A Usage-Based Model of Early Grammatical Development

Barend Beekhuizen[1,2]

[1]Leiden University Centre for Linguistics, Leiden University

[2]Institute for Logic, Language, and Computation, University of Amsterdam

12 December 2014

Coglingdays 6, Ghent

Motivation

- How does a constructicon emerge in language acquisition?

Theoretical issues

- Metaphors of abstraction: hypotheses vs. processing
- Starting big and starting small: the horizontal dimension
- The linking problem: 'all together now'

Model and experiments

- The Syntagmatic-Paradigmatic Learner
- Comprehension experiment
- Generation experiment

Motivation

- How does a constructicon emerge in language acquisition?

Theoretical issues

- Metaphors of abstraction: hypotheses vs. processing
- Starting big and starting small: the horizontal dimension
- The linking problem: 'all together now'

Model and experiments

- The Syntagmatic-Paradigmatic Learner
- Comprehension experiment
- Generation experiment

Motivation

- How does a constructicon emerge in language acquisition?

Theoretical issues

- Metaphors of abstraction: hypotheses vs. processing
- Starting big and starting small: the horizontal dimension
- The linking problem: 'all together now'

Model and experiments

- The Syntagmatic-Paradigmatic Learner
- Comprehension experiment
- Generation experiment

Understanding the acquisition of grammar

- Adult state (cf. Goldberg 1995): Constructions of various degrees of abstraction
- How to get there? (Tomasello 2003, Goldberg 2006)
- Clear sketch of processes: Langacker (2009):
  - Learning is a by-product of processing
  - 'Units' (constructions) emerge through selective reinforcement
  - Units are of the same make as the input items (conceptual and phonological structure)
  - Abstraction is not 'creating something new', but rather the potential that is immanent in a number of more concrete experiences

Three theoretical issues

What *is* abstraction

- Usage-based work blends two frames:
  - abstraction as hypotheses about generalizability
  - abstraction as a by-product of processing

What *is* abstraction

- Usage-based work blends two frames:
  - abstraction as hypotheses about generalizability
  - abstraction as a by-product of processing

Relevance of this?

- Hypothesis frame brings along:
  - Search for new hypotheses
  - Evaluating them against some data
  - Corroborating or rejecting them
- Legacy: too post-hoc/offline for usage-based view
- Chang (2006) is based on the hypothesis view.

Proposal #1: Abstraction

- Emphasize the 'by-product' view in explanation of abstraction

Chunks and syntagms: The horizontal dimension

- What is the width/arity/'length' of early units
- Impression: infant has built up (e.g.) full caused-motion utterances, then abstracts i.o to form paradigms.
- Starting big: infants learn chunks, later break them down and do 'blame assignment' (Tomasello 2003)
- Problems with starting big
  - Storage of unstructured phonological wholes is problematic
  - Profile early SVO/SV/VO-productions (Theakston et al. 2012)
  - Argument omission in early production

Chunks and syntagms: The horizontal dimension

- What is the width/arity/'length' of early units
- Impression: infant has built up (e.g.) full caused-motion utterances, then abstracts i.o to form paradigms.
- Starting big: infants learn chunks, later break them down and do 'blame assignment' (Tomasello 2003)
- Problems with starting big
  - Storage of unstructured phonological wholes is problematic
  - Profile early SVO/SV/VO-productions (Theakston et al. 2012)
  - Argument omission in early production

Proposal #2: Syntagmatization

- 'Longer' constructions emerge through a gradual build-up of the horizontal dimension of constructions as a by-product of processing

The linking problem

- (Beekhuizen, Bod & Verhagen 2014): All processes have to work at the same time: developmental continuity
  - Learning lexical *and* grammatical constructions
  - Syntagmatization and paradigmatization (abstraction)
- (Instrumental) assumption/idealization:
  Lexical cx > syntagmatize > paradigmatize (Chang 2006, Alishahi & Stevenson 2010)

The linking problem

- (Beekhuizen, Bod & Verhagen 2014): All processes have to work at the same time: developmental continuity
  - Learning lexical *and* grammatical constructions
  - Syntagmatization and paradigmatization (abstraction)
- (Instrumental) assumption/idealization:
  Lexical cx > syntagmatize > paradigmatize (Chang 2006, Alishahi & Stevenson 2010)
- In a sense true:
  - syntagmatization requires analyses using lexical constructions
  - paradigmatization requires syntagmatized constructions
- But not as consecutive stages.

The linking problem

- (Beekhuizen, Bod & Verhagen 2014): All processes have to work at the same time: developmental continuity
  - Learning lexical *and* grammatical constructions
  - Syntagmatization and paradigmatization (abstraction)
- (Instrumental) assumption/idealization:
  Lexical cx > syntagmatize > paradigmatize (Chang 2006, Alishahi & Stevenson 2010)
- In a sense true:
  - syntagmatization requires analyses using lexical constructions
  - paradigmatization requires syntagmatized constructions
- But not as consecutive stages.

Proposal #3: Processing continuity

- Mechanisms of abstraction, syntagmatization, and different varieties of form-meaning association operate simultaneously.

The model

The cycle of the model

- Model receives utterance paired with set of situations
- Model comes up with the best analysis for this utterance
- Used constructions are reinforced

Key innovations

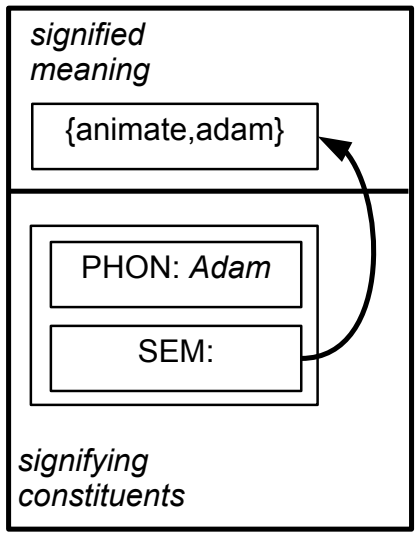
- All mechanisms of analysis and learning are available throughout time
- In analyzing, the model can concatenate multiple partial analyses
- These are starting point for novel syntagms
- Abstraction is a blind process
- (Stat.) pre-emption, semantic fit, prototype effects follow from same analysis mechanism.

- Representations learned from input:
  - Constructions, cf. construction grammar (Goldberg 1995)

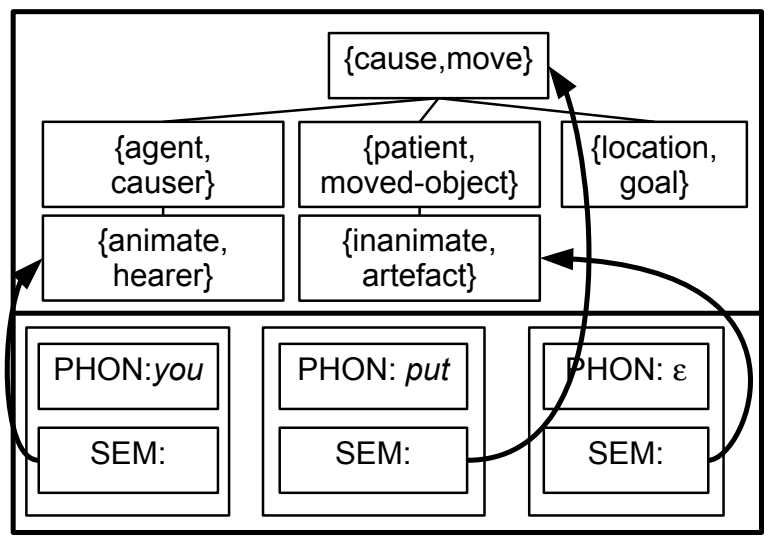
Definition

- Pairings of
  - a meaning
  - a string of constituents, each containing
    - a phonological form (possibly empty)
    - a semantic pointer

(a)



(b)



Analyzing

- Model tries to find which parts of utterance map to parts of a situation
- By creating derivations of constructions, using four interpretation mechanisms:
  - COMBINE: fill a constituent of one construction with another construction
  - CONCATENATE: create a list of derivations
  - BOOTSTRAP: fill a phonologically open constituent with an unknown word
  - IGNORE: don't integrate the word in the derivation
- Often many possibilities: select most probable one (see Beekhuizen et al. 2014)

Analyzing

- Model tries to find which parts of utterance map to parts of a situation
- By creating derivations of constructions, using four interpretation mechanisms:
  - COMBINE: fill a constituent of one construction with another construction
  - CONCATENATE: create a list of derivations
  - BOOTSTRAP: fill a phonologically open constituent with an unknown word
  - IGNORE: don't integrate the word in the derivation
- Often many possibilities: select most probable one (see Beekhuizen et al. 2014)

Analyzing

- Model tries to find which parts of utterance map to parts of a situation
- By creating derivations of constructions, using four interpretation mechanisms:
  - COMBINE: fill a constituent of one construction with another construction
  - CONCATENATE: create a list of derivations
  - BOOTSTRAP: fill a phonologically open constituent with an unknown word
  - IGNORE: don't integrate the word in the derivation
- Often many possibilities: select most probable one (see Beekhuizen et al. 2014)

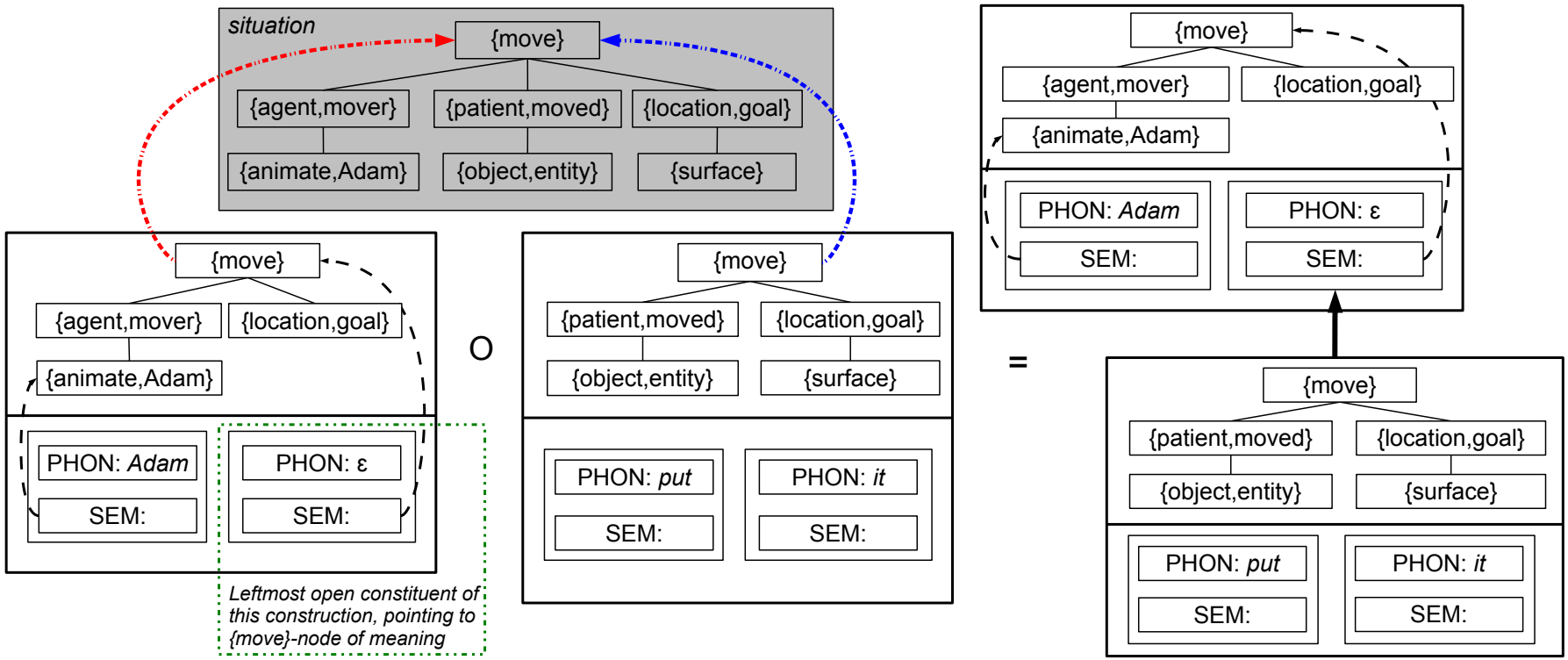


Figure: The COMBINATION mechanisms

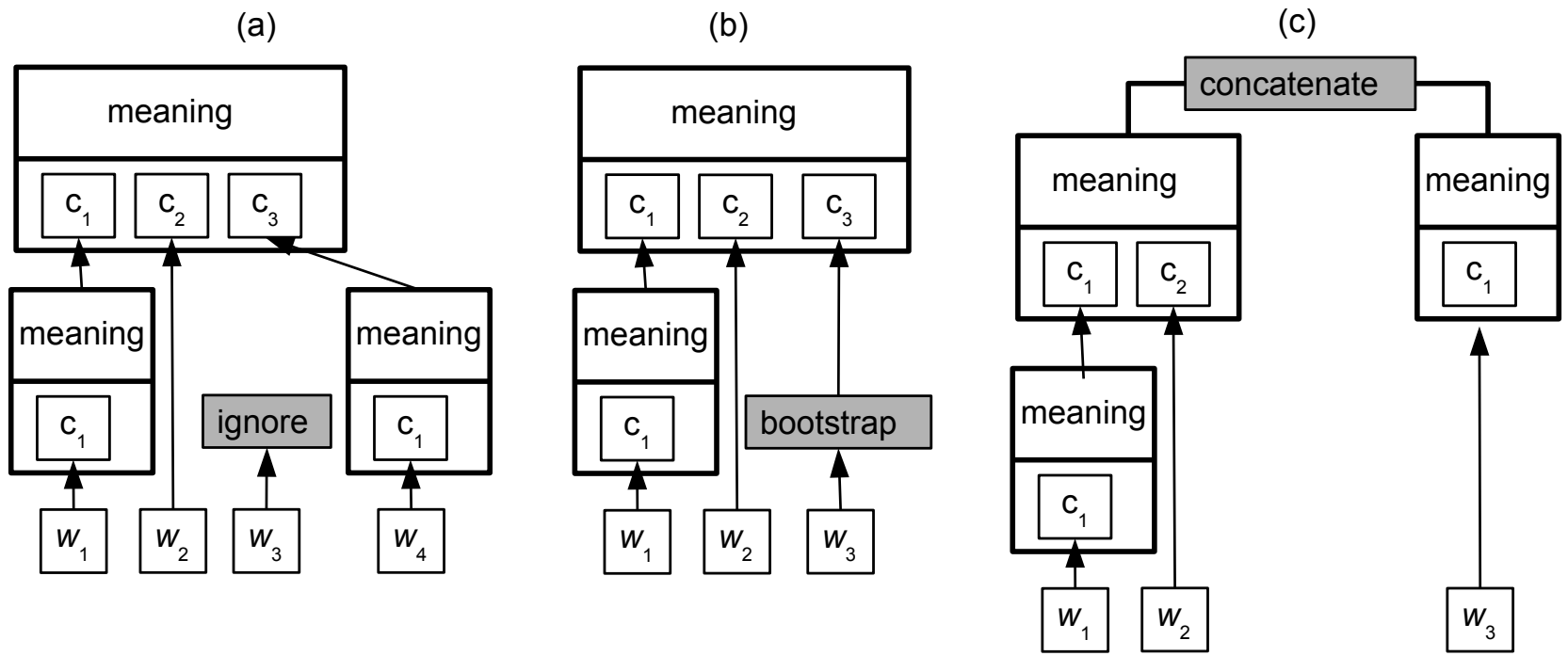


Figure: The IGNORE, BOOTSTRAP, and CONCATENATE mechanisms

- Learning on the basis of best analysis
- Idea of learning-as-processing (Langacker 2009)
- Idea of learned units being 'of the same matter' as input items (ibid.)

Four learning mechanisms

- ASSOCIATE parts of utterance and parts of a situation matching over recent experiences (cross-situational learning)
- REINFORCE used rules
- SYNTAGMATIZATION: store concatenation as a new construction
- PARADIGMATIZATION: add (more abstract) overlap between similar constructions to constructicon

- Learning on the basis of best analysis
- Idea of learning-as-processing (Langacker 2009)
- Idea of learned units being 'of the same matter' as input items (ibid.)

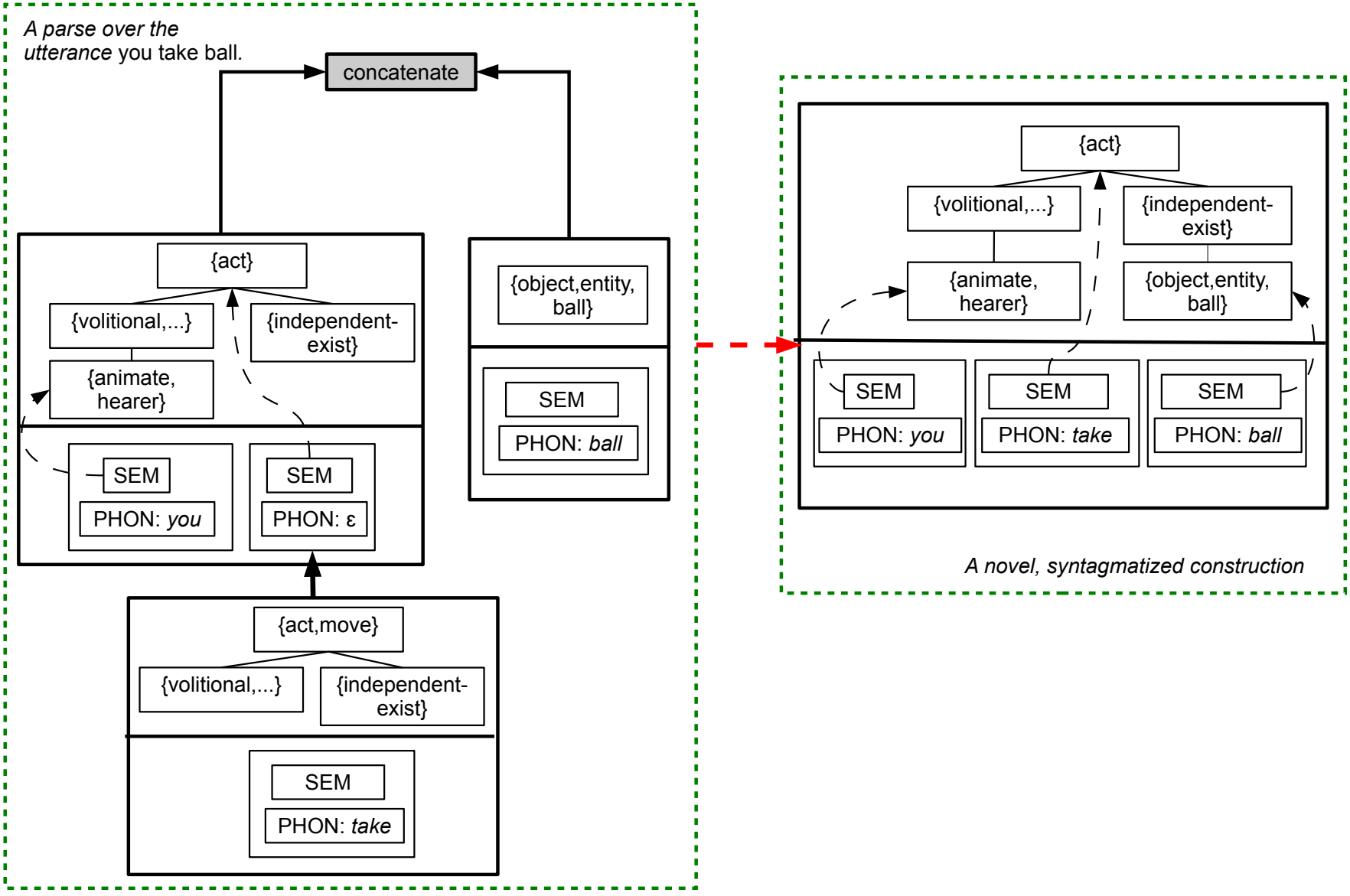
Four learning mechanisms

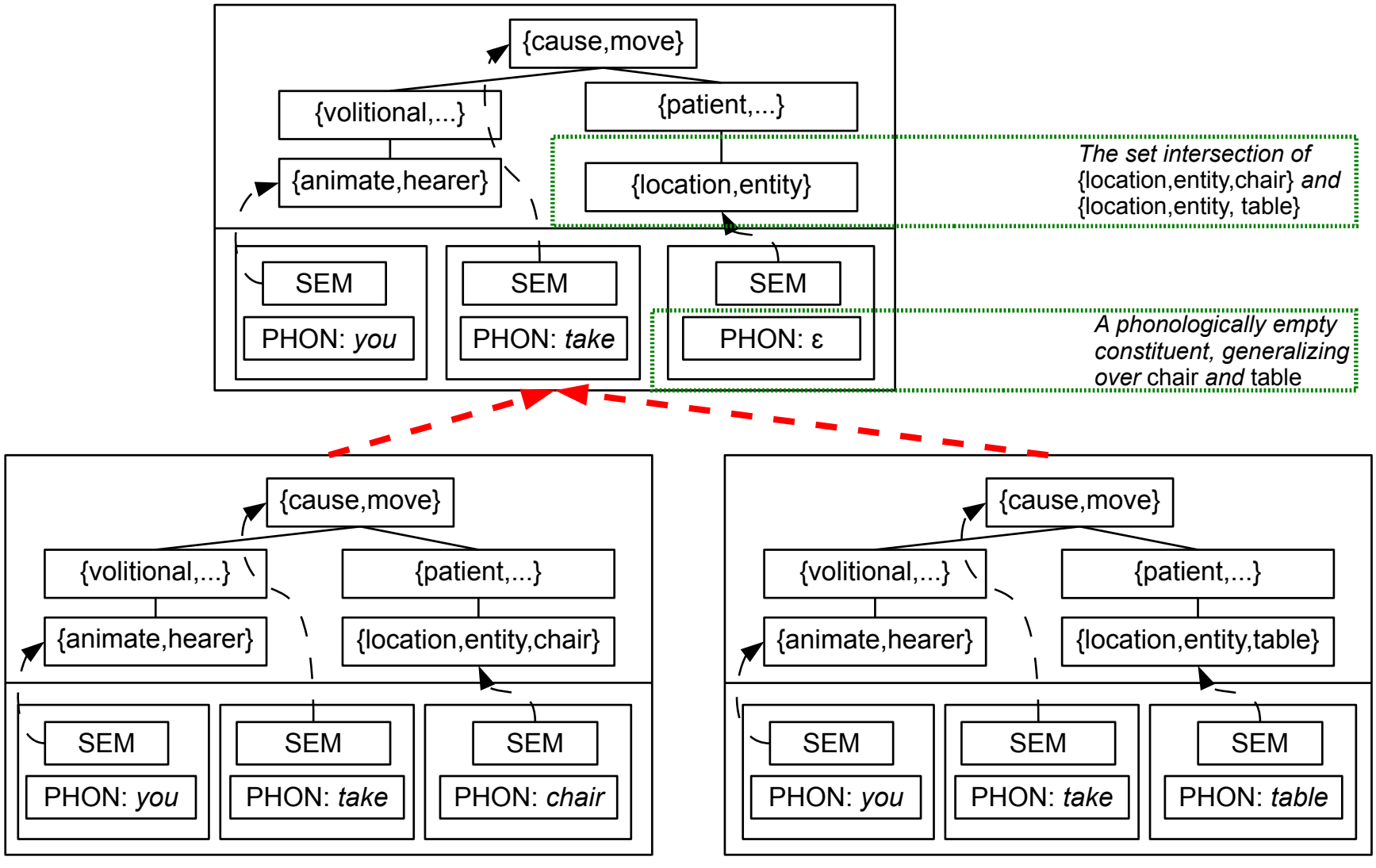
- Associate parts of utterance and parts of a situation matching over recent experiences (cross-situational learning)
- Reinforce used rules
- Syntagmatization: store concatenation as a new construction
- Paradigmatization: add (more abstract) overlap between similar constructions to constructicon

- Learning on the basis of best analysis
- Idea of learning-as-processing (Langacker 2009)
- Idea of learned units being 'of the same matter' as input items (ibid.)

Four learning mechanisms

- ASSOCIATE parts of utterance and parts of a situation matching over recent experiences (cross-situational learning)
- REINFORCE used rules
- SYNTAGMATIZATION: store concatenation as a new construction
- PARADIGMATIZATION: add (more abstract) overlap between similar constructions to constructicon


A parse over the utterance you take ball.
concatenate
{act}
{volitional,...}
{independent-exist}
{animate, hearer}
SEM
PHON: you
SEM
PHON: ε
{object,entity, ball}
SEM
PHON: ball
{act,move}
{volitional,...}
{independent-exist}
SEM
PHON: take
{act}
{volitional,...}
{independent-exist}
{animate, hearer}
{object,entity, ball}
SEM
PHON: you
SEM
PHON: take
SEM
PHON: ball
A novel, syntagmatized construction


{cause,move}
{volitional,...}
{patient,...}
{animate,hearer}
{location,entity}
The set intersection of {location,entity,chair} and {location,entity, table}
SEM
PHON: you
SEM
PHON: take
SEM
PHON: ε
A phonologically empty constituent, generalizing over chair and table
{cause,move}
{volitional,...}
{patient,...}
{animate,hearer}
{location,entity,chair}
SEM
PHON: you
SEM
PHON: take
SEM
PHON: chair
{cause,move}
{volitional,...}
{patient,...}
{animate,hearer}
{location,entity,table}
SEM
PHON: you
SEM
PHON: take
SEM
PHON: table

Experiments

Training

- Model incrementally presented with U, S pairs
- On the basis of Alishahi & Stevenson's (2010) generation procedure
- $|S| = 2$ (propositional uncertainty is 1)
- Non-correct $s \in S$ randomly generated
- 5 simulations of 2000 input items.

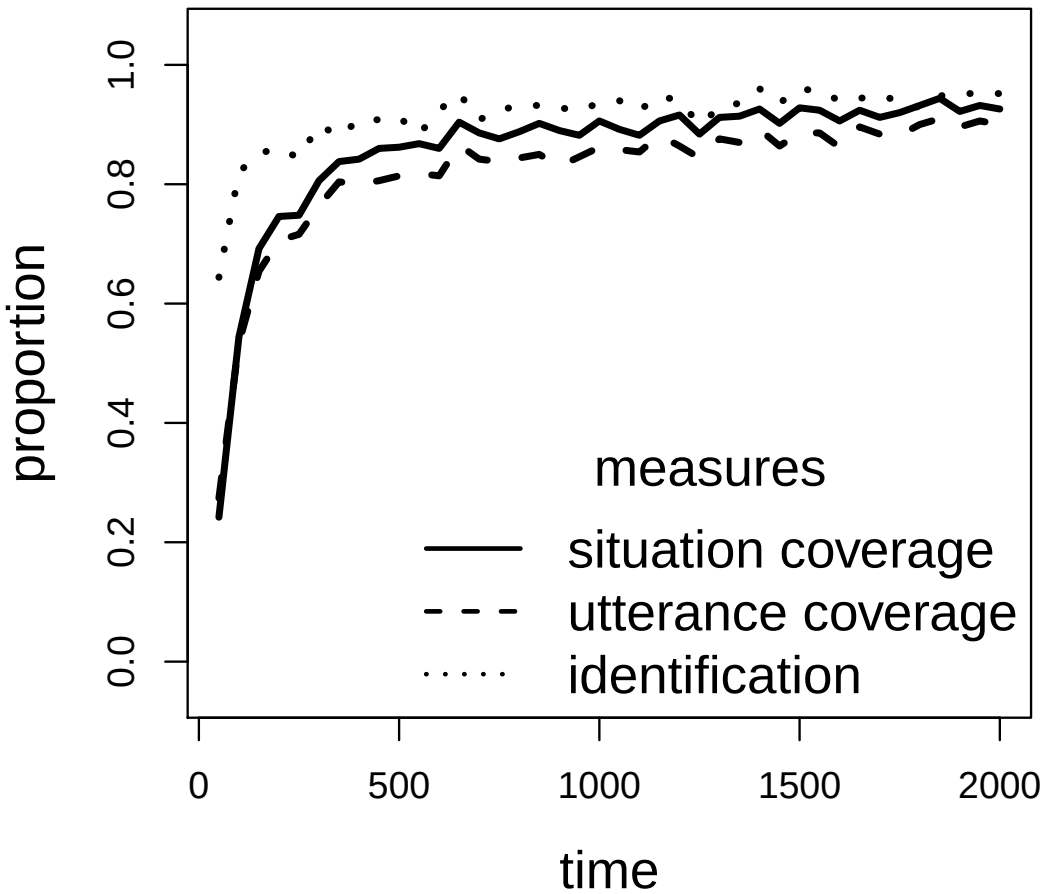


Figure: Comprehension scores over time.

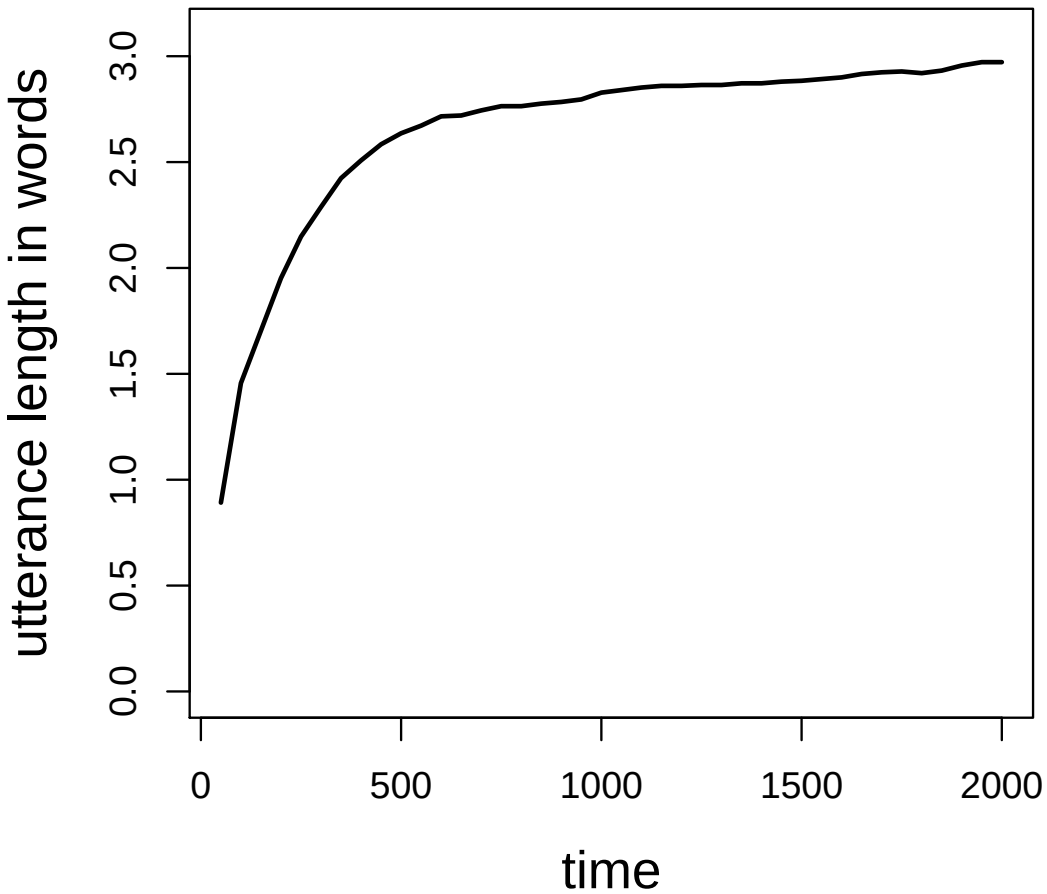


Figure: Mean length of U generated over time.

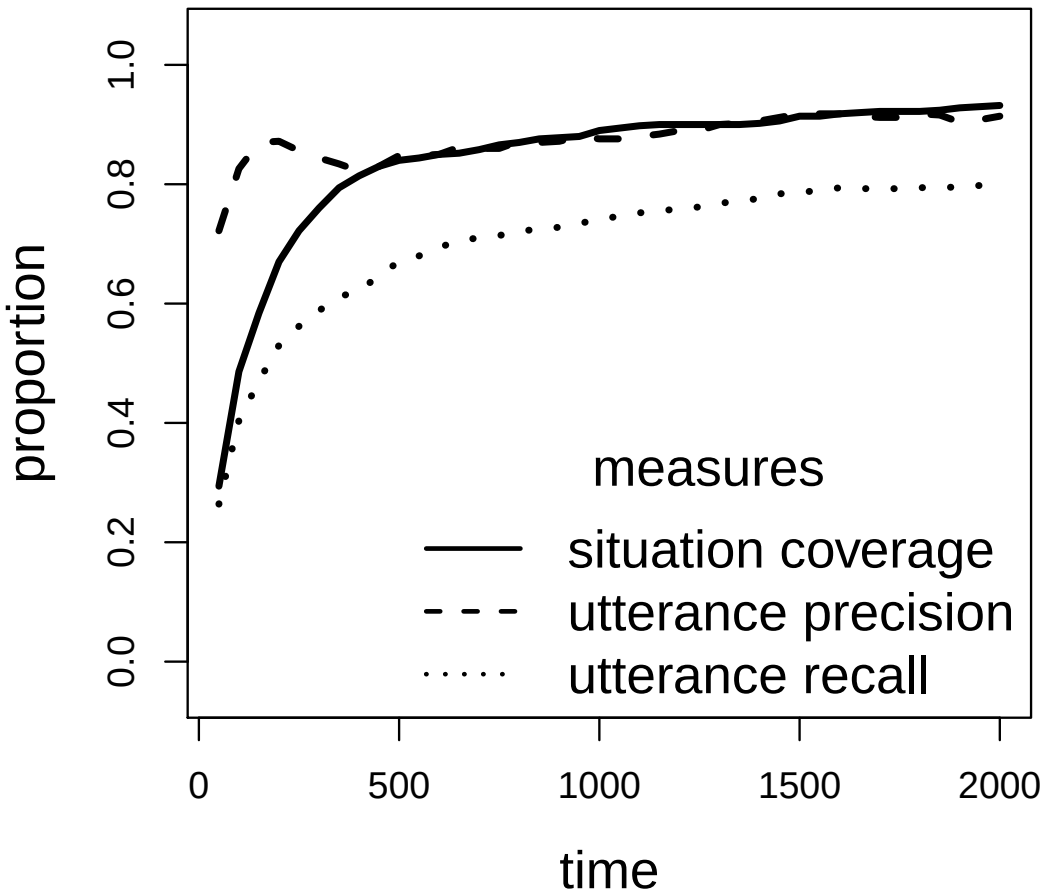


Figure: Generation scores over time.

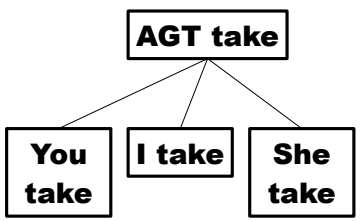
Table: Generations

time	generated production
50-450	[[she] put]
500	[she [put]]
550	[[she] [put] [in]]
600-900	[[she] put them [away]]
950	[[she] put [them]]
1000	[[she] put them [away]]
1050	[[she] put [them] away]
1400	[[she] put them away]

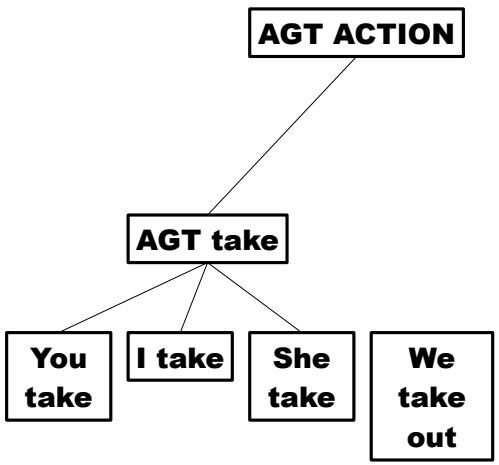
The growth of grammar

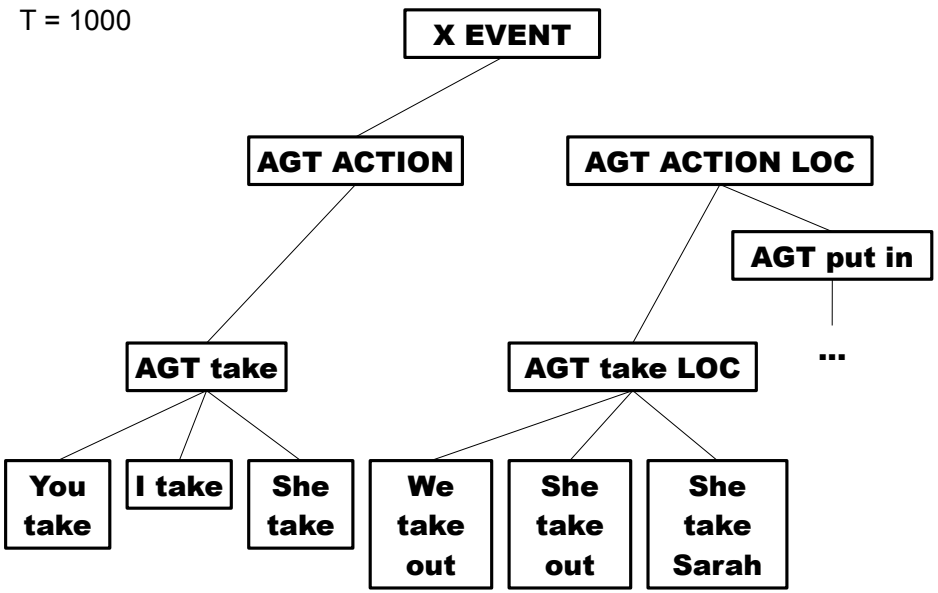
- A look under the hood
- How abstract are the constructions at a given point in time?
- How long are they?
- Case: constructions with *take* and their network

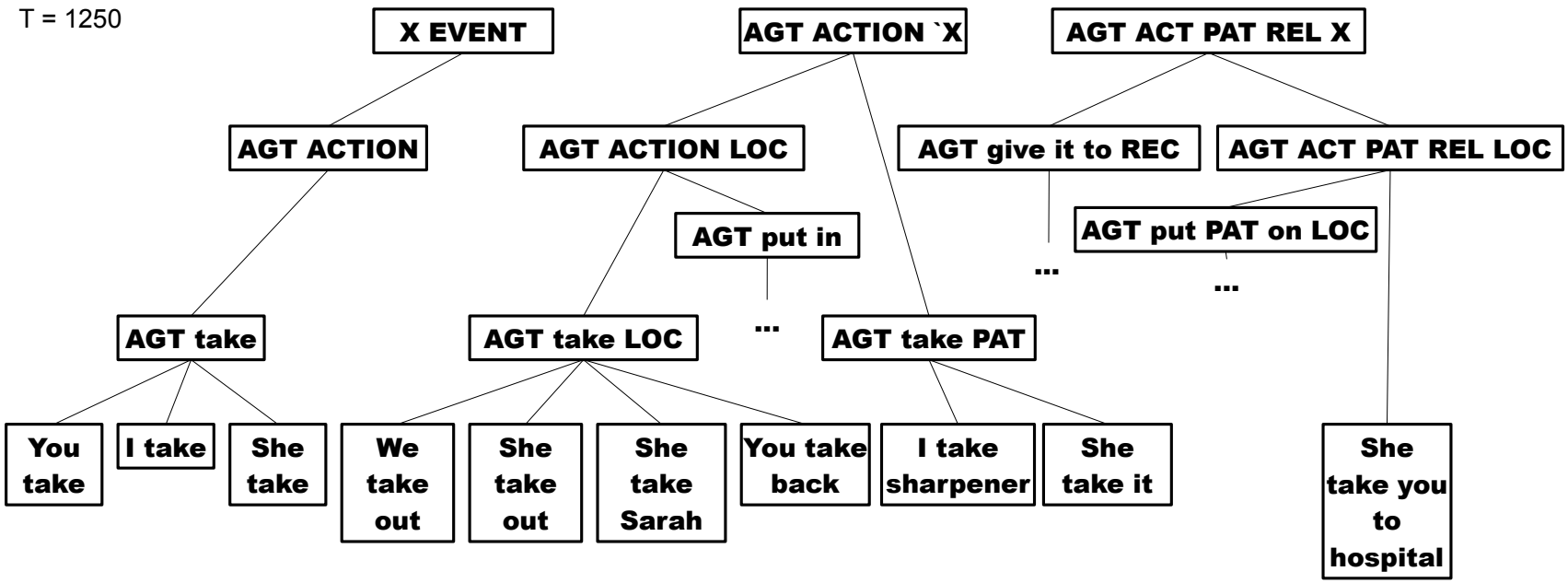
T = 500

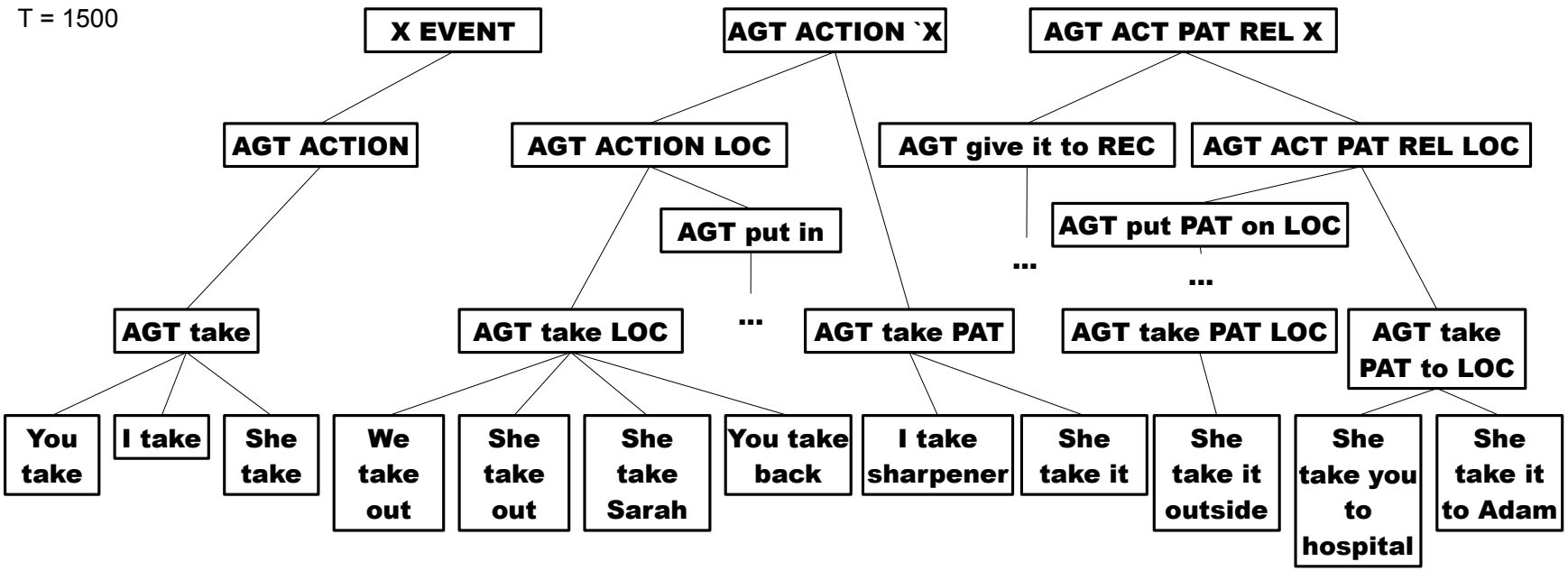


T = 750




T = 1000
X EVENT
AGT ACTION
AGT ACTION LOC
AGT put in
...
AGT take
AGT take LOC
You take
I take
She take
We take out
She take out
She take Sarah


T = 1250
X EVENT
AGT ACTION `X
AGT ACT PAT REL X
AGT ACTION
AGT ACTION LOC
AGT give it to REC
AGT ACT PAT REL LOC
AGT put in
AGT put PAT on LOC
...
...
AGT take
AGT take LOC
...
AGT take PAT
You take
I take
She take
We take out
She take out
She take Sarah
You take back
I take sharpener
She take it
She take you to hospital


T = 1500
X EVENT
AGT ACTION `X
AGT ACT PAT REL X
AGT ACTION
AGT ACTION LOC
AGT give it to REC
AGT ACT PAT REL LOC
AGT put in
AGT put PAT on LOC
...
...
AGT take
AGT take LOC
...
AGT take PAT
AGT take PAT LOC
AGT take PAT to LOC
You take
I take
She take
We take out
She take out
She take Sarah
You take back
I take sharpener
She take it
She take it outside
She take you to hospital
She take it to Adam

Theoretical points

- Abstraction as a by-product of processing
- Starting-small: syntagmatization
- Linking everything up

Theoretical points

- Abstraction as a by-product of processing
- Starting-small: syntagmatization
- Linking everything up

Empirical validation

- Gradual convergence in perception and production
- Constructions become increasingly long and abstract

Theoretical points

- Abstraction as a by-product of processing
- Starting-small: syntagmatization
- Linking everything up

Empirical validation

- Gradual convergence in perception and production
- Constructions become increasingly long and abstract

Raising new questions

- How is abstraction constrained under the 'by-product' view?
- Relation to Bybee's type frequency or Baayen's hapaxes
- Analyzing is now a rational decision making process: can we get rid of this legacy as well?

Thanks to:

- Rens Bod, Afsaneh Fazly, Suzanne Stevenson, and Arie Verhagen for many fruitful discussions
- The anonymous reviewers for their thoughtful comments,
- Afra Alishahi for providing me with the code of the generation procedure,
- NWO (Netherlands) for funding
- The organization and participants of the Coglingdays,