

# Modeling and theory

## Insights from the Syntagmatic-Paradigmatic Learner

Barend Beekhuizen

Leiden University & University of Amsterdam

22 September 2015

- 1 Overview of SPL
  - Representations
  - Processing
  - Learning
  
- 2 Main findings
  - Modeling issues, theoretical puzzles
  - Comprehension
  - Production
  
- 3 Competence and performance
  - Gaps in the theory
  - Comprehension
  - Representation
  - Production
  
- 4 Wrap-up

## 1 Overview of SPL

- Representations
- Processing
- Learning

## 2 Main findings

- Modeling issues, theoretical puzzles
- Comprehension
- Production

## 3 Competence and performance

- Gaps in the theory
- Comprehension
- Representation
- Production

## 4 Wrap-up

## The general problem

- Mapping form to meaning
- Acquisition:
  - Arriving at adult state
  - Explaining developmental waypoints
- Starting point: Usage-Based framework
- Computational cognitive model as method

## The Syntagmatic-Paradigmatic Learner

Flow of the model:

- 1 Model receives input item: pair of an **utterance** and a number of **situations**
- 2 Model tries to analyze using **processing mechanisms** and existing **representations**
- 3 Model updates grammar using **learning mechanisms** and **best analysis**
- 4 goto 1

## Representations

Constructions: pairings of signifiers and signifieds, both for 'grammar' and 'lexicon'

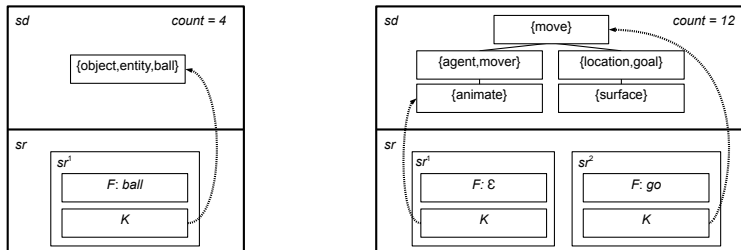


Figure: Constructions

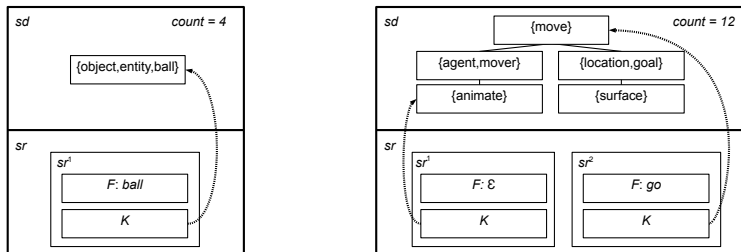


Figure: Constructions

1 [ BALL / *ball* ]

2 [ [ ANIMATE ] [ MOVE / *go* ] ] |

MOVE(AGENT(ANIMATE), LOCATION(SURFACE))



## Representations

Constructions: pairings of signifiers and signifieds, both for 'grammar' and 'lexicon'

## Processing

An utterance in a situational context is analyzed using the set of known constructions and processing mechanisms

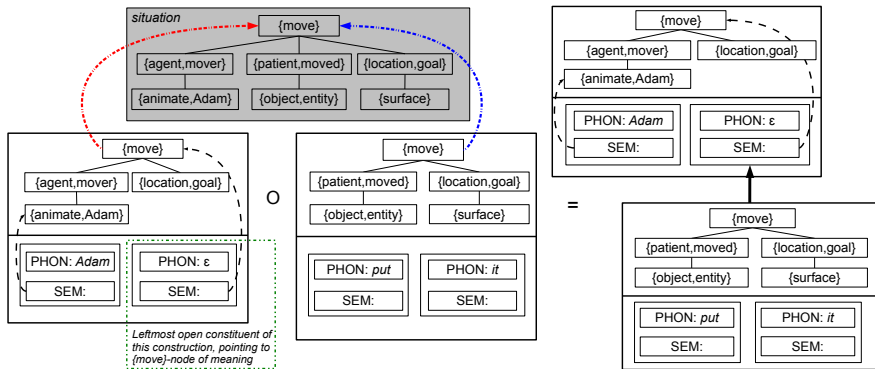
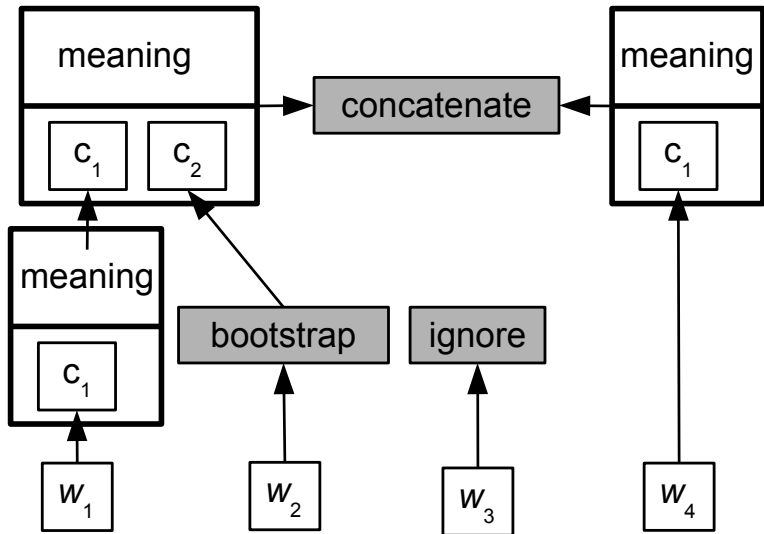


Figure: Combine



## Representations

Constructions: pairings of signifiers and signifieds, both for 'grammar' and 'lexicon'

## Processing

An utterance in a situational context is analyzed using the set of known constructions and processing mechanisms.

Often many analyses possible, so find best one:

- Most frequently encountered constructions
- With fewest concatenate, bootstrap, and ignore operations.

## Representations

Constructions: pairings of signifiers and signifieds, both for 'grammar' and 'lexicon'

## Processing

An utterance in a situational context is analyzed using the set of known constructions and processing mechanisms.

Often many analyses possible, so find best one:

- Most frequently encountered constructions
- With fewest concatenate, bootstrap, and ignore operations.

## Learning

Best analysis leaves trace in memory: 5 learning mechanisms.

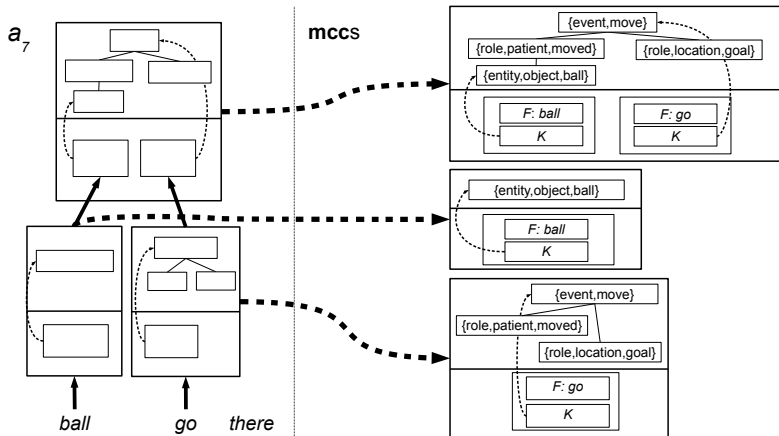


Figure: Adding most concrete constructions

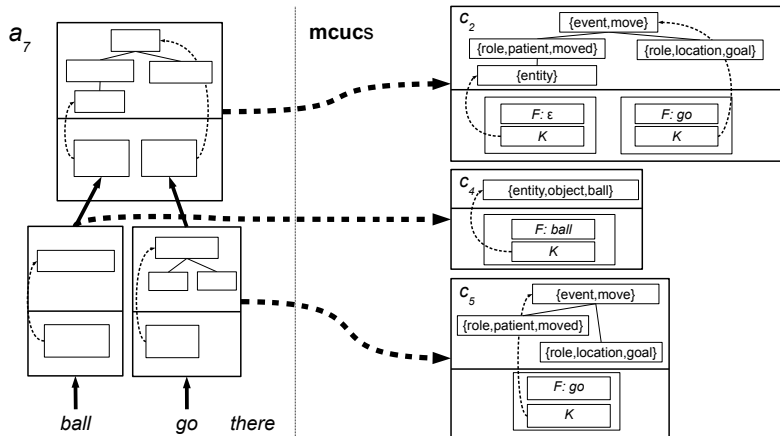


Figure: Updating the most concrete used constructions

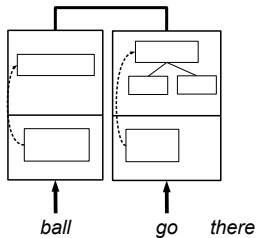
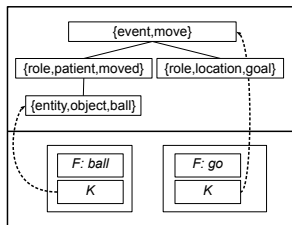
$a_5$  $c_{\text{syn}}$ 

Figure: Syntagmatization



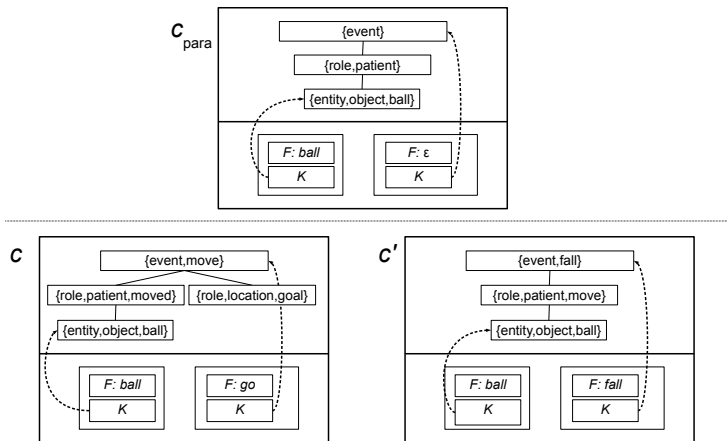


Figure: Paradigmatization

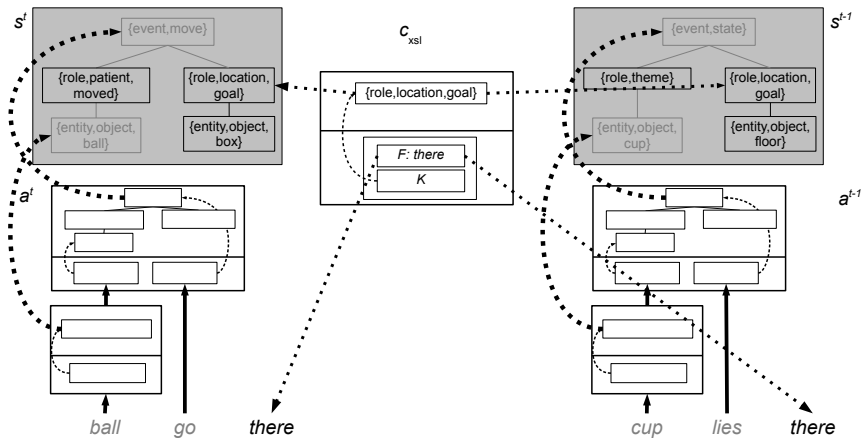


Figure: Cross-situational learning

- 1 Overview of SPL
  - Representations
  - Processing
  - Learning
  
- 2 Main findings
  - Modeling issues, theoretical puzzles
  - Comprehension
  - Production
  
- 3 Competence and performance
  - Gaps in the theory
  - Comprehension
  - Representation
  - Production
  
- 4 Wrap-up

## SPL resolves some a priori issues (chapter 2)

- **Comprehensiveness**: comprehension and production
- **Simultaneity**: lexical and grammatical constructions

## SPL resolves some a priori issues (chapter 2)

- Comprehensiveness: comprehension and production
- Simultaneity: lexical and grammatical constructions
- Reappraisal of the **starting-small** approach
- Learning as by-product of processing (**immanence**)
- Reappraisal of the **competence-performance** distinction

## Comprehension (chapter 5, 6)

- **Robustness**: making sense of utterance despite knowing little (using concatenation, bootstrapping)
- **Increasing coverage** of utterance and situation
- **Increasing accuracy** of picking out situation from 6 candidates
- **Varying mechanisms**: XSL precedes bootstrapping; bootstrapping dominates.

## Production (chapter 7)

- **Experiment:** give model situation, ask to produce utterance
- **Increasing length** of produced utterance
- **Hardly any errors of comission**

- 1 Overview of SPL
  - Representations
  - Processing
  - Learning
  
- 2 Main findings
  - Modeling issues, theoretical puzzles
  - Comprehension
  - Production
  
- 3 Competence and performance
  - Gaps in the theory
  - Comprehension
  - Representation
  - Production
  
- 4 Wrap-up



- Linguistic knowledge grounded in language use
- So we can reason from child's productions to its knowledge of language

- Linguistic knowledge grounded in language use
- So we can reason from child's productions to its knowledge of language
- However:
  - **Sample** may not contain reflection of full potential
  - **Other reasons** for not producing some linguistic item
  - **Interactivity** of components **invalidates** line of reasoning

- Linguistic knowledge grounded in language use
- So we can reason from child's productions to its knowledge of language
- However:
  - Sample may not contain reflection of full potential
  - Other reasons for not producing some linguistic item
  - Interactivity of components invalidates line of reasoning
- So: need to **account** for a linguistic competence and performance within Usage-Based framework.
- And show its **explanatory value**.
- **Not unique** to SPL: all UB computational models do so. However, **interaction lexical/grammatical** acquisition gives interesting effects

## Comprehension

- Early abstraction
- Increasing use of more concrete constructions
- Does not entail loss of abstraction (to the contrary)

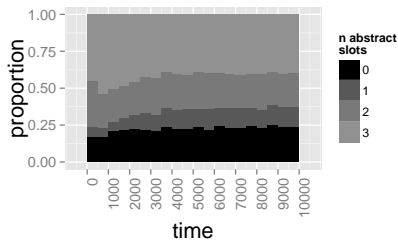
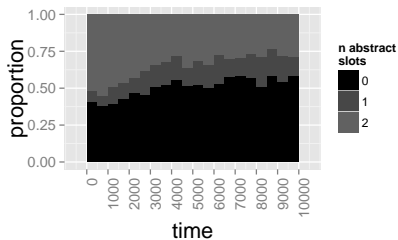


Figure: Abstraction of used length-2 and 3 constructions

## Reflections of use on representations

- Look under the hood to obtain a fuller understanding of representational potential
- Abstractions are there, but not used so much

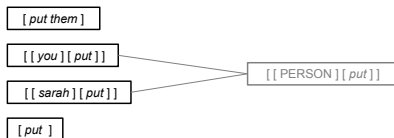


Figure: After 100 input items

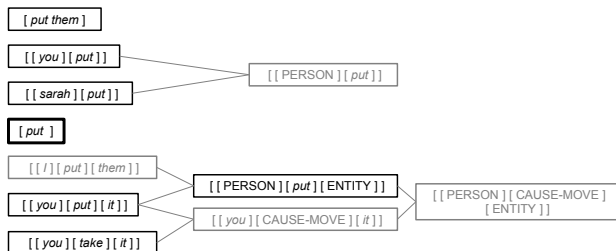


Figure: After 500 input items

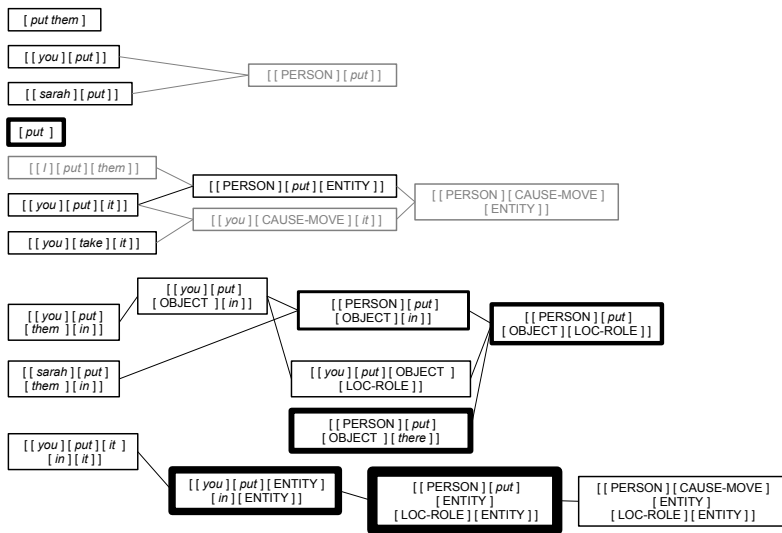


Figure: After 10000 input items



## Linguistic competence in production

- Wysiwyg?
- No:
  - Lexical items may be known but not produced because grammatical constructions are not known yet
  - Interaction: competition between grammatical constructions

## The unexpressed expressables

Words that are **known** but nonetheless **not produced**, because there is (1) an erroneous word outcompeting them or (2) there is no grammatical construction to 'host' them.

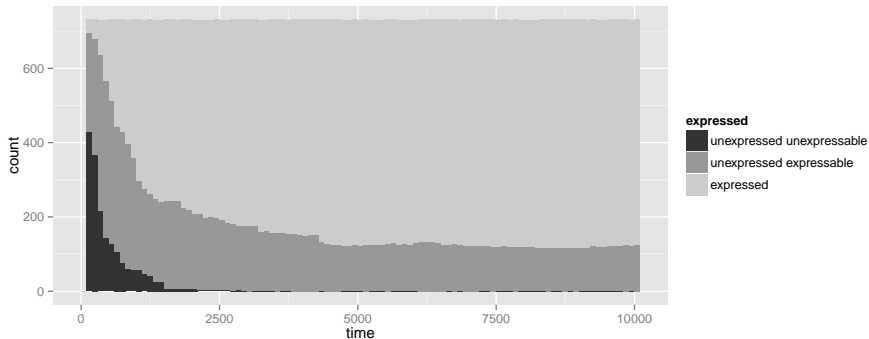


Figure: The expression of 'second arguments' over time

# Wrap-up

## Why we need to focus on competence/performance

- Corpora hide potential for abstraction
- Simultaneity effects hide potential
  - Lexical knowledge hidden (unexpressed expressables)
  - Paradoxal blocking effects

Thank you