

# Das Referenzkorpus Mittelniederdeutsch/ Niederrheinisch als Quelle für Studien zur Lexembildung Zur Varianz des Suffixes *-schop*

*Sarah Ilden und Ingrid Schröder*

## 1 Einleitung

Nachdem die großen historischen Grammatiken des (Hoch-)Deutschen der Lexembildung jeweils einen eigenen Band gewidmet hatten (vgl. Grimm 1826, Wilmanns 1899, Paul 1920), blieb dieser Grammatikbereich lange Zeit Einzelstudien vorbehalten (vgl. Kluge 1925, Henzen 1965) und fand bestenfalls marginal Eingang in die Gesamtgrammatik einer Sprachstufe. Die Feststellung von Klein/Solms/Wegera (2009: VII), dass die Beschreibung der Lexembildung „eine der großen Lakunen der mhd. Grammatik“ darstellt, lässt sich daher ohne Weiteres auch auf andere historische Sprachstufen übertragen, insbesondere auf das Mittelniederdeutsche. Bereits ein flüchtiger Blick in die älteren Grammatiken des Mittelniederdeutschen (vgl. Lübben 1882, Colliander 1912, Lasch 1914, Sarauw 1924) zeigt, dass diese die Lexembildung nahezu vollständig ausgespart haben. In Einzelstudien sind vor allem die Suffigierung abstrakter Substantive (vgl. Hoge 1912, Grunewald 1944, Dahlberg 1962), die Diminution (vgl. Latendorf 1878, Seelmann 1919, Seelmann 1920, Katara 1954), die Adverbbildung (vgl. Listen 1999, Mähl 2004) sowie die Feminin-Movierung (vgl. Schröder 1922, Werth 2015, Roolfs 2016, Möller 2017), schließlich auch einzelne Derivationsaffixe (vgl. Walther 1910 zum Adjektivsuffix *-em*, Ahlsson 1988 zum Substantivsuffix *-schap*, Seelmann 1927/28 und Klein 2004 zum Präfix *ge-*) thematisiert worden.

Dieser Befund führte vor gut zwei Jahrzehnten zum Plan einer systematischen Analyse der mittelniederdeutschen Lexembildung, wofür konzeptionelle Überlegungen und erste Analysen zur Derivation der Substantive, der Verben und der Adjektive von Möhn/Schröder (2003, 2007, 2009) veröffentlicht wurden. Diese Arbeiten verstanden sich zugleich als Vorstudien für eine neu zu konzipierende mittelniederdeutsche Grammatik, deren Realisierung jedoch erst jetzt mit der Erarbeitung der nominalen Flexionsmorphologie (Pilotphase 2020–2023, gefördert durch die DFG, Projekt-Nr.: 433078928) in Angriff genommen werden konnte.

Für die oben genannten Untersuchungen dienten als Materialbasis Belegsammlungen aus Texten (z. B. bei Grunewald 1944 aus Urkunden) oder das Mittelniederdeutsche Handwörterbuch (MNWB) mit seinem Archiv (z. B. bei Dahlberg 1962), das auch für die Analysen von Möhn/Schröder (2003, 2007, 2009) herangezogen wurde. Erst seit Kurzem ist es möglich, solche grammatischen Studien auch auf Basis digitaler Textkorpora durchzuführen. So wurde für den Band zur Wortbildung der Mittelhochdeutschen Grammatik (vgl. Klein/Solms/Wegera 2009) ein nach Zeit, Sprachraum und Textart strukturiertes Korpus mit insgesamt 102 Texten genutzt (vgl. Klein/Dipper 2016: 3), das durch weitere externe Quellen ergänzt wurde. Für Studien zum Mittelniederdeutschen und zum frühen Norddeutschen zogen Werth (2015) und Möller (2017) das Korpus der Hexenverhörprotokolle der Frühen Neuzeit (vgl. Macha et al. 2005) für die Analyse von Feminin-Movierungen insbesondere bei Namen heran. Da dieses Korpus lediglich fünf niederdeutsche Protokolle enthält, ist es nur für sehr begrenzte Fragestellungen zum Mittelniederdeutschen verwendbar. Erst das seit 2019 vollständig zur Verfügung stehende Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650) (Version 1.1 = ReN-Team 2021) bildet ein geeignetes empirisches Fundament für umfassende grammatische Analysen. Es eröffnet neue Möglichkeiten für eine variationssensitive Auswertung, die für Untersuchungen historischer Sprachstufen aufgrund der fehlenden Normierung unabdingbar ist. Variation besteht vor allem im Hinblick auf die Parameter Zeit, Sprachraum und Domäne/Textsorte, weniger ausgeprägt auch im Hinblick auf den individuellen Text bzw. Textproduzenten, kann aber auch durch sprachliche Kontakte oder sprechsprachlichen Einfluss begründet sein. Diese Faktoren gilt es bei einer grammatischen Beschreibung zu berücksichtigen.<sup>1</sup>

Ziel des vorliegenden Beitrags ist es, an einem Beispiel aus dem Bereich der Lexembildung, nämlich der Variation des Substantivsuffixes *-schop*, zu prüfen, auf welche Weise das Referenzkorpus Mittelniederdeutsch/Niederrheinisch (ReN) zur Untersuchung der diasystematischen Variation innerhalb der Grammatik dienen kann und welche methodischen Anforderungen dabei zu berücksichtigen sind. Einleitend wird das ReN vorgestellt, wobei die Korpusstruktur und die sich eröffnenden Nutzungsmöglichkeiten im Mittelpunkt stehen (Kap. 2). Einen Referenzpunkt

---

<sup>1</sup> Am 1.5.2024 konnte außerdem das Projekt ‚Wortfamilien diachron (Wodia)‘ (gefördert durch die DFG; Projektnummer: 524625420; Projektleitung: Dr. Thomas Burch, Prof. Dr. Jost Gippert, Dr. Sarah Ihden, Dr. Ralf Plate, Prof. Dr. Ingrid Schröder) seine Arbeit aufnehmen. Es zielt auf eine wortbildungsmorphologische Analyse zur Darstellung der Wortfamilien des historischen Hochdeutschen und Niederdeutschen. Als Basis dienen die Lemmalisten der jeweiligen Wörterbücher, die zusätzlich mit den Referenzkorpora verknüpft werden.

für die exemplarischen Analysen zum Suffix *-schop* und seinen Varianten stellt der Forschungsstand zum Mittelniederdeutschen dar (Kap. 3). Eine erste methodische Annäherung an die Identifizierung von Einflussfaktoren für die Variation bietet das statistische Modell der Random Forests (Kap. 4), das es ermöglicht, eine Hierarchisierung relevanter Faktoren vorzunehmen. Die Parameter werden einzeln untersucht, wobei die Bildung von Artefakten durch Überlagerung von Einflussfaktoren zu diskutieren ist (Kap. 5). Zusammenfassend lassen sich die Tendenzen raum-zeitlicher Variation für das Suffix *-schop* beschreiben (Kap. 6).

## 2 Das Referenzkorpus Mittelniederdeutsch/ Niederrheinisch (1200–1650) als Analysebasis

Eine Entscheidung über die Materialgrundlage für die Erarbeitung einer neuen Grammatik ist besonders virulent, wenn der Anspruch besteht, auf eine möglichst repräsentative und damit ausreichend große Datenbasis zurückgreifen zu können, die geeignet ist, diasystematische Variation jenseits der Lautebene hinsichtlich der diatopischen, diachronischen und diasituativen Vielfalt des mittelalterlichen und frühneuzeitlichen Textaufkommens abzubilden. Solche Anforderungen erfüllt nur ein strukturiertes und grammatisch annotiertes Korpus, das es erlaubt, grammatische Phänomene raum-zeitlich genau zu verorten und zusätzlich textsortenspezifische Schwerpunkte aufzudecken.

Mit dem ReN (vgl. als Überblicksdarstellungen Peters 2017, Schröder et al. 2017 sowie die Projektwebseite unter <https://www.slm.uni-hamburg.de/ren.html>) scheinen diese Erfordernisse auf den ersten Blick eingelöst. Es bietet in der zuletzt publizierten Korpusversion ReN 1.1 vom 06.01.2021 (vgl. ReN-Team 2021) 235 Texte mit ca. 2,3 Mio. Token, von denen ca. 1,5 Mio. Token mit dem Tagset HiNTS (Historisches Niederdeutsch-Tagset; vgl. Barteld et al. 2018) nach Wortarten und Flexionsformen annotiert sind.<sup>2</sup> Zusätzlich erfolgte eine Lemmatisierung nach den Regeln des MNWB.

Um der Vielfalt des Textaufkommens gerecht zu werden, wurde für die Textauswahl der Überlieferungszeitraum in 50-Jahres-Schritte unterteilt. Den Sprachraum repräsentieren neun Schreibsprachlandschaften, nämlich Nordniedersächsisch, Westfälisch, Ostfälisch und Elbostfälisch im ehemaligen altniederdeutschen Gebiet sowie Lübisches, Ostelbisches, baltisches Niederdeutsch (im Folgenden „Baltisch“) und Südmärkisch im Gebiet der hochmittelalterlichen Ostsiedlung. Das Niederrheinische tritt als sprachliche Übergangslandschaft zum rhein-maasländischen Sprachraum hinzu. Außerdem sind in soziopragmatischer Perspektive als Felder der Schriftlichkeit Verwaltung (v. a. Stadtbücher), Recht und Urkunden (Land-, Stadt-, Schiffsrechte, Burspraken, Zunftordnungen, Kontorordnungen, Testamente u. a.), Wissensvermittlung (Chronistik, moralische Didaxe, medizinische und andere alltagspraktische Schriften), Geistliche

---

<sup>2</sup> Im Unterschied zur Version ReN 1.0 enthält die Version 1.1 im annotierten Teil des Korpus hinzukommend 15 Urkunden aus dem lübischen, westfälischen, ostfälischen und elbostfälischen Sprachraum mit einem Gesamtumfang von 70.465 Token, die zuvor im ReN 1.0 lediglich in transkribierter Form vorlagen.

Schriftlichkeit/Religion (Bibeln, Evangeliare, Andachtstexte, religiöse Dichtungen), Literatur, Private Schriftlichkeit und Korrespondenz sowie Inschriften unterschieden worden, wobei eine klare Trennung der Felder nicht immer möglich ist. Alle diese Parameter sind als Metadaten mit den Texten verknüpft, sodass sie als unabhängige Variablen in eine quantitative Analyse einbezogen werden können.

Trotz dieser Vorteile des ReN bietet es sich bei einer Betrachtung vor allem der Lexemstrukturen an, als eine zweite lexembezogene Materialsammlung das MNWB mit seinen Archivbeständen heranzuziehen.<sup>3</sup> Mit ca. 1,2 Mio. Belegen und ca. 80.000 Lemmata stellt das Archiv die größte konventionelle Wortschatzsammlung zum Mittelniederdeutschen dar. Vergleicht man den Lexembestand von Korpus und Wörterbuch, so zeigt sich, dass trotz der großen Menge an Token im ReN die Zahl der Lemmata im Wörterbucharchiv wesentlich höher ist. Den 17.000 Lemmata im Korpus steht eine projektinterne Lemmaliste mit 72.000 Einträgen für bereits publizierte Wörterbuchartikel gegenüber.<sup>4</sup> Als Materialbasis für Lexembildungsanalysen im Rahmen einer neuen mittelniederdeutschen Grammatik ist somit eine Kombination von Wörterbuch und Referenzkorpus von Vorteil. Während mittels Wörterbuchdaten (v. a. der Lemmaliste) der Gesamtbestand der Lexembildungsmuster qualitativ sehr gut abgebildet werden kann, ist eine quantitative diastematische Analyse nur mithilfe des strukturierten Korpus möglich (vgl. dazu auch Schröder 2022). In diesem Aufsatz liegt der Fokus auf einer korpusbasierten Untersuchung, die eine wörterbuchgestützte Beschreibung des betreffenden Suffixes, wie sie Möhn/Schröder (2003) vorgelegt haben (s. dazu Kap. 3), sinnvoll ergänzen kann.

Die Repräsentativität von Sprachkorpora im Sinne einer exakten Abbildung der sprachlichen Verhältnisse durch eine Stichprobe stellt bekanntlich aufgrund der unklar umrissenen Grundgesamtheit ein nicht zu erreichendes Ideal dar. Daher ist für eine diastematische Betrachtung prinzipiell die Ausgewogenheit des Korpus (vgl. Lemnitzer/Zinsmeister 2010: 52f.) von besonderer Bedeutung. Dabei wird keine vollkommen gleiche Verteilung nach Raum, Zeit und Feld der Schriftlichkeit angestrebt, denn diese ist für historische Sprachstufen wie das Mittelniederdeutsche aufgrund der Struktur der Textüberlieferung, die nicht zuletzt von der Quantität der ursprünglichen Textproduktion bestimmt wird, ausgeschlossen. Während das 13. Jahrhundert davon gekennzeichnet ist, dass eine nur äußerst geringe Zahl an mittelniederdeutschen Texten entstanden und überliefert ist (vgl. Korlén 1945), ist die Überlieferungsdichte im 15. Jahrhundert besonders hoch. Für den Zeitraum von 1370 bis 1530 wird aufgrund des Geltungsbereichs im Ostseeraum weit über den norddeutschen Sprachraum hinaus und aufgrund seiner Funktion als überregionale Verkehrssprache vom „klassischen Mittelniederdeutsch“ (vgl. Peters 2000:

<sup>3</sup> Das Projekt ‚Wortbildung in der deutschen Urkundensprache des 13. Jahrhunderts‘ nutzt ebenfalls ein Wörterbuch als Basis, nämlich das Wörterbuch der mittelhochdeutschen Urkundensprache; vgl. Ganslmayer/Müller/Ring (2003: 335-338).

<sup>4</sup> Für eine Lemmaliste wurde die Printversion mit den Wortstrecken *a-* und *u* für die bisher unbearbeiteten Wortstrecken durch das ältere Handwörterbuch von Lübben/Walther (1888) ergänzt (zur Erstellung der Lemmaliste vgl. Kleymann/Nagel/Peters 2015). Mit der weiteren Ausarbeitung des MNWB wird die Zahl der Wörterbuchlemmata steigen, sodass die Schätzung von 80.000 Lemmata (vgl. oben) gerechtfertigt erscheint.

1483) gesprochen. Dass Lübeck im Hanseraum eine besondere wirtschaftliche und kulturelle Rolle übernahm, schlägt sich ebenfalls in der Überlieferung nieder. Diese Vorrangstellung Lübecks wird im Korpus dadurch berücksichtigt, dass Lübeck als eigener Raum repräsentiert ist. Die übrigen Räume sind an den strukturell (insbes. auf der phonologisch-graphematischen Ebene) unterscheidbaren dialektalen Schreibsprachlandschaften orientiert, wobei neben dem Lübischen das Nordniedersächsische, das Westfälische, das Ostfälische und das Ostelbische Kernräume der Überlieferung bilden, während auf der anderen Seite das Südmärkische insgesamt überlieferungsschwach geblieben ist, nicht zuletzt, weil es an Schreib- und später auch Druckorten mangelte. Dass innerhalb der Felder der Schriftlichkeit dem Recht, den Urkunden und den Verwaltungstexten eine besondere Bedeutung zukommt, hat Agathe Lasch (1925: 57) geradezu als das Spezifikum des mittelniederdeutschen gegenüber dem mittelhochdeutschen Textaufkommen angesehen. Alle diese Faktoren waren bei der Textauswahl zu berücksichtigen, auch wenn eine genaue Abbildung der tatsächlichen quantitativen Verhältnisse der historischen Textproduktion letztlich nicht erreicht werden konnte. Mit Blick auf die angestrebten grammatischen Analysen kommt hinzu, dass das Aufkommen der einzelnen grammatischen Phänomene im Korpus nicht gleich hoch ist. So lassen sich die graphematisch-phonologischen Verhältnisse recht gut mit beliebigen Textsorten erschließen. Bestimmte flexionsmorphologische Phänomene, z. B. das Vorkommen der 1. oder 2. Person Singular bei Verben, sind nicht nur verhältnismäßig selten anzutreffen (vgl. Klein 2007), sondern auch in hohem Maße abhängig von der Textsorte. Eine solche Textsortenspezifität ist auch für die Lexembildung zu veranschlagen, z. B. bei der Verteilung der Abstraktbildungen auf verschiedene Felder der Schriftlichkeit (z. B. Verwaltung vs. Geistliche Schriftlichkeit).

Abb. 1 zeigt die Verteilung der Texte im Raum-Zeit-Raster und signalisiert zugleich durch die Farbe und Größe der Kreise die Zugehörigkeit der Texte zu einem Feld der Schriftlichkeit sowie die Tokenzahl pro Text. Auf den ersten Blick wird sichtbar, dass die Kernzeit der Texterfassung von der 2. Hälfte des 13. Jahrhunderts bis zur 1. Hälfte des 16. Jahrhunderts reicht (218 Texte) und davon am meisten Texte (75) in der zweiten Hälfte des 15. Jahrhunderts entstanden sind. Betrachtet man die Sprachräume, so dominieren hinsichtlich der Textzahl das Ostfälische (40 Texte = 329.186 Token, davon 228.621 annotiert), das Nordniedersächsische (39 Texte = 504.921 Token, davon 402.578 annotiert) und das Westfälische (37 Texte = 404.527 Token, davon 151.045 annotiert). Diese drei Teilräume sind auch für diachrone Studien am besten geeignet, da für sie von der 2. Hälfte des 13. Jahrhunderts bis zur 1. Hälfte des 16. Jahrhunderts durchgehend Texte in vergleichsweise größerem Umfang vorhanden sind, für das Ostfälische auch in der 1. Hälfte des 13. Jahrhunderts, für das Nordniedersächsische bis zur 2. Hälfte des 17. Jahrhunderts. Hinsichtlich der Felder der Schriftlichkeit lässt sich folgende Verteilung feststellen: Im Rechtsbereich wurden 37 Texte aufgenommen (444.448 Token, davon 282.637 annotiert); Urkunden sind mit 70 Einheiten vertreten (216.936 Token, davon 172.943 annotiert), die Verwaltung mit 17 Texten (102.174 Token, davon 71.383 annotiert).

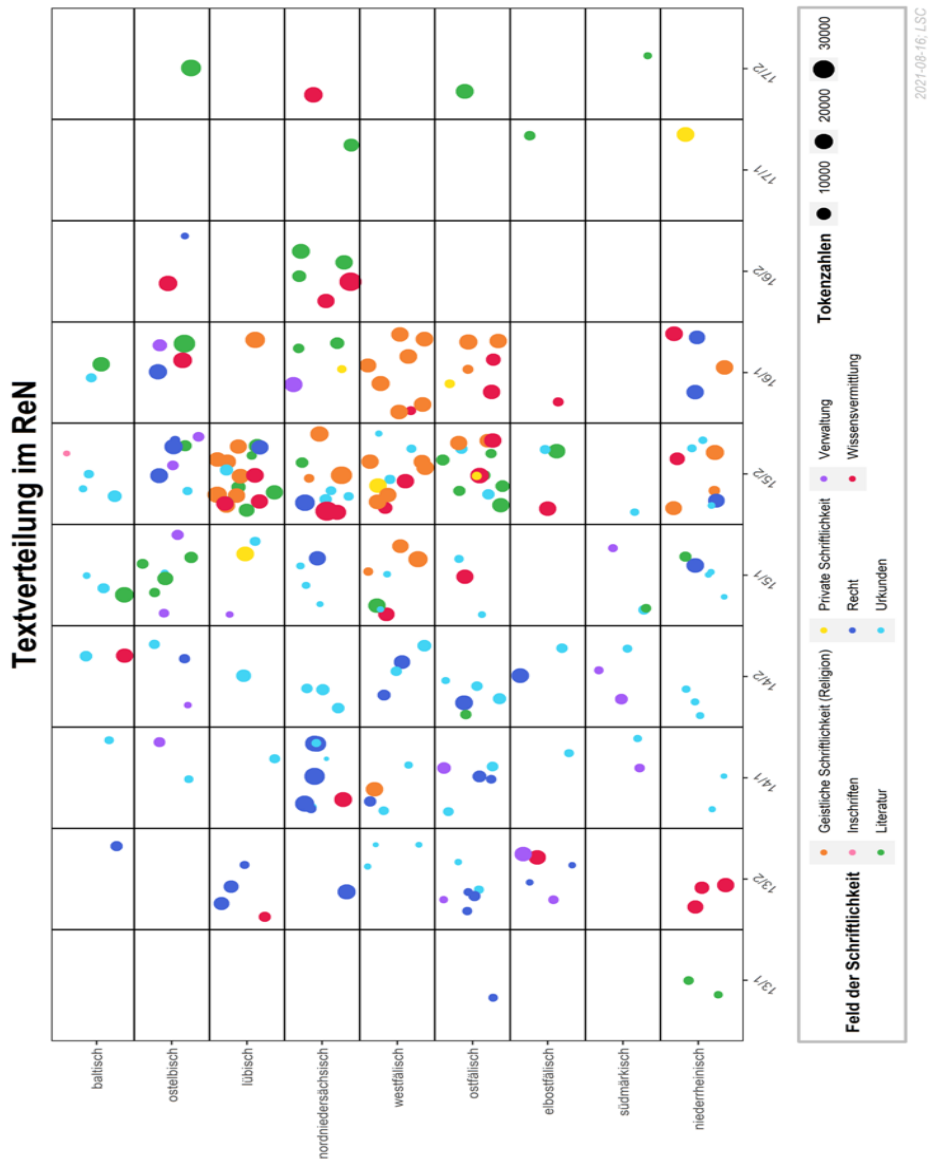


Abbildung 1: Textverteilung im ReN nach Raum, Zeit und Feld der Schriftlichkeit<sup>5</sup>

<sup>5</sup> Die Visualisierung der Textverteilung in Abb. 1 wurde dankenswerterweise von Lena Schnee erstellt, die zu diesem Zeitpunkt als studentische Hilfskraft im DFG-Projekt ‚Mittelniederdeutsche Grammatik: Flexionsmorphologie I. Nominalflexion‘ tätig war. Die Darstellung orientiert sich an den für die Analyse des Referenzkorpus Altdeutsch genutzten chronographischen Karten (vgl. Donhauser/Zeige 2019). Texte des ReN, deren Entstehungszeit auf mehrere Jahrhunderthälften fällt (s. in den Metadaten z. B. die Angabe 13/2-14/1 für einen um 1300 entstandenen Text), wurden in Abb. 1 jeweils dem späteren Zeitraum zugeordnet.

Während Urkunden und Verwaltungstexte im Zeitraum zwischen der 1. Hälfte des 14. Jahrhunderts und der 1. Hälfte des 16. Jahrhunderts annähernd gleichmäßig erfasst werden konnten, zeichnet sich für die Tokenzahl in den Rechtstexten ein klarer Schwerpunkt in der Zeit zwischen 1250 und 1350 ab (193.908 Token, davon 170.046 annotiert). Die Literatur (38 Texte mit 407.605 Token, davon 337.728 annotiert) ist erst ab der 1. Hälfte des 15. Jahrhunderts stärker vertreten, ebenso Texte der Wissensvermittlung (30 Texte mit 493.966 Token, davon 282.345 annotiert) und der Geistlichen Schriftlichkeit (36 Texte mit 597.408 Token, davon 312.444 annotiert). Mit einem zusätzlichen Blick auf den Sprachraum fällt auf, dass die meisten geistlichen Texte aus Westfalen stammen (16 Texte mit 266.615 Token, davon 89.505 annotiert). Diese Textverteilung im ReN ermöglicht es, in einigen Textsorten, z. B. den Urkunden, die diachrone Entwicklung und diatopische Verteilung eines sprachlichen Phänomens abzubilden. Für alle Sprachräume sind Urkunden aus mindestens einer wichtigen städtischen Kanzlei von 1301 bis 1500 erfasst worden. Auch die Gruppe der Stadtrechte (Braunschweig, Bremen, Duisburg, Goslar, Hamburg, Herford, Hildesheim, Kalkar, Kleve, Lübeck, Magdeburg, Rügen, Soest und Stade) sowie innerhalb der wissensvermittelnden Texte die Gruppe der Chroniken bieten gute Möglichkeiten für diachrone und diatopische Vergleiche. Mithilfe der Metadaten kann im ReN gezielt nach einem Phänomen in diesen Textsorten gesucht und eine entsprechende Analyse angeschlossen werden.

Grammatische Analysen mit dem ReN werden durch die den Texten hinzugefügten Annotationen wesentlich unterstützt. Das Korpus ist mithilfe von ANNIS (vgl. Krause/Zeldes 2016), einem Instrument zur Visualisierung und Analyse linguistisch annotierter Daten,<sup>6</sup> durchsuchbar. Hierfür können die verschiedenen Ebenen im Bereich des transkribierten Textes (z. B. *token*, *tok\_dipl* und *tok\_anno*) sowie sämtliche Annotationsebenen wie *lemma* (Lemma nach dem MNWB, aber ohne Homonymendifferenzierung), *pos* (Wortart) und *morph* (Flexionsform) genutzt werden. Die Ergebnisse einer Suchabfrage lassen sich mitsamt den Metadaten exportieren und in anderen Programmen weiterverarbeiten und ggf. um weitere Annotationen anreichern.

Da die Token im ReN nicht eigens nach ihrer Lexembildungsstruktur annotiert worden sind, ist für Lexembildungsanalysen eine generalisierte Suche nach beispielsweise allen Affixen als Bestandteilen von Lexembildungsprodukten nicht möglich. Stattdessen können jedoch mittels kombinierter Suchabfragen auf der

---

<sup>6</sup> Eine Benutzungsanleitung für ANNIS findet sich unter <http://korpling.github.io/ANNIS/3.6/user-guide/> (16.08.2021).

<sup>7</sup> Eine Erläuterung der Unterschiede zwischen den einzelnen Tokenisierungsebenen enthält das ReN-Handbuch unter [https://www.fdr.uni-hamburg.de/record/9195/files/ReN-Handbuch\\_2021-01-30.pdf?download=1](https://www.fdr.uni-hamburg.de/record/9195/files/ReN-Handbuch_2021-01-30.pdf?download=1) (16.08.2021).

Lemmaebene und der *pos*-Ebene einzelne Affixe wie *-schop* und seine Varianten ermittelt werden.<sup>8</sup> Hierbei ist zu berücksichtigen, dass derartige Suchabfragen auch False Positives enthalten können (im Fall von *-schop* z. B. das Lexem *bischof*), um welche die Daten bereinigt werden müssen. Für die Analyse können je nach Fragestellung zudem bestimmte weitere Annotationen notwendig sein, so z. B. die differenzierte Angabe der Varianten eines Suffixes wie im Fall von *-schop*.

### 3 Das Derivationsuffix *-schop* im Mittelniederdeutschen – Forschungsstand

Das Suffix *-schop* tritt im Mittelniederdeutschen durch seine Variantenvielfalt besonders hervor. Neben frequenten *-schop* und *-schap* stehen als weniger frequente Formen *-schup* und *-schep*, mit hochdeutschem Konsonantenstand auch *-schaf* und *-schafft*. Die wesentlichen Informationen zum Derivationsuffix *-schop* und seinen Varianten sind Einzelstudien, Überblicksartikeln zur Lexembildung im Mittelniederdeutschen (vgl. Cordes 1985, Cordes/Niebaum 2000) sowie den älteren mittelniederdeutschen Grammatiken von Lübben (1882), Colliander (1912), Lasch (1914) und Sarauw (1924) zu entnehmen.<sup>9</sup> Da Letztere einen starken Fokus auf Phonologie und Flexion legen, sind nur vereinzelt Angaben zur Verbreitung des hier untersuchten Suffixes und seiner Varianten zu finden. Colliander (1912: 146, § 268) liefert hierzu keine Informationen, sondern führt *-schap* und *-schop* nebeneinander im Kapitel zur Deklination der Feminina auf. Lübben (1882: 12, § 13) weist lediglich auf einen Wechsel von *u* und *o* im Suffix *-schup* hin und begründet diese Varianz mit der schwächeren Differenzierung der Vokalqualitäten in unbetonten Endsilben, vor allem in der gesprochenen Sprache. Sarauw (1924: 55) nennt *-schop* und *-schup* als lautlich jüngere Varianten gegenüber *-schap*; für das ältere und nur vereinzelt gebrauchte *-schep* führt er zwei Belege aus Goslar und Bremen an. Nur am Rande erwähnt er die Varianten *-schaf* und *-schafft*, die er als hochdeutsch klassifiziert (vgl. Sarauw 1924: 56). Zur raum-zeitlichen Verbreitung der Suffixvarianten gibt Lasch (1914: 118, § 213) an, dass *-schap* insbesondere im früheren Mittelniederdeutschen noch häufiger vorkommt. Während *-schop* für den gesamten mittelniederdeutschen Raum gilt, findet sich *-schup* vor allem im Ostfälischen und im Westen des nordniedersächsischen Sprachgebietes, daneben aber auch in Hamburg, Kiel, Göttingen und in Westfalen (vgl. Lasch 1914: 118, § 213 sowie 16, § 14).

Im Fokus der Einzelstudien zu mittelniederdeutschen Derivationsuffixen stehen die Beschreibung der Lexembildung unter besonderer Berücksichtigung der

<sup>8</sup> Vgl. z. B. die folgende Suchabfrage, mit der alle Substantive (NA) mit einem auf *-schap*, *-schop*, *-schup* oder *-schep* endenden Lemma ermittelt werden: lemma=/.+sch(a|o|u|e)p/\_=\_ pos="NA".

<sup>9</sup> Für das ältere Hochdeutsche liegt neben zahlreichen Einzeluntersuchungen zum Suffix *-schafft* (vgl. z. B. Messing 1917, Stoltenberg 1929, Piltz 1951, Meineke 1991, Döring/Eichler 1996, Brendel et al. 1997) auch eine korpusbasierte Beschreibung im Rahmen der neuen Mittelhochdeutschen Grammatik von Klein/Solms/Wegera (2009: 114-118, 172-182) vor.



Derivationsbasis und der Semantik (vgl. Möhn/Schröder 2003: 43-44, Hoge 1912: 36-40), aber auch Aspekte des Ursprungs (vgl. Ahlsson 1988). Umfassende Untersuchungen zum Gebrauch der Suffixvarianten im Mittelniederdeutschen und zu möglichen Einflussfaktoren wie Zeit und Raum fehlen bisher. Möhn/Schröder (2003: 44) machen in ihrer detaillierten Beschreibung von *-schap*, *-schop* und *-schup* keine Angaben zur zeitlichen oder regionalen Verbreitung. Für die Varianten *-schaft* und *-schaf* führen sie an, dass diese vorwiegend in Texten mit hochdeutschem Einfluss verwendet werden (vgl. Möhn/Schröder 2003: 44).

Zur Diachronie der Suffixvarianten findet sich in der Forschungsliteratur der Hinweis, dass *-schap* vor allem im älteren Mittelniederdeutschen vorkommt (vgl. Peters 1987: 84). Cordes/Niebaum (2000: 1465) halten fest, dass ab dem Ende des 14. Jahrhunderts die Variante *-schop* die stärkste Frequenz aufweist. Grunewald (1944: 181) macht darauf aufmerksam, dass der Gebrauch der Varianten *-schap*, *-schop* und *-schup* nicht ausschließlich zeitlich bedingt ist, sondern dass auch der Sprachraum eine gewisse Rolle spielt. In einer Untersuchung westfälischer, ostfälischer und nordniedersächsischer Urkunden des 14. und 15. Jahrhunderts belegt er, dass *-schop* insgesamt im Verlauf des untersuchten Zeitraums an Frequenz zunimmt und etwa ab der zweiten Hälfte des 14. Jahrhunderts *-schap* ablöst, dass jedoch das Verhältnis zwischen diesen beiden Varianten in den drei betrachteten Räumen unterschiedlich ausfällt (vgl. Grunewald 1944: 181). Dahlberg (1962: 24) führt in Ergänzung zu Grunewald (1944) weitere Belege auf und ordnet sie verschiedenen Sprachräumen bzw. Kombinationen von Sprachräumen zu, wobei er auch das Mittelniederländische berücksichtigt. Dass *-schap* aufgrund seiner Konzentration auf frühere Zeiträume nicht zwingend die ältere Variante darstellt, begründet Grunewald (1944: 181-182) damit, dass auch *-schop* in seinem Material bereits ab dem 14. Jahrhundert belegt ist. Korlén (1945: 230) liefert sogar Belege für *-schop* sowie *-schup* in einzelnen Quellen des 13. Jahrhunderts.

Eine weitere Quelle zum Gebrauch der Suffixvarianten stellt neben den Einzelstudien der ‚Atlas spätmittelalterlicher Schreibsprachen des niederdeutschen Altlandes und angrenzender Gebiete‘ (ASnA) von Peters/Fischer/Nagel (2017) dar, der in Karte 69 (S. 113) das Vorkommen der Varianten für hd. *-schaft* im Untersuchungszeitraum vom Beginn des 13. bis zum Ende des 15. Jahrhunderts abbildet. Insgesamt zeigt die Darstellung eine eher heterogene Verteilung in den gewählten Ortspunkten. Die Variante *-schap* konzentriert sich vor allem auf das Westfälische und das mittelniederländisch-mittelniederdeutsche Übergangsgebiet, zu dem beispielsweise Groningen zählt. In späterer Zeit, insbesondere ab dem 15. Jahrhundert ist in den südlichen und östlichen Gebieten des Westfälischen und den östlich davon liegenden Sprachräumen die Variante *-schop* vorherrschend, während sich im äußersten Westen *-schap* länger hält. Die Variante *-schup* ist auf einzelne Orts- und Zeitpunkte insbesondere des Ostfälischen begrenzt; lediglich in Bremen dominiert sie ab 1351 zu fast allen Zeitpunkten. Nur vereinzelt sind die Varianten *-schep* (Osnabrück, Utrecht), *-schef* (Köln), *-schafft* (Köln, Magdeburg) und *-skēp* (Utrecht) belegt. Die Suffixvariante *-schaf* konzentriert sich sehr stark auf Köln, wo sie von 1301

bis 1400 durchgehend dominiert; sie ist außerhalb von Köln nur vereinzelt belegt (z. B. in Paderborn). Diese das niederdeutsche Altland betreffenden Beobachtungen zur Verbreitung des Suffixes *-schop* und seiner Varianten sollen im Folgenden mithilfe der Daten des ReN in Quellen anderer Textsorten überprüft sowie um Informationen zu den ostniederdeutschen Sprachräumen ergänzt werden.

## 4 Zur Varianz des Suffixes *-schop* im Mittelniederdeutschen

### 4.1 Methodische Anmerkungen

Um Aussagen zum Gebrauch des Suffixes *-schop* und seiner Varianten sowie zu potentiellen Einflussfaktoren zu gewinnen, wurden die Belege für Lexembildungen mit dem Suffix und seinen Varianten im ReN in der Version 1.1 sowohl im annotierten als auch im lediglich transkribierten Teil des Korpus ermittelt. Im annotierten Korpus konnte für die Suchabfrage die Annotationsebene *lemma* (Lemma nach dem MNWB, aber ohne Homonymendifferenzierung) herangezogen und nach Lemmata auf *-schop* und Varianten gesucht werden. Im transkribierten Teil war die Suche auf die Tokenebene beschränkt, weshalb komplexe Suchabfragen unter Berücksichtigung aller potentiellen graphematischen Varianten des Suffixes und seiner Flexionsformen formuliert werden mussten. Sämtliche Belege wurden mitsamt den relevanten Metadaten (Text, Zeit, Sprachraum, Textgruppe) sowie dem sprachlichen Kontext exportiert und in einem Tabellenkalkulationsprogramm weiterbearbeitet. In einer ersten Sichtung wurden alle False Positives entfernt (z. B. das Lemma *bischop*). Anschließend wurden für die Belege aus dem transkribierten Teil des Korpus die entsprechenden Lemmata wie auf der Annotationsebene *lemma\_wsd* (Lemma mit Homonymendifferenzierung) nachannotiert. Für sämtliche Belege wurde zudem die tatsächlich realisierte Suffixvariante (z. B. *-schap*, *-schop*, *-schup* etc.) annotiert. Abzüglich der False Positives liegen der Untersuchung insgesamt 1.785 Belege für insgesamt 86 unterschiedliche Lemmata zugrunde, von denen 27 nur ein einziges Mal belegt sind.

Für eine Betrachtung des Vorkommens der Suffixvarianten in Abhängigkeit verschiedener außer- und innersprachlicher Faktoren wurde in einem ersten Zugriff, der möglichst objektiv erfolgen sollte, mithilfe von Random Forests ermittelt, welche Parameter einen stärkeren Einfluss als andere ausüben. Im Anschluss wurde für sämtliche potentiellen Faktoren das Vorkommen mithilfe von additiven Säulendiagrammen genauer betrachtet. Auf die Ergebnisse dieser unterschiedlichen Analysen soll im Folgenden genauer eingegangen werden.

## 4.2 Hierarchisierung möglicher Einflussfaktoren durch Random Forests

Das statistische Modell der Random Forests wird in den vergangenen Jahren zunehmend für korpuslinguistische Untersuchungen herangezogen (vgl. insbesondere Tagliamonte/Baayen 2012 sowie Fonteyn/Nini 2020 und Gries 2020; vgl. außerdem die Arbeiten von Hartmann 2018, Zeige 2019, Dücker 2020 und Schmitt 2020; einführend zu Random Forests für korpuslinguistische Analysen vgl. Levshina 2015: 291-300). Für historische Sprachdaten mit ihren besonderen Eigenschaften wie geringen Belegzahlen für einzelne Phänomene, einer Vielzahl potentieller Einflussfaktoren (unabhängiger Variablen) sowie einer starken Varianz bei den Phänomenen und folglich einer großen Menge an Ausprägungen einer abhängigen Variable (vgl. Zeige 2019: 45) bietet gerade dieses Verfahren eine geeignete Möglichkeit der statistischen Auswertung, wie Zeige (2019) in seiner Studie zur Präpositionalgruppe im Althochdeutschen belegt.<sup>10</sup> Statt die Signifikanz möglicher Einflussfaktoren einzeln zu prüfen, können mit einem multivariaten Verfahren wie den Random Forests mehrere Parameter gleichzeitig auf ihre Relevanz hin untersucht und entsprechend hierarchisiert werden (vgl. Zeige 2019: 251). Die Random Forests bauen auf dem Modell der Conditional Inference Trees auf, das die statistische Signifikanz durch die wiederholte zufällige Anordnung der Daten und den Abgleich der Vorhersagekraft der Zufallsanordnung einerseits und der tatsächlichen Verteilung im Datenset andererseits ermittelt (vgl. Levshina 2015: 292). Auf diese Weise bestimmen Conditional Inference Trees, ob ein Zusammenhang zwischen einer oder mehreren der unabhängigen Variablen und der abhängigen Variable besteht. Das Datenset wird immer an der Variablen mit dem stärksten Zusammenhang zwischen dieser unabhängigen Variable und der abhängigen Variable geteilt (daher die Baumstruktur), dies wird so oft wiederholt, bis keine signifikante Variable mehr übrig ist (vgl. Levshina 2015: 291; 2020: 612). Auf der Basis mehrerer Conditional Inference Trees werden die Random Forests gebildet, mit denen für jede unabhängige Variable der entsprechende Wert des statistischen Zusammenhangs ermittelt wird (vgl. Levshina 2015: 292).

In der vorliegenden Untersuchung wurden Random Forests eingesetzt, um vor einer detaillierten Analyse des Variantengebrauchs zunächst in einem objektiven Zugriff den unterschiedlich starken Einfluss verschiedener Faktoren auf die Verteilung der Suffixvarianten zu ermitteln. Berücksichtigt wurden die Parameter Zeit, Sprachraum und Textgruppe (entspricht im ReN dem Feld der Schriftlichkeit). Die Annotationen für diese Metadaten wurden aus dem ReN 1.1 übernommen (zur Korpusstruktur vgl. oben, Kap. 2). Abb. 2 liefert die Werte für das im Statistikprogramm R mithilfe des Pakets ‚party‘ (vgl. Hothorn et al. 2006, Strobl et al. 2007, Strobl et al. 2008) ermittelte Relevanzmaß der untersuchten Parameter. Die Reihenfolge der Werte für die verschiedenen unabhängigen Variablen entspricht der Reihenfolge in der Relevanz der Variablen. Im vorliegenden Fall haben vor allem die Zeit (time) mit einem Wert von 0,182 und der

---

<sup>10</sup> Zu Kritik am Einsatz des Modells, unter anderem die Rechenleistung betreffend, vgl. Zeige (2019: 252-254).

Sprachraum (*language\_area*) mit 0,140 einen Einfluss auf das Vorkommen der Suffixvarianten, die Relevanz des Faktors Textgruppe (*topic\_ReN*) fällt mit einem Wert von 0,071 im Vergleich dazu deutlich geringer aus.

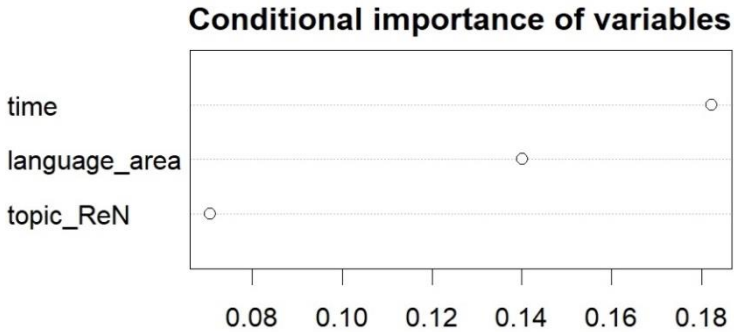


Abbildung 2: Relevanzmaß der Faktoren Zeit (*time*; 0,182), Sprachraum (*language\_area*; 0,140) und Textgruppe (*topic\_ReN*; 0,071)<sup>11</sup>

## 5 Der Einfluss außer- und innersprachlicher Faktoren auf die Verteilung der Suffixvarianten

Im Folgenden soll ausführlicher auf den sich hier andeutenden stärkeren Einfluss von Zeit und Sprachraum eingegangen werden, bevor anschließend ergänzend die Textgruppe sowie als weitere potentielle Faktoren der Text und das Lemma betrachtet werden.

### 5.1 Zur zeitlichen Verteilung der Suffixvarianten

Wie die statistische Analyse mithilfe der Random Forests nahelegt, wird der Gebrauch der Varianten von *-schop* vor allem durch den Zeitraum bestimmt. Wie genau sich dieser

<sup>11</sup> *seed* = 287, *ntree* = 1.000, *mtry* = 2; *accuracy* = 0.859944. Beim *seed* handelt es sich um eine zufällig ausgewählte Ziffer, die als Startpunkt für die Berechnungen der Random Forests dient. Auf Basis des *seed* wird ein Vektor von Ziffern erstellt, deren Verteilung einer zufälligen Verteilung ähnelt, jedoch algorithmisch, also reproduzierbar ist. *ntree* bezeichnet die Zahl der gebildeten Bäume. Je höher diese Zahl ausfällt, desto zuverlässiger sind das Modell und die berechneten Werte für das Relevanzmaß (vgl. Strobl/Malley/Tutz 2009: 343). In der Forschung werden üblicherweise 1.000 Bäume angesetzt (vgl. Levshina 2015: 297, Zeige 2019: 49). Der Wert *mtry* wird in der Regel als Wurzel aus der Anzahl unabhängiger Variablen bestimmt (vgl. Levshina 2015: 297). Der Wert von 0.859944 für die *accuracy* belegt eine sehr hohe Verlässlichkeit des Modells. Er besagt, dass 86% der Datenpunkte im vorliegenden Testset korrekt vorhergesagt wurden. Im Vergleich dazu beträgt bei einem auf Zufall beruhenden Modell der Wert bei sechs Ausprägungen lediglich ein Sechstel, d. h. 17%. In einem Modell, das durchgängig die frequenteste Ausprägung (hier *-schop*) wählt, liegt der Wert ebenfalls nur bei ca. 43%.

Einfluss entfaltet, das heißt, wie sich die einzelnen Varianten auf die verschiedenen Zeiträume verteilen, soll im Folgenden intensiver beleuchtet werden. Berücksichtigt werden dafür sämtliche mittelniederdeutschen Sprachräume; das Niederrheinische wird aus der Betrachtung ausgeschlossen. In Kap. 5.2 wird es in Ergänzung zum Mittelniederdeutschen am Rand mit betrachtet.

Abb. 3 zeigt, dass insgesamt sechs Varianten des Derivationssuffixes im Korpus belegt sind: *-schaf*, *-schaft*, *-schap*, *-schep*, *-schop* und *-schup*. Die Angabe der Zeiträume in Jahrhunderten, unterteilt in jeweils zwei Hälften, richtet sich nach den Metadaten im ReN 1.1. Wie in der Darstellung zur Textverteilung in Kap. 2 wurde auch hier bei nicht genauer datierbaren Texten mit einer Angabe wie *13/2-14/1* eine Zuordnung zum jeweils spätesten Zeitraum vorgenommen.

Deutlich seltener als *-schaf* und *-schop* erscheint *-schup* mit insgesamt 166 Belegen im Korpus. Eine besondere zeitliche Entwicklung im Gebrauch der Variante ist nicht erkennbar. *-schup* kommt in der zweiten Hälfte des 13. Jahrhunderts und der ersten Hälfte des 14. Jahrhunderts nur vereinzelt vor und weist von der zweiten Hälfte des 14. Jahrhunderts bis zur ersten Hälfte des 16. Jahrhunderts nahezu konstant einen Anteil von um die zehn Prozent auf. Ab der zweiten Hälfte des 16. Jahrhunderts ist es nicht mehr belegt.

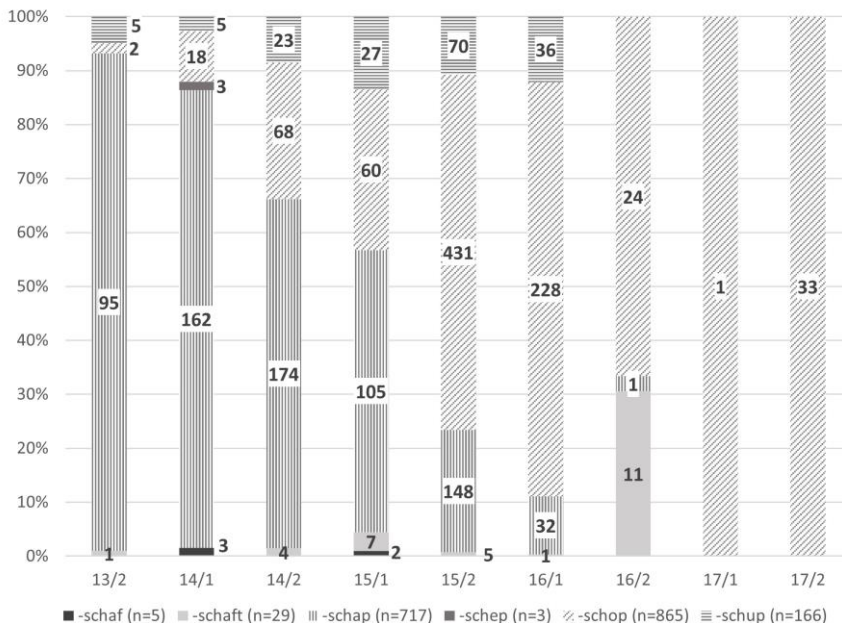


Abbildung 3: Zeitliche Verteilung der Suffixvarianten (n = 1.785)<sup>12</sup>

<sup>12</sup> Die zeitliche Verteilung der Suffixvarianten ist statistisch signifikant:  $\chi^2 = 882,83$ ,  $df = 40$ ,  $p < 0,001$ . Der Effekt ist mit einem Wert von 0,3145 für Cramers V mittelstark.

Die Variante *-schep* stellt mit nur drei Belegen, die sämtlich aus der Originalhandschrift des ‚Bremer Stadtrechts‘ von 1303/04 stammen, eine Randerscheinung dar.<sup>13</sup> Auch *-schaf* mit fünf und *-schaft* mit 29 Belegen sind im Korpus nur marginal vertreten. *-schaf* wird in fünf Texten der jeweils ersten Hälfte des 14. und 15. Jahrhunderts aus vier verschiedenen Sprachräumen verwendet, ist also nicht auf einen Zeit- oder Sprachraum oder auf eine bestimmte Quelle begrenzt. *-schaft* ist bis zur ersten Hälfte des 16. Jahrhunderts lediglich vereinzelt belegt und weist in der zweiten Hälfte des 16. Jahrhunderts einen höheren Anteil auf. Sämtliche Belege aus diesem Zeitraum stammen aus ‚Dre kortwilige Historien‘, drei in einem Hamburger Druck überlieferten Erzählungen, die Übersetzungen hochdeutscher Vorlagen darstellen.<sup>14</sup> Der vergleichsweise höhere Anteil an der Variante *-schaft* in der zweiten Hälfte des 16. Jahrhunderts ist daher auf den Einfluss des Hochdeutschen zurückzuführen. Dass die Variante *-schaft* nicht ausschließlich in hochdeutsch beeinflussten Quellen auftritt, zeigt sich in den Belegen vor der zweiten Hälfte des 16. Jahrhunderts, von denen mehrere aus Texten ohne hochdeutsche Vorlage, z. B. aus Urkunden, stammen.

Um zu prüfen, ob die hier dargestellte zeitliche Entwicklung im Gebrauch der Suffixvarianten nicht nur auf das Mittelniederdeutsche insgesamt zutrifft, sondern auch auf die einzelnen Sprachräume, wurde zusätzlich für jeden Raum die zeitliche Verteilung der Varianten betrachtet. Dabei offenbart sich zunächst das in Kap. 2 angesprochene Problem der ungleichen Überlieferungslage. Während für das Lübische, Nordniedersächsische, Ostelbische, Ostfälische und Westfälische eine ausreichend große Belegmenge zugrunde liegt, sind die Belegzahlen im Baltischen, Elbostfälischen und Südmärkischen vergleichsweise gering, vor allem in einzelnen Zeitabschnitten, weshalb Aussagen zur zeitlichen Verteilung der Varianten für diese Räume kaum möglich sind. In den Sprachräumen mit größerer Belegmenge wird das oben für das gesamte Mittelniederdeutsche gezeichnete Bild im Wesentlichen bestätigt, wobei sich jedoch je nach Sprachraum die Variante *-schop* zu unterschiedlichen Zeiten gegenüber *-schap* durchsetzt. Im Ostfälischen dominiert *-schop* bereits ab der ersten, im Nordniedersächsischen ab der zweiten Hälfte des 15. Jahrhunderts. Im Lübischen und Ostelbischen fällt die Verteilung heterogener aus. Hier überwiegt der Anteil der Variante *-schop* bereits ab der zweiten Hälfte des 14. Jahrhunderts, liegt jedoch in einem jeweils später liegenden Zeitabschnitt unter dem Anteil für *-schap*. Dies verdeutlicht, dass sich der diachrone Wechsel beider Suffixvarianten keineswegs in allen Räumen so glatt vollzieht, wie in Abb. 3 suggeriert wird. Im Gegensatz zu den übrigen Sprachräumen weist das Westfälische bis zur zweiten Hälfte des 15. Jahrhunderts durchgehend einen sehr hohen Anteil an der Variante *-schap* auf. Ab der ersten Hälfte des 16. Jahrhunderts dominiert hier *-schop*, daneben

<sup>13</sup> Korlén (1945: 34) weist *-schep* daneben auch in einem Zusatz zum ‚Goslarer Kramerrecht‘ vom Ende des 13. Jahrhunderts nach.

<sup>14</sup> Zu den Vorlagen und der Art der Bearbeitung der drei einzelnen Texte vgl. Beckers (1977: 22f.). Einen Vergleich der mittelniederdeutschen und der frühneuhochdeutschen Fassung des ‚Hörnen Sifrit‘ liefert Recker (2021).

ist aber *-schap* noch immer vergleichsweise frequent. Im Westfälischen scheint sich *-schap* damit stärker zu halten als im Rest des mittelniederdeutschen Sprachraums. Ob die sich hier bereits andeutende Relevanz des Sprachraums auch in der alle Zeitabschnitte zusammenfassenden räumlichen Verteilung der Suffixvarianten zutage tritt, wird im folgenden Kapitel näher beleuchtet.

## 5.2 Zur räumlichen Verteilung der Suffixvarianten

Wie das in Kap. 4.2 ermittelte Relevanzmaß zeigt, kommt neben dem Faktor Zeit auch dem Parameter Sprachraum eine gewisse Relevanz bei der Verteilung der Suffixvarianten zu. Abb. 4 veranschaulicht den Anteil aller Varianten in den mittelniederdeutschen Sprachräumen. Auf den ersten Blick sind zwei Gruppen zu unterscheiden: 1.) das Nordniedersächsische, Lübische, Ostelbische, Baltische und Ostfälische, in denen *-schop* dominiert, und 2.) das Südmärkische, Westfälische und Elbstfälische, wo *-schap* die frequenteste Variante darstellt.

Ein genauerer Blick auf die Verteilung der Belege in den einzelnen Sprachräumen offenbart, wie diese Gruppierung zustande kommt und wie sie zu interpretieren ist. Wie im vorangehenden Kapitel belegt wurde, besteht eine deutliche Korrelation zwischen der Entstehungszeit und der Verteilung der Suffixvariante: *-schap* ist nur bis zur ersten Hälfte des 15. Jahrhunderts dominant und wird anschließend von *-schop* verdrängt. Um nun zu prüfen, ob ungleiche Belegzahlen in früheren und späteren Zeitabschnitten die Verteilung der Varianten in bestimmten Sprachräumen beeinflussen, wurde die Summe der Belege bis zur ersten Hälfte des 15. Jahrhunderts mit der Summe für die sich anschließenden Zeiträume verglichen. Während im Westfälischen die Summe in beiden Abschnitten ähnlich hoch ausfällt, liefern das Elbstfälische und das Südmärkische für die Zeiträume bis zur ersten Hälfte des 15. Jahrhunderts, in denen die Variante *-schap* insgesamt stark vertreten ist, deutlich mehr Belege als für die folgenden Zeitabschnitte, in denen insgesamt im Korpus die Wahrscheinlichkeit für *schop*-Belege höher ist. Die Dominanz von *-schap* im Elbstfälischen und Südmärkischen ist daher auf diesen Zusammenhang zwischen der Belegzahl in den einzelnen Zeiträumen und der diachronen Verteilung der Varianten zurückzuführen. Im Westfälischen hingegen ist *-schap* unabhängig von der Beleglage insgesamt wesentlich frequenter als in den übrigen Sprachräumen. Von den fünf Sprachräumen, in denen *-schop* die häufigste Variante bildet, weisen das Nordniedersächsische und das Ostfälische einen annähernd ausgewogenen Anteil an Belegen bis zur ersten Hälfte des 15. Jahrhunderts und ab der zweiten Hälfte des 15. Jahrhunderts auf. Das Baltische, Lübische und Ostelbische hingegen enthalten deutlich mehr Belege im späteren Zeitabschnitt, weshalb sich die Verteilung in Abb. 4 hier aus dem Einfluss der Zeit in Verbindung mit der spezifischen Beleglage ergibt. Ein auf den Sprachraum zurückzuführender stärkerer Gebrauch von *-schop* ist somit lediglich für das Nordniedersächsische und das Ostfälische zu konstatieren.

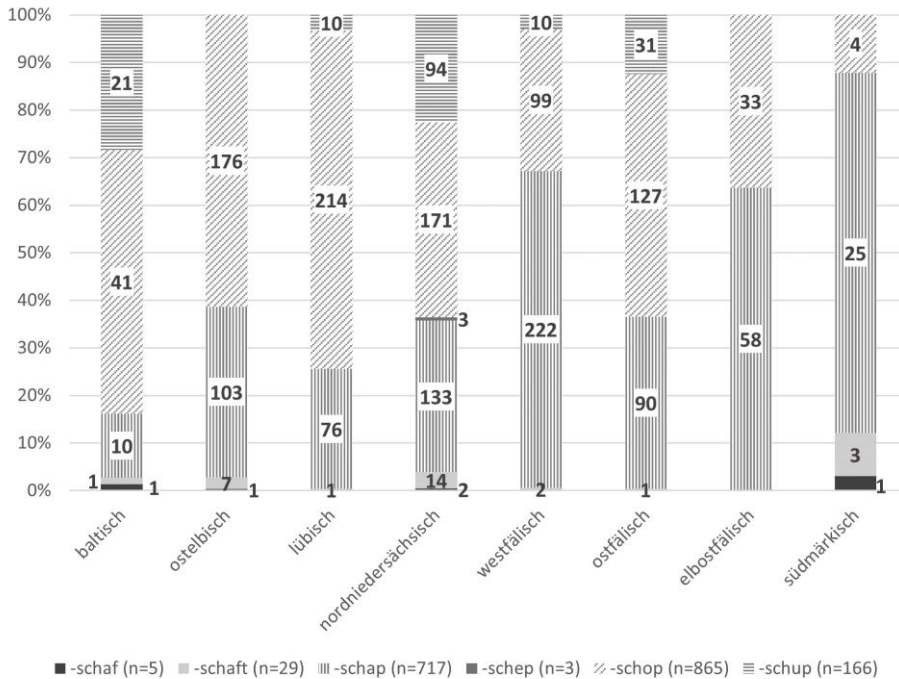


Abbildung 4: Räumliche Verteilung der Suffixvarianten (n = 1.785)<sup>15</sup>

Hinsichtlich des Niederrheinischen, das in den Auswertungen zur Varianz im Mittelniederdeutschen nicht abgebildet ist, sei ergänzend zum bisherigen Bild zu erwähnen, dass es einen deutlich höheren Anteil an der Variante *-schap* als die mittelniederdeutschen Gebiete aufweist. Da im Niederrheinischen die Summe der Belege bis zur ersten Hälfte des 15. Jahrhunderts und ab der zweiten Hälfte des 15. Jahrhunderts annähernd gleich hoch ausfällt, ist eine Verzerrung durch unterschiedliche Belegzahlen in den Zeiträumen auszuschließen.

Ein Blick auf die übrigen Varianten in Abb. 4 zeigt, dass *-schep* auf das Nordniedersächsische und dort, wie in Kap. 5.1 erwähnt wurde, auf einen einzelnen Text beschränkt ist. Die Variante *-schup* ist im Ostelbischen, Südmärkischen und Elbstfälischen nicht und im Lübischen und Westfälischen lediglich marginal belegt. Nur im Ostfälischen, Nordniedersächsischen und Baltischen kommt sie vergleichsweise häufiger vor, im Baltischen dabei sogar öfter als die Variante *-schap*. Im Ostfälischen stammt der überwiegende Teil der Belege für *-schup* aus der zweiten Hälfte des 15. Jahrhunderts, vor allem aus der ‚Gandersheimer Reimchronik‘, deren einziges überliefertes Manuskript diesem Zeitraum zuzuordnen ist. Mehrfach ist es außerdem in

<sup>15</sup> Die räumliche Verteilung der Suffixvarianten ist statistisch signifikant:  $\chi^2 = 432,39$ ,  $df = 35$ ,  $p < 0,001$ . Der Effekt ist mit einem Wert von 0,2201 für Cramers V mäßig stark.



den ‚Hildesheimer Urkunden‘ von 1401–1450 sowie 1451–1500 und im ‚Mirakel‘ von Arnt Buschmann in der Wolfenbütteler Handschrift belegt. Im Nordniedersächsischen stellt sich die zeitliche Verteilung mit Belegen für *-schup* von der zweiten Hälfte des 13. Jahrhunderts bis zur ersten Hälfte des 16. Jahrhunderts wesentlich ausgewogener dar. Die Belege stammen aus verschiedenen Stadtrechten (vor allem aus der ‚Bilderhandschrift des Hamburger Stadtrechts‘ von 1497 und aus dem ‚Bremer Stadtrecht‘ von 1428), einem Stadtbuch, einem kirchlichen Text und Urkunden (vor allem aus den ‚Bremer Urkunden‘ in den Abschnitten von 1351–1500). Es ist somit keine Konzentration auf eine bestimmte Quelle zu erkennen, aber auf kanzleisprachliche Rechts- und Verwaltungstexte. Im Baltischen stammen sämtliche Belege für *-schup* aus den ‚Revaler Schragen‘ von 1501–1550. Der frequentere Gebrauch der Variante *-schup* im Ostfälischen, Nordniedersächsischen und Baltischen ist daher vor allem auf Präferenzen in bestimmten Texten und auch Textgruppen zurückzuführen.

Die Variante *-schaf* wird im Nordniedersächsischen, Ostelbischen, Baltischen und Südmärkischen verwendet und konzentriert sich, wie in Kap. 5.1 gezeigt wurde, nicht auf einen bestimmten Zeitraum oder Text. Auch *-schafft* kommt in unterschiedlichen Quellen vor und ist in allen mittelniederdeutschen Sprachräumen außer dem Elbstfälischen belegt. Es findet sich, wie Kap. 5.1 darlegt, nicht ausschließlich in hochdeutsch beeinflussten Texten.

Insgesamt zeigt sich somit ein Einfluss des Sprachraums für das Westfälische und für das Niederrheinische mit einem vergleichsweise hohen Anteil der Variante *-schap* sowie für das Nordniedersächsische und das Ostfälische mit einem höheren Vorkommen an *-schop*. Bei *-schaf* und *-schafft* ist keine Korrelation mit dem Sprachraum erkennbar. Der höhere Anteil an *-schup* im Ostfälischen, Nordniedersächsischen und Baltischen wurde mit der dort beobachteten Konzentration der Variante auf bestimmte Quellen und Textgruppen erklärt. Die Beobachtungen Laschs zur räumlichen Verbreitung von *-schup* lassen sich daher nicht bestätigen. Der sich hier andeutende Einfluss weiterer Faktoren, der laut der Variablenziehung anhand der Random Forests (s. Abb. 2) geringer als die Relevanz von Zeit und Sprachraum ausfällt, soll im folgenden Kapitel eingehender geprüft werden.

### 5.3 Textgruppe, Text und Lemma als weitere potentielle Einflussfaktoren

Da zwischen der oben untersuchten Entstehungszeit und der Textgruppe insofern ein gewisser Zusammenhang besteht, als sich bestimmte Textgruppen hinsichtlich eines diachronen Wandels konservativer als andere verhalten (vgl. Ihden 2020), ist zu vermuten, dass auch dem Faktor Textgruppe eine gewisse Rolle bei der Verteilung der Suffixvarianten zukommt. Tatsächlich zeigen sich auf den ersten Blick in einigen Textgruppen sehr stark abweichende Frequenzen, die jedoch auf den zweiten Blick in den meisten Fällen durch die Beleglage motiviert sind. Den höchsten Anteil an der Variante *-schop* weisen die Bereiche Wissensvermittlung und Literatur

mit einem enormen Übergewicht an Belegen ab der zweiten Hälfte des 15. Jahrhunderts auf. Ein vergleichsweise starker Gebrauch von *-schap* ist in den Bereichen Recht sowie Private Schriftlichkeit und Korrespondenz auszumachen. Die Belege für Private Schriftlichkeit und Korrespondenz konzentrieren sich sehr stark auf drei Quellen der ersten und zweiten Hälfte des 15. Jahrhunderts, die jeweils deutlich entweder *-schap* oder *-schop* präferieren; hier kommt also dem einzelnen Text ein stärkerer Einfluss als der Textgruppe zu. Bei den Rechtstexten ist ein Übergewicht an Belegen bis zur ersten Hälfte des 15. Jahrhunderts festzustellen, was mit der im gesamten ReN feststellbaren Dominanz älterer Quellen in dieser Textgruppe (vgl. Kap. 2) zusammenhängt. Auffällig ist zudem, dass sämtliche ab der zweiten Hälfte des 15. Jahrhunderts belegten Bildungen mit *-schap* aus dem ‚Stendaler Sachsenspiegel‘ und der zugehörigen Glosse stammen, sodass sich auch hier ein gewisser Einfluss des einzelnen Textes manifestiert. Die einzige Textgruppe, die im ReN über mehrere Sprachräume hinweg annähernd ausgewogen vertreten ist, sind die Urkunden, sodass eine textgruppenspezifische Verwendung der Suffixvarianten, sofern vorhanden, hier am ehesten unbeeinflusst von den Faktoren Zeit und Raum beobachtet werden kann. Hinsichtlich der Verteilung der Varianten in den Urkunden fällt das in Kap. 5.2 erwähnte vergleichsweise starke Vorkommen von *-schup* auf. Einen im Vergleich dazu geringeren Anteil weisen die Bereiche Recht, Religion und Literatur auf; in den übrigen Textgruppen ist *-schup* nicht oder nur vereinzelt belegt. Während die Belege in den Bereichen Recht und Religion verschiedenen Quellen zuzuordnen sind, stammen die Belege für die Literatur lediglich aus einem Text, der ‚Gandersheimer Reimchronik‘.

Die Analyse des Variantengebrauchs in den einzelnen Texten ergibt zunächst, dass im überwiegenden Teil der Texte (128 von 188) ausschließlich eine Variante vorkommt. Die meisten der 128 Quellen spiegeln in ihrer Präferenz die diachrone Verteilung von *-schap* und *-schop* wider, das heißt, Texte bis zur ersten Hälfte des 15. Jahrhunderts enthalten in der Regel *-schap*, spätere Texte hingegen *-schop*. Viele der davon abweichenden Texte stammen aus dem 15. Jahrhundert, wo die Variation am stärksten ist und sich der Wechsel von *-schap* zu *-schop* vollzieht. Innerhalb der früheren und späteren Quellen fallen die Reimbearbeitung der ‚Disticha Catonis‘ von Meister Stephan (baltisch, 14/1-14/2, Wissensvermittlung), die ausschließlich *-schop* verwendet, sowie der ‚Verlorene Sohn‘ des Borchardt Waldis von 1527 (baltisch, 16/1, Literatur), in dem ausschließlich *-schap* vorkommt, auf. Diese beiden Texte verhalten sich ausgehend von der diachronen Entwicklung progressiver (‚Disticha Catonis‘) bzw. konservativer (‚Verlorener Sohn‘).

Von den 60 Quellen mit Variation im Gebrauch des Derivationssuffixes gehören 20 zu den Stadtbüchern und Urkunden, bei denen mit mehreren Schreibern zu rechnen ist und die Variation auf die Präferenz der verschiedenen Schreiber zurückzuführen sein könnte. Bei den übrigen 40 Texten ist in den meisten Fällen eine deutliche Dominanz einer Variante zu erkennen, die durch die Entstehungszeit des Textes vor dem Hintergrund der beschriebenen diachronen Entwicklung begründet

ist. Von dieser Beobachtung weicht innerhalb der mindestens zehn Belege liefernden Quellen vor allem die ‚Liudger-Vita‘ von Johannes Cincinnius aus dem Jahr 1512 (westfälisch, 16/1, Religion) ab, die überwiegend Bildungen mit *-schap* enthält. Hier zeichnet sich die in Kap. 5.1 formulierte Tendenz ab, dass westfälische Texte länger an der Variante *-schap* festhalten. Eine auffällige Verteilung zeigen zudem die ‚Bilderhandschrift des Hamburger Stadtrechts‘ von 1497 (nordniedersächsisch, 15/2, Recht), das ‚Bremer Stadtrecht‘ von 1428 (nordniedersächsisch, 15/1, Recht), die ‚Gandersheimer Reimchronik‘ (ostfälisch, 15/2, Literatur) und die ‚Bekentnisse van beyden Sacramenten‘ von Bernd Rothmann 1533 (westfälisch, 16/1, Religion) mit einer hohen Frequenz der Variante *-schup*. Der Suffixgebrauch in diesen Quellen unterschiedlicher Entstehungszeiten, Sprachräume und Textgruppen weist auf eine individuelle Präferenz der Variante *-schup* bei bestimmten Schreibern hin. Die stärkste Variation beinhaltet das ‚Bremer Stadtrecht‘ von 1303/04 in der Originalhandschrift (nordniedersächsisch, 14/1, Recht): Es enthält neben jeweils acht Belegen für *-schap* und *-schop* auch vier Bildungen mit *-schup* und drei mit *-schep*. In der Abschrift des ‚Bremer Stadtrechts‘ von 1303/04 ist diese starke Variation aufgelöst; neben *-schap*, das die deutliche Mehrheit der Belege stellt, findet sich lediglich noch *-schop*. Dieser Vergleich von Original und Abschrift unterstreicht deutlich die Rolle des Schreibers bei der Wahl der Suffixvariante.

Die Relevanz des jeweiligen Lemmas beim Gebrauch der Suffixvariante ist im Wesentlichen als gering zu bewerten. Von den 49 Lemmata, die häufiger als einmal vorkommen, ist der überwiegende Teil (41) mit unterschiedlichen Varianten belegt,<sup>16</sup> was nicht auf einen stark lemmaspezifischen Gebrauch schließen lässt. Um ein mögliches Zusammenwirken der Faktoren Lemma und Text zu prüfen, wurde die in den Lemmata auftretende Suffixvariation in der oben erwähnten Gruppe von 40 Texten mit Variation, die nicht zu den Stadtbüchern oder Urkunden gehören, genauer untersucht. Dabei wurden lediglich Quellen mit mindestens zehn Belegen berücksichtigt (25 Texte). Innerhalb dieser 25 Texte erscheint die Mehrheit der Lemmata mit nur einer Suffixvariante; in den meisten Fällen ist pro Text nur ein Lemma mit Variation vorhanden, die sich wiederum in der Regel auf zwei Varianten beschränkt. Einen verhältnismäßig größeren Anteil an Lemmata mit Suffixvariation zeigen die ‚Bilderhandschrift des Hamburger Stadtrechts‘ von 1497, in der fünf von den sieben mehrfach belegten Lemmata zwischen *-schop* und *-schup* variieren, und ‚Der Leyen Doctrinal‘, wo drei der sechs mehrfach vorkommenden Lemmata sowohl mit *-schap* als auch *-schop* erscheinen. Den Gegensatz hierzu bilden vor allem ‚De denscke kroneke‘, die Abschrift des ‚Bremer Stadtrechts‘ von 1303/04 sowie das ‚Bremer Stadtrecht‘ von 1428, in denen die Lemmata jeweils nur in einer Suffixvariante belegt sind. Insgesamt stellt die Variation innerhalb eines Lemmas und eines Textes somit eher eine Randerscheinung dar, die sich in stärkerem Ausmaß auf einzelne Quellen konzentriert.

---

<sup>16</sup> Von diesen 41 Lemmata erscheinen 16 mit zwei Varianten, weitere 16 mit drei, acht mit vier und ein Lemma (*bödeschop*) mit sogar fünf Varianten.

Den 41 Lemmata mit Variation im Korpus stehen die acht ausschließlich mit einer Variante verwendeten Lemmata *juncvrûw(e)~*<sup>17</sup> (zehnmal *-schop*), *börg(e)~* (siebenmal *-schop*), *drunken~* (fünfmal *-schap*), *vulkömen~* (viermal *-schap*), *unvrünt~* (viermal *-schop*), *gôgrêve~* (dreimal *-schap*), *kôpman~* (dreimal *-schap*) und *vrôt~* (zweimal *-schop*) gegenüber. Im Fall von *börg(e)~*, *drunken~*, *vulkömen~* und *vrôt~* ist das Vorkommen auf jeweils einen Text beschränkt. Die Wahl der Variante lässt sich daher hier nicht deutlich auf einen Einflussfaktor zurückführen, sie könnte durch die insgesamt stärkeren Einflussfaktoren Zeit und Sprachraum (die im Faktor Einzeltext Niederschlag finden) bedingt sein. Die Lemmata *juncvrûw(e)~*, *unvrünt~*, *gôgrêve~* und *kôpman~* sind zwar in verschiedenen Quellen ohne eine Konzentration auf einen Zeit- oder Sprachraum belegt, jedoch ist bei genauerer Betrachtung auch hier ein Einfluss des Lemmas eher auszuschließen: Die beiden mit *-schop* verwendeten Lemmata *juncvrûw(e)~* und *unvrünt~* stammen aus Quellen ab der zweiten Hälfte des 15. Jahrhunderts; die mit *-schap* vorkommenden Lemmata *gôgrêve~* und *kôpman~* finden sich vor allem in früheren Texten bis zur ersten Hälfte des 15. Jahrhunderts sowie im oben erwähnten ‚Stendaler Sachsenspiegel‘ mit seiner starken Präferenz der Variante *-schap*, sodass davon auszugehen ist, dass sich auch bei diesen Lemmata der Einfluss des Zeitraums sowie des individuellen Textes spiegelt. Auch ein Einfluss weiterer mit dem Lemma zusammenhängender sprachinterner Faktoren wie der Vokale und des Auslauts der Derivationsbasis konnten im Korpus nicht festgestellt werden.

## 6 Fazit

Eine variationssensitive Darstellung grammatischer Strukturen ist auf die Existenz eines strukturierten Korpus als Datengrundlage angewiesen. Dabei zeigt sich, dass für detaillierte Analysen von raum-zeitlichen Teilkorpora deren relative Größe zu berücksichtigen ist, die Textverteilung innerhalb des Korpus also in jedem Fall in die Diskussion der Ergebnisse einzubeziehen ist.

Für die Analyse zur Variation des Suffixes *-schop* wurde zunächst mithilfe von Random Forests das Relevanzmaß verschiedener Parameter ermittelt, die anschließend intensiver betrachtet wurden. Der starke Zusammenhang zwischen der Entstehungszeit und der Verteilung der Suffixvarianten, auf den auch in der Forschungsliteratur hingewiesen wird, konnte in den tiefergehenden Analysen belegt und konkretisiert werden. Von den beiden frequentesten Suffixvarianten konzentriert sich *-schap* auf die früheren Zeiträume und wird ab der zweiten Hälfte des 15. Jahrhunderts durch *-schop* abgelöst. Dieser diachrone Übergang von *-schap* zu *-schop* lässt sich auch in einzelnen Sprachräumen, im Lübschen, Nordniedersächsischen, Ostelbischen, Ostfälischen und Westfälischen,

<sup>17</sup> Die Lemmata werden hier nicht in der Form des MNWB dargestellt, das die Lemmata je nach Beleglage mit unterschiedlicher Suffixvariante angesetzt hat, sondern stattdessen mit einer Tilde an der Stelle des Derivationsuffixes.

die eine ausreichend große Belegmenge liefern, nachzeichnen. Dabei setzt sich allerdings die Variante *-schop* in den verschiedenen Sprachräumen zu unterschiedlichen Zeiten durch.

Bei der Untersuchung der räumlichen Verteilung musste für jeden Sprachraum die individuelle Beleglage in den unterschiedlichen Zeitabschnitten berücksichtigt werden, um eine indirekte Beeinflussung durch die Diachronie auszuschließen. Unter Berücksichtigung dieses Aspekts konnte eine Relevanz des Sprachraums nachgewiesen werden, die sich im Westfälischen und darüber hinaus im Niederrheinischen (das hier nicht näher behandelt werden konnte) in einem vergleichsweise hohen Vorkommen der Variante *-schop* sowie im Nordniedersächsischen und Ostfälischen in einem stärkeren Gebrauch von *-schop* manifestiert. Die Angabe von Lasch (1914: 118, § 213 sowie 16, § 14), dass *-schop* vor allem im Ostfälischen und im Westen des nordniedersächsischen Sprachgebietes begegnet, konnte hingegen durch die quantitativen Analysen nur eingeschränkt bestätigt werden, denn im Korpus konzentriert sich die Variante im Ostfälischen und Nordniedersächsischen wie auch im baltischen Niederdeutschen stark auf einzelne Quellen und Textgruppen.

Während sowohl für die Textgruppe als auch das Lemma kein relevanter Einfluss auf die Verteilung der Suffixvarianten festgestellt werden konnte, besteht ein nicht zu vernachlässigender Zusammenhang zwischen dem einzelnen Text und dem Gebrauch der Suffixvarianten, der sich unter anderem in einer besonders stark ausgeprägten Variation oder in einer deutlichen Präferenz einer Variante niederschlagen kann. Im Fall des ‚Bremer Stadtrechts‘ von 1303/04 hat der Vergleich von Originalhandschrift und Abschrift zudem aufgezeigt, welche Rolle dem Schreiber bei der Wahl der Suffixvariante zukommt.

Das methodische Vorgehen in der vorliegenden Studie hat verdeutlicht, dass eine rein quantitative Ermittlung der Verteilung von Varianten und die statistische Prüfung von Korrelationen zwischen unabhängigen und abhängigen Variablen nicht ausreicht, um die Ursachen für den Gebrauch von Varianten aufzudecken. Statistische Verfahren wie die Random Forests, die einen objektiven Blick auf potentielle Zusammenhänge ermöglichen, müssen stets um tiefergehende Analysen am Datenmaterial ergänzt werden. Dabei ist unter anderem zu prüfen, ob sich der Einfluss einer Variable in einer anderen spiegelt, um die tatsächliche Relevanz der betrachteten unabhängigen Variable korrekt bewerten zu können.

Die hier erfolgte Untersuchung leistet damit zugleich einen inhaltlichen wie auch einen methodischen Beitrag zu einer Darstellung der Lexembildung im Mittelniederdeutschen.

## Literaturverzeichnis

### Quellen

Die folgenden Quellen sind dem ReN entnommen.

- Bilderhandschrift des Hamburger Stadtrechts. Hamburg 1497. Staatsarchiv der Freien und Hansestadt Hamburg: Senat Cl. VII Lit. LaNr. 2 Vol. 1 c.
- Waldis, Borchardt: De Parabell vam vorlorn Szohn [...]. Magdeburg 1527: Heinrich Öttinger. Herzog August Bibliothek Wolfenbüttel: M: Lo 7812.
- Bremer Stadtrecht (Codex III). Bremen 1428. Staatsarchiv Bremen: 2-P.5.b.2.a.3.
- Bremer Stadtrecht, Abschrift. 1303/04. Staatsarchiv Bremen: 2-P.5.b.2.a.2.
- Bremer Stadtrecht, Originalhandschrift. 1303/04. Staatsarchiv Bremen: 2-P.5.b.2.a.1.
- Bremer Urkunden aus dem ASnA, 10 Urkunden von 1351–1400, 2 Urkunden von 1401–1450, 8 Urkunden von 1451–1500. Bremen.
- Buschmann, Amt: Mirakel, Wolfenbütteler Hs. 1476. Herzog August Bibliothek Wolfenbüttel: Cod. 1180 Helmst.
- Cicinnius, Johannes: Liudger-Vita. 1512. Münster, Landesarchiv Nordrhein-Westfalen, Abteilung Westfalen (LA NRW, Abt. WF): Altertumsverein Münster Msc. 136.
- Gandersheimer Reimchronik des Priesters Eberhard. Herzog August Bibliothek Wolfenbüttel: Cod. 503 Helmst.
- Hildesheimer Urkunden aus dem ASnA. 2 Urkunden von 1401–1450, 8 Urkunden von 1451–1500. Hildesheim.
- De Hörnen Sifrit, De Rese Sigenot, Köninck Laurin. Dre kortwilige Historien. Hamburg um 1560: Joachim Loew. Staatsbibliothek zu Berlin: Yf 7864 (5 in).
- Der Leyen Doctrinal. Lüneburg 15. Jh. Herzog August Bibliothek Wolfenbüttel: Cod. 127a Blankenburg.
- Revaler Handwerkerschragen. 4 Schragen von 1501–1550. Reval (Tallinn). National Archives of Estonia/Tallinn City Archives: REV6: TLA.190.2.663; REV14: TLA.190.2.668; REV2: TLA.190.2.473; REV7: TLA.190.2.664.
- Rothmann, Bernhard: Bekentnisse van beyden Sacramenten / Doepe unde Nachtmale / der predicanten tho / Munster. Münster 1533. Münster, Universitäts- und Landesbibliothek (ULB): F+3 3009.
- Sachsenspiegel, Stendaler Druck (nach Berl. Druck). Stendal 1488. Staatsbibliothek zu Berlin: 4° Inc 2083.
- Saxo Grammaticus: De densce kronke. Lübeck 1490: Matthäus Brandis. Kopenhagen, Kgl. Bib.: Pal. 5a 4°.
- Meister Stephan: Disticha Catonis, Reimbearbeitung, Dorpat um 1365. Bibliothek der Polnischen Akademie der Wissenschaften Danzig (BGPAN): Ms. 2416.

## Forschungsliteratur

- Ahlsson, Lars Erik: Mnd. -schap (-schop, -schup). Herkunft eines Suffixes. In: Korrespondenzblatt des Vereins für niederdeutsche Sprachforschung 95 (1988), S. 61-64.
- Barteld, Fabian/Ihden, Sarah/Dreessen, Katharina/Schröder, Ingrid: HiNTS: A Tagset for Middle Low German. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation (2018), S. 3940-3945. Online unter: <http://www.lrec-conf.org/proceedings/lrec2018/pdf/870.pdf> [Stand: 16.08.2021].
- Beckers, Hartmut: Mittelniederdeutsche Literatur – Versuch einer Bestandsaufnahme. In: Niederdeutsches Wort 17 (1977), S. 1-58.
- Brendel, Bettina/Frisch, Regina/Moser, Stephan/Wolf, Norbert Richard: Wort- und Begriffsbildung in frühneuhochdeutscher Wissensliteratur. Substantivische Affixbildung. Wiesbaden 1997 (Wissensliteratur im Mittelalter. 26).
- Colliander, Elof: Mittelniederdeutsches Elementarbuch. Photokopie der Druckfahnen des nicht zur Veröffentlichung gelangten Werkes, das mit dem § 365 abbricht. Heidelberg 1912.
- Cordes, Gerhard: Wortbildung des Mittelniederdeutschen. In: Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung. Hrsg. von Werner Besch/Oskar Reichmann/Stefan Sonderegger. Berlin/New York 1985 (Handbücher zur Sprach- und Kommunikationswissenschaft. 2.2), S. 1243-1247.
- Cordes, Gerhard/Niebaum, Hermann: Wortbildung des Mittelniederdeutschen. In: Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung. Hrsg. von Werner Besch/Anne Betten/Oskar Reichmann/Stefan Sonderegger. 2. Aufl. Berlin/New York 2000 (Handbücher zur Sprach- und Kommunikationswissenschaft. 2.2), S. 1463-1469.
- Dahlberg, Torsten: Mittelniederdeutsche Suffixabstrakta. Lexikalische und wortgeographische Randbemerkungen. Göteborg 1962 (Göteborgs Germanistische Forschungen. 6).
- Donhauser, Karin/Zeige, Lars Erik: Bildgebende Verfahren zur Analyse des Referenzkorpus Altdeutsch: Das Konzept der chronographischen Karte. In: *atbe in palice, atbe in anderu sumenuelicheru stedī*. Raum und Sprache. Festschrift für Elvira Glaser zum 65. Geburtstag. Hrsg. von Andreas Nievergelt/Ludwig Rübekeil unter Mitarb. von Andi Gredig. Heidelberg 2019 (Germanistische Bibliothek. 66), S. 73-82.
- Döring, Brigitte/Eichler, Birgit: Sprache und Begriffsbildung in Fachtexten des 16. Jahrhunderts. Wiesbaden 1996 (Wissensliteratur im Mittelalter. 24).

- Dücker, Lisa: Großschreibung von Personenbezeichnungen in Hexenverhörprotokollen – eine Analyse von Gender, Lemma und Rolle im Gerichtsprozess. In: Kolloquium Forschende Frauen 2019. Beiträge Bamberger Nachwuchswissenschaftlerinnen. Hrsg. von Astrid Schütz/Renata Szczepaniak/Mona Hess. Bamberg 2020, S. 31-50.
- Fonteyn, Lauren/Nini, Andrea: Individuality in syntactic variation: An investigation of the seventeenth-century gerund alternation. In: *Cognitive Linguistics* 31,2 (2020), S. 279-308.
- Ganslmayer, Christine/Müller, Peter O./Ring, Uli: Wortbildung in der deutschen Urkundensprache des 13. Jahrhunderts. Möglichkeiten, Probleme und Grenzen ihrer Erforschung. In: *Zeitschrift für deutsche Philologie* 122 (2003), S. 321-358.
- Gries, Stefan Th.: On classification trees and random forests in corpus linguistics: Some words of caution and suggestions for improvement. In: *Corpus Linguistics and Linguistic Theory* 16,3 (2020), S. 617-647.
- Grimm, Jacob: Deutsche Grammatik. Zweiter Theil. Göttingen 1826.
- Grunewald, Gottfried: Die mittelniederdeutschen Abstraktsuffixe. Kopenhagen/Lund 1944 (Lunder Germanistische Forschungen. 13).
- Hartmann, Stefan: Divergent theories, converging evidence: The constructional semantics of competing future constructions. In: *Yearbook of the German Cognitive Linguistics Association (GCLA)* 6 (2018), S. 91-114.
- Henzen, Walter: Deutsche Wortbildung. 3. Aufl. Tübingen 1965.
- Hoge, John: Die produktiven Abstraktsuffixe des Mittelniederdeutschen. Kahl 1912.
- Hothorn, Torsten/Bühlmann, Peter/Dudoit, Sandrine/Molinaro, Annette/Van Der Laan, Mark J.: Survival Ensembles. In: *Biostatistics* 7,3 (2006), S. 355-373. Online unter: <http://dx.doi.org/10.1093/biostatistics/kxj011> [Stand: 16.08.2021].
- Ihden, Sarah: Relativsätze im Mittelniederdeutschen. Korpuslinguistische Untersuchungen zu Struktur und Gebrauch. Berlin/Boston 2020 (*Lingua Historica Germanica*. 23).
- Katara, Pekka: Das Diminutivum bei Johannes Veghe. In: *Annales Academiae Scientiarum Fennicae* 84 (1954), S. 599-627.
- Klein, Thomas: Zum Schwund des Präfixes *ge-* im Frühmittelniederdeutschen. In: *Westfälische vnde sächsische Festgabe für Robert Peters zum 60. Geburtstag*. Hrsg. von Robert Damme/Norbert Nagel. Bielefeld 2004, S. 89-99.
- Klein, Thomas: Zur Dimensionierung historischer Textkorpora. In: *Sprache und Datenverarbeitung. International Journal for Language Data Processing* 31,1-2 (2007), S. 15-22.



- Klein, Thomas/Dipper, Stefanie: Handbuch zum Referenzkorpus Mittelhochdeutsch. Bochum 2016 (Bochumer Linguistische Arbeitsberichte. 19). Online unter: <https://linguistics.rub.de/forschung/arbeitsberichte/19.pdf> [Stand: 16.08.2021].
- Klein, Thomas/Solms, Hans-Joachim/Wegera, Klaus-Peter: Mittelhochdeutsche Grammatik. Teil III: Wortbildung. Tübingen 2009.
- Kleymann, Verena/Nagel, Norbert/Peters, Robert: Die digitale Lemmaliste für das Mittelniederdeutsche im DFG-Projekt „Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650)“. In: Korrespondenzblatt des Vereins für niederdeutsche Sprachforschung 122 (2015), S. 95-100.
- Kluge, Friedrich: Deutsche Sprachgeschichte. Werden und Wachsen unserer Muttersprache von ihren Anfängen bis zur Gegenwart. 2. Aufl. Leipzig 1925.
- Korlén, Gustav: Die mittelniederdeutschen Texte des 13. Jahrhunderts. Beiträge zur Quellenkunde und Grammatik des Frühmittelniederdeutschen. Lund/Kopenhagen 1945 (Lunder germanistische Forschungen. 19).
- Krause, Thomas/Zeldes, Amir: ANNIS3: A new architecture for generic corpus query and visualization. In: Digital Scholarship in the Humanities 31/2 (2016), S. 118-139. Online unter: <http://dsh.oxfordjournals.org/content/31/1/118> [Stand: 16.08.2021].
- Lasch, Agathe: Mittelniederdeutsche Grammatik. Tübingen 1914 (2. Aufl. 1974/ Nachdruck 2011) (Sammlung kurzer Grammatiken germanischer Dialekte. A Hauptreihe. 9).
- Lasch, Agathe: Vom Werden und Wesen des Mittelniederdeutschen. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 51 (1925), S. 55-76. Wieder abgedruckt in: Agathe Lasch: Ausgewählte Schriften zur niederdeutschen Philologie. Hrsg. von Robert Peters/Timothy Sodmann. Neumünster 1979, S. 232-253.
- Latendorf, Friedrich: Die Deminutiva der niederdeutschen Ausgabe von Agricola's Sprichwörtern. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 3 (1878), S. 101-103.
- Lemnitzer, Lothar/Zinsmeister, Heike: Korpuslinguistik. Eine Einführung. 2. Aufl. Tübingen 2010.
- Levshina, Natalia: How to do Linguistics with R. Data exploration and statistical analysis. Amsterdam 2015.
- Levshina, Natalia: Conditional Inference Trees and Random Forests. In: A Practical Handbook of Corpus Linguistics. Hrsg. von Magali Paquot/Stefan Th. Gries. Cham 2020, S. 611-643.
- Listen, Paul H.: Middle Low German adverbial *-liken*, *-likes*, *-lik*. In: Interdigitations. Essays for Irmengard Rauch. Hrsg. von Gerald F. Carr. New York/Frankfurt a. M. 1999, S. 295-304.

- Lübben, August: Mittelniederdeutsche Grammatik. Nebst Chrestomathie und Glossar. Leipzig 1882.
- Lübben, August/Walther, Christoph: Mittelniederdeutsches Handwörterbuch. Norden 1888.
- Macha, Jürgen/Topalović, Elvira/Hille, Iris/Nolting, Uta/Wilke, Anja: Deutsche Kanzleisprache in Hexenverhörprotokollen der Frühen Neuzeit. Bd. 1-2. Berlin/New York 2005.
- Mähl, Stefan: Studien zum mittelniederdeutschen Adverb. Köln/Weimar/Wien 2004 (Niederdeutsche Studien. 49).
- Meineke, Birgit: Althochdeutsche -scaf(t)-Bildungen. Göttingen 1991 (Studien zum Althochdeutschen. 17).
- Messing, E. E. J.: Das Suffix *-schaft* (Eng. *-ship*, Nl. *-schap*) nach Ursprung und Entwicklung; *Wirtschaft, Wirt.* In: Neophilologus 2 (1917), Teil I: S. 185-190, Teil II: S. 272-283.
- MNWB = Mittelniederdeutsches Handwörterbuch. Bd. 1 ff. Hrsg. von Agathe Lasch/Conrad Borchling/Gerhard Cordes/Dieter Möhn/Ingrid Schröder. Neumünster 1956 ff.
- Möhn, Dieter/Schröder, Ingrid: Vorstudien zu einer mittelniederdeutschen Grammatik I. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 126 (2003), S. 7-51.
- Möhn, Dieter/Schröder, Ingrid: Sprachbedarf und Lexembildung am Beispiel der Grammatik des Mittelniederdeutschen. In: Beiträge zur Morphologie. Germanisch, Baltisch, Ostseefinnisch. Hrsg. von Hans Fix-Bonner. Odense 2007, S. 227-258.
- Möhn, Dieter/Schröder, Ingrid: Lexembildung im Aufriss einer Grammatik des Mittelniederdeutschen. Das Adjektiv als Exempel. In: Low Saxon Dialects across Borders – Niedersächsische Dialekte über Grenzen hinweg. Hrsg. von Alexandra N. Lenz/Charlotte Gooskens/Siemon Reker. Stuttgart 2009 (ZDL-Beihefte. 138), S. 38-59.
- Möller, Robert: *Euphrosina kolerin, Beckhin vonn Painnten, die Berndt bonesche und andere beclagtinnen*. Feminin-Movierung von Appellativen und Namen in Hexenverhörprotokollen des 16./17. Jahrhunderts. In: Deutsch im 17. Jahrhundert. Studien zu Sprachkontakt, Sprachvariation und Sprachwandel. Gedenkschrift für Jürgen Macha. Hrsg. von Markus Denkler/Stephan Elspaß/Dagmar Hüpper/Elvira Topalović. Heidelberg 2017 (Sprache – Literatur und Geschichte. Studien zur Linguistik/Germanistik. 46), S. 129-159.
- Paul, Hermann: Deutsche Grammatik. Bd. 5, Teil 5: Wortbildungslehre. Halle a. d. S. 1920.
- Peters, Robert: Katalog sprachlicher Merkmale zur variablenlinguistischen Erforschung des Mittelniederdeutschen, Teil 1. In: Niederdeutsches Wort 27 (1987), S. 61-93.

- Peters, Robert: Die Diaglierung des Mittelniederdeutschen. In: Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung. Hrsg. von Werner Besch/Anne Betten/Oskar Reichmann/Stefan Sonderegger. 2. Aufl. Berlin/New York 2000 (Handbücher zur Sprach- und Kommunikationswissenschaft. 2.2), S. 1478-1490.
- Peters, Robert: Das „Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650)“. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 140 (2017), S. 35-42.
- Peters, Robert/Fischer, Christian/Nagel, Norbert: Atlas spätmittelalterlicher Schreibsprachen des niederdeutschen Altlandes und angrenzender Gebiete (ASnA). Band I: Einleitung, Karten; Band II: Verzeichnis der Belegtypen; Band III: Verzeichnis der Schreibformen und der Textzeugen (Ortspunktdokumentation). Berlin/Boston 2017.
- Piltz, Günter: Die Bedeutungsentwicklung der Substantiva auf -heit, -schaft und -tum. Diss. Hamburg 1951.
- Recker, Anabel: *De Hörnen Sijfrt*. Textvergleich der mittelniederdeutschen und der frühneuhochdeutschen Fassung mit Vorüberlegungen zu einer Edition des Hamburger Druckes (um 1560). In: Studien zur mittelniederdeutschen und frühneuhochdeutschen Sprache und Literatur. Hrsg. von Sarah Ihden/Katharina Dreessen/Robert Langhanke. Hildesheim/Zürich/New York 2021 (Kleine und regionale Sprachen. 6), S. 189-207.
- Roolfs, Friedel Helga: *Anna Bergmanns* und *Maria Witten*. Parentale Femininmovierung von Familiennamen in westfälischen Varietäten. In: Bäuerliche Familiennamen in Westfalen. Hrsg. von Friedel Helga Roolfs. Münster 2016, S. 57-70.
- Sarauw, Christian: Niederdeutsche Forschungen. Bd. 2. Die Flexionen der mittelniederdeutschen Sprache. Kopenhagen 1924 (Historisk-filologiske meddelelser. 10,1).
- Schmitt, Eleonore: „Lange dachte man, das deutsche Problem sei ein besonders hoher Anteil kinderloser Frauen“ – Eine korpusbasierte Analyse der Verwendungskontexte des Adjektivs *kinderlos*. In: Kolloquium Forschende Frauen 2019. Beiträge Bamberger Nachwuchswissenschaftlerinnen. Hrsg. von Astrid Schütz/Renata Szczepaniak/Mona Hess. Bamberg 2020, S. 151-178.
- Schröder, Edward: Die Nomina agentis auf -ster. (Zu Jahrb. XLVII S. 42). In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 48 (1922), S. 1-8.
- Schröder, Ingrid: Stand und Perspektiven mittelniederdeutscher Grammatikographie am Beispiel der expliziten Derivation. In: Beiträge zur historischen Wortbildung des Mittelniederdeutschen. Hrsg. von Markus Denkler/Stefan Mähl. Köln/Wien 2022 (Niederdeutsche Studien. 61), S. 13-36.

- Schröder, Ingrid/Barteld, Fabian/Dreessen, Katharina/Ihden, Sarah: Historische Sprachdaten als Herausforderung für die manuelle und automatische Annotation: Das ‚Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650)‘. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 140 (2017), S. 43-57.
- Seelmann, Wilhelm: Niederdeutsche Diminutive auf -el. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 45 (1919), S. 18-21.
- Seelmann, Wilhelm: Altsächsische und mittelniederdeutsche Diminutive. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 46 (1920), S. 51-57.
- Seelmann, Wilhelm: Mittelnd. und neund. Infinitive mit ge-Präfix. In: Korrespondenzblatt des Vereins für niederdeutsche Sprachforschung 41 (1927/28), S. 35-37.
- Stoltenberg, H. L.: Die Bezeichnung von Gruppen durch die Nachsilbe schaft. In: Zeitschrift für die gesamte Staatswissenschaft 86,1 (1929), S. 136-142.
- Strobl, Carolin/Boulesteix, Anne-Laure/Zeileis, Achim/Hothorn, Torsten: Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. In: BMC Bioinformatics 8/25 (2007). Online unter: <https://doi.org/10.1186/1471-2105-8-25> [Stand: 16.08.2021].
- Strobl, Carolin/Boulesteix, Anne-Laure/Kneib, Thomas/Augustin, Thomas/Zeileis, Achim: Conditional Variable Importance for Random Forests. In: BMC Bioinformatics 9/307 (2008). Online unter: <https://doi.org/10.1186/1471-2105-9-307> [Stand: 16.08.2021].
- Strobl, Carolin/Malley, James/Tutz, Gerhard: An Introduction to Recursive Partitioning: Rationale, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests. In: Psychological Methods 14,4 (2009), S. 323-348.
- Tagliamonte, Sali A./Baayen, R. Harald: Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. In: Language Variation and Change 24 (2012), S. 135-178.
- Walther, Christoph: Niederdeutsche Adjektive auf -ern. In: Korrespondenzblatt des Vereins für niederdeutsche Sprachforschung 31 (1910), S. 38-39.
- Werth, Alexander: *Gretie Dwengers*, genannt *die Dwengersche*. Formale und funktionale Aspekte morphologischer Sexusmarkierung (Movierung) in norddeutschen Hexenverhörprotokollen der Frühen Neuzeit. In: Jahrbuch des Vereins für niederdeutsche Sprachforschung 138 (2015), S. 53-76.
- Wilmanns, Wilhelm: Deutsche Grammatik. Gotisch, Alt-, Mittel- und Neuhochdeutsch. Abt. 2: Wortbildung. 2. Aufl. Straßburg 1899.
- Zeige, Lars Erik: Die Präpositionalgruppe im ältesten Deutsch. Korpus, Grammatik, Darstellung. Habilitationsschrift, Humboldt-Universität zu Berlin 2019.

## Online-Ressourcen

ANNIS, Userguide: <http://korpling.github.io/ANNIS/3.6/user-guide/> [Stand: 16.08.2021].

Barteld, Fabian/Dreessen, Katharina/Ihden, Sarah/Schröder, Ingrid: Handbuch zur Nutzung des Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650). 2021. [https://www.fdr.uni-hamburg.de/record/9195/files/ReN-Handbuch\\_2021-01-30.pdf?download=1](https://www.fdr.uni-hamburg.de/record/9195/files/ReN-Handbuch_2021-01-30.pdf?download=1) [Stand: 16.08.2021].

Referenzkorpus Mittelniederdeutsch/Niederrheinisch 1200–1650, Projektwebseite: <https://www.slm.uni-hamburg.de/ren.html> [Stand: 16.08.2021].

ReN-Team: Reference Corpus Middle Low German/Low Rhenish (1200–1650); Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650). Version 1.1. 2021. <https://doi.org/10.25592/uhhfdm.1668> [Stand: 16.08.2021].