

Erscheint in:

Coniglio, Marco; Anabel Recker; Heike Sahn (i.V.). *Mittelniederdeutsch an der Schnittstelle zwischen Literaturwissenschaft, Sprachwissenschaft und Digital Humanities*. Göttingen: Universitätsverlag.

Hannah Booth, Anne Breitbarth<sup>1</sup> und Melissa Farasyn

## Linke Satzperipherie, Mittelfeld und Nullsubjekte im Mittelniederdeutschen

Zur Untersuchung syntaktischer Phänomene mit dem CHLG

**Zusammenfassung:** Nachdem ihre Erforschung lange vernachlässigt wurde, erfährt die Syntax des Mittelniederdeutschen seit einigen Jahren endlich die Aufmerksamkeit, die sie verdient. Inzwischen stehen auch endlich annotierte elektronische Korpora zur Verfügung, die die weitere Erforschung noch mehr erleichtern werden. Im vorliegenden Beitrag soll eines davon vorgestellt werden, das sich insbesondere durch seine explizite syntaktische Annotation auszeichnet. Anhand ausgewählter Fallstudien soll gezeigt werden, dass eine solche Annotation für die Beantwortung vieler syntaktischer Fragestellungen unerlässlich ist.

**Schlagwörter:** Mittelniederdeutsch, Korpora, Parsing, Integration linksperipherer Nebensätze, Nullsubjekte, Scrambling

### 1 Einleitung

Die Syntax des Mittelniederdeutschen erfährt seit einigen Jahren erfreulicherweise die Aufmerksamkeit, die sie verdient, und die ihr so lange versagt geblieben ist (vgl. Peters 1973). Pioniere auf dem Gebiet ihrer Erforschung sind u. a. Rösler (1997), Mähl (2004, 2014), Tophinke (2009), Petrova (2012, 2013), Wallmeier (2013, 2015) und Breitbarth (2014). In den letzten Jahren sind auch eine Reihe von Dissertationen zur Syntax des 20 Mittelniederdeutschen entstanden (bspw. Farasyn 2018; Merten 2018; Ihden 2020), was für die Zukunft hoffen lässt. Im Rahmen der von der Deutschen Forschungsgemeinschaft geförderten Initiative *Deutsch Diachron Digital* (DDD) wurden in den letzten Jahren verschiedene große Referenzkorpora zu historischen Sprachstufen des Deutschen angelegt, die inzwischen auch alle öffentlich zugänglich sind. Auch für das Mittelniederdeutsche steht seit 2019 das Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200-1650) (ReN) (ReN-Team 2017) zur Verfügung. Seit Ende 2020 ist auch ein syntaktisch annotiertes Subkorpus des ReN online verfügbar, das Corpus of Historical Low German (CHLG; [www.chlg.ugent.be](http://www.chlg.ugent.be)). Dieses soll im vorliegenden Beitrag

---

<sup>1</sup> Korrespondenzautor: Anne Breitbarth, Vakgroep Taalkunde, Universiteit Gent, Blandijnberg 2, B-9000 Gent, Belgien, [anne.breitbarth@ugent.be](mailto:anne.breitbarth@ugent.be)  
Hannah Booth, Fachbereich Linguistik, Universität Konstanz, Fach 185, D-78457 Konstanz, Deutschland, [hannah.booth@uni-konstanz.de](mailto:hannah.booth@uni-konstanz.de)  
Melissa Farasyn, Vakgroep Taalkunde, Universiteit Gent, Blandijnberg 2, B-9000 Gent, Belgien, [melissa.farasyn@ugent.be](mailto:melissa.farasyn@ugent.be)

kurz vorgestellt und seine Besonderheiten im Vergleich zu anderen Korpora anhand ausgewählter syntaktischer Fallstudien illustriert werden.

## **2 Das CHLG**

### **2.1 Hintergrund**

Die beiden Projekte für ein Korpus des Mittelniederdeutschen, das ReN und das CHLG, wurden zunächst unabhängig voneinander konzipiert und beantragt. Als für beide etwa zeitgleich Ende 2013 eine finanzielle Förderung bewilligt wurde, im Falle des ReN durch die DFG, im Falle des CHLG durch die belgische Hercules-Stiftung (die 2016 in das neue Forschungsinfrastrukturprogramm des Fonds voor Wetenschappelijk Onderzoek-Vlaanderen (FWO) eingegliedert wurde), wurde auf Initiative von Robert Peters beschlossen, die Ressourcen zu bündeln. Die Mitarbeiter des CHLG halfen bei der Annotation der Wortarten (POS-tagging) und der Morphologie einiger Texte des ReN; im Gegenzug durfte das CHLG ausgewählte Texte des ReN syntaktisch annotieren.

Das ReN ist Teil des „Korpus historischer Texte des Deutschen“ der Initiative Deutsch Diachron Digital (DDD), zu dem auch die Referenzkorpora Altdeutsch (REA; <https://www.deutschdiachrondigital.de/>), Mittelhochdeutsch (REM; <http://www.linguistics.rub.de/rem/>) und Frühneuhochdeutsch (REF; <http://www.linguistics.rub.de/ref/>) zählen. Es bietet Zugang zu ca. 1,45 Millionen Wortformen in 145 diplomatisch transkribierten, lemmatisierten und grammatisch annotierten Texten. Das CHLG bietet zusätzlich zur morphologischen Annotation des ReN für achtzehn der Texte (ca. 187.500 Wortformen) eine syntaktische Annotation nach dem Schema der Penn Parsed Corpora of Historical English (<https://www.ling.upenn.edu/hist-corpora/>), vgl. Booth u. a. (2020). Es bietet somit auch eine mittelniederdeutsche Fortsetzung des HeliPaD-Korpus (Heliand Parsed Database, Walkden 2015, 2016), das auf Sievers' (1878) Edition der Handschrift Heliand C beruht. Beide, das CHLG und das HeliPaD, sind lemmatisiert und verwenden dasselbe Parsingschema, unterscheiden sich jedoch in der Annotation der Wortarten (POS-Tagging) und Morphologie, wie im Folgenden ausgeführt wird.

## 2.2 Aufbau und Annotation des CHLG

Da syntaktische Annotation historischer Texte sehr arbeitsaufwändig ist, konnte nur ein Teil der Texte des ReN für das CHLG geparst werden. Um dennoch eine ausgewogene Auswahl zu erreichen, wurde entschieden, das CHLG auf die Schreibsprachen Westfälisch, Ostfälisch, Nordniedersächsisch und Ostelbisch zu beschränken, und baltische, südmarkische sowie niederrheinische Texte auszuschließen. Aus diesen vier Schreibsprachen wurden vier bzw. fünf Texte für das CHLG geparst, vgl. Tab. 1. Wie auch im ReN besteht ein gewisses Ungleichgewicht in der Verteilung von Textsorten über den Zeitraum, was der Überlieferungslage geschuldet ist: typischerweise sind Rechts- und Urkundentexte früher und über den gesamten Zeitraum belegt, während andere Textsorten erst im 15. Jahrhundert hinzukommen.

Text	Zeitraum	Textsorte
Westfälisch		
Abdinghofer Arzneibuch	1451-1500	Wissenschaft
Herforder Rechtsbuch	1375	Recht
Soest, Schrae im Statutenbuch	1367	Recht
Rüthen, Statutarrecht	1300	Recht
Spiegel der Leyen	1444	Religion
Ostfälisch		
Braunschweiger Urkunden	1301-1500	Recht/Urkunden
Duderstadt, Braunschweiger Stadtrecht	1279	Recht
Engelhus, Chronica nova	1435	Literatur
Helmstedt, Zeno	1401-1450	Literatur
Nordniedersächsisch		
Agneta Willeken	1535	Privatbrief
Bremer Urkunden	1300-1350	Recht/Urkunden
Buxtehuder Evangeliar	1451-1500	Religion
Griseldis	1502	Literatur
Oldenburger Urkunden	1350-1500	Recht/Urkunden
Ostelbisch		
Flos unde Blankeflos	1401-1450	Literatur
Greifswalder Bürgersprache	1451	Recht
Rostocker Bürgersprache	1580	Recht
Schweriner Stadtbuch	1451-1500	Recht
Stralsunder Urkunden	1301-1500	Recht/Urkunden

Tab. 1: Textauswahl des CHLG

Aufgrund der Zusammenarbeit mit dem ReN benutzt das CHLG nicht das von den anderen

historischen Korpora der Penn-Familie gebrauchte POS-Tagset (Santorini 1990). Stattdessen setzt die syntaktische Annotation, die bis auf kleine Anpassungen dem Penn-Standard folgt, auf das HiNTS-Tagset des ReN auf,<sup>2</sup> das viel feinere Unterscheidungen zulässt und bereits eine ganze Reihe syntaktischer Merkmale mit einschließt. So unterscheidet das HiNTS zwischen 26 POS-Tags für determinative Elemente, wo das Penn-Tagset nur ein Tag D hat. Diese 26 Tags beginnen alle mit dem Grundlabel D, werden dann jedoch weiter nach verschiedenen Merkmalen ausdifferenziert, wie in Tab. 2 zu sehen ist.

Merkmal	möglicher Wert	Label
Art des Determinativs	definit	D
	indefinit	I
	negativ	NEG
	possessiv	POS
	relativisch	REL
	interrogativ	W
Artikelartigkeit	artikelartig	ART
	nicht artikelartig	—
Position	attributiv	A
	nachgestellt	N
	substituierend	S
	prädikativ	D

Tab. 2: Kodierung verschiedener Merkmale von Determinativen im HiNTS

Somit wird ein definites, artikelartiges Determinativ, das seinem Bezugsnomen vorangeht, wie z.B. *dat* in *dat land* im HiNTS als DDARTA getagt. Ein Determinativ, das negativ, nicht artikelartig und prädikativ ist (also als Kopf eines nominalen Prädikats fungiert), wie z.B. *niemand* in *dat ist niemand* wird als DNEGD getagt. Dieser Annotation der Wortart (sowie der Flexionsmorphologie und der Lemmata, die auch im ReN zur Verfügung stehen) fügt das CHLG nun eine eigene Ebene der syntaktischen Annotation zu. Dabei wurde entschieden, dem Penn-Annotationsstandard zu folgen. Dabei handelt es sich um einen konstituenzbasierten Standard, im Unterschied zu dependenzbasierten Standards wie den Universal Dependencies (Nivre u. a. 2016). Solche dependenzbasierten Ansätze wurden auch bereits für historische Korpora eingesetzt, bspw. für Familie der Pragmatic Resources in Old Indo-European Languages (PROIEL) (Haug und Jøhndal 2008). Beide Systeme sind gleich mächtig und prinzipiell isomorph (Haug 2015; Taylor 2020). Es ist wichtig, im Auge zu behalten, dass Baumbanken

<sup>2</sup> HiNTS steht für Historisches Niederdeutsch-Tagset (Barteld u. a. 2018). Es handelt sich um eine Weiterentwicklung des Historischen Tagsets (HiTS) (Dipper u. a. 2013), das wiederum auf dem Stuttgart- Tübingen Tagset (STTS) (Schiller u. a. 1999) aufbaut.

Werkzeuge für eine möglichst präzise und effiziente Extraktion von Daten sind, keine exakte sprachwissenschaftliche Analyse.<sup>3</sup> Die Wahl des Penn-Systems im Falle des CHLG ist hauptsächlich durch die Kompatibilität mit den zahlreichen anderen historischen Korpora motiviert, wodurch dieselben Suchanfragen benutzt werden können, um mehrere Korpora zu durchsuchen. International ist das System auch bekannter und somit zugänglicher als andere Annotationsschemata wie das für das Fnhd. gebrauchte. Zudem können Werkzeuge wie das Suchprogramm CorpusSearch (Randall 2005) oder das GUI für die syntaktische Annotation Annotald (Beck, Ecaj und Ingason 2015) ohne weitere Modifikationen verwendet werden. Für linguistische Argumente für die Verwendung des Penn-Systems verweisen wir auf Booth u. a. (2020).

### **2.3 Wozu ein geparstes Korpus?**

Historische Syntaxforschung war lange eine mühsame Disziplin vor allem durch die aufwändige Sammlung von Belegen aus Handschriften, Drucken oder Ausgaben. Elektronische Textsammlungen brachten nur bedingt Erleichterung, da sich syntaktische Phänomene einfachen Volltextsuchen entziehen, sobald sie über die Abfolge einzelner Wörter hinausgehen und sich auf die Abfolge oder Interaktion mehr oder weniger umfangreicher Satzglieder beziehen. Die Entwicklung syntaktisch annotierter Korpora in den letzten Jahrzehnten hat eine entscheidende Rolle dabei gespielt, quantitative Untersuchungen syntaktischer Veränderungen in vielen Sprachen zu erleichtern und reproduzierbare und zuverlässige Untersuchungsergebnisse zu ermöglichen. Pintzuk, Taylor und Warner (2017) beispielsweise bieten einen Überblick über diese Entwicklungen.

Wie in Abschnitt 2.2 bereits angesprochen, enthalten einige POS-Tags im HiNTS bereits gewisse syntaktische Informationen. Im oben diskutierten Beispiel der Determinative, aber auch bei anderen Wortarten, ist ein Teil des Tags für die Kodierung der Position des getagten Tokens hinsichtlich eines Bezugswortes (vorangestellt/attributiv, nachgestellt) reserviert. Somit ist im ReN allein schon auf der Ebene der Wortartannotation eine Bearbeitung gewisser syntaktischer Fragestellungen möglich. Die Anis-Suchanfragesprache erlaubt auch die Arbeit mit Abstandsoperatoren, um Wortformen, Lemmata, Wortart- oder morphologische Annotationen von Tokens im Kontext anderer solcher Elemente zu suchen. Zusätzlich ermöglicht die Annotation von Satzgrenzen (bound-sent) im ReN – eine sehr willkommene Innovation

---

<sup>3</sup> So werden im Penn-System beispielsweise weder VP-Knoten noch andere funktionale Projektionen im Mittelfeld annotiert; alle unmittelbaren Konstituenten eines Satzes sind direkte Töchter des Satzknotens. Dadurch können Konstituenten (Verben, Argumente und Ergänzungen) schnell und eindeutig über ihre Dominanz- und Präzedenzrelationen zueinander gefunden werden können.

gegenüber den anderen DDD-Korpora! – die Einschränkung einer Suche auf Einfachsätze. Syntax ist jedoch mehr als die direkte Adjazenz bzw. konkrete Abstände zwischen einzelnen Wörtern. Dem ReN wie auch den anderen DDD-Korpora fehlt die Möglichkeit, Konstituenten beliebigen Umfangs in Beziehung zueinander zu suchen. Die syntaktische Funktion nominaler Konstituenten kann eventuell über die morphologische Annotation der Kasusmarkierung angenähert werden, doch auch hier besteht keine Möglichkeit, mit Konstituenten unterschiedlichen Umfangs umzugehen. Im Folgenden sollen beispielhaft drei Phänomene besprochen werden, die einen Zugang zu linearen, hierarchischen und funktionalen Beziehungen zwischen Konstituenten unterschiedlichen Umfangs erfordern, und deren Untersuchung daher in entscheidender Weise von der konstituentenstrukturellen Annotation des CHLG Gebrauch machen kann.

### **3 Mehrfache XP-Voranstellung in Sätzen mit Verbbewegung**

Petrova (2012) untersucht an ausgewählten Texten die verschiedenen Möglichkeiten der Späterstellung des finiten Verbs im unabhängigen Satz im Mnd. Neben Abfolgen, die mit einer V2-Grammatik vereinbar sind und auch im Gegenwartsdeutschen möglich wären (z.B. Linksversetzungen, Freies Thema, scheinbare mehrfache Vorfelddbesetzung) bespricht sie sie auch Abfolgen, die nicht mit einer solchen Grammatik vereinbar sind. Dabei formuliert sie diagnostische Kriterien für die Feststellung von Verbbewegung in die linke Satzperipherie, um (unabhängige) Sätze ohne Verbbewegung, die in ihren Daten auch belegt sind, von solchen mit mehrfacher Vorfelddbesetzung zu unterscheiden. Im Wesentlichen identifiziert Petrova zwei größere Gruppen von V2-Verletzungen, einerseits Abfolgen, bei denen das Subjekt einer Adverbialbestimmung vorangeht (1a-1b) und andererseits solche, bei denen eine Adverbialbestimmung dem Subjekt vorangeht (1c-1e). Zudem scheinen direkte Objekte Subjekten im Vorfeld vorangehen zu können, (1f), allerdings ist nicht bei allen Belegen klar, ob Verbbewegung stattgefunden hat.

#### 1) a. Subjekt - temporale PP

[Constantinus] [des morgenes vro] segede dit sinen vrunden (SW 118, 24)

‘Des Morgens früh sagte Constantinus dies zu seinen Freunden.’

(Petrova 2012: 173)

#### b. Subjekt - kausale PP

[Kato de wise] [vor leide] dodede sic selven (SW 87, 27)

‘Cato der Weise tötete sich selbst aus Leid.’

(Petrova 2012: 173)

c. temporale Adverbialbestimmung – Subjekt

[Nicht langhe darna] [der sone] volghete na vor ludeke (LChr 112,3)

‘Nicht lange danach folgte der Sohn nach Richtung Lübeck.’

(Petrova 2012: 174)

d. temporaler Adverbialsatz - pronominales Subjekt

[Do Allexander vornam ere missedat], [he] let se bemuren (SW 78, 4)

‘Als Alexander ihre Missetat vernahm, ließ er sie einmauern.’

(Petrova 2012: 174)

e. temporaler Adverbialsatz - DP-Subjekt

[Do he starf], [de Romere] clageden in sere (SW 90,17) ‘Als er starb, beklagten die Römer ihn sehr.’

(Petrova 2012: 174)

f. direktes Objekt - Subjekt

[Eine warheit] [ich] wille dir sagen (SW 100, 29)

‘Eine Wahrheit möchte ich dir sagen.’

(Petrova 2012: 175)

Da Petrova nur die absoluten Zahlen der gefundenen Belege angibt<sup>4</sup> und auch nur eine eingeschränkte Textauswahl untersucht – vier der sieben Texte bzw. Textsammlungen sind Chroniken und einer ein Reisebericht<sup>5</sup> – stellt sich die Frage, wie der Befund in einem ausgewogeneren Korpus aussähe, und wie häufig solche V2-Verletzungen überhaupt in unabhängigen Sätzen vorkommen. Eine solche Frage stellt eine Herausforderung für das ReN dar. Zwar ist es möglich, durch Bezug auf die Satzspanne (bound-sent) nach Phänomenen innerhalb eines Einfachsatzes oder Gliedsatzes zu suchen, allerdings können nur individuelle Tokens, ihre Eigenschaften (Lemma, Wortart, morphologische Eigenschaften) und ihr Abstand zu anderen Tokens gesucht werden. Dadurch würde man verschiedene Suchanfragen benötigen, um nach Abfolgen wie denen in (1) zu suchen. Je nachdem, ob ein Subjekt oder Objekt intern komplex ist (volle DP wie in (1c), evtl. noch mit Attributen erweitert, vs. pronominal wie in (1d)), oder eine adverbiale PP mehr als nur Präposition und Nomen enthält, müsste bspw. mit verschiedenen Abstandsangaben bzw. Abstandsbereichen zwischen Tokens gearbeitet werden.

---

<sup>4</sup> Relative Belegzahlen finden sich nur in Donhauser und Petrova (2009: 299), wo die Anzahl der uninvertierten Nachsätze nach vorangestellten Temporalsätzen in zwei mnd. Texten, die auch Teil von Petrovas (2012) Korpus ausmachen, nämlich in der ‚Sächsischen Weltchronik‘ (81,3% V3 nach Temporalsätzen) und ‚Ludolfs von Sudheim Reise ins Heilige Land‘ (4,62%).

<sup>5</sup> Hinzu kommen eine Predigtsammlung sowie eine Sammlung von Rechtstexten aus Stadtbüchern des 14.-16. Jhs.

Die Gefahr, hierbei Belege zu übersehen, liegt auf der Hand. Für Abfolgen mit linksperipheren Adverbialsätzen wie (1e) ist eine Annis-Anfrage eigentlich gar nicht zu formulieren, da zwei Satzspannen nicht innerhalb einer Anfrage miteinander in Beziehung gesetzt werden können.

Mit einer konstituentenbasierten syntaktischen Annotation wie der des CHLG ist hingegen eine einzige Suchanfrage ausreichend. Man beschränkt die Suche auf unabhängige Sätze (node = IP-MAT\*), spezifiziert, dass die dritte Tochter dieses Satzknottens ein finites Verb sein soll ((IP-MAT\* iDomsNum 3 V\*FIN)), und stellt über weitere Spezifikationen sicher, dass die Anzahl der falschen Positive möglichst gering ausfällt.<sup>6</sup> Da CorpusSearch auch immer die Gesamtzahl der Belege in der spezifizierten Umgebung (hier node = IP-MAT\*) angibt, bekommt man zudem schnell einen Überblick über die Häufigkeit eines Phänomens.

Im CHLG finden sich alle von Petrova beschriebenen Abfolgen, (2).<sup>7</sup>

## 2) a. Subjekt-temporale PP

[De engel godes] [na der tijd] stech neder an den dik  
(Buxtehuder Evangeliar, 1451-1500)

‘Nach der Zeit stieg der Engel Gottes nieder an den Deich

## b. Subjekt – temporaler Adverbialsatz

[Enoch] [alse he was lxx iar olt] ghewan Matuselem (Engelhus, Chronica nova, 1435)  
‘Enoch, als er 65 Jahre alt war, gewann Matuselem.’

## c. temporale Adverbialbestimmung - Subjekt

[Dan] [de richtere] sal den man kuden de in dem gherichte bescreghen ys to 5  
dreynevyrteyn nachten. (Rüthener Statutarrecht, 1300)  
‘Dann soll der Richter den Mann künden, der in dem Gericht zu dreimal 14 Nächten  
angeklagt ist.’

---

<sup>6</sup> So muss man bspw. spezifizieren, dass die erste Konstituente keine koordinierende Konjunktion, META-Tags, die elliptische Konjunktionen über Indizes verbinden, oder textstrukturierende Zeichen (z.B. <I) enthält ((IP-MAT\* iDomsNum 1 ! KON|META|CP-ADV\*)), und es sich bei der zweiten Konstituente nicht etwa um die präverbale Negationspartikel ne/en handelt ((IP-MAT\* iDomsNum 2 ! PTKNEG)), die als Klitik/Affix für die Berechnung der Verbposition keine Rolle spielt (vgl. Breitbarth 2014). Da Linksversetzungsstrukturen nicht als Verletzungen der V2-Regel aufgefasst werden, wurden Linksversetzungen (1a), von einem Resumptivum aufgenommene linksperiphere Adverbialsätze - oftmals korrelative Strukturen (1b-1c) - manuell aussortiert.

- (1) a. [Alle deghene de to dussen iarmarkedem comet in de stat to Ruden] de solen hebben syker gheleyde. (Rüthener Statutarrecht, WF, 14.1)  
‘Alle diejenigen, die zu diesen Jahrmärkten kommen in der Stadt zu Rüthen, die sollen sicheres Geleit haben.’  
b. [Do dat kint vorstarf] do wart der moder mit rechte syn erue toghedelet. (Herford, WF, 14.2) ‘Als das Kind verstarb, da wurde der Mutter mit Recht sein Erbe zugeteilt.’  
c. [also my myn vader sant heft] also sende ik iu (Buxtehuder Evangeliar, NNS, 15.2) ‘Wie mich mein Vater gesandt hat, so sende ich euch.’  
d. [Swe den Dotslagh eder de wunnynghede doyt.] De ne scal dar nerghene vrede hebben (Oldenburger Urkunden, NNS, 14.2-15)  
‘Wer immer den Totschlag oder die Verwundung tut, der soll dort nirgends Frieden haben.’

Freie Relativsätze, die sehr häufig in Linksversetzungsstrukturen vorkommen, die auch im Gegenwartsdeutschen möglich wären (1d), wurden von vornherein in der Suchanfrage ausgeschlossen ((IP-MAT\* Doms !CP-FRL\*)).

<sup>7</sup> Aufgrund stark abweichender Wortstellungsmuster (z.B. Häufung von Verbendstellung) wurden gereimte Texte (Flos unde Blankeflos, Spiegel der Leyen, Zeno) allerdings von dieser Untersuchung ausgeschlossen.



d. temporaler Adverbialsatz - pronominales Subjekt

[Do dat geruchte vor Griseldis quam] [se] was trurich als yk meyne (Griseldis, 1502)

‘Als das Gerücht vor Griseldis kam, war sie traurig, wie ich meine.’

e. temporaler Adverbialsatz - DP-Subjekt

[DO Albert stur starf] [sin eldeste sone] nam sin herwede. (Herforder Rechtsbuch, 1375).

‘Als Albert Stur starb, nahm sein ältester Sohn sein Heergewäte.’

f. direktes Objekt - Subjekt

[Der moder rade] [de paepe] nympt ghelike dele der suster in der moder rade (Herforder Rechtsbuch, 1375).

‘(was) die Gerade der Mutter (betrifft), der Pape nimmt gleiche Teile der Gerade 20 der Schwester und der Mutter.’

Diese verteilen sich nicht ganz gleichmäßig über die Texte, wie aus Tab. 3 ersichtlich wird, in der eine grobe Unterteilung in Rechtstexte i.w.S. und nicht-Rechtstexte vorgenommen wird.

Eine Besonderheit v.a. der mnd. Rechtssprache sind vorangestellte Konditionalsätze, bei denen eine Konstituente des Konditionalsatzes von einem resumptiven d-Pronomen wie in einer Linksversetzungsstruktur wieder aufgenommen wird (3). Man könnte solche Fälle als „Teilresumption“ bezeichnen. Es kann sich dabei um ein Subjekt oder ein Objekt handeln.

3) a. Teilresumption – Subjekt

[Wert iemandt dar ouer beargwonet] die schal sick mit seinem Eide Purgiren Bei Straffe des Rades (Rostocker Bürgersprache 1580)

‘Wird jemand darüber verdächtigt, der soll sich mit seinem Eid für unschuldig erklären bei Strafe des Rates.’

b. Teilresumption – Objekt

[Were th dath Ienich tymmerholt tho schepe queme vor de Stadt edder an de hauene] dat schall nen vorkoper kopen bynnen eneme daghe vnd ener nacht (Greifswalder Bürgersprache 1451)

‘Wäre es, dass jegliches Zimmerholz zu Schiff käme vor die Stadt oder an den Hafen, das soll kein Verkäufer kaufen innerhalb eines Tages und einer Nacht.’

Solche Konstruktionen wurden vorläufig, wie auch bei Donhauser/Petrova (2009), mitgezählt, ebenso wie Konstruktionen, in denen eine Konstituente des vorangestellten Adverbialsatzes von einem Personalpronomen wiederaufgenommen werden, ähnlich wie bei einem Freien Thema

(vgl. (2d)). In Tab. 3 werden diese Belege jedoch in Klammern angegeben, um sie ggfs. von anderen mehrfach besetzten Vorfeldern trennen zu können.<sup>8</sup>

Text	Mehrfache XP-Voranstellung			Total
	Su-XP-Vf	XP-Su-Vf	XP-XP-Vf	
Rechtstexte u. Urkunden				
Braunschweiger Stadtrecht 1279	0	0	0	132
Braunschweiger Urkunden 1301-1500	0	0	0	110
Rüthener Statutarrecht 1300	0	0(5)	0(5)	265
Soester Schrae 1367	0	1	1	262
Herforder Rechtsbuch 1375	1	2(2)	0(6)	830
Bremer Urkunden 1300-1350	0	0	0	48
Oldenburger Urkunden 1300-1530	0	0	0	231
Stralsunder Urkunden 1301-1500	1	0	0(2)	314
Greifswalder Bürgersprache 1451	0	0(6)	0(3)	148
Schweriner Stadtbuch 1451-1500	0	0	0(1)	327
Rostocker Bürgersprache 1580	0	0	0	44
<b>Gesamt</b>	<b>2</b>	<b>3(+13)</b>	<b>1(+17)</b>	<b>2711</b>
Literatur/Religion/Wissenschaft/Privatbrief				
Engelhus, Chronica 1435	15	0(2)	2	1532
Abdinghofer Arzneibuch 1451-1500	3	1(1)	0	1770
Buxtehuder Evangeliar 1451-1500	2	0(2)	0(2)	2038
Griseldis 1502	0	2	0	537
Agneta Willeken 1535	0	0(1)	0	188
<b>Gesamt</b>	<b>20</b>	<b>3(+6)</b>	<b>2(+2)</b>	<b>6065</b>

Tab. 3: Mehrfache XP-Voranstellung in den Texten des CHLG (in Klammern: Teilresumption)

Allgemein lässt sich anhand der relativen Zahlen feststellen, dass mehrfache XP-Voranstellung nur sehr selten vorkommt. Allerdings fallen zwei Dinge ins Auge. Erstens ist mehrfache XP-Voranstellung mit 0,5% aller unabhängiger Deklarativsätze in nicht-Rechtstexten weniger häufig belegt als in Rechtstexten und Urkunden (8,7%). Das Bild verändert sich jedoch, wenn die v.a. in Rechtstexten häufigen Teilresumptionen nach vorangestellten Adverbialsätzen nicht

<sup>8</sup> In Breitbarth (2018) wurden Argumente dafür angeführt, solche Strukturen als weiteres Zwischenstadium zu dem Weg zur Entwicklung von V2, zusätzlich zu Resumption (für das Mnd. vgl. Tophinke 2009), zu interpretieren.

mitgezählt werden: dann ist mehrfache XP-Voranstellung mit 0,4% doppelt so häufig in nicht-Rechtstexten belegt wie in Rechtstexten und Urkunden (0,2%). Zweitens ist die Art der mehrfachen XP-Voranstellung deutlich asymmetrisch verteilt. Abgesehen von den Teilresumptionsstrukturen ist die Abfolge XP-Su-Vf wie in (2d) und (2e), die für die Gegenwartssprache vor allem für das Kiezdeutsche (Wiese, Öncü und Bracker 2017; Wiese und Müller 2018; Wiese u. a. 2020), zunehmend jedoch auch für standardnahes gesprochenes Deutsch beschrieben wurde (Schalowski 2015; Wiese und Müller 2018; Breitbarth i.Dr.), im Mnd. extrem selten (insgesamt 0,06%). Falls es hier also eine historische Kontinuität geben sollte, läge sie bestenfalls im Bereich der gesprochenen Sprache und somit außerhalb des Zugriffs der historischen Korpuslinguistik.<sup>9</sup>

Die im Gegenwartsdeutschen nicht mögliche Abfolge Su-XP-Vf ist viel stärker in nicht-Rechtstexten vertreten, und dort v.a. in einem Text, nämlich der Chronik von Engelhus (wo sie allerdings auch nur in 1% der unabhängigen Sätze vorkommt). Das ist vor allem deshalb interessant, weil Petrova (2012) v.a. Chroniktexte untersucht. Diese Konstruktion verbindet eindeutig eine Art Freies Thema mit einem Rahmensetzer: Ein Subjekt wird eingeführt (was X betrifft), dann folgt ein V2-Satz mit bspw. Rahmensetzer im Vorfeld (2a-2b).

Zusammenfassend kann festgestellt werden, dass die konstituentenbasierte syntaktische Annotation des CHLG eine schnelle und zuverlässige Überprüfung, Aktualisierung und Präzisierung älterer Forschungsergebnisse zur mehrfachen Vorfeldbesetzung erlaubt, sowie die Annäherung an neue Fragen wie bspw. zum Alter und der historischen (Dis)Kontinuität bestimmter mehrfacher Vorfeldbesetzungen ermöglicht.

## 4 Spiegeleffekte im Mittel- und Nachfeld

Sogenannte ‘Spiegeleffekte’, bei denen die im Mittelfeld (MF) bevorzugte Wortstellung im Nachfeld (NF) in umgekehrter Abfolge erscheint, sind in verschiedenen modernen westgermanischen Sprachen nachgewiesen worden (Koster 1974, 2001; Barbiers 1995; Haider 2000; Pittner 2004). Im modernen Niederländischen manifestiert sich dieser Effekt in der Wortstellung von PP-Argumenten und PP-Adjunkten: die bevorzugte Wortstellung im MF ist Adjunkt-Argument (4a), und im NF Argument-Adjunkt (4b).

---

<sup>9</sup> Die Tatsache, dass im Gegenwartsdeutschen (innerhalb ebenso wie außerhalb des Kiezdeutschen) keine Teilresumtionen belegt sind, sowie weiterhin die Tatsache, dass Inversion nach vorangestellten Adverbialbestimmungen im Mnd. kaum belegt ist – Donhauser & Petrova (2009:299) finden 0,93% in der ‚Sächsischen Weltchronik‘ und 1,54% in ‚Ludolf‘; in den hier untersuchten Texten im CHLG sind es ca. 7% aller Sätze mit vorangestellten Adverbialbestimmungen, allerdings stammen zwei Drittel dieser Belege (8/12) aus einem einzigen Text (Griseldis) – weist darauf hin, dass die mehrfachen Vorfeldbesetzungen im Mnd, inklusive der Abfolge XP-Su-Vf, eher darauf zurückzuführen sind, dass die V2-Stellung nach vorangestellten Adverbialbestimmungen noch nicht vollständig etabliert war.

- 4) a. Hij heeft [tijdens de pauze] [aan zijn vader] gedacht  
 b. Hij heeft gedacht [aan zijn vader] [tijdens de pauze]  
 ‘Er hat während der Pause an seinen Vater gedacht’ (Koster 2001: 1)

Ein ähnlicher Effekt ist in Bezug auf die Abfolge mehrerer semantisch unterschiedlichen Adjunkte belegt, wie Haider (2000) für verschiedene westgermanische Sprachen zeigt, vgl. (4) (adaptiert aus Haider 2000: 96). Für das Neuhochdeutsche insbesondere gilt 5 die Wortstellung in (5a) für das MF und die invertierte Wortstellung in (5b) für das NF (siehe auch Pittner 2004).

- 5) a. Zeit < Ort < Art&Weise - V  
 b. V - Art&Weise < Ort < Zeit

Das Niederdeutsche befindet sich im kontinentalwestgermanischen Dialektkontinuum typologisch und geographisch gesehen zwischen dem Niederländischen und dem Hochdeutschen. Aufgrund dessen würde man erwarten, dass sich solche Spiegeleffekte auch im Niederdeutschen identifizieren lassen. Das ist bisher allerdings noch nicht untersucht, weder synchron noch diachron. Das Mnd. bietet sich nun als Ausgangspunkt für eine erste Untersuchung dieser Frage an, da in der in der Literatur bereits gezeigt wurde, dass die Verbal- und Satzklammer, die zu einem strukturell abgegrenzten MF (und NF) führt, im Laufe der Periode immer weiter ausgebaut wird, dahingehend, dass zunehmend mehr Konstituenten im Mittelfeld stehen und weniger ausgeklammert werden (Rösler 1997; Mähl 2014; Dreessen und Ihden 2015). Petrova (2013) zeigt, dass im Mnd. zumindest oberflächlich sowohl OV- als auch VO-Wortstellungen nachgewiesen sind. Daher findet man im Mnd. Sätze mit mehreren Konstituenten im MF (6a), sowie Sätze mit mehreren Konstituenten im NF (6b).

- 6) a. Nemant schall [ock] [des nachtes] [mit lichten ane lucerne] [In den stall] gan  
 (Greifswalder Büegersprache, 1451)  
 ‘Niemand soll auch des Nachtes mit Lichtern ohne Laterne in den Stall gehen’  
 b. vnde se schal sitten [openbare] [ane bode] [in deme hus] (Stralsunder Urkunden, 1301–1500)  
 ‚und sie soll sitzen öffentlich ohne Geselle in dem Haus’

Allerdings ist die diachrone Entwicklung dieses Phänomens im Westgermanischen bisher noch nicht untersucht worden. Es ist zwar bekannt, dass die Wortstellung in älteren germanischen Sprachstufen im Vergleich zu den modernen Sprachen generell freier ist (bspw. Hinterhölzl und

Petrova 2009; Bech und Eide 2014). Der Status der Wortstellung im MF ist jedoch bei mehreren älteren Sprachstufen umstritten, wie beispielsweise Frey 2003 und Demske und Patak 2016 zu Adjunkten im MF im Frühneuhochdeutschen zeigen. Das Mnd. kann daher sowohl eine historische als auch eine typologische Lücke füllen.

Im vorliegenden Beitrag untersuchen wir, ob die oben genannten Spiegeleffekte im Mnd. zu erkennen sind. Das CHLG bietet eine ideale Gelegenheit, diese Frage zu untersuchen. Die hier präsentierte Vorstudie ist auf Hauptsätze mit einem strukturell abgegrenzten MF und NF, d.h. Hauptsätze mit einer Verbalklammer aus einem Hilfsverb und einem nicht-finiten Verb beschränkt.<sup>10</sup> Aufgrund der konstituentenbasierten syntaktischen Annotation des CHLG und der Abfragesprache CorpusSearch (Randall 2005) lassen sich MF und NF gut identifizieren und isolieren. Dazu wird spezifiziert, welche Art(en) von Konstituenten im MF oder NF stehen sollen und auch jeweils wie viele davon. Um zum Beispiel die Abfolge von PP-Argumenten und PP-Adjunkten im MF zu untersuchen, kann eine Suchanfrage nach Hauptsätzen mit zwei oder mehr PPs im MF formuliert werden.<sup>11</sup> Solche Suchanfragen sind nur durch die konstituentenbasierte Annotation möglich; als Voraussetzung muss man mehrere ganze phrasale Konstituenten (bspw. PPs und ADVPs) beschreiben können, die durch reine POS-tags wie bei der Annotation des ReNs nicht identifizierbar sind. Durch solche Anfragen kann man also zuverlässige Datensätze mit sehr spezifischen und subtilen syntaktischen Merkmalen relativ schnell aus dem CHLG abrufen, die sich gut für weitere quantitative oder qualitative Untersuchungen eignen.

Als Erstes wurde untersucht, ob die Abfolge von PP-Argumenten und PP-Adjunkten im NF im Vergleich zur Abfolge im MF invertiert ist, wie im Niederländischen (4). Dazu wurden alle Hauptsätze mit mindestens zwei PPs im MF oder im NF herausgesucht. Aus den so gefundenen 81 Hauptsätzen mit mindestens zwei PPs im MF wurden manuell die 18/81 Sätze mit einem PP-Argument und einem PP-Adjunkt im MF herausgefiltert (in den übrigen Fällen handelte es sich um rein adverbiale PPs). Die Ausgangshypothese war, dass das Mnd. sich wie das Niederländische verhält, also im MF die Abfolge Adjunkt-Argument zeigt. Bei allen 18 Sätzen wurde diese Hypothese bestätigt. Einige Beispiele werden in (7) gezeigt.

- 7) a. vnd \*con\* wart dar [na sime ccclxv iare] [in den paradis] ghevort (Engelhus, Chronica nova, 1435)  
‘und (er) wurde dar nach seinem 365. Jahr in das Paradies geführt’

---

<sup>10</sup> query: (IP-MAT\* iDoms VAFIN|VMFIN) AND (IP-MAT\* iDoms VVINF|VVPP) AND (VAFIN|VMFIN precedes VVINF|VVPP)

<sup>11</sup> query: (IP-MAT\* iDoms VAFIN|VMFIN) AND (IP-MAT\* iDoms VVINF|VVPP) AND (VAFIN|VMFIN precedes VVINF|VVPP) AND (IP-MAT\* iDoms PP) AND (PP hasSister PP\*) AND (VAFIN|VMFIN precedes PP) AND (PP precedes VVINF|VVPP) AND (VAFIN|VMFIN precedes PP\*) AND (PP\* precedes VVINF|VVPP)

- b. vnde sik ne mach [mit rechte] nemant [to der vormundescap] teen (Herforder Rechtsbuch, 1375)  
‘und niemand mach sich mit rechte in die Vormundschaft begeben’
- c. so en sculden se [binnen dessen Ses weken]... nigt vppe de straten komen (Oldenburger Urkunden, 1350–1500)  
‘so sollten sie binnen diesen sechs Wochen nicht auf die Strasse kommen’
- d. Se scolde ot ok [to den hilgen dagen To pinxten] [in de kerken] dragen ‘Sie sollte es auch an den heiligen Tagen zu Pfingsten in die Kirche tragen’ (Zeno, 1401–1450)

Was das Nachfeld betrifft, wurden 172 Hauptsätze mit mindestens zwei ausgeklammerten PPs gefunden. Auch diese Daten wurden manuell geprüft, um alle Belege mit einem PP-Adjunkt und einem PP-Argument im NF herauszufiltern ( $n=15$ ). 14/15 Sätze zeigen die gespiegelte Abfolge im NF, d.h. Argument-Adjunkt, vgl. die Beispiele in (8).

- 8) a. dar schal ze vore ghaen [to rome] [dorch myner zele willen] (Stralsunder Urkunden, 1301–1500)  
‘da soll sie vorgehen zu Rom durch Willen meiner Seele’
- b. Ock schall dar nymant veghelsze mes edder ander vnstedicheit bringen lathen 10 [vp dat grote market] [by iij punden] (Greifswalder Bürgersprache, 1451)  
‘Auch soll da niemand Abfall, Mist oder anderen Unrat bringen lassen auf den großen Markt bei (Strafe von) drei Pfund’

Bei einem Beispiel weicht die Abfolge der PPs im NF allerdings von diesem Muster ab, siehe (9). Es ist jedoch anzumerken, dass in diesem Beleg vier PPs im NF stehen, wobei das Argument von den Adjunkten flankiert wird, die Mehrzahl der Adjunkte jedoch, wie erwartet, dem Argument folgt (Adjunkt-Argument-Adjunkt-Adjunkt).

- 9) Vnde \*pro\* wart do geuoret [vp eynen witten telder] [yn dat pallas] [myt groter vroude] [myt volghende volke]. (Griseldis, 1502)  
‘Und sie wurde do geführt auf einem Zelter in den Palast mit großer Freude mit 20 folgendem Volk’

Da (9) die einzige Ausnahme ist, lässt sich schließen, dass die überwiegende Wortstellung im NF Adjunkt-Argument ist und daher, dass tatsächlich ein Spiegeleffekt bezüglich der Abfolge von PP-Adjunkten und PP-Argumenten im Mnd. vorhanden ist.

Um zu untersuchen, ob im Mnd. auch ein Spiegeleffekt in Bezug auf die Abfolge von Temporal-, Modal- und Lokaladverbien (wie in (4)) belegt ist, wurden alle Hauptsätze herausgesucht, in denen mindestens zwei der folgenden Konstituenten entweder im MF oder im NF vorhanden sind: AdvP (ADVP), PP (PP), NP-Adjunkt (NP-TMP|NP-ADT). Auf diese Weise wurden 292 Sätze mit einem mehrfach belegten MF gefunden, und 214 mit einem mehrfach belegten NF. Alle diese Sätze wurden manuell geprüft und solche mit mindestens zwei semantisch verschiedenen Adjunktclassen (Zeit, Ort, Art&Weise) im MF oder NF zurückgehalten.<sup>12</sup> Für die Abfolge im MF gab es 25 relevante Belege. Die Mehrheit (22/25) entspricht dem Wortstellungsmuster in (4a), d.h. Zeit-Ort-Art&Weise; Temporaladverbien gehen Lokal- und Modaladverbien vorwiegend voran (10a), während Lokaladverbien vorwiegend Modaladverbien vorangehen (10b).

- 10) a. vnde dat sal hey io [ses weken vore] [in allen kerken] dun kundigen (Soester Schrae, 1367)  
 ‘und das soll er jo 6 Wochen vorher in allen Kirchen kündigen lassen’
- b. So hebbe ik myn jngesegele [vor my vnde myne erue] [vor myne husfrowen vnde eren eruen] [witliken] gehenghet (Stralsunder Urkunden, 1301–1500)  
 ‘So habe ich mein Siegel vor mir und meinen Erben, vor meiner Frau und ihren Erben wissentlich gehängt’

Diese Wortstellungstendenzen finden sich im NF in umgekehrter Reihenfolge; die Mehrheit der relevanten Belege (19/22) entspricht dem Muster in (4b); Modaladverbien stehen vorwiegend vor Lokal- und Temporaladjunkten (11a) und Lokaladjunkte vorwiegend vor Zeitadjunkten (11b).

- 11) a. man sal it don [in gemene hant] [iar vnde dach]. (Braunschweiger Urkunden, 1301–1500)  
 ‘man soll es tun in gemeiner Hand Jahr und Tag’
- b. vnde \*pro\* is ghegheuen [in der stad to me sunde] [na ghodes bord dusent yar drehundert yar] [yn deme achte vnde sostighesten yare] [yn deme Daghe sunte Johannes baptisten]... (Stralsunder Urkunden, 1301–1500)  
 ‘und (dies) ist gegeben in der Stadt bei dem Sund 1300 Jahre nach Gottes Geburt in dem 68. Jahre an dem Tag von Sankt Johann Baptist’

Mithilfe der syntaktisch annotierten Daten im CHLG lässt sich also zeigen, dass zwei

---

<sup>12</sup> Pronominaladverbien wie *daran*, *darinne*, die im CHLG auf der syntaktischen Ebene auch als PP annotiert sind, wurden hier ausgeschlossen.

verschiedene MF/NF-Spiegeleffekte im Mnd. belegt sind. Damit ist ein erster Schritt zur Untersuchung des diachronischen Status solcher Effekte im Westgermanischen gemacht. Wie mit der vorliegenden Vorstudie gezeigt wurde, haben solche Effekte - zumindest im Niederdeutschen - eine relativ lange Geschichte und scheinen in dieser Hinsicht im Westgermanischen tief verwurzelt zu sein. Weitere Untersuchungen müssen ausweisen, wodurch solche Effekte zustande kommen und ob es tieferliegende strukturelle Gründe für ihre scheinbare diachrone und sprachübergreifende Stabilität gibt. Zumindest ist gezeigt, dass Spiegeleffekte es auf jeden Fall wert sind, weiter erforscht zu werden, zumal über die theoretischen Grundlagen in Bezug auf Adjunkte, Wortstellung und syntaktische Struktur noch lange keine Einigkeit herrscht.

## 5 Nullsubjekte

Wie in Abschnitt 2 bereits besprochen wurde, haben geparste Korpora den Vorteil, dass sie im Gegensatz zu Korpora, die nur auf der Wortebene annotiert sind, die vollständige Satzstruktur berücksichtigen. Das bedeutet, dass nicht nur Information auf Wortebene, sondern auch die Beziehung oder Abhängigkeit von Konstituenten annotiert und durchsuchbar gemacht wird. Dadurch ist es möglich, einen Platz für „leere“ Kategorien innerhalb von Sätzen zu schaffen, was es ermöglicht, auch elliptische Strukturen im Korpus zu finden. Das ist ein durchaus konventioneller Schritt, bspw. werden der syntaktisch annotierten Potsdamer Baumbank des Frühneuhochdeutschen (Demske 2019) afinite Konstruktionen (Ellipse des finiten Hilfsverbs) mit einem separaten Knotenlabel (AS) markiert, da sie beispielsweise mit einer Volltextsuche nicht auffindbar wären.

Im CHLG werden aufgrund ähnlicher Erwägungen elliptische Kategorien mit einem Platzhalter an der Stelle des zu erwartenden lexikalischen Materials annotiert. Auf diese Art und Weise können beispielsweise referenzielle Nullsubjekte leichter gefunden werden, die mit \*pro\* annotiert werden, vgl. auch (11b) im vorigen Abschnitt.<sup>13</sup>

---

<sup>13</sup> Darüberhinaus werden nicht overt realisierte Expletiva und Korrelate (1a) sowie Konjunktionsreduktionen (1b) annotiert.

(1) a. ock \*exp\* ys better dat men dat crud in cleyne stücke dar vnder lege (Abdinghofer Arzneibuch 1451–1500)

‘[Es] ist auch besser, das Gewürz in kleinen Stücken darunter zu legen.’

b. Jtem so scalt u eme verbeden dat he nicht to vro bade vnde dat he nicht en ete to harde spyse vnde \*con\* drincke neynen bosen dranck (Abdinghofer Arzneibuch 1451–1500)

‘Genauso solltest du darauf achten, dass er nicht zu früh badet und dass er keine zu harten Speisen esse und [er] trinke kein schlechtes Getränk.’

Das Nullelement wird, der Tradition der Penn Parsed Corpora folgend, immer gleich am Satzanfang eingefügt, ungeachtet der vermutlichen tatsächlichen Position (in (1a) würde man vermutlich *auch ist es besser* ... annehmen). Die Annotation \*con\* wird nur verwendet, wenn das ausgelassene Subjekt koreferentiell und in der Zahl identisch mit dem Subjekt des vorangehenden Konjunktts ist (1b).



Obwohl referentielle Nullsubjekte in den germanischen Sprachen Gegenstand etlicher Studien sind (siehe bspw. Rosenkvist 2009 für eine Übersicht), wurde erst kürzlich entdeckt, dass sie auch im Mittelniederländischen vorkommen (Farasyn und Breitbarth 2016, 2021). Allerdings machen sie in diesen Studien, bei denen die Daten manuell annotiert wurden, nur einen sehr kleinen Teil der Gesamtzahl der pronominalen Subjekte im Korpus aus. Das CHLG ermöglicht es nun, viel Zeit bei der Suche nach diesem seltenen, aber hochinteressanten linguistischen Phänomen einzusparen, indem man im Korpus nach dem Platzhalter \*pro\* sucht. Die Studien von Farasyn und Breitbarth (2016, 2021) hatten ferner gezeigt, dass referentielle Nullsubjekte asymmetrisch über Haupt- (12a) und Nebensätze (12b) verteilt sind. Auch das kann ohne Probleme in einer Suchanfrage im CHLG formuliert werden, wo als Suchdomäne IP-MAT\* bzw. IP-SUB\* spezifiziert werden kann.

- 12) a So we in my nicht ne bliuet de scal uthgheworpen warden Vnde scal dorren also de winlode vnde se scolen ene to hope binden vnde werpen ene in dat vur Vnde \*pro\* scal vorbernen. (Buxtehuder Evangeliar, 1451–1500)  
 ‚Wer sich meinem Urteil nicht unterwirft, wird verbannt und verdorrt wie die Reben. Und sie sollen ihn zusammenbinden und ihn ins Feuer werfen. Und [er] soll verbrennen‘.
- b. forder szo schole gy weten dath \*pro\* yuwen breff den ghy my by yuwen steffader senden entfangen hebbe (Agneta Willeken 1535)  
 ‚Außerdem solltest du wissen, dass [ich] deinen Brief erhalten habe, den du mir bei deinem Stiefvater abliefern lassen hast‘

Eine Suche im CHLG liefert dabei ähnliche Ergebnisse wie die von Farasyn und Breitbarth (2016) beschriebenen: referentielle Nullsubjekte treten im Mnd. vorzugsweise in Hauptsätzen auf, vgl. Tab. 4, was unterstreicht, dass die alte Asymmetrie, die bereits für das Althochdeutsche und Altsächsische charakteristisch war, im Mnd. bewahrt bleibt (vgl. Axel 2005, Walkden 2014).

	*pro*	Overt	Total <sup>14</sup>
Hauptsatz	241 (2,2%)	10.526 (97,8%)	10.767

<sup>14</sup> Zusätzlich gibt es auch noch Nullexpletiva (\*exp\*) und Konjunktionsreduktionen, die hier jedoch nicht mitgezählt wurden.

Nebensatz	67 (0,7%)	9.248 (99,3%)	9.315
<b>Total</b>	<b>308 (1,5%)</b>	<b>19.774 (98,5%)</b>	<b>20.082</b>

Tab. 4: Referentielle Nullsubjekte in Haupt- und Nebensätzen im CHLG

Referentielle Nullsubjekte kommen besonders häufig in Koordinationsstrukturen, sowohl Haupt- als auch Nebensätzen, ohne dass es sich dabei um Konjunktionsreduktion (\*con\*) handeln würde, da die Verben hinsichtlich der Kongruenzmerkmale nicht miteinander übereinstimmen. In (13) beispielsweise ist das Antezedens des referentiellen Nullsubjekts (3. Person Singular) im letzten Teilsatz das direkte Objekt des vorangehenden Satzes (*de iukfrowe*), nicht mit dem Subjekt. Denn dieses, zudem in der 3. Person Plural, ist unter Koordinationsellipse ausgelassen (\*con\*), und ist mit dem Subjekt des ersten Teilsatzes koreferent (*se*).

- 13) *Se<sub>i</sub> gingen henin myt groten wunder des volkes vnde \*con\*<sub>i</sub> vunden de iukfrowe<sub>j</sub> schaffen in deme huseken Vnde \*pro\*<sub>j</sub> erschrak sere van deme seltzen gaste. (Griseldis 1502)*  
 ‚Sie gingen unter großem Erstaunen des Volkes hinein und [sie] fanden die Jungfrau arbeiten in dem Häuschen und [sie] erschrak sehr vor dem sonderbaren Gast.‘

Obwohl Farasyn und Breitbarth (2016, 2020) anmerken, dass es im Mittelniederdeutschen grundsätzlich zwei Arten von Nullsubjekten gibt, nämlich Nulltopiks in Spez,CP (14a) und Null-Wackernagelklitika rechts von C / in Spez,TP (14b), werden diese im CHLG nicht gesondert annotiert. Zunächst einmal werden Hauptsätze in Übereinstimmung mit dem Penn-Format ohnehin nicht als CPs, sondern als IP-MAT annotiert. Nur in Sätzen mit w-Bewegung und in Nebensätzen gibt es eine CP-Ebene. Die syntaktische Annotation stimmt also – wie auch bei der flachen Annotation des Mittelfeldes ohne VP-Knoten oder evtl. funktionale Projektionen – nicht mit der in der Literatur gängigen Analyse der Satzstruktur überein, sondern ist ausschließlich dazu bestimmt, das Auffinden von Strukturen zu erleichtern, nicht, eine vollständige theoretisch untermauerte Strukturanalyse zu liefern. Das gibt Nutzern des Korpus die Möglichkeit, eine eigene Analyse zu formulieren, statt sie von der Annotation aufgezwungen zu bekommen. Dennoch ist es möglich, mit einer differenzierten Suchanfrage die beiden Typen aufgrund ihrer Position in der Struktur zu unterscheiden.<sup>15</sup> Auf diese Art und Weise lassen sich 131 (43%)

<sup>15</sup> In Hauptsätzen kann man \*pro\* im Vorfeld / in Spez,CP mit Anfrage (1) suchen. Referenzielle Nullsubjekte in Spez,TP findet man im Hauptsatz Nebensatz mit Anfrage (2) und im Nebensatz mit Anfrage (3).

(1)	node: IP-MAT*	(2)	node: IP-MAT*	(3)	node IP-SUB*
	query: (IP-MAT* iDoms NP-SBJ)		query: (IP-MAT* iDoms NP-SBJ)		query: (IP-SUB* iDomsNum 1 NP-SBJ)
	AND (IP-MAT* iDoms VVFIN)		AND (IP-MAT* iDoms VVFIN)		AND (NP-SBJ dominates \*pro\*)
	AND (NP-SBJ dominates \*pro\*)		AND (NP-SBJ dominates \*pro\*)		
	AND (NP-SBJ precedes VVFIN)		AND (VVFIN precedes NP-SBJ)		

Belege mit einem referentiellen Nullsubjekt in Spez,CP finden; die anderen 177 Belege befinden sich in der Position nach C (Spez,TP).

- 16) a. Des vor wunderde sek dat her vnd \*pro\* hedden gerne wust wat he dar mede meynde  
(Engelhus, Chronica nova, 1435)  
, Der Herr war davon sehr überrascht und hätte gerne gewusst, was er damit meinte‘
- b. Also lef heuet god de werlt ghehad dat \*pro\* sinen enbornen sone uor se geue vppe  
dat en iewelk de in eme louet nicht ne uorderue. (Buxtehuder Evangeliar, 1451–1500)  
‘So sehr hat Gott die Welt geliebt, dass er seinen eingeborenen Sohn für sie geopfert hat,  
damit alle, die an ihn glauben, nicht untergehen.‘

Es mag eingewendet werden, dass die Annotation nicht-overten Materials vielleicht dem Anspruch des Penn-Schemas widerspricht, so wenig Analyse wie möglich in der selbst Annotation zu verankern. Die Abfragesprache CorpusSearch erlaubt den Gebrauch eines Negationsoperators, man könnte also im Prinzip nach Sätzen ohne NP-SBJ-Zweig suchen.<sup>16</sup> Allerdings müsste in diesem Fall manuell zwischen referentiellen Nullsubjekten, Nullexpletiva bzw. Nullkorrelaten und Konjunktionsreduktionen unterschieden werden, was v.a. im Mnd. bei der verhältnismäßig geringen Anzahl von referentiellen Nullsubjekten und der sehr großen Anzahl von Koordinationsellipsen zu größerer Fehleranfälligkeit führen dürfte.

## 6 Schlussbetrachtung

Im vorliegenden Beitrag sollte anhand dreier empirischer Phänomene gezeigt werden, dass die syntaktische Annotation des CHLG es erlaubt, schnell, zuverlässig und reproduzierbar Strukturen aufzufinden, die mit einem syntaktisch nicht annotierten Korpus wie dem ReN nicht ohne weiteres zu finden wären. Es ist unsere Hoffnung, dass möglichst viele Nutzer diese Möglichkeiten zu schätzen wissen werden.

---

<sup>16</sup> Beispielsweise  
(1) node: IP\*  
query: (IP\* dominates ! NP-SBJ)

**Danksagung:** Das CHLG kam mithilfe einer großzügigen Infrastrukturförderung durch die Hercules-Stiftung (01.07.2014-31.12.2015, AUGÉ/13/02) und den Fonds voor Wetenschappelijk Onderzoek – Vlaanderen (01.01.2016-31.12.2020, FWO G0F2614N) zustande. Besonderer Dank ist den Mitarbeitern des ReN-Projekts für die Ermöglichung unserer Zusammenarbeit geschuldet, insbesondere Robert Peters und Ingrid Schröder, sowie Fabian Barteld, Katharina Dreessen, Sarah Ihden und Norbert Nagel für die praktische Umsetzung dieser Zusammenarbeit. Die Autorinnen danken weiterhin den anderen (ehemaligen) CHLG-Projektmitgliedern Aaron Ecay, Liliane Haegeman, Veronique Hoste und Mariya Koleva, sowie den externen Beratern des Projekts, Sheila Watts, George Walkden und Svetlana Petrova, für ihren jeweiligen Beitrag zu diesem Projekt.

## Literatur

- Axel, Katrin. 2005. Null subjects and verb placement in Old High German, in: Stephan Kepser u. Marga Reis (eds.): *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*, Berlin, 27-48.
- Barbiers, Sjef. 1995. *The Syntax of Interpretation*. Leiden, NL: Rijksuniversiteit Leiden Diss.
- Barteld, Fabian, Sarah Ihden, Katharina Dreessen & Ingrid Schröder. 2018. HiNTS: A Tagset for Middle Low German. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, 3940-3945.
- Bech, Kristin & Kristine Gunn Eide. 2014. *Information Structure and Syntactic Change in Germanic and Romance languages*. Amsterdam: John Benjamins.
- Beck, Jana, Aaron Ecay & Anton Karl Ingason. 2015. *Annotald. Version 1.3*. 7.
- Booth, Hannah, Anne Breitbarth, Aaron Ecay & Melissa Farasyn. 2020. A Penn-style Treebank of Middle Low German. In *Proceedings of the Twelfth International Conference on Language Resources and Evaluation (LREC-2020)*, 766-775. <http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.96.pdf>.
- Breitbarth, Anne. 2014. *The history of Low German negation*. Oxford University Press.
- Breitbarth, Anne. 2018. Degrees of integration: Resumption after left-peripheral conditional clauses in Middle Low German. Vortrag DiGS XX (York) und SLE 2018 (Tallinn).
- Breitbarth, Anne. i.Dr.. Prosodie, Syntax und Diskursfunktion von V>2 in gesprochenem Deutsch. *Deutsche Sprache*.
- Demske, Ulrike. 2019. Referenzkorpus Frühneuhochdeutsch: Baumbank. UP. Universität Potsdam: Institut für Germanistik (<https://hdl.handle.net/11022/0000-0007-EAF7-B>)

- Demske, Ulrike & Marianna Patak. 2016. The syntax of adverbials in Early New High German. Handout, SLE 2016, Naples.
- Dipper, Stefanie, Karin Donhauser, Thomas Klein, Sonja Linde, Stefan Müller & Klaus-Peter Wegera. 2013. HiTS: ein Tagset für historische Sprachstufen des Deutschen. *Journal for Language Technology and Computational Linguistics* 28(1). 85-137.
- Donhauser, Karin & Svetlana Petrova. 2009. Syntaktische Variation in Nachsatz-Konstruktionen des Mittelhochdeutschen. Sprachliche Evidenzen zur Überlieferungsgeschichtlichen Diskontinuität im mhd. Prosa-Lancelot. In Doris Wagner, Tuomo Fonsen & Henrik Nikula (Hrsg.), *Germanitisk zwischen Baum und Borke. Festschrift für Kari Keinistö zum 60. Geburtstag*, 289-303. Helsinki: Societe Neophilologique.
- Dreessen, Katharina & Sarah Ihdn. 2015. Korpuslinguistische Studien zur mittelniederdeutschen Syntax. *Jahrbuch für Germanistische Sprachgeschichte* 6(1). 249-275.
- Farasyn, Melissa & Anne Breitbarth. 2016. Nullsubjekte im Mittelniederdeutschen. *Beiträge zur Geschichte der deutschen Sprache und Literatur*, 138(4), 524-559.
- Farasyn, Melissa & Anne Breitbarth. 2021. Null subjects in Middle Low German : diachronic stability and change. In J. Gísli Jónsson & T. Eythórsson (Hrsg.), *Syntactic features and the limits of syntactic change*, 84-110. Oxford University Press.
- Farasyn, Melissa. 2018. Fitting in or standing out? Subject agreement phenomena in Middle Low German. Ghent University Diss.
- Frey, Werner. 2003. Syntactic conditions on adjunct classes. In E Lang, C Maienborn & C Fabricius-Hansen (Hrsg.), *Modifying adjuncts*, Bd. 163-209, 209. Berlin: de Gruyter.
- Haider, Hubert. 2000. Adverb Placement-Convergence of Structure and Licensing. *Theoretical Linguistics* 26(1-2). 95-134.
- Haug, Dag. 2015. Treebanks in historical linguistic research. In Carlotta Viti (Hrsg.), *Perspectives on Historical Syntax*. Benjamins.
- Haug, Dag T & Marius Jøhndal. 2008. Creating a parallel treebank of the old Indo-European Bible translations. In *Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008)*, 27-34.
- Hinterhölzl, Roland & Svetlana Petrova. 2009. *Information Structure and Language Change: New Approaches to Word Order Variation in Germanic*. Berlin: de Gruyter.
- Ihdn, Sarah. 2020. *Relativsätze im Mittelniederdeutschen. Korpuslinguistische Untersuchungen zu Struktur und Gebrauch*. De Gruyter.
- Koster, Jan. 1974. Hetwerkwwoord als spiegelcentrum. *Spektator* 3. 601-618.

- Koster, Jan. 2001. Mirror symmetry in Dutch. In Marc van Oostendorp & Elena Anagnostopoulou (Hrsg.), *Grammar in Progress: Articles Presented at the 20th Anniversary of the Comparison of Grammatical Models Group in Tilburg*. Amsterdam.
- Mähl, Stefan. 2004. *Studien zum mittelniederdeutschen Adverb*. Böhlau.
- Mähl, Stefan. 2014. *Mehrgliedrige Verbalkomplexe im Mittelniederdeutschen: ein Beitrag zu einer historischen Syntax des Deutschen*. Böhlau.
- Merten, Marie-Luis. 2018. *Literater Sprachausbau kognitiv-funktional. Funktionswort-Konstruktionen in der historischen Rechtsschriftlichkeit*. De Gruyter.
- Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty & Daniel Zeman. 2016. *Universal Dependencies v1: A Multilingual Treebank Collection*. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 1659-1666. Portoroz, Slovenia: European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/L16-1262>.
- Peters, Robert. 1973. *Mittelniederdeutsche Sprache*. In Jan Goossens (Hrsg.), *Niederdeutsch - Sprache und Literatur*. Band 1: *Sprache*, 66-115. Wachholtz.
- Petrova, Svetlana. 2012. Multiple XP-fronting in Middle Low German root clauses. *The Journal of Comparative Germanic Linguistics* 15(2). 157-188.
- Petrova, Svetlana. 2013. *The Syntax of Middle Low German*. Habilitation, Humboldt-Universität zu Berlin.
- Pintzuk, Susan, Ann Taylor & Anthony Warner. 2017. *Corpora and Quantitative Methods*. In Adam Ledgeway & Ian Roberts (Hrsg.), *The Cambridge Handbook of Historical Syntax*, 218-240. Cambridge University Press.
- Pittner, Karin. 2004. „Where syntax and semantics meet: adverbial positions in the German middle field“. In JR Austin, S. Engelberg & G. Rauh (Hrsgg.), *Adverbials: the interplay between meaning, context, and syntactic structure*, Amsterdam/Philadelphia, John Benjamins. 253-287.
- Randall, Beth. 2005. *CorpusSearch2 User's Guide*. Philadelphia: Dept. of Linguistics, University of Pennsylvania. <http://corpussearch.sourceforge.net>.
- ReN-Team. 2017. *Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200-1650)*. Archived in Hamburger Zentrum für Sprachkorpora. Version 0.3. Publication date 2017-06-15. <http://hdl.handle.net/11022/0000-0006-473B-9>.
- Rosenkvist, Henrik. 2009. Referential Null Subjects in Germanic Languages – an Overview, in: *Working Papers in Scandinavian Syntax* 84, p. 151–180.

- Rösler, Irmtraud. 1997. Satz, Text, Sprachhandeln: Syntaktische Normen der mittelniederdeutschen Sprache und ihre soziofunktionalen Determinanten. Universitätsverlag 20 Winter.
- Santorini, Beatrice. 1990. Part-of-speech tagging guidelines for the Penn Treebank Project (3rd revision). Technical Report, University of Pennsylvania Department of Computer & Information Science.
- Schalowski, Sören. 2015. Wortstellungsvariation aus informationsstruktureller Perspektive: 25 Eine Untersuchung der linken Satzperipherie im gesprochenen Deutsch. Working Papers of the SFB 632. Interdisciplinary Studies on Information Structure (ISIS), vol. 18. Potsdam: Universitätsverlag Potsdam.
- Schiller, Anne, Simone Teufel, Christine Stöckert & Christine Thielen. 1999. Guidelines für das Tagging deutscher Textcorpora mit STTS (Kleines und großes Tagset). Uni-30 Universität Stuttgart: Institut für maschinelle Sprachverarbeitung. <http://www.sfs.uni-tuebingen.de/resources/stts-1999>.
- Sievers, Eduard. 1878. Heliand. Verlag der Buchhandlung des Waisenhauses.
- Taylor, Ann. 2020. Treebanks in Historical Syntax. *Annual Review of Linguistics* 6(1). 195-212. <https://doi.org/10.1146/annurev-linguistics-011619-030515>.
- Tophinke, Doris. 2009. Vom Vorlesetext zum Lesetext: Zur Syntax mittelniederdeutscher Rechtsverordnungen im Spätmittelalter. In A. Linke & H. Feilke (Hrsg.), *Oberfläche und Performanz. Untersuchungen zur Sprache als dynamischer Gestalt*, 161-183. Tübingen: Niemeyer.
- Walkden, George 2014: *Syntactic Reconstruction and Proto-Germanic*, Oxford.
- Walkden, George. 2015. HeliPaD: the Heliand Parsed Database. Version 0.9. <http://www.chlg.ac.uk/helipad/>.
- Walkden, George. 2016. The HeliPaD. A parsed corpus of Old Saxon. *International Journal of Corpus Linguistics* 21(4). 559-571. <https://doi.org/https://doi.org/10.1075/ijcl.21.4.05wal>.
- Wallmeier, Nadine. 2013. Sprachliche Muster in Der Mittelniederdeutschen Rechtssprache: Zum Sachsenspiegel Und Zu Stadtrechtsaufzeichnungen Des 13. Bis 16. Jahrhunderts. Winter.
- Wallmeier, Nadine. 2015. Konditionale Adverbialsätze und konkurrierende Konstruktionen in mittelniederdeutschen Rechtstexten. *Niederdeutsches Jahrbuch* 138. 7-26.
- Wiese, Heike & Hans G. Müller. 2018. The hidden life of V3: An overlooked word order variant on verb-second. In Mailin Antomo & Sonja Müller (Hrsg.), *Non-canonical verb positioning in main clauses*, Bd. *Linguistische Berichte Sonderheft* 25, 201-223. Hamburg:

Buske.

Wiese, Heike, Mehmet Tahir Öncü & Philipp Bracker. 2017. Verb-dritt-Stellung im türkisch-deutschen Sprachkontakt: Informationsstrukturelle Linearisierungen ein- und mehrsprachiger Sprecher/innen. *Deutsche Sprache* 2017/1. 31-52.

Wiese, Heike, Mehmet Tahir Öncü, Hans G. Müller & Eva Wittenberg. 2020. Verb third in spoken German: A natural order of information? In Rebecca Woods & Sam Wolfe (Hrsg.), *Rethinking Verb Second*, 682-699. Oxford: Oxford University Press.