Language Acquisition

Fall 2010/Winter 2011

Syntax-Semantics Interface

Afra Alishahi, Heiner Drenhaus

Computational Linguistics and Phonetics
Saarland University

Linking Syntax to Semantics

- How is the surface structure (syntax) linked to the underlying meaning (semantics)?
  - Alternative 1: syntax is learned independently of semantics; later the two are linked together
  - Alternative 2: syntax and semantics are learned simultaneously
- Central unit: verb argument structure
  - Relationship btw the semantics of a verb and its syntactic form
    - Number and type of the arguments that the verb takes
    - Semantic roles that the arguments receive in an event
    - Syntactic realization of the verb and its arguments

How to Convey a Relational Meaning?

chimp

apple

eat

The chimp is eating an apple

How to Convey a Relational Meaning?

[Fisher'94]

This is blicking!

The rabbit is blicking the duck

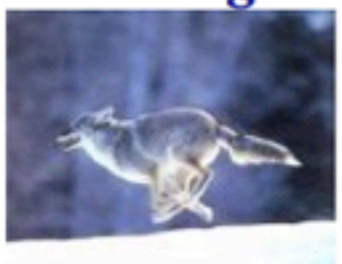
How to Convey a Relational Meaning?

She is dropping the vase.

The vase is falling.

**She is falling the vase.*

running



drinking

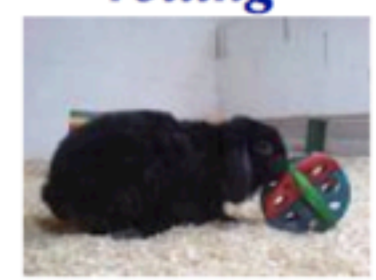
breaking

looking

reading

eating

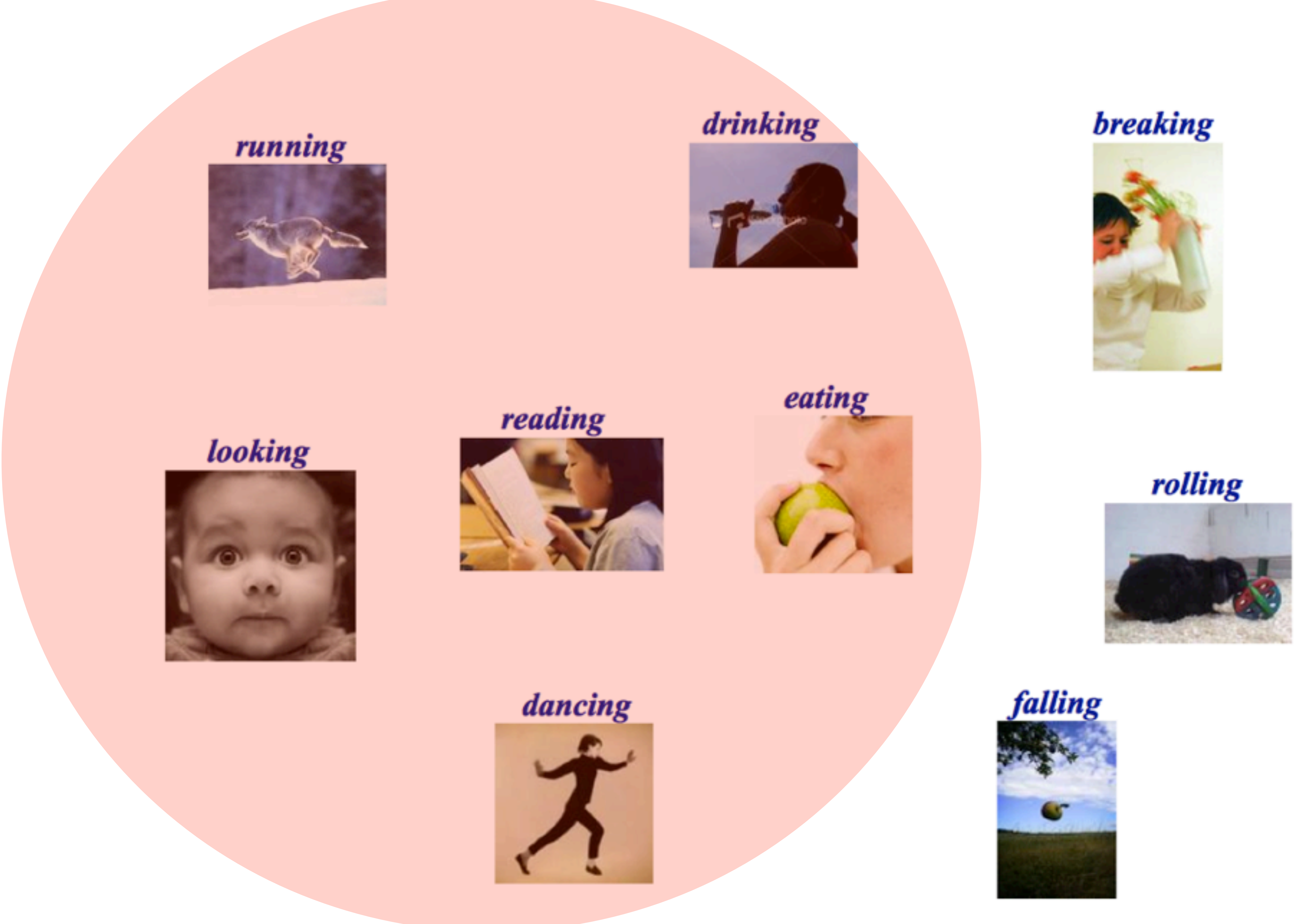
rolling



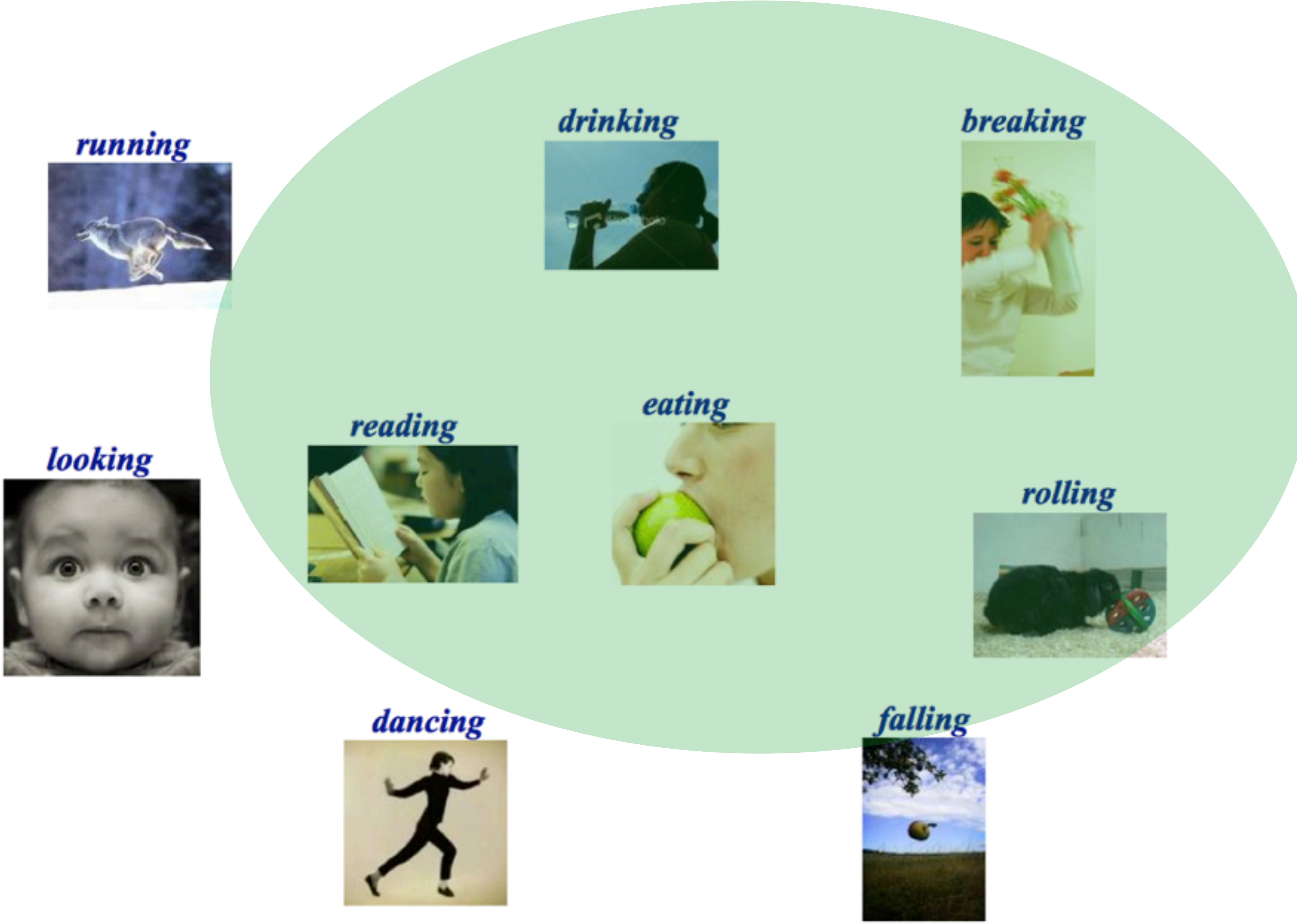
dancing

falling

AGENT *is* VERB*ing*

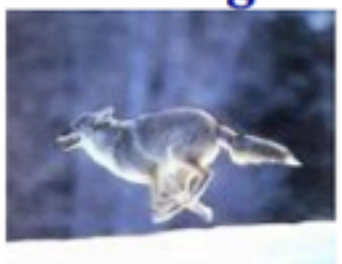


AGENT *is* VERB*ing* THEME



THEME *is* VERB*ing*

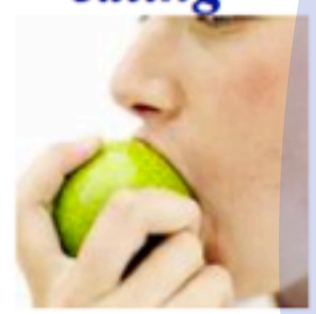
running



drinking

breaking

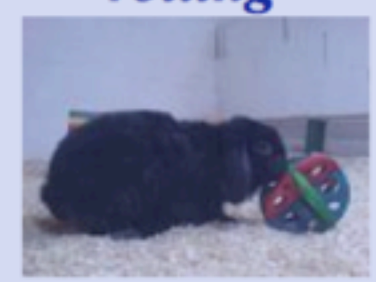
eating



reading

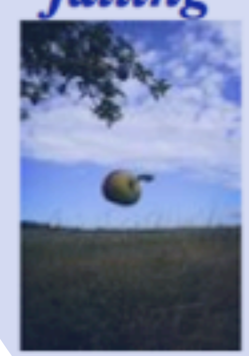
looking

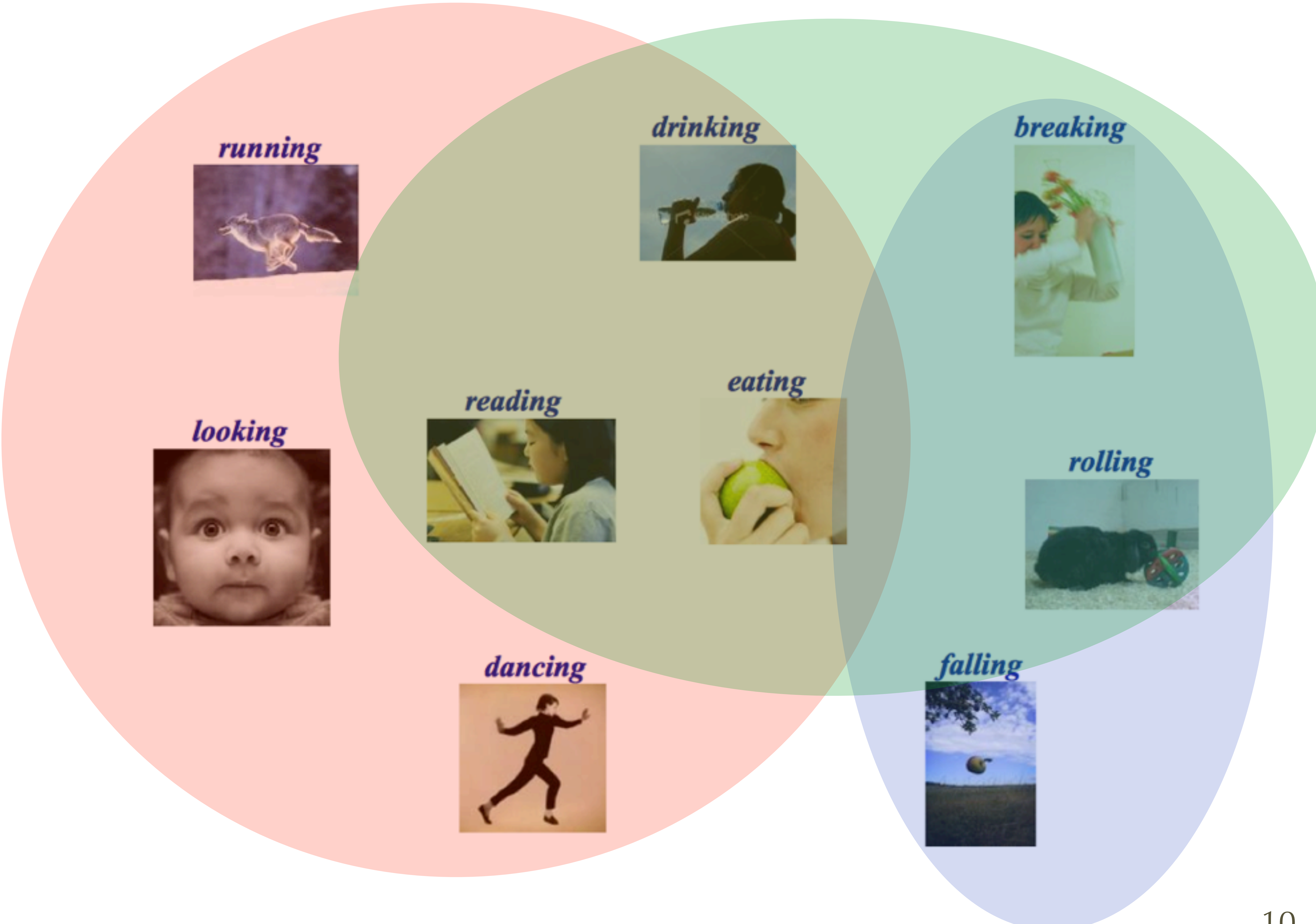
rolling



dancing

falling




running
drinking
breaking
looking
reading
eating
rolling
dancing
falling

Acquisition of Verb Argument Structure

- General patterns
  - Young children are sensitive to argument structure regularities

bunny gorped duck ⇒ causal action?

kitty blicked down the street ⇒ manner of motion?

- Idiosyncrasies
  - Semantically similar verbs can have different syntactic behaviour

*I filled the glass with water, *I filled water into the glass*

**They loaded the truck with hay, They loaded hay into the truck*

- A U-shaped behavioural pattern is observed for children's argument structure acquisition

Semantic Bootstrapping

- Semantic Bootstrapping (Pinker, 1984)
  - Syntactic behaviour of a verb is innately determined by the decompositional representation of its meaning

```
Agent is 1st argument of CAUSE

Patient is 2nd argument of CAUSE

Theme is 1st argument of GO and BE
```

  - With the innate knowledge of mapping between semantics and syntax, a child can predict the correct mapping once she knows what a verb means

Argument Structure Constructions

- Construction Grammar
  - Meaning may be directly associated with syntactic forms
  - Lakoff 1987, Fillmore et al. 1988, Langacker 1999
- Argument structure construction (Goldberg, 1995)
  - A mapping between underlying verb-argument relations and the syntactic form that is used to express them

> Subj V Obj Obj2 ⇔ `X cause Y receive Z`
>
> Example: *Pat faxed Bill the letter.*

> Subj V Oblique ⇔ `X move Y`
>
> Example: *The fly buzzed into the room.*

How are Constructions Learned?

- Tomasello (1991):
  - Argument structure patterns are acquired on a verb-by-verb basis
  - Abstract constructions learned through categorization and generalization of common patterns
- Goldberg (1995):
  - Constructional meaning is formed around the meanings of highly frequent light verbs
  - E.g., the construction "Subj V Oblique" paired with the meaning **`"X moves Y"`** corresponds to the light verb ***go***

Computational Modeling of Constructions

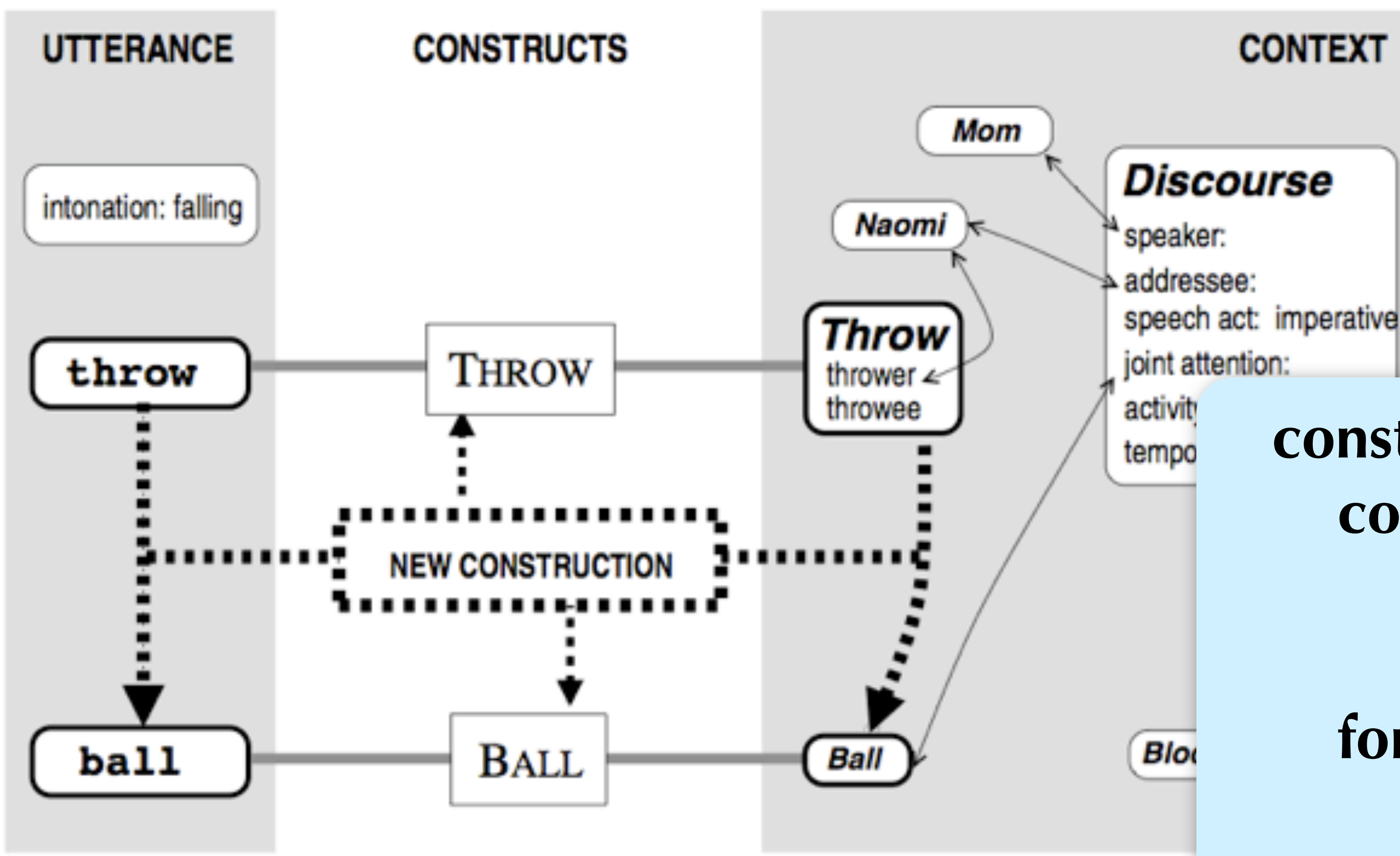
- FrameNet (Baker, Fillmore, Low, 1998)
  - A database of lexical constructions (or frames)
- The acquisition of constructions
  - Learning lexical constructions via structure mapping (Chang, 2004)
  - Learning verb meaning from image data (Dominey, 2003; Dominey & Inui, 2004)
  - Learning abstract constructions from verb usage data (Alishahi & Stevenson, 2008)

Case Study: Chang (2004)

- Learning lexical-based constructions from child-directed data
  - Goal: learning associations between form relations (word order) and meaning relations (role-filler bindings)
  - Search space: grammars defined by a unification-based formalism (Embodied Construction Grammar, ECG)
  - Form and meaning representations: subgraphs of elements and relations among them
  - Construction representation: a mapping between two subgraphs
  - Learning task: finding the best grammar to fit the observed data

Case Study: Chang (2004)

- Learning lexical-based constructions from child-directed data



construction THROW-BALL
constituents
t1: THROW
t2: BALL
form
t1*f* before t2*f*
meaning
t1*m*.throwee t2*m*

Case Study: Chang (2004)

- The model makes generalizations at the lexical level:

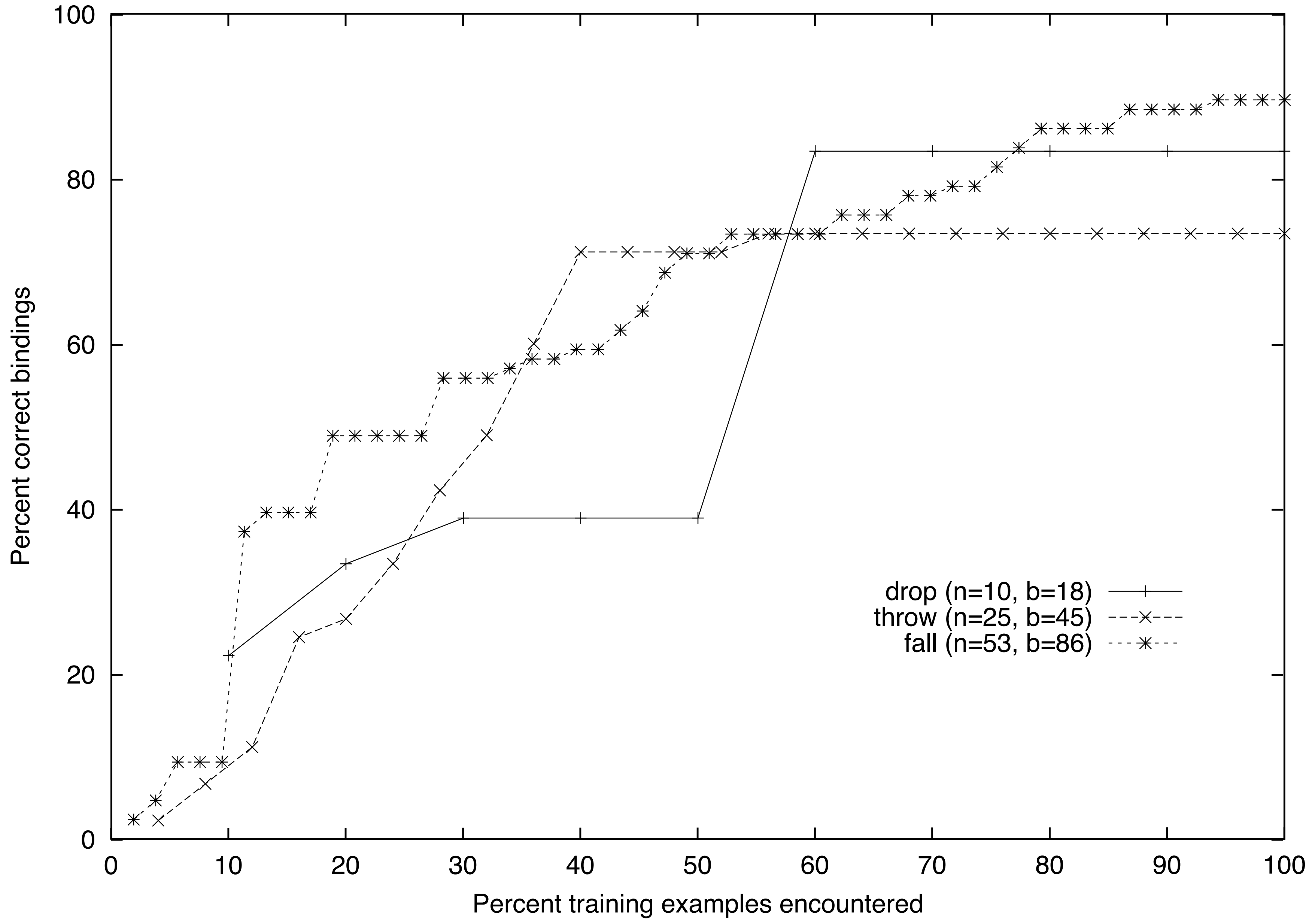
construction THROW-BALL
  constituents
    t1: THROW
    t2: BALL
  form
    $t1_f$ before $t2_f$
  meaning
    $t1_m$.throwee t2

construction THROW-BLOCK
  constituents
    t1: THROW
    t2: BLOCK
  form
    $t1_f$ before $t2_f$
  meaning
    …owee $t2_m$

construction THROW-OBJECT
  constituents
    t1: THROW
    t2: OBJECT
  form
    $t1_f$ before $t2_f$
  meaning
    $t1_m$.throwee $t2_m$

Case Study: Chang (2004)

- The model makes generalizations at the lexical level:



Case Study: Alishahi & Stevenson (2008)

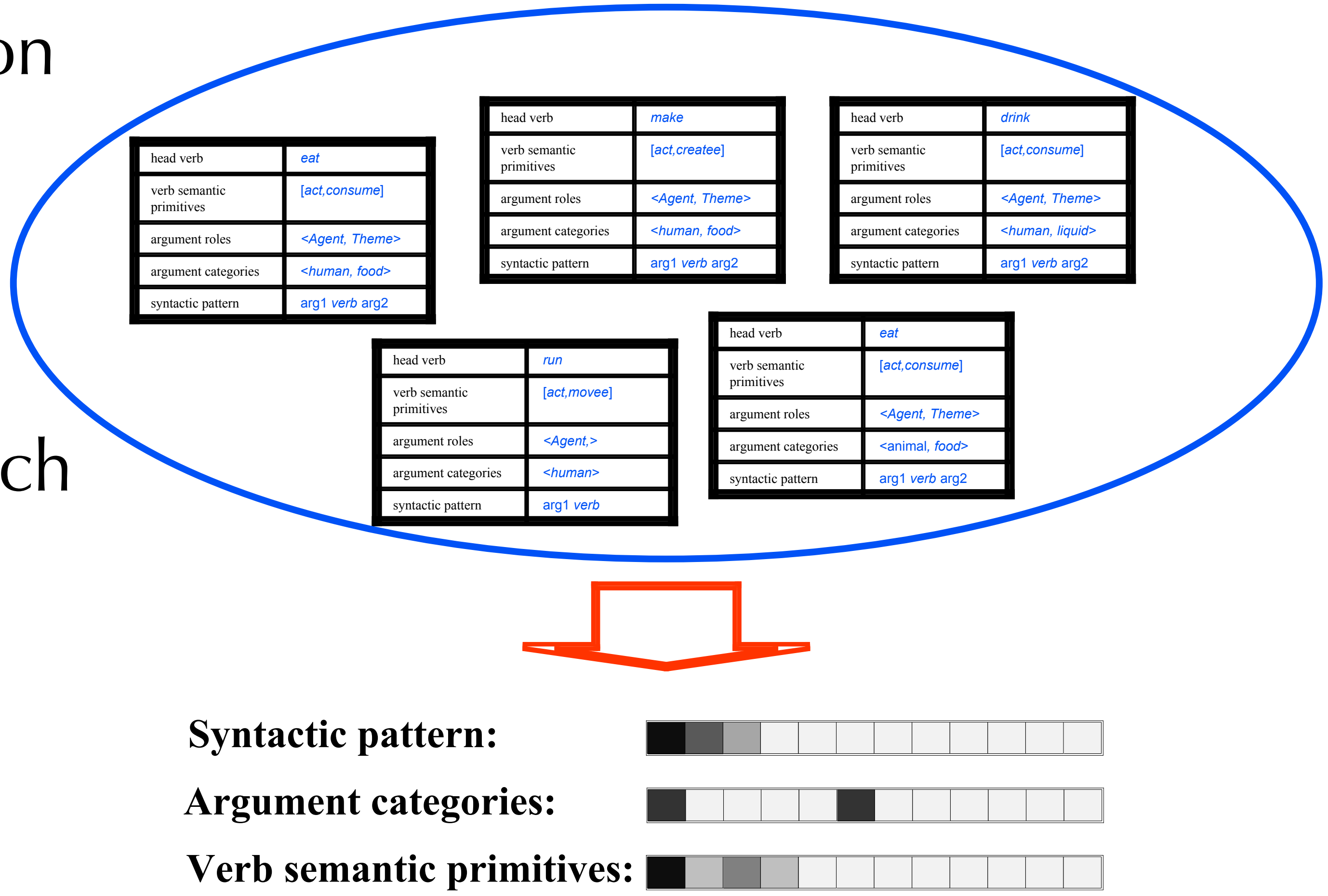
- A Bayesian model of early argument structure acquisition
  - Each verb usage is represented as a set of features

Sara is eating an apple

head verb	*eat*
verb semantic primitives	[*act*,*consume*]
argument roles	*<Agent, Theme>*
argument categories	*<human, food>*
syntactic pattern	arg1 *verb* arg2

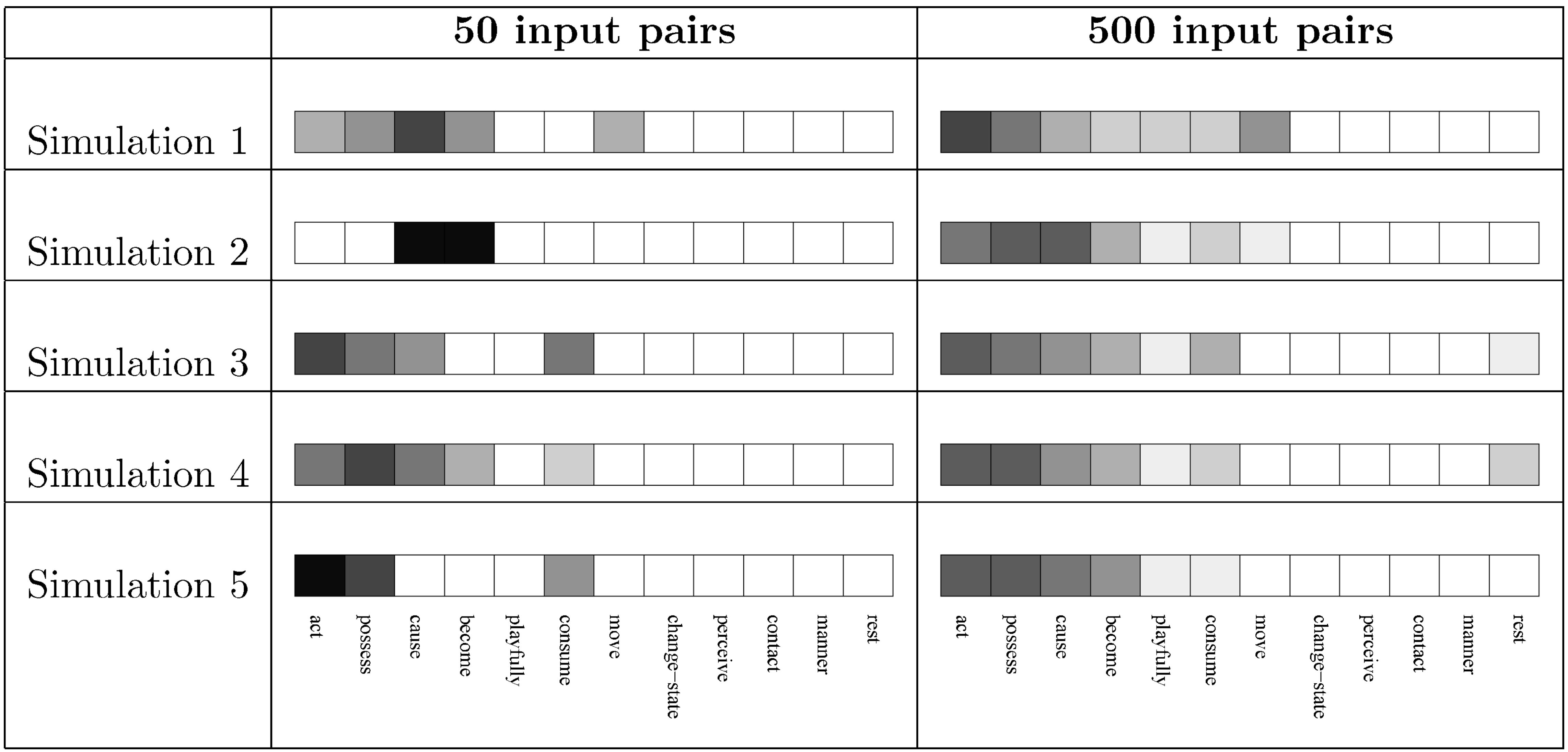
Case Study: Alishahi & Stevenson (2008)

- A Bayesian model of early argument structure acquisition
  - Each verb usage is represented as a set of features
  - Each construction is a cluster of verb usages
    - A probability distribution over feature values
  - The best construction is found for each new usage through a Bayesian approach



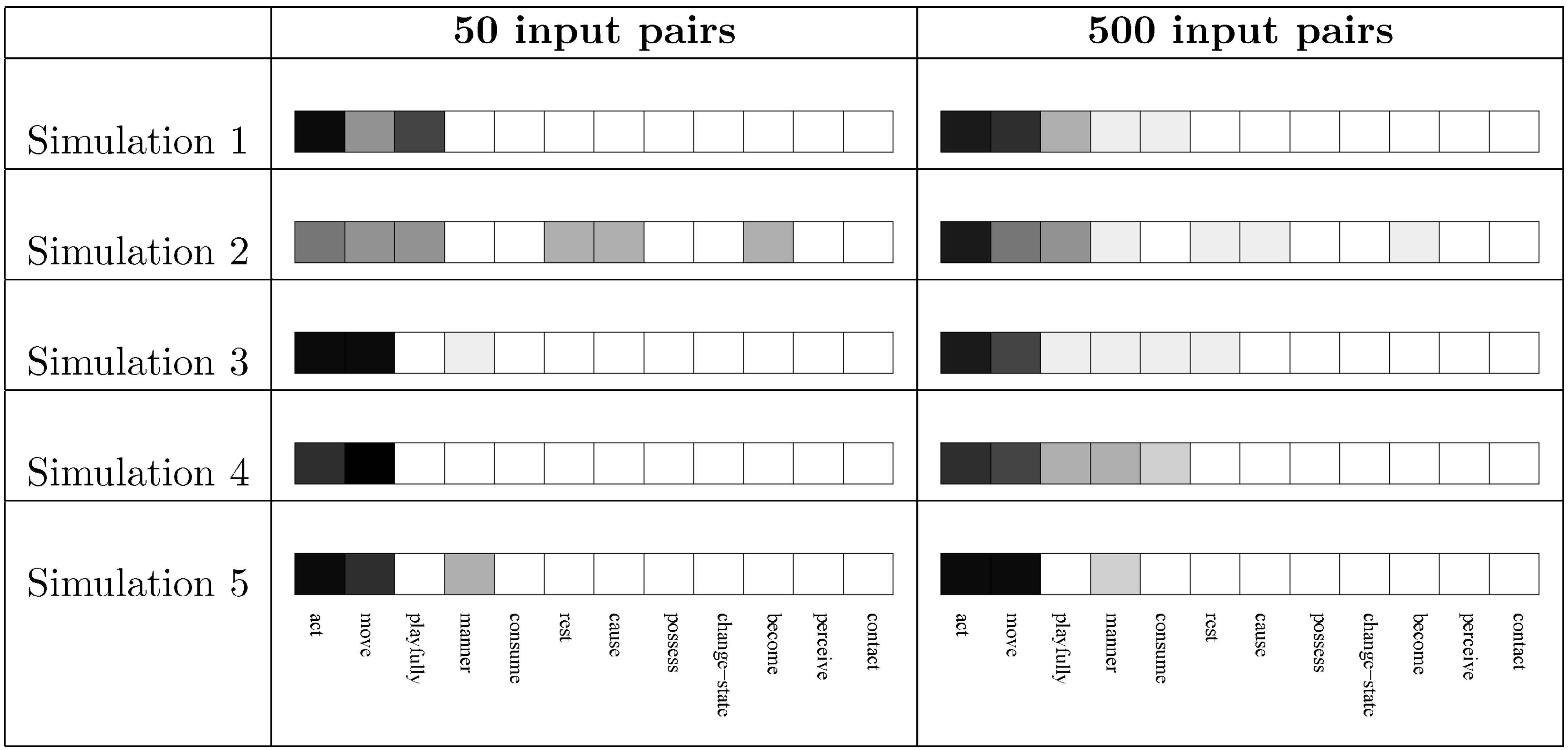
Sample Constructions

- Verb semantic primitives for Transitive Construction:



Sample Constructions

- Verb semantic primitives for Intransitive Construction:

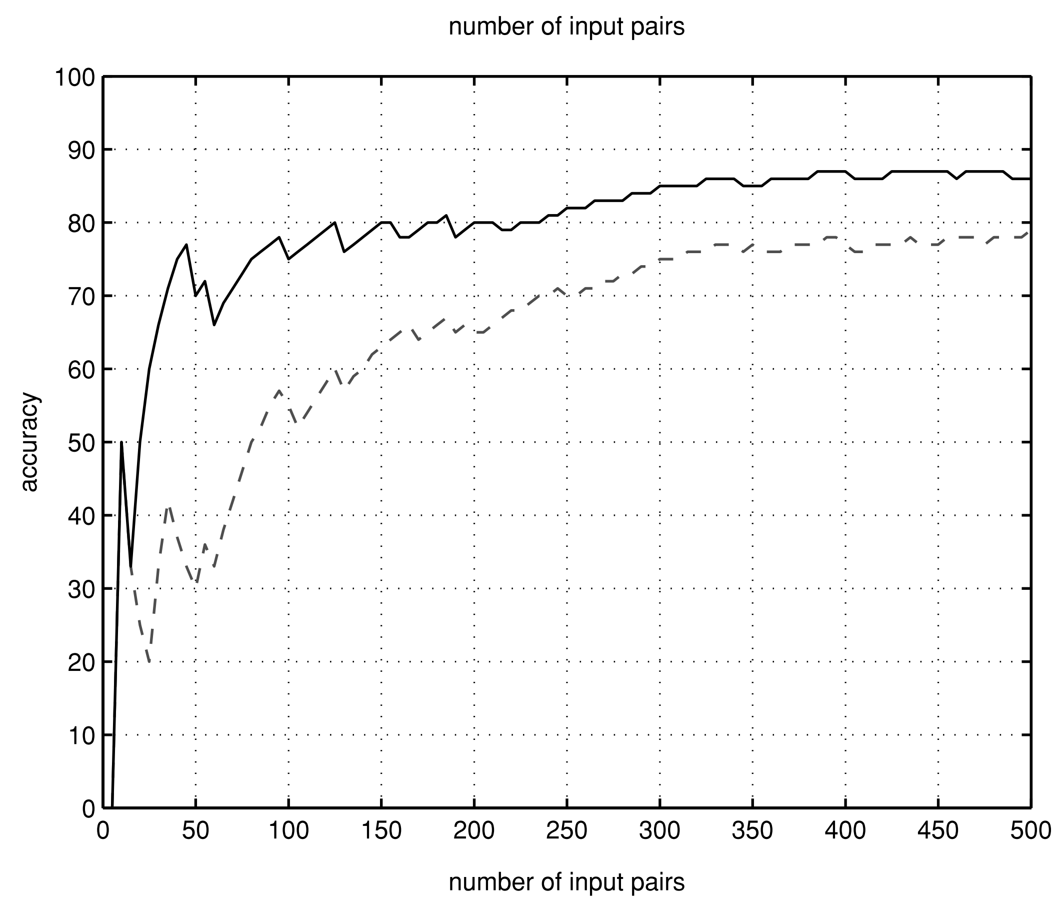


Sample Sentence Production

- First eight usages of *fall* by Adam (CHILDES) and by one of the simulations of our model

Adam	Our model
go fall!	*John fall ball*
no no fall no!	*toy fall*
no fall!	*Mary fall book*
oh Adam fall.	*toy fall*
Adam fall toy.	*cookie fall*
Adam fall toy.	*kitty fall*
oh fall.	*spoon fall*
I not fall.	*ball and toy fall*

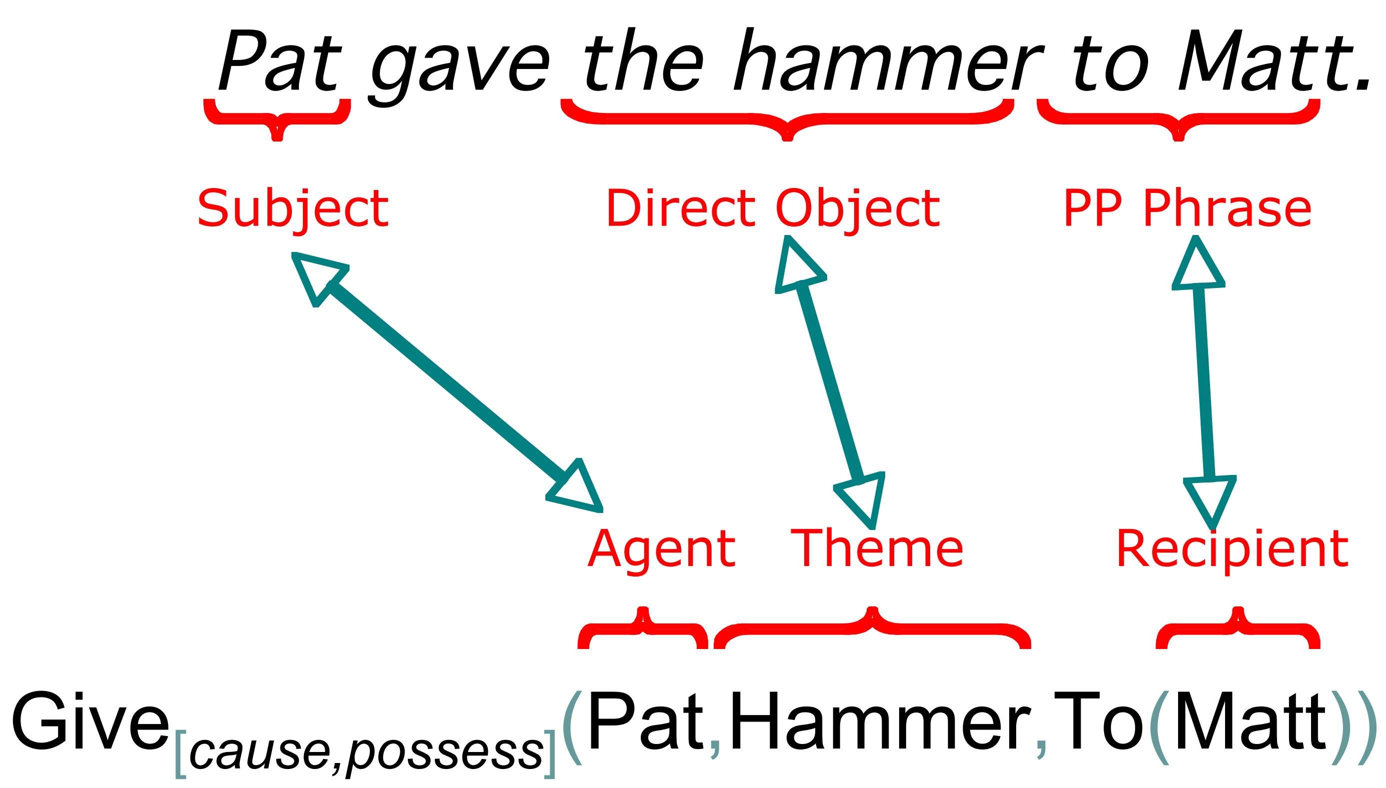
Learning Curves



- Learning phases are successfully simulated:
  - Imitation
  - Overgeneralization and recovery
  - Productive generalization

Verb Semantic Roles

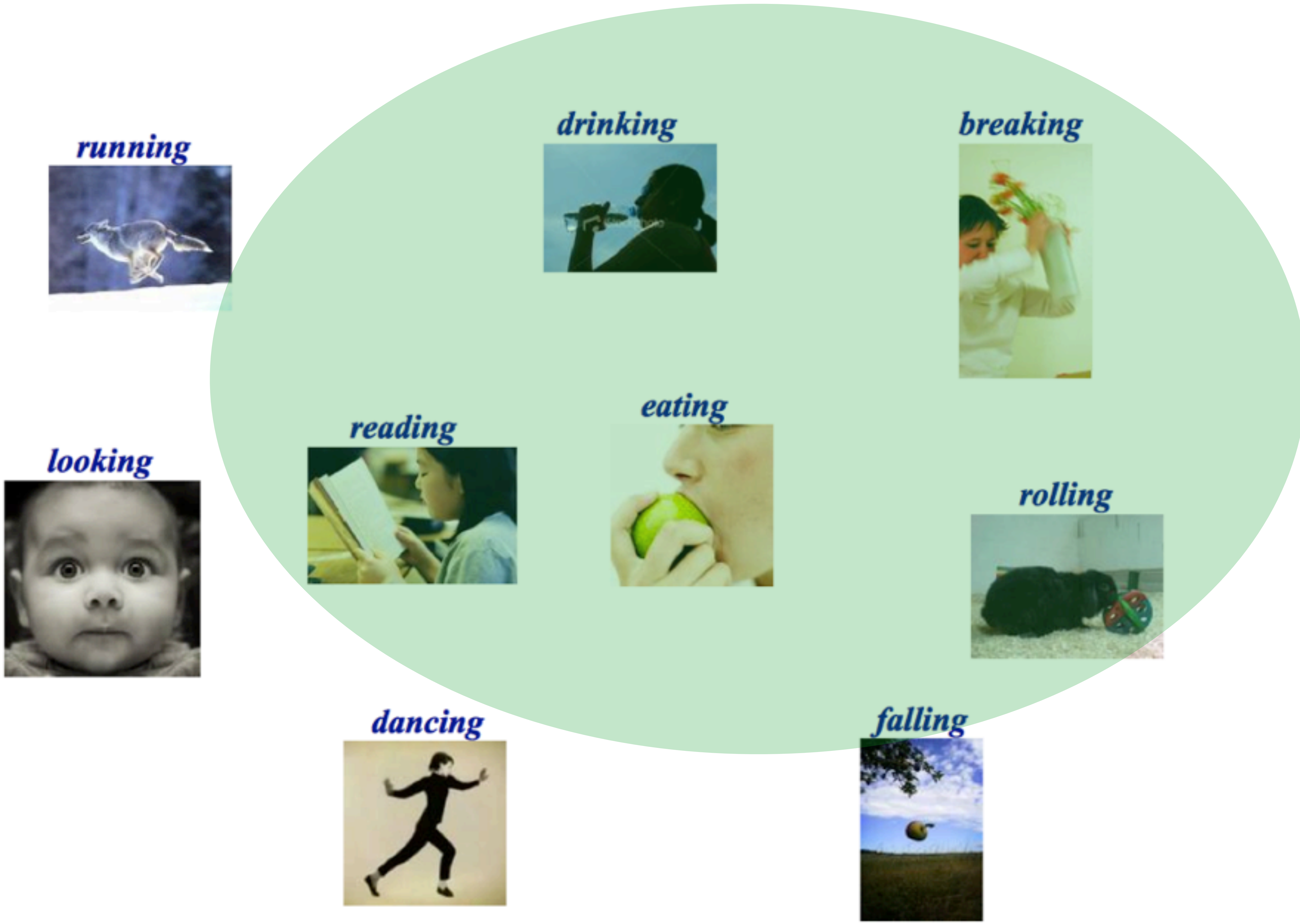
- Semantic (thematic) roles indicate the relations of the participants in an event to the main predicate



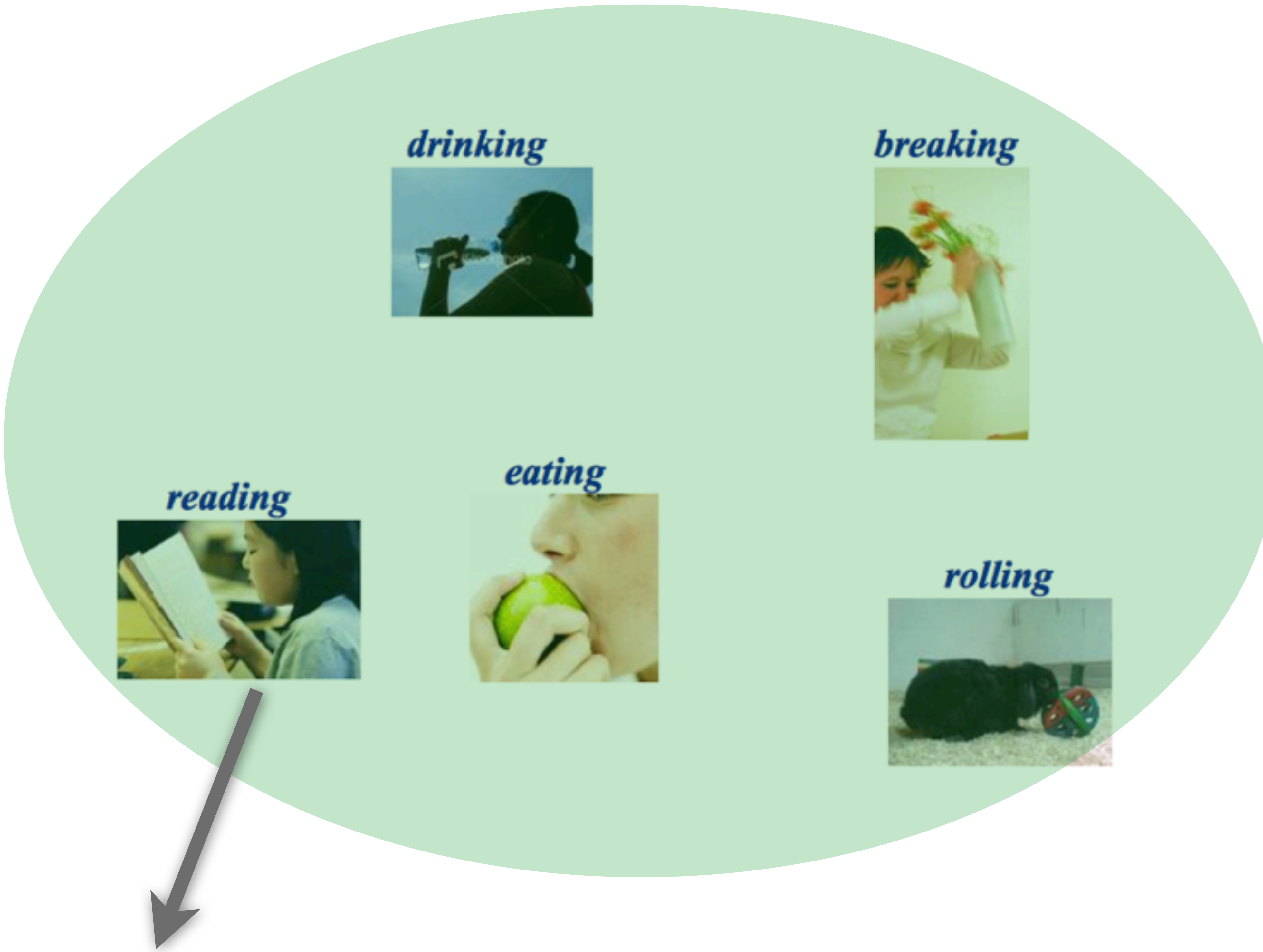
Theoretical Questions

- What is the nature of semantic roles?
  - Traditional view: roles are atomic and universal, such as Agent, Theme, Goal, … (e.g., Jackendoff 1990)
  - Proto-role Hypothesis (Dowty, 1991): roles are a set of properties, such as volitional, affecting, animate
- Where do they come from?
  - Traditional view: roles and their link to syntactic positions are innate (e.g., Pinker 1989)
  - Alternative view: they are gradually learned from verb usages (e.g., Tomasello 2000)

AGENT *is* VERB*ing* THEME

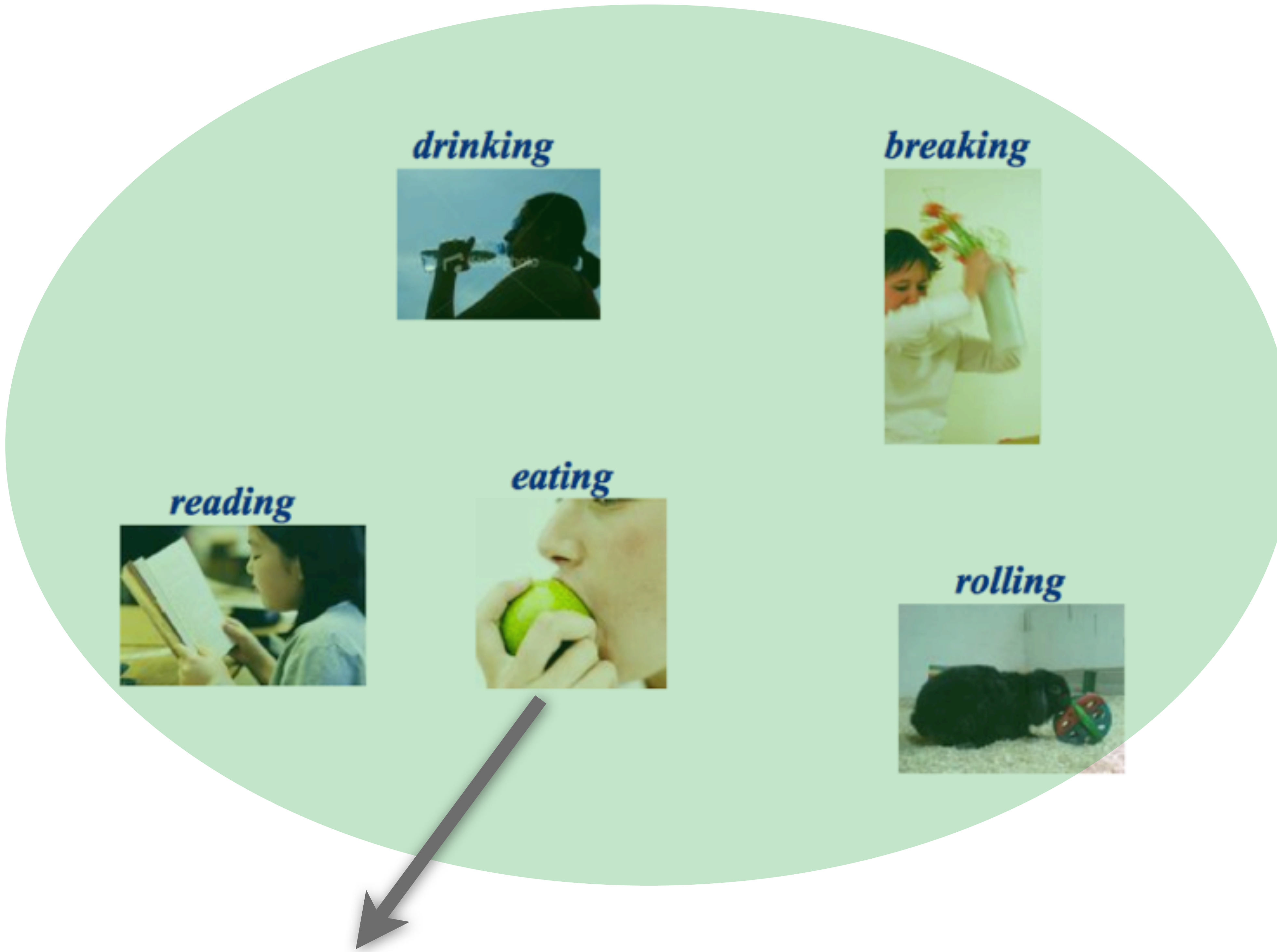


AGENT *is* VERB*ing* THEME



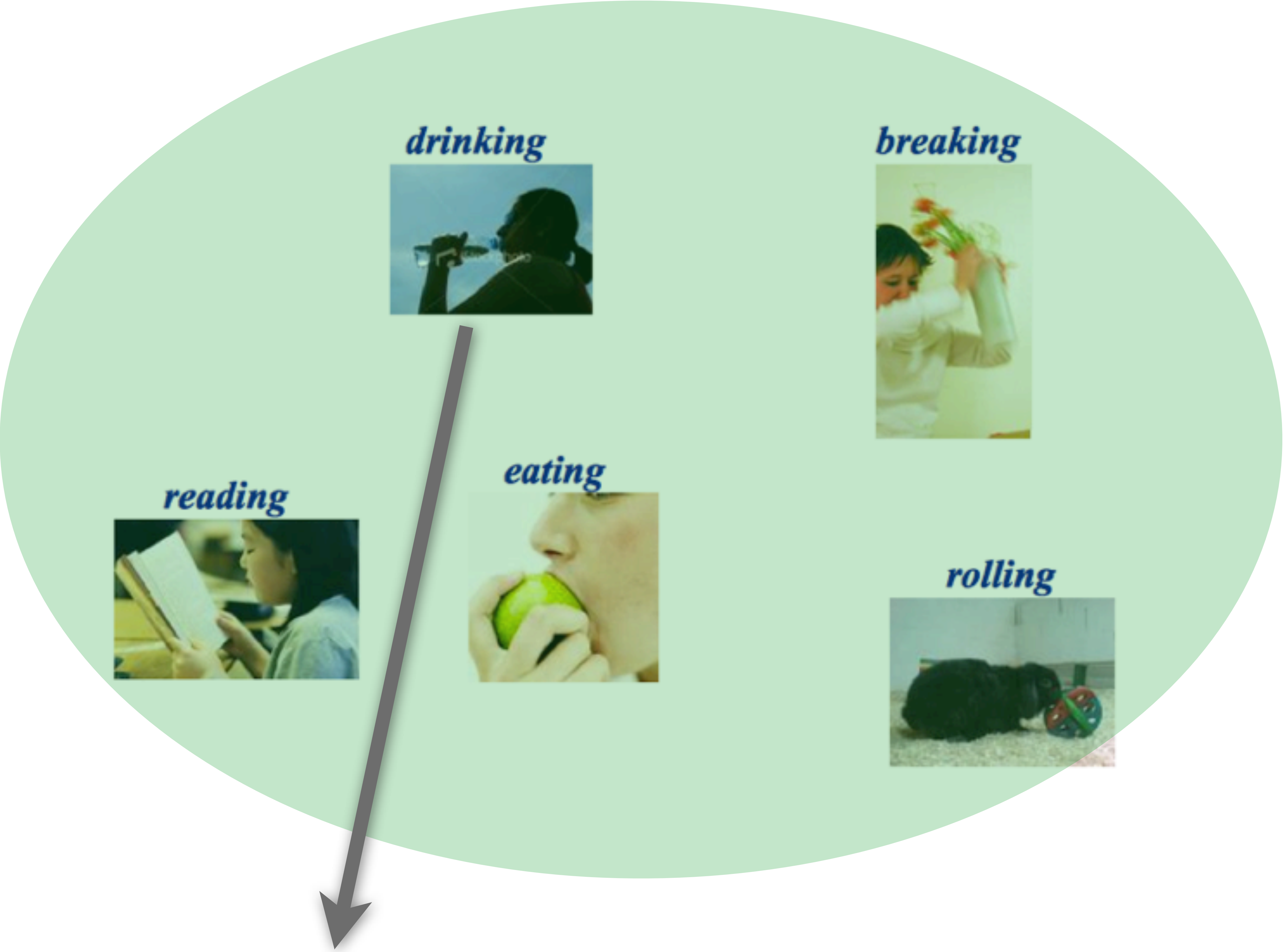
reader is reading text

AGENT *is* VERB*ing* THEME



eater is eating food

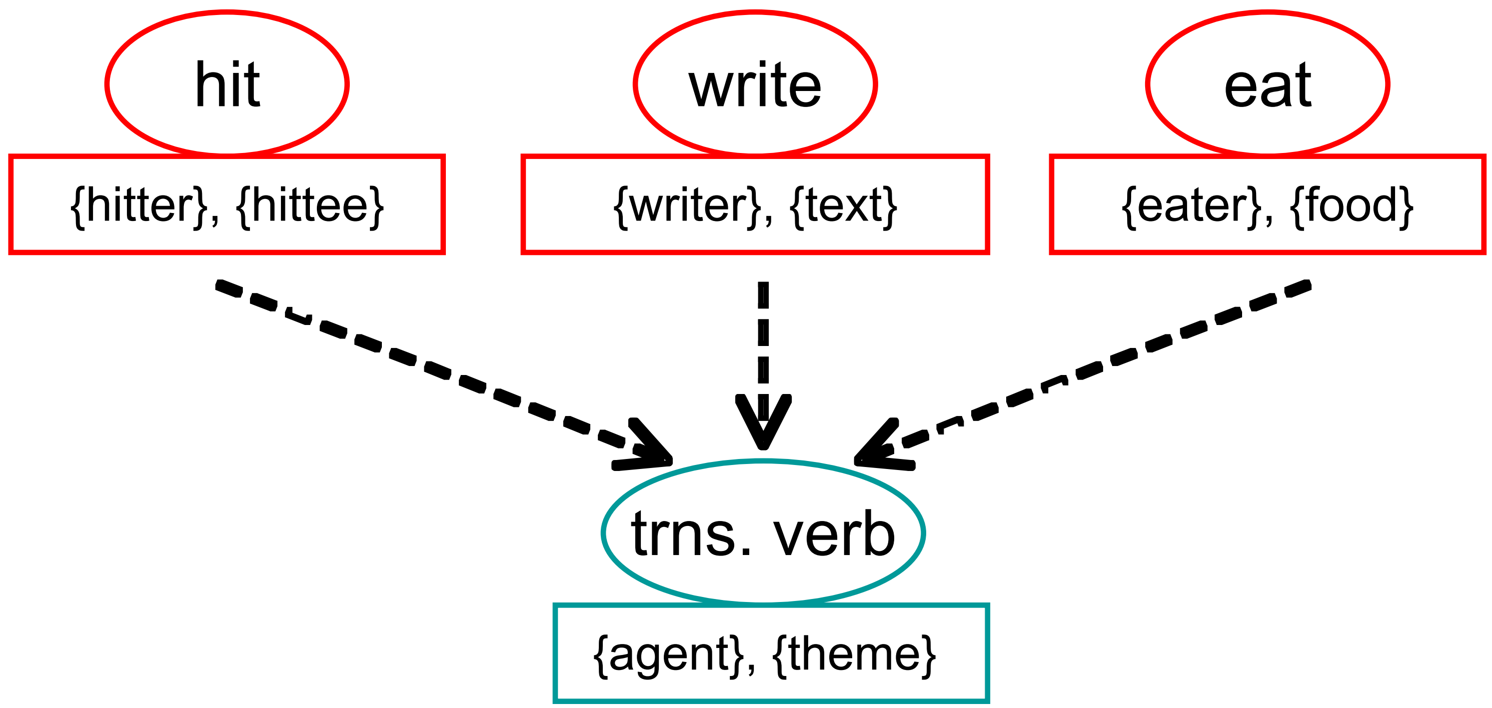
AGENT *is* VERB*ing* THEME



drinker is drinking liquid

Learnability of Semantic Roles

- Usage-based account: verb-specific roles change to general roles over time



- Experimental evidence confirms that access to general roles such as Agent and Theme is age-dependent (Shayan & Gershkoff-Stow, 2007)

Linking Semantic Roles to Grammatical Functions

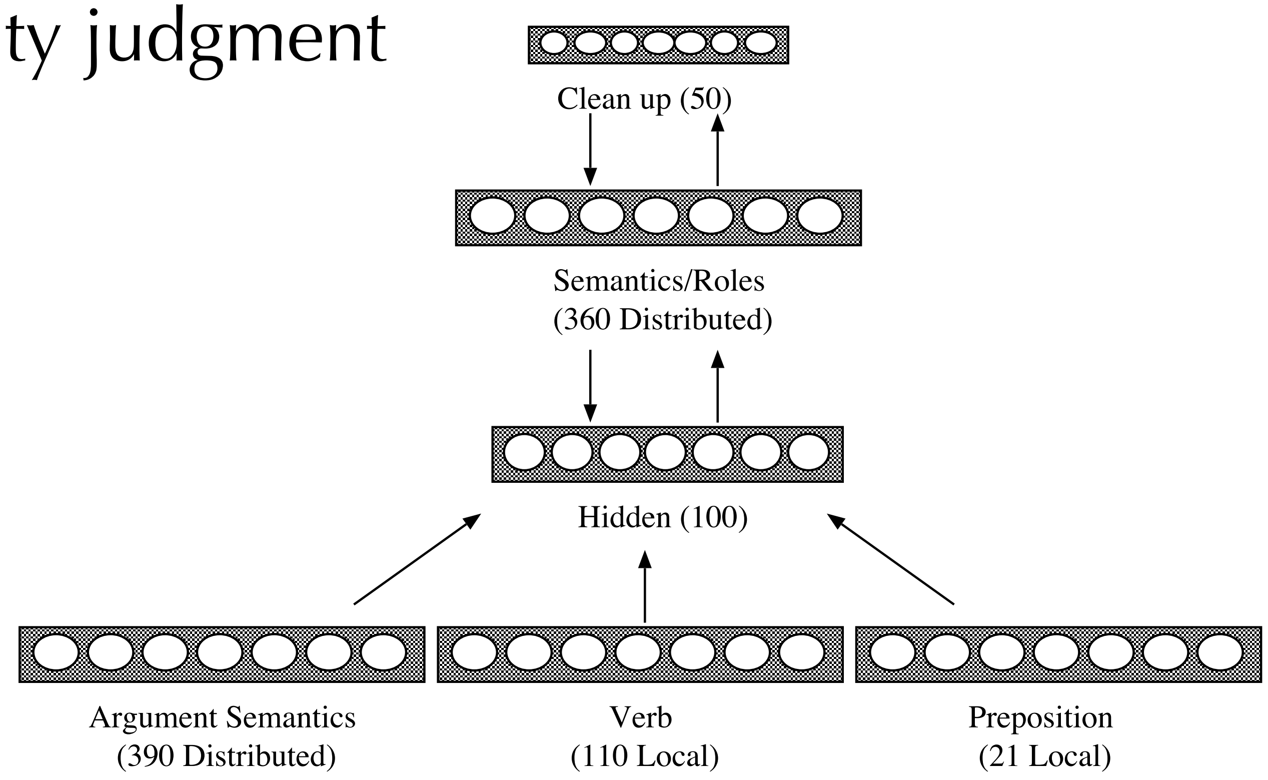
- Semantic roles are linked to syntactic positions early on
  - Children are sensitive to the association between semantic roles (e.g. Agent) and grammatical functions (e.g. Subject) from an early age (Fisher 1994, 1996; Nation et al., 2003)
- Nativist account
  - Innate linking rules that map roles to sentence structure enable children to infer associations between role properties and syntactic positions (e.g., Pinker, 1989)

Computational Studies of Roles

- Assignment of general pre-defined roles to sentence constituents
  - E.g., McClelland and Kawamoto (1986), Allen (1997)
- Role learning
  - Learning verb-specific roles from annotated data (Chang 2004)
  - Discovering relational concepts from unstructured examples (Kemp et al., 2006; Doumas et al., 2008)
  - Acquiring semantic profiles for general roles from verb usages (Alishahi & Stevenson, 2008)

Case Study: Allen (1997)

- A connectionist model of thematic role assignment
  - Integrates syntax, semantics and lexical information
  - Is trained on usages of most frequent verbs in CHILDES
  - Predicts semantic properties of verbs and nouns
  - Simulates grammaticality judgment



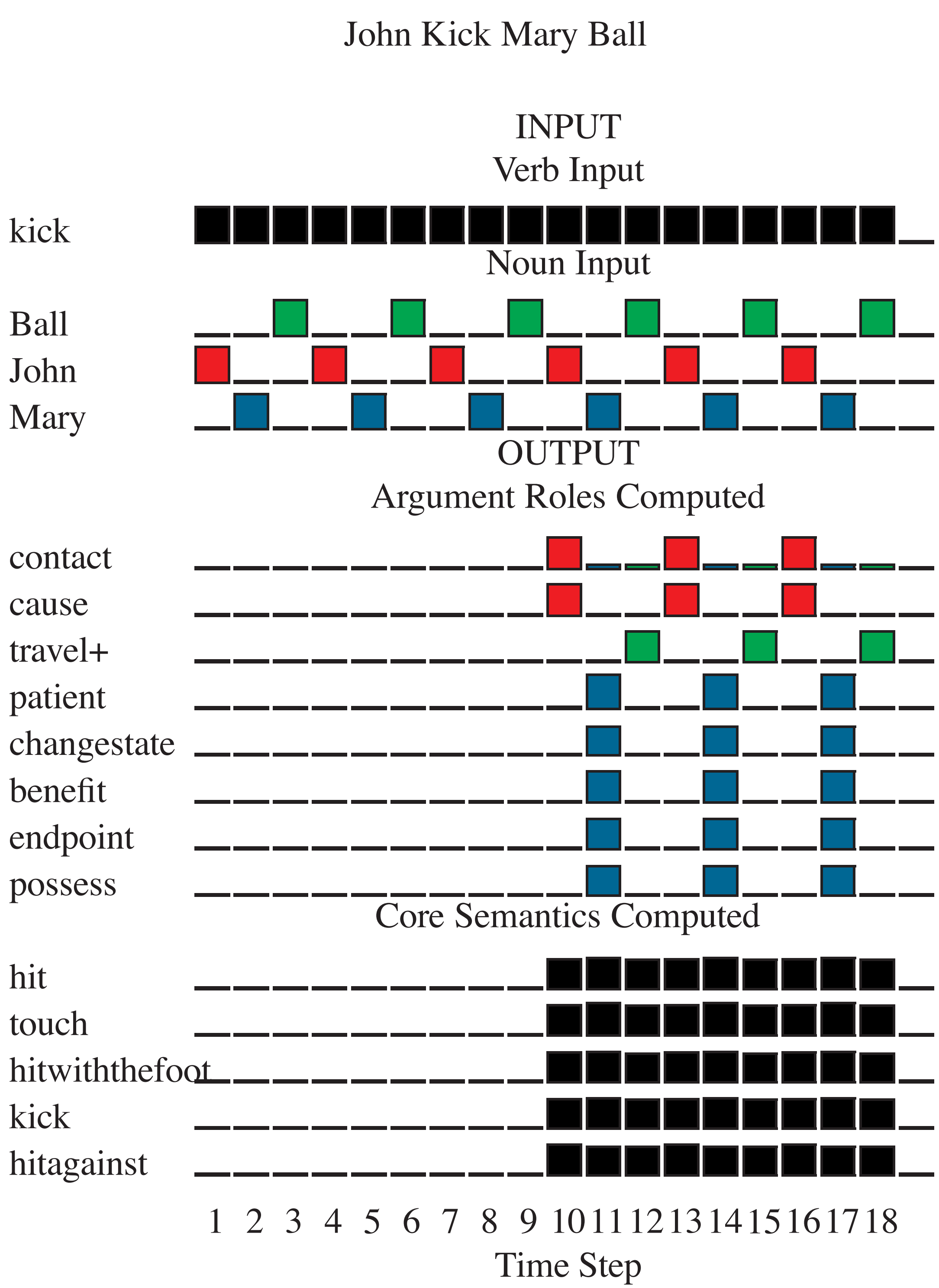
Role Representation

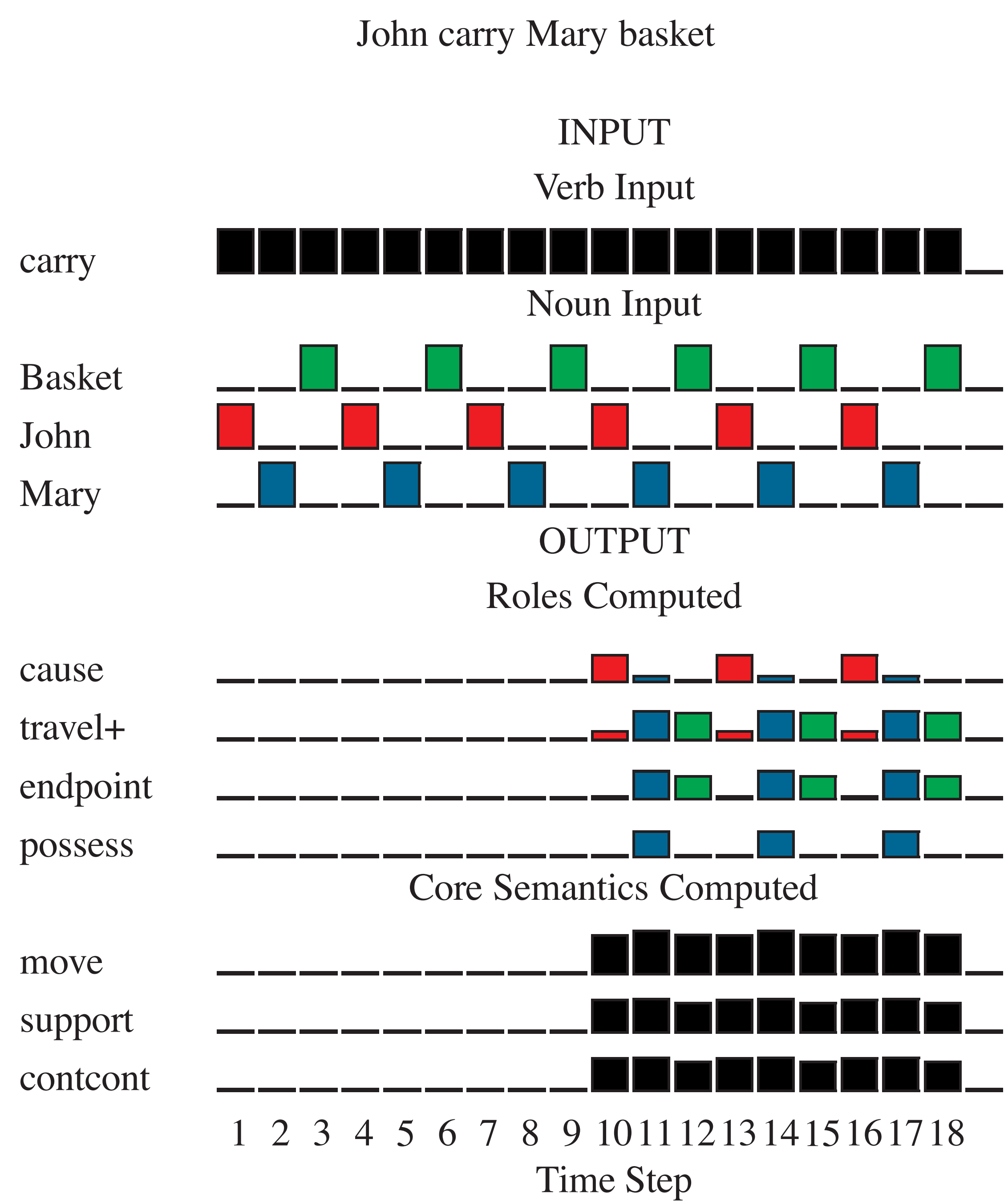
- Each basic role was elaborated by a set of subroles, or proto-role properties:

Cause →
patient
change of state
motion
travel
location
experiencer
possessor
instrument
path

Causation Subtypes
apply force
action
direct cause
allow
help
impede
instrumental
author
agent
internal cause

Semantic Prediction & Grammaticality





Case Study: Alishahi & Stevenson (2010)

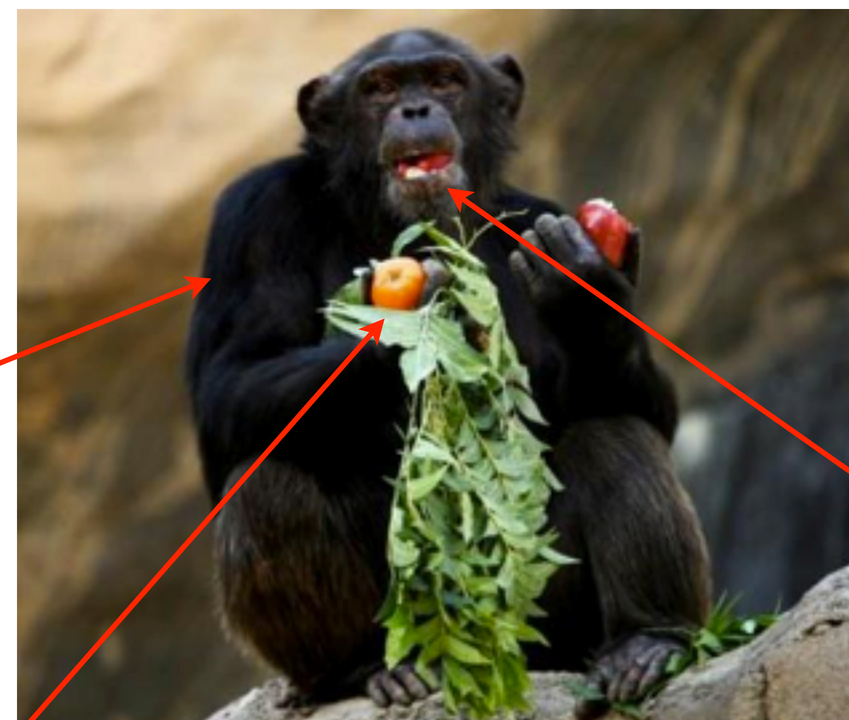
- A Bayesian model of early verb learning can learn
  - general conceptions of roles based only on exposure to individual verb usages
  - associations between general semantic properties and the syntactic positions of the arguments
- The acquired semantic roles
  - naturally metamorphose from verb-specific to general properties
  - are an intuitive match to the expected properties of various roles
  - are useful in guiding comprehension in the face of ambiguity

Distributional Representations of Roles

Number of arguments: 2

The chimp is eating an apple

arg1 verb arg2



Lexical properties:
{living thing, animal, chimp, ...}

Event-based properties:
{volitional, affecting, animate, ...}

Event (Verb): Eat

Event properties:
action, consumption

Lexical properties:
{entity, object, fruit, ...}

Event-based properties:
{non-independently exist, affected, change, ...}

Event-based Properties of Transitive Arguments

Argument 1 (Agent)	
Probability	Event-based property
0.048	independently exist
0.048	sentient
0.035	animate
0.035	change
0.035	affected
0.035	change emotional
0.035	becoming
0.013	volitional
0.013	possessing
0.013	getting

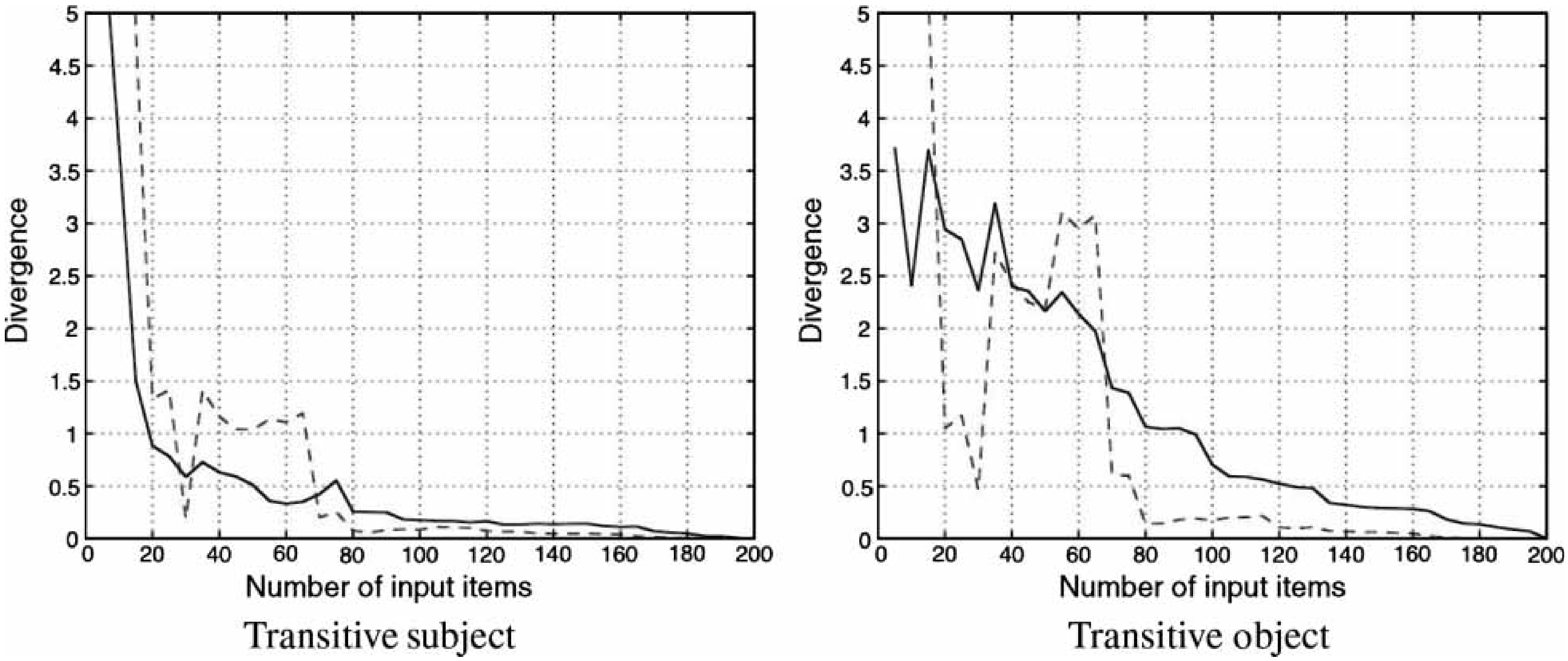
Argument 2 (Theme)	
Probability	Event-based property
0.086	state
0.031	independently exist
0.031	change
0.031	change possession

Lexical Properties of Transitive Arguments

Argument 1 (Agent)	
Probability	Lexical property
0.054	entity
0.040	object
0.040	physical object
0.026	being
0.026	organism
0.026	living thing
0.026	animate thing
0.015	person
0.015	individual
0.015	someone
0.015	somebody
0.015	mortal
0.015	human
0.015	soul
0.015	causal agent

Argument 2 (Theme)	
Probability	Lexical property
0.056	entity
0.037	object
0.037	physical object
0.023	unit
0.023	artifact
0.023	artefact
0.023	whole
0.023	whole thing
0.018	abstraction
0.014	being
0.014	organism
0.014	living thing
0.014	animate thing
0.014	person
0.014	individual

Learning Curves for Semantic Profiles



Verb Selectional Restrictions

- Most verbs impose semantic restrictions on the arguments that they take, which
  - affect the acceptability of natural language sentences: *eating food, drinking water, *eating water, *drinking food*
  - facilitate language comprehension and word learning
- Earlier theories view selectional constraints as defining features of the arguments:

> ***hit*** (**Subj** , **Obj**)
>
> **Subj**: HUMAN or HIGHER ANIMAL
>
> **Obj**: PHYSICAL OBJECT

- Identifying necessary and sufficient restrictions is a challenge

Verb Selectional Preferences

- Resnik (1993) proposed an alternative view: verbs have *preferences* for the type of arguments they allow for
  - World knowledge is represented as a semantic class hierarchy
  - Selectional preferences are viewed as probability distributions over various semantic classes
- Verbs have different degrees of preference
  - e.g. *eat* and *sing* have strong preferences for the direct object position, but *put* and *make* do not

Computational Modeling of Selectional Preferences

- Most of the existing computational models are influenced by the information-theoretic model of Resnik (1993,1996)
  - Represent preference for an argument position of a verb as a mapping of each semantic class to a real number
  - Model the induction of a verb's preferences as estimating that number, using a training data set
  - Examples: Li and Abe (1998), Abney and Light (1999), Ciaramita and Johnson (2000), Clark and Weir (2002)
- Different approach: Erk (2007)
  - Estimate preferences for a head word based on the similarities between that word and other head words observed in a corpus.

Cognitive Modeling and NLP

- Early NLP viewed itself as building models of human understanding
- Modern NLP has shifted emphasis
  - Focus on applications: do limited tasks accurately and robustly, often without real understanding
  - Emphasis on representations, coverage and efficiency, not concerned with cognitive plausibility
- However, cognitive modeling of language is heavily informed by research in NLP
  - Modeling of human language acquisition is influenced by specialized machine learning techniques

Open Questions

- How various aspects of language acquisition interact with each other?
  - Various learning procedures are most probably interleaved (e.g., word leaning and syntax acquisition)
  - Most of the existing models of language acquisition focus on one aspect, and simplify the problem
- How to evaluate the models on realistic data?
  - Large collections of child-directed utterances/speech are available, but no such collection of semantic input
  - A wide-spread evaluation approach is lacking in the community