



Slavische Sprachen für (Computer-) Linguisten

"Slavic languages are sufficiently similar and sufficiently different to provide an attractive research laboratory." (Corbett 1998)

Eigenschaften & Phänomene



z.B.	Bulgarisch	Polnisch	Russisch	Serbo-Kroatisch	Slowenisch	Sorbisch	Tschechisch
morph.Kasus	nein	ja	ja	ja	ja	ja	ja
Gen.d.Neg.	---	ja	ja, meist lokal	nein	ja	fakultativ im Niedersorb. etwas öfter	Nein (veralt.)
Artikel	ja	nein	nein	nein	nein	nein	nein
Infinitiv	nein	ja	ja	ja	ja	ja	ja
Verbkomplex	ja	nein	nein	nein	nein	nein	nein
Null-Kopula	nein	jein	ja	nein	nein	nein	nein
präd.Klitika	verbal	frei (auch Quasi-Klitika)	keine	2P nach der 1.Konst o. dem 1.Wort	2P nach der 1.Konst.	frei	2P nach der 1.Konst. (auch Quasi-Klitika)
cl.cluster	ja	(nein)	---	ja	ja	(nein)	ja
cl.replication	ja	nein	---	nein	nein	nein	nein
cl.haplology	nein	ja	---	?	?	?	ja
cl.scrambling	nein	ja	---	?	?	?	ja
IS-Faktoren	Wortstell. plast.Inton. Definitheit cl.doubling	Wortstell. plast.Inton. IS-Marker	Wortstell. plast.Inton. IS-Marker	Wortstell. ?plast.Inton.	Wortstell. ?plast.Inton.	Wortstell. ?plast.Inton.	Wortstell.

Besonderheiten der slavischen Klitisierung



● Background (Zwicky 1977)

- A *simple clitic* is an element of some basic word class and occupies the normal syntactic position of a non-clitic word of its category.
- A *special clitic* occupies an “unusual” position: special principles and mechanisms must be invoked for obtaining the correct placement.
- *While simple clitics are only phonologically deficient, special clitics have additional restrictions which require them to cluster.*

● Common Slavic inventory of special predicative clitics

- *pro-forms cross-referencing core arguments* of a predicate
- defined in terms of *morphosyntactic features*
- resulting in a distinguished module of inflecting clitics
- subject to parameterisation across Slavic languages



- *Bulgarisch* → **Complete Set**
 - reflexiv: **cu** & **ce**
 - pronominalisierte Objekte (*“Kongruenzmarker” bei Objektverdopplung*)
 - Person-Numerus-Auxiliar (*Vollparadigma, Subjektkongruenz*)
- *Tschechisch* → **Restricted Set**
 - reflexiv: **si** & **se** (*Haplologie, keine multiple Reflexivmarkierung*)
 - pronominalisierte Objekte (*z.T. Quasi-Klitika*)
 - Person-Numerus-Auxiliar (1&2, *Subjektkongruenz*)
- *Polnisch* → **Reduced Restricted Set**
 - reflexiv: nur **się** (*Haplologie*)
 - pronominalisierte Objekte (*“simple clitics”, z.T. keine Klitika*)
 - bewegliche Auxiliar-Flexion (1&2, *Subjektkongruenz*)
- *Russian* → **empty**
 - Reflexivflexion: **-ся** / **-сь**

Verbkomplex & Verbalklitika (Bulgarisch)



Togava si štjala da si mi gi pokazala veče.
then AUX-2sg FUT-AUX-sg.f PRT AUX-2sg DAT-1sg ACC-3pl shown-sg.f already

➤ *'You would have already shown them to me by that time (reportedly).'*

Az na Ivan pokazvala li sym mu gi njakoga dokumentite?
I-1sg to Ivan-3sg.m shown-sg.f Q AUX-1sg DAT-3sg.m ACC-3pl ever paper-3pl.def

➤ *'Have I ever shown the papers to John?'*

● host-category-bound

➤ *belong to the syntactic domain of a particular (morphosyntactic) category which also hosts them prosodically [Av'97]*

● Gruppierung im Cluster

➤ $AUX_{non3sg} < DAT < ACC < AUX_{3sg}$

Wackernagelklitika (Tschechisch)



● “second” position

➤ *What is in the “first” position?*

Minulý týden v neděli ráno doma v koupelně se Jan při holení nebezpečně zranil.

last week on Sunday morning at-home in bathroom REFL Jan at shaving dangerously hurt

● prosodically autonomous clitics

➤ *inherently related to a “second” position rather than to a prosodic host of a particular (morphosyntactic) category [Av’97]*

● pure clitics vs. semi-clitics

Wackernagelklitika (Slowenisch)



- “second” position

- *(after the “first phrase”)*

- prosodically neutral clitics

- *after an intonational pause*

- *clitic-initial sentences (“topic-drop”)*

Si ga videl?

AUX-2sg ACC-3sg.m saw

(‘Have you seen him?’)

Se mi je smejal.

REFL DAT-1sg AUX-3sg laughed

(‘He was laughing at me.’)

Wackernagelklitika (Serbisch-Kroatisch)



● “second” position

➤ *(after the “first word” or the “first phrase”)*

Taj čovjek joj ga je danas kupio.

that man DAT-3sg.f ACC-3sg.n AUX-3sg today bought

Taj joj ga je čovjek danas kupio.

● prosodically enclitics

Ona mu ga je dala, a i ja sam mu ga dala.

she DAT-3sg.m ACC-3sg.n AUX-3sg gave and also I AUX-1sg DAT-3sg.m ACC-3sg.n gave

Ona mu ga je dala, a i ja sam mu ga.

Ona mu ga je dala, a i ja sam mu.

Ona mu ga je dala, a i ja sam.

Gruppierungsmuster im 2P-Cluster



● Tschechisch

- $AUX < ETH.DAT < REFL(dat/acc) < DAT < ACC (GEN)$
- *& co-occurrence restrictions: *two phonologically identical clitics*
**two reflexive clitics*

● Serbo-Kroatisch

- $AUX_{non3sg} < DAT < ACC/GEN < REFL < AUX_{3sg}$

Klitika im Polishen



● argument / reflexive clitics

- “free” clitics (*special clitics with simple clitic behaviour*)
- optional clitic climbing
- haplology (*się*)

● “floating” weak auxiliaries

- “auxiliary inflection” on verbs but also hosted by other categories

● no clitic-cluster formation

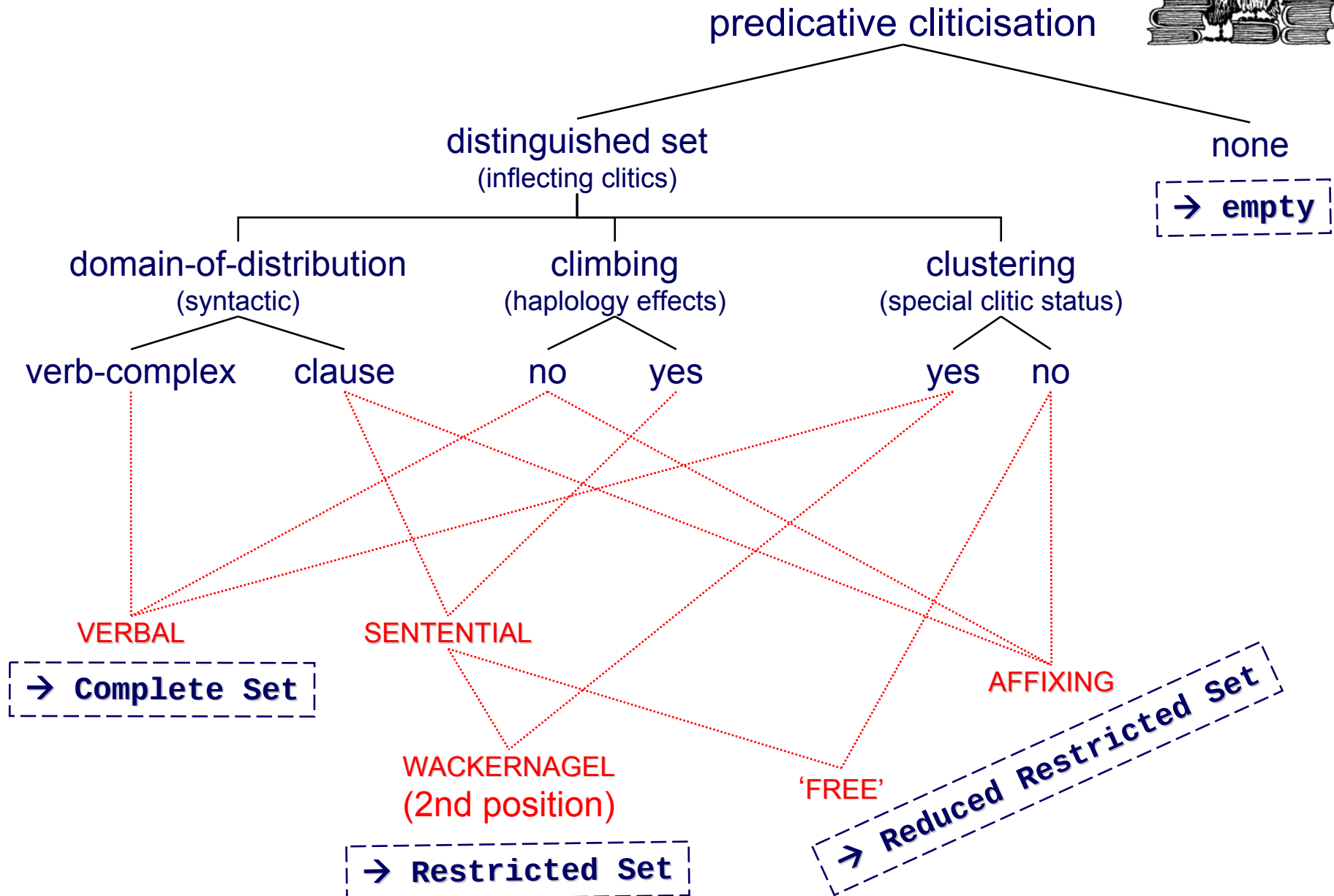
(?!) Pokazali cię mu wczoraj.

showed ACC-2sg DAT-3sg.m yesterday

● “second position” effects

- non-initial in the clausal domain
- resist non-immediate post-verbal positioning: left-edge tendency

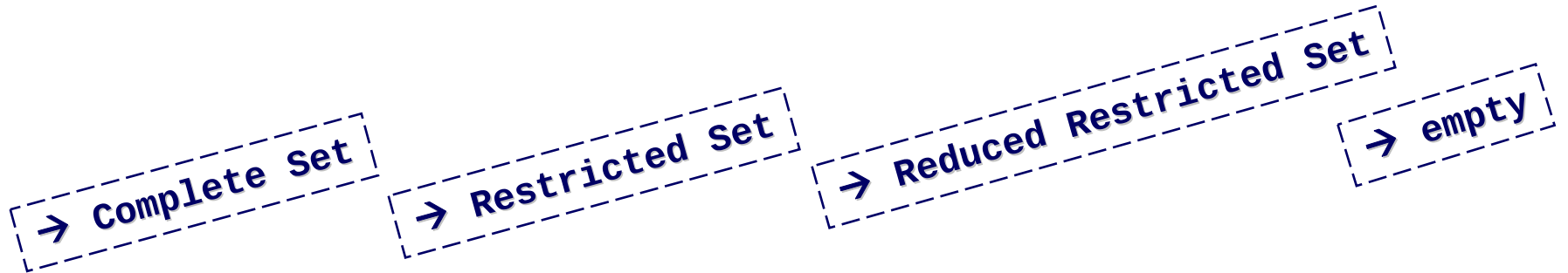
Cross-Slavic parameterisation



Beobachtete Tendenz ➤ Süd ➤ West ➤ Ost



Bulgarisch / Tschechisch / Polnisch / Russisch



Umgekehrte Tendenz ➤ Ost ➤ West ➤ Süd



Russisch / Polnisch / Tschechisch / Bulgarisch

- elaborate case system → case syncretism → case remnants
- copula-less predication → morphosyntactic verb complex
- infinitival complements → finite complements



Freie Wortstellung (?!?)



(1) *Colorless green ideas sleep furiously.*

Бесцветные зелёные идеи спят яростно

(2) **Furiously sleep ideas green colorless.*

Яростно спят идеи зелёные бесцветные

● Generell (auf Satzebene „syntaktisch frei“):

- *Die Wortstellung in den slavischen Sprachen ist nicht frei, sondern von der jeweiligen Informationsstruktur (IS) abhängig.*
- *Insofern kann von Wohlgeformtheit von Sätzen in ihren jeweiligen Kontexten gesprochen werden.*

● Speziell (auf Satzgliedebebene weniger flexibel)

- *z. T. grammatikalisiert*
- *Umstellungen und diskontinuierliche Einheiten möglich (durch IS gesteuert)*



Informationsstruktur des Satzes

● Fokus (Rhema)

- *bringt die Kommunikation voran*
- *ist die (neue) Information*
- *trägt etwas zum Kontext bei*
- *ist kontextunabhängig*

● Hintergrund: *link* („Topik“) / *tail* („Füllung“)

- *erstellt eine Verknüpfung mit dem vorhergehenden Text / Diskurs (link!)*
- *beinhaltet bekannte Daten*
- *trägt nicht zum Kontext bei*
- *ist kontextabhängig*
- *wird vorausgesetzt*



Informationsstruktur des Satzes

● Fokussierung als Default-Informationszustand

- *unmarkierter Fall*
- *neutrale (= am meisten informative) Sätze $\Leftrightarrow$ kanonische Konfigurationen (syntaktisch und prosodisch)*
- *Fokus ist mit kanonischen strukturellen Realisierungen assoziiert*

● Mittel der Informationsstrukturierung

- *Linearisierung syntaktischer Einheiten*
- *Beitrag der Intonation*
- *lexikalische Mittel, IS-Partikel, ...*
- *Determiniertheit und Duplizieren durch Klitika (vgl. Bulgarisch)*
- *usw.*

Linearisierung



● Morphosyntaktische Besonderheiten

- *Betreff. „head-marking“ vs. „dependent-marking“*
- *Kasussystem & Klitika*
- *Kongruenzphänomene*

● Beitrag der Intonation



T. Avgustinova, *Word order and clitics in Bulgarian*. Saarbrücken
Dissertations in Computational Linguistics and Language Technology
(Universität des Saarlandes / DFKI, Saarbrücken, 1997), vol. 5.

T. Avgustinova, *Language Family Oriented Perspective in Multilingual
Grammar Design*. Linguistik International 17 (Peter Lang - Europäischer
Verlag der Wissenschaft, Frankfurt am Main, Berlin, Bern, Bruxelles,
New York, Oxford, Wien, 2007).