

Automatische Wortarten- und Syntaxannotation bei modifizierenden Wortarten des Deutschen

Hagen Hirschmann

Ines Rehbein

Forschungskolloquium
Korpuslinguistik

19. 11. 2014

Plan

1. Motivation & Ziele
2. Umklassifizierung von ADV & ADJD (STTS) und neue Kategorien
3. Evaluation des manuellen Taggings
4. Evaluation des automatischen Taggings
5. Evaluation des automatischen Parsings
6. Fazit

Motivation

- Viele theoretische Fragestellungen zu 1. Adverb(ial)- und 2. Partikelklassen:
Kategorisierung, Verarbeitung, Universalität, ...
- Aktuell viel untersuchte Klassen: Frameadverbien ([Störzer & Stolterfoht 2013](#)), Typen lokaler Adverbien ([Maienborn 2001](#); [Bücking 2013](#)), subjekt- vs. sprecher- vs. domainorientierte Adverbien & Syntax vs. Semantik von Adverbien ([Ernst 2002, 2007, 2008](#); Fokuspartikeln ([Sudhoff 2010](#); [Spalek & Gotzner & Wartenburger erscheint](#)), Modalpartikeln ([Degand & Cornillie & Pietrandrea 2013](#); [Döring 2013](#))

Motivation

- Fragestellungen zu einzelnen Klassen
 - Welches sind die syntaktischen und semantischen Besonderheiten?
 - Welche Bedeutungskomponenten (satzsemantischer Beitrag vs. Folgerungen, Implikaturen usw.)?
 - Was sagt uns die Grammatik der Elemente über die Grammatik im Allgemeinen?
 - Kognitive Realität (z. B. Fokusalternativen als kognitiv nachweisbare Größe)?
 - Erwerbsmuster in L1; Lernbarkeit in L2?

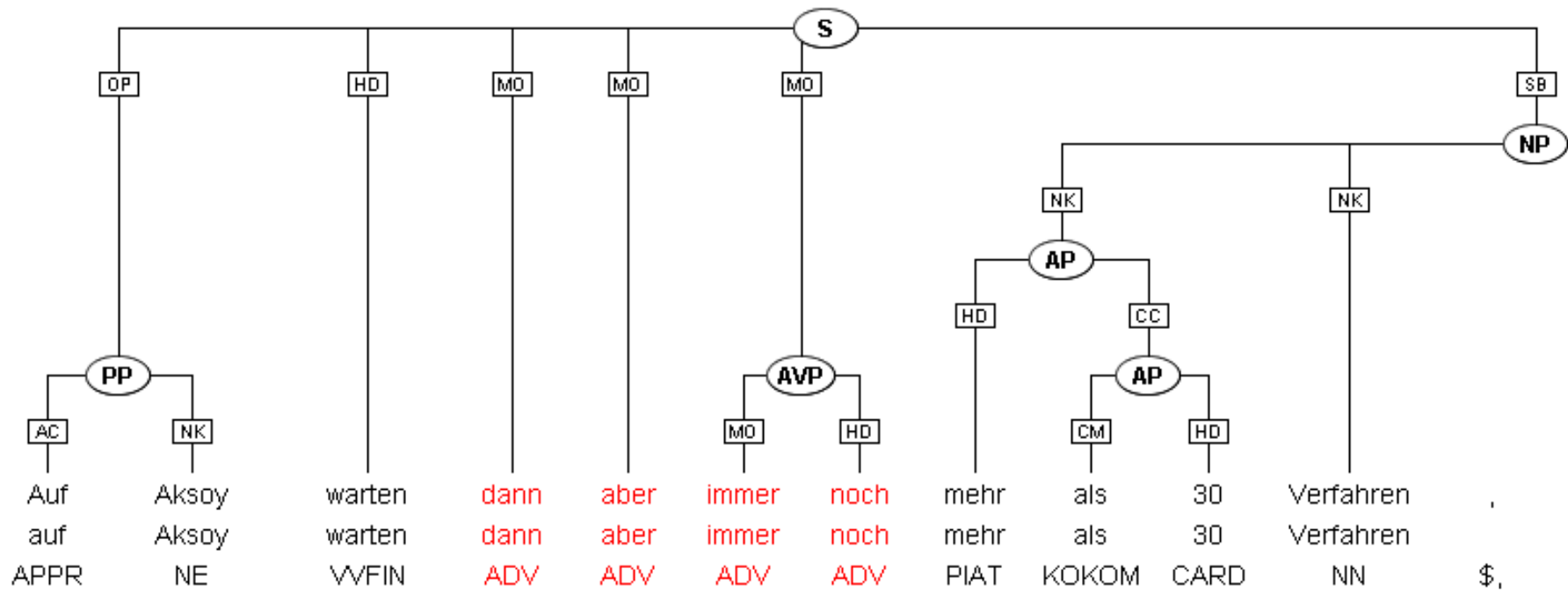
Motivation

- Ziel: Korpuslinguistische Unterstützung vorhandener Theorien und Hypothesen
 - Topologische Restriktionen von Adverb- und Partikelklassen?
 - Assoziation von Fokuspartikeln mit Alternativenmengen?
 - ...

Motivation

- Problem: Identifizieren syntaktischer und semantischer Klassen in Korpusdaten
 - Von der Klasse zum Token:
Finde alle Fokuspartikeln, Satzadverbien, ... ???
 - Wortarten-Tags:
PTK..., **ADV**, **ADJ.** (STTS & RFTagger)
 - Funktionen in Syntaxschemata:
MO (Tiger); **MOD**, **...MOD** (TüBa)
 - Vom Token zur Klasse:
 - *auch* als Fokuspartikel
 - *schon* als Modalpartikel
 - ...

Motivation



TIGER07, s17263

Motivation

- Ziele
 - Startpunkt: STTS-Tagset ([Schiller et al. 1999](#))
 - Anreicherung des Kategoriensystems um manuell und automatisch annotierbare Modifikatorklassen
 - von theoretischer Relevanz
 - syntaktisch motiviert
 - Abstimmung der Wortartenkonzepte mit syntaktischen Annotationsschemata (Tiger-Konstituentenstrukturen, Tiger-Dependenzstrukturen)
 - Erstellung/Erweiterung von Korpusressourcen mit diesem erweiterten System

Motivation

- Fragen heute
 - Welches sind die neuen Wortarten?
 - Wie gut lassen sich diese manuell taggen?
 - Wie gut lassen sich diese automatisch taggen?
 - Helfen diese syntaktsichen Klassen bei der automatischen Syntaxanalyse?
 - Konstituentenparsing
 - Dependenzparsing

'Neue' Klassen

- 2 Typen
 1. Konzeptionelle Änderung vorhandener Klassen ("Syntaktifizierung")
 2. Unterklassifizierung vorhandener Klassen
- Für jede Klasse Benennung von
 - Kriterien
 - Tests
 - Beispiele

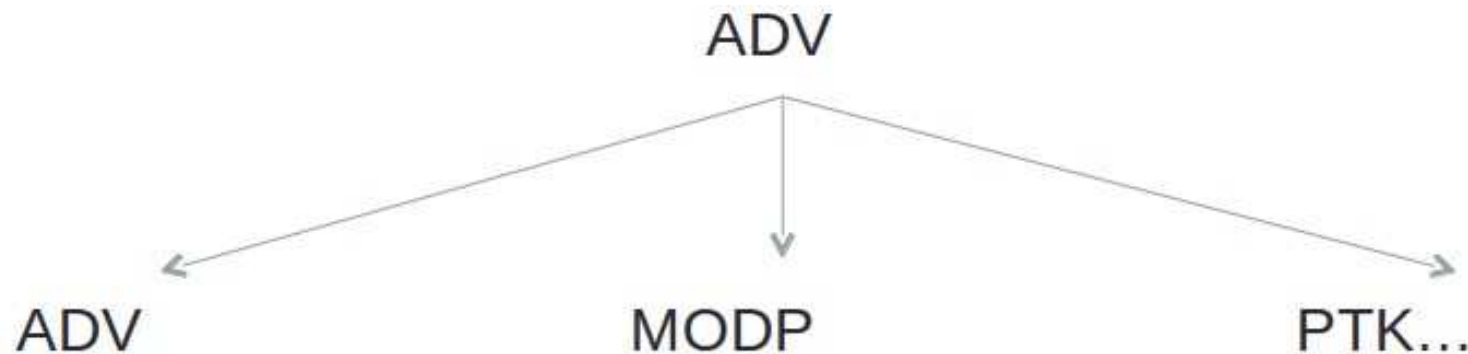
Umdefinierung von ADV ADJD

- ADV-ADJD-Unterscheidung im STTS
 - (...) **vielleicht/ADV** *wäre es ihm ähnlich ergangen (...)*
“**Perhaps** he would have experienced something similar”
(TIGER07, s9814)
 - (...) **wahrscheinlich/ADJD** *wird er nicht einmal gebilligt (...)*
“**Probably**, he will not even be approved”
(TIGER07, s17581)
[also concerns past participles (**gesteigert/VVPP** increased)]

Umdefinierung von ADV ADJD

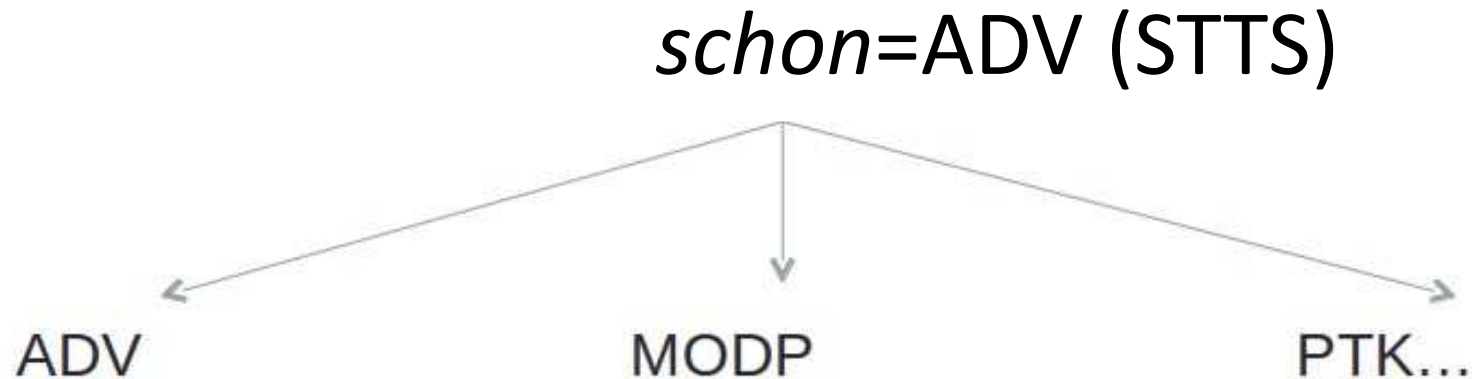
- Syntaktische Definition:
 - ADJD: Nomenmodifizierer
(Kriterium: Komplement von Kopulaverb)
 - ADV: Verb- oder Satzmodifizierer
(Kriterium: Alle übrigen Satzkonstituenten)

Unterklassen von ADV: MODP & PTK...



- Adverb
- Kriterium: Verb- oder Satzmodifizierer mit Satzgliedwert
- Alleinige Vorfeldposition
- Modalpartikel
- Kriterium: Satzmodifizierer mit Mittelfeldbeschränkung
- Test: no pre-field position
- Partikel
- Kriterium: Modifizierer in Satzkonstituente
- Test: Vorfeldposition innerhalb von Satzkonstituente

Unterklassen von ADV: MODP & PTK...



- *Bjerg fängt den ersten Text an (...), und **schon** lacht einer. (s2207)*
- *Übers Internet hat sie allerdings **schon** Stellung bezogen (s740)*
- *Da der Landesverband (...) nur einen " ganz kleinen Haushalt " hat (...), hätten Summen von (...) 165.000 Mark **schon** auffallen müssen. (s21)
Die Enttäuschung vom letzten Sonntag ist **schon** noch groß (...) (s723)*
- *Kulisammler aufgepaßt : Es gibt einen neuen Kuli auf dem Markt , der heute **schon** als Rarität zu bezeichnen ist . (s128)
Schon Thomas Mann befand , er sei ein " religiöser Humorist " . (s223)*

PTK...

- PTKFO: ***Nur Peter gewinnt***
 - Klasse: Fokuspartikel
 - Kriterium: Spezifikation eines Sets an Alternativen
 - Test: Alternativen benennen
- PTKINT: ***Sehr oft geschieht das***
 - Klasse: Intensivierer
 - Kriterium: Graduation oder Quantifizierung eines Kopfs
 - Test: Äquivalenten Gradausdruck nennen
- PTKLEX: ***Immer noch regnet es***
 - Klasse: Teil eines nicht-kompositionellen Mehrwortausdrucks
 - Kriterium: Lexikale Bedeutung passt nicht zur Bedeutung in Phrase
 - Test: Vergleich der verschiedenen Kontexte

Zusammenfassung: Gesamtinventar modifizierende Wortarten

Veränderungen am STTS:

- ADV
- ADJD
- MODP
- PTKFO
- PTKINT
- PTKLEX

Nicht verändert:

- ADJA, PTKNEG, PTKVZ, PTKA, PROAV/PAV

Inter-annotator Agreement

POS	# STTS	# new	# agr.	Fleiss' κ
VAPP	21	21	21	1.000
VVPP	173	172	172	0.989
ADJD	191	74	63	0.891
ADV	445	378	343	0.800
PTKFO	-	80	67	0.797
PTKINT	-	63	49	0.788
PTKLEX	-	33	17	0.594
MODP	-	12	6	0.515
gesamt	830	833	88.3%	0.838

Table: Verteilung (orig. STTS, modifiziertes Schema) und Übereinstimmung (prozentual und Fleiss' κ) für die verschiedenen Tags

POS-Tagging-Experimente – Setup

- Tagger ähnlich zu FLORS (Schnabel & Schütze 2014)
 - Linearer Klassifizierer (SVM)
(LIBLINEAR, Fan et al. 2008)
⇒ lerne einen separaten Tagger für jedes POS-Tag

POS-Tagging-Experimente – Setup

- Tagger ähnlich zu FLORS (Schnabel & Schütze 2014)
 - Linearer Klassifizierer (SVM)
(LIBLINEAR, Fan et al. 2008)
⇒ lerne einen separaten Tagger für jedes POS-Tag
 - Features:
 - Formfeatures
 - Präfix-/Suffixfeatures
 - Kontextfeatures
 - distributionelle Features (distributionelle Semantik)

POS-Tagging-Experimente – Setup

- Tagger ähnlich zu FLORS (Schnabel & Schütze 2014)
 - Linearer Klassifizierer (SVM)
(LIBLINEAR, Fan et al. 2008)
⇒ lerne einen separaten Tagger für jedes POS-Tag
 - Features:
 - Formfeatures
 - Präfix-/Suffixfeatures
 - Kontextfeatures
 - distributionelle Features (distributionelle Semantik)
 - Im Gegensatz zu Schnabel & Schütze (2014) verwenden wir zusätzlich POS-Kontextfeatures (automatisch getaggt mit Hunpos, einer Version des TNT-Taggers von Brants 2000)

POS-Tagging-Experimente

Ergebnisse (gesamt)

Tagger	Setting	Akk.
Baseline1 Hunpos	orig STTS	96.11
Baseline2 Hunpos	neue Tags	94.78
unser Tagger (ohne POS-Kontext)	neue Tags	94.91
unser Tagger (mit POS-Kontext)	neue Tags	96.68

Table: Akkuratheit für POS-Tagging (Hunpos Tagger und unser Tagger, ohne/mit POS-Kontextfeatures) auf dem Goldstandard (500 Sätze)

POS-Tagging-Experimente

Ergebnisse für die neuen Tags

TAG	Precision	Recall	F-score	Anzahl
ADV	87.25	89.76	88.49	(6276/6992)
PTKFO	82.32	84.73	83.51	(1243/1467)
ADJD	81.10	75.46	78.18	(824/1092)
PTKINT	80.81	72.87	76.63	(779/1069)
MODP	85.37	69.31	76.51	(70/101)
PTKLEX	81.87	67.62	74.07	(307/454)

Akkuratheit auf allen 58 Tags: 97.0%

Table: POS-Tagging-Akkuratheit auf dem größeren Datenset
(10.000 Sätze)

Parsingexperimente

Setup

- Zwei datengesteuerte, sprachunabhängige Parser

Parsingexperimente

Setup

- Zwei datengesteuerte, sprachunabhängige Parser



MATE Parser
(Bohnet 2010)

Parsingexperimente

Setup

- Zwei datengesteuerte, sprachunabhängige Parser



MATE Parser
(Bohnet 2010)



Berkeley Parser
(Petrov et al. 2006)

Parsingexperimente

Setup

- Zwei datengesteuerte, sprachunabhängige Parser



MATE Parser
(Bohnet 2010)



Berkeley Parser
(Petrov et al. 2006)

- Trainiert auf 10.000 Sätzen
 - mit original STTS-Tags
 - mit den neuen Tags
- Evaluation: 10-fache Kreuzvalidierung

Konstituentenparsing

Ergebnisse

		<i>original STTS</i>			Rec	<i>neue Tags</i>	
		Rec.	Prec	F-score		Prec	F-score
fold 1- 10	(\emptyset .)	75.38	75.12	75.25	75.45	75.64	75.55

Table: 10-fache Kreuzvalidierung, ohne grammatikal. Funktionen (GF)

Konstituentenparsing

Ergebnisse

		<i>original STTS</i>			Rec	<i>neue Tags</i>	
		Rec.	Prec	F-score		Prec	F-score
fold 1- 10	(\emptyset .)	75.38	75.12	75.25	75.45	75.64	75.55

Table: 10-fache Kreuzvalidierung, ohne grammatikal. Funktionen (GF)

- bescheidene, aber statistisch signifikante Verbesserung von 0.3% F-score (ohne GF)
- mit GF: 0.2% höherer F-score

Dependenzparsing

Ergebnisse

fold 1-10 (\emptyset)	<i>UAS orig.</i>	<i>neu</i>	<i>LAS orig.</i>	<i>neu</i>
gold POS	91.9	92.2	90.0	90.5
auto POS	89.7	89.8	86.9	87.1

Table: Parsingergebnisse (MATE Parser) für original STTS und für neue Tags

Dependenzparsing

Ergebnisse

fold 1-10 (\emptyset)	<i>UAS orig.</i>	<i>neu</i>	<i>LAS orig.</i>	<i>neu</i>
gold POS	91.9	92.2	90.0	90.5
auto POS	89.7	89.8	86.9	87.1

Table: Parsingergebnisse (MATE Parser) für original STTS und für neue Tags

- Verbesserung auf Goldtags: 0.3% (UAS) und 0.5% (LAS)
- Verbesserung auf auto. Tags: 0.1% (UAS) und 0.2% (LAS)

Dependenzparsing

Ergebnisse

fold 1-10 (\emptyset)	<i>UAS orig.</i>	<i>neu</i>	<i>LAS orig.</i>	<i>neu</i>
gold POS	91.9	92.2	90.0	90.5
auto POS	89.7	89.8	86.9	87.1

Table: Parsingergebnisse (MATE Parser) für original STTS und für neue Tags

- Verbesserung auf Goldtags: 0.3% (UAS) und 0.5% (LAS)
- Verbesserung auf auto. Tags: 0.1% (UAS) und 0.2% (LAS)
- Von 40 Dependenzrelationen:
 - Verbesserung für 23 Dependenzen + korrekter Kopf
 - gleiche Ergebnisse für 3 Dependenzen + korrekter Kopf
 - Verschlechterung für 14 Dependenzen + korrekter Kopf

Dependenzparsing

Verbesserung für einzelne Abhängigkeiten + korrekter Kopf (F-score)

DEPREL	freq.	orig f-score	neu f-score	+ (%)
NK	55439	97.4	97.5	0.1
AG	4929	92.9	93.0	0.1
CP	1759	90.8	92.9	2.1
OC	8133	86.9	88.1	1.2
SB	13065	86.4	87.1	0.7
OA	6678	79.7	80.2	0.5
AC	172	78.3	79.0	0.7
NG	1097	76.0	76.1	0.1
MO	22378	74.9	75.2	0.3
RC	1395	66.4	68.4	2.0
MNR	5227	67.0	67.5	0.5
DA	1045	57.1	59.6	2.5
APP	796	55.9	56.5	0.6
OP	1597	51.5	51.7	0.2
...	...			

Dependenzparsing

Verbesserung für einzelne Abhängigkeiten + korrekter Kopf (F-score)

DEPREL	freq.	orig f-score	neu f-score	+ (%)
NK	55439	97.4	97.5	0.1
AG	4929	92.9	93.0	0.1
CP	1759	90.8	92.9	2.1
OC	8133	86.9	88.1	1.2
SB	13065	86.4	87.1	0.7
OA	6678	79.7	80.2	0.5
AC	172	78.3	79.0	0.7
NG	1097	76.0	76.1	0.1
MO	22378	74.9	75.2	0.3
RC	1395	66.4	68.4	2.0
MNR	5227	67.0	67.5	0.5
DA	1045	57.1	59.6	2.5
APP	796	55.9	56.5	0.6
OP	1597	51.5	51.7	0.2
...	...			

Dependenzparsing

Verbesserung für einzelne Abhängigkeiten + korrekter Kopf (F-score)

DEPREL	freq.	orig f-score	neu f-score	+ (%)
NK	55439	97.4	97.5	0.1
AG	4929	92.9	93.0	0.1
CP	1759	90.8	92.9	2.1
OC	8133	86.9	88.1	1.2
SB	13065	86.4	87.1	0.7
OA	6678	79.7	80.2	0.5
AC	172	78.3	79.0	0.7
NG	1097	76.0	76.1	0.1
MO	22378	74.9	75.2	0.3
RC	1395	66.4	68.4	2.0
MNR	5227	67.0	67.5	0.5
DA	1045	57.1	59.6	2.5
APP	796	55.9	56.5	0.6
OP	1597	51.5	51.7	0.2
...	...			

Dependenzparsing

Verbesserung für einzelne Abhängigkeiten + korrekter Kopf (F-score)

DEPREL	freq.	orig f-score	neu f-score	+ (%)
DA	1045	57.1	59.6	2.5
CP	1759	90.8	92.9	2.1
RC	1395	66.4	68.4	2.0
OC	8133	86.9	88.1	1.2
SB	13065	86.4	87.1	0.7
AC	172	78.3	79.0	0.7
APP	796	55.9	56.5	0.6
OA	6678	79.7	80.2	0.5
MNR	5227	67.0	67.5	0.5
MO	22378	74.9	75.2	0.3
OP	1597	51.5	51.7	0.2
NK	55439	97.4	97.5	0.1
AG	4929	92.9	93.0	0.1
NG	1097	76.0	76.1	0.1
...	...			

Dependenzparsing

Verbesserung für einzelne Abhängigkeiten + korrekter Kopf (F-score)

DEPREL	freq.	orig f-score	neu f-score	+ (%)
DA	1045	57.1	59.6	2.5
CP	1759	90.8	92.9	2.1
RC	1395	66.4	68.4	2.0
OC	8133	86.9	88.1	1.2
SB	13065	86.4	87.1	0.7
AC	172	78.3	79.0	0.7
APP	796	55.9	56.5	0.6
OA	6678	79.7	80.2	0.5
MNR	5227	67.0	67.5	0.5
MO	22378	74.9	75.2	0.3
OP	1597	51.5	51.7	0.2
NK	55439	97.4	97.5	0.1
AG	4929	92.9	93.0	0.1
NG	1097	76.0	76.1	0.1
...	...			

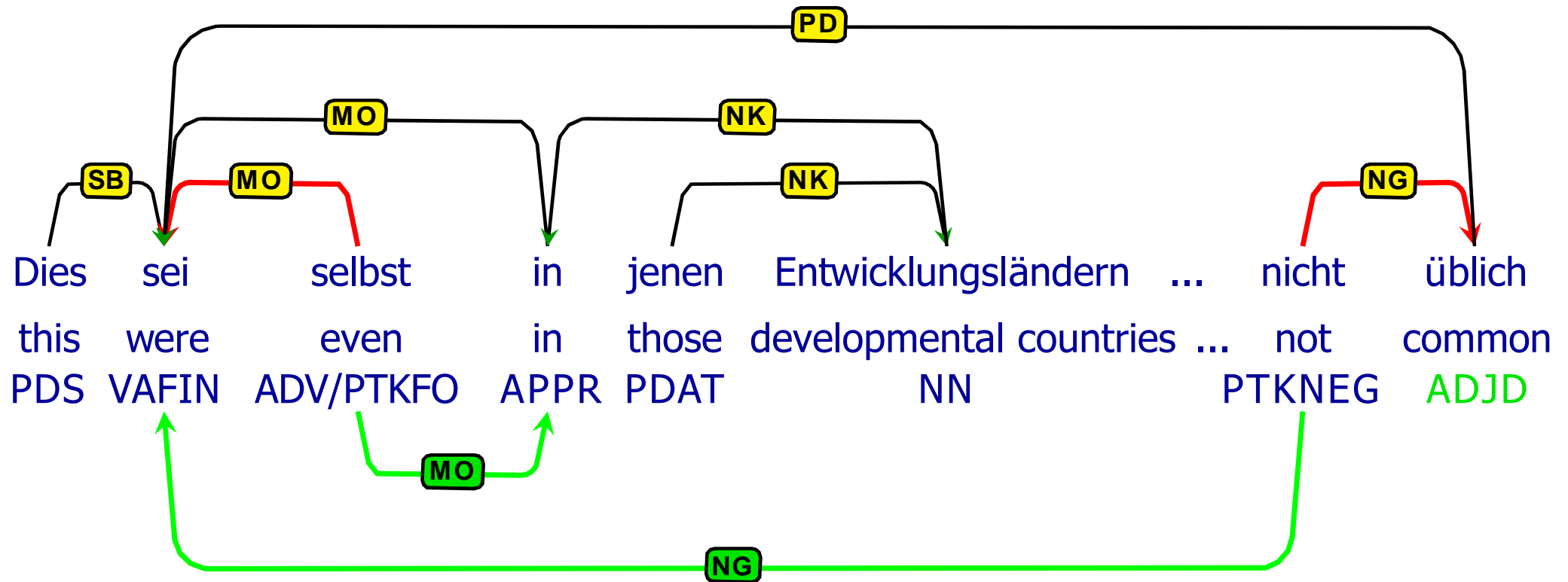
Dependenzparsing

Verbesserung für einzelne Abhängigkeiten + korrekter Kopf (F-score)

DEPREL	freq.	orig f-score	neu f-score	+ (%)
DA	1045	57.1	59.6	2.5
CP	1759	90.8	92.9	2.1
RC	1395	66.4	68.4	2.0
OC	8133	86.9	88.1	1.2
SB	13065	86.4	87.1	0.7
AC	172	78.3	79.0	0.7
APP	796	55.9	56.5	0.6
OA	6678	79.7	80.2	0.5
MNR	5227	67.0	67.5	0.5
MO	22378	74.9	75.2	0.3
OP	1597	51.5	51.7	0.2
NK	55439	97.4	97.5	0.1
AG	4929	92.9	93.0	0.1
NG	1097	76.0	76.1	0.1
...	...			

Dependenzparsing

Fehleranalyse -exemplarisch



Parser-Output für orig. (oben) and new (unten) Tags;
 rote Linien=falsche Anhängung; grüne Linien=korrekte Anhängung

Dependenzparsing

Evaluation der Parses bei den geänderten und neuen Tags

- Frage: Führen die 6 'neuen' syntaktischen Kategorien (ADV, ADJD, MODP, PTKFO, PTKINT, PTKLEX) zu einer besseren Anbindung der entsprechenden Token im Parse?
- Betrachtung der Gold-Tags in dem New-Datenset und der Entsprechungen in dem Old-Datenset
- Betrachtung der Anbindungen, nicht der Funktionen

Dependenzparsing

Evaluation der Parses bei den geänderten und neuen Tags

- Datenformat:

tok	word	New_gold_pos	New_gold_dep	Old_gold_pos	Old_gold_dep	TIGER_pos	TIGER_dep
9	davon	PROAV	14	PROAV	14	PROAV	14
10	sind	VAFIN	0	VAFIN	0	VAFIN	0
11	selbst	PTKFO	13	ADV	14	ADV	13
12	seine	PPOSAT	13	PPOSAT	13	PPOSAT	13
13	Kritiker	NN	10	NN	10	NN	10
14	überzeugt	VVPP	10	VVPP	10	VVPP	10

Dependenzparsing

Evaluation der Parses bei den geänderten und neuen Tags

- Datenformat:

tok	word	New_gold_pos	New_gold_dep	Old_gold_pos	Old_gold_dep	TIGER_pos	TIGER_dep
9	davon	PROAV	14	PROAV	14	PROAV	14
10	sind	VAFIN	0	VAFIN	0	VAFIN	0
11	selbst	PTKFO	13	ADV	14	ADV	13
12	seine	PPOSAT	13	PPOSAT	13	PPOSAT	13
13	Kritiker	NN	10	NN	10	NN	10
14	überzeugt	VVPP	10	VVPP	10	VVPP	10

- Wo unterscheiden sich New und Old bei New-Tag X?
- Hat der Parser bei New die korrekte Entscheidung gefällt?
- Hat der Parser bei Old die korrekte Entscheidung gefällt?
- Liegt der Parser bei beiden falsch?

Dependenzparsing

Evaluation der Parses bei den geänderten und neuen Tags

Kategorie	Vork. NEW	Disagree	NEW_korr	OLD_korr	both_inkorr
ADV	6992	1047	553	317	177
PTKFO	1467	280	196	44	41
ADJD	1092	88	37	27	24
PTKINT	1069	114	89	17	9
PTKLEX	454	80	33	24	22
MODP	101	33	13	11	9

Absolute (vergleichbare) Zahlen der Vorkommen und Nicht-Übereinstimmungen zw. NEW-OLD sowie NEW-GOLD und OLD-GOLD

Dependenzparsing

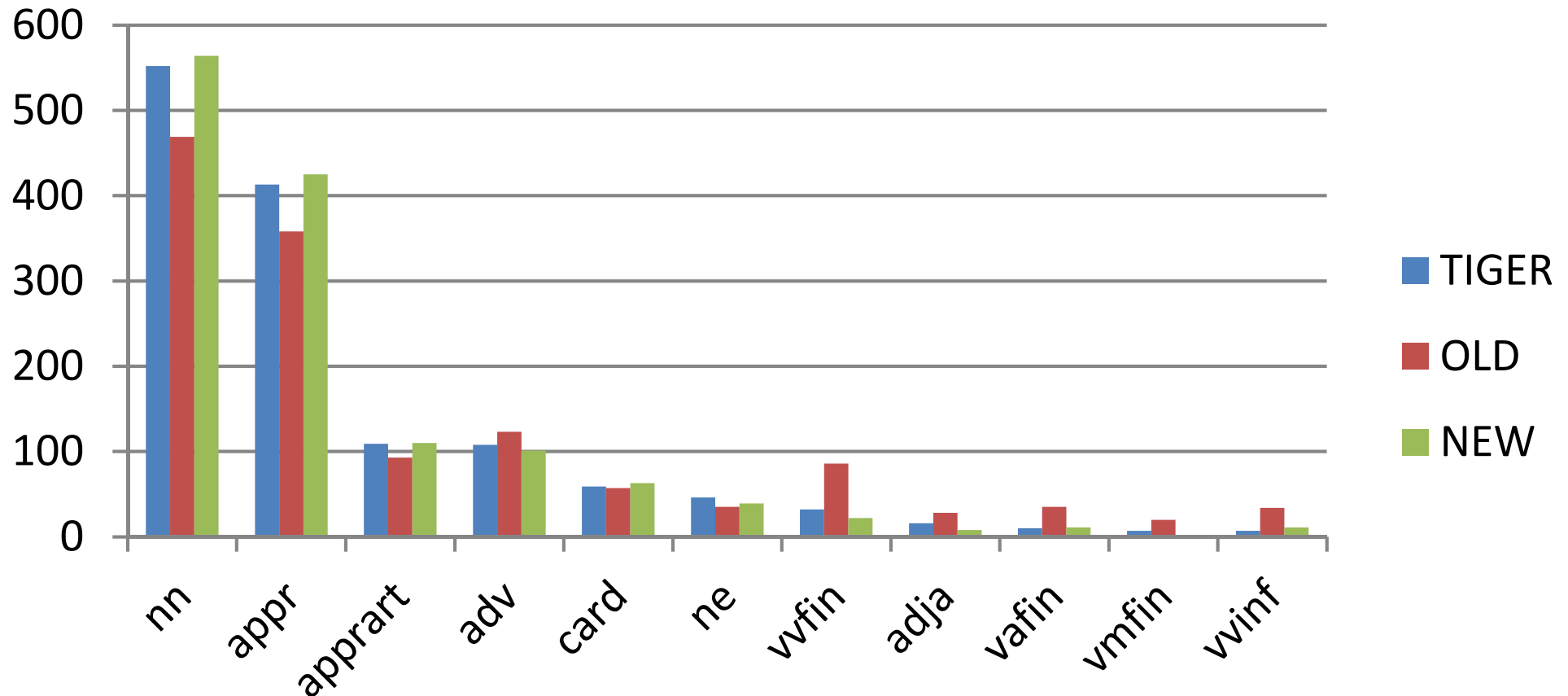
Evaluation der Parses bei den geänderten und neuen Tags

Kategorie	Disagree %	NEW_korr %	OLD_korr %	NEW/OLD
PTKINT	10,7	78,1	14,9	5,2
PTKFO	19,1	70,0	15,7	4,5
ADV	15,0	52,8	30,3	1,7
ADJD	8,1	42,0	30,7	1,4
PTKLEX	17,6	41,3	30,0	1,4
MODP	32,7	39,4	33,3	1,2

Anteil der Nicht-Übereinstimmungen am Gesamtvorkommen; Anteil der korrekten Anbindungen an den Nicht-Übereinstimmungen bei NEW und OLD in %; Verhältnis NEW/OLD

Dependenzparsing

Evaluation der Parses nach Dependenzzielen bei PTKFO



Ab mindestens 20 Vorkommen

Standardabweichung: OLD 221,3 NEW 51,6

Zusammenfassung

Ranking der einzelnen Tags

Manuelles Tagging (IAA):

- ADJD > ADV > PTKFO > PTKINT > PTKLEX > MODP

Automatisches Tagging (F-Scores):

- ADV > PTKFO > ADJD > PTKINT > MODP > PTKLEX

"Gewinn" bei automatischem Parsing (Gold-Tags):

- PTKINT > PTKFO > ADV > ADJD > PTKLEX > MODP

Zusammenfassung

POS-Tagset-Verfeinerung

- Erweiterung des STTS → linguistisch motivierte Klassen und detailliertere Analyse von Modifikation
- Neue Tagunterscheidungen können automatisch getaggt werden mit F-scores zwischen 74–88%

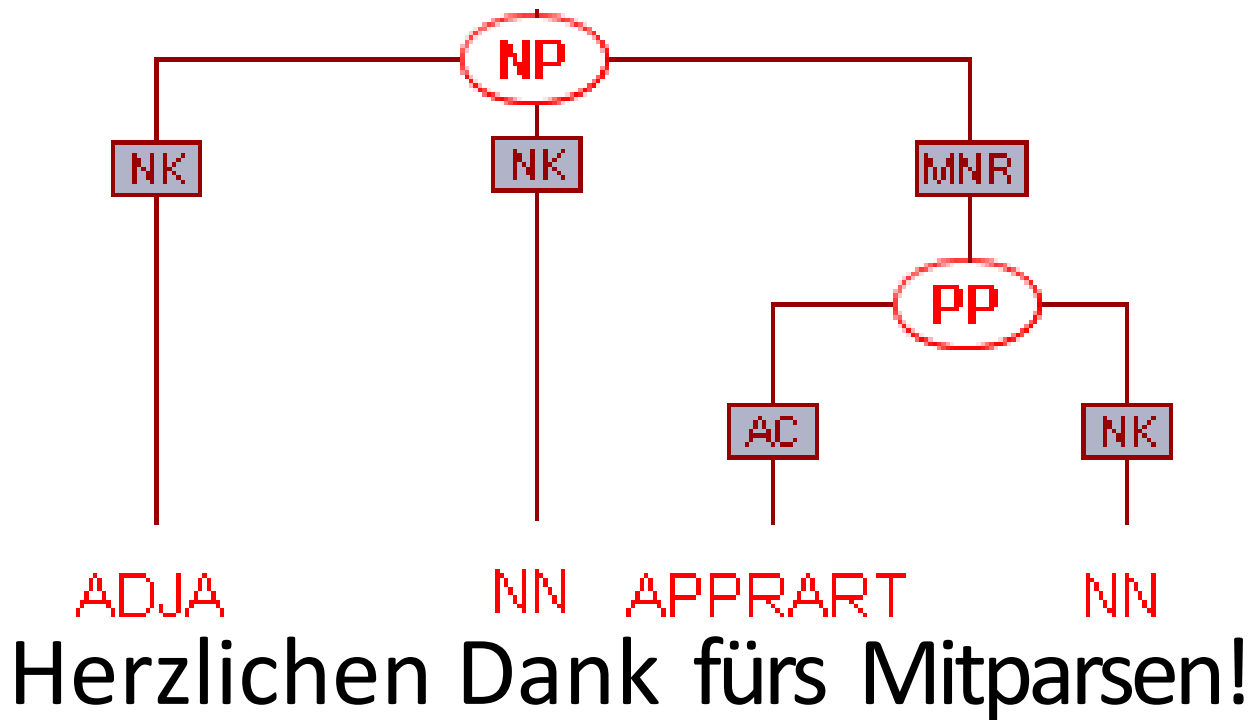
Einfluss auf statistisches Parsing

- Parsing mit neuen Tagunterscheidungen erzielt leichte Verbesserung über einen Parser, der auf dem Original-STTS trainiert wurde
(für Goldtags und für automatisch getaggtten Input)
- Verbesserte Ergebnisse für die wichtigsten grammatikalischen Funktionen/Argumente (SB, OA, DA, AG)

Zusammenfassung

ToDo

- Verbesserung des POS-Taggings durch linguistisch motivierte Features
- Veröffentlichung von Guidelines zu den neuen Tag-Kategorien
- Schaffung eines Goldstandards auf dem gesamten Tiger-Korpus oder einem gemischten Korpus (wie NoSta-D)
- Abstimmung der syntaktischen Wortklassen mit der Syntaxannotation im Tiger-Korpus bzw. gegebenen Korpus



Fragen?

hirschhx@hu-berlin.de

irehbein@uni-potsdam.de

Referenzen

- Brants, S., Dipper, S., Eisenberg, P., Hansen, S., König, E., Lezius, W., Rohrer, C., Smith, G., and Uszkoreit, H. (2004). TIGER: Linguistic Interpretation of a German Corpus. *Journal of Language and Computation*, 2004 (2), 597-620.
- Brill, E. (1992). A simple rule-based part of speech tagger. *3rd conference on Applied natural language processing (ANLC'92)*, Trento, Italy.
- Dickinson, M. and Meurers, D. W. (2003). Detecting Errors in Part-of-Speech Annotation. *10th Conference of the European Chapter of the Association for Computational Linguistics (EACL-03)*, Budapest, Hungary.
- Dickinson, M. (2006). An Investigation into Improving Part-of-Speech Tagging. *Proceedings of the Third Midwest Computational Linguistics Colloquium (MCLC-06)*, Urbana-Champaign, IL.
- Dligach, D. and Palmer, M. (2011). Reducing the Need for Double Annotation. *Proceedings of the 5th Linguistic Annotation Workshop (LAW V '11)*, Portland, Oregon.
- Eskin, E. (2000). Automatic Corpus Correction with Anomaly Detection. In *1st Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, Seattle, Washington.
- Květoň, P. and Oliva, K. (2002). (Semi-)Automatic Detection of Errors in PoS-Tagged Corpora, *19th International Conference on Computational Linguistics (COLING-02)*.
- Loftsson, H. (2009). Correcting a POS-Tagged Corpus Using Three Complementary Methods. *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, Athens, Greece.

Referenzen

(2)

- Manning, C. D. (2011). Part-of-speech Tagging from 97% to 100%: Is It Time for Some Linguistics?. *Proceedings of the 12th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing'11)*, Tokyo, Japan.
- Rehbein, I., Schalowski, S. and Wiese, H. (2014). The KiezDeutsch Korpus (KiDKo) Release 1.0. *The 9th International Conference on Language Resources and Evaluation (LREC-14)*, Reykjavik, Iceland.
- Rocio, V. Silva, J. and Lopes, G. (2007). Detection of Strange and Wrong Automatic Part-of-speech Tagging. *Proceedings of the Artificial Intelligence 13th Portuguese Conference on Progress in Artificial Intelligence (EPIA'07)*, Guimares, Portugal.
- Schmid, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. *International Conference on New Methods in Language Processing*, Manchester, UK.
- Schiller, A., Teufel, S., Stöckert, C. and Thielen, C. (1999). *Guidelines für das Tagging deutscher Textkorpora mit STTS*, Universität Stuttgart, Universität Tübingen.
<http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf>.
- Toutanova, K. and Manning, C. D. (2000). Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. *Proceedings of the conference on Empirical methods in natural language processing and very large corpora (EMNLP '00)*, Hong Kong.
- van Halteren, H. (2000). The Detection of Inconsistency in Manually Tagged Text. *Proceedings of the COLING-2000 Workshop on Linguistically Interpreted Corpora*, Centre Universitaire, Luxembourg.

Referenzen

(3)

- Bücking, S. (2012). Kompositional flexibel. Partizipanten und Modifikatoren in der Nominaldomäne. Tübingen; Stauffenburg.
- Degand, L.; Cornillie, B.; Pietrandrea, P. (2013). Discourse markers and modal particles: two sides of the same coin? L. Degand, B. Cornillie, P. Pietrandrea (Eds.) Discourse markers and modal particles: categorization and description. Amsterdam-Philadelphia; John Benjamins.
- Döring, S. (2013). Modal Particles and Context Shift. In D. Gutzmann, & H. - M. Gärtner (Eds.), Beyond expressives: Explorations in use-conditional meaning (pp. 95–123). Leiden: Brill.
- Ernst, T. 2002. The syntax of adjuncts. Cambridge: Cambridge University Press.
- Ernst, T. 2007. On the role of semantics in a theory of adverb syntax. *Lingua* 117.1008-1033.
- Ernst, T. 2008. Speaker-oriented adverbs. University of Massachusetts, Amherst, and Dartmouth College ms.
- Maienborn, C. (2001). On the Position and Interpretation of Locative Modifiers. *Natural Language Semantics* 9/2: 191-240.
- Spalek, K., Gotzner, N., & Wartenburger, I. (erscheint). Small words, big impact: How the focus particles “even” and “only” improve memory for sentence content. Atomium Culture.
- Störzer, M. & Stolterfoht, B. (2013). Syntactic Base Positions for Adjuncts? *Psycholinguistic Studies on Frame and Sentence Adverbials. Questions and Answers in Linguistics* 1(2), 57-72.
- Sudhoff, S. (2010). Focus Particles in German. *Syntax, Prosody, and Information Structure*. Amsterdam; John Benjamins.