Computational Linguistics

Miniblock: Language Sciences

2009/2010

Organisation

- Session 1:
  - basic theoretical tools of computational linguistics
  - Natural Language Generation
- Session 2:
  - basic theoretical tools of computational linguistics
  - Natural Language Analysis
- Session 3:
  - basic theoretical tools of computational linguistics
  - Dialogue Systems

Accompanying Reading

- Dale / Moisl / Somers (Hrgg)(2000) *A handbook of natural language processing: techniques and applications for the processing of language as text.* Dekker.

  Chapters:

  Preface, pp.iii-viii
  5. Semantic Analysis (Poesio)
  7. Natural Language Generation (McDonald)
  13. Machine Translation (Somers)
  14. Dialogue Systems: from theory to practice (Allen)

- Mitkov (Hrg)(2002) *Oxford Handbook of Computational Linguistics.* OUP.

  Chapter:

  25. Ontologies (Vossen)

- Meyer (2002) *Synchronic English Linguistics: An introduction.* GNV.

  Chapter:

  4.5 Formal Semantics (pp118—132)

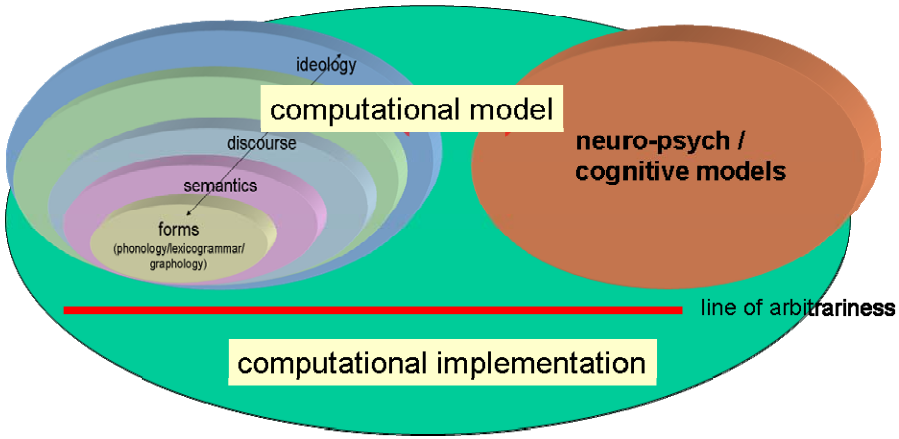
Session 1

- basic theoretical tools of computational linguistics
  - logic: introduction
  - event semantics
  - ontology
- Natural Language Generation

Basic computational modelling tools

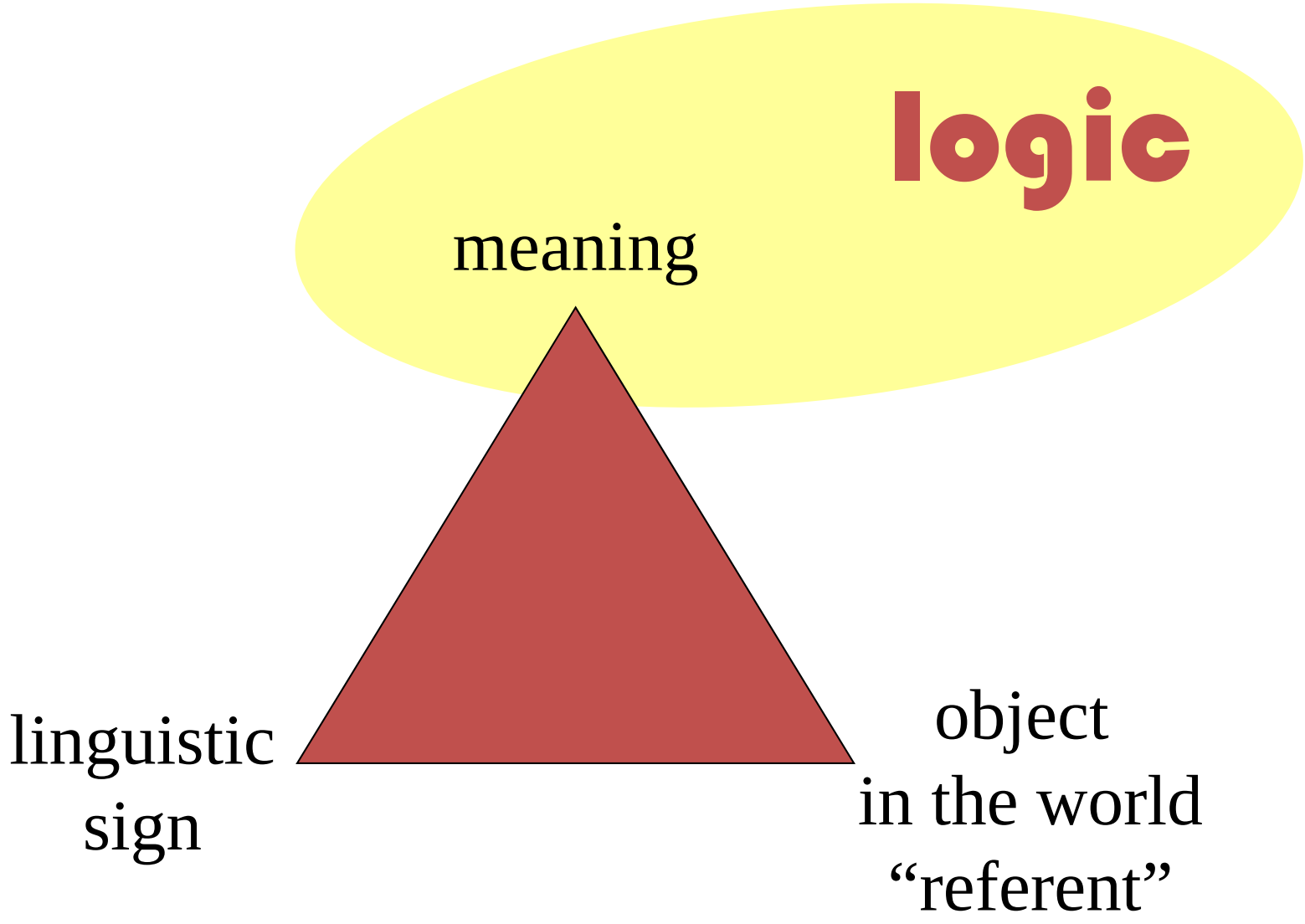
- basic computational linguistic tools for representing **semantics**
  - **logics**
  - **event-based semantics**
  - **ontologies**
  - lambda expressions

- basic computational linguistic tools for representing **information**
  - feature structures
  - unification



Logic and Semantics

The semiotic triangle



Logic

Predicates

- **“one place”**
  - door (x)
  - accountant (x)
  - book (x)
  - run (x)
- **“two place”**
  - eat (x, y)
  - chase (x, y)
  - read (x, y)

Connectives

‘and’ : $\wedge$ ‘or’: $\vee$ ‘not’: $\neg$ ‘implies’ $\rightarrow$

Logical formulae

- “If someone chases someone else then both people run”

what predicates?
what connectives?

Logical formulae

- "If someone chases someone else then both people run"

$$\text{chase}(x, y) \rightarrow \text{run}(x) \wedge \text{run}(y)$$

$$\text{person}(x) \wedge \text{person}(y) \wedge \text{chase}(x, y) \rightarrow \text{run}(x) \wedge \text{run}(y)$$

Logic

Quantifiers

- **existence:** ∃
- **for all:** ∀

All men are mortal.
Socrates is a man.
Therefore Socrates is mortal.

Logic

Quantifiers

- **existence:** ∃
- **for all:** ∀

All men are mortal.

Socrates is a man.

Therefore Socrates is mortal.

- **∀x: man (x) → mortal (x)**
- **man (Socrates)**

→ **mortal (Socrates)**

Logic

Quantifiers

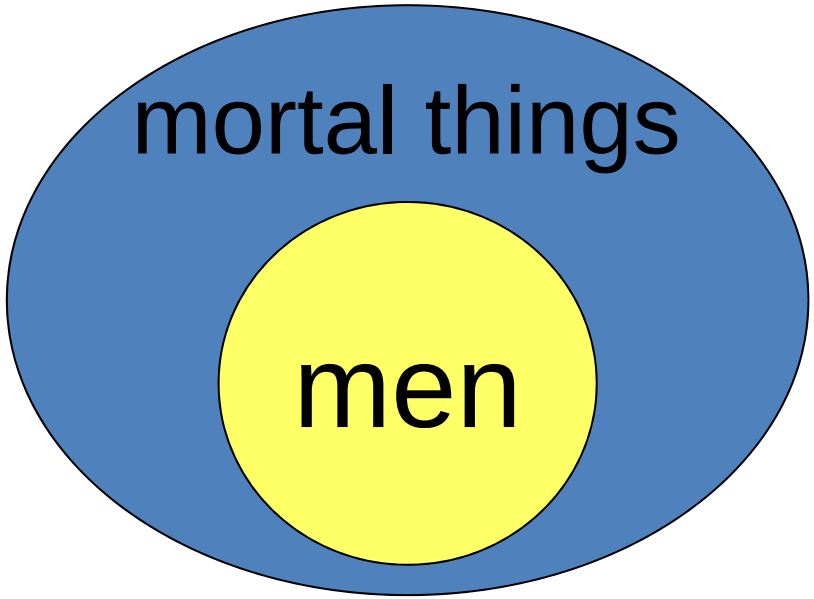
- **existence:** ∃
- **for all:** ∀

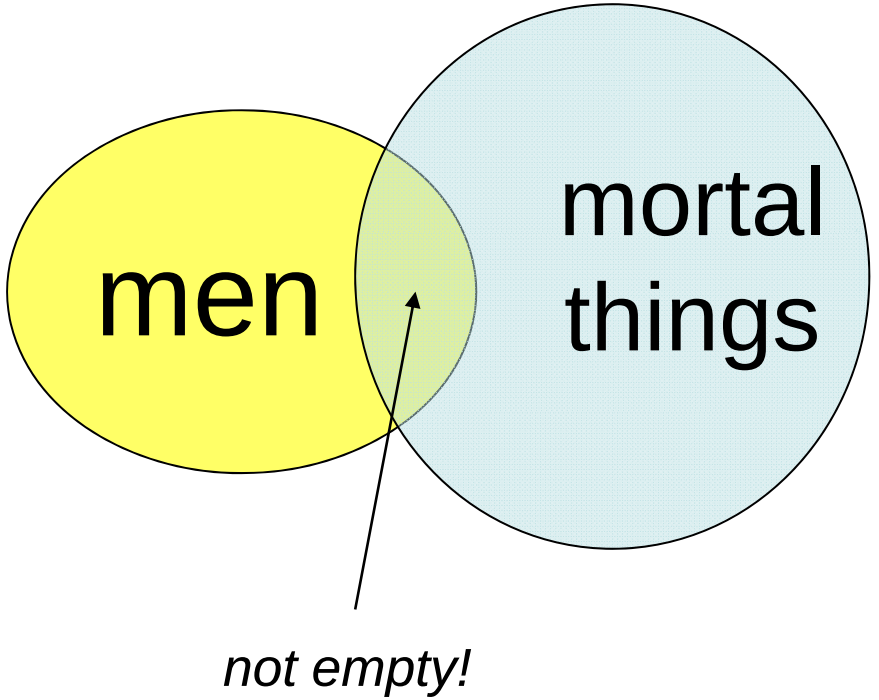
All men are mortal.

Some man is mortal.

- $\forall$**x: man (x)** $\rightarrow$ **mortal (x)**
- $\exists$ **x: man (x)** $\wedge$ **mortal (x)**

Logic: Venn diagrams





All men are mortal.

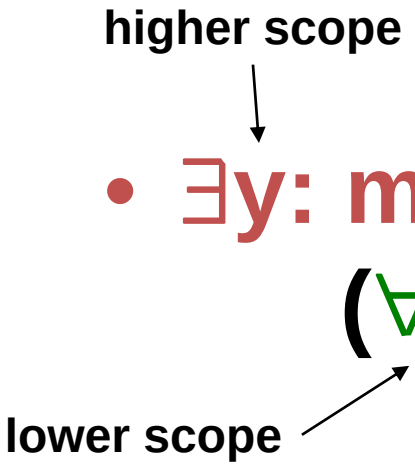
$\forall$**x: man (x)** $\rightarrow$ **mortal (x)**

Some man is mortal

$\exists$ **x: man (x)** $\wedge$ **mortal (x)**

Logical Scope

All men like some mountain.



- $\exists y$: mountain (y)
  ($\forall x$: man (x) $\rightarrow$ like (x, y))

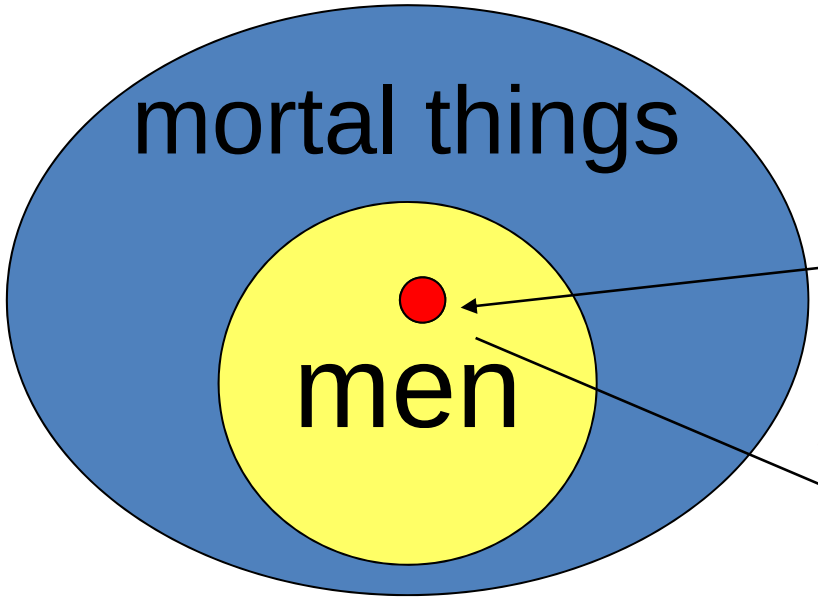
- $\forall x$: man (x)
  ($\exists y$: mountain (y) $\rightarrow$ like (x, y))

Logic: capturing ambiguity

1. Every person in this room speaks two languages.
2. Two languages are spoken by everyone in this room.

what predicates?
what connectives?
what quantifiers?

Logic: Venn diagrams



Socrates is a man.

man (Socrates)

therefore

Socrates is mortal

mortal (Socrates)

All men are mortal.

$\forall$**x: man (x)** $\rightarrow$ **mortal (x)**

VALID LOGICAL DEDUCTION

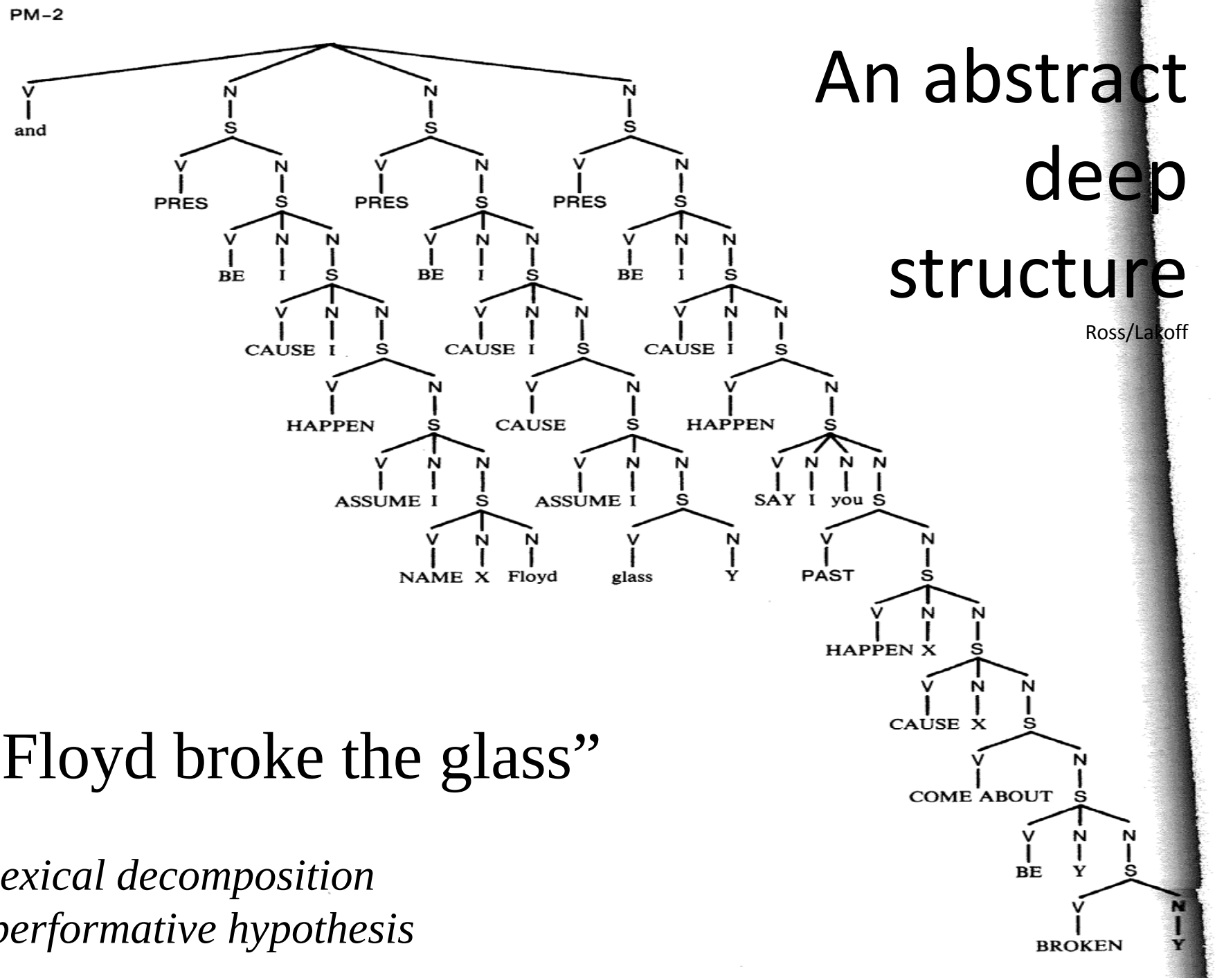
Componential Meanings for verbs…

- Dowty…
  - The door is open. open(x)
  - The door opens. become(open(x))
  - John opened the door. cause(John, become(open(x)))

–Statives, inchoatives, causatives…

An abstract deep structure

Ross/Lakoff



"Floyd broke the glass"

- *lexical decomposition*
- *performative hypothesis*

Linguistic Semantics is not just logic though…

- Typically bundled into ‘frames’

John kicked the ball on Tuesday

- **Frame** semantics (Fillmore)

‘Davidsonian’ semantics

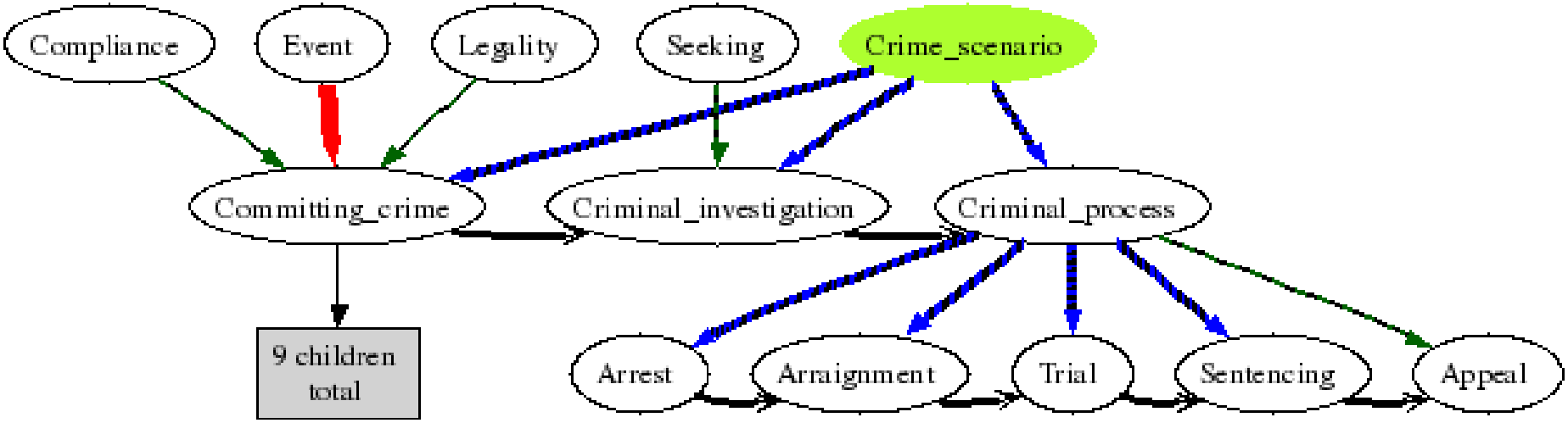
Charles Fillmore: ‘case grammar’

John kicked the ball on Tuesday

kick (j,b,t)

- doesn’t really help us put the meaning together out of the parts
- doesn’t really seem ‘ontologically’ appropriate: we know more than is explicitly represented here…

Frame Semantics

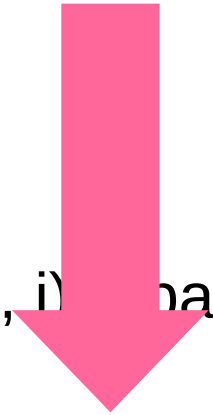


e.g, FrameNet

‘Davidsonian’ semantics

Charles Fillmore: ‘case grammar’

John kicked the ball on Tuesday



$\exists$ **e** event (e) ∧ actor (e, j) ∧ patient (e, b) ∧ time (e, t)

event ∧ has.actor (j) ∧ has.patient (b) ∧ has.time (t)

Description Logic

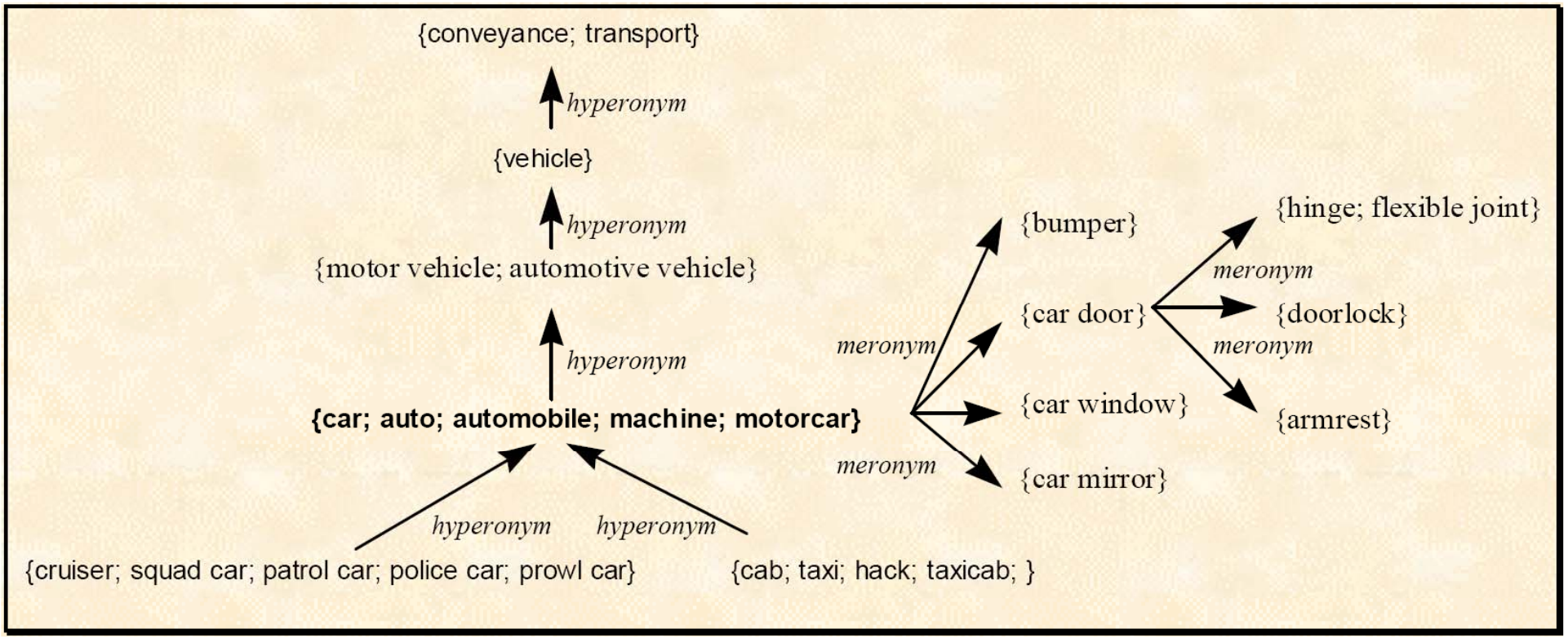
Using a Description Logic

“Person all of whose children are Doctors and/or have a child who is a Doctor”

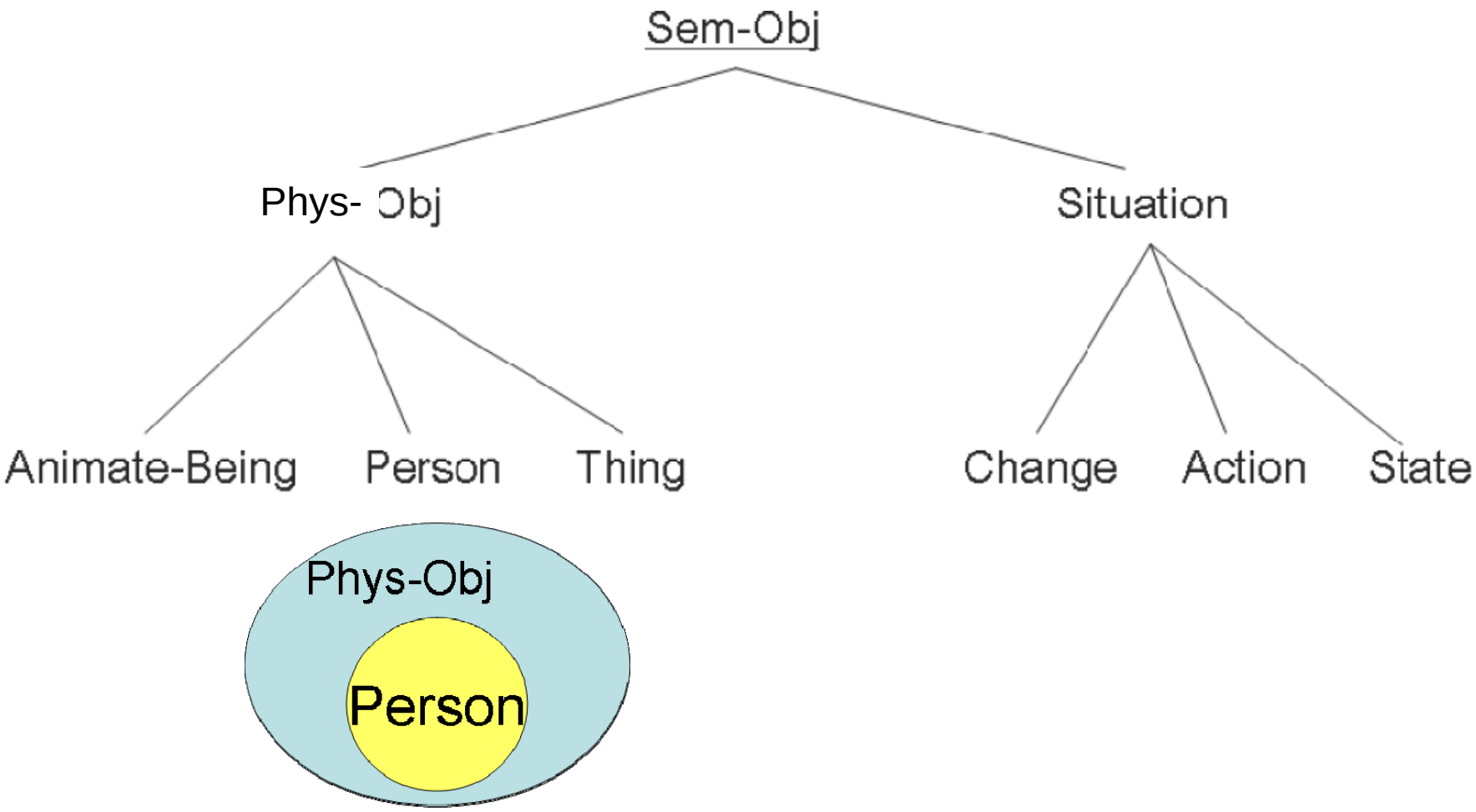
$$\text{Person} \sqcap \forall \text{hasChild}.(\text{Doctor} \sqcup \exists \text{hasChild}.\text{Doctor})$$

person ∧ has.Child (j)
j: doctor ∨ has.Child (b)
b: doctor

WordNet organisation



Semantic Hierarchies: Ontologies

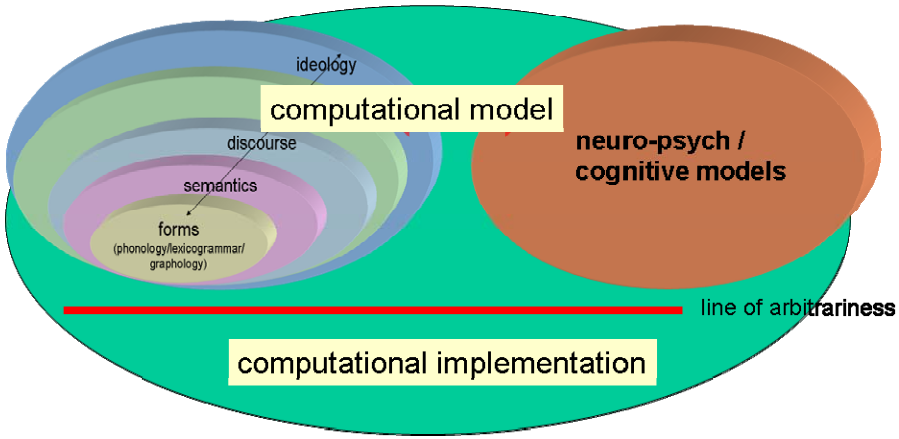


Basic computational modelling tools

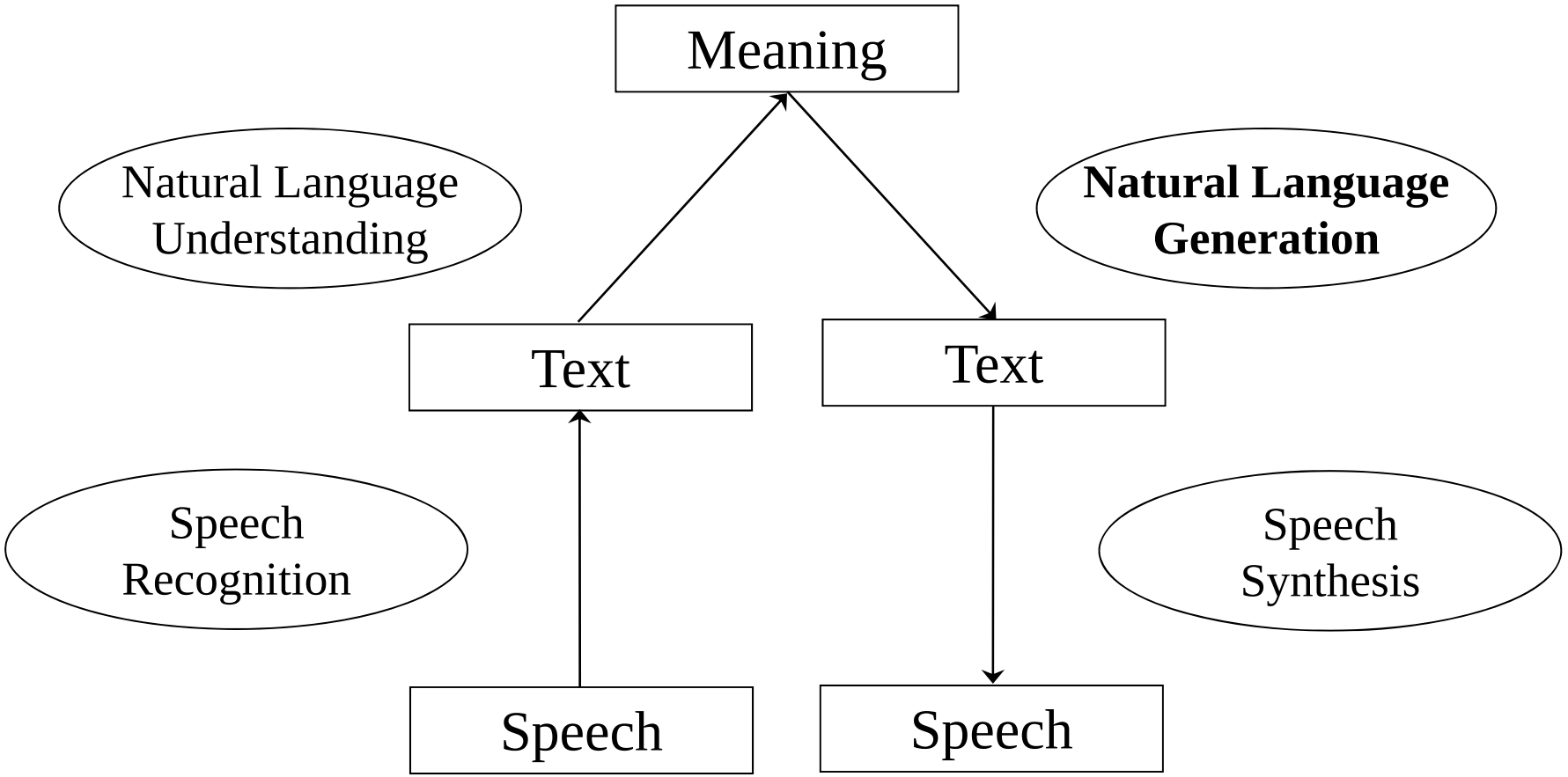
- basic computational linguistic tools for representing **semantics**
  - **logic**
  - **event-based semantics**
  - **ontologies**
  - lambda expressions

- generation
- analysis

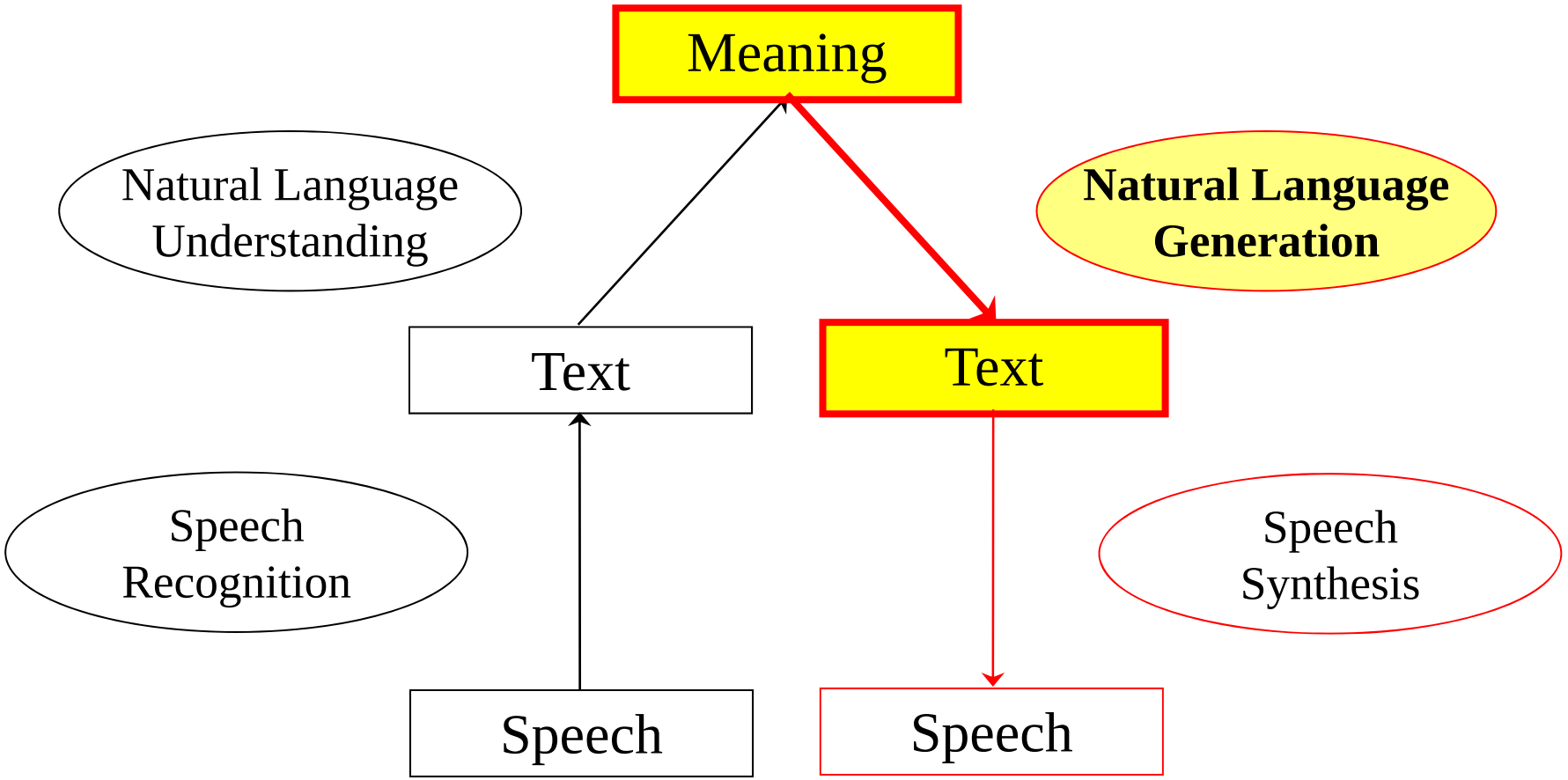
- basic computational linguistic tools for representing **information**
  - feature structures
  - unification



Language Processing Tasks



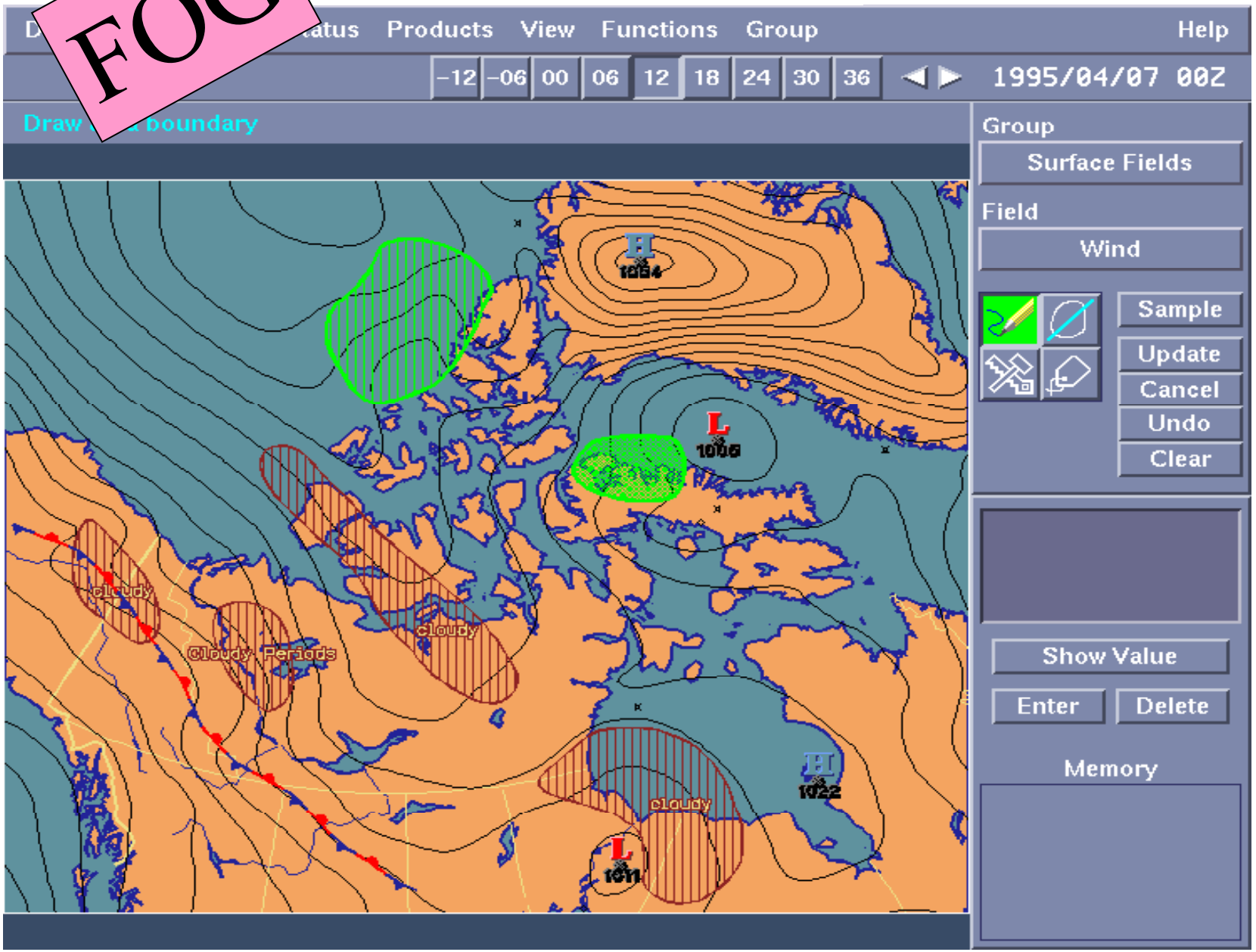
Language Processing Tasks



Early System: FoG

- Function:
  - Produces textual weather reports in English and French
- Input:
  - Graphical/numerical weather depiction
- User:
  - Environment Canada (Canadian Weather Service)
- Developer:
  - CoGenTex
- Status:
  - Fielded, in operational use 1992 onwards

FoG: Input


FOG
Products
View
Functions
Group
Help
-12
-06
00
06
12
18
24
30
36
1995/04/07 00Z
Boundary
Group
Surface Fields
Field
Wind
Sample
Update
Cancel
Undo
Clear
Show Value
Enter
Delete
Memory
H
1024
L
1006
cloudy
Cloudy Periods
cloudy
H
1022
cloudy
L
1011

FoG: Output

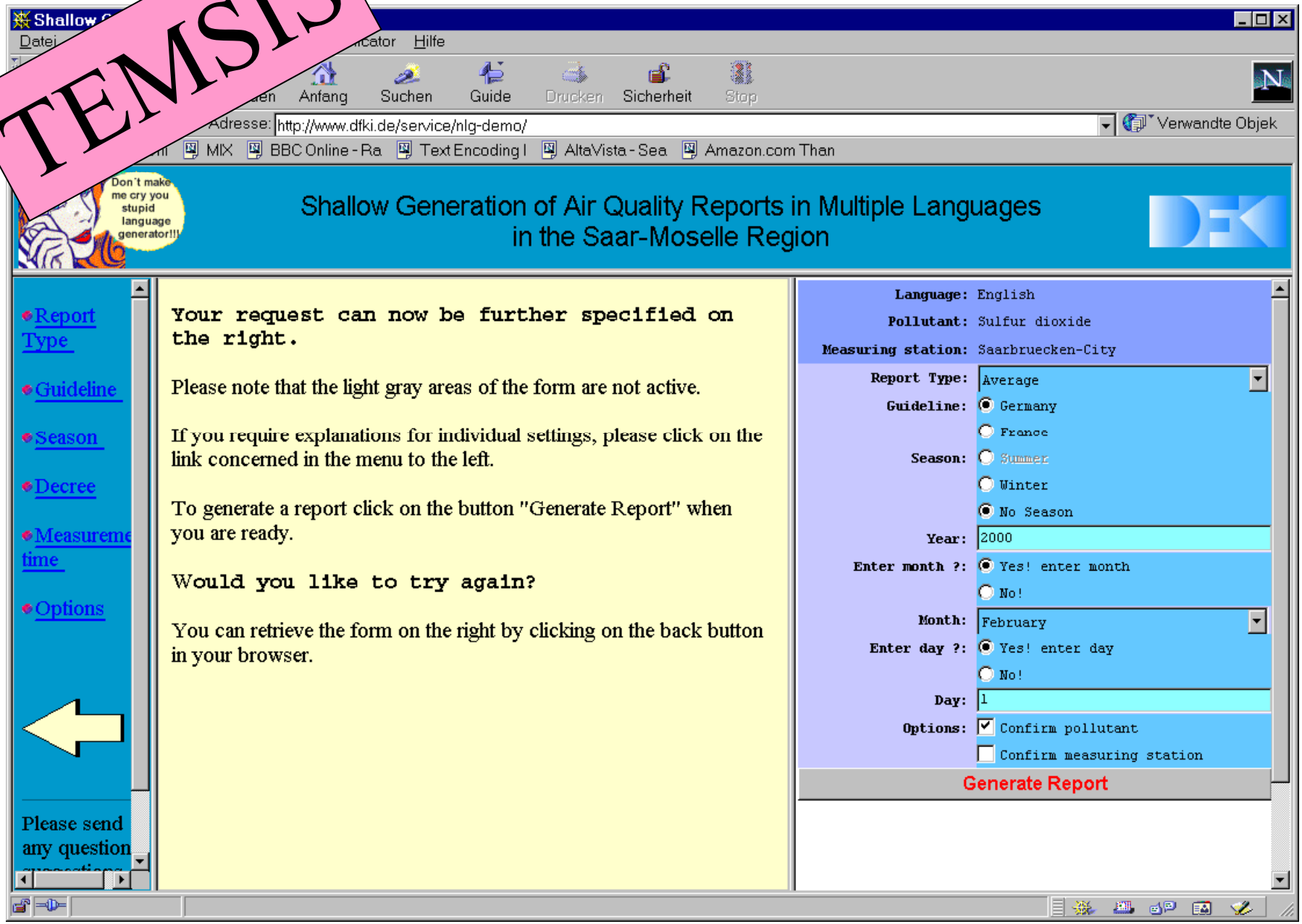

FOG
-NOT RELEASED
FOR ARCTIC WATERS ISSUED BY THE ARCTIC WEATHER CENTRE
CANADA AT 05.00 PM MDT SATURDAY 15 APRIL 1995 FOR TONIGH
WITH AN OUTLOOK FOR MONDAY.
THE NEXT SCHEDULED FORECAST WILL BE ISSUED AT 05.00 AM MDT.
WINDS ARE IN KNOTS.
FOG IMPLIES VISIBILITY LESS THAN 5/8 NM.
MIST IMPLIES VISIBILITY 5/8 TO 6 NM.
GREAT SLAVE LAKE.
WINDS LIGHT TONIGHT AND SUNDAY. SNOW ENDING NEAR MIDNIGHT. VISIBILITIES
NEAR 2 NM IN SNOW.
OUTLOOK FOR MONDAY... LIGHT WINDS.
GREAT BEAR LAKE.
FREEZING SPRAY WARNING ISSUED.
WINDS EAST 20 TO 25 TONIGHT AND SUNDAY. FREEZING SPRAY.
OUTLOOK FOR MONDAY... WINDS EASTERLY 20 TO 25.
MACKENZIE RIVER FROM MILE 0 TO MILE 100.
WINDS LIGHT TONIGHT AND SUNDAY. SNOW ENDING THIS EVENING. VISIBILITIES
NEAR 2 NM IN SNOW.
OUTLOOK FOR MONDAY... LIGHT WINDS.
MACKENZIE RIVER FROM MILE 100 TO MILE 300.
WINDS LIGHT STRENGTHENING TO SOUTHEAST 15 SUNDAY AFTERNOON. SNOW ENDING
EARLY THIS EVENING. VISIBILITIES NEAR 2 NM IN SNOW.
OUTLOOK FOR MONDAY... WINDS SOUTHEASTERLY 15.
Forecasts
-Marine--
* ARWC **
FPCN20
FPCN21
FPCN22/74
FPCN23/75
FPCN24/76
FPCN25/77
UL 22/83
-Public--
FPCN15
Set Element Priority ...
Set Active Areas ...
Source
Working Version
Official Release
Forecast Rollup
Language
English
French
Generate
Update
Edit ...
Release
Print
Close
Help

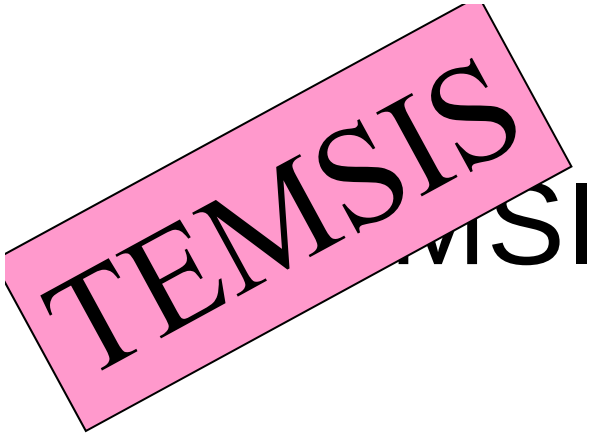
Example System: TEMSIS

- Function:
  - Summarises pollutant information for environmental officials
- Input:
  - Environmental data + a specific query
- User:
  - Regional environmental agencies in France and Germany
- Developer:
  - DFKI GmbH
- Status:
  - Prototype developed; requirements for fielded system being analysed

TEMSIS

http://www.dfki.de/service/nlg-demo/





TEMSIS: Output Summary

- Le 21/7/1998 à la station de mesure de Völklingen - City, la valeur moyenne maximale d'une demi-heure (Halbstundenmittelwert) pour l'ozone atteignait 104.0 µg/m³. Par conséquent, selon le decret MIK (MIK-Verordnung), la valeur limite autorisée de 120 µg/m³ n'a pas été dépassée.

- Der höchste Halbstundenmittelwert für Ozon an der Meßstation Völklingen -City erreichte am 21. 7. 1998 104.0 µg/m³, womit der gesetzlich zulässige Grenzwert nach MIK-Verordnung von 120 µg/m³ nicht überschritten wurde.

There are many NLG systems; some are available online...

- For example, the following system is a simple weather report generator:

- SUMTIME
  - generation of weather reports
  - http://www.csd.abdn.ac.uk/~ssripada/cgi_bin/StartSMT.html

Basic Generation Problem

- How to go from an abstract semantic input to a concrete linguistic form that is
  - semantically correct
  - stylistically appropriate
  - textually appropriate

???

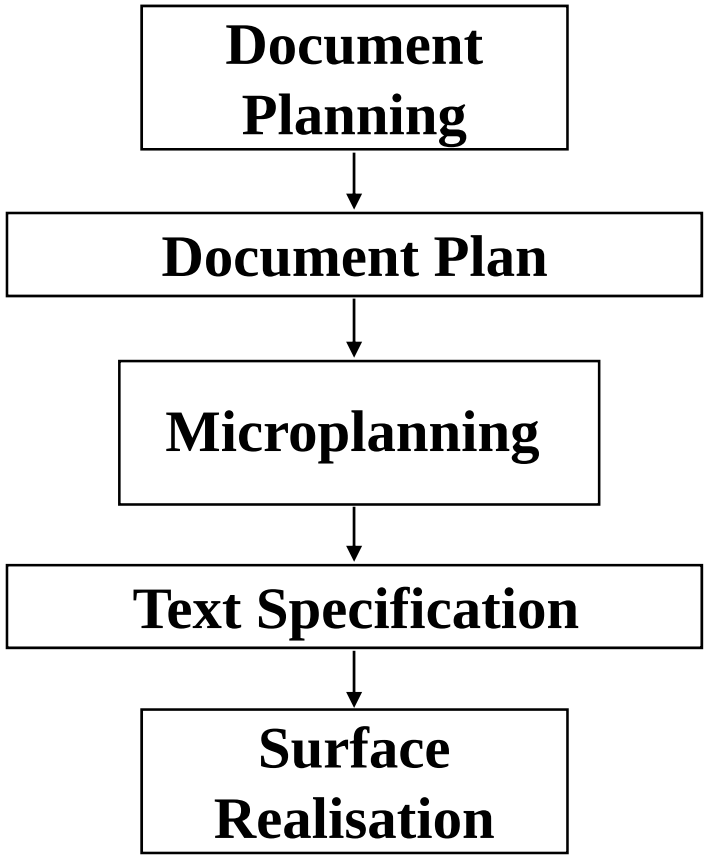
What is NLG?

Natural language generation is the process of automatically constructing a natural language text in order to meet specified communicative goals.

[McDonald 1992]

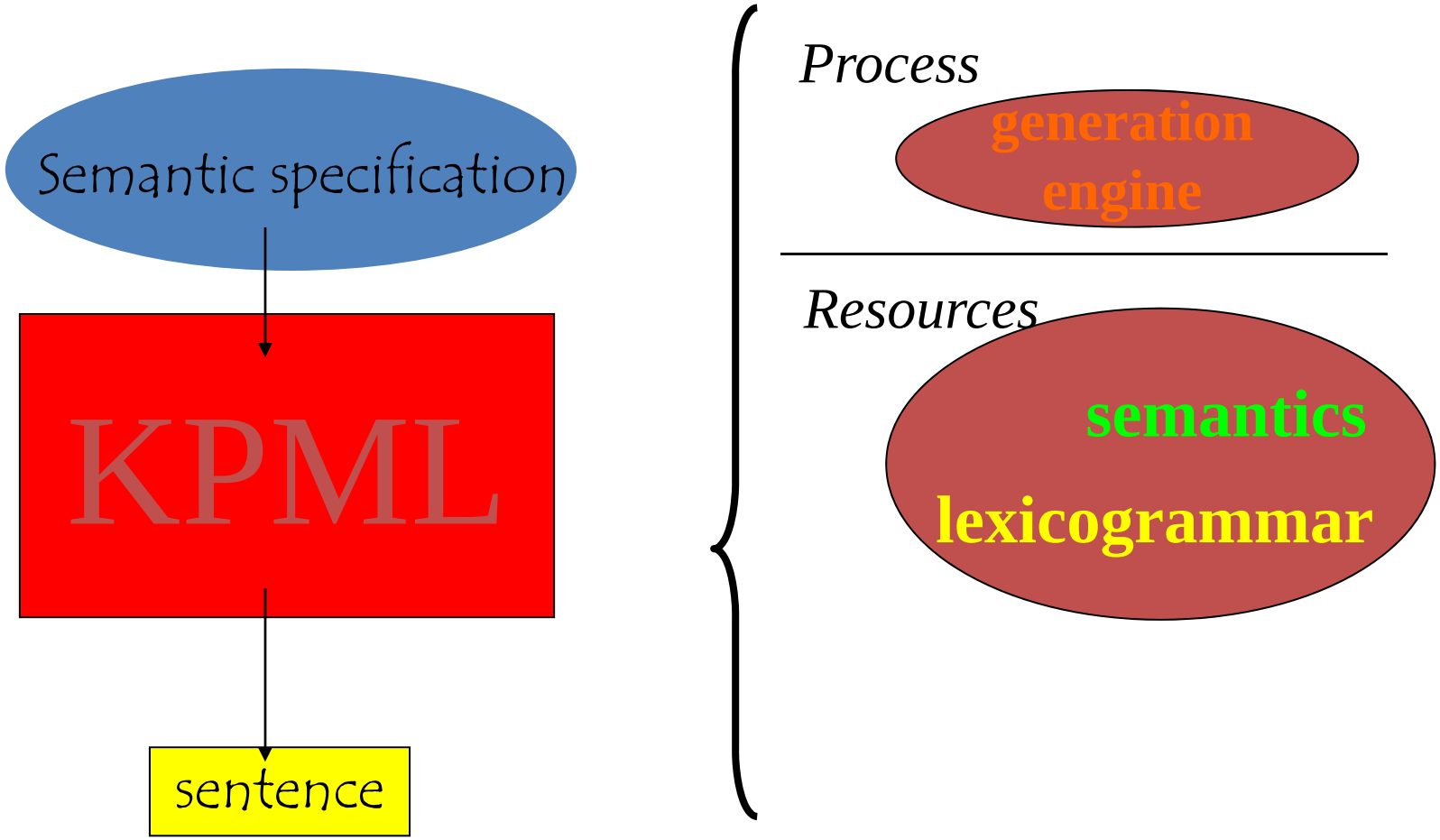
NLG is a process of choice under specified constraints

Standard Pipelined Architecture

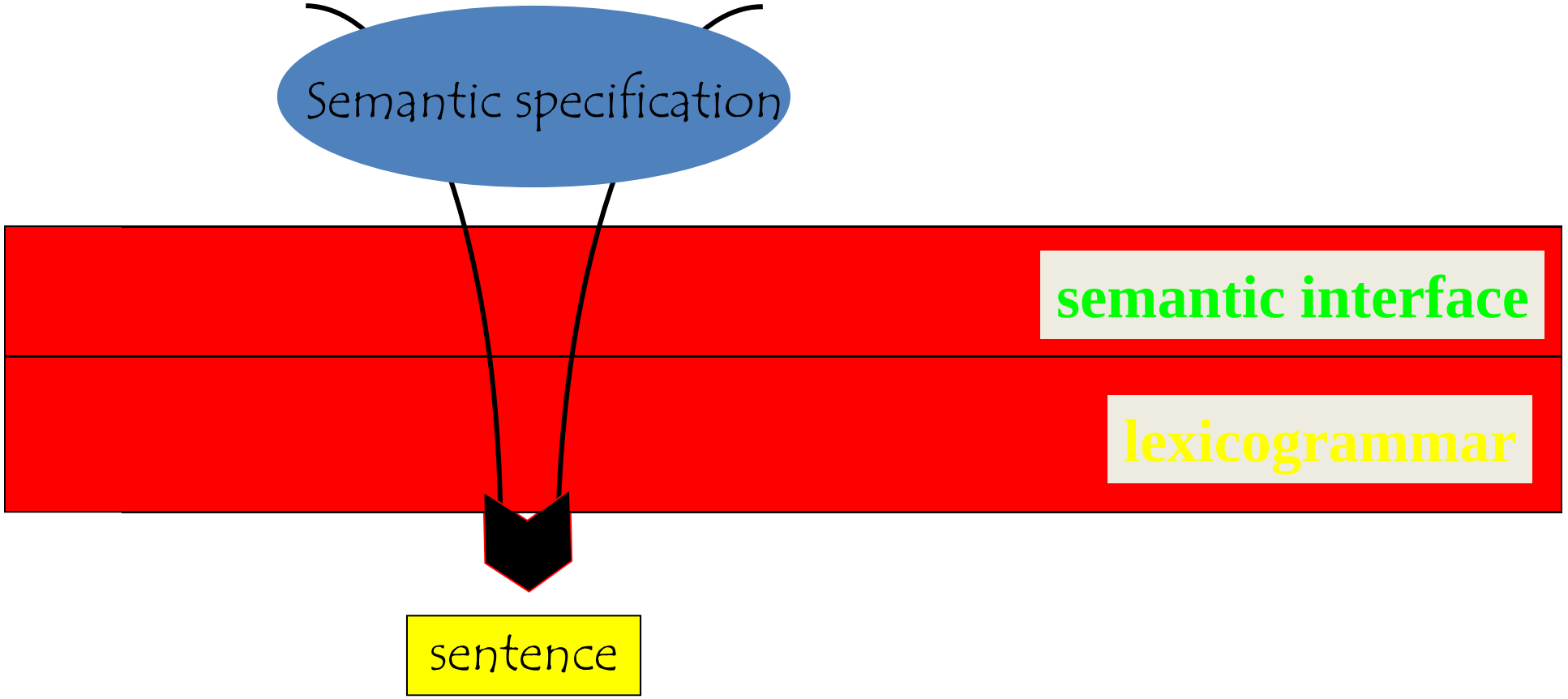


TACTICAL GENERATION

KPML is a # TACTICAL GENERATOR



TACTICAL GENERATION

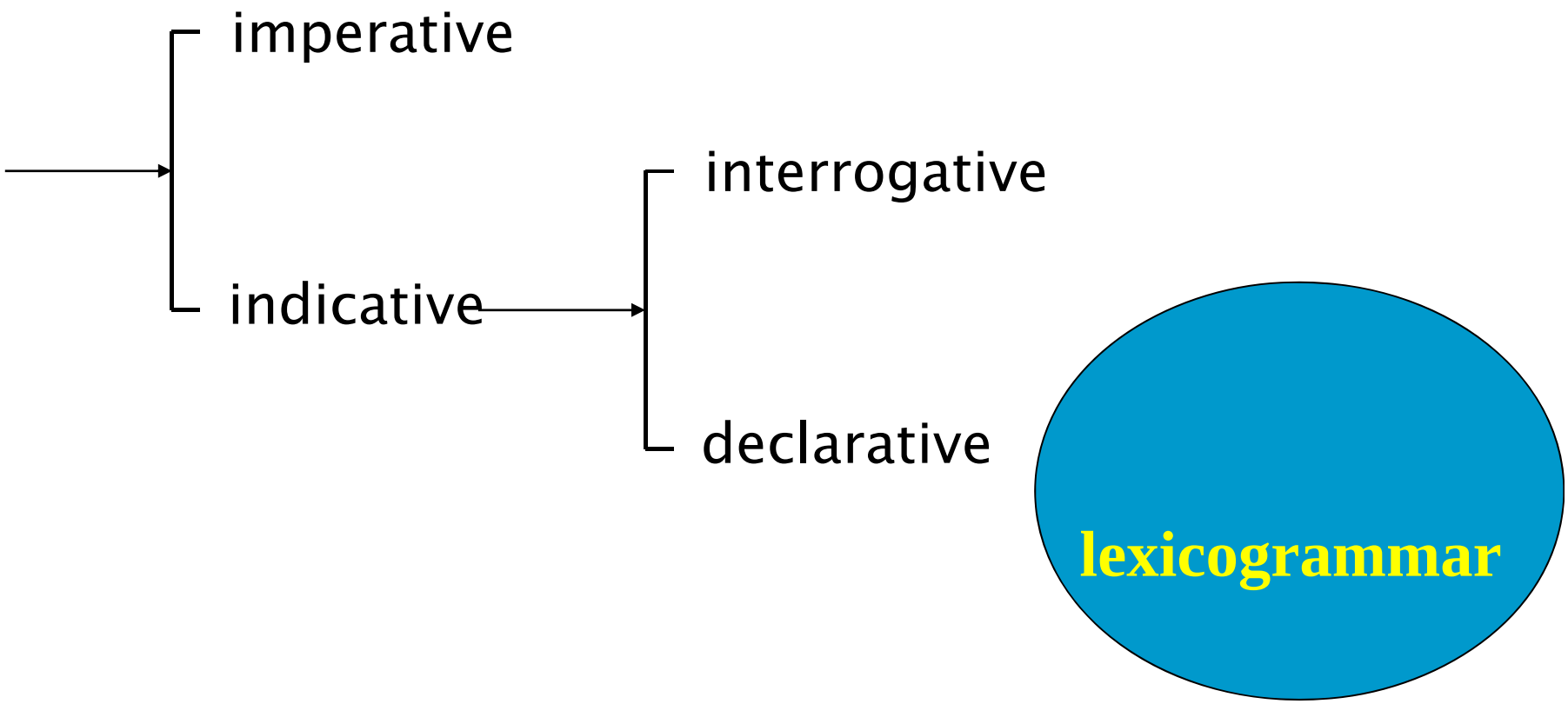


What is NLG?

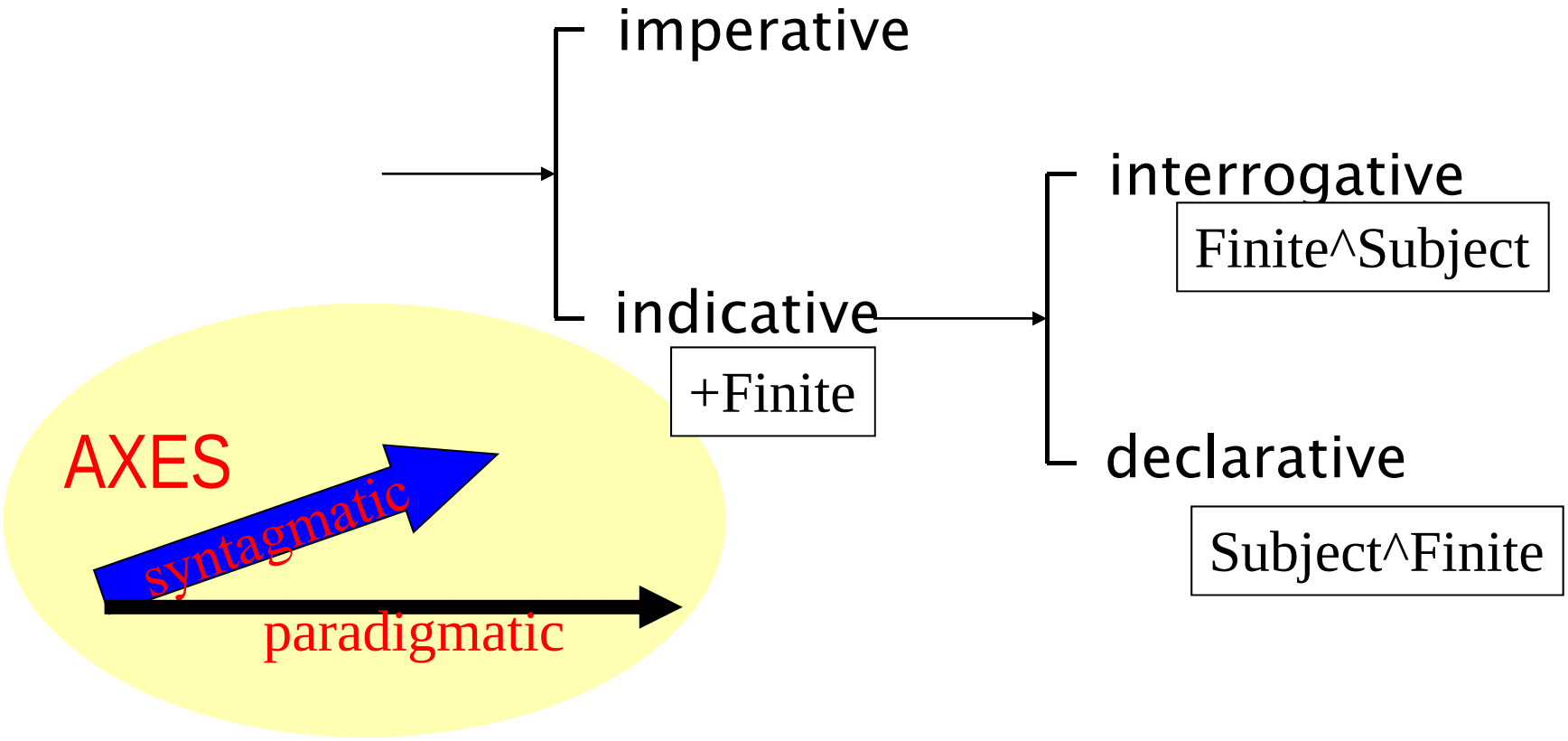
NLG is a process of choice under specified constraints

[McDonald]

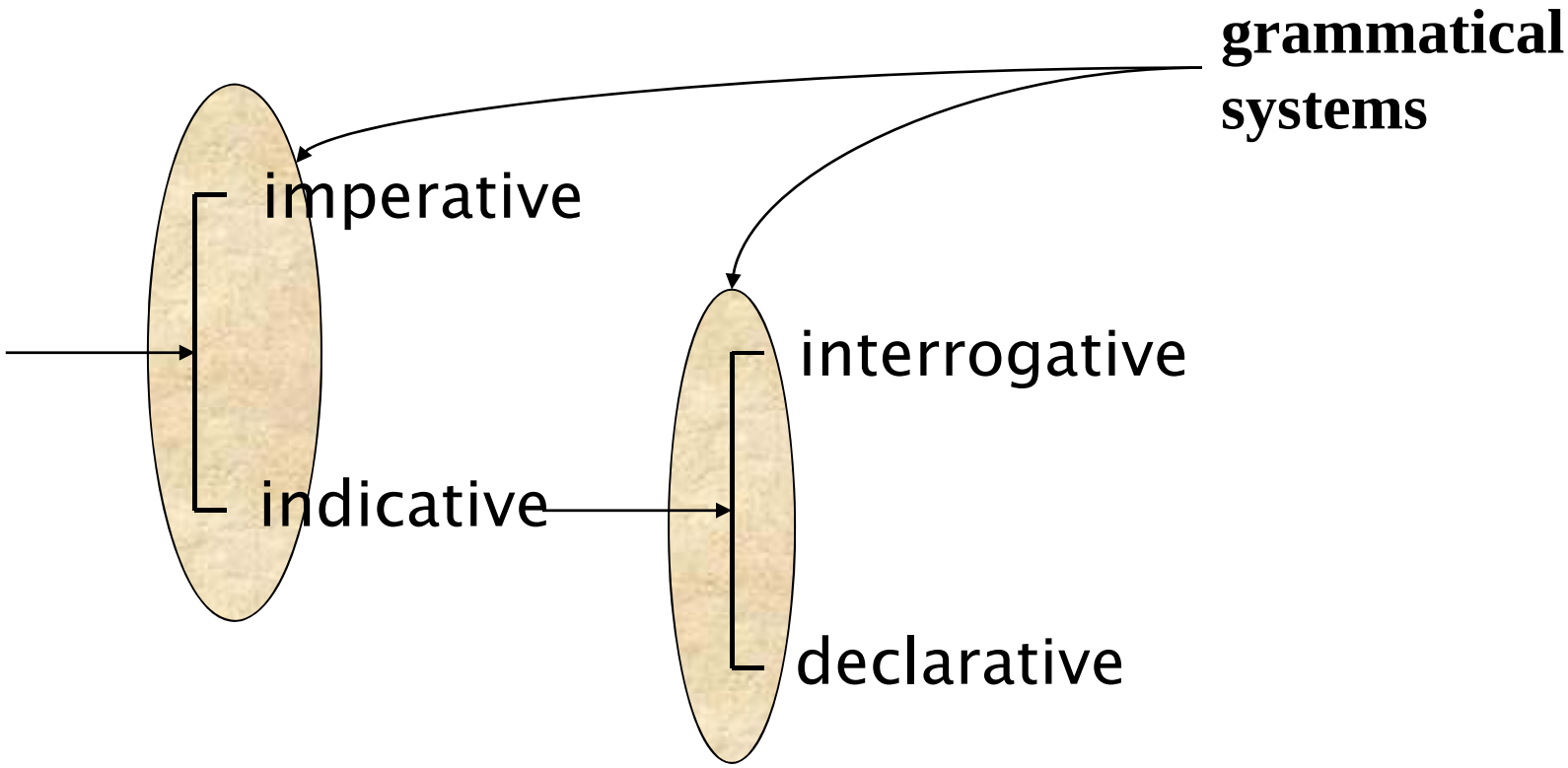
Resource Architecture in KPML: system networks



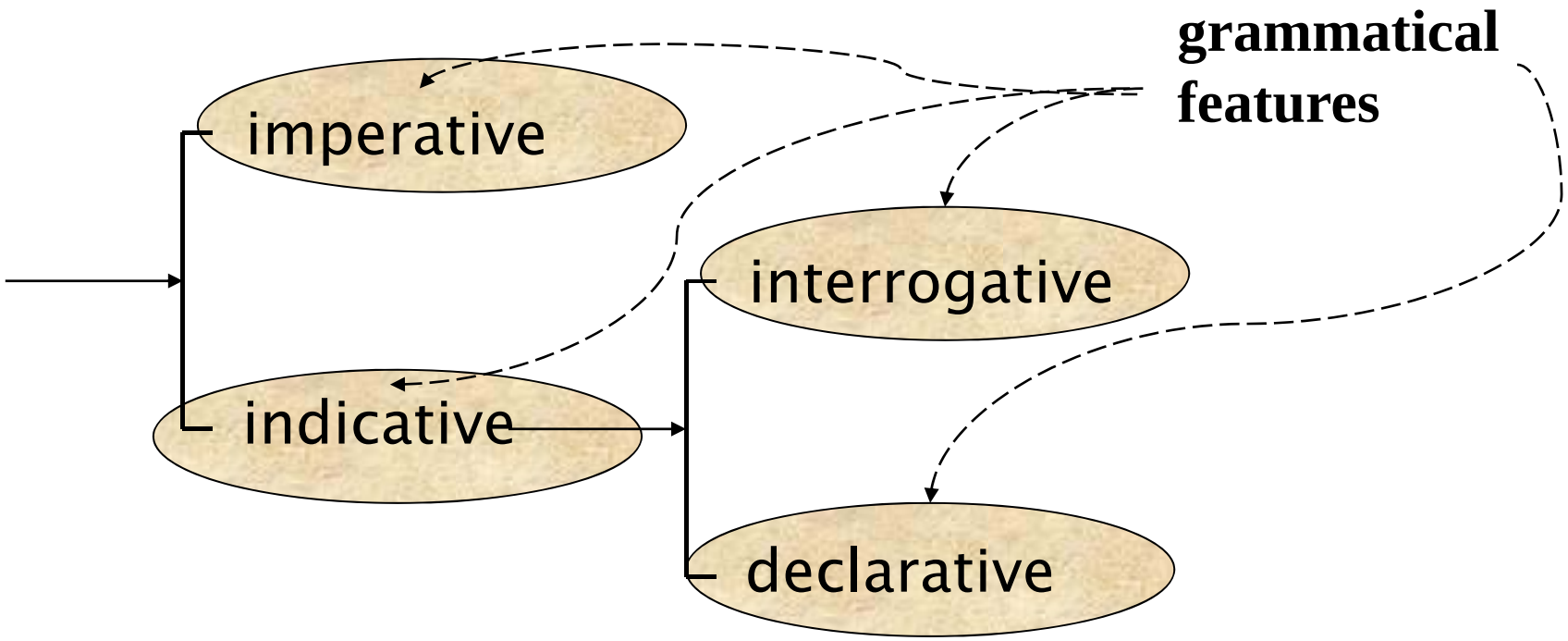
Linguistic Description with Networks of Systems of Choice



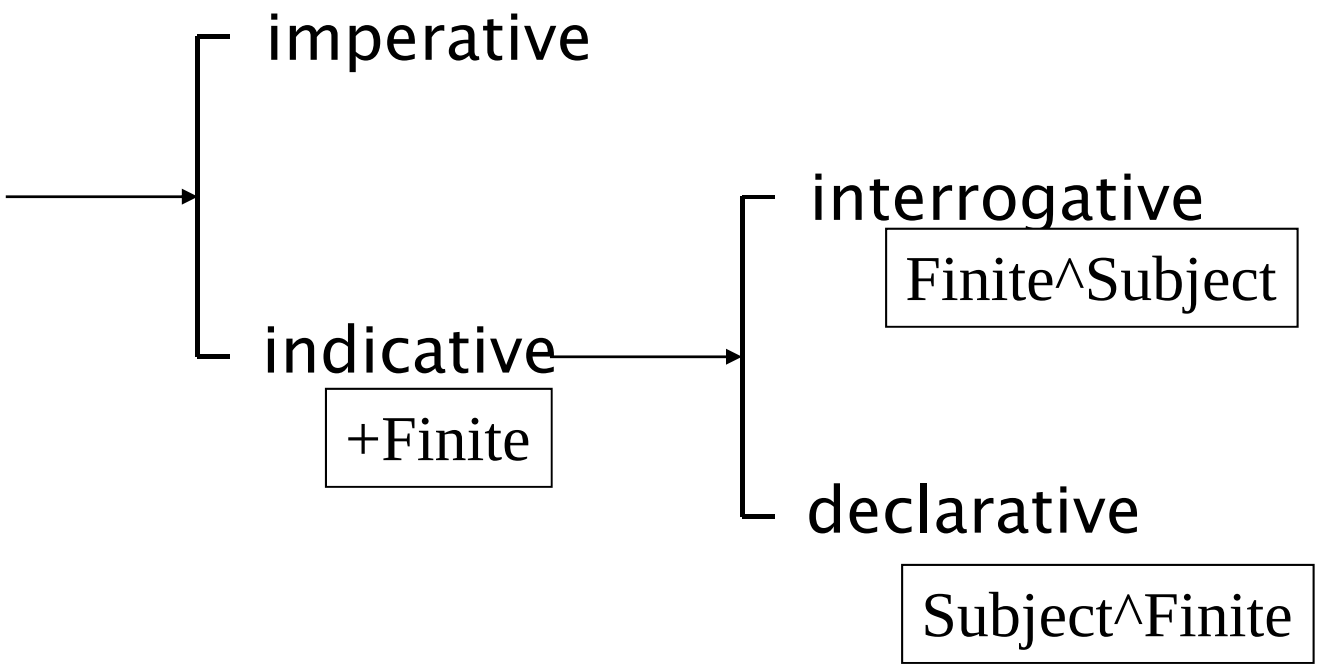
Linguistic Description with Networks of Systems of Choice



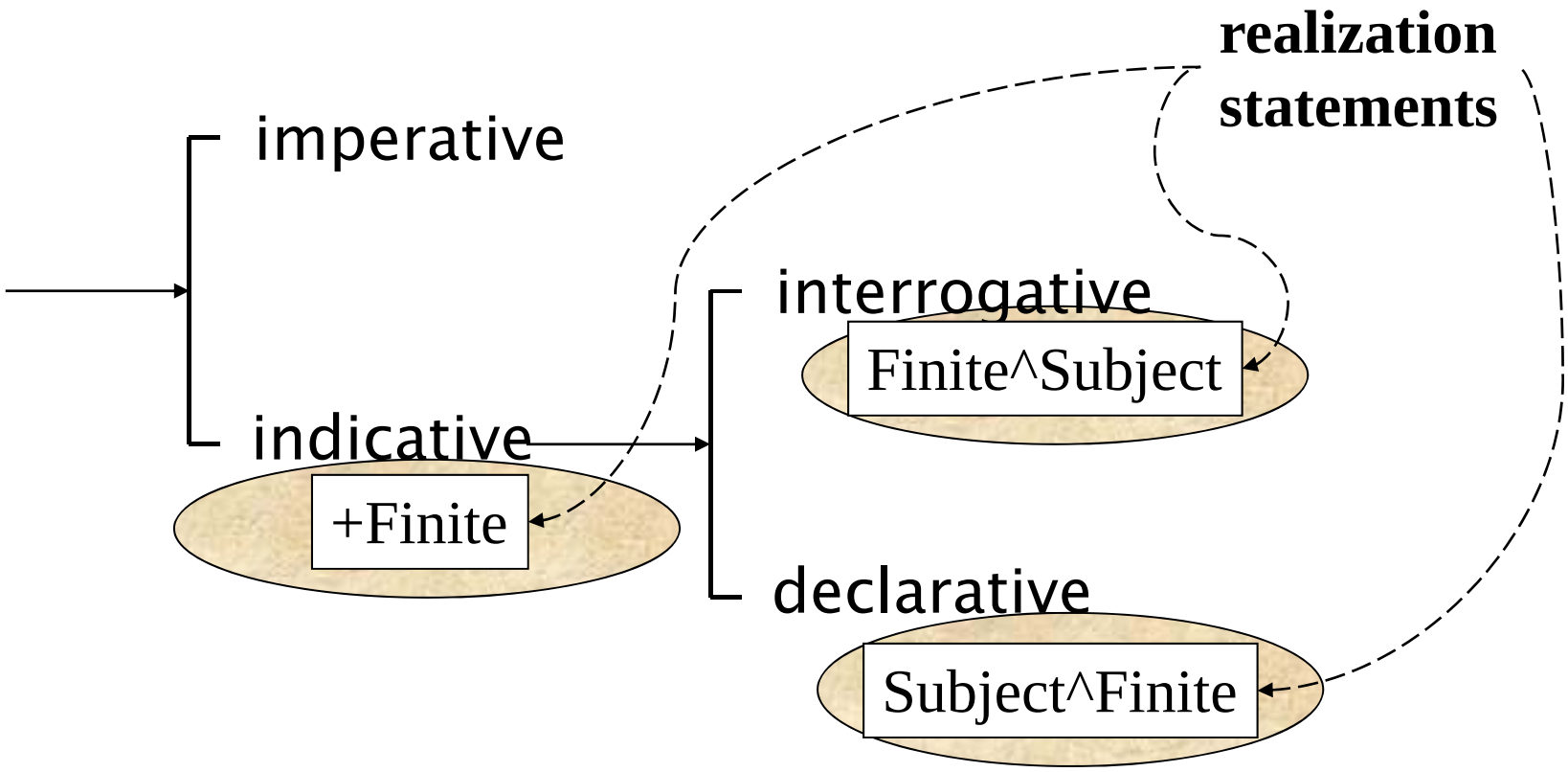
Linguistic Description with Networks of Systems of Choice



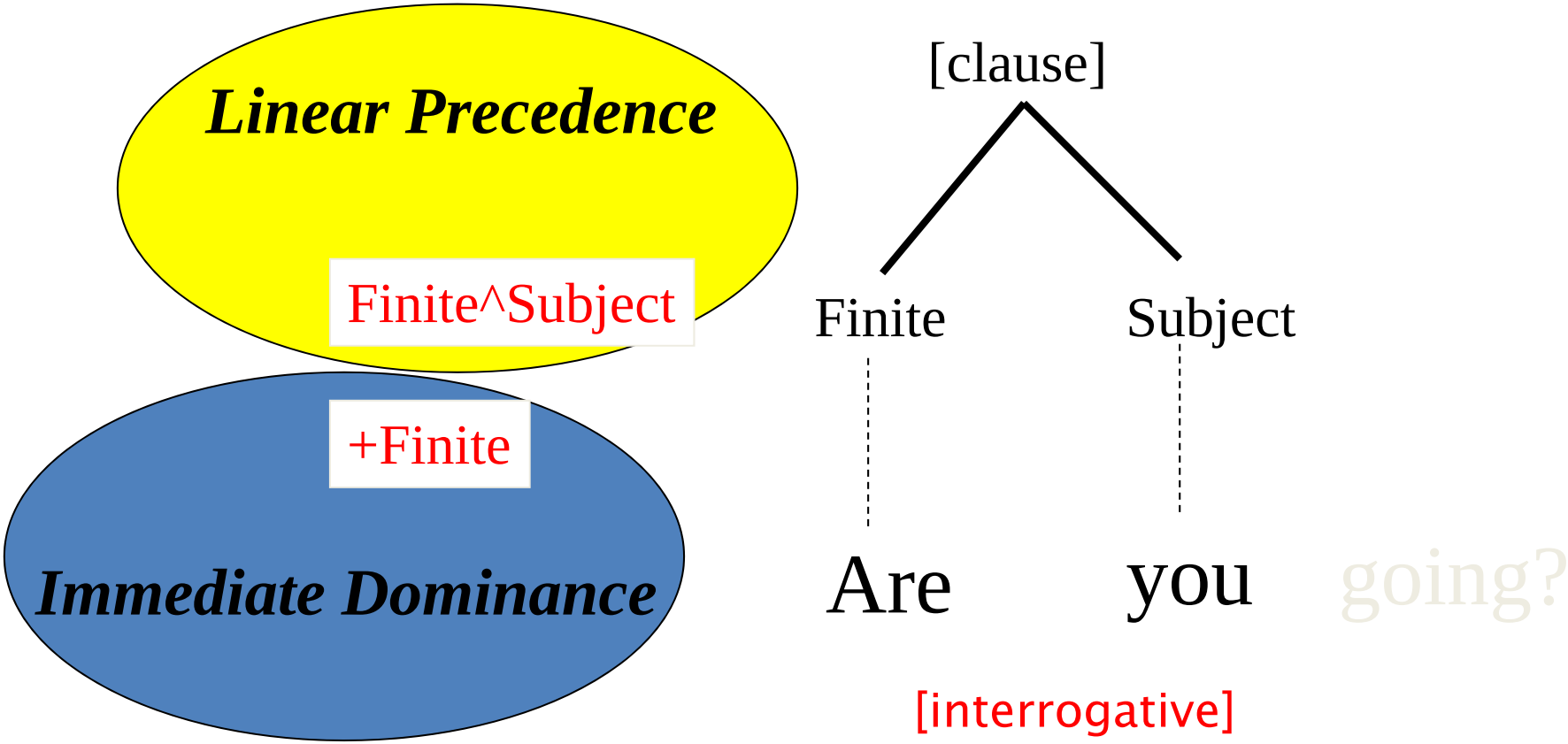
Linguistic Description with Networks of Systems of Choice



Linguistic Description with Networks of Systems of Choice



Generation Process: realization statements

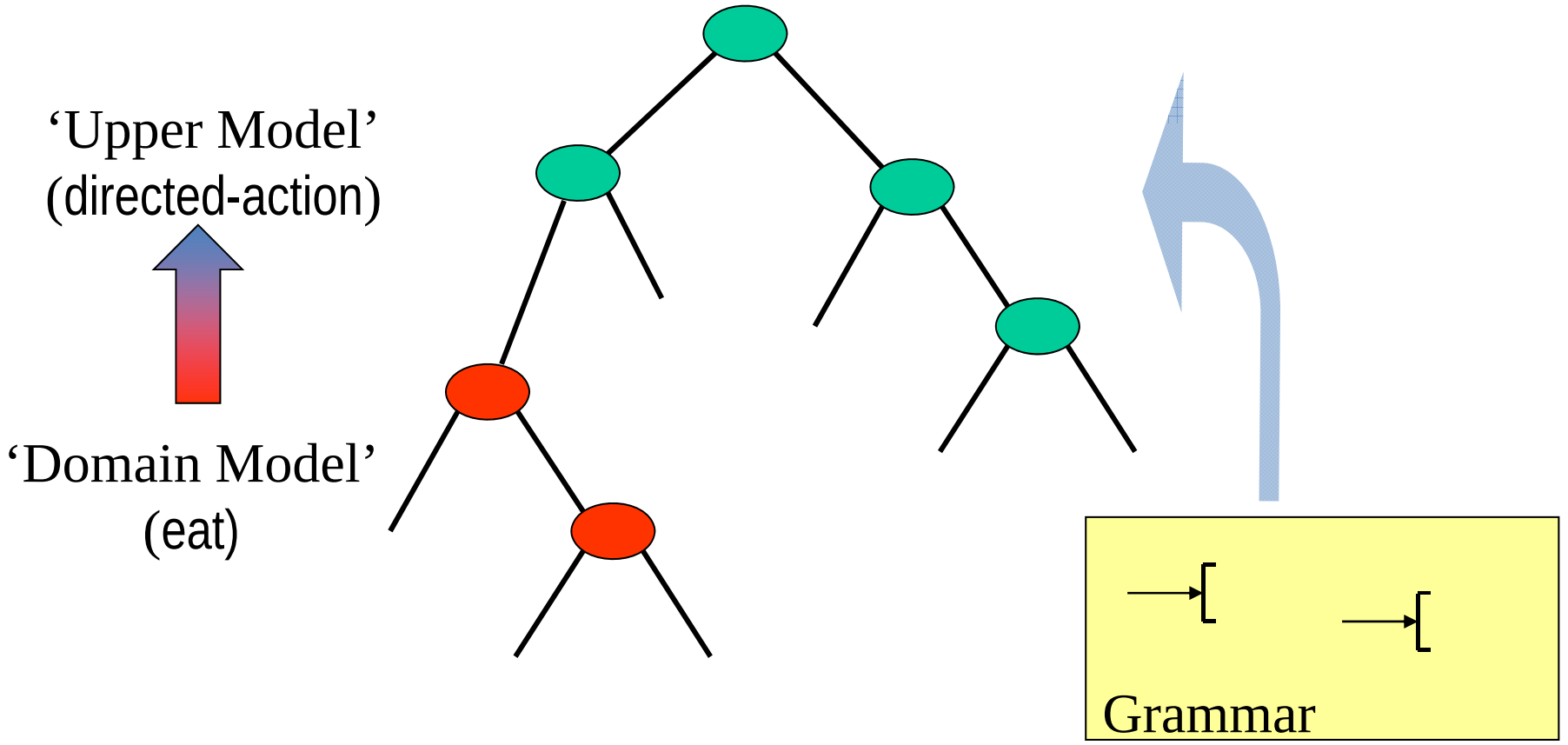


Types of Realization Statements

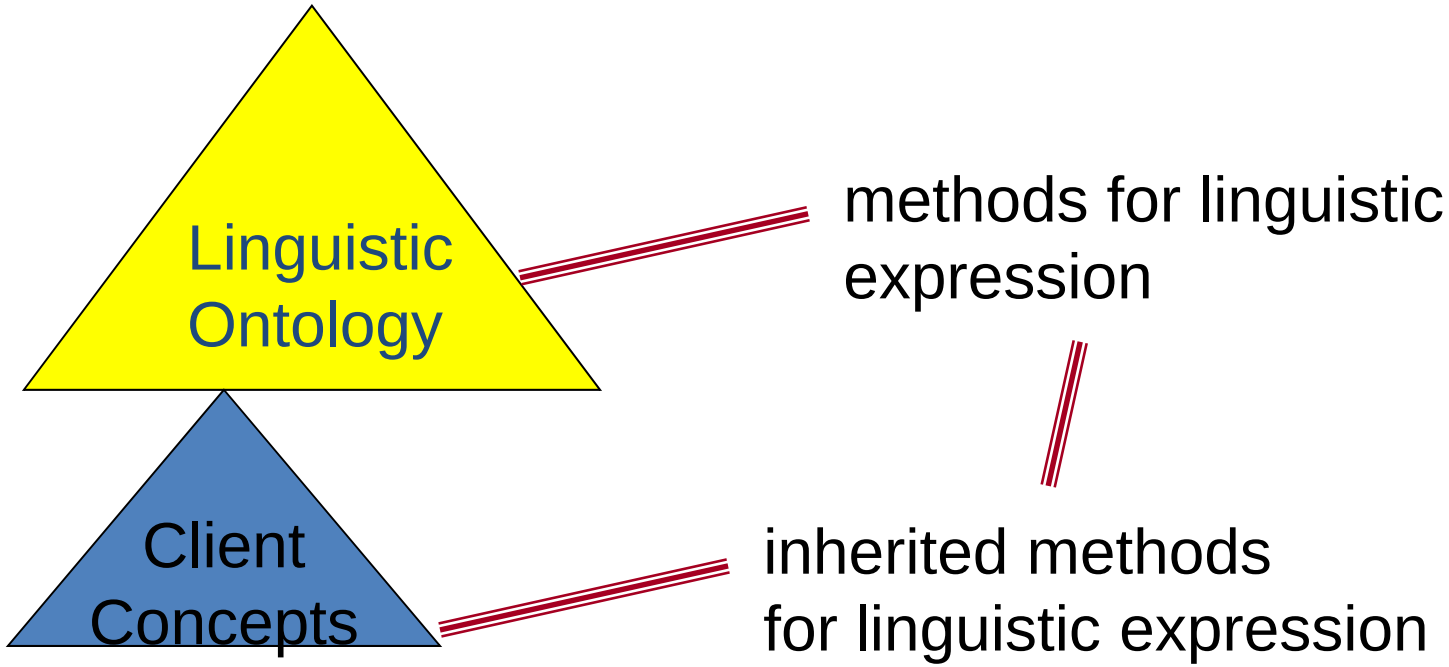
- Ordering (immediate, relative)
- Structure building
- Lexicalization

How to control the generation process…

Generation Control from an ontology

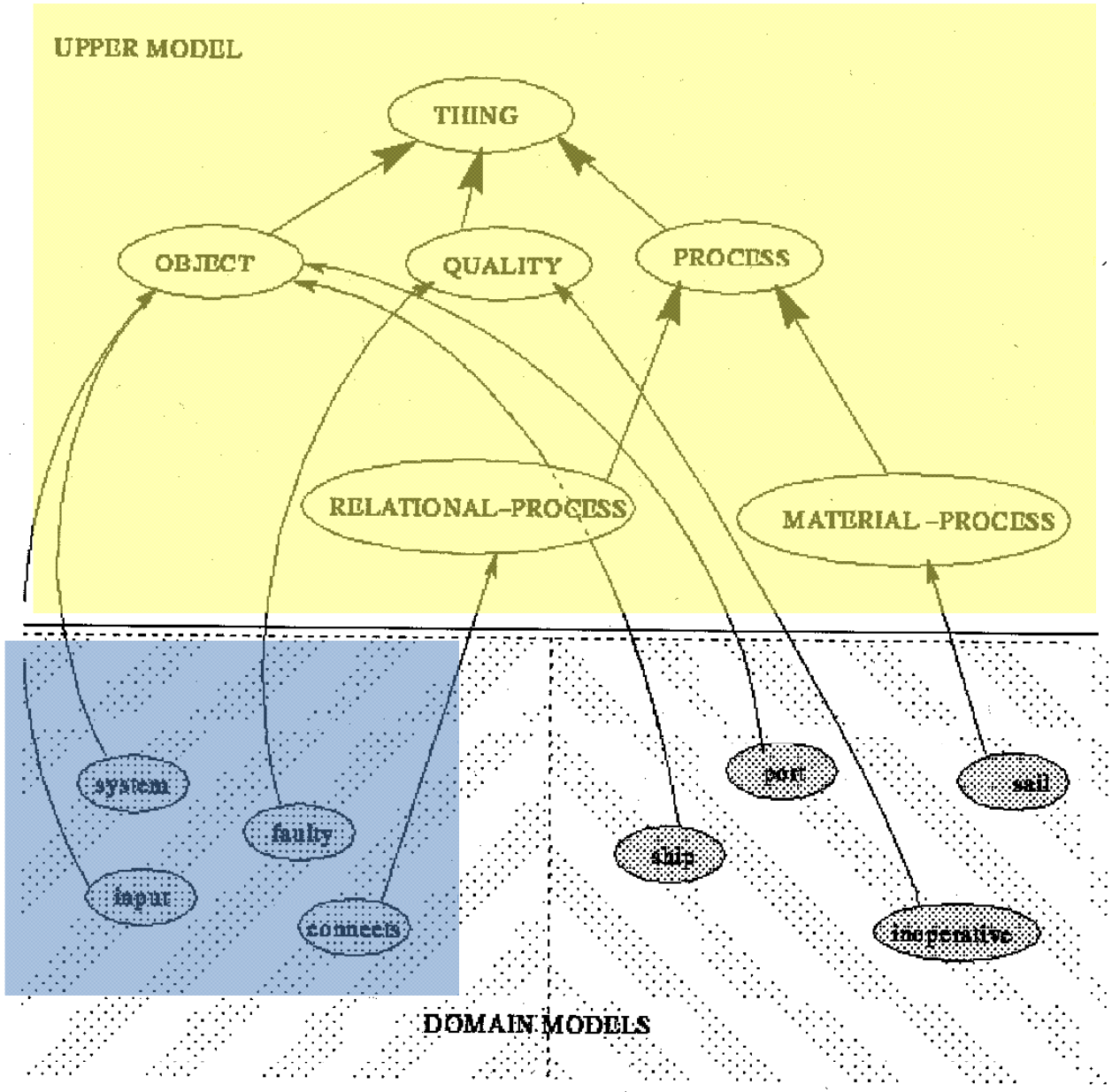


Relations between ontology and client knowledge representation



Relation between upper model and domain models

(1988-1990)



Semantic Input:
Sentence Plan Language (SPL)

"The robot moves to the red block"

```
(x / nondirected-action :lex move
     :actor (r / object :lex robot)
     :destination
          (p / object :lex block
               :color-property-ascription
                    (c / color-quality :lex red))))
```

Upper model concepts

Upper model relations

Example of use: semantic specifications

concepts and relations drawn from the upper model only

"The robot moves to the red block"

```
(x / nondirected-action :lex move
     :actor (r / object :lex robot)
     :destination
          (p / object :lex block
               :color-property-ascription
                    (c / color-quality :lex red))))
```

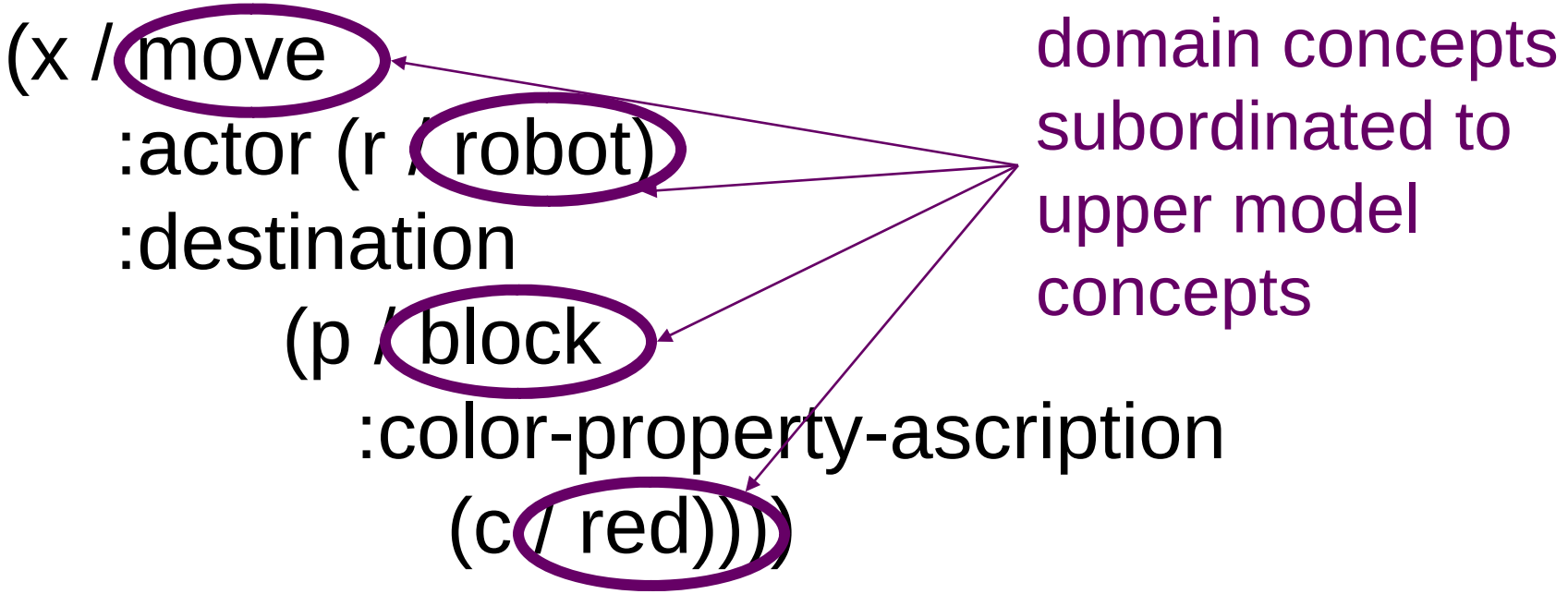
Upper model concepts

Upper model relations

Example of use: semantic specifications

concepts and relations drawn from the domain model

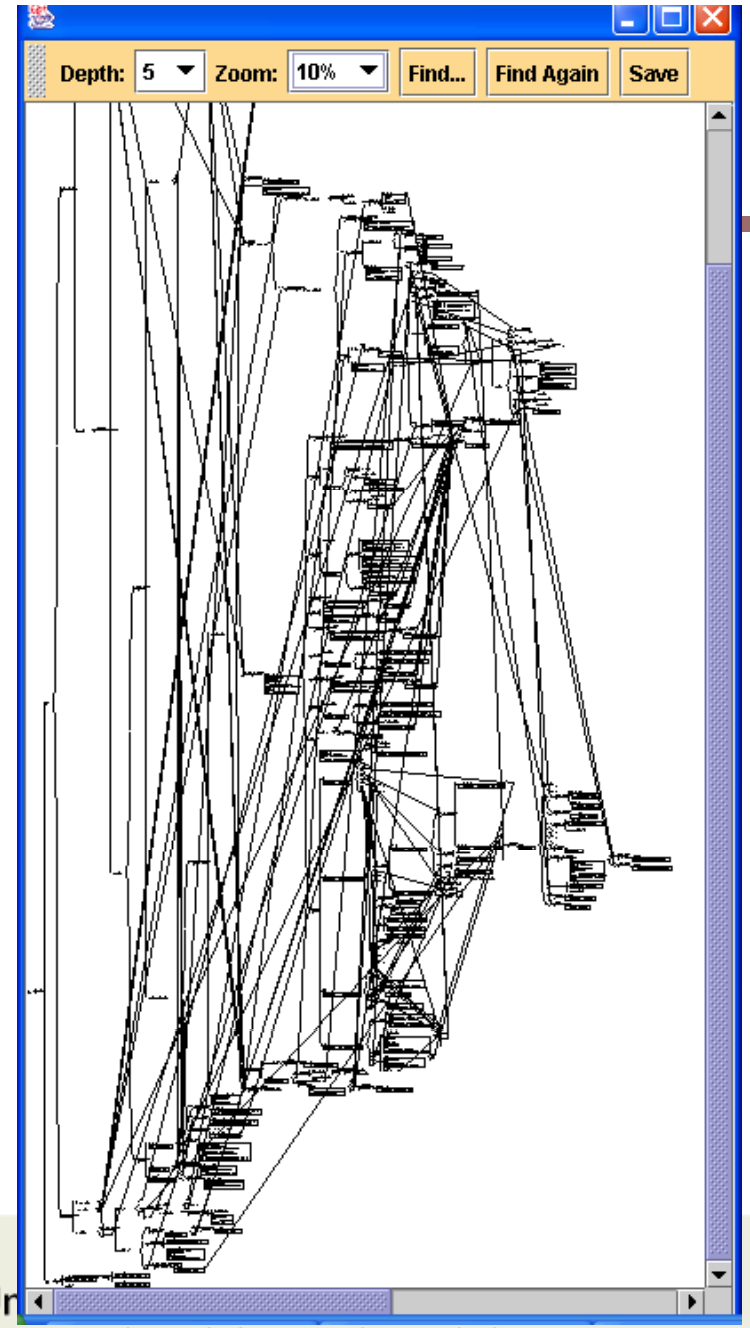
"The robot moves to the red block"



event ∧ has.actor (j) ∧ has.patient (b) ∧ has.time (t)

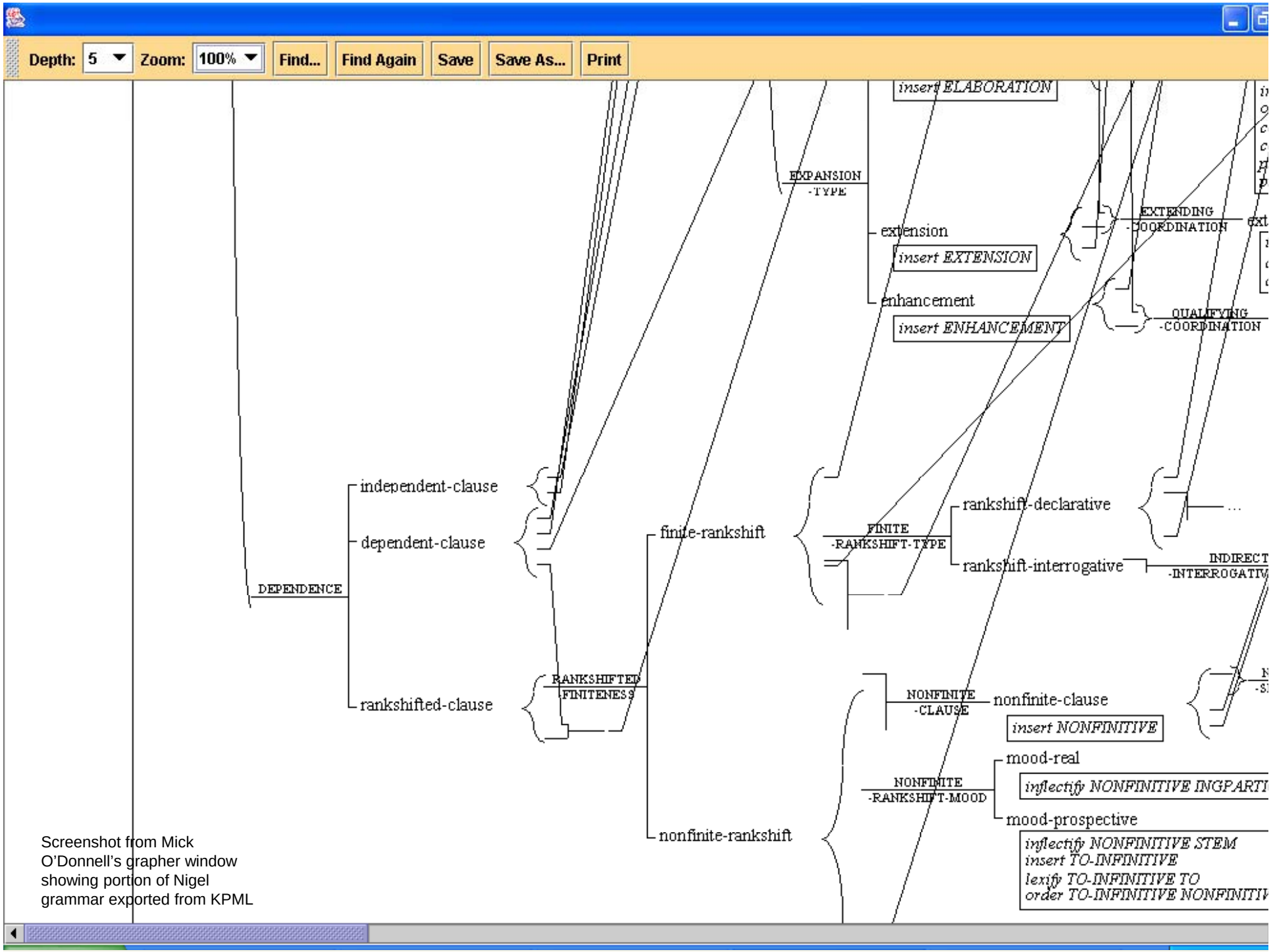
Generating Sentences

- basic event structure
- reliance on ontological knowledge
- directed/non-directed actions
- temporal information
- polarity
- speech acts
- textual appropriateness
  - :identifiability-q identifiable
  - :empty-number-q empty



Paradigmatic organisation is extensive

(for large grammars around 1000 features)

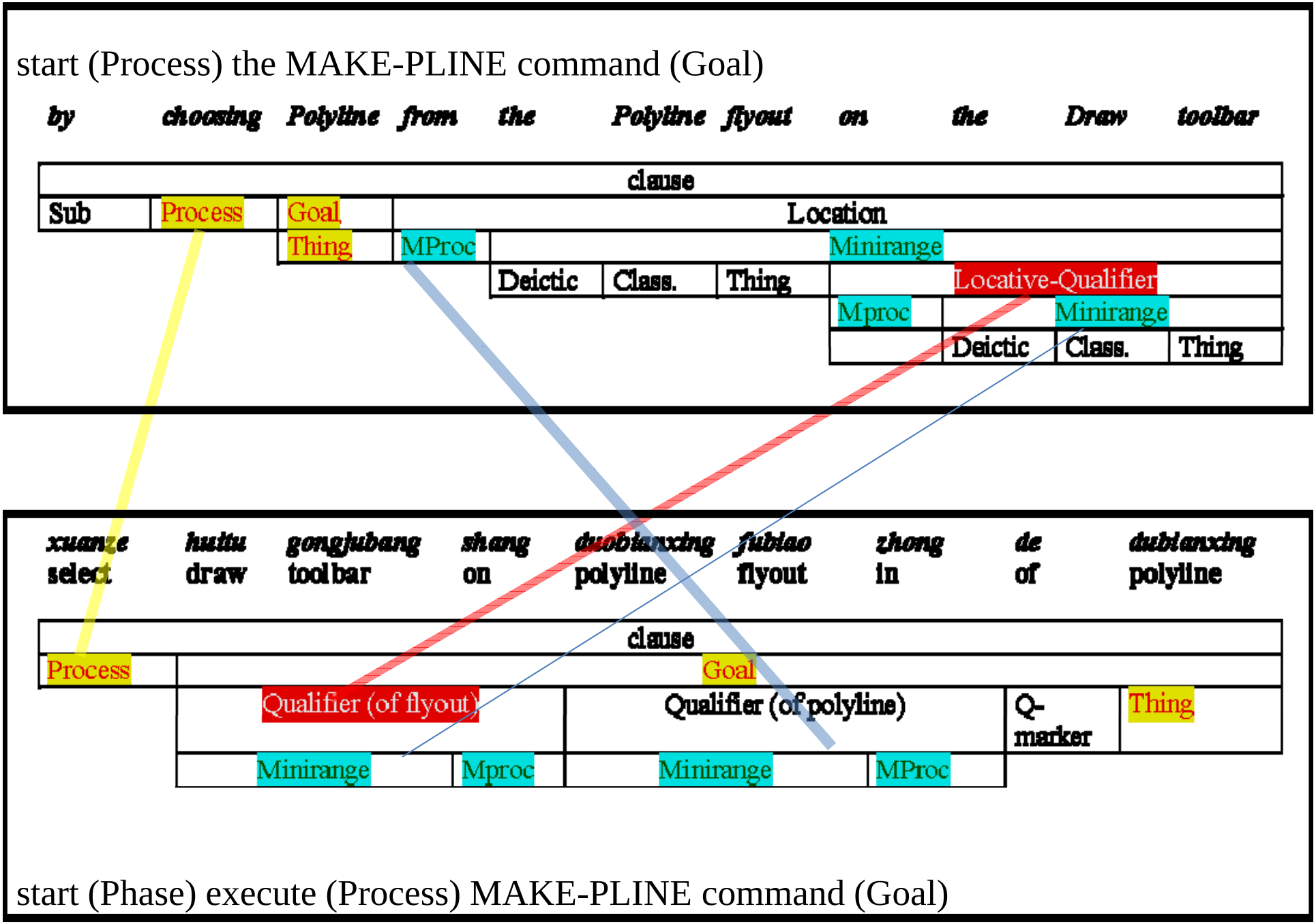


Screenshot from Mick O'Donnell's grapher window showing portion of Nigel grammar exported from KPML

Various Investigations

- Multilingual linguistic description
- Stylistic variation
- Dialogue-appropriate utterances
- …

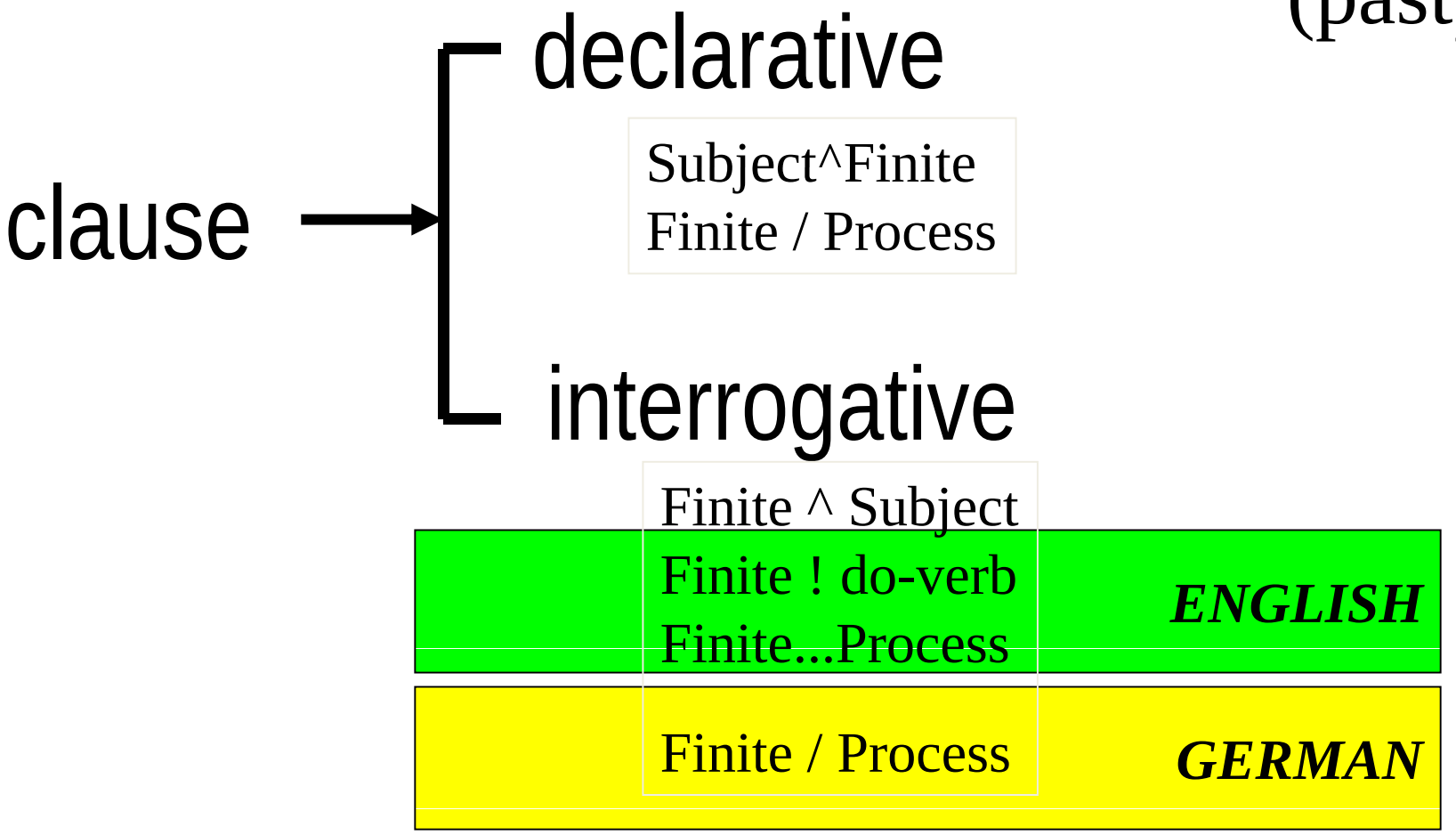
start (Process) the MAKE-PLINE command (Goal)



start (Phase) execute (Process) MAKE-PLINE command (Goal)

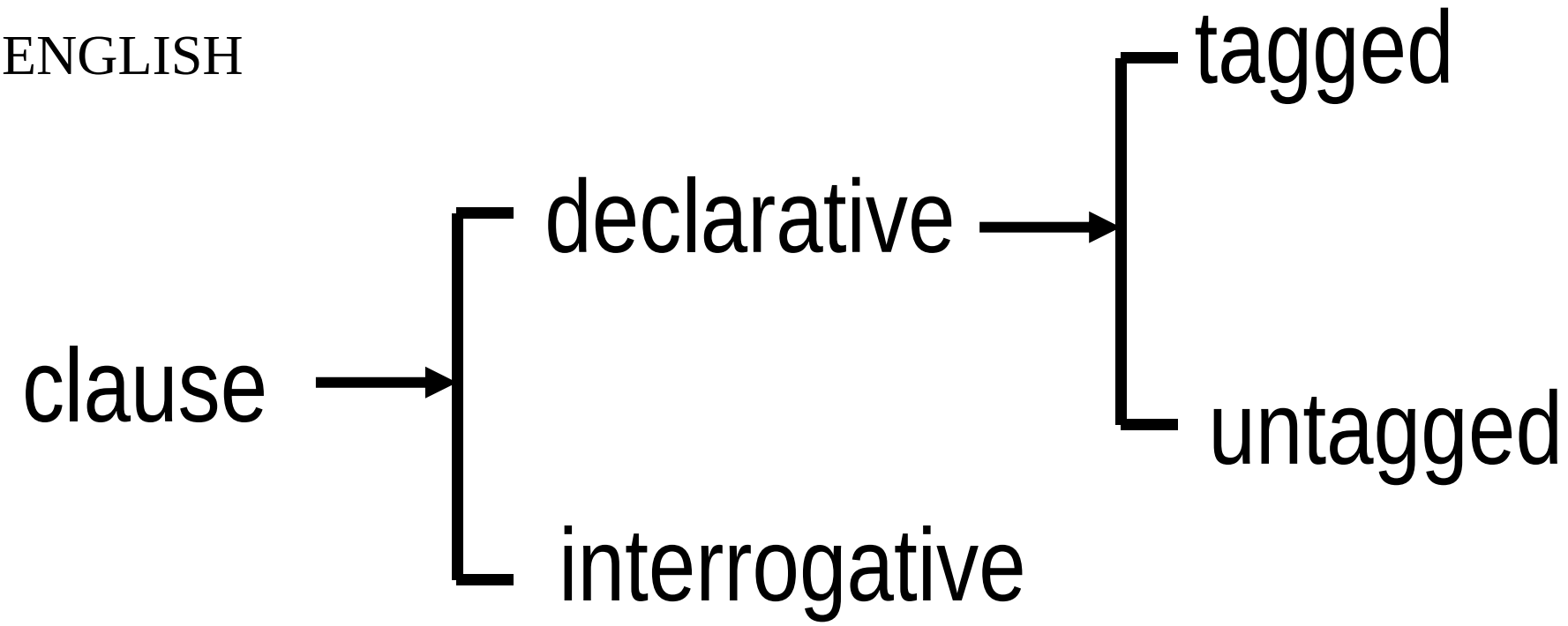
Partial congruence of syntagmatic specifications (potential)

(past)



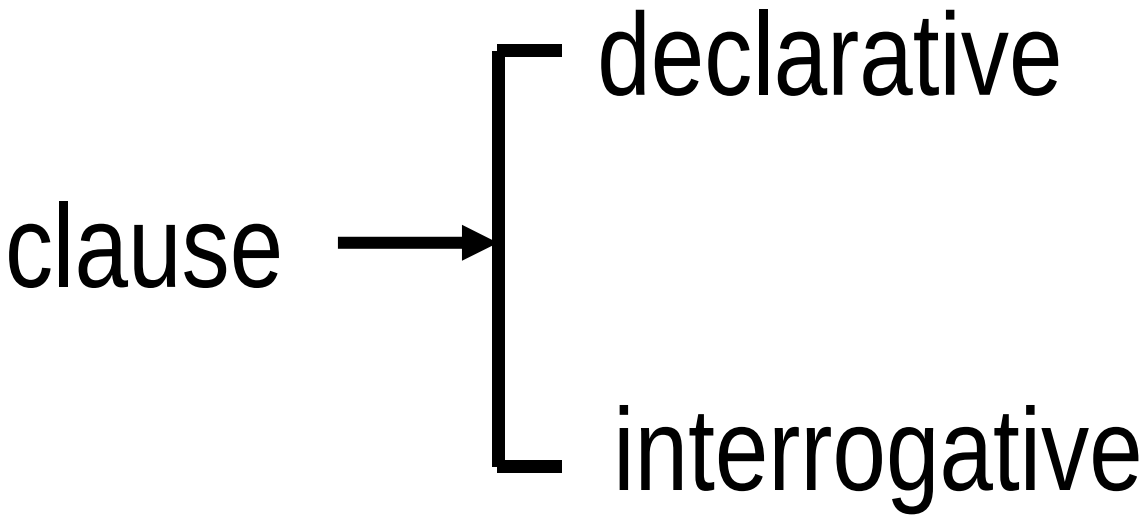
Partial congruence of paradigmatic specifications (potential)

ENGLISH

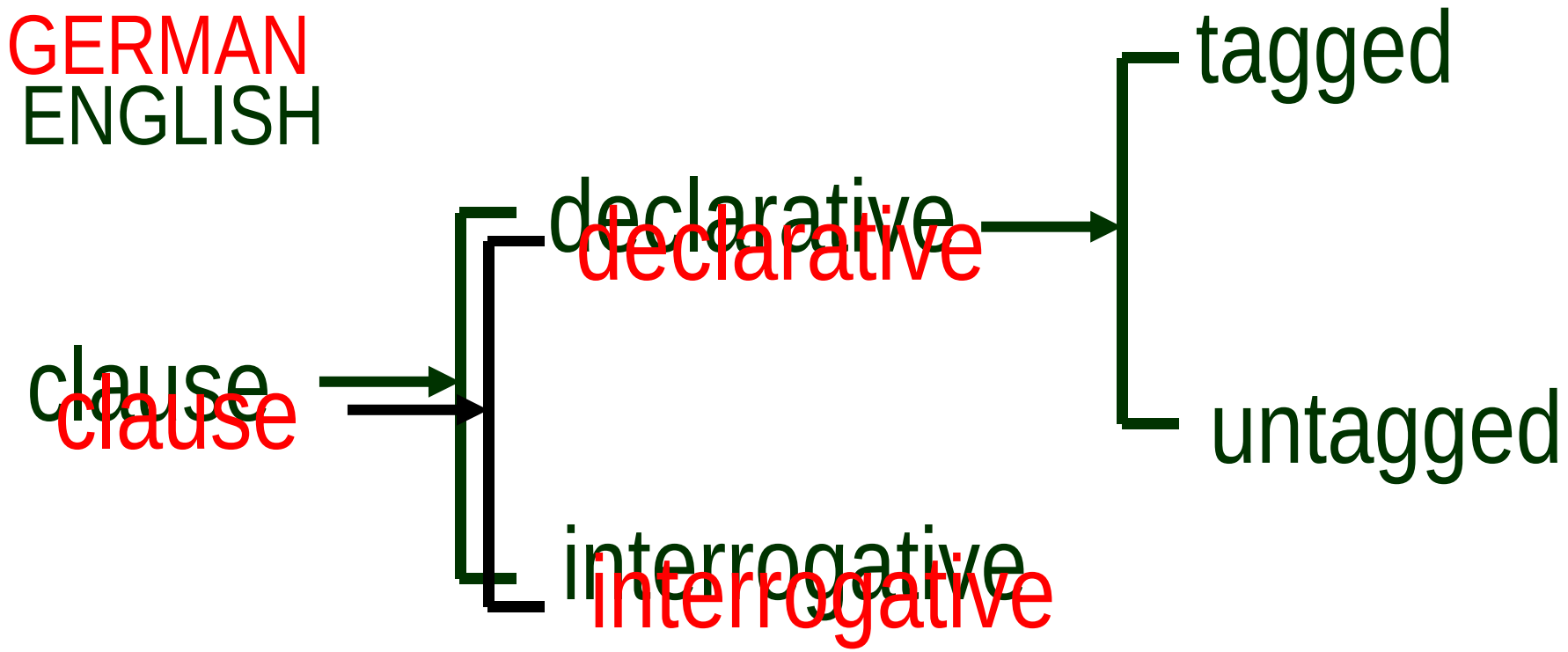


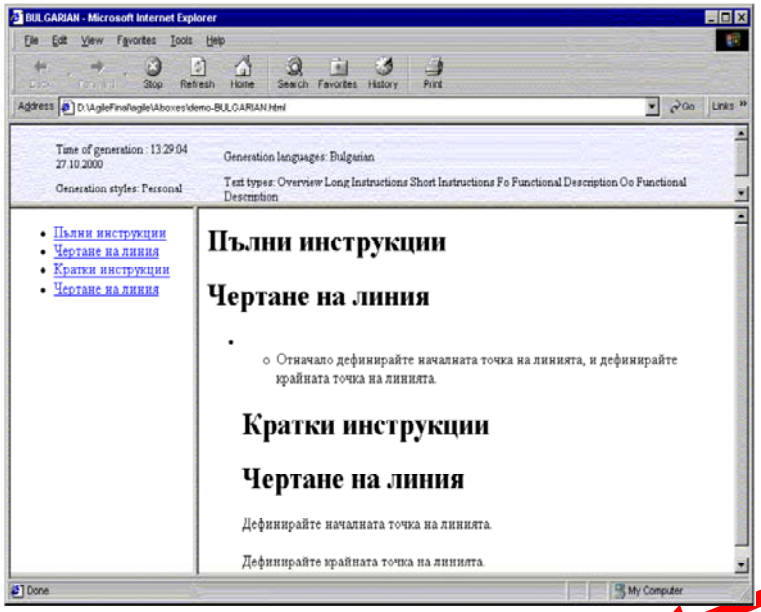
Partial congruence of paradigmatic specifications (potential)

GERMAN

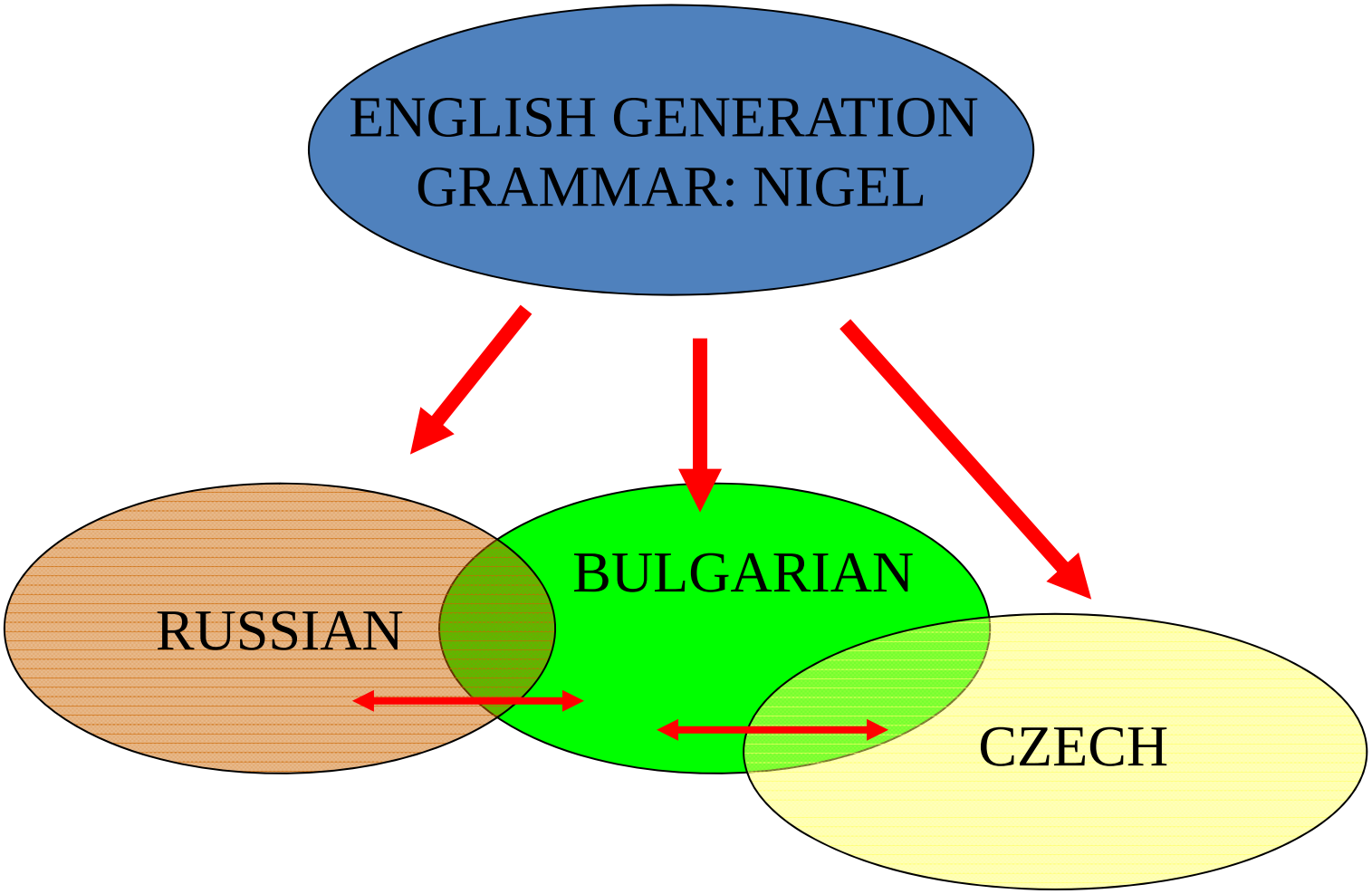


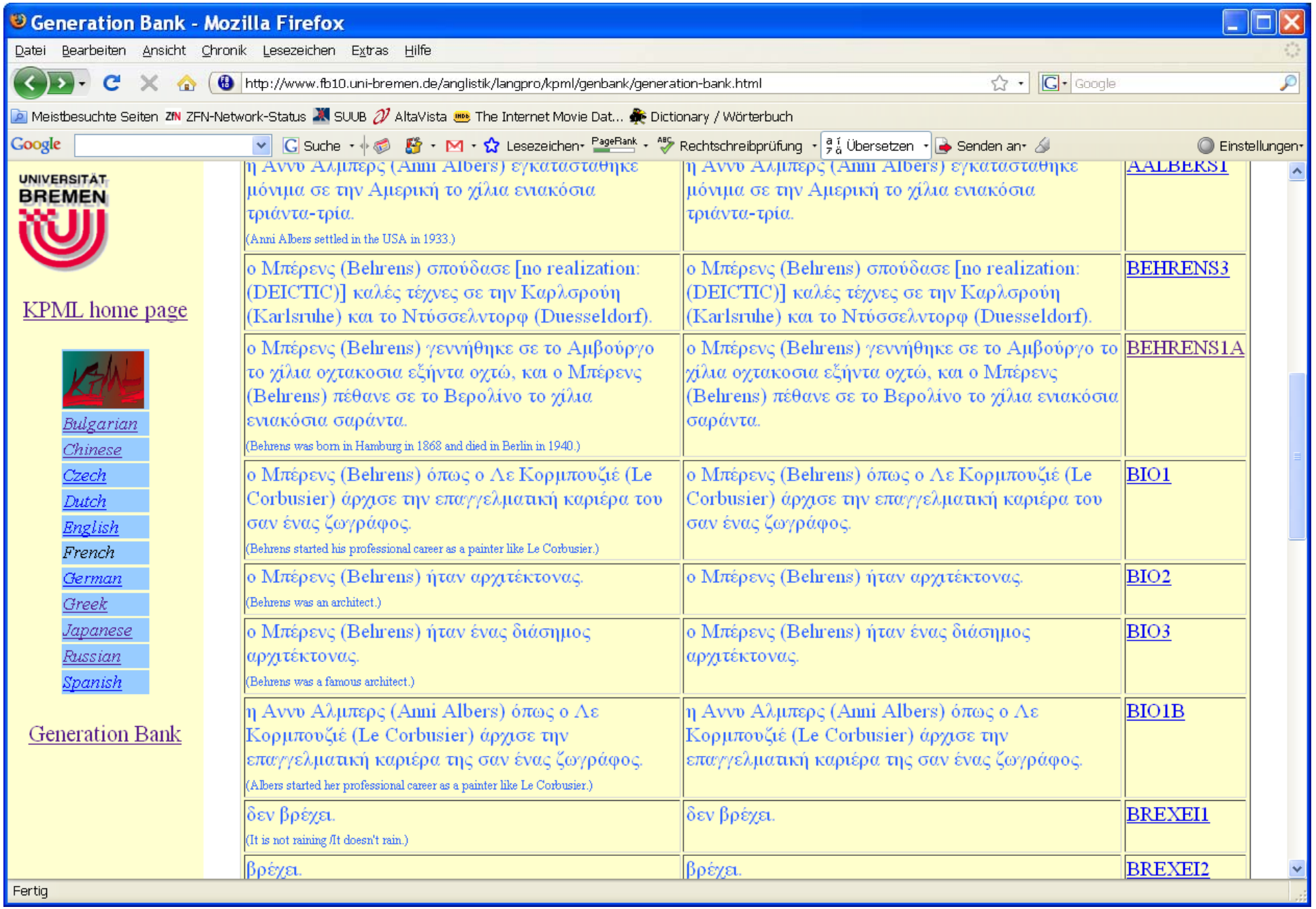
Partial congruence of paradigmatic specifications (potential)

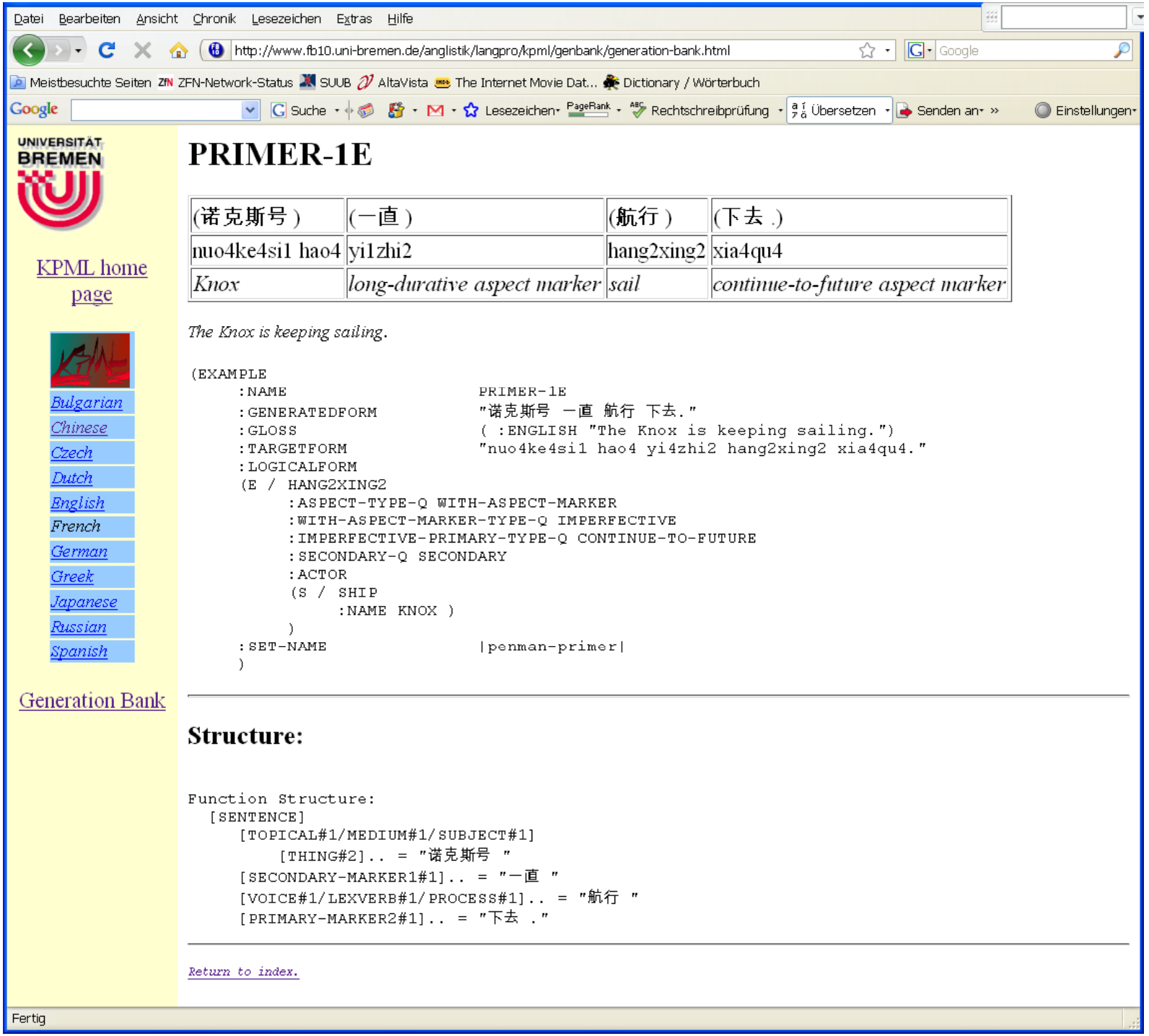




Agile Project: late 1990s


ENGLISH GENERATION GRAMMAR: NIGEL
RUSSIAN
BULGARIAN
CZECH


Generation Bank - Mozilla Firefox
Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
http://www.fb10.uni-bremen.de/anglistik/langpro/kpml/genbank/generation-bank.html
Meistbesuchte Seiten ZFN-Network-Status SUUB AltaVista The Internet Movie Dat... Dictionary / Wörterbuch
Suche Lesezeichen PageRank Rechtschreibprüfung Übersetzen Senden an Einstellungen
UNIVERSITÄT BREMEN
KPML home page
Bulgarian
Chinese
Czech
Dutch
English
French
German
Greek
Japanese
Russian
Spanish
Generation Bank
η Αννυ Αλμπερς (Anni Albers) εγκαταστάθηκε μόνιμα σε την Αμερική το χίλια ενιακόσια τριάντα-τρία.
(Anni Albers settled in the USA in 1933.)
η Αννυ Αλμπερς (Anni Albers) εγκαταστάθηκε μόνιμα σε την Αμερική το χίλια ενιακόσια τριάντα-τρία.
AALBERS1
ο Μπέρενς (Behrens) σπούδασε [no realization: (DEICTIC)] καλές τέχνες σε την Καρλσρούη (Karlsruhe) και το Ντύσσελντορφ (Duesseldorf).
ο Μπέρενς (Behrens) σπούδασε [no realization: (DEICTIC)] καλές τέχνες σε την Καρλσρούη (Karlsruhe) και το Ντύσσελντορφ (Duesseldorf).
BEHRENS3
ο Μπέρενς (Behrens) γεννήθηκε σε το Αμβούργο το χίλια οχτακοσια εξήντα οχτώ, και ο Μπέρενς (Behrens) πέθανε σε το Βερολίνο το χίλια ενιακόσια σαράντα.
(Behrens was born in Hamburg in 1868 and died in Berlin in 1940.)
ο Μπέρενς (Behrens) γεννήθηκε σε το Αμβούργο το χίλια οχτακοσια εξήντα οχτώ, και ο Μπέρενς (Behrens) πέθανε σε το Βερολίνο το χίλια ενιακόσια σαράντα.
BEHRENS1A
ο Μπέρενς (Behrens) όπως ο Λε Κορμπουζιέ (Le Corbusier) άρχισε την επαγγελματική καριέρα του σαν ένας ζωγράφος.
(Behrens started his professional career as a painter like Le Corbusier.)
ο Μπέρενς (Behrens) όπως ο Λε Κορμπουζιέ (Le Corbusier) άρχισε την επαγγελματική καριέρα του σαν ένας ζωγράφος.
BIO1
ο Μπέρενς (Behrens) ήταν αρχιτέκτονας.
(Behrens was an architect.)
ο Μπέρενς (Behrens) ήταν αρχιτέκτονας.
BIO2
ο Μπέρενς (Behrens) ήταν ένας διάσημος αρχιτέκτονας.
(Behrens was a famous architect.)
ο Μπέρενς (Behrens) ήταν ένας διάσημος αρχιτέκτονας.
BIO3
η Αννυ Αλμπερς (Anni Albers) όπως ο Λε Κορμπουζιέ (Le Corbusier) άρχισε την επαγγελματική καριέρα της σαν ένας ζωγράφος.
(Albers started her professional career as a painter like Le Corbusier.)
η Αννυ Αλμπερς (Anni Albers) όπως ο Λε Κορμπουζιέ (Le Corbusier) άρχισε την επαγγελματική καριέρα της σαν ένας ζωγράφος.
BIO1B
δεν βρέχει.
(It is not raining /It doesn't rain.)
δεν βρέχει.
BREXEI1
βρέχει.
βρέχει.
BREXEI2
Fertig


Datei Bearbeiten Ansicht Chronik Lesezeichen Extras Hilfe
http://www.fb10.uni-bremen.de/anglistik/langpro/kpml/genbank/generation-bank.html
Google
Meistbesuchte Seiten ZFN-Network-Status SUUB AltaVista The Internet Movie Dat... Dictionary / Wörterbuch
Google Suche Lesezeichen PageRank Rechtschreibprüfung Übersetzen Senden an Einstellungen
UNIVERSITÄT BREMEN
KPML home page
Bulgarian
Chinese
Czech
Dutch
English
French
German
Greek
Japanese
Russian
Spanish
Generation Bank
PRIMER-1E
(诺克斯号) (一直) (航行) (下去 .)
nuo4ke4si1 hao4 yi1zhi2 hang2xing2 xia4qu4
Knox long-durative aspect marker sail continue-to-future aspect marker
The Knox is keeping sailing.
(EXAMPLE
:NAME PRIMER-1E
:GENERATEDFORM "诺克斯号 一直 航行 下去."
:GLOSS (:ENGLISH "The Knox is keeping sailing.")
:TARGETFORM "nuo4ke4si1 hao4 yi4zhi2 hang2xing2 xia4qu4."
:LOGICALFORM
(E / HANG2XING2
:ASPECT-TYPE-Q WITH-ASPECT-MARKER
:WITH-ASPECT-MARKER-TYPE-Q IMPERFECTIVE
:IMPERFECTIVE-PRIMARY-TYPE-Q CONTINUE-TO-FUTURE
:SECONDARY-Q SECONDARY
:ACTOR
(S / SHIP
:NAME KNOX)
)
:SET-NAME |penman-primer|
)
Structure:
Function Structure:
[SENTENCE]
[TOPICAL#1/MEDIUM#1/SUBJECT#1]
[THING#2].. = "诺克斯号 "
[SECONDARY-MARKER1#1].. = "一直 "
[VOICE#1/LEXVERB#1/PROCESS#1].. = "航行 "
[PRIMARY-MARKER2#1].. = "下去 ."
Return to index.
Fertig

Research&Development Questions

- working out new grammars for new languages for generation in those languages
- extending grammars into areas not covered in sufficient detail
- using generation grammars for generating reports
- using generation grammars in dialogue systems

Short Documentation / Information on writing semantics specifications

http://www.fb10.uni-bremen.de/anglistik/langpro/kpml/Doc/tutorial/tutorial.htm