

# STTS-Analysen von Lernertexten

Marc Reznicek und Heike Zinsmeister

3. STTS-Workshop

Tübingen, 6. Mai 2013

# Forschungsfragen

- Wie gut kann man mit Standard-POS-Taggern das STTS auf Lernaltersprache abbilden?
- Geben Unstimmigkeiten unterschiedlicher Tagger Hinweise auf Probleme in der Beschreibbarkeit von Lernaltersprache mithilfe des STTS?
- Lassen sich daraus Verbesserungsvorschläge für das STTS ziehen?

# Tagging von Lernersprache

Wortartenkategorien sind Eigenschaftsbündel von

- **syntaktischer** Verteilung (Distribution)
- **morphologischer** Markierung
- **lexikalischer** Information (Díaz Negrillo et al. 2010)

Lernersprache weicht systematisch von der Zielsprache ab

*Jeden Tag viele **Kriminal/NN** Aktivitäten **passiert/VVPP** in der **Heutzutager/NN** Gesellschaft.* (TreeTagger@FalkoEssayL2v2.3)

→ Hinweise auf den unterschiedlichen Ebenen können sich widersprechen!

# Normalisierung & Zielhypothese

## STTS -Guidelines

"Wenn der Sinn erkennbar ist, wird die Wortform verbessert, und es wird so getaggt, wie die richtige Wortform ausgesehen hätte." (Schiller et al. 1999:10)

**Minimale Zielhypothese (ZH1)** korrigiert lediglich Morphosyntax und Orthografie. (Lüdeling et al. 2005, Reznicek et al. to appear)

*Jeden Tag **passiert/VVFIN** viel **kriminelle/ADJA** Aktivität in der **heutigen/ADJA** Gesellschaft.*

## Tagging-Qualität auf Zielhypothese (ZH1):

- 98.8 % korrekte POS-Tags (RFTagger) (Rehbein et al. 2012)

# Minimale Zielhypothese (ZH1)

## Label für Abweichungen

- ZH1 = Normalisierter Paralleltext zum Lernertext
- Unterschiede im Text bzw. dessen Annotationen werden automatisch mit **Abweichungslabeln** markiert

### Label

INS
DEL
CHA
MERGE
SPLIT
MOVS
MOVT

# Minimale Zielhypothese (ZH1)

## Label für Abweichungen

Label
INS
DEL
CHA
MERGE
SPLIT
MOVS
MOVT

tok	pos	ZH1	ZH1pos	ZH1Diff	ZH1posDiff
Jeden	PIAT	Jeden	PIAT		
Tag	NN	Tag	NN		
viele	ADJA			MOVS	MOVS
Kriminal	NN			MOVS	MOVS
Aktivitäten	NN			MOVS	MOVS
passiert	VVPP	passiert	VVFIN		CHA
		viel	ADV	MOVT	MOVT
		kriminelle	ADJA	MOVT	MOVT
		Aktivität	NN	MOVT	MOVT
in	APPR	in	APPR		
der	ART	der	ART		
Heutzutage	NN	heutigen	ADJA	CHA	CHA
Gesellschaft	NN	Gesellschaft	NN		

# Forschungsfragen

- Wie gut kann man mit Standard-POS-Taggern das STTS auf Lernaltersprache abbilden?
- Geben Unstimmigkeiten unterschiedlicher Tagger Hinweise auf Probleme in der Beschreibbarkeit von Lernaltersprache mithilfe des STTS?
- Lassen sich daraus Verbesserungsvorschläge für das STTS ziehen?

# Forschungsfragen

- Wie gut kann man mit Standard-POS-Taggern das STTS auf Lernaltersprache abbilden?
  - Kann Ensemble-Tagging eingesetzt werden, um die manuelle Korrekturarbeit zu minimieren?
- Geben Unstimmigkeiten unterschiedlicher Tagger Hinweise auf Probleme in der Beschreibbarkeit von Lernaltersprache mithilfe des STTS?
- Lassen sich daraus Verbesserungsvorschläge für das STTS ziehen?



# Ensemble-Tagging

## Idee

- Parallele Annotation mit mehreren Taggern
- Einheitsentscheidung ist wahrscheinlich korrekt (Van Halteren et al. 2001)
- Unterschiede weisen auf problematische Instanzen hin

# Ensemble-Tagging

## Idee

- Parallele Annotation mit mehreren Taggern
- Einheitsentscheidung ist wahrscheinlich korrekt  
(Van Halteren et al. 2001)
- Unterschiede weisen auf problematische Instanzen hin

## Unser Tagger-Ensemble

- TreeTagger (TIGER) (Schmid 1995)
- RFTagger (TIGER) (Schmid & Laws 2008)
  - Hauptkategorien und Morphologie gemappt auf STTS
- Stanford Tagger (NEGRA + TT-Output) (Toutanova & Manning 2000)

# Daten fürs Tagging

## **Kobalt-Korpus** ([www.kobalt-daf.de](http://www.kobalt-daf.de))

- Aufsätze von fortgeschrittenen Lernern (OnDaF: ~B2)

## **Kleiner Goldstandard**

- 4x3 Texte L1 (BEL, CMN, SWE, DEU)
- Korrektur des Taggings der Zielhypothese
- 2 Annotatoren und Konsolidierung

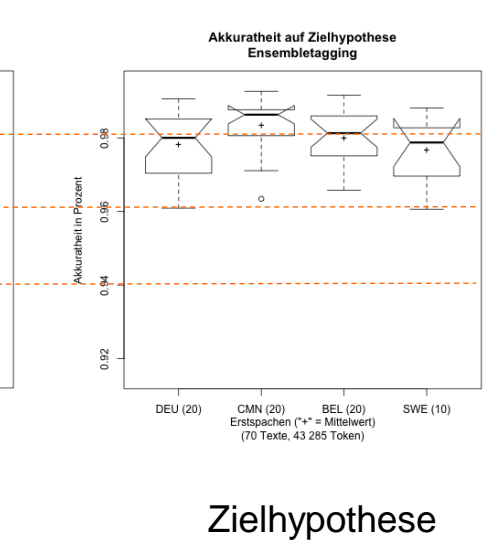
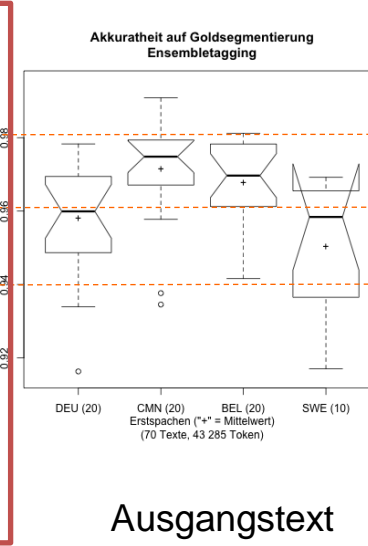
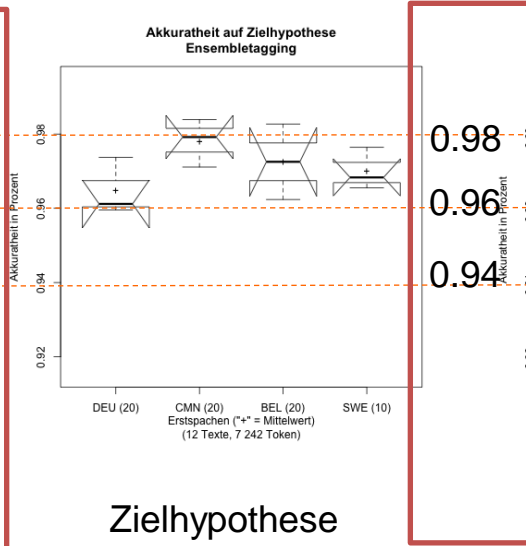
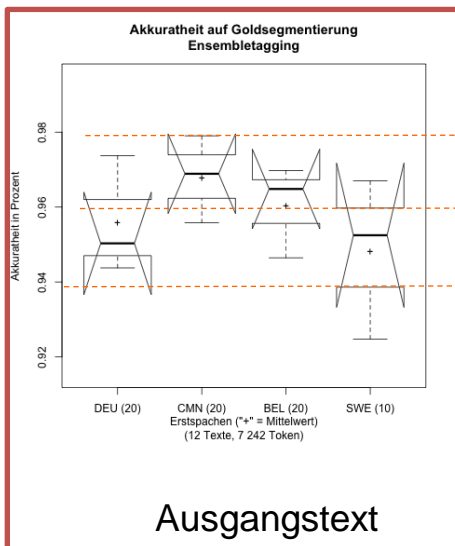
## **Großer Goldstandard**

- 70 Texte
- Ensemble-Tagging mit Mehrheitsvotum (Default: Treetagger)
- Manuelle Korrektur von Differenz beim Ensemble-Tagging

# Ergebnisse: Tagging-Akkuratheit

## Kleiner Goldstandard

## Großer Goldstandard



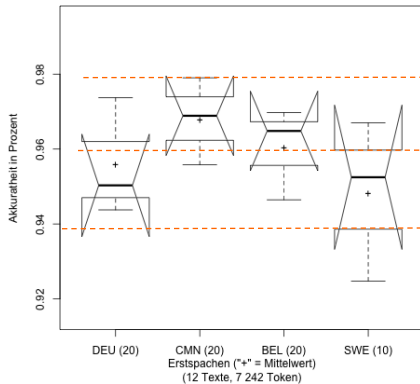
- Ausgangstext (4x3 Texte)
  - Unterschied zu den Gold-Tags
  - Akkuratheit um 96%

# Ergebnisse: Tagging-Akkurtheit

## Kleiner Goldstandard

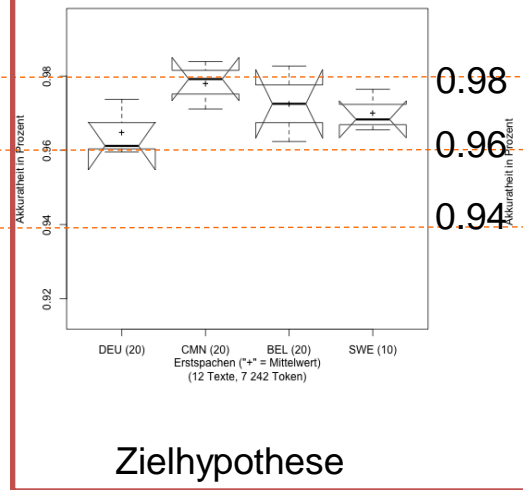
## Großer Goldstandard

Akkurtheit auf Goldsegmentierung  
Ensembletagging



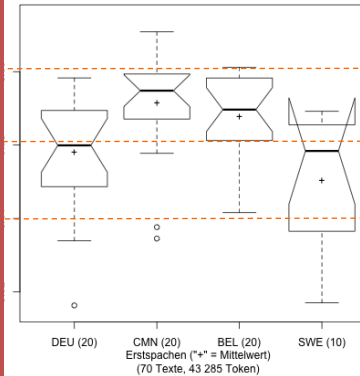
Ausgangstext

Akkurtheit auf Zielhypothese  
Ensembletagging



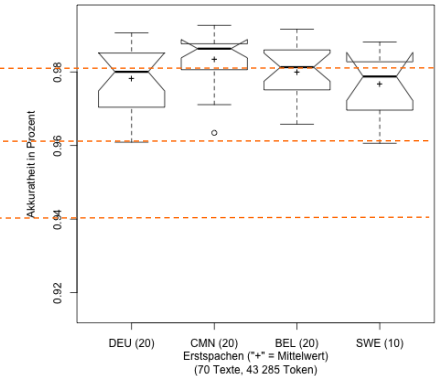
Zielhypothese

Akkurtheit auf Goldsegmentierung  
Ensembletagging



Ausgangstext

Akkurtheit auf Zielhypothese  
Ensembletagging



Zielhypothese

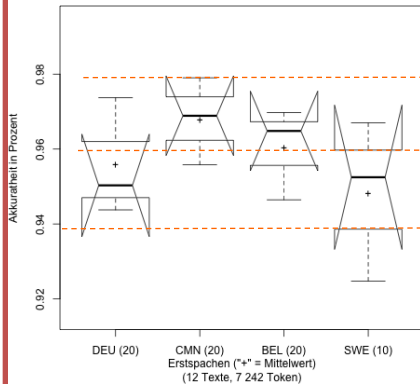
## ■ Zielhypothese

- Upper Bound der Taggerleistung auf normalisierten Texten (Textsorte)
- Ca. 2-4% falsche Tags

# Ergebnisse: Tagging-Akkuratheit

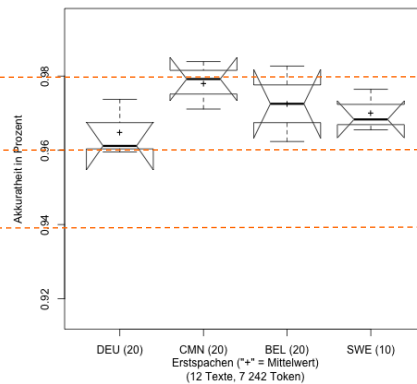
## Kleiner Goldstandard

Akkuratheit auf Goldsegmentierung  
Ensembletagging



Ausgangstext

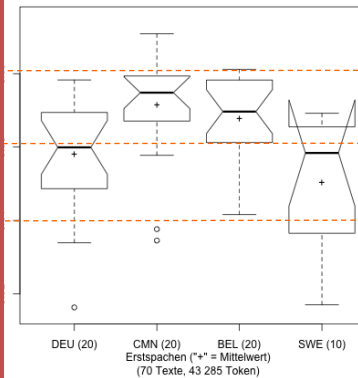
Akkuratheit auf Zielhypothese  
Ensembletagging



Zielhypothese

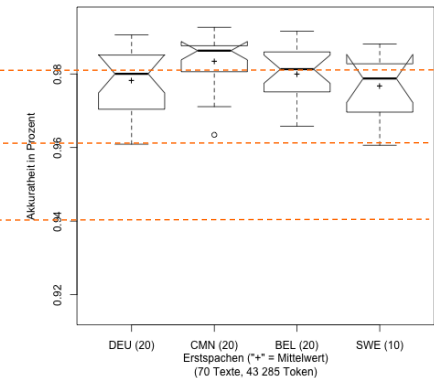
## Großer Goldstandard

Akkuratheit auf Goldsegmentierung  
Ensembletagging



Ausgangstext

Akkuratheit auf Zielhypothese  
Ensembletagging



Zielhypothese

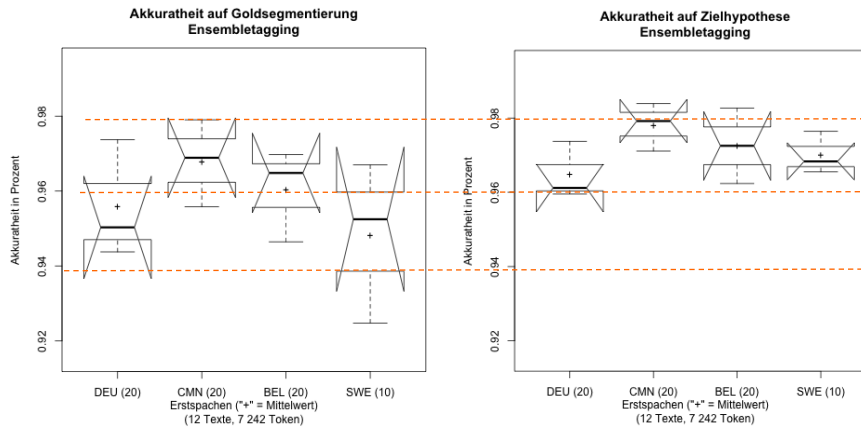
## Ensembletagging als Präprozessierung

- Manuelle Korrektur nur bei Nicht-Harmonie: 828 von 7 424 (11,43%) der Tokens
- “Noise”: 96 von 7 242 Taggingfehler (1,33%) werden “übersehen”

➤ Abwägung: Zeitinvestition versus “Noise”

# Ergebnisse: Tagging-Akkuratheit

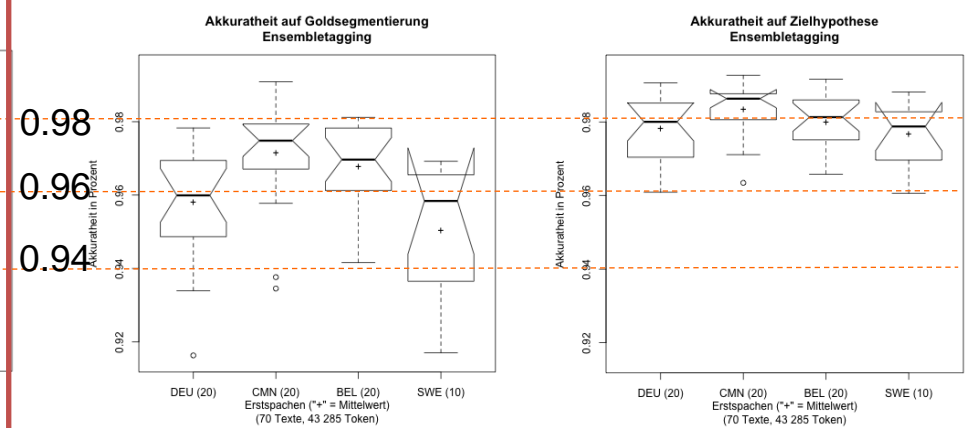
## Kleiner Goldstandard



Ausgangstext

Zielhypothese

## Großer Goldstandard



Ausgangstext

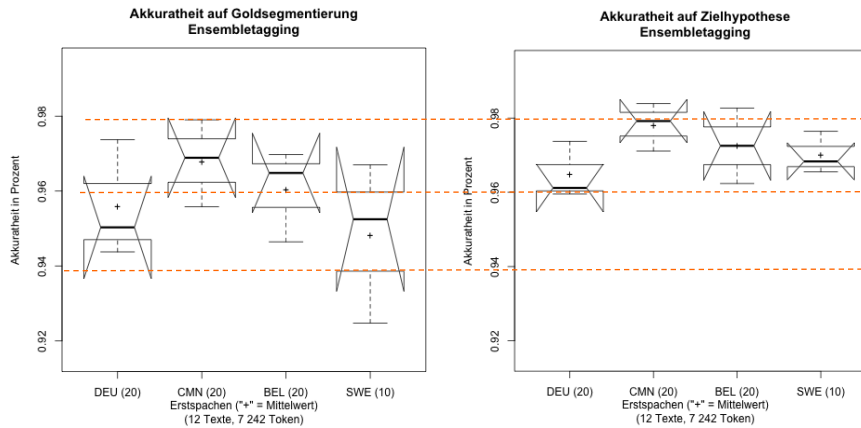
Zielhypothese

Gold = Korrektur nur bei Nicht-Harmonie

- Ergebnis ist zu optimistisch

# Ergebnisse: Tagging-Akkuratheit

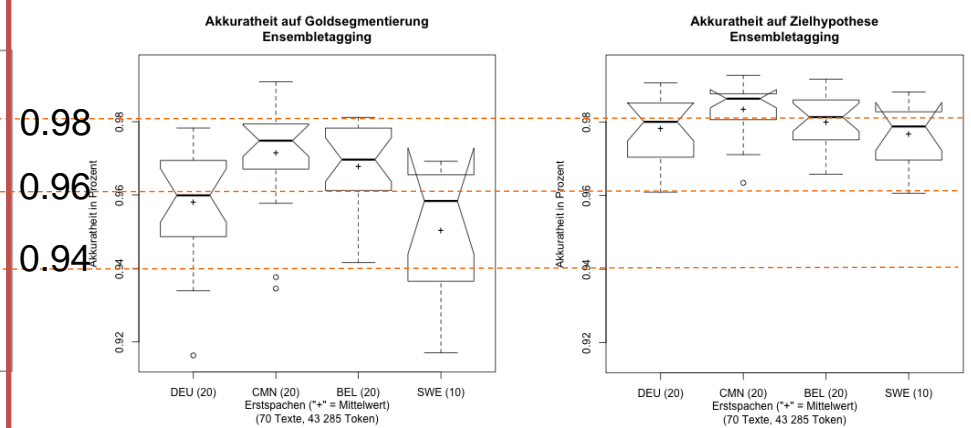
## Kleiner Goldstandard



Ausgangstext

Zielhypothese

## Großer Goldstandard



Ausgangstext

Zielhypothese

## Sprachabhängige Akkuratheit und Varianz

- Verhältnisse aus dem kleinen Goldstandard wiederholen sich (robustes Bild)
- L2-Lernertexte nicht schlechter als deutsche L1-Texte

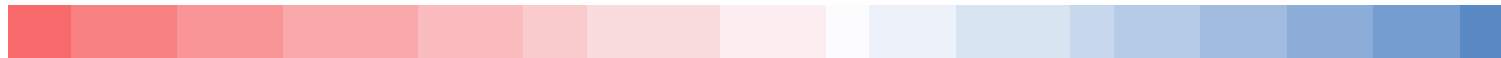


# Konfusionsdifferenz L1 vs. L2

		Gold																	
		ADJA	ADJD	ADV	APPR	KOKOM	KON	KOUS	NN	PAV	PIAT	PIS	PPER	PRELS	PRF	PTKA	VVFIN	VVINFL	VVPP
ctok	ADJA				-0,5		0,2	-0,2	-1		-0,2	-1					0,5		-0,2
	ADJD	0,9		0,4	-1,4		-1		0,6	-0		0,2					0,4	-0,1	-0,6
	ADV	-0,1	1		0,2		0,1		0,3	-3	1,2	1,1			0,3	0,1	-0,2	-0,4	0,1
	APPR	-0,2	0,2	0,6		-0,9	0,1	-0,7		-1						0,8	-0,2		0,1
	APPRART			0,3	1,6											-0,2			
	ART	0,1	-0,2		0,2	-0,2	0,1	-0,7				0,8	0,1	-2,7					
	KOKOM			0,1	0,1		0,3	-0,9			0,1								
	KON			0,1	-0,3	0,2		-0,1											
	NN	-0,4	-0,5	-0,2	-0,5			-0		0,2		-0					0,4		-0,3
	PDS			0,1				-1,1							-0,4				
	PIS	0,2		0,4					0,1		0,1		0,1	-0,3					
	PPER											0,3			0,4			-0,2	
	PRF			0,1									0,6						
	PTKNEG										0,5								
	PTKVZ				-0,6					-0							0,2	-0,4	
	PTKZU		-0,1		0,2											0,1		-0,2	
	PWS											0,1		0,5					
	VVFIN	-0,1		-0,9	0,2		-0	-0,1	-0		0,6		0,1					0,9	0,1
	VVINFL			0,2	-0,1				-1									3	0,8
	VVPP	-0,1	0,2	0,1					0,1								-0,5	1,1	

mehr Korrekturen bei L1

mehr Korrekturen bei L2

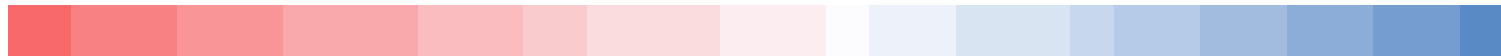


# Konfusionsdifferenz L1 vs. L2

		Gold																	
		ADJA	ADJD	ADV	APPR	KOKOM	KON	KOUS	NN	PAV	PIAT	PIS	PPER	PRELS	PRF	PTKA	VVFIN	VVINP	VVPP
ctok	ADJA				-0,5			0,2	-0,2	-1	-0,2	-1					0,5		-0,2
	ADJD	0,9		0,4	-1,4			-1		0,6	-0		0,2				0,4	-0,1	-0,6
	ADV	0,1	1		0,2			0,1		0,3	-3	1,2	1,1		0,3	0,1	-0,2	-0,4	0,1
	APPR	-0,2	0,2	0,6		-0,9	0,1	-0,7		-1						0,8	-0,2		0,1
	APPRART			0,3	1,6											-0,2			
	ART	0,1	-0,2		0,2	-0,2	0,1	-0,7				0,8	0,1	-2,7					
	KOKOM			0,1	0,1		0,3	-0,9			0,1								
	KON			0,1	-0,3	0,2		-0,1											
	NN	-0,4	-0,5	-0,2	-0,5			-0		0,2		-0					0,4		-0,3
	PDS			0,1				-1,1							-0,4				
	PIS	0,2		0,4					0,1		0,1		0,1	-0,3					
	PPER											0,3			0,4			-0,2	
	PRF			0,1									0,6						
	PTKNEG										0,5								
	PTKVZ				-0,6					-0							0,2	-0,4	
	PTKZU		-0,1	0,2												0,1		-0,2	
	PWS											0,1		0,5					
	VVFIN	-0,1		-0,9	0,2			-0	-0,1	-0		0,6		0,1				0,9	0,1
	VVINP		0,2	-0,1					-1								3		0,8
	VVPP	-0,1	0,2	0,1					0,1								-0,5	1,1	

höher bei L1

höher bei L2



# L2: ADJD<sub>L2</sub> → ADJA<sub>Gold</sub>

## Standard-Distribution & Nichtstandard-Morphologie

- *Andererseits, so argumentieren sie, hat die jüngere Generation zum Glück eine Welt mit relativ Frieden und Freiheit.* (Kobalt\_CMN\_008\_2011\_03)
- *Der Jugend heute hat ein klar Bild von der Welt mit seiner eigenen Meinung.* (Kobalt\_CMN\_021\_2011\_03)

## Standard-Morphologie & Nichtstandard-Distribution (Kontext)

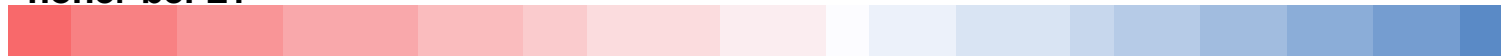
- *... weil das ein Ergebnis auch geographischer auch zeitlicher Bedingungen ist.* (Kobalt\_SWE\_006\_2011\_12)

# Konfusionsdifferenz L1 vs. L2

		Gold																	
		ADJA	ADJD	ADV	APPR	KOKOM	KON	KOUS	NN	PAV	PIAT	PIS	PPER	PRELS	PRF	PTKA	VVFIN	VVINF	VVPP
ctok	ADJA				-0,5		0,2	-0,2	-1		-0,2	-1					0,5		-0,2
	ADJD	0,9		0,4	-1,4		-1		0,6	-0		0,2					0,4	-0,1	-0,6
	ADV	-0,2	1		0,2		0,1		0,3	-3	1,2	1,1			0,3	0,1	-0,2	-0,4	0,1
	APPR	-0,2	0,2	0,6		-0,9	0,1	-0,7		-1						0,8	-0,2		0,1
	APPRART			0,3	1,6											-0,2			
	ART	0,1	-0,2		0,2	-0,2	0,1	-0,7				0,8	0,1	-2,7					
	KOKOM			0,1	0,1		0,3	-0,9			0,1								
	KON			0,1	-0,3	0,2		-0,1											
	NN	-0,4	-0,5	-0,2	-0,5			-0		0,2		-0					0,4		-0,3
	PDS			0,1				-1,1							-0,4				
	PIS	0,2		0,4					0,1		0,1		0,1	-0,3					
	PPER											0,3			0,4			-0,2	
	PRF			0,1									0,6						
	PTKNEG										0,5								
	PTKVZ				-0,6					-0							0,2	-0,4	
	PTKZU		-0,1		0,2											0,1		-0,2	
	PWS											0,1		0,5					
	VVFIN	-0,1		-0,9	0,2		-0	-0,1	-0		0,6		0,1					0,9	0,1
	VVINF			0,2	-0,1				-1								3		0,8
VVPP	-0,1		0,2	0,1				0,1								-0,5	1,1		

höher bei L1

höher bei L2



# L2: ADJD<sub>Gold</sub> → ADV<sub>ctok</sub>

## Standard-Distribution & Nichtstandard-Lexik

- *Wir machen es aber für uns einfach und nehmen Schweden, das kleine Land ganz viel (weit) oben auf dem Erdball. (Kobalt\_SWE\_007\_2011\_12)*

## Standard-Lexik & Nichtstandard-Distribution

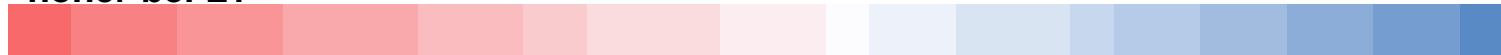
- *Das Verhältniss der Jügend sieht nicht gleich in alle Länder aus (Kobalt\_SWE\_010\_2012\_05)*

# Konfusionsdifferenz L1 vs. L2

		Gold																		
		ADJA	ADJD	ADV	APPR	KOKOM	KON	KOUS	NN	PAV	PIAT	PIS	PPER	PRELS	PRF	PTKA	VVFIN	VVINFL	VVPP	
ctok	ADJA				-0,5		0,2	-0,2	-1		-0,2	-1					0,5		-0,2	
	ADJD	0,9		0,4	-1,4		-1		0,6	-0		0,2					0,4	-0,1	-0,6	
	ADV	-0,1	1		0,2		0,1		0,3	-3	1,2	1,1			0,3	0,1	-0,2	-0,4	0,1	
	APPR	-0,2	0,2	0,6		-0,9	0,1	-0,7		-1						0,8	-0,2		0,1	
	APPRART			0,3	1,6											-0,2				
	ART	0,1	-0,2		0,2	-0,2	0,1	-0,7				0,8	0,1	-2,7						
	KOKOM			0,1	0,1		0,3	-0,9			0,1									
	KON			0,1	-0,3	0,2		-0,1												
	NN	-0,4	-0,5	-0,2	-0,5		-0			0,2		-0					0,4		-0,3	
	PDS			0,1				-1,1										-0,4		
	PIS	0,2		0,4					0,1		0,1		0,1	-0,3						
	PPER											0,3			0,4				-0,2	
	PRF			0,1									0,6							
	PTKNEG										0,5									
	PTKVZ				-0,6					-0								0,2	-0,4	
	PTKZU		-0,1		0,2											0,1			-0,2	
	PWS											0,1		0,5						
	VVFIN	-0,1		-0,9	0,2		-0	-0,1	-0		0,6		0,1						0,9	0,1
	VVINFL		0,2	-0,1					-1									3		0,8
	VVPP	-0,1	0,2	0,1					0,1									-0,5	1,1	

höher bei L1

höher bei L2



# L2: APPRART<sub>L2</sub> → APPR<sub>Gold</sub>

## Standard-Distribution & Nichtstandard-Morphologie

- *Die Kinder werden vom Geburt an von den Eltern und den Großeltern verwöhnt* (Kobalt\_CMN\_010\_2011\_03)

## Nichtstandard-Distribution & Nichtstandard-Morphologie

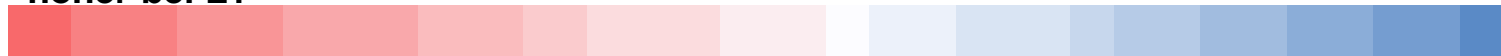
- *früher hatte man einfach nicht die Möglichkeit herumzusitzen, fern zu schauen, vom Computer stundenlang zu sitzen*. (Kobalt\_SWE\_010\_2012\_05)

# Konfusionsdifferenz L1 vs. L2

		Gold																	
		ADJA	ADJD	ADV	APPR	KOKOM	KON	KOUS	NN	PAV	PIAT	PIS	PPER	PRELS	PRF	PTKA	VVFIN	VVINFL	VVPP
ctok	ADJA				-0,5		0,2	-0,2	-1		-0,2	-1					0,5		-0,2
	ADJD	0,9		0,4	-1,4		-1		0,6	-0	0,2						0,4	-0,1	-0,6
	ADV	-0,1	1		0,2		0,1		0,3	3	1,2	1,1			0,3	0,1	-0,2	-0,4	0,1
	APPR	-0,2	0,2	0,6		-0,9	0,1	-0,7		-1						0,8	-0,2		0,1
	APPRART			0,3	1,6											-0,2			
	ART	0,1	-0,2		0,2	-0,2	0,1	-0,7				0,8	0,1	-2,7					
	KOKOM			0,1	0,1		0,3	-0,9			0,1								
	KON			0,1	-0,3	0,2		-0,1											
	NN	-0,4	-0,5	-0,2	-0,5		-0			0,2		-0					0,4		-0,3
	PDS			0,1				-1,1							-0,4				
	PIS	0,2		0,4					0,1		0,1		0,1	-0,3					
	PPER											0,3			0,4			-0,2	
	PRF			0,1									0,6						
	PTKNEG										0,5								
	PTKVZ				-0,6				-0								0,2	-0,4	
	PTKZU		-0,1		0,2											0,1		-0,2	
	PWS											0,1		0,5					
VVFIN	-0,1		-0,9	0,2		-0	-0,1	-0		0,6		0,1					0,9	0,1	
VVINFL		0,2	-0,1					-1									3		0,8
VVPP	-0,1	0,2	0,1					0,1									-0,5	1,1	

höher bei L1

höher bei L2





# L2: ADV/PRELS<sub>L2</sub> → PIAT<sub>Gold</sub>

## Standard-Distribution & Nichtstandard-Morphologie

- *Erstens bekommt man heute **mehre** Information als früher...* (Kobalt\_CMN\_06\_2011\_03 )

## Nichtstandard-Distribution & Nichtstandard-Lexik

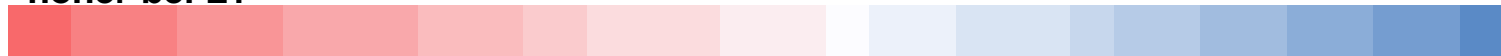
- in ehemaligen Zeiten gab es **was** solches nicht.  
(Kobalt\_SWE\_06\_2012\_05)

# Konfusionsdifferenz L1 vs. L2

		Gold																	
		ADJA	ADJD	ADV	APPR	KOKOM	KON	KOUS	NN	PAV	PIAT	PIS	PPER	PRELS	PRF	PTKA	VVFIN	VVINFL	VVPP
ctok	ADJA				-0,5		0,2	-0,2	-1		-0,2	-1					0,5		-0,2
	ADJD	0,9		0,4	-1,4		-1		0,6	-0		0,2					0,4	-0,1	-0,6
	ADV	-0,1	1		0,2		0,1		0,3	-3	1,2	1,1			0,3	0,1	-0,2	-0,4	0,1
	APPR	-0,2	0,2	0,6		-0,9	0,1	-0,7		-1						0,8	-0,2		0,1
	APPRART			0,3	1,6											-0,2			
	ART	0,1	-0,2		0,2	-0,2	0,1	-0,7				0,8	0,1	-2,7					
	KOKOM			0,1	0,1		0,3	-0,9			0,1								
	KON			0,1	-0,3	0,2		-0,1											
	NN	-0,4	-0,5	-0,2	-0,5			-0		0,2		-0					0,4		-0,3
	PDS			0,1				-1,1							-0,4				
	PIS	0,2		0,4					0,1		0,1		0,1	-0,3					
	PPER											0,3			0,4			-0,2	
	PRF			0,1									0,6						
	PTKNEG										0,5								
	PTKVZ				-0,6					-0							0,2	-0,4	
	PTKZU		-0,1		0,2											0,1		-0,2	
	PWS											0,1		0,5					
	VVFIN	-0,1		-0,9	0,2		-0	-0,1	-0		0,6		0,1					0,9	0,1
	VVINFL			0,2	-0,1				-1									3	0,8
VVPP	-0,1	0,2	0,1					0,1								0,5	1,1		

höher bei L1

höher bei L2



# L2: VVINFL<sub>L2</sub> → VVFIN<sub>Gold</sub>

## Standard-Distribution & Nichtstandard-Morphologie

- *...sind neue Welte, auf denen der Jugend heute **leben**.*  
(Kobalt\_CMN\_06\_2011\_03 )
- *Die chinesische Regierung hat durch ihre Reform das Ziel erreicht , dass das Volk nicht mehr unter Armut leidet , sondern ein gutes Leben **führen**.* (Kobalt\_CMN\_002\_2011\_03 )

ABER:

Verbalbereich ist generell ein problematischer

# Zwischenfazit (2)

## Ausgangsfrage

- Geben Unstimmigkeiten unterschiedlicher Tagger Hinweise auf Probleme in der Beschreibbarkeit von Lernaltersprache mithilfe des STTS?

## Ergebnisse

- Unterschiedliche Problemklassen
- Anhand von Konfusionmatrix differenzierbar
- Filter: Differenz L1 zu L2 Konfusionen zeigt lernalterspezifische Probleme

# Forschungsfragen

- Wie gut kann man mit Standard-POS-Taggern das STTS auf Lernaltersprache abbilden?
- Geben Unstimmigkeiten unterschiedlicher Tagger Hinweise auf Probleme in der Beschreibbarkeit von Lernaltersprache mithilfe des STTS?
- Lassen sich daraus Verbesserungsvorschläge für das STTS ziehen?

# Mögliche Konsequenzen

## STTS-Tags auf mehreren Ebenen sind ausreichend

- Mehrebenensuche
  - Lerner-POS || Diff-Tag || Zielhypothese\_POS
  - ADJD || change || ADJA
  
- Auf eine Ebene abbilden: Portmonteau-Tags
  - Lerner-POS\_Zielhypothese\_POS
  - ADJD\_ADJA

# Zusammenfassung

- Tagger sind einsetzbar auf Lernersprache
  - Ensembletagging als Vorverarbeitung für manuelle Korrektur hilfreich
- Ensemble-Mismatch gibt Hinweise auf interessante Abweichungstypen
- STTS-'Erweiterung': Portmonteau-Tags

# Ausblick (Artikel)

- Feinere Analyse der Taggerfähigkeiten
  - im Lexikon / nicht im Lexikon
  - Einfluss von falscher Distribution
  - Einfluss Satzlängen
- Feinere Analyse der Verwechslungskategorien
  - Fehler trotz Tagger-Harmonie
- Training
  - Klassifikator auf Portmonteau-Tags



# Referenzen

- Díaz+Negrillo, Ana; Meurers, Walt Detmar; Valera, Salvador; Wunsch, Holger (2010):** Towards Interlanguage POS Annotation for Effective Learner Corpora in SLA and FLT. In: *Language Forum*.
- Giesbrecht, Eugenie & Evert, Stefan (2009):** Part+of+speech Tagging + a Solved Task? An Evaluation of POS Taggers for the Web as Corpus. Alegria, I.; Leturia, I. & Sharoff, S. (ed.) Proceedings of the 5th Web as Corpus Workshop (WAC5)
- Lüdeling, Anke; Walter, Maik; Kroymann, Emil; Adolphs, Peter (2005):** Multi+level Error Annotation in Learner Corpora. In: Proceedings of Corpus Linguistics 2005. Birmingham.
- Rehbein, Ines; Hirschmann, Hagen; Lüdeling, Anke; Reznicek, Marc (2012):** Better Tags Give Better Trees or do they? In: *LiLT* 7 (10).
- Reznicek, Marc; Lüdeling, Anke; Hirschmann, Hagen (to appear):** Competing Target Hypotheses in the Falko Corpus. A Flexible Multi+Layer Corpus Architecture. In: Ana Díaz+Negrillo (ed.): Automatic Treatment and Analysis of Learner Corpus Data: John Benjamins.
- Schiller, Anne; Teufel, Simone; Stöckert, Christine; Thielen, Christine (1999):** Guidelines für das Tagging deutscher Textkorpora mit STTS. Technical Report. University of Stuttgart; University of Tübingen.
- Schmid, Helmut (1995):** Improvements in Part+of+Speech Tagging with an Application to German. In: Proceedings of the ACL SIGDAT+Workshop. Dublin, Ireland.
- Schmid, Helmut; Laws, Florian (2008):** Estimation of Conditional Probabilities with Decision Trees and an Application to Fine+grained POS Tagging. In: Donia Scott (ed.): 22nd International Conference on Computational Linguistics. Coling 2008. International Conference on Computational Linguistics. Manchester, United Kingdom, 18 + 22 August 2008. COLING. Stroudsburg, Pa: Association for Computational Linguistics (ACL), S. 777–784.
- Toutanova, Kristina; Manning, Christopher D. (ed.) (2000):** Enriching the Knowledge Sources Used in a Maximum Entropy Part+of+Speech Tagger. Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora: Association for Computational Linguistics (ACL).
- Van Halteren, Hans, Walter Daelemans, and Jakub Zavrel (2001):** "Improving Accuracy in Word Class Tagging through the Combination of Machine Learning Systems." *Computational Linguistics* 27.2.199-229.APA