Approximating disambiguation of some German nominalizations by use of weak structural, lexical and corpus information

Kurt Eberle, Gertrud Faaß and Ulrich Heid
University of Stuttgart

(Sonderforschungsbereich/special research area 732)

Overview ...

- Nominalizations investigated
  - background of the study
- Disambiguation methods
  - frequency studies, semantic deduction, hybrid methods
- Approximation model
  - shallow text representations
  - deduction of disambiguation indicators
  - automatic indicator annotation
  - computation of disambiguation preferences
  - bootstrapping
- Results
  - reliability of annotation
  - reliability of disambiguation

Nominalizations ...

-ung nominalizations

- Derivation: verb → verb stem + affix *-ung*

  sperr- / to *block* -> *Sperrung (blocking/-age)*

  + particle: *ab-sperr* -> *Absperrung*

- Meaning: sortally ambiguous between ***event***, ***state, object***

  - *Absperrung*: event (**e** - *blocking*), state (**s** *blockage*), object (**o** *barrier*)
  - *Teilung* : event (**e** – *dividing*) or result state (**s** - *partition*)
  - *Messung*: event (**e** - *measuring*) or object (data),(**o** – *measurement*)

Hypotheses ...

sortal diversity depends on...

- verb semantics [Roßdeutscher et al. 2007,...]: e.g.
  * deadjectival, property-denoting: → **e,s**
    (*trocknen, lähmen* – ***dry, lame***)
  * deverbal, effected object: → **e,o**
    (*sammeln, bilden* – ***collect, create***)
  * denominal, affected object: → **e,s,o**
    (*pflastern* – ***pavement***)

Background ...

- General work about German nominalization

  * Osswald/Helbig (1990), Ehrich/Rapp (2000), ...

- Sonderforschungsbereich 732 of University of Stuttgart

  Incremental Specification in Context

  * Rossdeutscher (2007, 2010), Hamm/Kamp (2009),..

- SFB 732: subproject B3:

  Nominalizations in corpora, extraction of linguistic data and disambiguation

  * Spranger Heid (2007), Eberle et al. (2008, 2009),...

Hypotheses ...

must be tested...

- corpus data
- frequency studies
- occurrences of readings

→ determine readings of words in context

- how?

→ manually: time consuming, costly

→ automatically: if possible

Computation of disambiguations ...

- purely statistical methods

  N.B.: meaning differences are subtle

  → big, balanced corpora

  → big amount of manual annotation for learning classifiers

- purely semantic methods (logic deduction)

  → very costly, (nearly) untractable

➢ hybrid methods

consider frequency,
use shallow representations,
simple deductions (cf. 'light-weight semantics' [Marek (2009)])

Approximation model ...

- Detecting disambiguation criteria:
  - Fine-grained representation of typical occurrences of the different sortal readings

    Framework: *Discourse Representation Theory*

    (Kamp 1982, Kamp/Reyle 1993)
  - Deduction of disambiguation indicators from semantic patterns
- Automatic identification of indicators
  - Shallow representations of reference sentences in corpora

    ***F**lat **u**nderspecified **d**iscourse **r**epresentation **s**tructures*

    ('FUDRSs' , Eberle 1997, 2004; 'UDRT' Reyle 1993)
  - recognition of predefined indicator patterns

Approximation model ...

- Maximize reliability of automatic indicator annotation
  - Guidelines for human annotators
  - Comparison between human & automatic annotation
  - Adjustment of guidelines & computation algorithm
- Computation of disambiguation preferences
  - Basic weighting of indicators
  - Computation of disambiguation weights for test corpus
  - Human annotation of disambiguation preferences
  - Comparison & adjustment of weights (maximum entropy model)
- Bootstrapping
  - Learn from analyses, revise & adjust criteria

Approximation model ...

- Implementation
  - via cooperation University Stuttgart & Lingenio GmbH
  - adaptation: research prototype of Lingenio GmbH
    (http://lingenio.de/english/research/prototypes.htm)
    adaptation of Machine Translation product *translate*
    for the analysis of German

  (Eberle et al 2008, 2009)
- Evaluation
  - *Deutsches Web-as-Corpus (DeWaC)*, (Baroni/Kilgarriff 2006)
  - Example: nominalizations of verbs of saying...

Example ...

PPs with *nach* and nominalizations of verbs of saying

.... are ambiguous

(1) La Rosas eigenes Leben scheint ***nach den Erklärungen*** *nicht sicherer* DeWaC-Korpus

a) **After the statements** La Rosa's life doesn't seem to be safer

b) **According to the statements** La Rosa's life doesn't seem to be safer

nach + nominalizations of verbs of saying are ambiguous ...

- preposition ***nach*** is ambiguous:
  - **temporal relation: *after/après/después***
  - **discourse relation: *according to/selon; d'après/segons***
- nominalizations of verbs of saying (***announcement, declaration***,. . .) are ambiguous:
  - **the event of announcing s.th**
  - **the contents/message/proposition of the announcement**
- disambiguations are mutually dependent :
  - **after an event**
  - **according to a proposition**

Selectional restrictions ...

- Event reading:

  *kurz nach den Erklärungen / **shortly** after*

  *nach den **erfolgten** Erklärungen / the statements **made***

- Propositional reading

  *nach den **aktuellen** Erklärungen / **current** statements*

→ hard criteria

→ relatively few occurrences

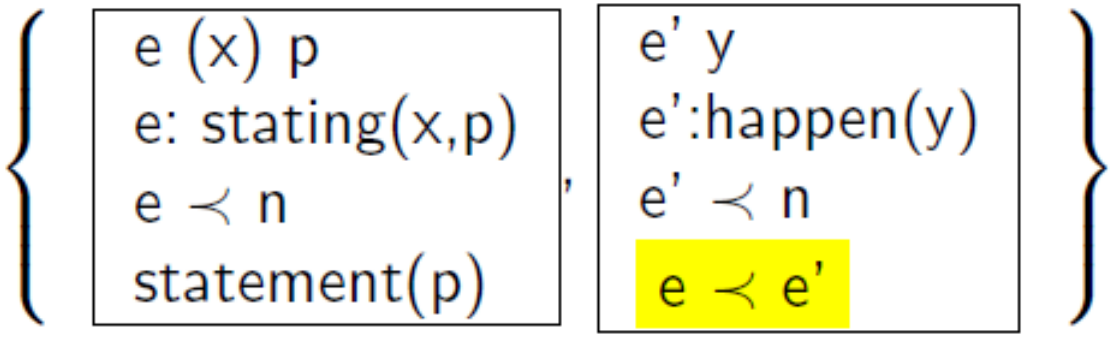
→ ‚weak' criteria: indicators which give hints !

→ deduce weak criteria from representational schemes !

Representations ...

- Event reading:

  after the statement, *s.th. happened/came into being*

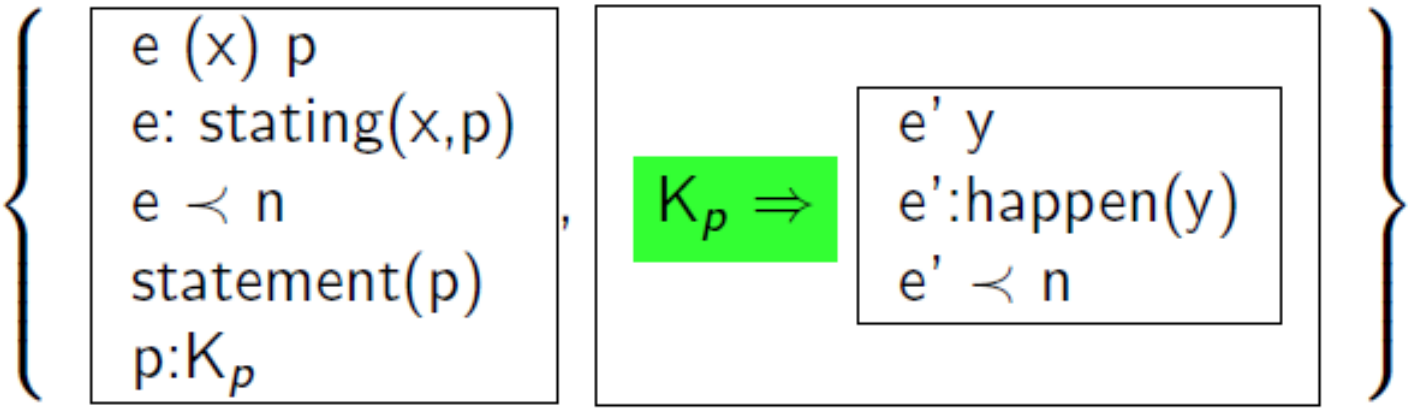


→ e = reference event of e', the event introduced by the main clause!

Representations ...

- Propositional reading

  according to the statement, *s.th. happened/came into being*



→ p = content of the event of stating, e, according to which the main clause statement ($K_{sentence}$)holds!

Consequences ...

Criteria for disambiguation:

→ use expectations connected to the different readings !

Example: **Event reading**

- *Event e of PP with* nach = *reference event of main clause*

→ not likely that there is a second temporal reference description for the main clause

(***reference time criterion***)

Criteria for disambiguation...

Example: **Event reading**

- *Event e of PP with nach = reference event of main clause*

→ e should characterize a relatively precise time, this means:

a single event (*e*) should be better suited than a (unspecified) sum of events (*E*)

- singular > numeral > quantified >bare plural

a known event should be better suited than an unknown

- definite vs. indefinite description

(***determination criterion***)

Theme ...

Example: Propositional reading

→ *Main clause reports the content of the statement of the 'according to' PP*

→ *not likely that PP reports (substantial parts of the) theme !*

*(2) Nach den Erklärungen **zu Fujimoris Vorgehen gegen die Opposition** scheint La Rosas Leben nicht sicherer.*

According to the statements about Fujimori's actions against the opposition La Rosa's life doesn't seem to be safer.

Criteria for disambiguation...

8 indicators

- Determination
- Reference time
- Reference location
- Aktionsart
- Tense
- Frame
- Agent
- Theme

Definition of a criterion ...

- Criterion must be checkable by a tool!
- and explicable to a non-expert annotator!

Definition of a criterion ...

Example: *Theme*

Guidelines: *Does the considered NP show:*

(a) a genitive phrase, (b) a prepositional phrase, (c) an adjective or (d) a subclause

that informs about the theme of the considered event of saying in a substantial way?

yes	→	Annotate: *theme criterion* *(temporal reading: 2, propositional reading: -2)*
no	→	Stop
unclear	→	Mark/extract corresponding phrase for later evaluation and stop (Mark&Stop)

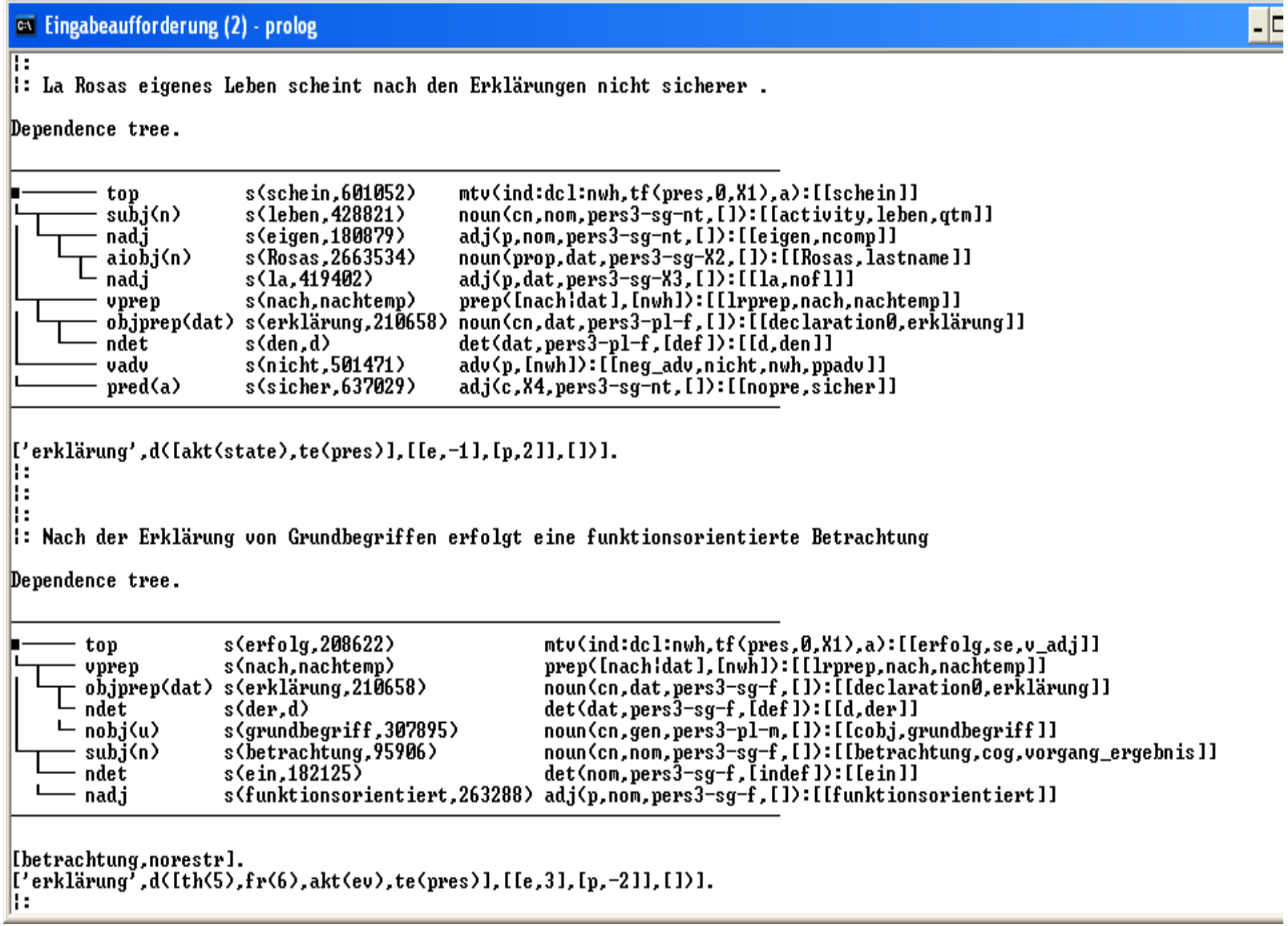
Definition of a criterion ...

Example: *Theme*

Implementation:

```
plausible_theme(WORD_ID,WORD_SEMTYPE) :-
        plausible_themes_w(WPTs), plausible_themes_t(TPTs),
        sem_subsumed(WORD_ID,WORD_SEMTYPE,WPTs,TPTs), !.

plausible_themes_t([liv\human,cog,doc,propos,situat]).
plausible_themes_w(PT_WORDLIST).
```


Eingabeaufforderung (2) - prolog
!:
!: La Rosas eigenes Leben scheint nach den Erklärungen nicht sicherer .
Dependence tree.
top s(schein,601052) mtv(ind:dcl:nwh,tf(pres,0,X1),a):[[schein]]
subj(n) s(leben,428821) noun(cn,nom,pers3-sg-nt,[]):[[activity,leben,qtm]]
nadj s(eigen,180879) adj(p,nom,pers3-sg-nt,[]):[[eigen,ncomp]]
aiobj(n) s(Rosas,2663534) noun(prop,dat,pers3-sg-X2,[]):[[Rosas,lastname]]
nadj s(la,419402) adj(p,dat,pers3-sg-X3,[]):[[la,nof1]]
vprep s(nach,nachtemp) prep([nach!dat],[nwh]):[[lrprep,nach,nachtemp]]
objprep(dat) s(erklärung,210658) noun(cn,dat,pers3-pl-f,[]):[[declaration0,erklärung]]
ndet s(den,d) det(dat,pers3-pl-f,[def]):[[d,den]]
vadv s(nicht,501471) adv(p,[nwh]):[[neg_adv,nicht,nwh,ppadv]]
pred(a) s(sicher,637029) adj(c,X4,pers3-sg-nt,[]):[[nopre,sicher]]
['erklärung',d([akt(state),te(pres)],[[e,-1],[p,2]],[])].
!:
!:
!:
!: Nach der Erklärung von Grundbegriffen erfolgt eine funktionsorientierte Betrachtung
Dependence tree.
top s(erfolg,208622) mtv(ind:dcl:nwh,tf(pres,0,X1),a):[[erfolg,se,v_adj]]
vprep s(nach,nachtemp) prep([nach!dat],[nwh]):[[lrprep,nach,nachtemp]]
objprep(dat) s(erklärung,210658) noun(cn,dat,pers3-sg-f,[]):[[declaration0,erklärung]]
ndet s(der,d) det(dat,pers3-sg-f,[def]):[[d,der]]
nobj(u) s(grundbegriff,307895) noun(cn,gen,pers3-pl-m,[]):[[cobj,grundbegriff]]
subj(n) s(betrachtung,95906) noun(cn,nom,pers3-sg-f,[]):[[betrachtung,cog,vorgang_ergebnis]]
ndet s(ein,182125) det(nom,pers3-sg-f,[indef]):[[ein]]
nadj s(funktionsorientiert,263288) adj(p,nom,pers3-sg-f,[]):[[funktionsorientiert]]
[betrachtung,norestr].
['erklärung',d([th(5),fr(6),akt(ev),te(pres)],[[e,3],[p,-2]],[])].
!:

Results ...

(1) Correctness of **annotation**

Test case: 100 sentences of DeWaC

informed tool (correct semantic typing)

8 weak criteria (algorithmic definitions & guidelines)

semantic selection constraints (hard criterion)

→ 9 decisions per sentence

→ 900 identical choices = 100 % agreement

Results ...

(1) Correctness of **annotation**

→ 900 identical choices = 100 % agreement

→ Human annotators (HA) : expert vs. no experience

→ Automatic annotation (AA)

Comparison	Agreement
∅ HA expert /HA non-expert	90,5 %
∅ HA non-expert /AA	87 %
∅ HA expert /AA	93,5%

Annotation ...

Typical problems:

(3) *Bei den am 11. November* ***in London*** *beginnenden Besprechungen .. werden,* ***nach Meldungen der Südena*** *, nur .. die Vollmachten..* ***zur Erörterung stehen****.*

According to announcements of Südena , only the authorisations .. **will be discussed** during the meetings beginning on Nov 11 **in London**

Sent	Ann	Agent	Theme	Det	Frame	Rspace	Rtime	Tense	Akt	Hard
5	HA1	+	-	no	-	+	+	F	nh	-
	HA2	+	-	no	+	-	+	F	h	-

Annotation ...

Typical problems:

(3) *Bei den am 11. November **in London** beginnenden Besprechungen .. werden, **nach Meldungen der Südena** , nur .. die Vollmachten.. **zur Erörterung stehen**.*

According to announcements of Südena , only the authorisations .. **will be discussed** during the meetings beginning on Nov 11 **in London**

HA: diverging structural and classifactory understanding

AA: wrong syntactic analysis

Results ...

(2) Correctness of disambiguation:

Test case: 100 sentences of DeWaC

informed tool (correct semantic typing)

8 weak criteria (algorithmic definitions & guidelines)

semantic selection constraints (hard criterion)

→ 9 decisions per sentence

Comparison	Agreement
HA expert / HA non-expert	94 %
AA with syntactic 'noise'/ ∅ HA	> 80 %
without (=core sentence analysis) / ∅ HA	> 90 % (∅ HA)

Results ...

(2) Correctness of disambiguation:

Test case: 100 sentences of DeWaC

with hard indicators (HI)	35 %
propositional reading (∅ HA)	64 %
event reading (∅ HA)	36 %
propositional reading without HI (∅ HA)	67 %
event reading without HI (∅ HA)	33 %

Comparison	Agreement
without HI + core sentence analysis / ∅ HA	> 85 % (∅ HA)

Conclusion, next steps ...

- approximation model for sortal disambiguation
- shallow semantic analysis
- mixture of statistical & semantic knowledge
- tractable: analysis tool (approx. < 20 sec per sentence)
- good precision

→ extension to complete DeWaC references (7864)
→ informed model ---> naive model
→ what is the tool's level of knowledge in the general case?
→ still tractable?
→ optimize weightings (maximum entropy model)
→ optimize criteria
→ extend to other phenomena

Thank you !