



Word order variants as pragmatic constructions: a corpus-based case study of German

Kris Heylen



University of Leuven
RU Quantitative Lexicology and Variational Linguistics

Purpose of the talk

Show for a type of word order variation in German that

- word order variants can be described as **highly schematic constructions** with a **pragmatic** function
- these highly schematic constructions have to be characterized in **relation both to each other** and **to less schematic constructions** within the *Constructicon*

Overview

1. Introduction

- Word order variation in the *Mittelfeld*
- Specific type of variation

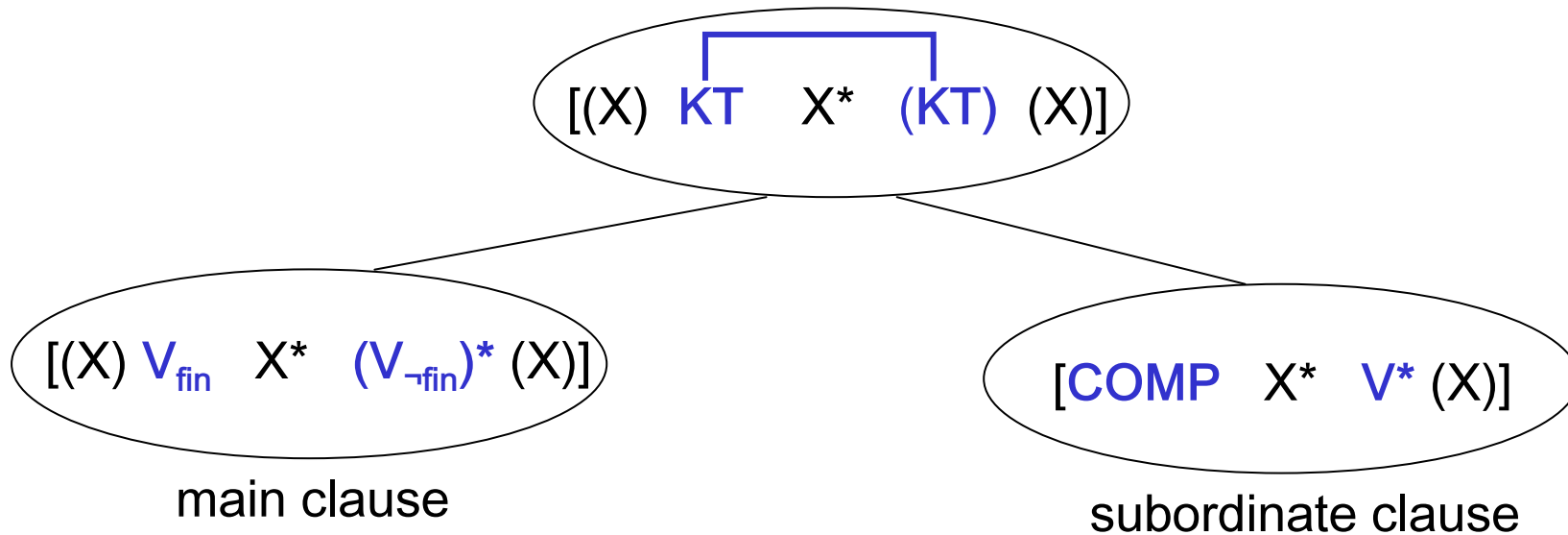
2. Corpus analysis

- Multivariate Analysis
- Pragmatic function: highlighting salient referents (in a set)
- Relation to less schematic constructions

3. Conclusions

- word order constructions in the *Constructicon*

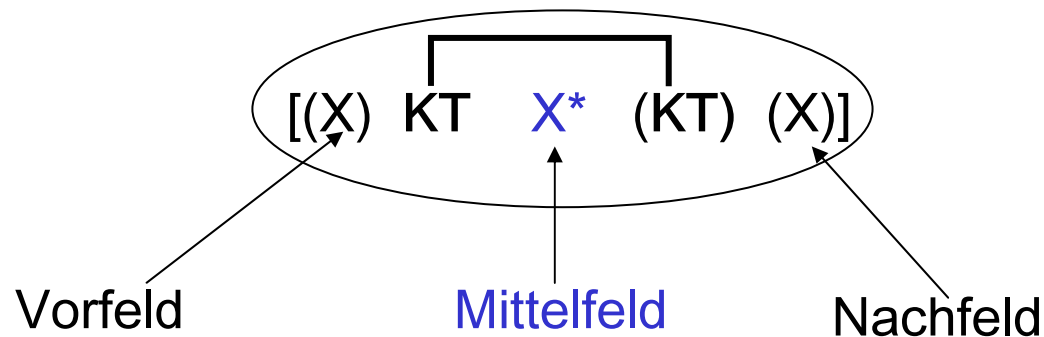
Introduction: Klammer-construction



Ich **habe** schlecht **geschlafen**, weil
Er **kennt** jemand , der
Schläfst du auch fast **ein** ?

weil es zu viel Lärm **gab**
der ihm **helfen will**, es zu tun

Introduction: Mittelfeld variation



Can contain multiple constituents
e.g. verbarguments, but not
always in the same order!

NP_{nom}	NP_{acc}	$><$	NP_{acc}	NP_{nom}
NP_{nom}	NP_{dat}	$><$	NP_{dat}	NP_{nom}
NP_{dat}	NP_{acc}	$><$	NP_{acc}	NP_{dat}
...				

Introduction: specific variation type

Nominal subject and pronominal object in the Mittelfeld

NP(pron)_{acc/dat} NP(full)_{nom} >< NP(full)_{nom} NP(pron)_{acc/dat}

		Mittelfeld		
Object-first	Später nahm	ihn	der Chef	beiseite
Subject-first	...wie	viele Frauen	sich	wünschen

examples from the Negra corpus

- In both main and subordinate clauses, with personal and reflexive pronouns, accusative and dative objects
- Object-first is most common, but there are no clear differences in meaning or grammaticality

Introduction: specific variation type

- Variation is well known in reference grammars, and fairly common but little studied in the word order literature:
- Generative paradigm
 - Lenerz (1994): optional movement of S_{nom} to SpecIP
 - Müller (2002): optional S-raising to SpecT (matched constraints)
 - ⇒ no motivation/trigger for movement identified
- Functional paradigm
 - Hoberg (1997): marked subject-first seems to occur more often with proper noun subjects, but no clear tendencies
 - ⇒ Free variation or different usage constraints?

Corpus analysis: multivariate analysis

- NEGRA corpus (400K, compiled at Saarbrücken)
 - Newspaper materials (*Frankfurter Rundschau*)
 - Morpho-syntactically annotated
- Semi-automatic extraction of relevant instances
 - Clauses with a nominal subject and one pronominal object in the Mittelfeld
 - 995 observations
- Coded for 17 language internal explanatory variables

#		VARIABLE	VALUES
1	Cl.	Clause type	<main>, <subordinate>
2		Clause coda	<following element>, <empty>
3	Pron. object	Pronoun type	<personal>, <reflexive>
4		Case	<accusative>, <dative>
5		Case ambiguity	<i>5 point ordinal scale</i>
6		Valency	<lexicalized>, <not lexicalized>
7		Animacy	<animate>, <inanimate>
8	Nom. Subject	Noun type	<Demonstrative>, <Proper>, <Appellative>, <Indefinite>
9		Length	<i>Lengthdifference subject/object</i>
10		Definiteness	<definite>, <indefinite>
11		Semantic role	<Agent>, <Experiencer>, <Theme>
12		Animacy	<animate>, <inanimate>
13		Accessibility	<i>10 point ordinal scale</i>
14		Set of referents activated	<set activated>, <set not activated>
15	Verb	Position of lexical full verb	<1 st Klammer>, <2 nd Klammer>
16		Genus Verbi	<active>, <passive>
17		Semantic verb class	<existential>, <not-existential>
18	Ext.	Medium	<written>, <spoken>
19		Region	<D>, <A>, <CH>

Corpus analysis: multivariate analysis

- NEGRA corpus (400K, compiled at Saarbrücken)
 - Newspaper materials (*Frankfurter Rundschau*)
 - Morpho-syntactically annotated
- Semi-automatic extraction of relevant instances
 - Clauses with a nominal subject and one pronominal object in the Mittelfeld
 - 995 observations
- Coded for 17 language internal explanatory variables
- Statistical analysis:
 - Separate effect of each explanatory variable
 - Combined effect in multivariate model: Logistic Regression model

Corpus analysis: multivariate analysis

- Distribution of the two variants

Subj.-first	Object-first
106 (11%)	889 (89%)

- Object-first clearly unmarked variant and Subject-first marked variant
- Specific usage conditions for *marked* variant?
- Statistical Logistic regression model
 - models probability of marked order as function of expl. Variables
 - Only retains variables that have a significant effect
 - Quantifies the strength of each variable's effect

Corpus analysis: multivariate analysis

step	Variable	Odss Ratios
1	SET OF SUBJ. REFERENTS (yes >< no)	21
2	CLAUSE TYPE (Subordinate >< main)	7.5
3	SUBJ. SEM. ROLE (agent >< non-agent)	4
4	SUBJ. ACCESSIBILITY (high >< low)	3
5	OBJ. CASE AMBIGUITY (ambiguous >< not ambig.)	2.5
6	SUBJECT LENGTH (short >< long)	2

Logistic Regression
stepwise selection
(subj-first modelled)
Model sign. Wald < .01
Goodness of fit > .2

Corpus analysis: multivariate analysis

step	Variable	Odss Ratios
1	SET OF SUBJ. REFERENTS (yes >< no)	21
2	CLAUSE TYPE (Subordinate >< main)	7.5
3	SUBJ. SEM. ROLE (agent >< non-agent)	4
4	SUBJ. ACCESSIBILITY (high >< low)	3
5	OBJ. CASE AMBIGUITY (ambiguous >< not ambig.)	2.5
6	SUBJECT LENGTH (short >< long)	2

Fairly weak factors
that do not raise
frequency of marked
order above 15%
(secondary effects)

Corpus analysis: Set activation

	Subj.-first	Object-first
Set activated	67 (52%)	63 (48%)
Set not activated	39 (5%)	889 (95%)

- In the discourse context, a set of alternative referents for the subject referent is activated, either explicitly or implicitly. The set can, but need not be interpreted contrastively (see examples)
- Related to
 - Givón (1983): Potential Interference
 - Chafe (1976)/Rooth (1992): Focus as list of alternatives
 - Birner & Wald (1998): Partially ordered set relations
- Tentative operationalization as yet, but attempt at ‘unprejudiced’ coding by masking order

Corpus analysis: Set activation

- Das schon für Montag geplante Treffen **Vassiliou** mit **Denktash** mußte verschoben werden , weil **Denktash sich** weigerte , eine Einladung zu akzeptieren , auf der er nur als " Führer " der türkischen Zyperer bezeichnet , sein Kollege Vassiliou hingegen als Präsident der griechischen Zyperer aufgeführt war .

SET "Cypriote leaders"
[Vassilou, Denktash]

Corpus analysis: Set activation

- Das schon für Montag geplante Treffen Vassiliou mit Denktash mußte verschoben werden , weil Denktash sich weigerte eine Einladung zu akzeptieren , auf der er nur als " Führer " der türkischen Zyprier bezeichnet , sein Kollege Vassiliou hingegen als Präsident der griechischen Zyprier aufgeführt war .
- Im Deutschen Gewerkschaftsbund (DGB) gibt es keine einhellige Meinung zu den jüngsten Vorschlägen einer sogenannten Gesundheitsreform , die von CDU / CSU , FDP und SPD gemeinsam erarbeitet wurden .Während **das Vorstandsmitglied der IG Metall, Peter Kirch,** es bei einer sozialpolitischen DGB-Veranstaltung im Main-Kinzig-Kreis als positiv empfand , daß die SPD den früheren Vorschlägen zur Kostendämpfung " einige Giftzähne " gezogen habe , sieht **der Kreisvorsitzende Sepp Sigulla** den Kompromiß wesentlich kritischer .

SET "Union Leaders"
[Kirch, Sigulla]

Corpus analysis: Set activation

- Das schon für Montag geplante Treffen Vassilious mit Denktash mußte verschoben werden , weil Denktash sich weigerte , eine Einladung zu akzeptieren , auf der er nur als " Führer " der türkischen Zyprer bezeichnet , sein Kollege Vassiliou hingegen als Präsident der griechischen Zyprer aufgeführt war .
- Im Deutschen Gewerkschaftsbund (DGB) gibt es keine einhellige Meinung zu den jüngsten Vorschlägen einer sogenannten Gesundheitsreform , die von CDU / CSU , FDP und SPD gemeinsam erarbeitet wurden .Während das Vorstandsmitglied der IG Metall , Peter Kirch, sich bei einer sozialpolitischen DGB-Veranstaltung im Main-Kinzig-Kreis als positiv empfand , daß die SPD den früheren Vorschlägen zur Kostendämpfung " einige Giftzähne " gezogen habe , sieht der Kreisvorsitzende Sepp Sigulla den Kompromiß wesentlich kritischer .
- Morgens um 9.00 Uhr spannt die 21jährige zwei Sonnenschirme auf , stellt die Rosen darunter und wartet auf Kundschaft - bis nachmittags um 17.00 Uhr. Dann löst ihr Bruder sie für die letzten drei Stunden ab

SET "Members of family business"
[21-Jährige, Bruder]

Corpus analysis: Set activation

- Die Weltwirtschaftsordnung , die die transnationalen Unternehmen und Banken verwirklicht haben , funktioniert . Sie entspricht zwar nicht den Vorstellungen , die insbesondere die Dritte Welt sich in den siebziger Jahren davon gemacht hat

SET “economic players”
[Unternehmen, Banken, Dritte Welt]

Corpus analysis: Set activation

- Die Weltwirtschaftsordnung , die die transnationalen Unternehmen und Banken verwirklicht haben , funktioniert .Sie entspricht zwar nicht den Vorstellungen , die insbesondere die Dritte Welt sich in den siebziger Jahren davon gemacht hat
- Die percussive Version des zu Tode genudelten " Bolero " , fiel **dermaßen** obskur aus , daß **selbst Ravel sich** amüsiert hätte

**IMPLICIT SCALED SET "critics of Bolero performance"
[Ravel.....]**

Corpus analysis: Set activation

- Die Weltwirtschaftsordnung , die die transnationalen Unternehmen und Banken verwirklicht haben , funktioniert .Sie entspricht zwar nicht den Vorstellungen , die insbesondere die Dritte Welt sich in den siebziger Jahren davon gemacht hat
- Die percussive Version des zu Tode genudelten " Bolero " , fiel dermaßen obskur aus , daß selbst Ravel sich amüsiert hätte
- Das Bild der Welt als wirtschaftlicher Handlungszusammenhang rundet sich , wenn man bedenkt , daß ein großer Teil des Welthandels sich innerhalb der transnationalen Firmen abspielt , also eigentlich schon zum Binnenhandel geworden ist .

IMPLICIT SET "subparts of world trade"
[ein großer Teil]

Corpus analysis: Set activation

- Die Weltwirtschaftsordnung , die die transnationalen Unternehmen und Banken verwirklicht haben , funktioniert .Sie entspricht zwar nicht den Vorstellungen , die insbesondere die Dritte Welt sich in den siebziger Jahren davon gemacht hat
- Die percussive Version des zu Tode genudelten " Bolero " , fiel dermaßen obskur aus , daß selbst Ravel sich amüsiert hätte
- Das Bild der Welt als wirtschaftlicher Handlungszusammenhang rundet sich , wenn man bedenkt , daß ein großer Teil des Welthandels sich innerhalb der transnationalen Märkte abspielt , also eigentlich schon zum Binnenhandel geworden ist

SET "programmes on on telly"

[andere Programme, Waschmittelwerbung]

BUT

- Denn erstens waren die anderen Programme auch nicht interessanter , und zweitens hatte sich die Waschmittel-Werbung erstaunlich weiter entwickelt .

Corpus analysis: Set activation

- The marked word order seems to be used to give prominence or salience to a referent that is part of a contextually activated set of referents
- Marked order is a construction with a *pragmatic* function:
 - Information structural: drawing attention of the hearer/reader to a salient referent
 - Construal: the salience that is conveyed by the marked order is partially construed subjectively by the speaker/writer
- **Highly schematic** word order patterns function as *pragmatic constructions* in the sense of Goldberg 2004 or Lambrecht 2005: Information structuring

Corpus analysis: relation to less schematic Cx

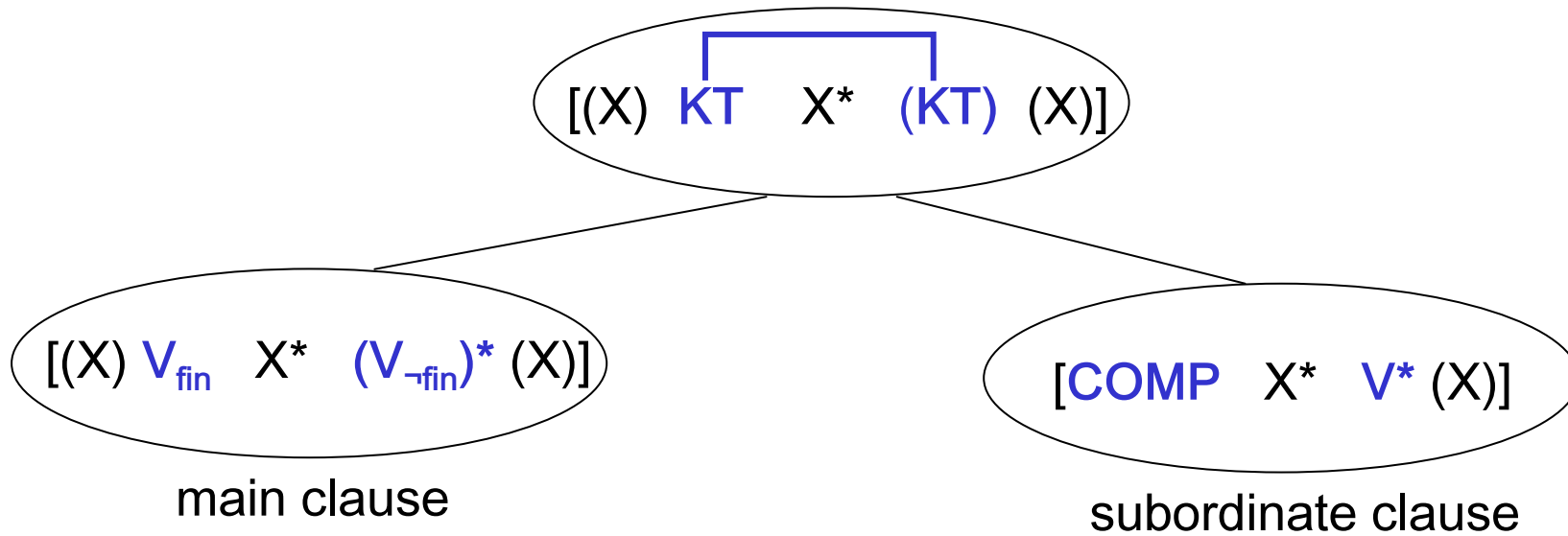
- The pragmatic function of the marked word order (conveying prominence of set member) is realized in **combination with less schematic constructions:**
 - Specific verbs *sich weigern*
 - Concessive '*während*'
 - Focussing particle + NP (*selbst, insbesondere*)
 - *dermassen* ADJ, *dass* NP VP

- Das schon für Montag geplante Treffen Vassilious mit Denktash mußte verschoben werden , weil Denktash sich **weigerte** , eine Einladung zu
- **Während** das Vorstandsmitglied der IG Metall , Peter Kirch, es bei einer sozialpolitischen DGB-Veranstaltung im Main-Kinzig-Kreis als positiv empfand , ... , sieht der Kreisvorsitzende Sepp Sigulla den Kompromiß
- Morgens um 9.00 Uhr spannt die 21jährige ...**Dann** löst ihr Bruder sie für die letzten drei Stunden ab
- Sie entspricht zwar nicht den Vorstellungen , die **insbesondere** die Dritte Welt sich in den siebziger Jahren davon gemacht hat
- Die percussive Version des zu Tode genudelten " Bolero " , fiel **dermaßen** obskur aus , **daß selbst** Ravel sich amüsiert hätte

Corpus analysis: relation to less schematic Cx

- The pragmatic function of the marked word order (conveying prominence of set member) is realized in combination with less schematic constructions:
 - Specific verbs *sich weigern*
 - Concessive 'während'
 - Focussing particle + NP (*selbst, insbesondere*)
 - *dermassen* ADJ, *dass* NP VP
- Could also partially explain the **unexpected effect of clause type**: order not regulated at highest level of schematicity

Klammer-construction



Corpus analysis: relation to less schematic Cx

- The pragmatic function of the marked word order (conveying prominence of set member) is realized in combination with less schematic constructions:
 - Specific verbs *sich weigern*
 - Concessive ‘*während*’
 - Focussing particle + NP (*selbst, insbesondere*)
 - *dermassen* ADJ, *dass* NP VP
- Could also partially explain the unexpected effect of clause type: order not regulated at highest level of schematicity
- Lower level of schematicity also confirmed by **distinctive verb collexeme analysis** (cf. Gries & Stefanowitsch 2004)
Verb forms that cooccur signif. more often with marked order than on average



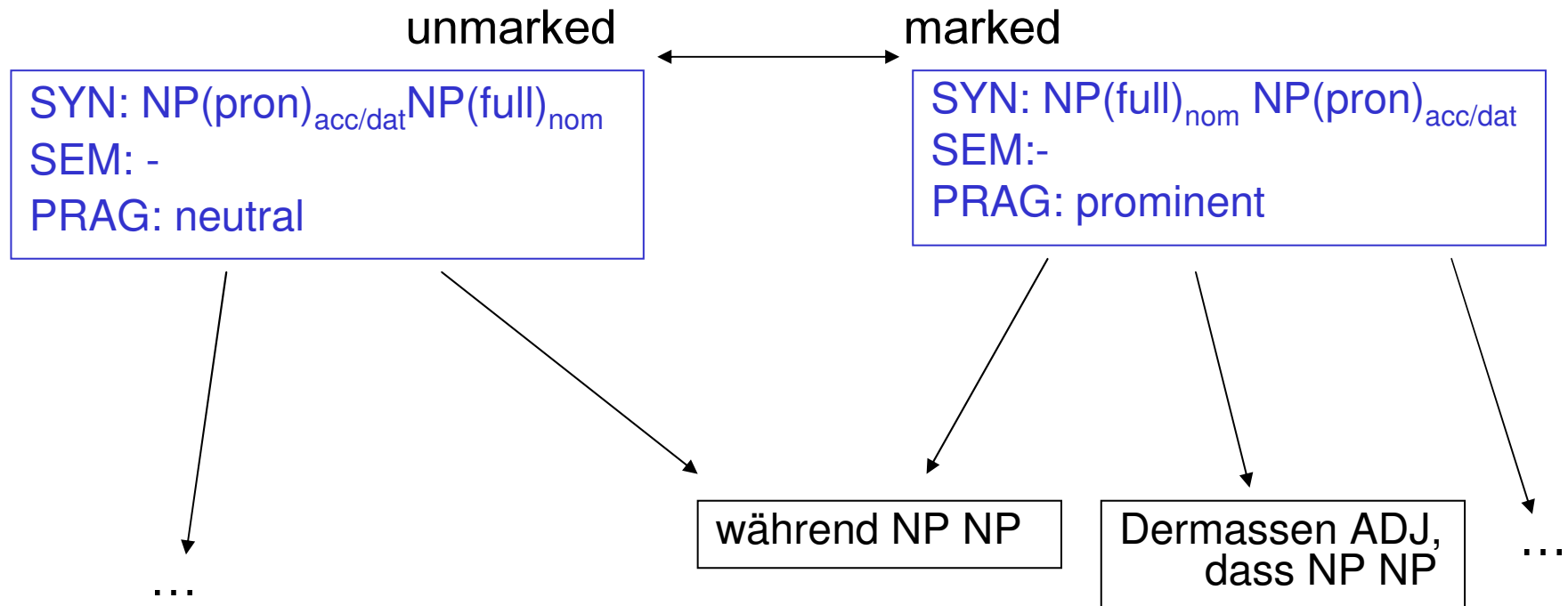
Verbforms ~ subject-first

VERB	FREQ	LLRSTAT	1-P	DIRECTIO	MINEXP
weigert	64	60.94733	1	1	26.68951
weigerte	73	57.94463	1	1	32.8281
weigern	40	30.98813	1	1	18.14887
scheidet	15	28.56955	1	1	4.537217
geweigert	54	24.33906	0.999999	1	29.62536
fuehlt	40	21.24114	0.999996	1	20.81782
ueberlasse	10	20.34415	0.999994	1	2.935847
erklaert	50	20.32326	0.999993	1	28.29089
gezwungen	20	19.36195	0.999989	1	8.27375
trauen	27	18.60719	0.999984	1	12.81097
lieben	17	18.55003	0.999983	1	6.672379
wehren	40	17.46434	0.999971	1	22.1523
bemuehen	35	17.19364	0.999966	1	18.68266
weigerten	36	16.93308	0.999961	1	19.48335
einsetzen	28	16.21783	0.999944	1	14.14544
entschlossen	30	15.51607	0.999918	1	15.74681
fuehlen	48	15.18327	0.999902	1	29.09157
bemueht	36	14.17789	0.999834	1	20.55093
helfen	22	13.31667	0.999737	1	10.9427

Conclusions

- Word order variants can be described as highly schematic constructions

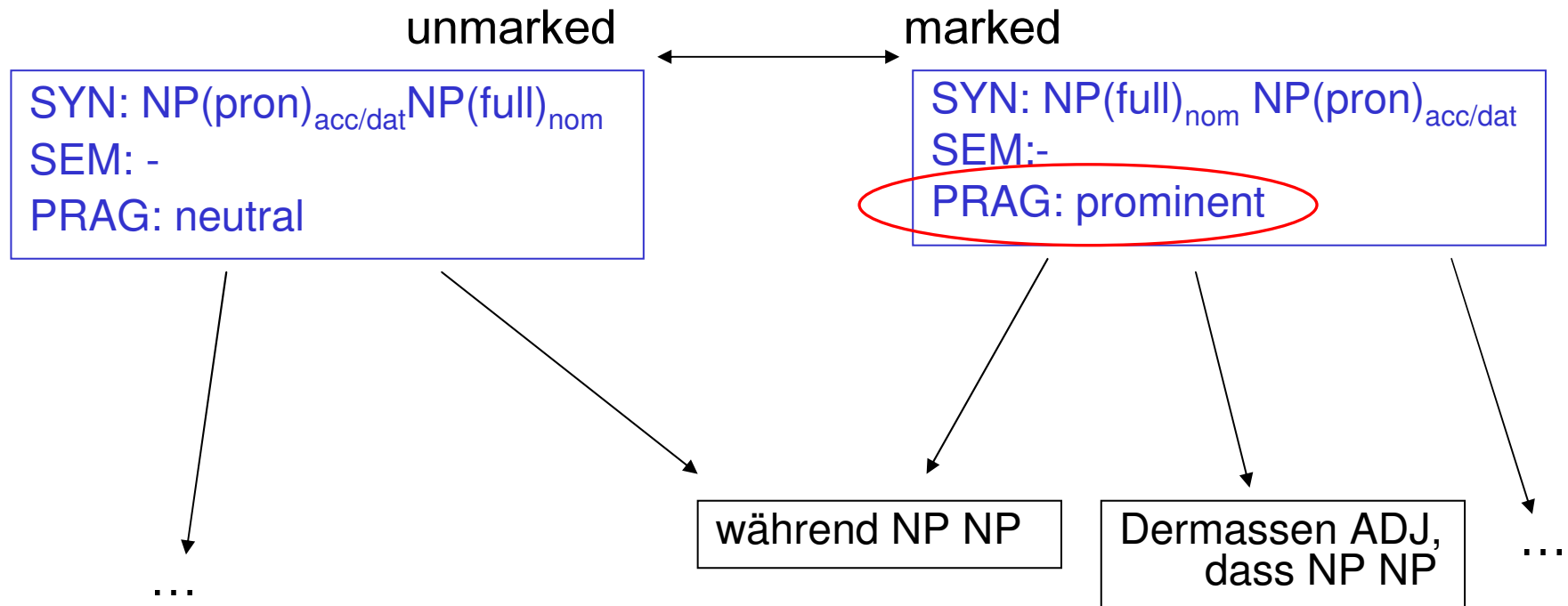
Conclusions



Conclusions

- Word order variants can be described as highly schematic constructions
- The subject-first word order has a pragmatic function, whereas the object-first order is pragmatically neutral

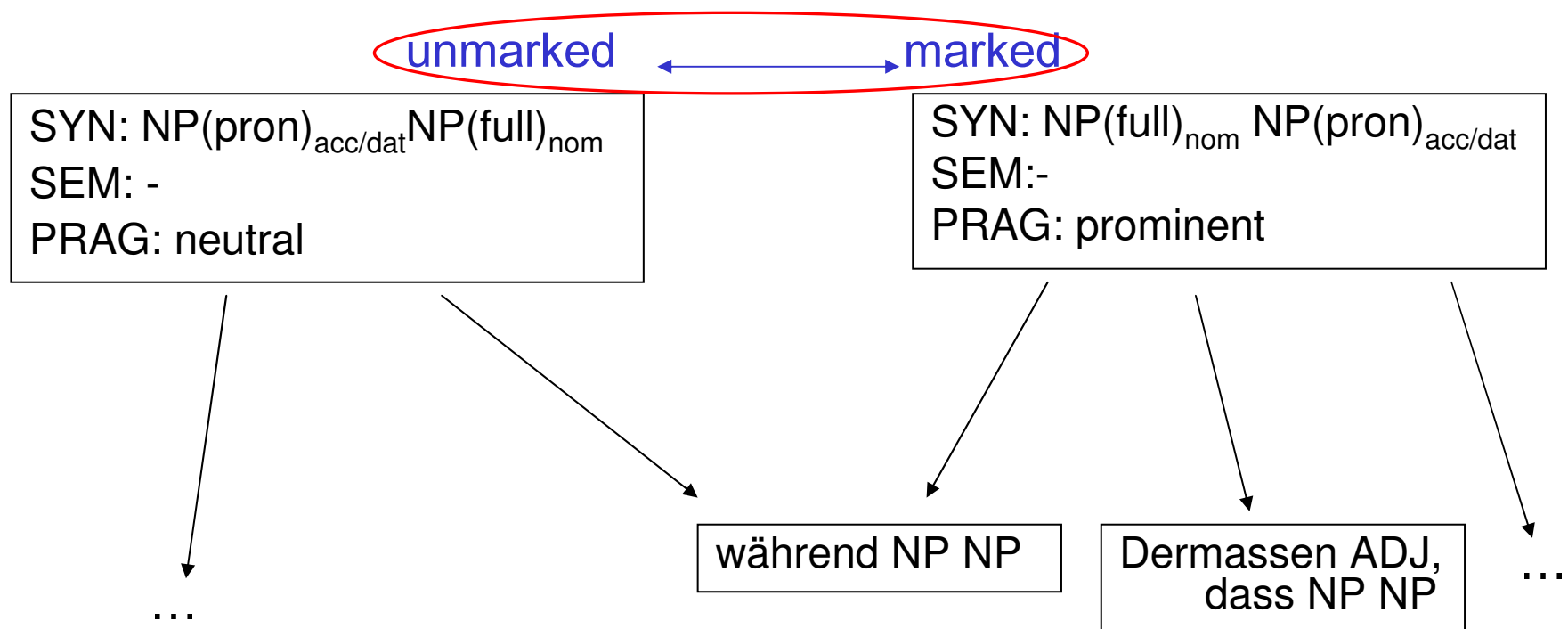
Conclusions



Conclusions

- Word order variants can be described as highly schematic constructions
- The subject-first word order has a pragmatic function whereas the object-first order is pragmatically neutral
- There is a markedness relation between the two word order constructions, which directly correlates with their pragmatic functions

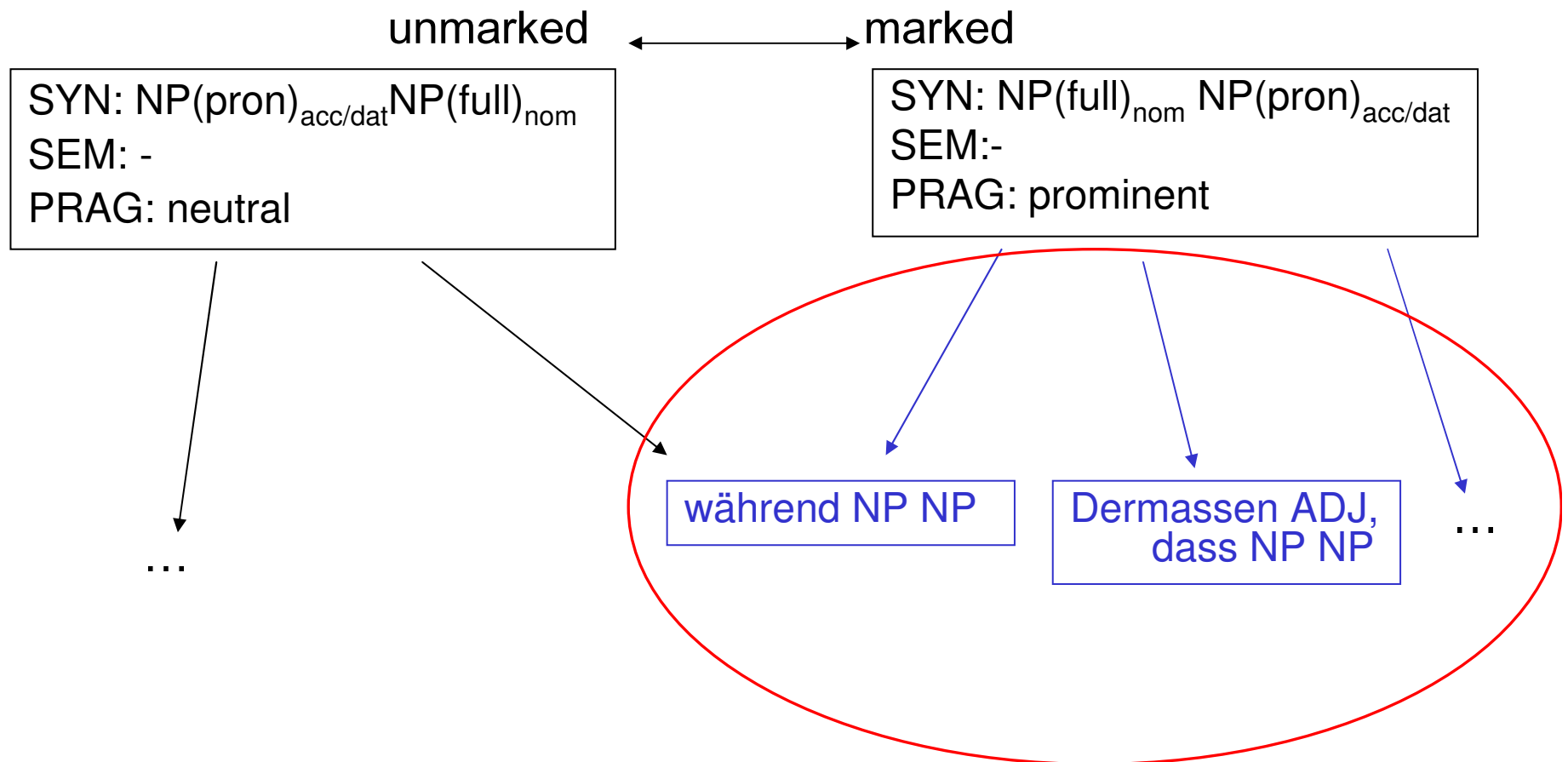
Conclusions



Conclusions

- Word order variants can be described as highly schematic constructions
- The subject-first word order has a pragmatic function, whereas the object-first order is pragmatically neutral
- There is a markedness relation between the two word order constructions, which directly correlates with their pragmatic functions
- The subject-first word order has relations within the construction to less schematic constructions (inheritance/instantiation vs. schematization/projection)

Conclusions





for further information:

kris.heylen@arts.kuleuven.be

<http://www.ling.arts.kuleuven.ac.be/qvl/>