The Parallel Meaning Bank

Oslo, 29 May 2018

Johan Bos

University of Groningen


The Parallel Meaning Bank

Lost in Translation – Found in Meaning

NWO vici project: 2016-2020

Kilian Evang
(Postdoc)

Johan Bos
(Project Leader)

Talita Anthonio
(Student Assistant)

Johannes Bjerva
(PhD Student)

Lasha Abzianidze
(Postdoc)

Rik van Noord
(PhD Student)

Hessel Haagsma
(PhD Student)

Jaap Nanninga
(Student Assistant)

Parallel Meaning Bank

university of groningen

NWO
Netherlands Organisation for Scientific Research

Objectives of the project

- Create a parallel corpus of sentences annotated with formal meaning representations
- Produce semantic tools for languages other than English
- Learn more about (cross-lingual) compositional semantics

Long-term Goals

- Learn about (human) translations
- Verify translations
- Improve (machine) translation?

The basic idea…

This school was founded in 1650.

x1 e1 t1

school(x1)
time(t1)
  YearOfCentury(t1, 1650)
  t1 < now
establish(e1)
  Time(e1, t1)
  Theme(e1, x1)

Diese Schule wurde 1650 gegründet.

x1 e1 t1

school(x1)
time(t1)
  t1 < now
  YearOfCentury(t1, 1650)
establish(e1)
  Time(e1, t1)
  Theme(e1, x1)

Integrating Lexical and Formal Semantics

Lexical Semantics:
"long words"

- Thematic Roles
- Word Senses
- Named Entities
- Metonomy
- MWEs
- Coercion
- Prepositions

Formal Semantics:
"short words"

- Quantification
- Pronouns
- Negation
- Ellipsis
- Presupposition
- Rhetorical relations
- Quotation
- Tense

The Parallel Meaning Bank

11,5M word tokens

Dutch Parallel Corpus

TATOEBA BETA project

INTERSECT

qtleap

OPUS

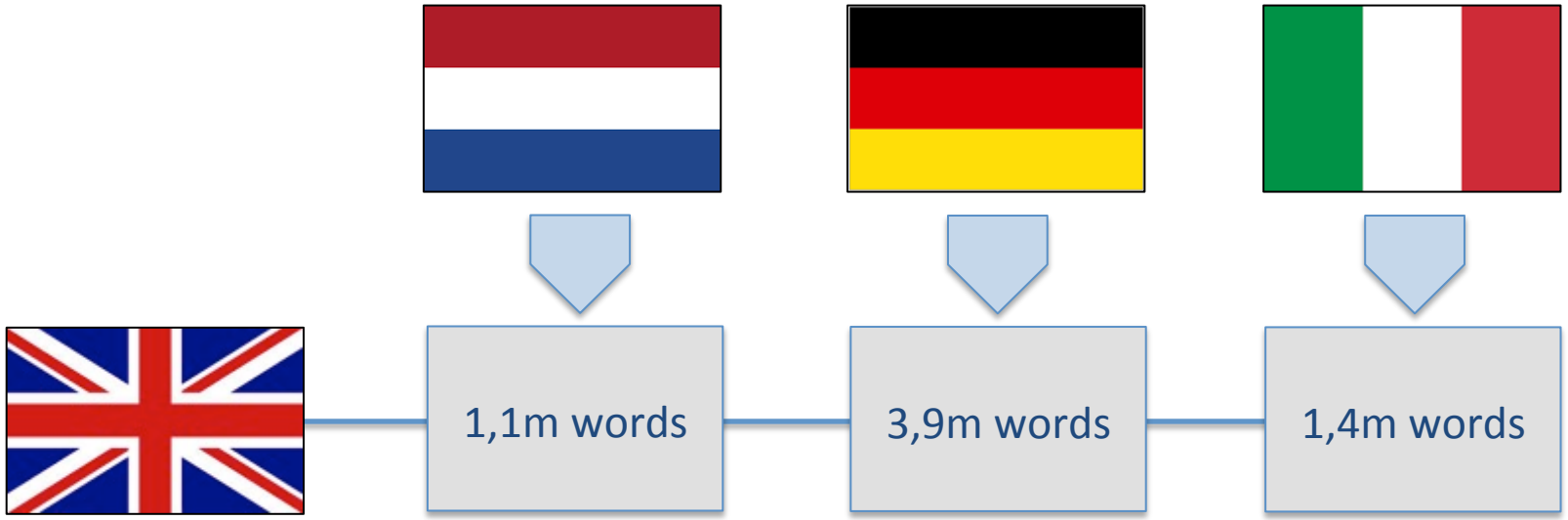
LONWEB VOLUNTEERS program

WWW.LONWEB.ORG

QA@CLEF-2004

PASCAL2 Pattern Analysis, Statistical Modelling and Computational Learning

TED



The Parallel Meaning Bank

English as pivot language (5 million words)

(ca. 10,000 documents for all four languages)

Design

- Several (but not too many) annotation layers
- Collaborative annotation:
  experts, the crowd, machines
- Distinguish Bronze/Silver/Gold standard

Annotation Standards

Annotation layer **completely** approved by a human

Annotation layer **partly** approved by a human

Annotation layer **not** approved by a human

Tools

Language-Neutral Semantic Analysis

- Segmentation (tokenisation)
- Symbolization
- Semantic tagging
- Parsing (CCG syntactic analysis)
- Boxing (DRT semantic analysis)

Segmentation

- Splitting texts into sentences and “words”
- Multi-word expressions…

What are the atoms of meaning?

Segmentation Cases

San Diego, Secretary of State, Royal Bank of Scotland, …

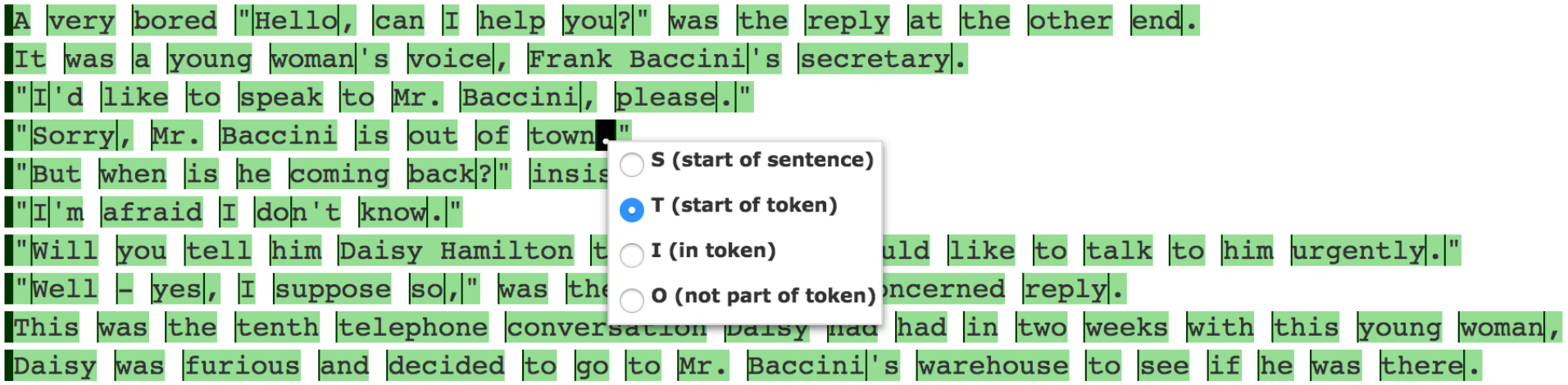
baseball club, knitting needles, pair of scissors …

unhappy, impossible, disagree, …

ten-year-old, data-driven, New York-based …

as well as, instead of, again and again, …

Improving Segmentation (IOB tagging)



elephant

K. Evang, V. Basile, G. Chrupała and J. Bos (2013): ***Elephant: Sequence Labeling for Word and Sentence Segmentation***. Proc. of EMNLP 2013: Conference on Empirical Methods in Natural Language Processing, Seattle, United States

Symbolization

- Mapping words to non-logical symbols
- Morphological analysis (lemmatization)
- Normalisation

token	**symbol**
third	3
men	man
played	play
2:30 pm	14:30
2,5 million	2500000
km	kilometer

Semantic tagging instead of POS-tagging

- POS-tagging required, but: language-specific
- Moreover: not fine-grained enough and some categories are irrelevant for semantics
- Named entity recognition

Semantic tagging *uses a semantically rather than syntactically motivated tagset and is language neutral*

Universal Semantic Tagging

ANA	PRO	pronoun
	DEF	definite
	HAS	possessive
	REF	reflexive
	EMP	emphasizing
ACT	GRE	greeting
	ITJ	interjection
	HES	hesitation
	QUE	interrogative
ATT	QUA	quantity
	UOM	measurement
	IST	intersective
	REL	relation
	RLI	rel. inv. scope
	SST	subsective
	PRI	privative
	INT	intensifier
	SCO	score
LOG	ALT	alternative
	EXC	exclusive
	NIL	empty
	DIS	disjunct./exist.
	IMP	implication
	AND	conjunct./univ.
	BUT	contrast

COM	EQA	equative
	MOR	comparative pos.
	LES	comparative neg.
	TOP	pos. superlative
	BOT	neg. superlative
	ORD	ordinal
DEM	PRX	proximal
	MED	medial
	DST	distal
DIS	SUB	subordinate
	COO	coordinate
	APP	appositional
MOD	NOT	negation
	NEC	necessity
	POS	possibility
ENT	CON	concept
	ROL	role
NAM	GPE	geo-political ent.
	PER	person
	LOC	location
	ORG	organisation
	ART	artifact
	NAT	natural obj./phen.
	HAP	happening
	URL	url

EVE	EXS	untensed simple
	ENS	present simple
	EPS	past simple
	EFS	future simple
	EXG	untensed prog.
	ENG	present prog.
	EPG	past prog.
	EFG	future prog.
	EXT	untensed perfect
	ENT	present perfect
	EPT	past perfect
	EFT	future perfect
	ETG	perfect prog.
	ETV	perfect passive
	EXV	passive
TNS	NOW	present tense
	PST	past tense
	FUT	future tense
TIM	DOM	day of month
	YOC	year of century
	DOW	day of week
	MOY	month of year
	DEC	decade
	CLO	clocktime

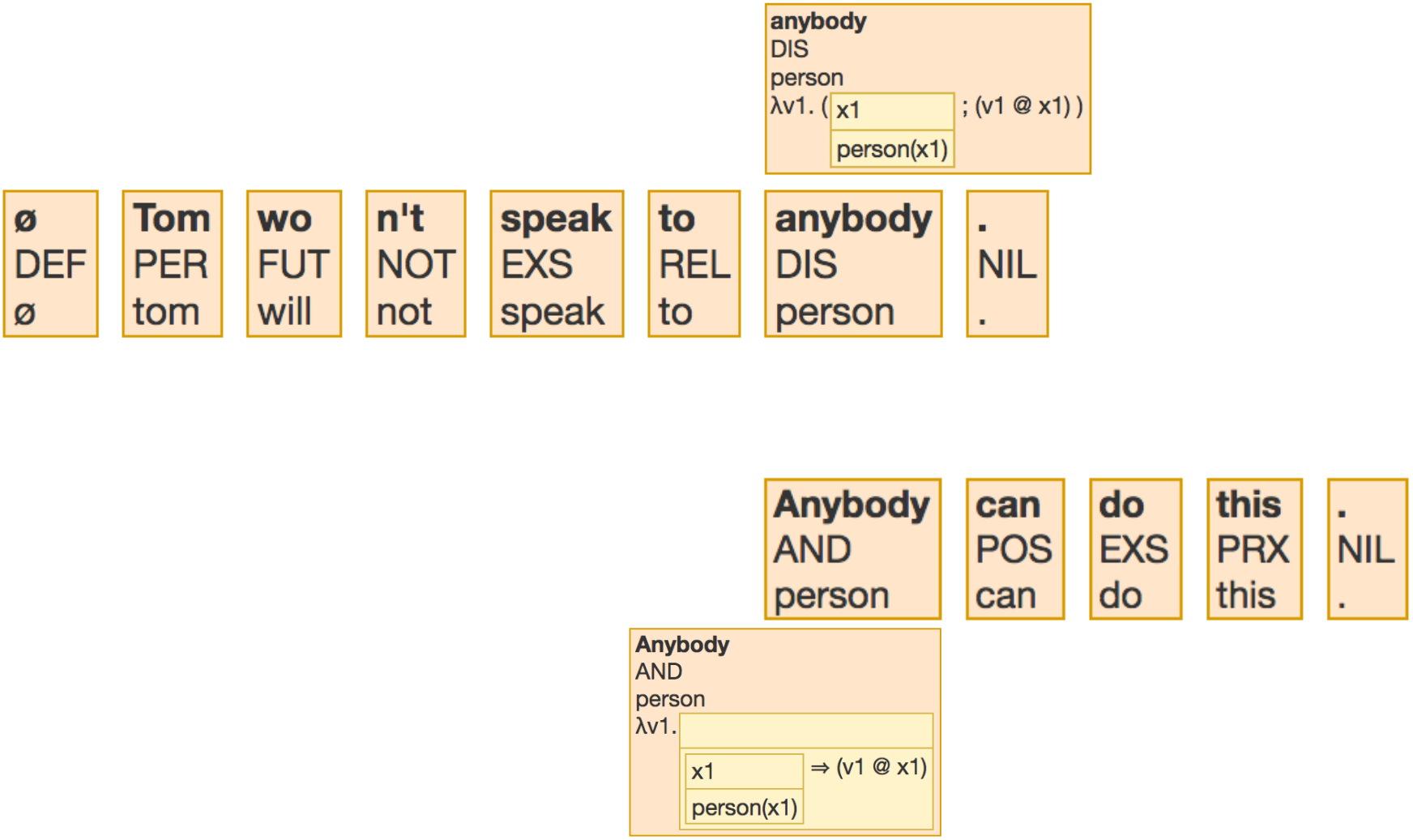
Bjerva, Plank & Bos (2016): Semantic Tagging with Deep Residual Networks. COLING.

Universal Semantic Tagging

ø	Tom	wo	n't	speak	to	anybody	.
DEF	PER	FUT	NOT	EXS	REL	DIS	NIL
ø	tom	will	not	speak	to	person	.

Anybody	can	do	this	.
AND	POS	EXS	PRX	NIL
person	can	do	this	.

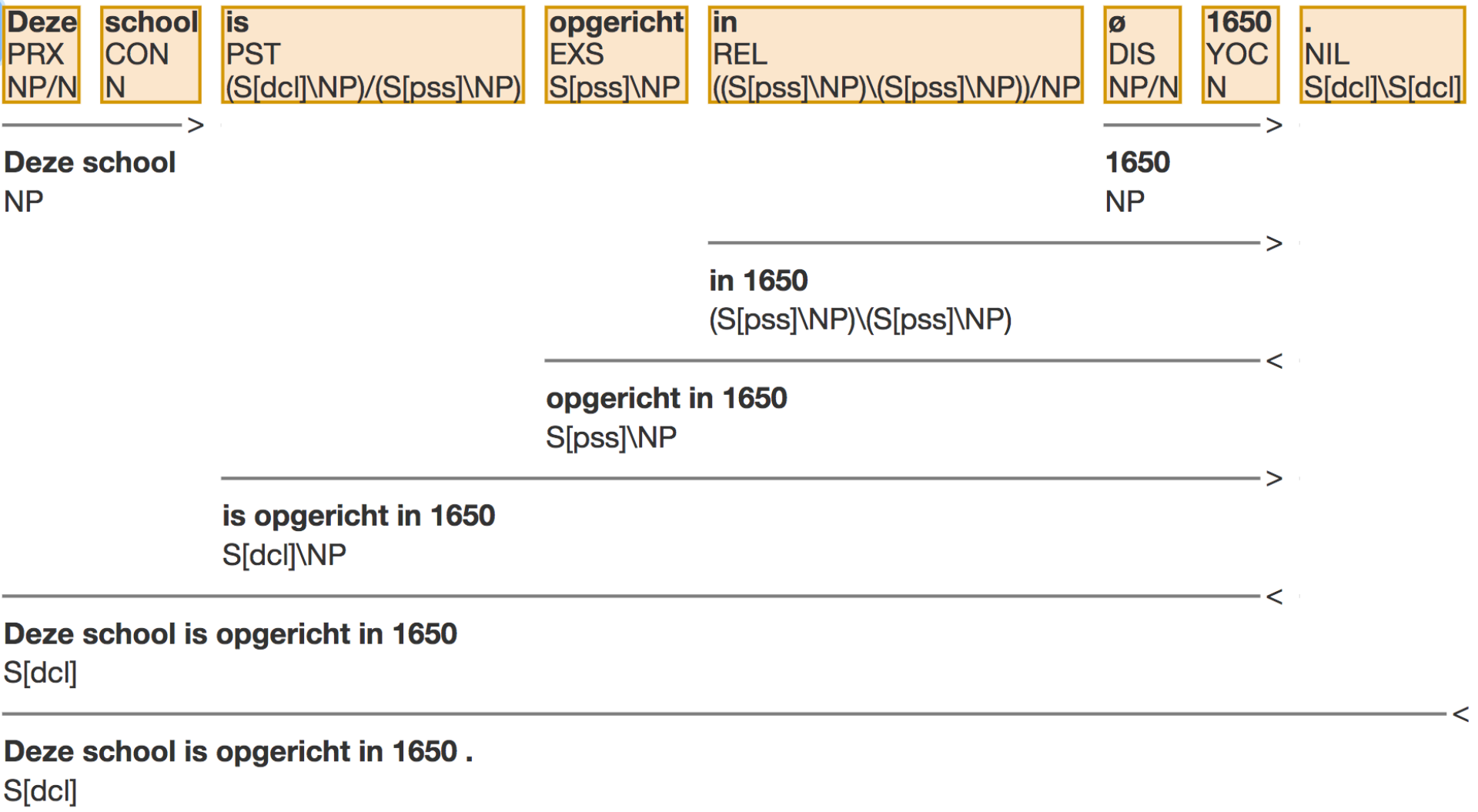
Universal Semantic Tagging



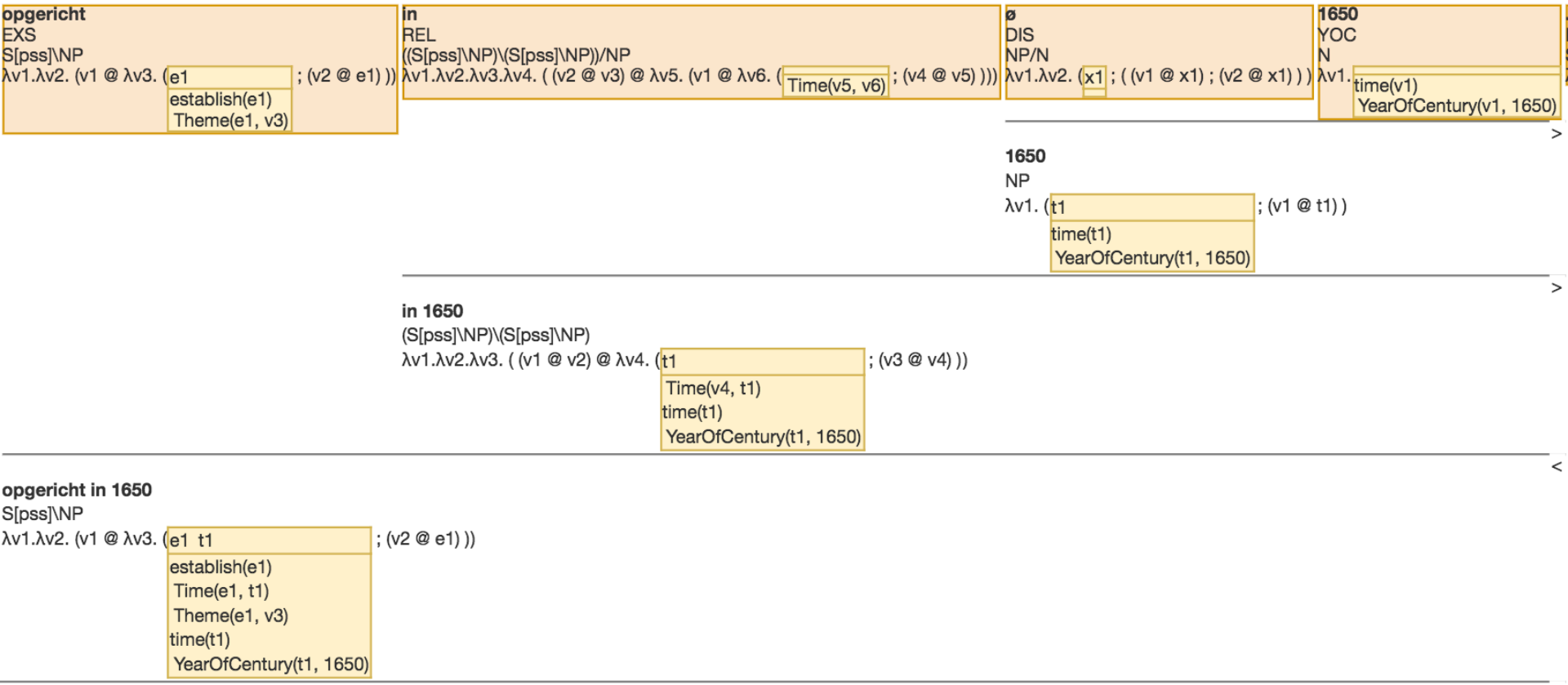
Syntactic Analysis

- Need syntactic analysis for compositional semantics
- CCG, Combinatory Categorial Grammar
- Efficient and wide-coverage parsers available
- Easy to train new languages
- Lexically-driven formalism

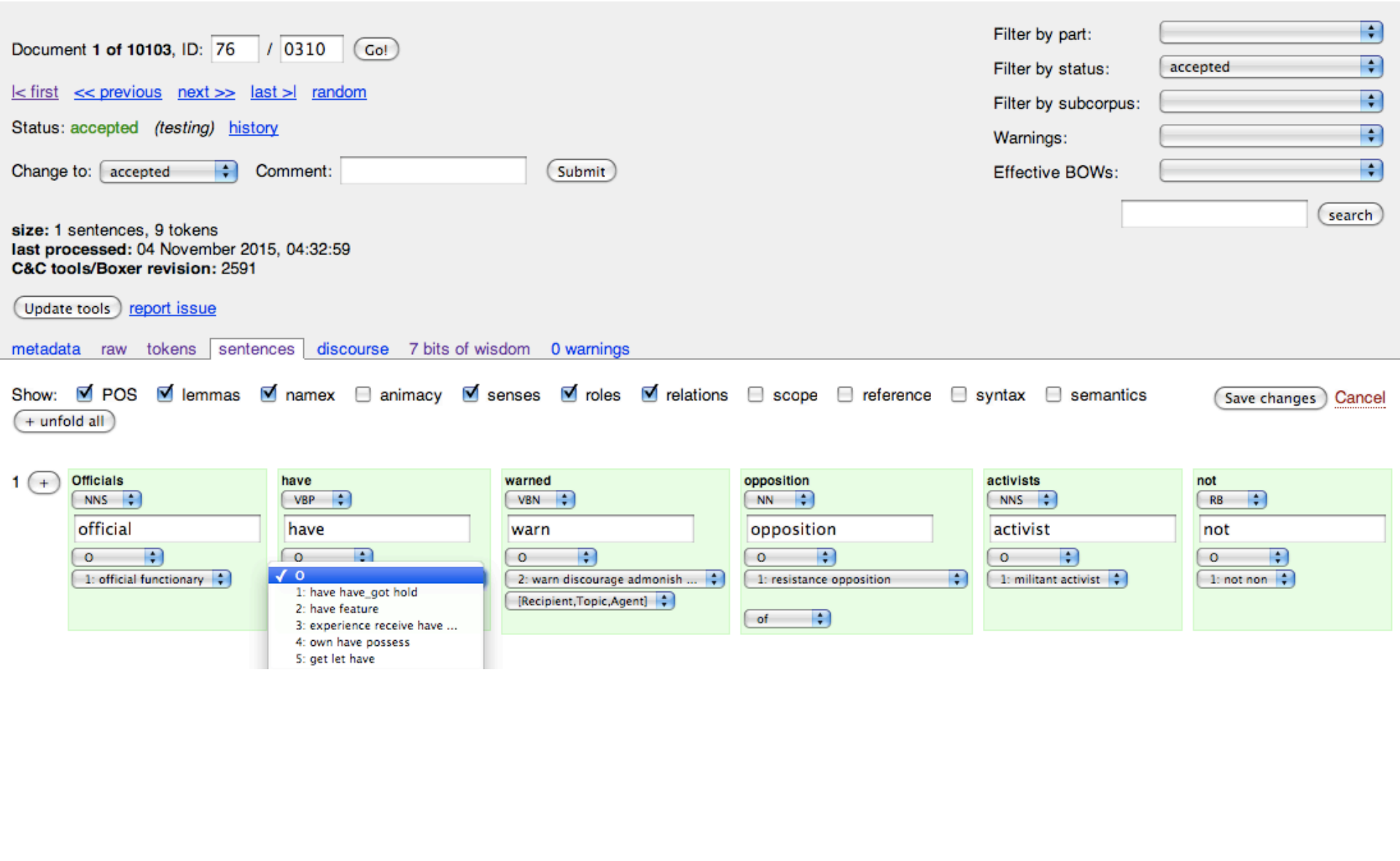
Syntactic Analysis (CCG)



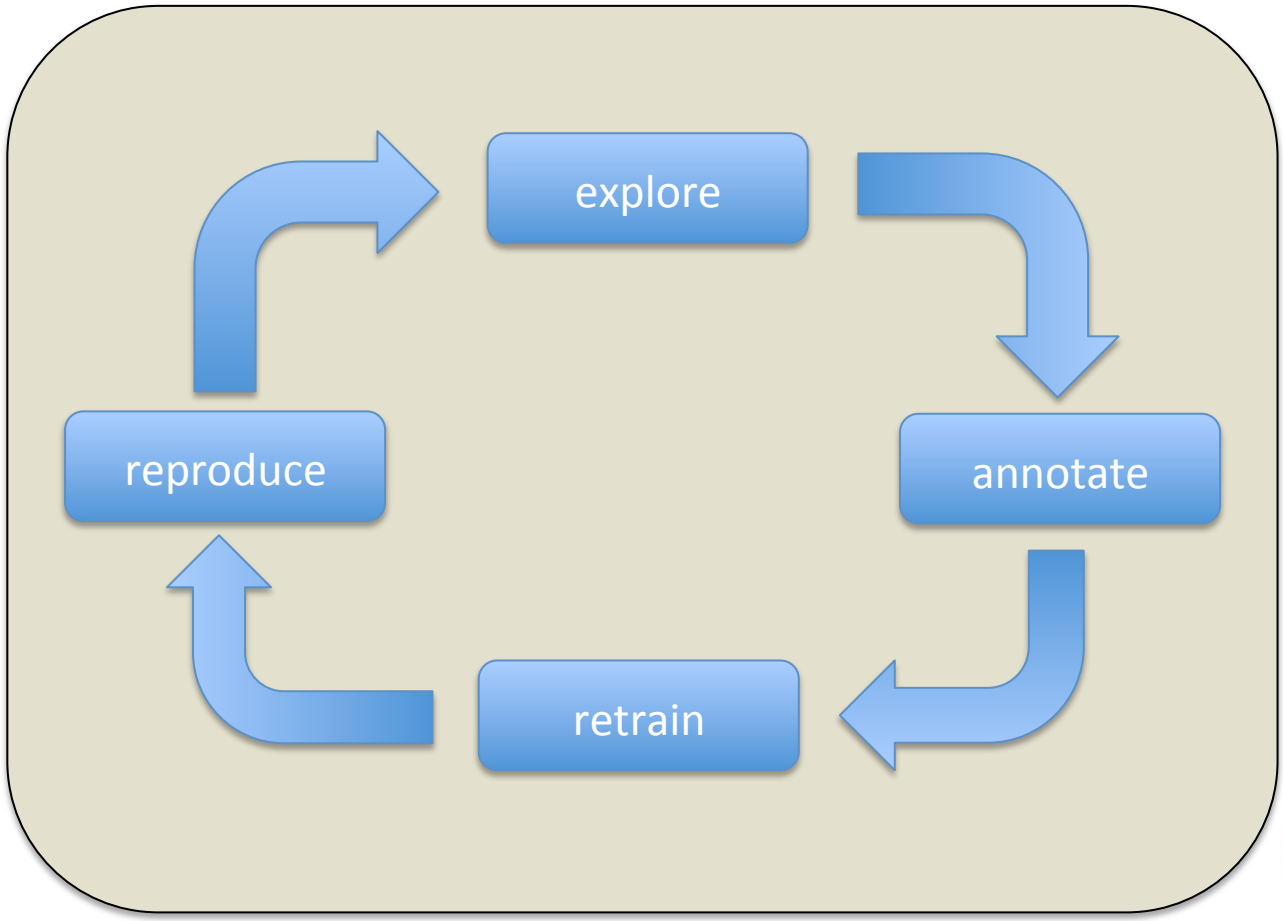
Compositional Semantics (λ-DRT)



The PMB explorer

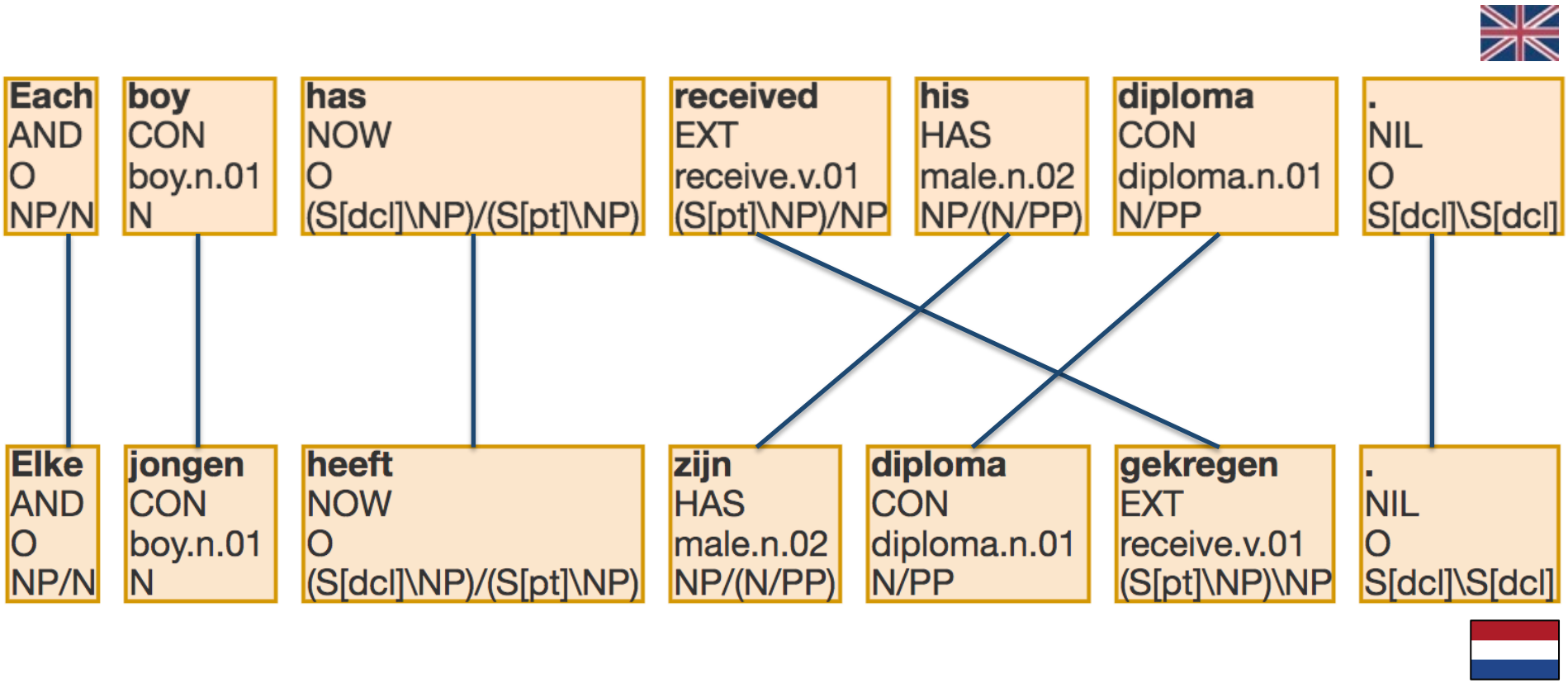


Semantic Annotation: the REAR cycle



Parallel MEANING BANK

Projection: EN → NL (PMB 19/0830)

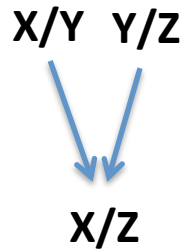


see: **Evang & Bos (COLING 2016)**

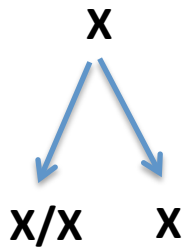
Copy, Merge & Split

- **Copy**:
  transfer of category from source to target

- **Merge**:
  two source categories merge into one target category (composition)



- **Split**:
  one source category into two target categories (de-composition)



Semantic Analysis: DRT

- Discourse Representation Theory
- Meaning representation: **DRS**
  (Discourse Representation Structure)
- There is not “one” DRT/DRS!

Difference with Kamp's DRT

- neo-Davidsonian event analysis (VerbNet)
- No analysis of plurals
- Just three tenses (no aspect)
- WordNet synsets as concepts
- Presupposition as anaphora (Van der Sandt)

Hans Kamp about Boxer

"But in the meantime, so this DRT-based system that Johan Bos has put in place is actually on a number of cases performs just as well or outperforms these largely statistically-based systems. So there is hope there, it is quite impressive, he has this constructor of these representations, so-called Discourse Representation Structures or DRSs, so he let's it run on one part of the Wall Street Journal, DRSs rush by, you can't even see them, they are not all perfect if you look close at it, but they are good enough."

Interview with Hans Kamp
University of Chicago, Elucidations (podcast)

Meaning Representation

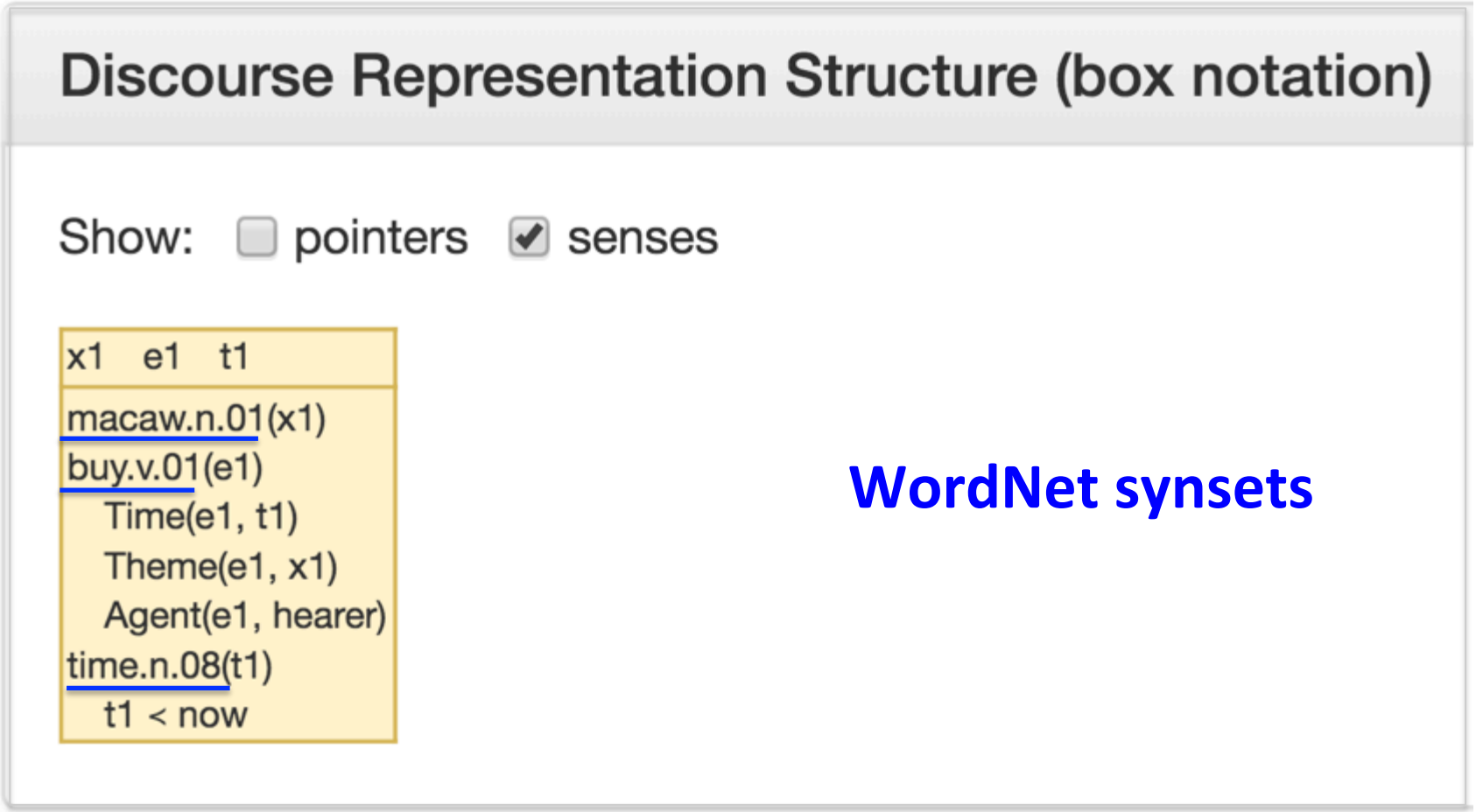
- Logical symbols
  - Comparison operators
  - Boolean operators
  - Variables (quantification)
- Non-logical symbols
  - Concepts (WordNet)
  - Relations (VerbNet)
  - Normalisation of numbers, time, names, etc.

Discourse Representation Structure (box notation)

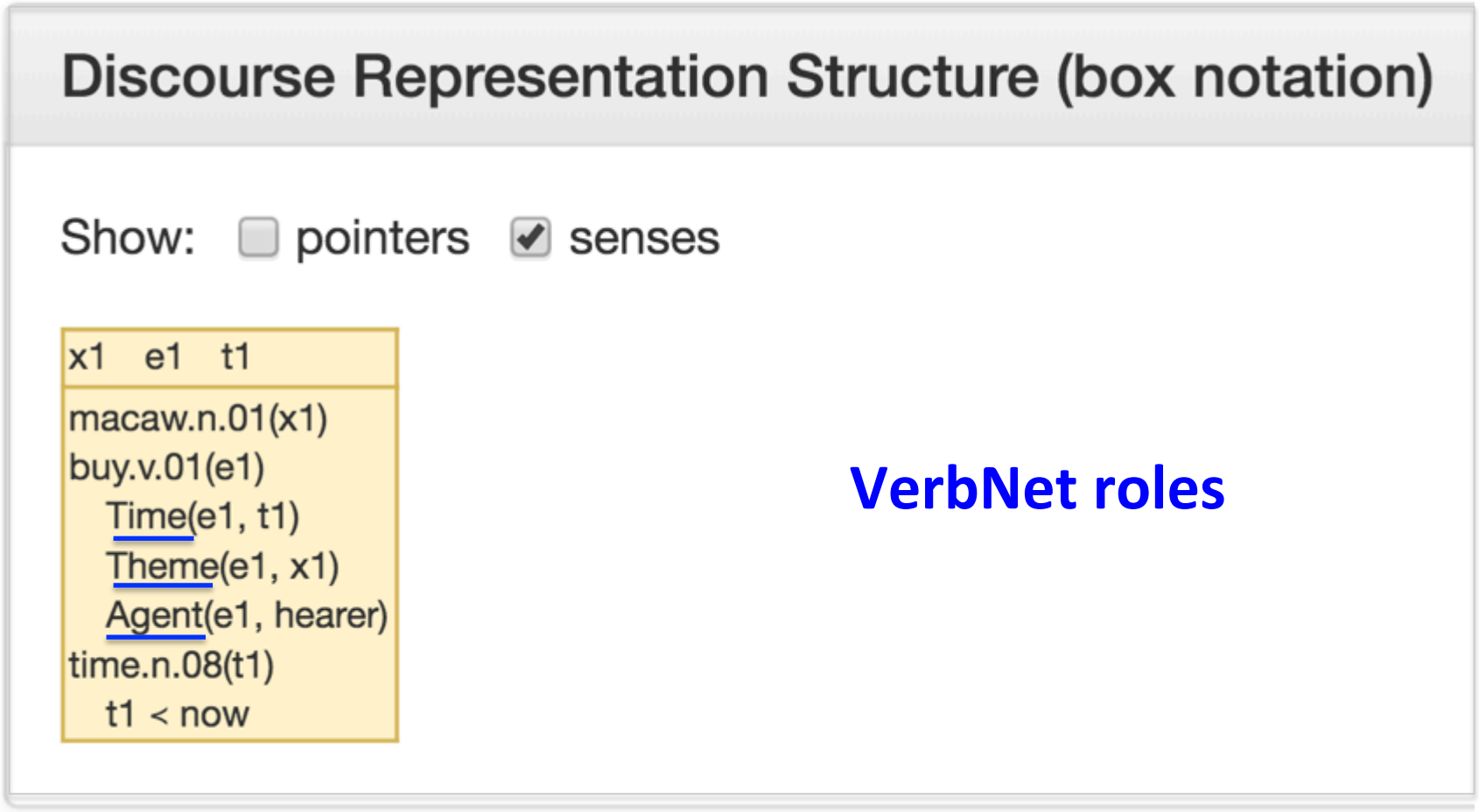
Show: ☐ pointers ☐ senses

x1 e1 t1
macaw(x1) buy(e1) Time(e1, t1) Theme(e1, x1) Agent(e1, hearer) time(t1) t1 < now

PMB 12/0949: **U heeft een ara gekocht.**

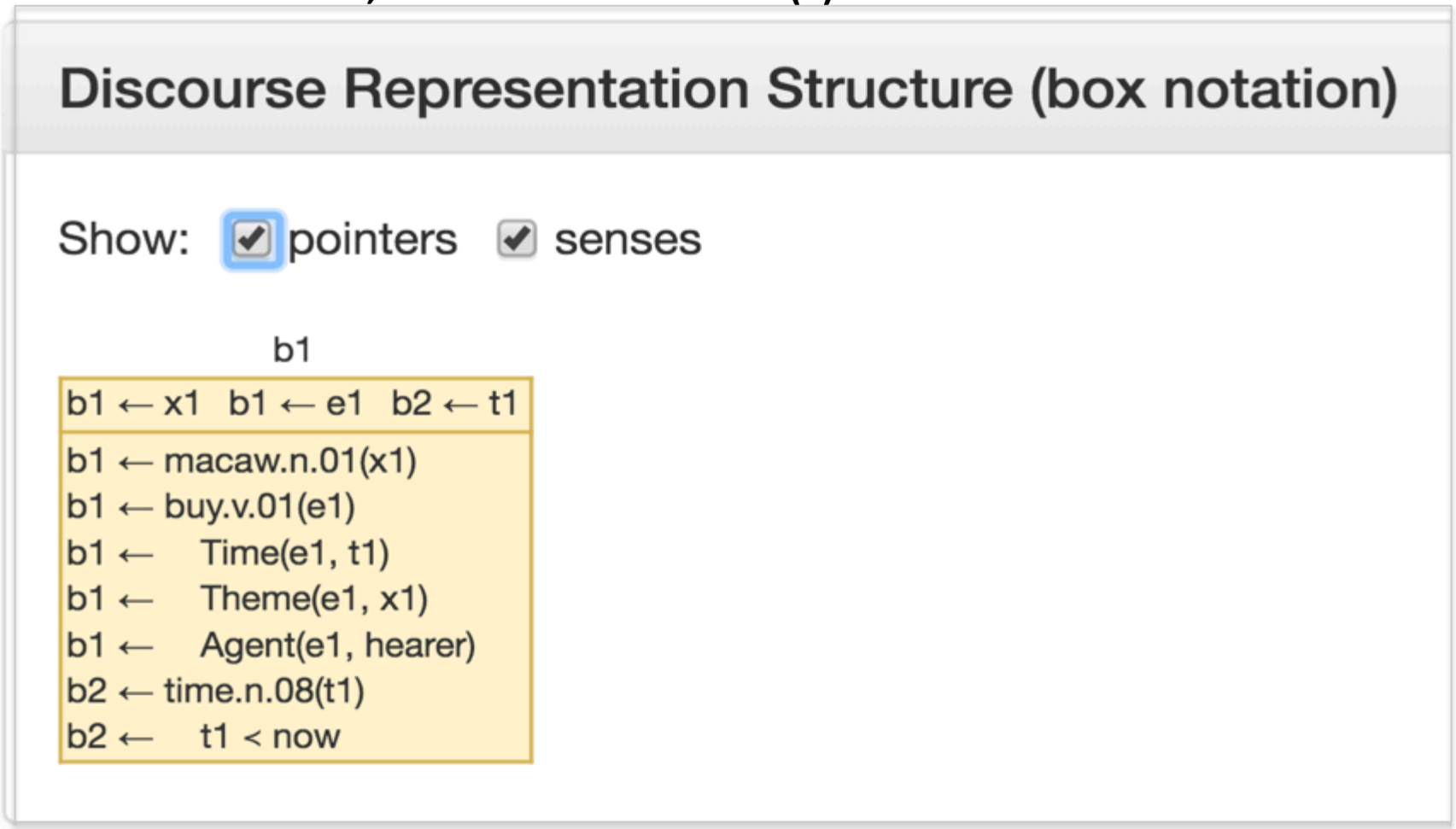


PMB 12/0949: **U heeft een ara gekocht.**



PMB 12/0949: **U heeft een ara gekocht.**

Venhuizen notation, Journal of Semantics 35(1)

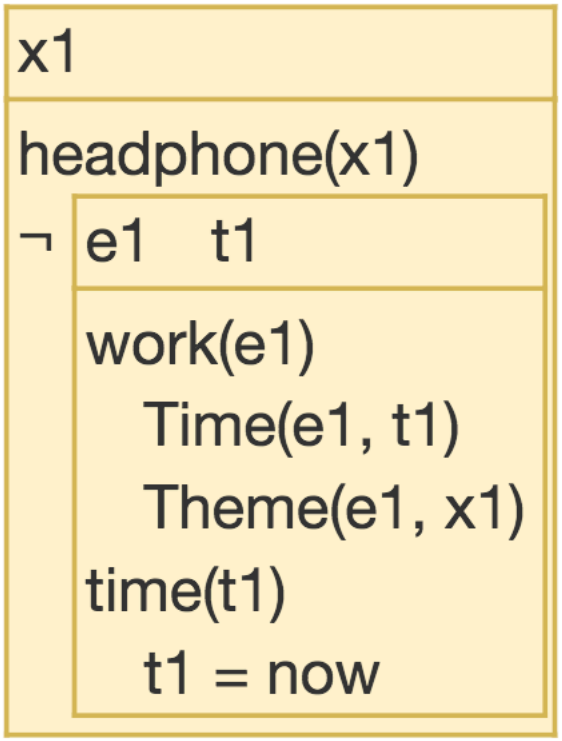


PMB 12/0949: **U heeft een ara gekocht.**

DRS (box notation)

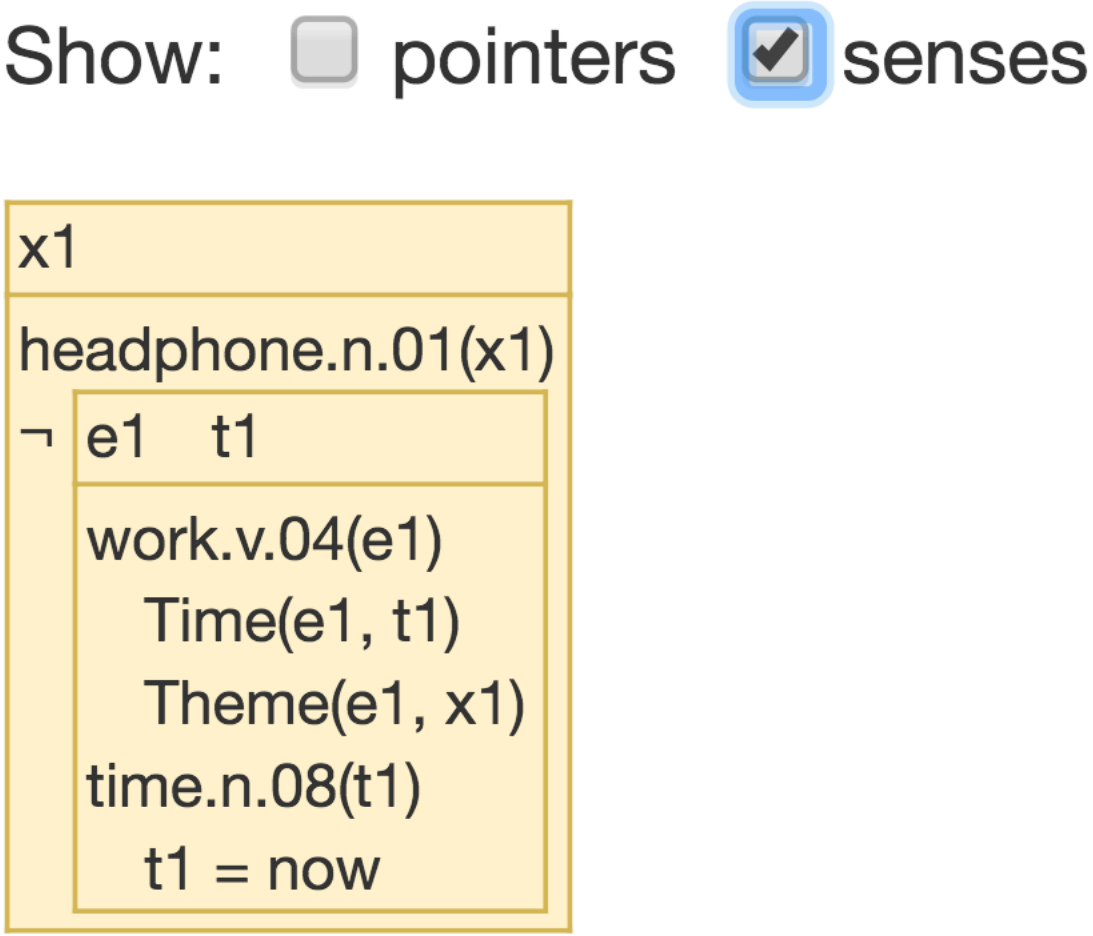
96/2544 These headphones don't work.

Show: ☐ pointers ☐ senses



DRS (box notation)

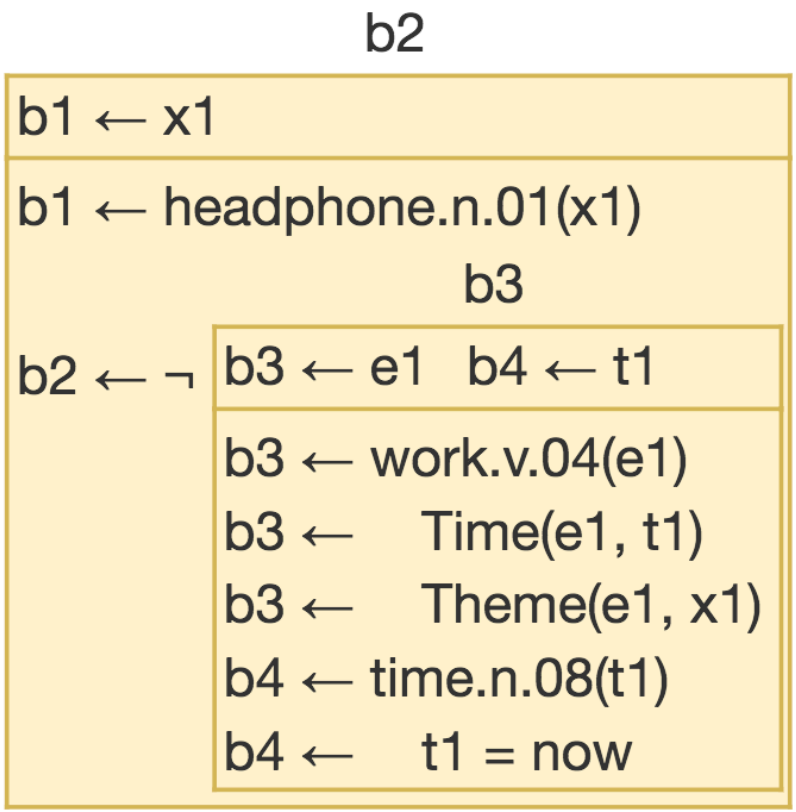
96/2544 These headphones don't work.



DRS (Venhuizen notation)

96/2544 These headphones don't work.

Show: ☑ pointers ☑ senses



DRS (clause notation)

96/2544 These headphones don't work.

```
b1 REF x1
b1 headphone "n.01" x1
b3 Time e1 t1
b4 REF t1
b4 EQU t1 "now"
b4 time "n.08" t1
b2 NOT b3
b3 REF e1
b3 Theme e1 x1
b3 work "v.04" e1
```

DRS (word-aligned clause notation)

96/2544 These headphones don't work.

```
b1 REF x1                    % These [0...5]
b1 headphone "n.01" x1       % headphones [6...16]
b3 Time e1 t1                % do [17...19]
b4 REF t1                    % do [17...19]
b4 EQU t1 "now"              % do [17...19]
b4 time "n.08" t1            % do [17...19]
b2 NOT b3                    % n't [19...22]
b3 REF e1                    % work [23...27]
b3 Theme e1 x1               % work [23...27]
b3 work "v.04" e1            % work [23...27]
                             % . [27...28]
```

Most probable interpretation

41/2289: Tom is stuck in his sleeping bag.

sleeping_bag.n.01(x)

bag.n.01(x) & sleep.v.01(e) & Agent(e,x)

Most probable interpretation

41/2289: Tom is stuck in his sleeping bag.

sleeping_bag.n.01(x)

in his sleeping~bag

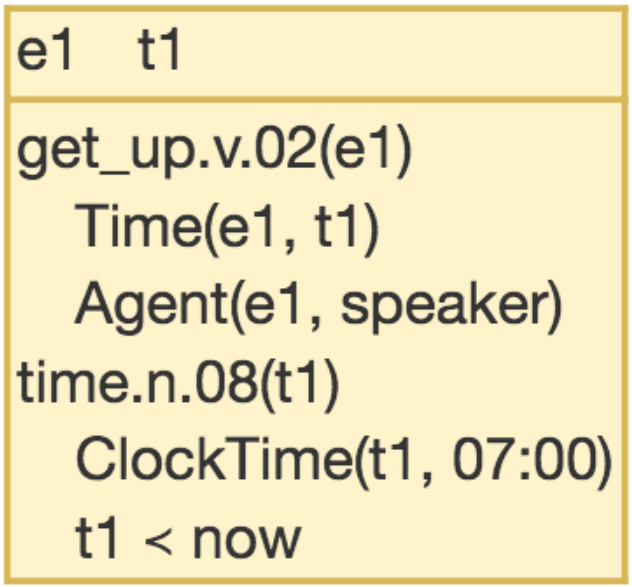
PP

λv1.

x1 x2
Location(v1, x2) male.n.02(x1) sleeping_bag.n.01(x2) User(x2, x1)

Most probable interpretation

19/3282: I got up at seven.



Back to the drawing board

- Proper names
- Agent nouns
- Vague numeral expressions

Proper Names

“Hillary Clinton”

x
female.n.02(x)
  Name(x, ”hillary~clinton”)

Agent nouns/Role nouns

“inventor”

x y
person.n.01(x)
  Role(x,y)
inventor.n.01(y)

Agent nouns/Role nouns

"old musician"

x y s
person.n.01(x)
  Role(x,y)
musician.n.01(y)
old.a.01(s)
  Theme(s,x)

Agent nouns/Role nouns

“old friend”

x y s
person.n.01(x)
  Role(x,y)
friend.n.01(y)
old.a.01(s)
  Theme(s,y)

Numeral expressions

“20 geese”

x y
goose.n.01(x)
  Quantity(x,y)
   y = 20

Numeral expressions

“ca. 20 geese”

x y z

goose.n.01(x)

Quantity(x,y)

$y \approx z$

$z = 20$

Noun-noun compounds

toy doll	toy.n.01(x) & doll.n.01(y) & y=x
toy car	toy.n.01(x) & car.n.01(y) & y≈x
toy store	toy.n.01(x) & store.n.01(y) & Theme(y,x)

bottle opener	bottle~opener.n.01(x)
peanut butter	peanut~butter.n.01(x)

Comparison

	GMB	AMR	PMB
scope	yes	no	yes
roles	VerbNet	PropBank	VerbNet
gold	no	yes	yes
senses	no	“verbs” only	wordnet
plurals	no	no	no
tense	yes	no	yes
aspect	no	no	no
wikification	no	yes	no
languages	EN	EN, ?	EN, DE, NL, IT

What now?

1. From AMR to DRS in 9 steps (fairly cool)
2. Neural semantic parsing (very cool)
3. Demo Semantic Annotation (extremely cool)

From AMR to DRS in nine easy steps

```
(m/man
   :ARG0-of (l/love
               :ARG1 (w/woman)))
```

Step 1: add variables to concepts

```
(m/man(m)
   :ARG0-of (l/love(l)
               :ARG1 (w/woman(w))))
```

Step 2: add variables to relations

```
(m/man(m)
   :ARG0-of(m,l) (l/love(l)
                      :ARG1(l,w) (w/woman(w))))
```

Step 3: re-inverse relations

```
(m/man(m)
   :ARG0(l,m) (l/love(l)
                  :ARG1(l,w) (w/woman(w))))
```

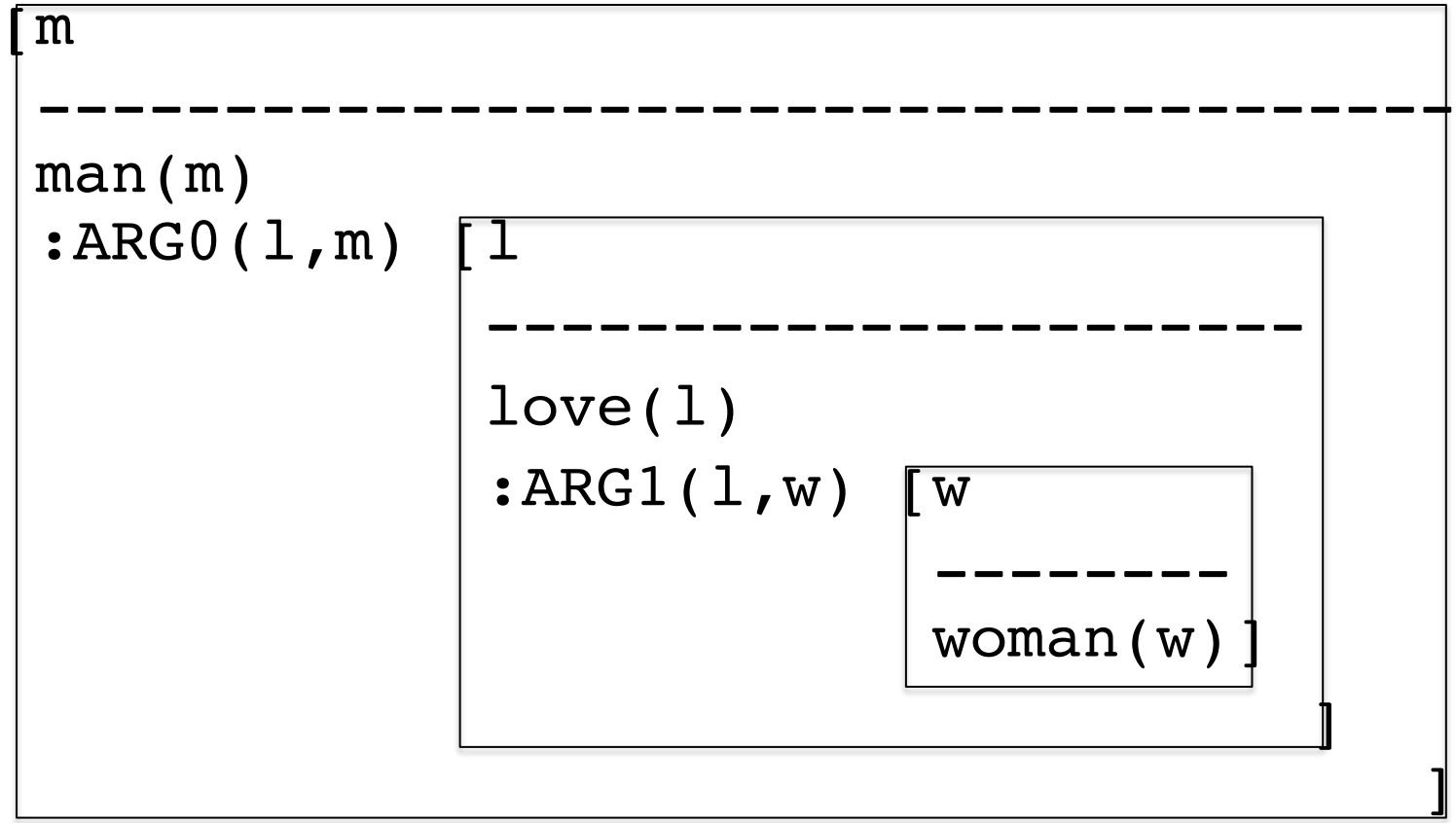
Step 4: square round brackets

```
[m/man(m)
   :ARG0(l,m) [l/love(l)
                  :ARG1(l,w) [w/woman(w)]]]
```

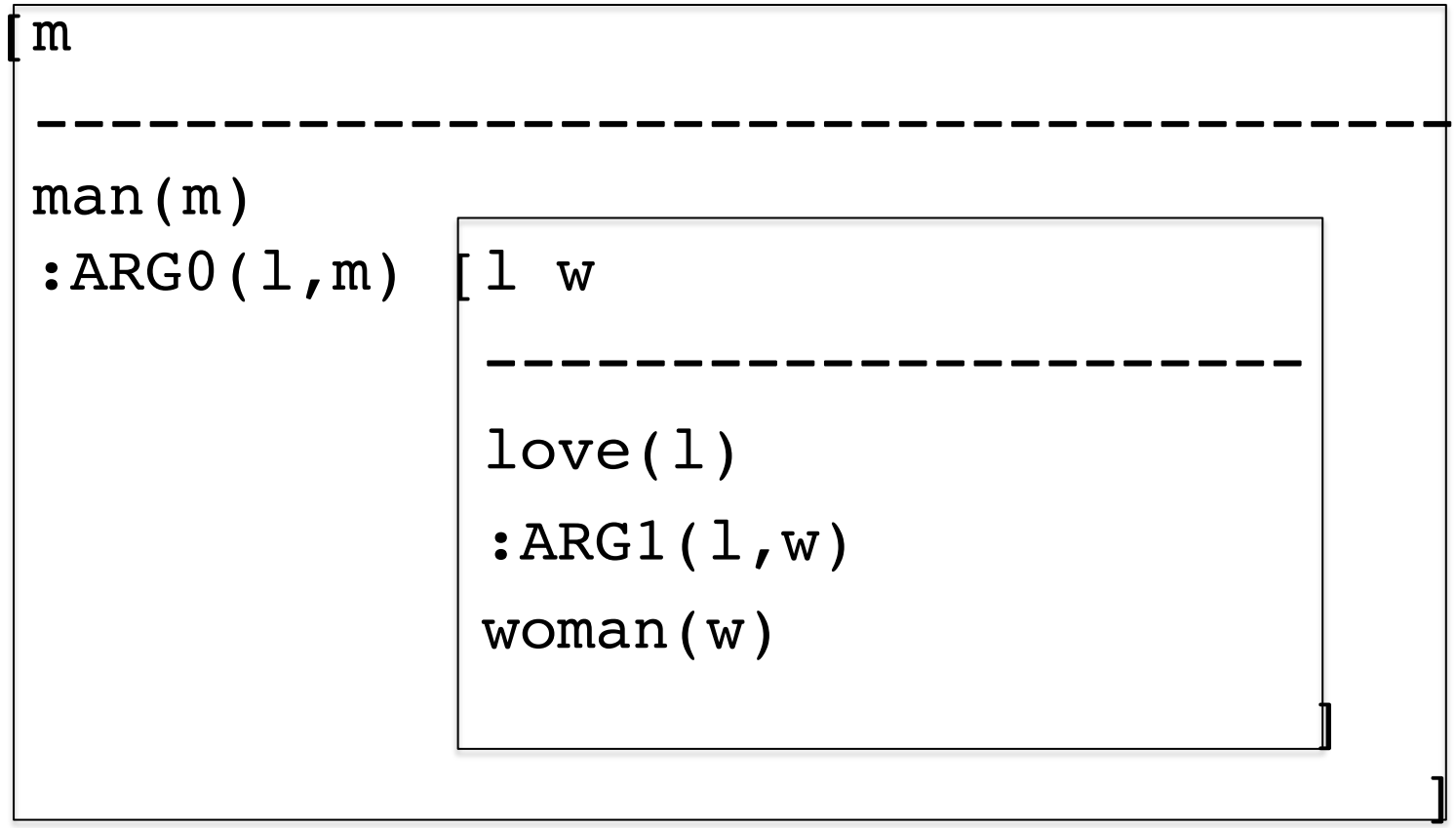
Step 5: add horizontal lines

```
[m
 ------------------------------------
 man(m)
 :ARG0(l,m) [l
             ------------------------
             love(l)
             :ARG1(l,w) [w
                         --------
                         woman(w)]
                                 ]
                                    ]
```

Step 6: draw boxes



Step 7: merge boxes



Step 7: merge boxes

```
[m l w
-----------------------------------
man(m)
:ARG0(l,m)
love(l)
:ARG1(l,w)
woman(w)
```

Step 8: replace relations

```
m l w
-------------------------------------
man(m)
  Experiencer(l,m)
love(l)
  Stimulus(l,w)
woman(w)
```

Step 9: use sorted variables

```
x e y
-----------------------------------
man(x)
  Experiencer(e,x)
love(e)
  Stimulus(e,y)
woman(y)
```

Neural Semantic Parsing

Lost in Translation – Found in Meaning

NWO vici project: 2016-2020

Kilian Evang
(Postdoc)

Johan Bos
(Project Leader)

Talita Anthonio
(Student Assistant)

Johannes Bjerva
(PhD Student)

Lasha Abzianidze
(Postdoc)

Rik van Noord
(PhD Student)

Hessel Haagsma
(PhD Student)

Jaap Nanninga
(Student Assistant)

Parallel Meaning Bank

university of groningen

NWO
Netherlands Organisation for Scientific Research

COUNTER (Rik van Noord)

- Syntactic comparison of two DRSs
  - Compare meanings of translations
  - Compare system output with gold
- Take DRSs in clause notation
- Compute precision and recall of matching clauses (using best mapping of variables)
- Based on AMR/Smatch

COUNTER: example

PMB: 96/3505

EN: Tom was moaning in pain.

NL: Tom kreunde van de pijn.

```
8 out of 9 clauses match

F-score  : 0.8889

Matching clauses:

b1 Name x1 "tom"   % Tom [0...3]      | b1 Name x1 "tom"   % Tom [0...3]
b1 male "n.02" x1  % Tom [0...3]      | b1 male "n.02" x1  % Tom [0...3]
b0 Time e1 t1      % kreunde [4...11] | b0 Time e1 t1      % was [4...7]
b4 TPR t1 "now"    % kreunde [4...11] | b3 TPR t1 "now"    % was [4...7]
b4 time "n.08" t1  % kreunde [4...11] | b3 time "n.08" t1  % was [4...7]
b0 Agent e1 x1     % kreunde [4...11] | b0 Agent e1 x1     % moaning [8...15]
b0 moan "v.01" e1  % kreunde [4...11] | b0 moan "v.01" e1  % moaning [8...15]
b0 Theme e1 x2     % van [12...15]    | b0 Theme e1 x2     % in [16...18]

Non-matching clauses:

b3 pain "n.01" x2  % pijn [19...23] | b0 pain "n.01" x2  % pain [19...23]

Concepts normalized to synset ID:

moan.v.01 -> groan.v.01
time.n.08 -> fourth_dimension.n.01
```

PMB: 18/3348

EN: My nephew is allergic to eggs.

IT: Mio nipote è allergico alle uova.

```
10 out of 11 clauses match

F-score  : 0.9091

Matching clauses:

b1 Of x2 "speaker"     % Mio [0...3]          | b1 Of x2 "speaker"     % My [0...2]
b1 Role x1 x2          % nipote [4...10]      | b1 Role x1 x2          % nephew [3...9]
b1 nephew "n.01" x2    % nipote [4...10]      | b1 nephew "n.01" x2    % nephew [3...9]
b1 person "n.01" x1    % nipote [4...10]      | b1 person "n.01" x1    % nephew [3...9]
b0 Time s1 t1          % è [11...12]          | b0 Time s1 t1          % is [10...12]
b4 EQU t1 "now"        % è [11...12]          | b3 EQU t1 "now"        % is [10...12]
b4 time "n.08" t1      % è [11...12]          | b3 time "n.08" t1      % is [10...12]
b0 Experiencer s1 x1   % allergico [13...22] | b0 Experiencer s1 x1   % allergic [13...21]
b0 allergic "a.02" s1  % allergico [13...22] | b0 allergic "a.02" s1  % allergic [13...21]
b0 Stimulus s1 x3      % alle [23...27]       | b0 Stimulus s1 x3      % to [22...24]

Non-matching clauses:

b3 egg "n.02" x3       % uova [28...32] | b0 egg "n.02" x3       % eggs [25...29]

Concepts normalized to synset ID:

time.n.08     -> fourth_dimension.n.01
allergic.a.02 -> allergic.s.02
```

Boxer vs Neural Boxer

Boxer

- Tokenisation
- Semantic tagging
- Syntactic analysis (CCG)
- Thematic role labelling
- Word sense disambiguation
- Pronoun/Presupposition resolution
- α-conversion + β-conversion

Neural Semantic Parsing

- Barzdins & Gosko (2016)
- AMR parsing (state-of-the art)
- Van Noord & Bos (2017)
  Neural Semantic Parsing by Character-based Translation: Experiments with Abstract Meaning Representations
- OpenNMT, bi-LSTM with general attention
- Use this approach for DRS (scoped)

seq2seq, character-based

- Input: “She showers every morning”

- Output:

  b3 REF x1
  b3 female “n.02” x1
  b4 REF e1
  b4 shower “v.03” e1
  b4 Agent e1 x1
  b4 Time e1 x2
  b2 REF x2
  b2 morning “n.01” x2
  b0 IMP b2 b4

bi-directional De Bruyn notation

- Input: “She showers every morning”

- Output:

  $0 REF
  $0 female “n.02” @0
  $1 REF
  $1 shower “v.03” @0
  $1 Agent @0 @1
  $1 Time @0 @-1
  $2 REF
  $2 morning “n.01” @0
  $3 IMP $2 $1

Dealing with Variables

- Variables can have any name!
- Use De Bruyn index notation
- R. Van Noord, J. Bos (2017):
  Dealing with Co-reference in Neural Semantic Parsing. Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2), pp 49–58, Montpellier, France

seq2seq, character-based

S h e + s h o w e r s + e v e r y + m o r n i n g .

$0 + REF |||
$0 + f e m a l e + "n.02" + @0 |||
$1 + REF |||
$1 + s h o w e r + "v.03" + @0 |||
$1 + Agent + @1 + @0 |||
$1 + Time + @-1 + @0 |||
$2 + REF |||
$2 + m o r n i n g + "n.01" + @0 |||
$3 + IMP + $2 + $1 |||

Preliminary Results

- Baseline: 42%
- Old Boxer:
- Neural Boxer:

Preliminary Results

- Baseline: 42%
- Old Boxer: 76%
- Neural Boxer:

Preliminary Results

- Baseline: 42%
- Old Boxer: 76%
- Neural Boxer: 85%

Practical Results

- Project in progress
- Ca. 5,000 short sentences with gold standard annotation
- PMB Release 2.0.0 freely available (next one coming soon!)
- Counter: comparing scoped meaning representations