

Experiments on grammaticalization

Remi van Trijp & Pieter Wellens
<http://emergent-languages.org/>

Summer Atelier on Computational Construction Grammar
Erice, Italy - 18 July 2007

Sony CSL

Sony Computer Science Laboratory Paris
sony computer science laboratory Paris

Contents

- Language systematicity [\sim (self-)organization of the construction]
- Pattern formation
 - Luc Steels, Remi van Trijp & Pieter Wellens (in press). **Multi-level selection in the emergence of language systematicity**. In: F. Almeida e Costa et al. Proceedings of the ninth ECAL conference. LNAI 4648. Berlin: Springer-Verlag.
- Emergence of semantic roles
(experiments on case grammar)

A selectionist model...

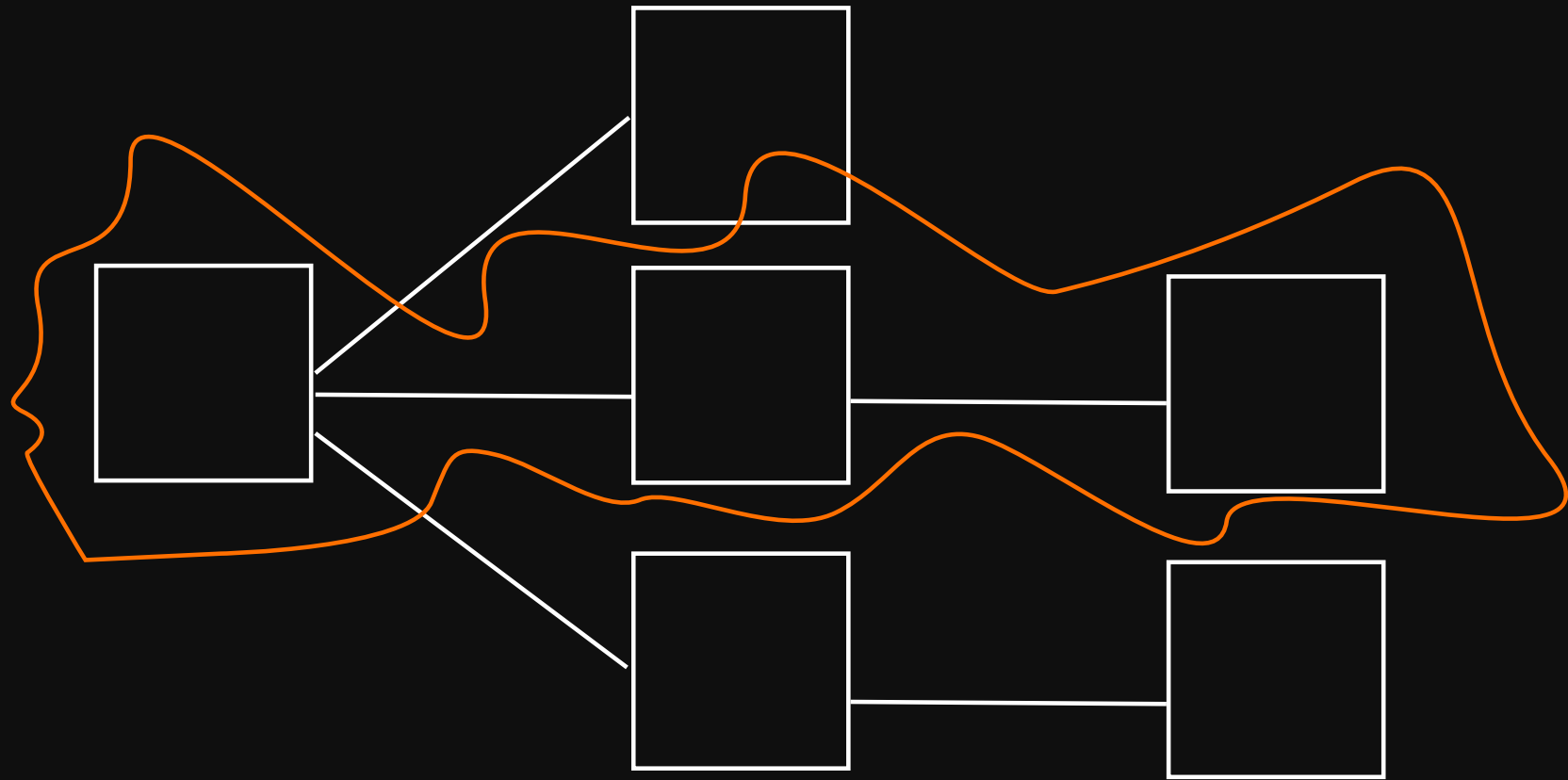
- Genetic transmission
- Cultural transmission (across generations)

Simon Kirby (2000). **Syntax without natural selection: how compositionality emerges from vocabulary in a population of learners.**

- Cultural transmission (across interactions)
 - > Language change at **each** interaction
 - > functional and frequency-based forces
 - > Language as a problem-solving activity

J. De Beule & B. Bergen (2006). **On the emergence of compositionality.**

Pattern formation



Demo!

Pattern formation

(book ?obj)

gerofi

(take ?x ?y)

takka

(take ?x ?y)
(book ?y)

gerofi
prendra

(take ?x ?y)
(book ?y)

wogglesplat
prendra

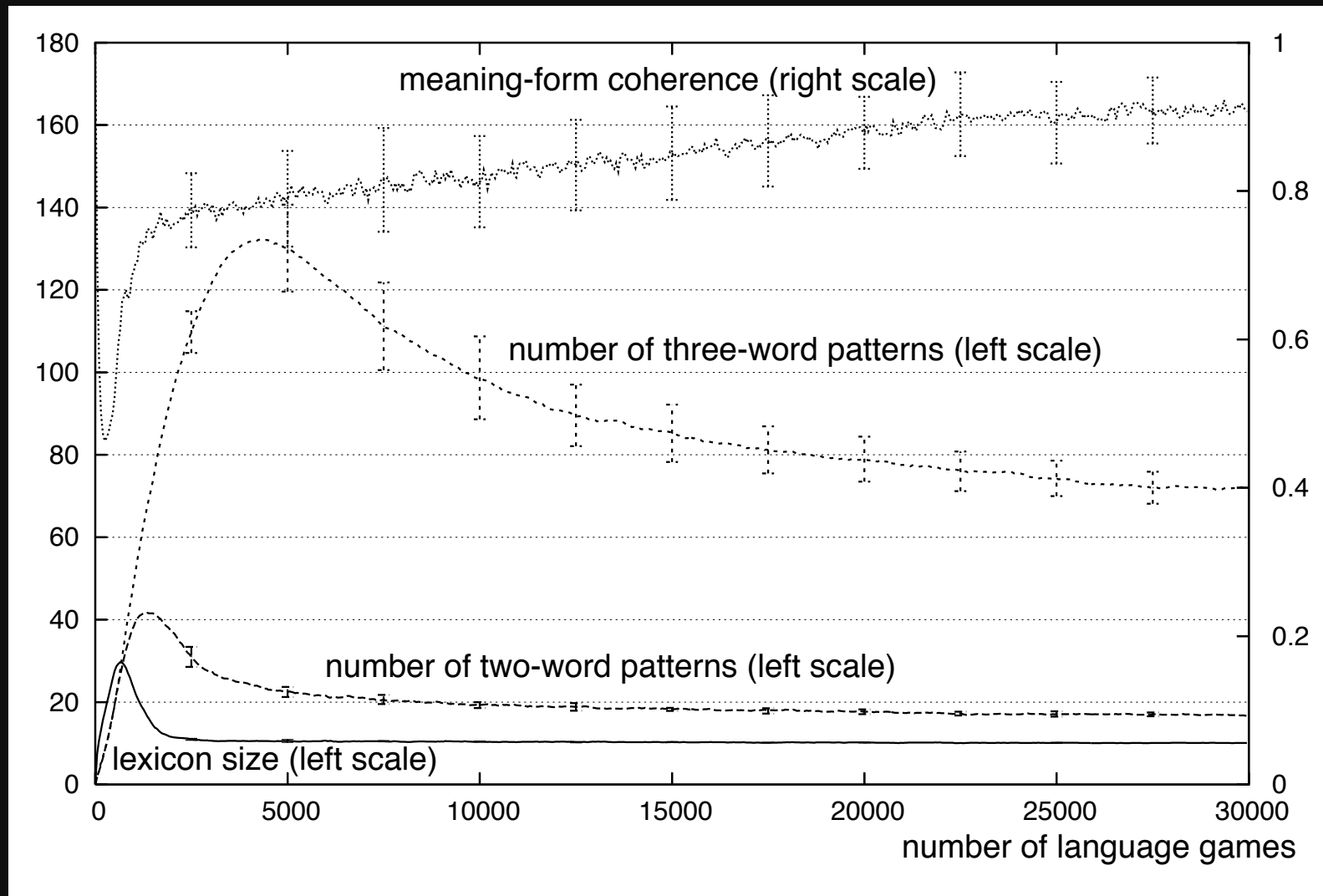
(book ?obj)

wogglesplat

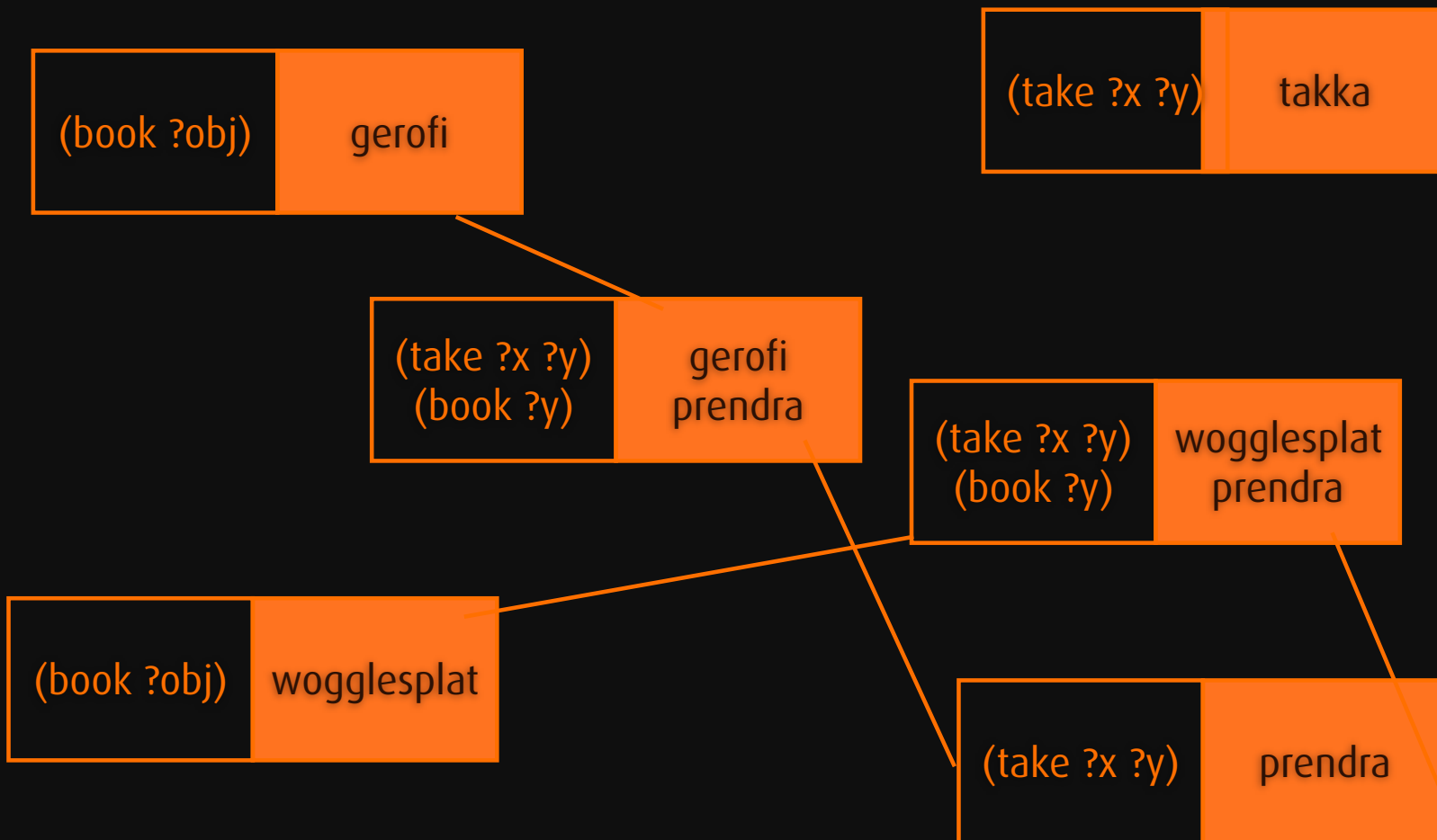
(take ?x ?y)

prendra

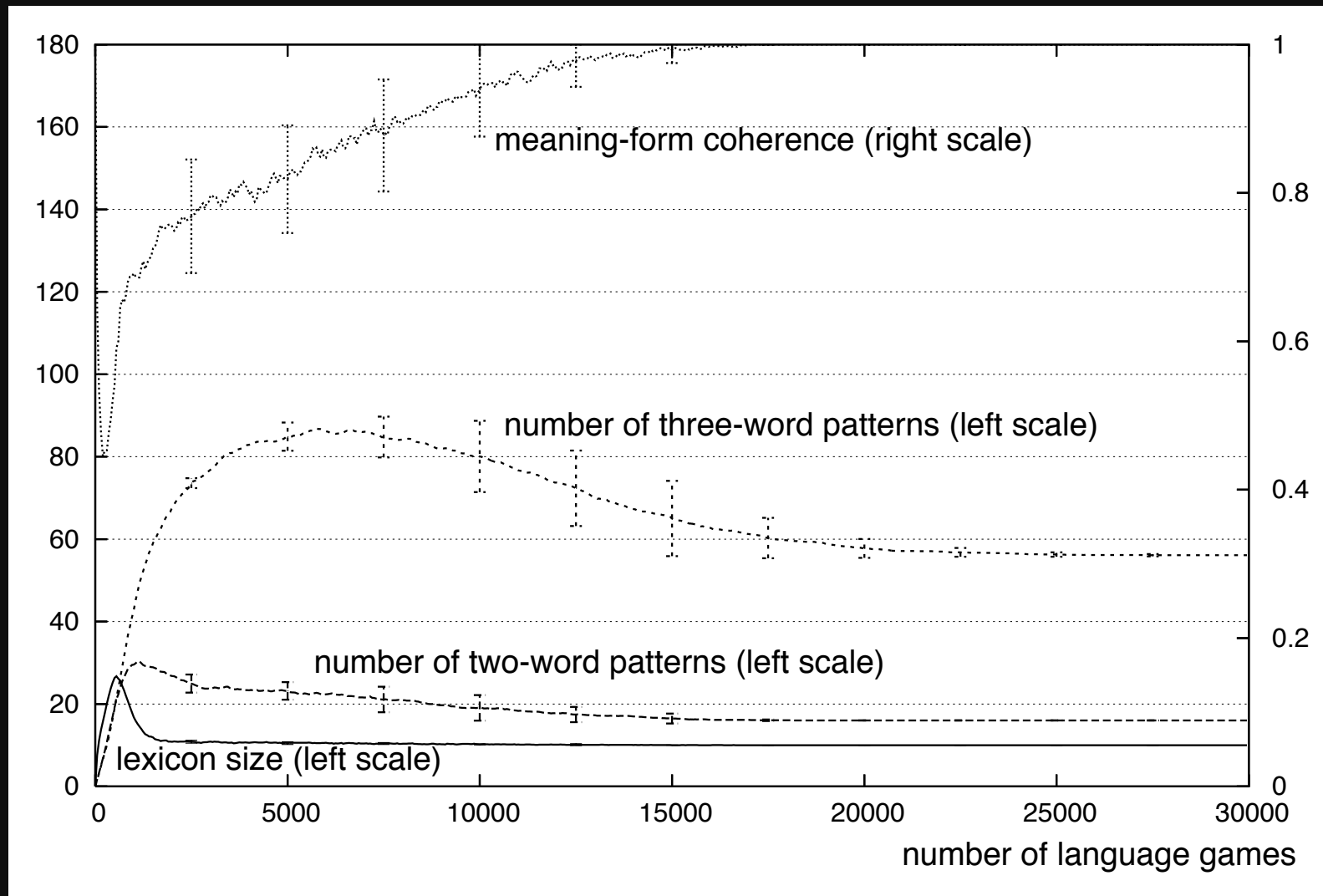
Pattern formation



Pattern formation



Pattern formation



Organization of the constructicon

- Early experiments on lexicon formation:
 - > Lateral inhibition of direct competitors
- Conflicts in rule application
- Fine-grained scoring within rules
- Multiple levels of linguistic organization:
 - > Sub- and super-links need to be kept
- Other experiments/types of links:
 - > co-occurrence links (preferred paths)
 - > schematicity links (future work)

Case Grammar

- Marking case relations is useful for...
 - > avoiding ambiguity / multiple interpretations
 - > decontextualizing
 - > optimizing language processing
- “The relation between form and meaning is rather indirect and multi-layered [...] First, there are formal influences of valency and constructions on cases, which can also be observed diachronically. Second, the dative/IO can have not only pragmatic or discourse meaning but also a whole range of submeanings varying cross-linguistically and diachronically. It is not always easy to link these various senses, let alone find a unifying principle.”
(Van Langendonck, W. and W. Van Belle (eds.) (1998) *The Dative*. Volume 1. Volume 2. *Theoretical and Contrastive Studies*. p. xvi

Case Grammar

- [T]aken to its logical conclusion, a fine-grained analysis of thematic roles will result in a unique case frame for almost every verb in a natural language. I believe that this reductio ad absurdum is the correct grounding point for the analysis of case, however: what it implies is that the definitions of the thematic roles must mesh in some natural way with the lexical semantics of the verbs that govern them. It is obvious that, although semantic thematic ROLES may be verb specific, surface morphosyntactic case MARKINGS in natural languages gather together large classes of thematic roles since there are thousands of verbs with thousands of roles but never more than fifty to eighty case markers and at the very most just a dozen or so case markers other than the spatiotemporal ones. Thus, surface case marking imposes structure on thematic relations to an even more abstract degree than verb roots impose structure on the human experience of events.

> *William Croft (1991). Syntactic categories and grammatical relations.*
p. 151

The vision system

- Luc Steels & Jean-Christophe Baillie (2003). **Shared grounding of event descriptions by autonomous robots.**



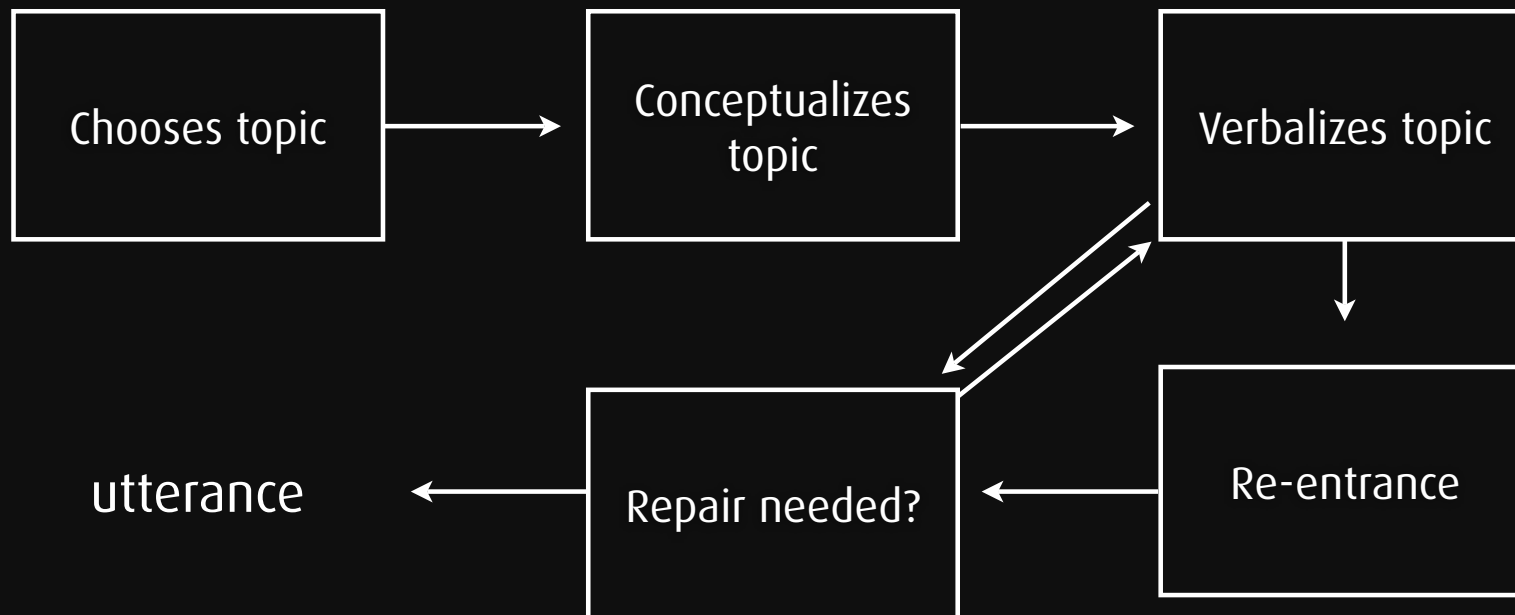
((take ev-155626 true) (take-1 ev-155626 Jack)
(take-2 ev-155626 block-1) (take-3 ev-155626 Jill))

The vision system / world model

- See demo-file...

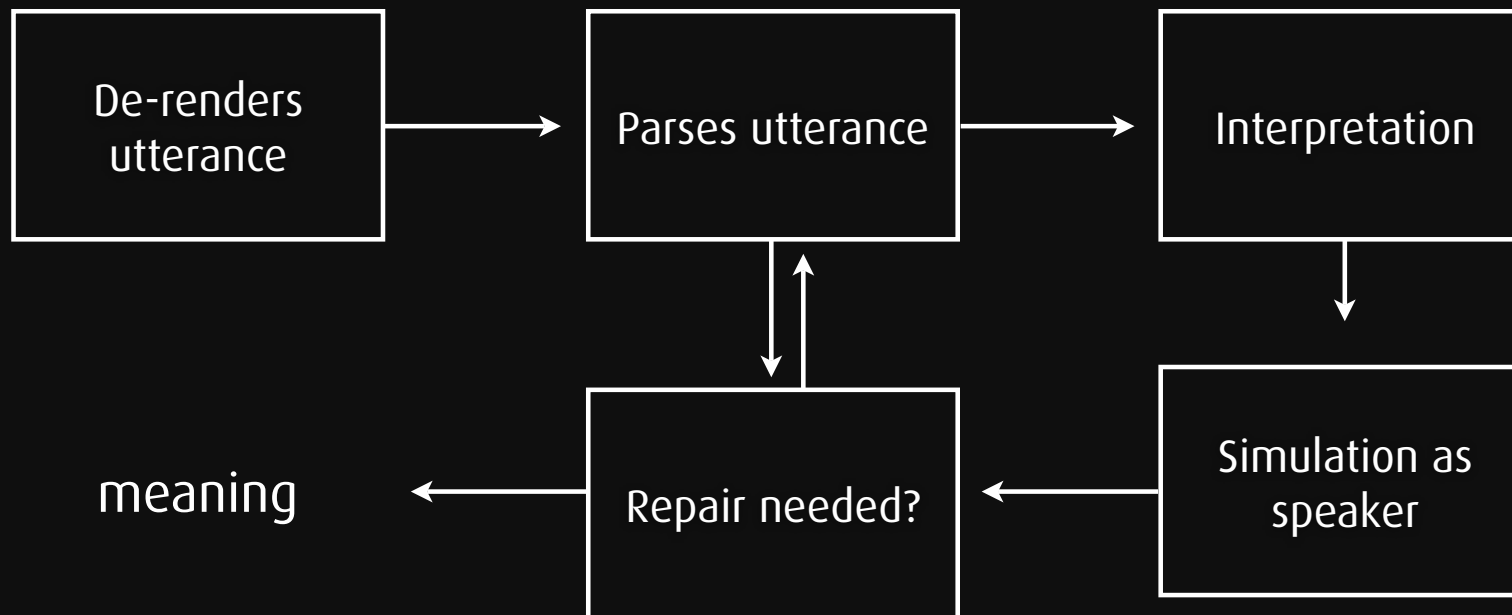
Description games

Speaker



Description games

Hearer



Speaker strategy

- Given a specific event-object relation R and word W introducing the object...
- If there is an existing marker for marking this relation, use it...
- If there is none, try to reuse other existing markers through analogical reasoning or invent a new one...

First time marker...

- Participant role under investigation: move-outside-1
- Take all micro-events from the event structure that involve the filler of move-outside-1



(visible-1 ev-1 jill)
(visible-1 ev-2 jill)
(increase-distance-1 ev-3 jill)
(move-1 ev-4 jill)
(touch-1 ev-5 jill)
(touch-1 ev-6 jill)

Reuse through analogy



move-outside-1

move-inside-1

appear

move-away

approach

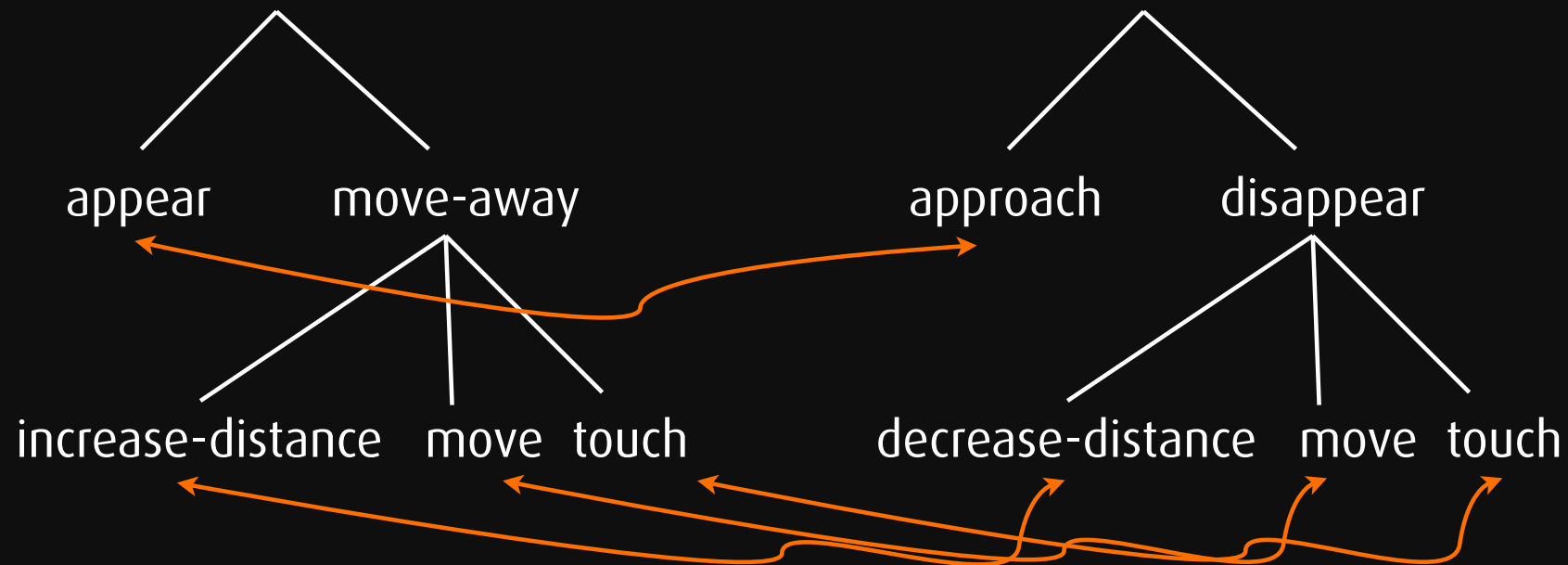
disappear

increase-distance

move touch

decrease-distance

move touch



Challenges increase enormously...

- “Jack bo walk.”
- These markers do not “float around” somewhere, they occur **within constructions**

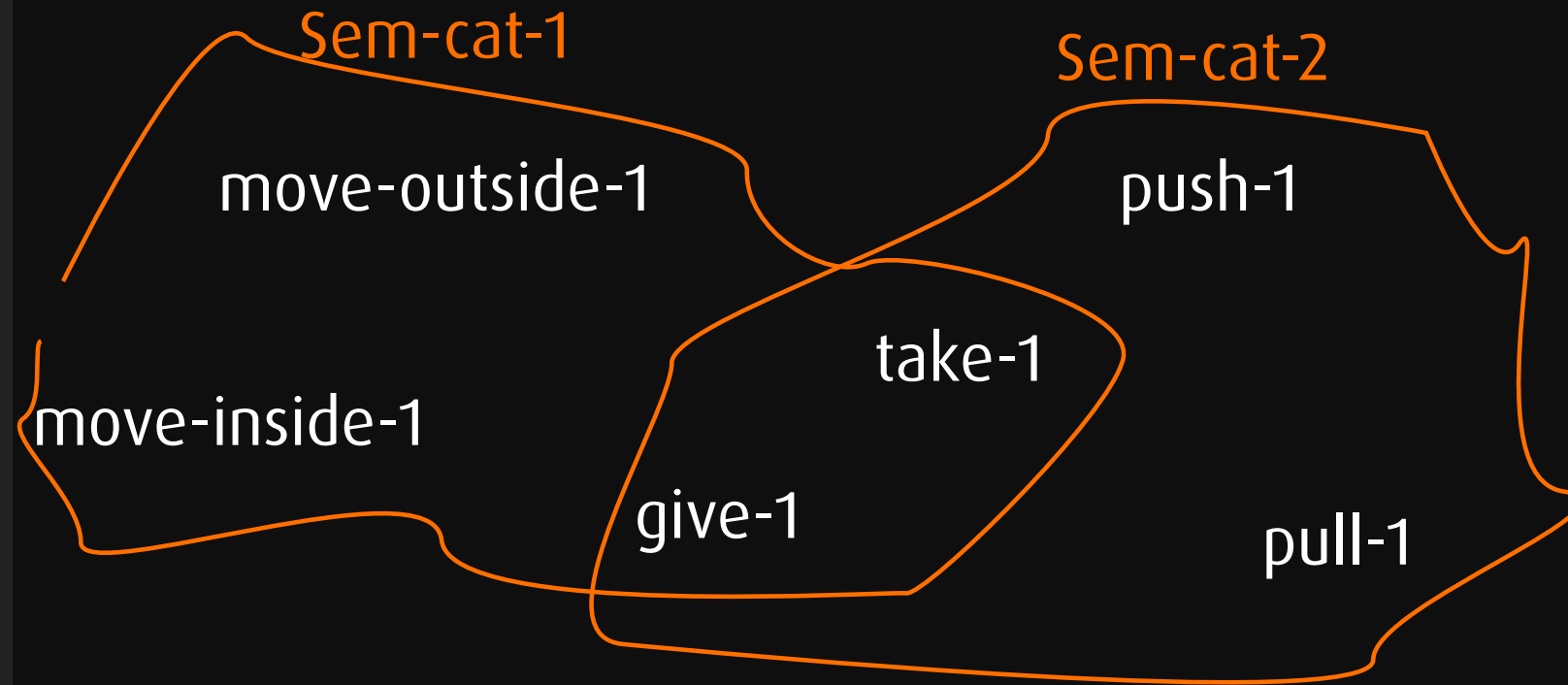
<pre>((?top (sem-subunits (== ?unit-1 ?unit-2))) (?unit-1 (case-frame (== (sem-cat-1 ?obj-x)))) (?unit-2 (referent ?obj-x)) ((J ?new-unit ?top)) ((J ?marker-unit ?new-unit)))</pre>	<-->	<pre>((?top (syn-subunits (== ?unit-1 ?unit-2 ?marker-unit)) (form (== (meets ?unit-2 ?marker-unit)))) (?marker-unit (syn-cat syn-cat-3)) ((J ?new-unit ?top)))</pre>
--	------	---

Challenges increase enormously...

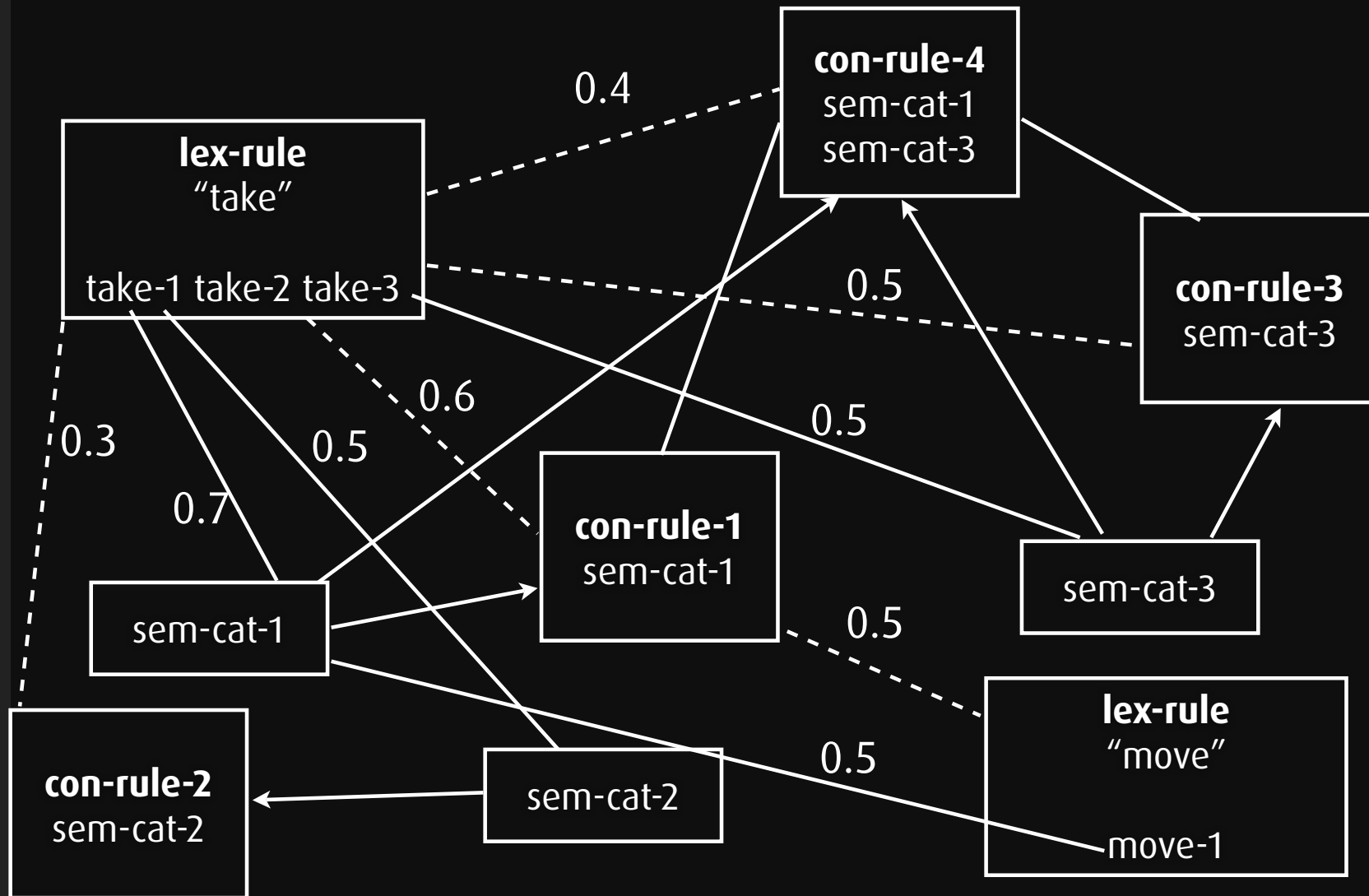
- “Jack bo move-outside house ka.”
- Multilevel selection needed when combining constructions into larger constructions...
- Additionally, when two semantic roles co-occur in one pattern, they lose their independency of one another...
- So combining constructions involves these roles adapting their “meaning/function” with respect to each other...

Challenges increase enormously...

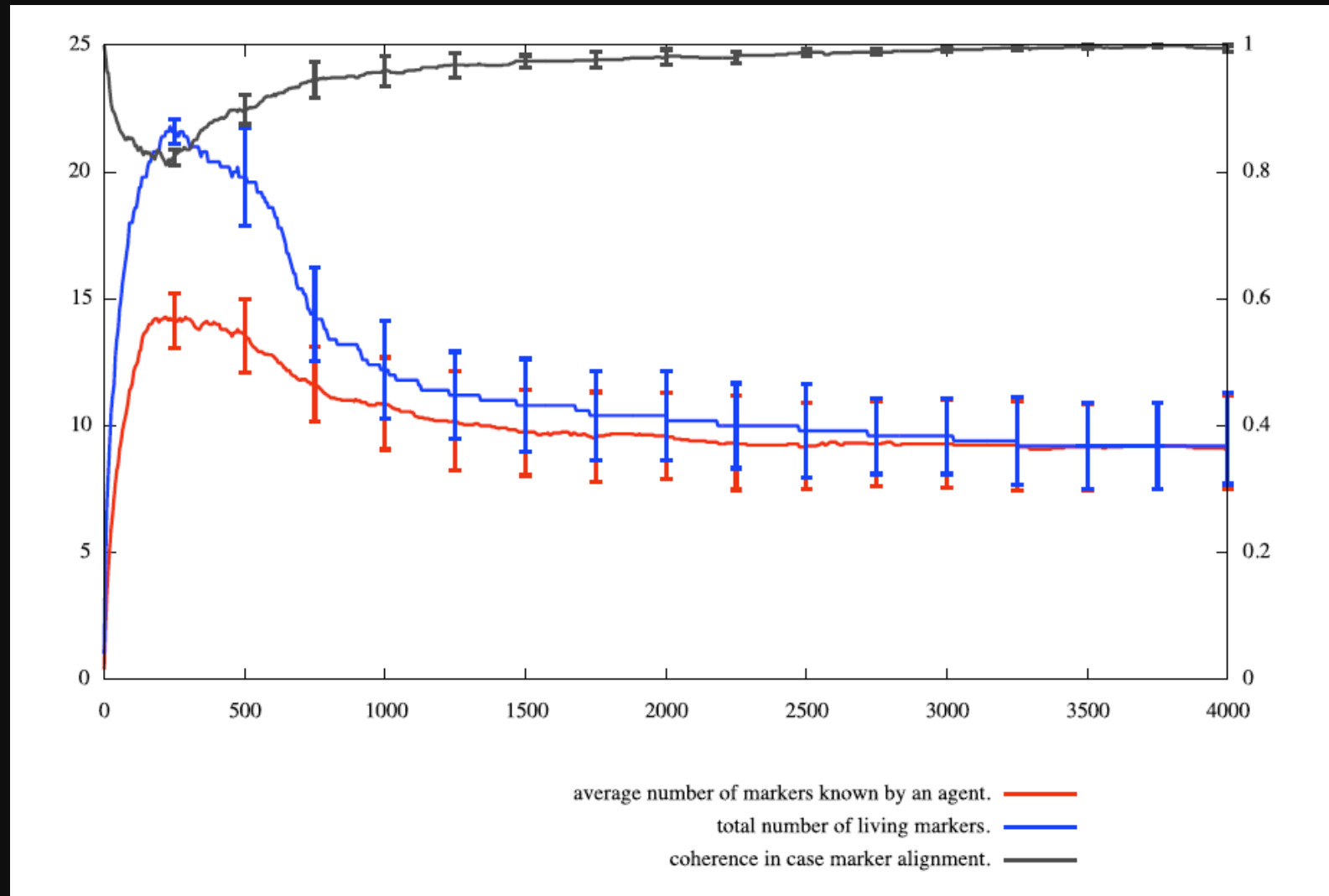
- Moreover, roles may be in conflict with each other...



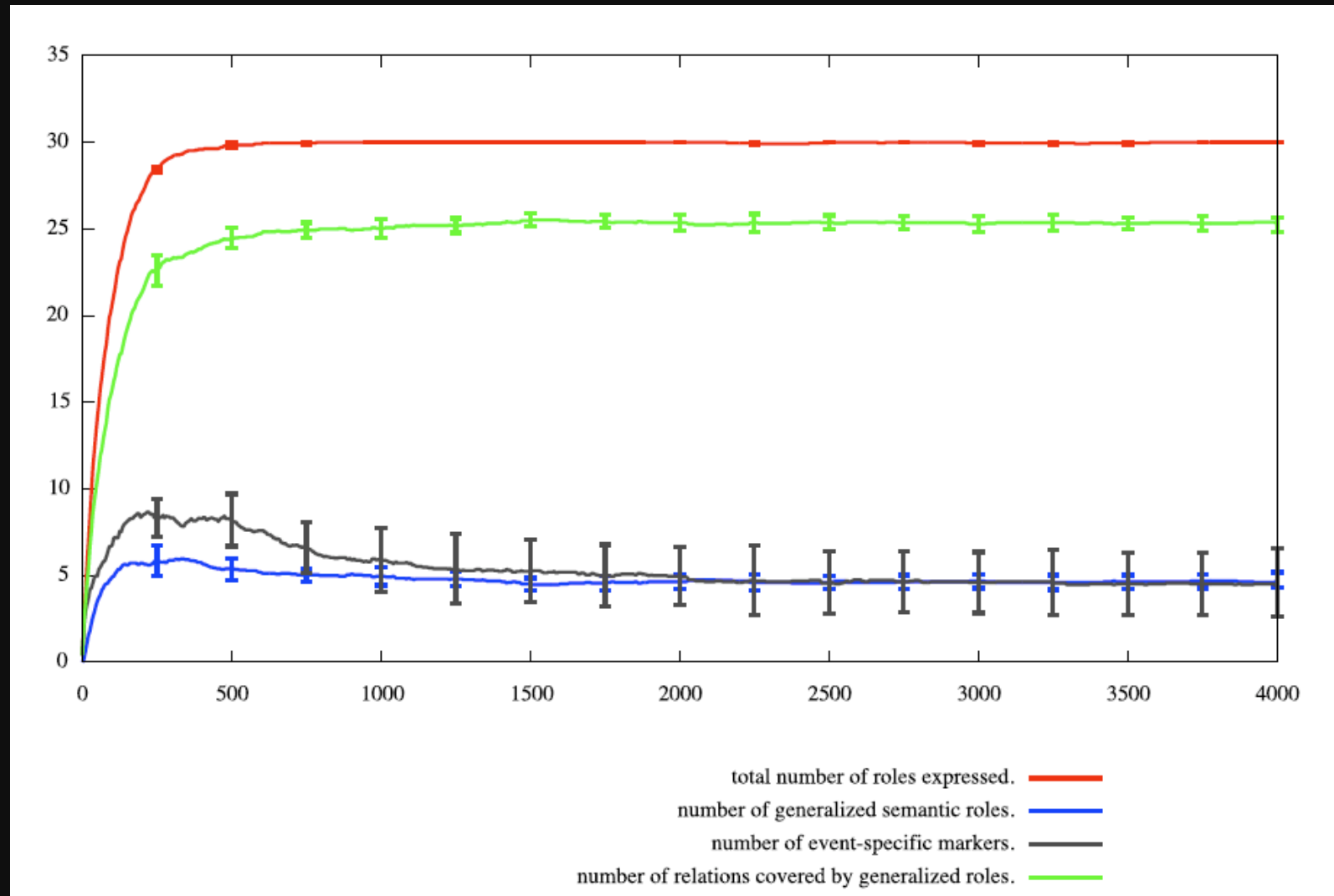
Complex network/constructicon



Some results...



Some results...



Future work

Event-object relations
touch - toucher

context



Semantic role
agent

context



Case
nominative

marker
"-us"

context



Conclusions

- Change is pervasive in all layers of language
- Constructing a language = dealing with variety and uncertainty
- complex networks emerge incrementally through problem-solving activities in local interactions
 - > Learning strategies: diagnostics, repairs, consolidation
- After the master class: time for more details...

Thank you for your attention!

- remi@csl.sony.fr - <http://www.csl.sony.fr/>
- Acknowledgements: ECAgents Project
- Thanks to my PhD supervisors:
 - > Luc Steels (University of Brussels - Sony CSL Paris)
 - > Walter Daelemans (University of Antwerp)
- ... and my colleagues:
Joris Bleys, Joachim De Beule, Martin Loetzsch, Michael Spranger, Wouter Van den Broeck and Pieter Wellens