Hagen Hirschmann
hirschhx@hu-berlin.de

Berechnung von Lernschwierigkeiten im Deutschen als Fremdsprache am Beispiel von Adverbien und Adverbialen

Forschungskolloquium
Korpuslinguistik
08.12.2010

Fragestellungen

- Welche Faktoren beeinflussen den Erwerb einer gegebenen Fremdsprache (L2)?
- (Was sind limitierende Faktoren?)
- Welche korpuslinguistische Methoden können zum Ermitteln dieser Faktoren beitragen?

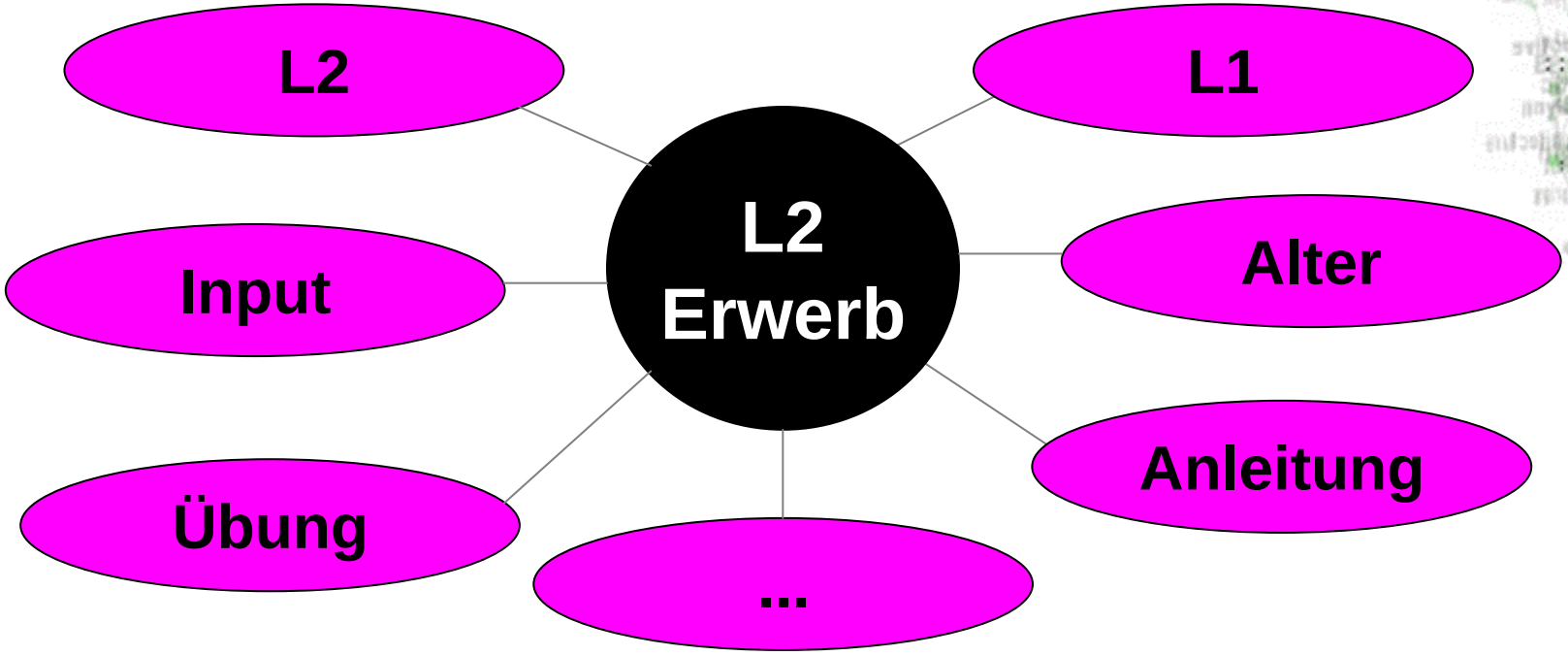
Gliederung

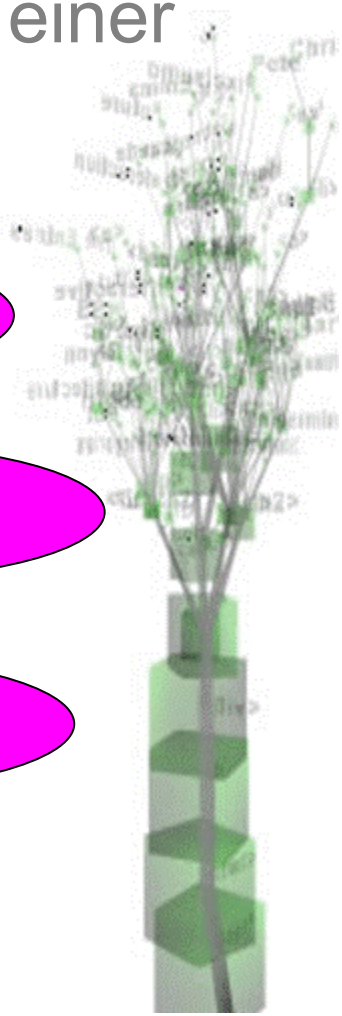
- Welche Faktoren beeinflussen den Erwerb einer gegebenen Fremdsprache (L2)?
  - Eingrenzung des Forschungsgegenstands
  - Forschungsziel
  - Methodische Überlegungen
- Welche korpuslinguistische Methoden können zum Ermitteln dieser Faktoren beitragen?
  - Ergebnisse aus Underuse-Studien zwischen Lexik uns Syntax
  - Ausblick: anstehende Studien

Gliederung

Welche Faktoren beeinflussen den Erwerb einer gegebenen Fremdsprache (L2)?

Einflussdomänen





L2-Einfluss auf den Spracherwerb

- Überblick über mögliche Einflussdomänen auf L2-Erwerb s. Gass&Selinker 2008
- Nur erschwerende o. limitierende Faktoren, die aus der L2 selbst hervorgehen

➢ Keine Interferenzen, kognitiven, psychischen, didaktischen Faktoren

- Einfluss der L2 unstrittig, Stärke dieses Einflusses unterschiedlich bewertet
- Modifizierte Fragestellung: Wie groß ist der Einfluss unterschiedlicher Eigenschaften der L2 auf den Spracherwerb?

L2-Einfluss auf den Spracherwerb

- Suche nach solchen grammatischen Elementen oder Strukturen in der Zielsprache, die von der gesamten Lernerpopulation schlecht, spät oder nicht erworben werden.
- Bekannte Phänomene: **Nominalflexion** (z.B. Diehl et al. 1991); **Wortstellung** (z.B. viele Arbeiten von Pienemann o. Pienemann&Clahsen, Boss 2004); **Modalpartikeln** (sehr viele Arbeiten, korpuslinguistisch z.B. Möllering 2004, Vyatkina 2007)
- Bezeichnung bzw. Terminus?
- Bisher: "strukturelle Lernschwierigkeiten"
- Aber: L1-Interferenzphänomene sind auch strukturell und lernerschwerend, fehlender authentischer L2-Input auch usw...
- Vorschlag: "L2-bedingte erwerblimitierende Faktoren"; kurz: "L2-Schwierigkeiten"

Wie ermittelt man L2-Schwierigkeiten?

Vorschlag 1

Marzari, Robert (2010): *Leichtes Englisch, schwieriges Französisch, kompliziertes Russisch. Evaluation der Schwierigkeiten des Englischen, Deutschen, Französischen, Italienischen, Spanischen, Russischen und Polnischen als Fremdsprache*

	E	D	F	I	S	R	P
1.1 Phoneminventar	4	4	3	2	2	4	4
1.2 Schwierige Laute	4	4	4	3	2	4	5
1.3 Schwierige Lautkombinationen	4	4	3	1	1	4	5
2.1 Schriftsystem	1	1	1	1	1	5	1
2.2 Diakritische Zeichen	1	5	4	3	2	1	5
2.3 Phonem-Graphem-Beziehung	5	3	5	2	1	4	4
3.1 Artikel	1	4	3	3	2	1	1
3.2 Substantivdeklination	1	4	2	2	1	5	5
3.3 Adjektivdeklination	1	5	2	2	1	4	5
3.4 Konjugation	1	3	5	4	4	3	4
3.5 Gebrauch der Tempora und Modi	3	3	4	4	4	2	2
3.6 Aspekt	4	1	1	1	1	5	5
4.1 Eigenheiten des GWS	1	4	2	3	3	4	4
4.2 Zugang zum Grundwortschatz	2	5	4	4	4	5	4

LESEN

4.3 Zugang zum Lesewortschatz	2	4	3	3	3	4	4
4.4 Wortschatzumfang / Synonymie	5	4	2	4	3	4	4
5.1 Satzbau	4	5	3	3	2	4	4
5.2 Textstrukturen	2	5	2	3	3	3	4

Grammatische und lexikalische Grundkompetenz:

	S	E	I	F	D	R	P
Σ	29	33	35	43	50	51	54
Ø	2,1	2,4	2,5	3,1	3,6	3,6	3,9

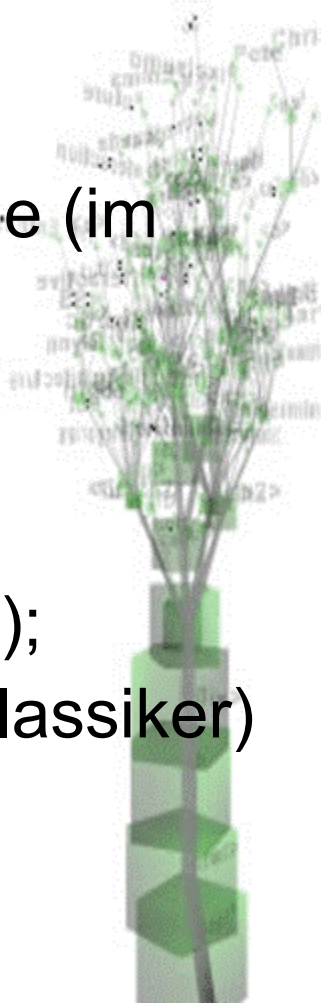
Lesekompetenz:

	S	E	I	F	R	D	P
Σ	40	46	48	53	66	68	70
Ø	2,2	2,6	2,7	2,9	3,7	3,8	3,9

➢Hypothese 1: "Was ich als L2-Schwierigkeit empfinde, ist eine L2-Schwierigkeit"

Wie ermittelt man L2-Schwierigkeiten?

- <u>Vorschlag 2 (systematischer):</u>
- Konstrativer (konfrontativer) Ansatz
  - Solche Elemente ermitteln, die in einer Sprache (im Vergleich zu anderen Sprachen) spezifisch ('besonders', 'auffällig') sind
  - ➢ Hypothese 2: Konstrastiv idiosynkratische Eigenschaften sind L2-Schwierigkeiten.
  - Vgl. Putzer 1994 (guter methodolog. Überblick);
  - Sternemann (z.B.) 1984, Helbig (z.B.) 1986 (Klassiker)
  - Götze&Helbig 2001 (HSK DaF)



Wie ermittelt man L2-Schwierigkeiten?

- <u>Vorschlag 3:</u>
- Fehleranalyse (in der Lernerkorpusforschung einer von zwei grundlegenden Ansätzen: EA; Granger (z.B.) 2002)
- Fehler kategorisieren, zählen, vergleichen

➢ Hypothese 3: Im Spracherwerb häufig produzierte Fehler sind L2-Schwierigkeiten.

- z.B. Diehl et al. 1991 (Behandlung der Nominalflexion im DaF auf Grundlage einer Fehlerstudie, die Deklinationsfehler als häufigste Fehlerursache ermittelt)

Wie ermittelt man L2-Schwierigkeiten?

- <u>Vorschlag 4 (wir nehmen diesen):</u>
- http://www.youtube.com/watch?v=gmOTpIVxji8
- Bsp: "*I sink, I'm sick.*"
- Engl. [θ] kontrastiv *relativ* selten → L2-Schwierigk.
- [s] statt [θ] gilt nachweisbar als typischer Lernerfehler → "
- Schauen wir nicht auf die Fehler bspw. von deutschen Englischlernern, sondern auf **<u>alle</u>** ihre Äußerungseinheiten, und zwar im Kontrast zu engl. Muttersprachlern:
- [θ] wird zu selten auftreten, [s] zu häufig.

Grundlage für alles Folgende: CIA-Underuse-Hypothese

- Hypothese 4: In der Lernersprache seltene Elemente deuten auf L2-Schwierigkeiten hin.
- Methodik: CIA (Contrastive Interlaguage Analysis, der zweite der beiden Ansätze Granger (z.B.) 2002)
- CIA hier nur als Vergleich von Muttersprachlern und Lernern
- Overuse (erst einmal) vernachlässigt
- Einschränkungen von H4:
  - **H 4 bezieht sich auf fortgeschrittene Lerner, die eigentlich alles ausdrücken können.**
  - **H 4 bezieht sich auf Elemente, die generell in den Lernerdaten vorkommen (es ist uns egal, wenn "*ob*" als Präposition mindergebraucht wird).**

Untersuchungsgebiete

Morphologie

Syntax

Ausgangsgebiet:
Wortarten

Phonologie

Semantik

Lexik

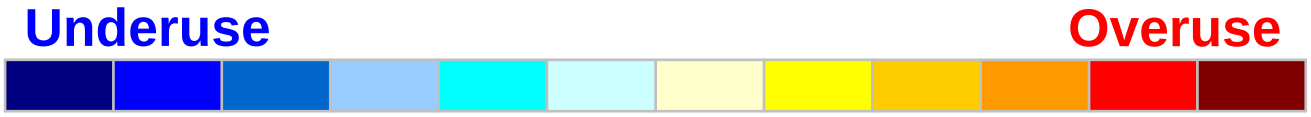
Datengrundlage

- Falko L2 und L1
- Für adverbielle Wortarten: manuell nachannotierte (subkategorisierte) pos-Daten
- Für Syntax: automatisch geparste Daten (Berkeley Parser)

L2 ADV_pos	L1 ADV_pos
37840 Token	**16772 Token**
L2 Berkeley (ZH1)	**L1 Berkeley (ZH1)**
Number of tokens: 123998 **Number of clauses:** 13586 **Number of corpus graphs:** 6749 **Average number of tokens:** 18.4 **Number of edges:** 182850	**Number of tokens:** 69313 **Number of clauses:** 7407 **Number of corpus graphs:** 3274 **Average number of tokens:** 21.2 **Number of edges:** 101358

CIA-Diagnostik in Falko

- Diagnosemethode:
- ➢ Vergleich normalisierter Frequenzen von beliebigen Elementen (Wörter, Wortarten, Phrasen, ...) zwischen Falko L2 und Falko L1
- Visualisierung: Stärke des Over- bzw. Underuses wird farbkodiert



(Amir Zeldes)

Excel Under/Overuse Addin: http://korpling.german.hu-berlin.de/~amir/uoaddin.htm

Bsp.:
Vergleich der Frequenzen von Wortartenketten (pos-Bigramme)

bigram	de	da	en	fr	pl	ru
$.-PPER	0.005297	0.009748	0.007963	0.006166	0.005801	0.007409
VVFIN-$,	0.006457	0.00776	0.006343	0.006937	0.006243	0.008391
PPOSAT-NN	0.008058	0.007247	0.007269	0.007066	0.006298	0.005802
ADV-ADV	**0.012858**	**0.010518**	**0.006111**	**0.006166**	**0.003094**	**0.002856**
ADV-APPR	0.009117	0.008016	0.005324	0.007837	0.004807	0.004642
PDAT-NN	0.005409	0.004233	0.005509	0.007837	0.007735	0.008837
ADV-ART	0.007629	0.006349	0.006898	0.005653	0.006133	0.004463

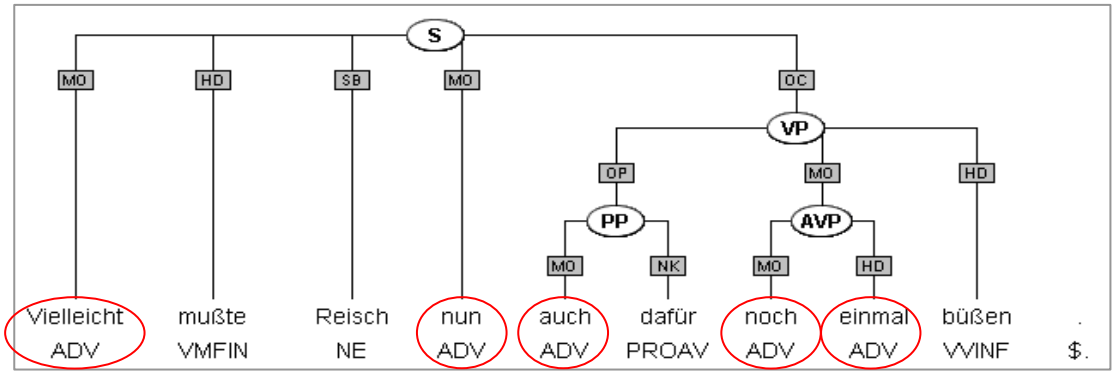
Aufeinanderfolgende ADVs werden von den Lernern unabhängig von ihrer L1 mindergebraucht.

ADV-Studie

- Korpusbefund:
  - "ADV" (STTS) am meisten mindergebrauchte Wortart
  - Dies gilt unabhängig von der Muttersprache der Lerner
  - Ketten von "ADV" (Bigramme und Trigramme) signifikant mehr mindergebraucht, als Unigramme vorhersagen
- Schlussfolgerungen?
  - Warum sollten Adverbien schwer sein (keine Flexion, einfache Morphologie)?
  - Schwere Syntax (Wortstellung, Komplexität von Strukturen, syntaktisch ambige Elemente)?
  - Schwere Semantik (schwer zu erfassende epistemische Bedeutungen, Polysemie)?

ADV-Studie

- Das Problem "ADV" (STTS):
  - Die Kategorie "ADV" in dem verwendenten STTS-Tagset ist eine zu grobe Klasse, in der alle großen adverbiellen Wortarten vereint sind. Zur Illustration: ADV-Treffer in einem Satz des TIGER-Korpus:



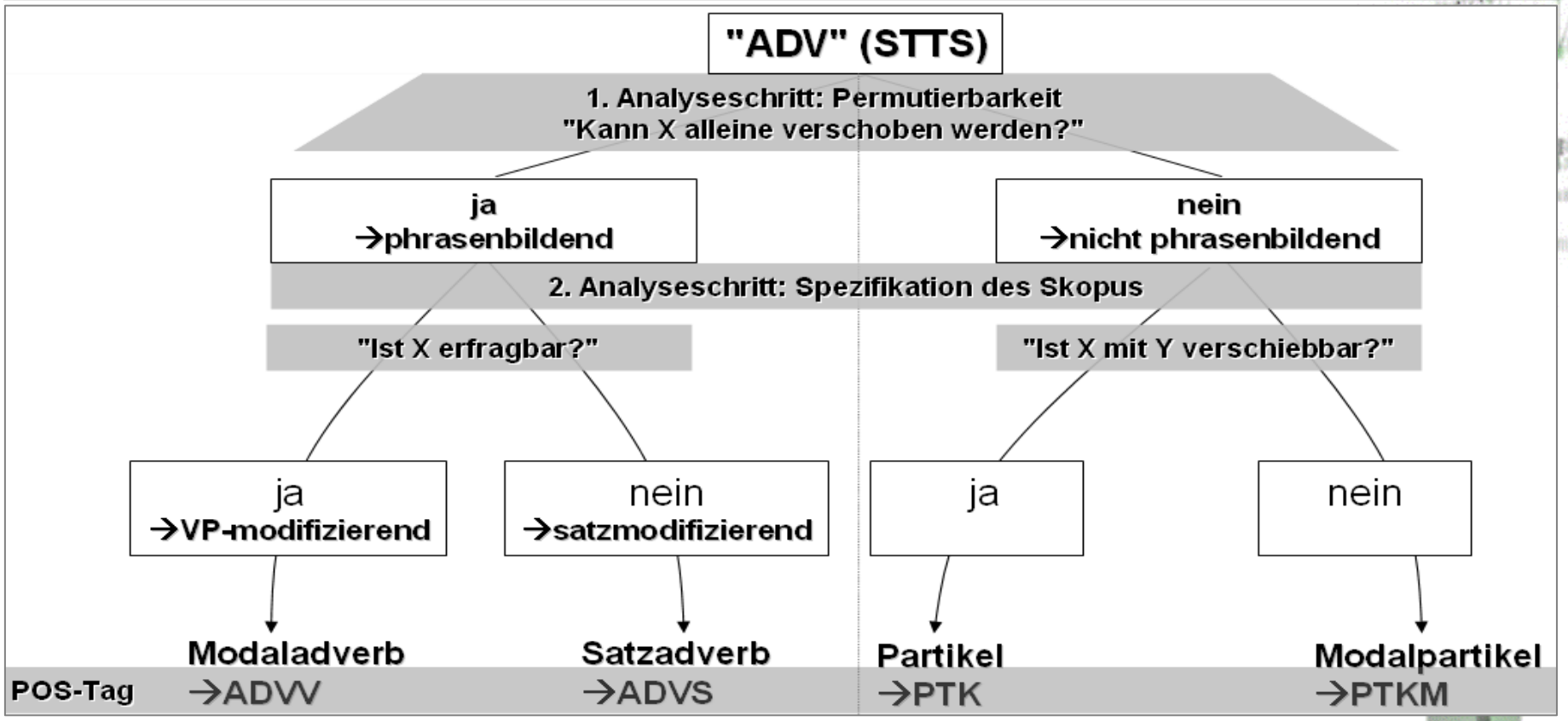
ADV-Studie

Aufgabe:

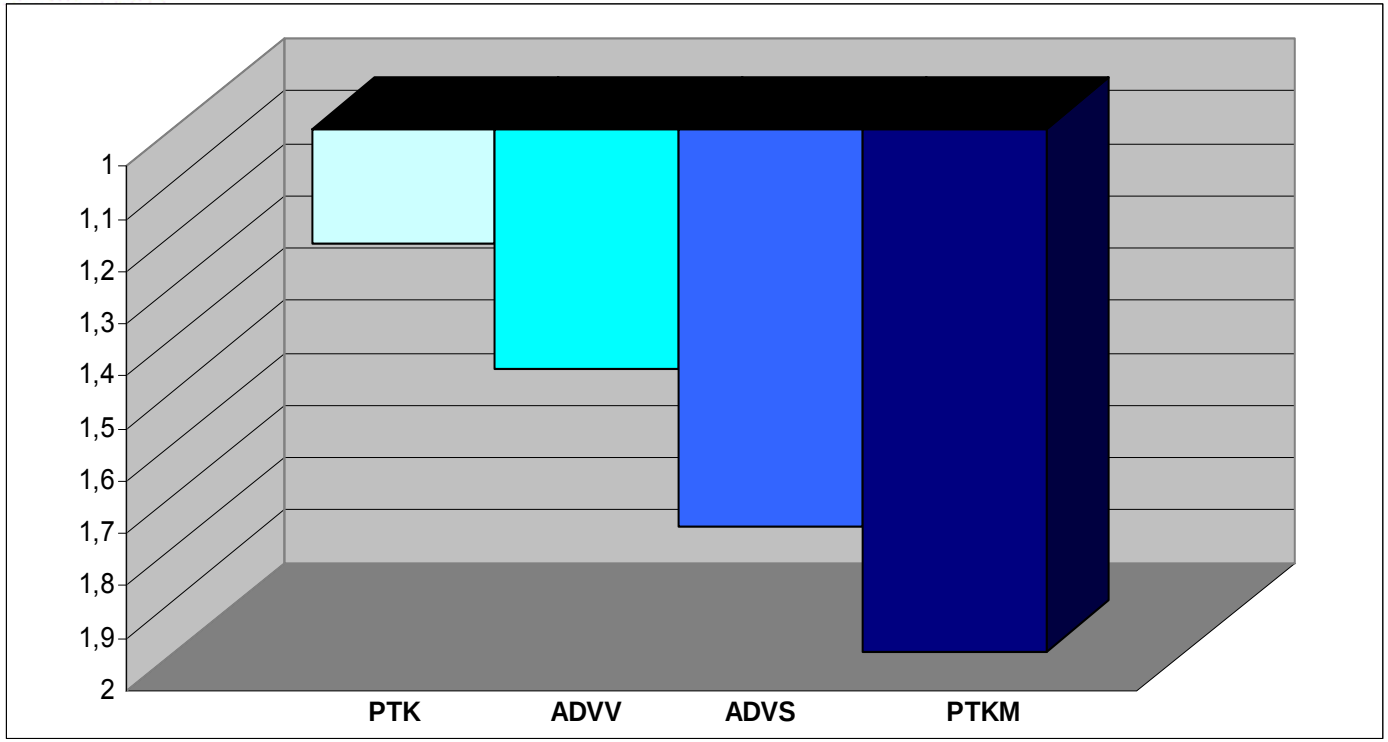
- Ausdifferenzierung von "ADV" nach syntaktischen Kriterien:
  - a) topologisch-distributionell
  - b) funktional
- Manuelle Annotation eines Teils von Falko
- Auswertung auf die Fragestellung: Zeigen die syntaktischen Klassen einen messbaren/signifikanten Unterschied im Mindergebrauch?
- Hypothese: Subklasse Modalpartikeln (als bekanntes Lernproblem) sollten die mindergebrauchteste Klasse sein.

ADV-Studie

Annotationsschema
zur syntaktischen Differenzierung von "ADV"



Ergebnisse



PTK: Partikel (***sehr*** *gut -* ***fast*** *drei Kilometer*)

ADVV: Modaladverb (***Bald*** *schneit es –* ***Dort*** *beginnt Bayern*)

ADVS: Satzadverbien (***Bestimmt*** *schneit es bald –* ***Leider*** *ist er nicht bescheuert*)

PTKM: Modalpartikeln (*Es schneit* ***wohl*** *gerade – Hast du* ***vielleicht*** *nen Euro?*)

Zur Terminologie u.a. Pittner 1999

Ergebnisse

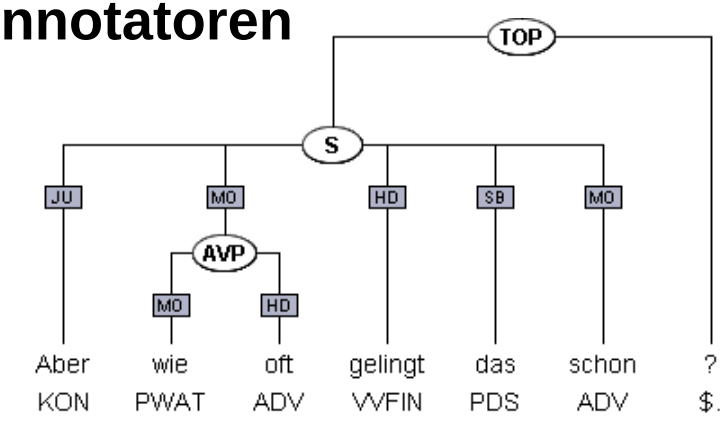
- Es gibt einen signifikanten Unterschied zwischen den einzelnen Wortarten.
- Die Syntax hat einen messbaren Effekt auf den Gebrauch adverbieller Wörter.
- Die Ergebnisse aus den Korpusdaten stimmen mit den Annahmen aus der bisherigen DaF-Forschung überein.

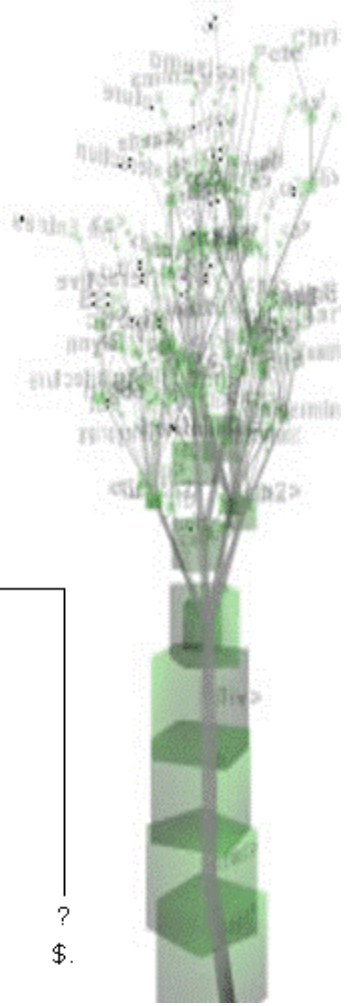
Schlussfolgerung

- Wenn Syntax relevant für den Gebrauch von Wörtern ist, müssen wir jetzt Syntax machen.
- Wir können Syntax nicht am Token/Wort analysieren.
- Wir nehmen den Aufwand auf uns, Falko syntaktisch zu annotieren.

Falko – Syntaxannotation

- **Zielhypothese 1 des L1- und L2-Korpus**
- **Berkeley parser**
- **Trainingsdaten: 48473 Graphen der TiGer-Baumbank**
- **Parserevaluation:**
- **Manuelle Erzeugung eines Goldstandards**
  - **200 Sätze zufällig aus L1 und L2**
  - **Durchschnittl. Satzlänge im Goldstandard ist repräsentativ für L1/L2-Daten (L1 all: 21.1 / L1 gold: 20.8) (L2 all: 18.3 / L2 gold: 18.3)**
  - **Manuelle Korrektur zweier Annotatoren**

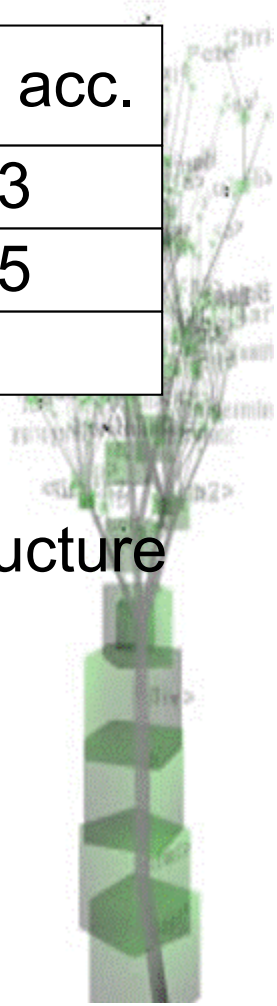




Parserevaluation: Vertrauen wir den Parserdaten?

- evaluation of constituent structure (evalb)

>40	Precision	Recall	F-Score	Tagging acc.
L1	73.61	74.00	73.80	91.93
L2	77.59	79.04	78.31	92.85
Negra*	80.01	80.01	80.01	



- L2 easier to parse than L1
- possible reasons: sentence length / L1 syntactic structure might be more complex

➢ we can use parser output to compare L1 and L2

- *Berkeley results on the Negra Treebank (Petrov & Klein, 2007)

Grundlage der Falko-Syntaxstudien

- Fragestellung I:
  - Gibt es eine nachweisbare reinen syntaktische Ursache des Mindergebrauchs von ADVs?
    - Die Lerner könnten den Mindergebrauch von ADVs (bzw. wortförmigen Modifikatoren) kompensieren, indem sie auf andere Elemente (Phrasen) ausweichen.

Grundlage der Falko-Syntaxstudien

- Fragestellung II:
  - Welche syntaktischen Gegebenheiten wirken sich auf den Gebrauch von Modifikation und Modifikatoren aus?
    - Stellungsfelder?
    - Satzart?
    - Komplexität
  - Hypothese: Modifikation im Vorfeld ist einfacher als im Mittelfeld.

➢ Gibt es aber messbare Effekte?

Zusammenfassung der Ergebnisse

VF Matrixsatz	L1 absolut	L1 norm	L2 absolut	L2 norm
MO_all	1107	207,77027	2211	223,69486
ADV	485	91,0285285	759	76,790773
PP	338	63,4384384	899	90,9550789
AVP	81	15,2027027	99	10,0161878
ADJD	33	6,19369369	64	6,47511129
PROAV	112	21,021021	287	29,0368272
AP	11	2,06456456	12	1,21408337

Nebensätze	L1 absolut	L1 norm	L2 absolut	L2 norm
MO_all	2467	916,759569	3251	831,245206
ADV	935	347,454478	998	255,177704
PP	905	336,306206	1248	319,099974
AVP	84	31,2151616	116	29,6599335
ADJD	177	65,7748049	270	69,0360522
PROAV	41	15,2359718	84	21,4778829
AP	63	23,4113712	77	19,6880593

MF Matrixsatz	L1 absolut	L1 norm	L2 absolut	L2 norm
MO_all	665	247,12003	733	187,420097
ADV	357	132,664437	381	97,4175403
PP	156	57,9710145	168	42,9557658
AVP	32	11,8914902	25	6,39222705
ADJD	47	17,4656262	57	14,5742777
PROAV	23	8,54700855	35	8,94911787
AP	25	9,29022668	17	4,3467144

Zusammenfassung

Zu den Forschungsfragen:

Es lassen sich rein syntaktische Effekte nachweisen:

- Modifikation ist ganz allgemein mindergebraucht (Mindergebrauch des Funktionslabels 'MO')
- Modifikation im Vorfeld ist häufiger als Modifikation im Mittelfeld (immer verglichen zum L1-Standard)
- Die Komplexität ('Phrasenhaftigkeit') unterschiedlicher Modifikatoren scheint kein messbarer Faktor zu sein.

Nebenbemerkung: Lexik

- Wir haben mit ANNIS, CQP) immer die Möglichkeit, die abstrakten syntaktischen (oder semantischen oder morphologischen...) Klassen wieder auf Lexeme herunterzubrechen.
- Bsp.: Underuse-Statistik auf frequenten zweigliedrigen Adverbialphrasen:

B	C	D	E	F
L1 norm	L1 tot	form	L2 tot	L2 norm
21,5315236	487	ALL	547	13,0720616
0,92846406	21	gar_nicht_	12	0,28677261
0,53055089	12	immer_mehr_	9	0,21507946
0,70740118	16	immer_noch_	23	0,54964751
0,22106287	5	immer_weiter	5	0,11948859
0,92846406	21	immer_wiede	8	0,19118174
0,13263772	3	nicht_ganz_	7	0,16728402
0,26527544	6	nicht_immer_	21	0,50185207
2,52011672	57	nicht_mehr_	41	0,97980643
1,10531435	25	nicht_nur_	91	2,17469232
0,1768503	4	nicht_so_	12	0,28677261
0,13263772	3	noch_einmal_	9	0,21507946
0,26527544	6	noch_nicht_	13	0,31067033
0,13263772	3	nur_dann_	4	0,09559087

Zusammenfassung

Allgemein:

- Wir haben eine Methode zur Ermittlung von L2-Schwierigkeiten entwickelt. Die Daten zeigen, dass die Methode taugt.
- Diese Methode hat auf der Wortarten- und Syntaxebene zu interessanten Ergebnissen und Forschungshypothesen geführt.
- Diese Methode lässt sich auf alle grammatischen Elemente/Phänomene anwenden.
- →Ausblick:

Ausblick

- Messung semantischer und morphologischer Einflussfaktoren auf den Gebrauch modifizierender Wörter
- Zusammenführung der Ergenisse: Welche Eigenschaften von Wörtern und der Syntax sind relevant für den Erwerb und den Gebrauch von Modifikatoren?

Danke!

Falko-Mitwirkende:

Torsten Andreas, Jia Wei Chan, Seanna Doolittle, Thomas Krause, Cedric Krummes, Anke Lüdeling, Marc Reznicek, Karin Schmidt, Maik Walter, Amir Zeldes, Florian Zipser

email: hirschhx@hu-berlin.de