask, care, know, see, ...
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

These verb behaviors yield a number of verb classes at each age
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

These verb behaviors yield a number of verb classes at each age

Example classes

<3yrs

[+passive]: carry, chase, crash, drop, eat, hit, hold, hurt, jump, kick, kiss, knock, lick, punch, push, scratch, shake, turn, wash, watch

[-passive]: believe, remember

[+non-finite to]: ask, have, need, start, suppose, teach, try, use, want

[+that-comp]: bet, hope, think, wish

[+passive, +non-finite to]: like

[+passive, +that-comp]: see
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

These verb behaviors yield a number of verb classes at each age

Example classes

<4yrs

[+passive]: bite, bump, carry, chase, crash, drop, find, hit, hold, hurt, jump, kick, kill, kiss, knock, lick, pull, punch, push, ride, scratch, shake, shoot, turn, wash, watch

[-passive]: believe, remember

[+that-comp]: bet, hope, think, wish

[+non-finite to, +raising-obj]: need

[+non-finite to, +raising-obj, +control-subj]: want

[+passive, +non-finite to, +psych-subj]: like

[+passive, +that-comp]: see
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

These verb behaviors yield a number of verb classes at each age

Example classes

[+passive]: bite, bump, carry, chase, crash, drop, find, hit, hold, hurt, jump, kick, kill, kiss, knock, lick pull, push, ride, scratch, shake, shoot, turn, wash, watch

[-passive]: believe, remember

[+that-comp]: bet, dream, guess, hope, lie, pretend, think, wish

[+non-finite to, +raising-obj]: need

[+non-finite to, +raising-obj, +control-subj]: want

[+passive, +non-finite to, +psych-subj]: like

[+passive, +that-comp, +whether/if-comp]: see
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

These verb behaviors yield a number of verb classes at each age:

- **<3yrs**
  - 15 classes of 60 verbs total

- **<4yrs**
  - 23 classes of 76 verbs total

- **<5yrs**
  - 24 classes of 82 verbs total
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Evaluation:
How well did the modeled learner do at finding these verb classes?
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Implementation: Random Index

Intuition: Get credit for putting things together that belong together and keeping things apart that don’t belong together.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

Implementation:
Random Index

\[ 0.0 \leq RI \leq 1.0 \]

For each pair of verbs in the inferred classes:

<table>
<thead>
<tr>
<th>True</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same class</td>
<td>Same class</td>
</tr>
<tr>
<td>Different class</td>
<td>Different class</td>
</tr>
</tbody>
</table>

Intuition: Get credit for putting things together that belong together and keeping things apart that don’t belong together.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 yrs</td>
<td>15</td>
</tr>
<tr>
<td>&lt;4 yrs</td>
<td>23</td>
</tr>
<tr>
<td>&lt;5 yrs</td>
<td>24</td>
</tr>
</tbody>
</table>

Implementation: Random Index

0.0 <= RI <= 1.0

For each pair of verbs in the inferred classes:

<table>
<thead>
<tr>
<th>Inferred Class</th>
<th>True</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same class</td>
<td>True Positive</td>
<td>False Positive</td>
</tr>
<tr>
<td>Different class</td>
<td>False Positive</td>
<td>True Negative</td>
</tr>
</tbody>
</table>

Intuition: Get credit for putting things together that belong together and keeping things apart that don’t belong together.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

15 classes
23 classes
24 classes

Implementation: Random Index
0.0 \leq RI \leq 1.0

Intuition: Get credit for putting things together that belong together and keeping things apart that don’t belong together.

But how do we know we’re doing better than chance?
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs 15 classes
<4yrs 23 classes
<5yrs 24 classes

Implementation: Adjusted Random Index

-1.0 \leq ARI \leq 1.0

Compared against the expected value of the Random Index:

1.0 = perfect classification
>0 = better than chance
0 = chance performance
<0 = worse than chance
-1.0 = perfectly awful classification
Verb classes

The penguin climbed.

Computational modeling

Object

Subject

Indirect Object

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Results & implications

The ice melted.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

-1.0 <= ARI <= 1.0

<3yrs: 15 classes
<4yrs: 23 classes
<5yrs: 24 classes
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs 15 classes
<4yrs 23 classes
<5yrs 24 classes

-1.0 <= ARI <= 1.0

Animacy

+animate
The penguin tried to climb.

-animate
The ice seemed to melt.
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

<3yrs  
15 classes

<4yrs  
23 classes

<5yrs  
24 classes

Animacy

Syntactic frame

The ice seemed to melt.

NP ___ S_{nonfinite} -surfmorph
NP ___+past S_{nonfinite} +surfmorph
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

-1.0 <= ARI <= 1.0

Animacy
Syntactic frame

+surfmorph

-surfmorph

Thematic roles and how to use them

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

UTAH

rUTAH

Subject
Object
Indirect Object

movement?
Goal: Model the developmental trajectory from 3 to 4 to 5 years old

-1.0 ≤ ARI ≤ 1.0

**Classification**

<3yrs 15 classes
<4yrs 23 classes
<5yrs 24 classes

**Thematic roles and how to use them**

<table>
<thead>
<tr>
<th>+surfmorph</th>
<th>-surfmorph</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTAH +expmap</td>
<td>UTAH -expmap</td>
</tr>
<tr>
<td>rUTAH +expmap</td>
<td>rUTAH -expmap</td>
</tr>
</tbody>
</table>

**Animacy**

<table>
<thead>
<tr>
<th>+expmap</th>
<th>-expmap</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANIMACY</td>
<td>ANIMACY</td>
</tr>
</tbody>
</table>

**Syntactic frame**

<table>
<thead>
<tr>
<th>+expmap</th>
<th>-expmap</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTAH</td>
<td>UTAH</td>
</tr>
<tr>
<td>rUTAH</td>
<td>rUTAH</td>
</tr>
</tbody>
</table>
Animacy
Syntactic frame

-1.0 <= ARI <= 1.0

Thematic roles and how to use them

+surfmorph

UTAH
+expmap
-expmap

rUTAH
+expmap
-expmap

UTAH
+expmap
-expmap

rUTAH
+expmap
-expmap

3yrs

15 classes

<4yrs
23 classes

<5yrs
24 classes
All learning strategies are doing better than chance...
But some are clearly doing better than others.
All learning strategies are doing better than chance...
But one is clearly doing better than others.
All learning strategies are doing better than chance…
But two are clearly doing better than others.
Animacy

Syntactic frame

-1.0 <= ARI <= 1.0

Thematic roles and how to use them

What does this mean?
Using animacy, syntactic frame, and thematic role information is a pretty good match for what children seem to be doing when creating verb classes.
Before 3, children ignore verb morphology but may be using any of the options for thematic roles.

The ice seemed to melt.

NP ___  S_{nonfinite} -surf\text{morph}
By 4, children heed verb morphology, and how to use them are using the UTAH intermediate representation, and don’t expect a mapping a priori.
By 5, children still heed verb morphology, but now may be using either the UTAH or rUTAH representation and expect a mapping.

4yrs

UtAH
+expmap

-utfmorp

5yrs

UtAH
+expmap

rUtAH
+expmap

-utfmorp

-utfmorp

-utfmorp

Thematic roles and how to use them

Animacy
Syntactic frame

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Subject
Object
Indirect Object

movement?
The ice seemed to melt.

NP ____ $S_{\text{nonfinite}}$

3yrs

-animate

4yrs

+surfmorph

UTAH -expmap

5yrs

+surfmorph

UTAH rUTAH +expmap
The ice seemed to melt.

NP ___+past S\text{nonfinite}

**3yrs**

- surfmorph

**4yrs**

+ surfmorph

UTAH

- expmap

**5yrs**

+ surfmorph

UTAH rUTAH

+ expmap
The ice seemed to melt.

Subject: animate

NP: +movement

past S
nonfinite

Theme: -movement

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)
This is the first articulation of the trajectory of learning assumptions children may have that causes them to group verbs into useful classes the way we observe.
The ice seemed to melt.

It suggests there are different timelines for
- ignoring vs. heeding surface morphology on verbs ...
It suggests there are different timelines for
- ignoring vs. heeding surface morphology on verbs
- a simpler vs. more flexible intermediate thematic representation...
It suggests there are different timelines for
- ignoring vs. heeding surface morphology on verbs
- a simpler vs. more flexible intermediate thematic representation
- not expecting vs. expecting a mapping between that intermediate thematic representation and syntactic positions
What does this mean for linguistic theory?
What seems to develop earlier (perhaps because it’s easy to derive from existing biases):

-surfmorph: Preference to ignore surface morphology (perhaps due to processing limitations)
What seems to develop somewhat earlier (perhaps because it’s easy to derive from existing biases):

UTAH: Simpler thematic representation
What seems to develop somewhat earlier (perhaps because it’s easy to derive from existing biases):

-expmap: No prior expectation about how to map — learn this from the intake
What seems to develop later (perhaps building on prior knowledge):

rUTAH: more sophisticated thematic representation

Subject   Object   Indirect Object

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)
What seems to develop later (perhaps building on prior knowledge and the intake):

+expmap: an expectation for a mapping between that representation and grammatical positions
Bigger theoretical takeaway:

Everyone’s right about the representation at some stage of development.
So now what?
So now what?

3yrs
- surfmorph

4yrs
+ surfmorph
UTAH
- expmap

5yrs
+ surfmorph
UTAH
rUTAH
+ expmap

(1) A broader assessment of children’s verb class knowledge
(1) A broader assessment of children’s verb class knowledge

We need more observable behavior for verbs in children’s input to match modeling results against.

<3yrs
239 verbs
15 classes of 60 verbs

<4yrs
267 verbs
23 classes of 76 verbs

<5yrs
284 verbs
24 classes of 82 verbs

Input
Children’s behavior
(1) A broader assessment of children’s verb class knowledge
This will further test these theoretical proposals, and validate (or not) the current findings.

So now what?

3yrs
- surfmorph

4yrs
+ surfmorph
UTAH
- expmap

5yrs
+ surfmorph
UTAH rUTAH
+ expmap

<3yrs
239 verbs
15 classes
of 60 verbs

<4yrs
267 verbs
23 classes
of 76 verbs

<5yrs
284 verbs
24 classes
of 82 verbs

Input
Children’s behavior
So now what?

3yrs
-surfmorph

4yrs
+surfmorph
-UTAH
-expmap

5yrs
+surfmorph
-UTAH rUTAH
+expmap

(1) A broader assessment of children’s verb class knowledge

(a) More verbs

<3yrs

239 verbs
15 classes of 60 verbs

<4yrs

267 verbs
23 classes of 76 verbs

<5yrs

284 verbs
24 classes of 82 verbs

Input
Children’s behavior
(1) A broader assessment of children’s verb class knowledge

(a) More verbs

<3yrs
- 239 verbs
- 15 classes of 60 verbs

<4yrs
- 267 verbs
- 23 classes of 76 verbs

<5yrs
- 284 verbs
- 24 classes of 82 verbs

(b) More behaviors

So now what?

- 3yrs -surfmorph
- 4yrs +surfmorph
- 5yrs +surfmorph

UTAH -expmap
rUTAH +expmap

Input

Children’s behavior
So now what?

**1.** A broader assessment of children’s verb class knowledge

**2.** Models incorporating more cognitively plausible assumptions

This may give us an even better match to children’s observed verb class behavior.
So now what?

(1) A broader assessment of children’s verb class knowledge

(2) Models incorporating more cognitively plausible assumptions

What happens when we embed these theories in a learning model that learns incrementally and has age-appropriate memory & processing limitations?
So now what?

(1) A broader assessment of children’s verb class knowledge

What kinds of child behavior does the model predict in the experimental scenarios already available, based on its internal representations?

(2) Models incorporating more cognitively plausible assumptions
So now what?

(1) A broader assessment of children’s verb class knowledge

(2) Models incorporating more cognitively plausible assumptions

What other types of information may be available, especially throughout development as children learn from their intake?
So now what?

(1) A broader assessment of children’s verb class knowledge

(2) Models incorporating more cognitively plausible assumptions

(3) Other theories of representation

Are there other options for linking thematic role information to syntactic structure that we can explore in this framework?
So now what?

(1) A broader assessment of children’s verb class knowledge

(2) Models incorporating more cognitively plausible assumptions

(3) Other theories of representation

Subject Object Indirect Object

3yrs -surfsmorph
4yrs +surfsmorph UTAH exmap
5yrs +surfsmorph UTAH rUTAH exmap
What we saw today

Verb classes

done-to
The ice melted.
The penguin climbed.
doer

Computational modeling

Agent > Experiencer > Theme > Patient > (Source, Goal, Instrument)

Results & implications

What we saw today

The ice melted.
What we saw today

Verb classes: An example of complex linguistic knowledge that children develop, involving several theoretical options for the representations they may be using.

The penguin climbed.

Computational modeling

Results & implications
**What we saw today**

**Verb classes:** complex linguistic knowledge involving several theoretical options for representations

Computational modeling: A way to explicitly test these theories by implementing them concretely in an empirically grounded model of the acquisition process.

**Results & implications**
What we saw today

**Verb classes:** complex linguistic knowledge involving several theoretical options for representations

**Computational modeling:** explicitly test these theories

**Results & implications:**
Articulating the trajectory of representations and learning assumptions children have at different stages of development
Verb classes: complex linguistic knowledge involving several theoretical options for representations

Computational modeling: explicitly test these theories

Results & implications: Articulating the representational trajectory over development

This approach allows us to connect theories of linguistic representation and theories of language acquisition to understand more about both.
Thank you!

Jon Sprouse

SynLinks workshop 2016
McGill University
Linguistics 2016

This work was supported in part by NSF grant BCS-1347028.

Special thanks to Abbie Thornton, Alandi Bates, Emily Yang, and BreAnna Silva for CHILDES Treebank corpus annotation.
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis


Each thematic role maps to a specific syntactic position (grammatical role).

She tried to melt the ice.

*It tried that she melted the ice.

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis


Each thematic role maps to a specific syntactic position (grammatical role).

Agent-like = grammatical subject

<table>
<thead>
<tr>
<th>Agent</th>
<th>Causer</th>
<th>Experiencer</th>
<th>Possessor</th>
</tr>
</thead>
</table>

("internal cause“ = Rappaport-Hovav 1995)

She tried to melt the ice.

*It tried that she melted the ice.*

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis


Each thematic role maps to a specific syntactic position (grammatical role).

Agent-like = grammatical subject

Agent
Causer
Experiencer (*Baker: only when subject)
Possessor

(“internal cause” = Rappaport-Hovav 1995)

She fears spiders.
Experiencer

Spiders frighten her.
Experiencer

She tried to melt the ice.
doer
done-to

*It tried that she melted the ice.
doer

doer
done-to

The penguin seemed to climb the hill.
doer
done-to

It seemed that the penguin climbed the hill.
doer
done-to

control

UG knowledge

(r)UTAH: U of T Theta Assignment Hypothesis


Each thematic role maps to a specific syntactic position (grammatical role).

Agent-like = grammatical subject

Agent
Causer
Experiencer (*Baker: only when subject)
Possessor

(“internal cause” = Rappaport-Hovav 1995)
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis


Each thematic role maps to a specific syntactic position (grammatical role).

Agent-like = grammatical subject
Patient-like = grammatical object

Patient
Theme
Experiencer
Subject Matter

(“external cause”)

She tried to melt the ice.

*It tried that she melted the ice.

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis

Each thematic role maps to a specific syntactic position (grammatical role).

Agent-like = grammatical subject
Patient-like = grammatical object

Patient
Theme
Experiencer (*Baker: only when not subject)
Subject Matter

("external cause")

She fears spiders.
Experiencer
Spiders frighten her.
Experiencer

She tried to melt the ice.
doer
done-to
*It tried that she melted the ice.
doer
done-to

The penguin seemed to climb the hill.
doer
done-to
It seemed that the penguin climbed the hill.
doer
done-to

She fears spiders.
Experiencer
Spiders frighten her.
Experiencer

control

UG knowledge
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis


Each thematic role maps to a specific syntactic position (grammatical role).

**Agent-like** = grammatical subject

**Patient-like** = grammatical object

**Goal-like** = grammatical indirect object

Location
Source
Goal
Benefactor
Instrument

---

She tried to melt the ice with a blow dryer.

*It tried that she melted the ice with a blow dryer.*

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.
Thematic roles & how to use them

**Syntax**

She **melted** the **ice** with a **blow dryer**.

**Subject** **Object** **Indirect Object**

**Mapping to Syntax**

The **Uniformity of Theta Assignment Hypothesis**:

**Intermediate representations**

**UTAH**

Thematic roles map to one of three categories.

(likely derived from lower level conceptual info) = Agent, Experiencer, Patient, Theme, Goal, Source, Instrument...
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) **Uniformity of Theta Assignment Hypothesis**

**rUTAH**: Larson 1988, Larson 1990

Thematic roles are ordered relative to each other, with the highest thematic role mapping to the highest grammatical role (subject > object > indirect object).

Control

She tried to melt the ice with a blow dryer.

*It tried that she melted the ice with a blow dryer.

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis

\[ \text{rUTAH: Larson 1988, Larson 1990} \]

Thematic roles are ordered relative to each other, with the highest thematic role mapping to the highest grammatical role (subject > object > indirect object).

Basic intuition:

- \text{doer (Agent-like)}
- \text{done-to (Patient-like)}
- \text{done-for/with (Goal-like)}

\[ \begin{align*}
\text{control} \\
\text{She tried to melt the ice with a blow dryer.} \\
& \text{doer} \quad \text{done-to} \quad \text{done-with} \\
& \text{She tried that she melted the ice with a blow dryer.} \\
& \text{doer} \quad \text{done-to} \quad \text{done-with}
\end{align*} \]

\[ \begin{align*}
\text{The penguin seemed to climb the hill.} \\
& \text{doer} \quad \text{done-to} \\
& \text{It seemed that the penguin climbed the hill.} \\
& \text{doer} \quad \text{done-to}
\end{align*} \]
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis

**rUTAH**: Larson 1988, Larson 1990

Thematic roles are ordered relative to each other, with the highest thematic role mapping to the highest grammatical role (subject > object > indirect object).

Basic intuition:

- **doer** (Agent-like) >
  - **done-to** (Patient-like) >
    - **done-for/with** (Goal-like)

An example implementation:

- **Agent** > **Causer** > **Experiencer** > **Possessor** >
  - Subject Matter > **Causee** > **Theme** > **Patient** >
    - (Location, Source, Goal, Benefactor, Instrument)

**control**

She tried to melt the ice with a blow dryer.

*It tried that she melted the ice with a blow dryer.*

The penguin seemed to climb the hill.

It seemed that the penguin climbed the hill.

*Pearl & Sprouse in progress*
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis

rUTAH: Larson 1988, Larson 1990

Thematic roles are ordered relative to each other, with the highest thematic role mapping to the highest grammatical role (subject > object > indirect object).

Basic intuition:
doer (Agent-like) >
done-to (Patient-like) >
done-for/with (Goal-like)

An example implementation:
Agent > Causer > Experiencer > Possessor >
Subject Matter > Causee > Theme > Patient >
(Location, Source, Goal, Benefactor, Instrument)

Note: You don’t need to have every role relatively ranked. If some are unranked with respect to each other, the order in which they get mapped to grammatical positions doesn’t matter.
Thematic roles & how to use them

One idea about how children could use thematic role information: (r)UTAH.

The (relativized) Uniformity of Theta Assignment Hypothesis

rUTAH: Larson 1988, Larson 1990

Thematic roles are ordered relative to each other, with the highest thematic role mapping to the highest grammatical role (subject > object > indirect object).

Basic intuition:

*doer (Agent-like) > done-to (Patient-like) > done-for/with (Goal-like)*

This relative ranking can help deal with certain situations, like those involving Experiencers.

An example implementation:

*Agent > Causer > Experiencer > Possessor >
Subject Matter > Causee > Theme > Patient >
(Location, Source, Goal, Benefactor, Instrument)*

---

She fears spiders.

Experiencer  Subject Matter
Experiencer  Subject Matter

Causer  Experiencer
Subject  Object

Spiders frighten her.

Causer  Experiencer
Subject  Object

---

*Pearl & Sprouse in progress*