

Zeitschrift für Wortbildung / Journal of Word Formation

eine Open-Access-Zeitschrift /
an Open Access Journal

1 / 2018

herausgegeben von / edited by

Petra M. Vogel, Elke Donalies, Ludwig M. Eichinger,
Mechthild Habermann, Jörg Meibauer,
Barbara Schlücker, Hans-Joachim Solms,
Pavol Štekauer und Salvador Valera Hernández



PETER LANG

Zeitschrift für Wortbildung
Journal of Word Formation

2018 • Jahrgang / Volume 2 • Heft / Issue 1

Zeitschrift für Wortbildung Journal of Word Formation

Wortbildung – historisch, mehrsprachig,
kontrastiv

Hrsg. von / Edited by

Peter Gilles
Heinz Sieburg

**Petra M. Vogel • Elke Donalies • Ludwig M. Eichinger
• Mechthild Habermann • Jörg Meibauer
• Barbara Schlücker • Hans-Joachim Solms
• Pavol Štekauer • Salvador Valera Hernández (eds.)**

Wissenschaftlicher Beirat / Advisory Board: Werner Abraham (University of Vienna & Munich University) • Aleksandra Bagasheva (Sofia University) • Irmhild Barz (University of Leipzig) • Geert Booij (University of Leiden) • Jan Čermák (Charles University Prague) • Gerrit J. Dimmendaal (University of Cologne) • Jan Don (University of Amsterdam) • Nanna Fuhrhop (University of Oldenburg) • Livio Gaeta (University of Turin) • Luise Kempf (Mainz University) • Lívía Körtvélyessy (Pavol Jozef Šafárik University Košice) • Elisabeth Leiss (Munich University) • Hans Christian Luschützky (University of Vienna) • Francesca Masini (University of Bologna) • Franz Rainer (Vienna University of Economics and Business) • Anna Thornton (University of L'Aquila) • Carola Trips (University of Mannheim) • Kristel van Goethem (Catholic University of Louvain) • Martina Werner (Austrian Center for Digital Humanities)

„Zeitschrift für Wortbildung / Journal of Word Formation“ (ZJWF) ist eine internationale Open-Access-Zeitschrift mit einem Double-blind-Begutachtungsverfahren. Das Spektrum der Zeitschrift umfasst Wortbildungsphänomene in allen Sprachen und in allen linguistischen Bereichen, z. B. Morphologie, Syntax, Lexikologie, Phonologie, Semantik, Pragmatik, Sprachgeschichte, Typologie, Dialektologie, Spracherwerb und Sprachkontakt. Die Zeitschrift erscheint online mit zwei Ausgaben pro Jahr. Sie enthält Artikel, Rezensionen und allgemeine Informationen wie z. B. Tagungsankündigungen. Sonderhefte zu wichtigen Themen der Wortbildung werden in unregelmäßigen Abständen erscheinen. Manuskripte können in Deutsch, Englisch, Französisch und Spanisch eingereicht werden.

Manuskripte sind an die federführende Herausgeberin unter petra.vogel@uni-siegen.de zu senden.

Redaktion: Dr. Carolin Baumann, Universität Siegen, Fakultät I, Germanistisches Seminar, Hölderlinstraße 3, D-57076 Siegen, Tel.: +(49) (0)271/740-2924

E-Mail: ZJWF@germanistik.uni-siegen.de

Homepage: www.peterlang.com/view/journals/zjwf/zjwf-overview.xml

“Zeitschrift für Wortbildung / Journal of Word Formation” (ZJWF) is an open access and double blind peer reviewed international journal. ZJWF publishes papers on all aspects of word-formation with respect to any language and linguistic field, e.g. morphology, syntax, lexicology, phonology, semantics, pragmatics, language history, typology, dialectology, language acquisition, language contact. The journal is published online and releases two issues a year. It contains original papers, reviews and general information such as announcements of conferences, meetings, workshops, etc. Special issues devoted to important topics in word-formation will occasionally be published. Manuscripts are accepted in English, German, French and Spanish.

Please send your manuscript to the editor-in-chief petra.vogel@uni-siegen.de.

Editorial Assistant: Dr. Carolin Baumann, Universität Siegen, Fakultät I, Germanistisches Seminar, Hölderlinstraße 3, D-57076 Siegen, Tel.: +(49) (0)271/740-2924

E-Mail: ZJWF@germanistik.uni-siegen.de

Homepage: www.peterlang.com/view/journals/zjwf/zjwf-overview.xml

ISSN 2367-3877

Peter Lang GmbH

Internationaler Verlag der Wissenschaften

Schlüterstraße 42, 10707 Berlin

<http://www.peterlang.com>

PETER LANG



This is an open access publication. Except where otherwise noted, content can be used under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

For details go to <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Inhalt / Contents

Vorwort.....7

Beiträge / Papers

Thomas Klein

Mittelhochdeutsche Wortfamilien: Ermittlung und Perspektiven11

Lisa Dücker

Die Getrennt- und Zusammenschreibung von Komposita in frnhd.
Hexenverhörprotokollen33

Kristin Kopf

The role of syntax in the productivity of German N+N compounds.
A diachronic corpus study.....61

Luise Kempf & Matthias Eitelmann

Von *diutisk* zu *dynamisch*, von *englisc* zu *anything-ish*. *-is(c)h*
kontrastiv diachron.....93

Eva Büthe-Scheider

Verwandtschaft oder Entlehnung? Die Beziehung der ripuarischen
Kollektiva mit *ge + ze* zu den mittelniederländischen mit *ge + te*135

Sören Stumpf

Textsortenorientierte Wortbildungsforschung.
Desiderate, Perspektiven und Beispielanalysen165

Joshgun Sirajzade

Korpusbasierte Untersuchung der Wortbildungsaffixe im
Luxemburgischen. Technische Herausforderungen und
linguistische Analyse am Beispiel der Produktivität195

Rezensionen / Reviews

Pius ten Hacken

Carola Trips & Jaklin Kornfilt: Further investigations into the
nature of phrasal compounding.....219

Stefanie Stricker

Nicolaus Janos Raag: Substantivkomposita und Sinngebung im
Kontext frühmittelalterlicher Wissensvermittlung225

Varia / Miscellaneous

Edwige Dugas-Michaud

Tagungsbericht / Conference report: “Morphology Days”
(Louvain-la-Neuve (Belgium), 22.–24.11.2017)237

Tagungsankündigung / Conference Announcement:
18th International Morphology Meeting
(Budapest/Hungary, May 2018)243

Tagungsankündigung / Conference Announcement:
Word-Formation Theories III & Typology and Universals in
Word-Formation (Košice/Slovakia, June 2018)245

Tagungsankündigung / Conference Announcement:
Morphosyntactic misfits: Clitics, particles, and non-canonical
affixes in the languages of the Americas (Lima/Perú, July 2018)247

Vorwort

Der vorliegende Band basiert auf einer Wortbildungs-Tagung, die im Oktober 2016 an der Universität Luxemburg stattfand. Ein Ziel dabei war, bisher noch eher wenig berücksichtigte Gegenstandsbereiche der Wortbildungsforschung aufzugreifen und deren Relevanz durch exemplarische Analysen deutlich zu machen. Im Zentrum standen Aspekte der historischen Wortbildungsforschung sowie der kontrastiven Analyse. Ziel war auch, mögliche Brückenschläge zu angrenzenden Forschungsfeldern aufzuzeigen. Diesem Anspruch ist auch die Zusammenstellung der Beiträge im vorliegenden Band verpflichtet.

Die hier versammelten sieben Beiträge greifen bewusst unterschiedliche Aspekte der Wortbildungsforschung auf und werfen ein Schlaglicht auf jeweils spezifische Fragestellungen. Ablesbar ist das zunächst an den Artikeln zur historischen Wortbildung, die am Anfang stehen. Dazu zählt der Beitrag von *Thomas Klein*, dem es nicht nur darum geht, die Erforschung der Wortfamilienstrukturen des Mittelhochdeutschen als Desiderat zu markieren. Vielmehr wird gleichzeitig ein Verfahren zu deren Erfassung vorgestellt, – von dem auch die weitere Untersuchung der mittelhochdeutschen Wortbildung profitieren kann.

Die beiden sich anschließenden Beiträge konzentrieren sich auf das Frühneuhochdeutsche. *Lisa Dücker* wendet sich der virulenten Frage der Getrennt- und Zusammenschreibung von Komposita zu, indem sie Regularien der Spatiensetzung im Vergleich zwischen Drucken und Handschriften untersucht. Komposita sind auch Gegenstand des Beitrags von *Kristin Kopf*. Im Kern geht es hier um den Einfluss bestimmter syntaktischer Strukturen auf die Bildung von N+N-Zusammensetzungen.

Wie bei Kopf basiert auch der Artikel von *Luise Kempf* und *Matthias Eitelmann* auf einer Corpusanalyse, nur dass diese hier als Grundlage einer kontrastiv ausgerichteten Untersuchung gewählt wurde. Gegenstand des Beitrages ist das deutsche Derivationssuffix *-isch* in Relation zum englischen Pendant *-ish*. Dem Prinzip einer wechselseitig erhellenden kontrastiven Gegenüberstellung folgt auch *Eva Bütthe-Scheider*. Ihr geht es um den Aufweis der Relation zwischen ripuarischen *ge + ze*-Kollektiva und den *ge + te*-Entsprechungen im Mittelniederländischen.

Der sich anschließende Beitrag von *Sören Stumpf* unternimmt einen Brückenschlag zwischen Wortbildung und Textsortenspezifikation. Das Potential dieses Zugriffs kann nicht zuletzt durch einige Beispielanalysen der Textsorte *battle rap* demonstriert werden.

Der letzte Artikel widmet sich, bedingt durch den Veranstaltungsort der zugrundeliegenden Fachtagung, dem Luxemburgischen. Darin befasst sich

Joshgun Sirajzade mit der computer-technischen Seite der Untersuchung Luxemburger Wortbildung. Hintergrund hierfür ist ein von den Herausgebern dieses Bandes initiiertes luxemburgistisch-germanistisches Tandemprojekt zur Corpuserstellung und -analyse der historischen und rezenten Wortbildung im luxemburgisch-moselfränkischen Sprachraum.

Alle Beiträge wurden gemäß den Vorgaben der *Zeitschrift für Wortbildung* begutachtet und überarbeitet. Den Beiträgern dieses Bandes gilt unser Dank. Zu danken ist auch den Herausgebern der vorliegenden Zeitschrift für das freundliche Angebot zur Publikation, wobei Petra Vogel und Carolin Baumann für die konstruktive und effiziente Zusammenarbeit besondere Anerkennung gebührt. Zu danken ist aber auch all den namentlich Ungenannten, die zum Gelingen der Fachtagung, insbesondere aber zum Erscheinen dieses Bandes beigetragen haben.

Luxemburg, im Januar 2018
Peter Gilles und Heinz Sieburg

Beiträge / Papers

Thomas Klein

Mittelhochdeutsche Wortfamilien: Ermittlung und Perspektiven

Abstract: An inventory of the Middle High German word families is still missing whereas the Old High German and New High German word families are recorded by the dictionaries of J. Splett. In this paper a semi-automatic method is represented which can help to find and analyze the Middle High German word families. By several scripts a combined list of MHG and OHG lemmata is transformed and expanded to a table containing among other things a column with a simplified variation of Splett's word formation formulas and a column with the common base of the word family the lemma probably belongs to. In a labour-intensive last step, these proposals have to be manually checked and corrected.

1. Mittelhochdeutsche Wortbildung: Stand und Aufgaben

Zwar gibt es eine neuere Gesamtdarstellung der mhd. Wortbildung (Klein, Solms, und Wegera 2009) und eine Reihe wichtiger und gewichtiger Monographien zur Wortbildung der einzelnen Wortarten (Herbers 2002; Leipold 2006; Ring 2008; Ganslmayer 2012). Doch bleibt hier durchaus noch einiges zu tun. Die gravierendste Lücke scheint mir zu sein, dass die mhd. Wortfamilien noch weithin im Dunkeln liegen. Weder ihre ungefähre Zahl noch Umfang und Struktur der einzelnen Wortfamilien sind bekannt, noch lassen sich bisher Aussagen machen über Wachstum und Schrumpfen von Wortfamilien innerhalb der mhd. Zeit oder im Vergleich mit dem Althochdeutschen und Neuhochdeutschen.¹ Für diese Sprachepochen sind die Wortfamilien von Jochen Splett bereits umfassend ermittelt und dokumentiert worden (Splett 1993; 2009).

Spletts Wörterbücher bieten

- eine strukturierte Darstellung aller Wortfamilien; die Binnenstruktur der Wortfamilie wird durch eine dezimale Gliederung verdeutlicht (s. Abb. 1);
- die Analyse der Wortbildungsstruktur aller Wörter in Gestalt von „Strukturformeln“, wobei von der Form, Funktion und Semantik der beteiligten Affixe und Kernwörter abgesehen wird.

1 Angeregt durch zwei Vorträge, die ich dazu und zur Verknüpfung digitaler Lemmalisten historischer Sprachstufen des Deutschen 2013 und 2016 bei den „Arbeitsgesprächen zur historischen Lexikographie“ in Bullay (<https://www.uni-trier.de/index.php?id=14673>) gehalten habe, gibt es erfreulicherweise im Projekt ZHistLex („eHumanities-Zentrum für Historische Lexikographie“) ein Arbeitspaket, das den „Epochenübergreifenden Wortfamilienstrukturen“ in ahd. und mhd. Zeit gewidmet ist, vgl. <http://zhistlex.de/ziele/arbeitspakete/>.

Abb. 1: Wortfamilie *ahd. BREMAN* in Splett (1993, I.1: 101).

BREMAN (1)		
1.0	breman ** st.V.	wV/ 'brüllen'
1.1.3	brāmen * sw.V.	(wV)Vjan/ 'brüllen'
1.1.4	brummen * sw.V.	(wV)Vjan/ 'brünstig grunzen'
1.2.1	brema * sw.M.	(wV)San/ 'Bremse [Insekt]'
	brema-krūt ? ** st.N.	((wV)San) (wS)/ 'Taubnessel' (2)
	-stal * st.M.	/ 'Sommergehege [Ort zum Schutz der Herde gegen Bremsen]'
	bremen-krūt ** st.N.	/ 'Taubnessel' (2)
	brema * sw.F.	(wV)Són/ 'Bremse [Insekt]'
1.2.2	bremma * F.	(wV)S/ 'Bremse [Insekt]' (3)
2.2.1	brim-issa * F.	((wV)San) sS/ 'Bremse [Insekt]' (4)
2.3.4	brum-ft-ig * Adj. ²	[(wV)sS]sA [((wV)Vjan) sS]sA/ 'brünstig'

Zu einer „Grundlegung einer zukünftigen Strukturgeschichte des deutschen Wortschatzes“, die der Untertitel von Spletts Wörterbüchern (Splett 1993; 2009) als Ziel benennt, wäre aber noch eine vergleichbare Erfassung der mhd. und frühhd. Wortfamilien wünschenswert. Der dafür nötige Arbeitsaufwand ist jedoch immens. Es ist aber möglich, ihn ausgehend von Spletts Wortfamilienwörterbüchern und mit (semi)automatischen Verfahren ein gutes Stück zu verringern. Dies soll im Folgenden für das Mittelhochdeutsche gezeigt werden. Schon vorab aber ist zu betonen, dass die hier vorgeschlagene Verfahrensweise eine linguistische „händische“ Wortfamilienklassifikation nicht ersetzen, sondern nur erheblich erleichtern kann.

2. Voraussetzungen und Vorarbeiten für die Ermittlung mittelhochdeutscher Wortfamilien

2.1. Althochdeutsche und mittelhochdeutsche Lemmalisten

Es liegt nahe, für eine Ermittlung der mhd. Wortfamilien Spletts Wortfamilienwörterbücher als Muster und Ausgangspunkt zu nutzen. Voraussetzung dafür ist allerdings, dass diese Wörterbücher in digitaler Form verfügbar sind. Für den Wortindex zu Spletts Althochdeutschem Wörterbuch (Splett 1993) ist das der Fall. Auf der mhd. Seite gibt es die Bonner Lemmaliste des Projekts „Mhd. Grammatik“. Im Umfang deckt sich diese Liste weitgehend mit dem Lexembestand von Lexers Mhd. Wörterbuch, vermehrt um den Bestand des Trierer Findebuchs (Gärtner et al. 1992). Die Lexemansätze der Bonner Lemmaliste entsprechen denen des Korpus der Mhd. Grammatik (MiGraKo). Diese Ansätze orientieren sich jedoch stärker als die Lexers oder gar des MWB an grammati(kographi)schen Bedürfnissen und sind daher auch für Untersuchungen zur Wortbildungsstrukturen weit geeigneter. Mit Lexers Wörterbuch und

dem Findebuch teilt die Bonner Lemmaliste andererseits den Nachteil, viele Lexeme zu enthalten, die ausschließlich in frühneuhochdeutschen Quellen oder in Quellen aus der spätmittelhochdeutsch-frühmittelhochdeutschen Übergangszeit belegt und damit für den engeren mhd. Zeitraum vom späten 11. bis zur Mitte des 14. Jahrhunderts nicht gesichert sind. Besser wäre es daher, diese Lexeme von einer Untersuchung der mhd. Wortfamilien auszuschließen. Für die bisher erschienene Alphabetstrecke (*a – hanfken*) sollte dafür der Lexembestand des MWB maßgeblich sein. Für die Erprobung des hier vorgestellten Verfahrens musste jedoch noch die unveränderte Bonner Lemmaliste zugrunde gelegt werden.

Von den Lexemansätzen der mhd. Wörterbücher weichen die internen Ansätze des Projekts „Mhd. Grammatik“ und damit auch der Bonner Lemmaliste aus grammatikographischen Gründen in folgenden Punkten ab:

- Die Ansätze sind frühmhd., also ohne Auslautverhärtung, ohne Lenisierung *-nt- > -nd*, z. B. *bald, lib, tag, winten* ‚winden‘.
- Anders als auch in den ahd. Wörterbüchern werden Geminaten im Auslaut nicht gekürzt, z. B. *ball, sinn, spott, schiff*.
- Die mhd. Schwa-Apokopierungen und Synkopierungen werden nur insoweit berücksichtigt, als das „normalmhd.“ getilgte Schwa in Rundklammern erscheint, z. B. *ar(e)* ‚Adler‘.
- Bei den kurzen e-Lauten werden unterschieden:
 - *ë* für altes (germ.) *e*,
 - *è* für (Primär- und Sekundär-)Umlaut-*e*, als Diphthongkomponente in *iè* und *üè* und in *oè* für langes *œ*,
 - *e* steht ausschließlich für Schwa.
- Die Lexeme sind wie bei Lexer morphologisch segmentiert, anders als bei Lexer aber konsequent auch in Fällen wie *ver-ant-würten* (Lexer *ant-würten*, aber *ver-antwürten*). Abgetrennt werden wie bei Lexer nur die „neuen“ Suffixe wie *-heit, -schaft, -tuom*, nicht auch die „alten“ wie *-ig, -(e)de < (-ida, -idi), -inne* usw.

2.2. Verknüpfte Lemmaliste Ahd. – Mhd.

Für interne Zwecke des Projekts Mhd. Grammatik wurde bereits vor mehr als 10 Jahren in Bonn der „Alphabetische Wortindex“ aus Spletts Ahd. Wörterbuch (Splett 1993 II: 389–620) mit der Bonner Lemmaliste verknüpft. Dazu setzte ein Script die ahd. Lexeme provisorisch ins Mhd. um und fügte das künstlich erzeugte mhd. Lexem jeweils dem Datensatz des ahd. Lexems hinzu. Sodann prüfte das Script, ob das provisorische mhd. Lexem in der mhd. Bonner Lemmaliste enthalten ist; wenn nein, wurde es mit einem initialen Stern markiert. Jedoch konnten so nur bei weniger als 40% der ahd. Lexeme die mhd. Nachfolger automatisch zugeordnet werden. Dies liegt in

der Hauptsache daran, dass sich nur knapp die Hälfte des ahd. Wortschatzes noch im Mhd. belegen lässt. Außerdem wich in nicht wenigen Fällen die automatisch erzeugte mhd. Form von der Gestalt des mhd. Nachfolgelexems in der mhd. Bonner Lemmaliste ab, sodass die Zuordnung unterblieb. Aus diesem Grund musste insbesondere bei allen Fällen mit sternmarkiertem mhd. Lexem geprüft werden, ob nicht doch ein mhd. Nachfolgelemma existiert. Das Ergebnis dieser Arbeiten ist eine verknüpfte ahd.-mhd. Lemmaliste (s. Tab. 1).

Tab. 1: Ausschnitt aus der Bonner Lemmaliste Ahd. – Mhd.

ahd. Lexem	Beleglage	mhd. Lexem ¹⁾	ahd. Wortfamilien- kernwort
bald, Adj.	(+*)	bald, adj	BALD
balda, st.F.[(j)ô]	*	balde, f	BALD
baldên, sw.V.	(*)	balden, swv	BALD
bald-lîh, Adj.	*	bald-lich, adj	BALD
bald-lîhho, Adv.	(*)	bald-lîche, adv	BALD
bald-lîhhûn, Adv.	*	bald-lîche, adv	BALD
baldo, Adv.	(+*)	balde, adv	BALD
belden, sw.V.	(+*)	bêlden, swv	BALD
beldî, st.F.[î]	(*)	bêlde, f	BALD
beldida, st.F.[ô]	(*)	*beldede, f	BALD
beldirîh, st.M.[a/i]	*	balderich, m	BALD
fra-bald, Adj.	*	*vra-bald, adj	BALD
fra-bald-lîh, Adj.	*	*vra-bald-lich, adj	BALD
fra-bald-lîhho, Adv.	*	*vra-bald-lîche, adv	BALD
fra-baldo, Adv.	*	*vra-balde, adv	BALD
fra-beldî, st.F.[î]	*	*vra-bêlde, f	BALD
gi-bald, Adj.	*	*ge-bald, adj	BALD
gi-baldôn, sw.V.	*	ge-balden, swv	BALD
ir-baldên, sw.V.	(*)	er-balden, swv	BALD
ir-baldnissi, st.N./F.[jaî]	*	er-baldnisse, f	BALD
ir-baldôn, sw.V.	()	er-balden, swv	BALD
ir-belden, sw.V.	*	*er-bêlden, swv	BALD
un-bald, Adj.	()	*un-bald, adj	BALD
un-baldo, Adv.	+	*un-balde, adv	BALD
un-beldî, st.F.[î]	()	*un-bêlde, f	BALD
bora-baldo, Adv.	()	*bore-balde, adv	BALD + BOR
kurzi-bold, st.M.[a/i]	*	kurze-bold, m	BALD + KURZ

¹⁾ * vor dem Lemma = mhd. nicht belegt, automatisch konstruiert

In die mhd. Bonner Lemmaliste wurde dann dort, wo für ein mhd. Lexem sein ahd. Vorläufer belegt ist, dieses ahd. Lexem nebst grammatischen Angaben und das zugehörige Wortfamilienkernwort aus Spletts Wortindex inseriert (s. Tab. 2).

Statistisch gesehen deutet die ahd.-mhd. Lemmaliste auf ein beträchtliches Maß an lexikalischer Diskontinuität: Nur für knapp die Hälfte der ahd. Lexeme (ca. 47%) finden sich Nachfolger in der Bonner Lemmaliste und nur etwa ein Sechstel (ca. 17%) der mhd. Lexeme dieser Liste haben Vorläufer, die in ahd. Quellen belegt sind. Zu bedenken ist allerdings, dass in der vergleichsweise spärlichen ahd. Überlieferung sehr viele ahd. Lexeme, die sich im Mhd. fortsetzen, zufällig nicht belegt sein dürften. Zudem sind in die Bonner Lemmaliste viele Lexeme aus Lexer und dem Findebuch (Gärtner et al. 1992) übernommen, die erst in frühmhd. Zeit belegt sind und daher eigentlich nicht einbezogen werden dürften (s. oben 2.1). Andererseits enthält die Liste auch Wortgut, das ausschließlich in Quellen der ahd.-frühmhd. Übergangszeit belegt ist. Berücksichtigt man diese Faktoren, so dürfte der tatsächliche Anteil der mhd. Lexeme mit ahd. Vorgeschichte erheblich höher liegen als 17%.

Nicht ganz selten sind zwei ahd. Lexeme ausdrucksseitig in einem mhd. Lexem zusammengefallen. Die Regel ist das bei den ahd. *-ên-* und *-ôn-* Verben, die im Ahd. oft auch semantisch weitgehend übereinstimmen; vgl. z. B. ahd. *ir-baldên, ir-baldôn* ‚sich erkühnen, wagen‘ → mhd. *er-balden* (s. Tab. 1). In der Bonner Lemmaliste sind dann dem mhd. Lexem beide ahd. Vorgängerlexeme zugeordnet (s. Tab. 2). Ähnlich wurde verfahren, wenn zwei semantisch verschiedene ahd. Lexeme im Mhd. ausdrucksseitig zusammengefallen sind und so zu einem polysemen mhd. Nachfolgelexem geführt haben, vgl. *bleichen* in Tab. 2.

Tab. 2: Ausschnitte aus der erweiterten Bonner mhd. Lemmaliste.

MiGraKo Lexem POS	(Lexer/MWB) MWB-ID	Althochdeutsch	
		Lexem POS	Wortfamilien- Kernwörter
bal-rât m<u>	bal-rât stm 10836000	balo-rât st.M.	BALO + RâTAN
bal-rêdære m	bal-reder stm 10842000	–	
...
balas m	balas stm 10683000	–	
bald adj	balt a 10959000	bald Adj.	BALD
bald-hêt f	balt-heit stf 10965000	–	
bald-lich adj	balt-lich a 10968000	bald-lih Adj.	BALD

<i>MiGraKo</i>	<i>(Lexer/MWB)</i>	<i>MWB-ID</i>	<i>Althochdeutsch</i>	
Lexem POS			Lexem POS	Wortfamilien-Kernwörter
bald-lîche adv	balt-lîche adv	10974000	bald-lîhho Adv./ bald-lîhhûn Adv.	BALD
bald-lôs adj	balt-lôs a	10977000	–	
bald-spræche adj	balt-spræche a	10980000	–	
bald-spræche f	–	–	–	
balde adv	balde adv	10704000	baldo Adv.	BALD
balde f	balde stf	10707000	balda st.F.	BALD
baldekîn mn	baldekîn stm	10719000	–	
...
blëichen swv {1} ‚bleich werden‘, {2} ‚bleich machen‘	bleichen swv	22158000	{1} bleihhên sw.V., {2} bleihhen sw.V.	BLÎHHAN
...
er-balden swv	er-balden swv	40434000	ir-baldên sw.V., ir-baldôn sw.V.	BALD

2.3. Übertragung der Strukturformeln auf die mhd. Nachfolgelexeme

Wenn Spletts Ahd. Wörterbuch komplett digitalisiert vorliegt, so lassen sich auch die Strukturformeln von den ahd. Lexemen auf ihre mhd. Nachfolger übertragen. Dazu müsste ein Script lediglich die ahd. Lexeme aus der vierten Spalte der Bonner Lemmaliste im Digitalisat von Splett Bd. I,1–2 aufsuchen, dort die Strukturformel entnehmen und in einer zusätzlichen Spalte der Bonner Lemmaliste einfügen. Ebenso könnten die Angaben zur Binnenstruktur der Wortfamilie in Gestalt der Dezimalgliederung (s. Abb. 1) übernommen werden. In den Strukturformeln wären die Stammbildungsangaben, die für das Mhd. nicht mehr von Belang sind, zu löschen, z. B. mhd. *balden* swv → ahd. *baldên* → ahd. Strukturformel (wA)Vên → mhd. (wA)V. Wo ein mhd. Lexem mehrere ahd. Vorgänger hat, werden deren Strukturformeln dadurch in der Regel zusammenfallen, z. B. mhd. *er-balden* swv → ahd. *ir-baldên*, *ir-baldôn* → ahd. Strukturformel p(wA)Vên, p(wA)Vôn → mhd. p(wA)V, p(wA)V → p(wA)V.

Für die Frage, welche Teile der ahd. Wortfamilien sich im Mhd. fortsetzen, würde sich dagegen die umgekehrte Blickrichtung empfehlen: Ausgehend von Spletts Wörterbuch würden in der ahd.-mhd. Lemmaliste die mhd. Nachfolgelexeme gesucht und erweiterten Datensätzen des Wörterbuchs hinzugefügt. Ist

das mhd. Lexem *besternt* (also nicht belegt, sondern nur konstruiert), so wäre auch das ahd. Lexem entsprechend zu markieren (etwa durch ein Sterbekreuz). Man würde so auch sehen, in welcher Weise die Binnenstruktur der Wortfamilien (in Gestalt der dezimalen Gliederung in Spletts Wörterbuch, s. Abb. 1) durch die Lexemverluste beeinträchtigt ist. Da im Schnitt nur die Hälfte der belegten ahd. Lexeme im Mhd. überlebt hat, dürften diese „Erosionsschäden“ teils beträchtlich sein.

Für die mhd. Seite wäre durch diese Übertragung der Wortbildungsstrukturdaten aus dem Splett-Wörterbuch zunächst wenig gewonnen. Denn erfasst werden damit nur die 17% der mhd. Lexeme, die einen belegten ahd. Vorläufer haben. Es fragt sich also, wie auch für den gesamten Rest mit vertretbarem Aufwand eine Zuordnung zu Wortfamilien erreichbar wäre. Im Folgenden wird ein semiautomatisches Verfahren vorgestellt, das diese Zuordnung erleichtert und dabei zugleich (vereinfachte) Strukturformeln konstruiert.

3. Konstruktion von mhd. Strukturformeln und Kernwortextraktion

3.1. Übersicht

Der automatische Teil des Verfahrens soll zu jedem mhd. Lexem eine (vereinfachte) Strukturformel erzeugen und die in ihm enthaltenen ‚Kernwörter‘² (Grundmorpheme) ihrer Wortfamilie (kurz: WF) zuweisen. In der Übersicht am Beispiel *bēte-gültig* ‚steuerpflichtig‘:

Ausgangslexem	bēte	– gültig	
	↓	↓	
Segmentierung (Suffixabtrennung)	bēt-e	– gült-ig	
	↓	↓	
Extraktion u. Strukturformelerzeugung	bēt	gült	→ wXlsS+wXlsA
	↓	↓	
Umlautbereinigung	bēt	gult	
	↓	↓	
Alternanzbereinigung	bit	gēlt	
	↓	↓	
Sortieren	WF	WF	
	BITTEN	GELTEN	

2 Dieser Terminus nach Splett (1993: XXIII).

3.2. Vorbereitende Schritte

3.2.1. Suffixabtrennung

In der Bonner Lemmaliste sind im Gefolge Lexers nur die Präfixe und die noch erkennbar aus Kompositionsgliedern entstandenen „Neu-Suffixe“ wie *heit*, *schaft*, *bære*, *haft* usw. durch „-“ abgetrennt und außerdem Kompositionsfugen ebenso bezeichnet. Für eine Extraktion der selbst möglichst nicht mehr suffigierten Kernwörter ist aber auch die Abtrennung der „Alt-Suffixe“ wie z. B. *ig*, *ing*, *inne*, *nisse*, *e*, *en*, *ede* usw. erforderlich, z. B. *aller-diè-müètigest* → *aller-diè-müèt-ig-est* (*müèt* ist hier umgelautete Variante des Kernworts *muot*).

Diese Abtrennung der „Alt-Suffixe“ ist einer der kritischsten Punkte des gesamten Verfahrens, denn es ist unvermeidlich, dass es dadurch auch zu unerwünschten Kernwort-Homonymien kommt. Dieser Schaden sollte durch a) eine optimale Auswahl der abzutrennenden Suffixe und b) eine geeignete Formulierung der Abtrennungsregeln aber so klein wie möglich gehalten werden. Beispiele:

- Wenn etwa *er* global abgetrennt würde, so wäre der Schaden vermutlich größer als der Gewinn, da die Zahl unerwünschter Kernwort-Homonymien sehr groß würde, vgl. z. B. (*⇒ steht für falsche Wortfamilienzuordnung):

muoter → *muot-er* *⇒ Wortfamilie MUOT; *meister* → *meist-er* *⇒ MEIST; *bitter* → *bitt-er* *⇒ BITTEN; *laster* → *last-er* *⇒ LADEN usw.

Daher wird bei der aktuellen Script-Version auf eine Abtrennung von *er* ganz verzichtet.

- Von einer globalen Suffixabtrennung können bestimmte Kernwörter oder Kernwortvarianten vorab durch ihre Großschreibung in der Lemmaliste ausgenommen werden, da das Script Groß- und Kleingeschriebenes strikt unterscheidet, z. B. (-)adel(-) → (-)ADEL(-), (-)èdel(-) → (-)ÈDEL(-); ADEL ist dann von der automatischen Abtrennung des Suffixes *-el* nicht mehr betroffen.
- Es darf nicht abgetrennt werden, wenn das danach links verbleibende Segment nicht silbisch ist, z. B. *bringen* → *bring-en*, nicht → **br-ing-en**.
- Wird ein mit *i* anlautendes Suffix abgetrennt, so darf das links verbleibende Segment nicht auf *e* enden, da sonst der Diphthong *ei* zertrennt würde, z. B. *abe-neig-en* nicht → *abe-ne-ig-en*.
- Wichtig auch: Suffixe dürfen nicht ihrerseits noch einmal segmentiert werden, weil sie dann später nicht mehr als Suffix erkannt werden, also z. B. *burg-mæze* ‚burgmässig‘ nicht → * *burg-mæz-e*. Das hat bei Homonymie von Suffix und Kernwort(variante) (s.u. 3.2.3) den Nachteil, dass der Segmentierungsschritt auch bei den (wenigen) Bildungen mit dem Kernwort

unterbleibt und damit auch die Kernwortextraktion und Zuordnung zur Wortfamilie nicht zum Ziel führt, z. B. nicht *un-mæze* → *un-mæz-e* → *mæz* → *mâz* → *mëz* ⇒ MËZZEN.

3.2.2. Affixliste

Wenn festgelegt ist, was als Affix zu betrachten ist, werden Präfixe und Suffixe in einer zweispaltigen Liste zusammengefasst, in der den Affixen jeweils das Symbol zugeordnet ist, durch das sie in der Strukturformel vertreten sind. Für Präfixe ist das „p“, für Suffixe „s“. Wenn das Suffix wortarteindeutig ist, wird das Wortartsymbol hinzugefügt: „sS“ Substantiv-, „sA“ Adjektiv-, „sAD“ Adverbsuffix; wenn sich das Suffix mit mehreren Wortarten verbinden kann, sollten möglichst alle aufgeführt werden, z. B. *-în* sS, sA, *-haft* sA, sAD; alternativ kann „sX“ ohne Festlegung der Wortart zugeordnet werden. Bei genuseindeutigen Substantivsuffixen wird das Genus in eckigen Klammern hinzugefügt, z. B. *-schaft* sS[f].

Tab. 3: Beispiele aus der Affixliste.

Präfix	Struktursymbol	Suffix	Struktursymbol
be	p	ig	sA
ge	p	lich	sA
ent	p	liche	sAD
abe	p	hêt	sS[f]
ane	p	(e)de	sS
gègen	p	schaft	sS[f]
		haft	sA, sAD
		e	sX

3.2.3. Mehrdeutige Morpheme

Besondere Schwierigkeiten bereiten bei dem späteren Ersatz der Affixe durch ihre Struktursymbole (s. unten 3.2.5) Affixe, die mit Kernwörtern homonym sind. In den meisten Fällen wird das fragliche Morphem den ersten oder letzten Teil eines komplexen Lexems bilden. Hier gibt es algorithmisierbare Regeln für die Entscheidung zwischen Affix und Kernwort. Zunächst hilft der simple Umstand, dass Präfixe nicht den letzten Teil und Suffixe nicht den ersten Teil des Lexems bilden können. Für die Entscheidung zwischen Suffix und Kernwort lässt sich weiterhin nutzen, dass das Script der Affixliste auch die Wortart und bei Substantiven das Genus des Lexems entnommen hat:

- Stimmt die vom Suffix geforderte Wortart mit der des Lexems überein, so handelt es bei dem fraglichen Morphem um das Suffix, sonst um das Kernwort, z. B. ist *anker-haft* ‚Ankergrund‘ Substantiv, während das

Suffix *-haft* nur Adjektive oder Adverbien bilden kann; also kann *haft* in *anker-haft* nur Kernwort und nicht Suffix sein.

- Stimmt das von einem Substantivsuffix geforderte Genus mit dem des Lexems überein, so liegt das Suffix vor, sonst das Kernwort. So werden mit *-schaft* nur feminine Substantive gebildet; daher handelt es bei *schaft* in *spër-schaft* mask. ‚Speerschaft‘ um das Kernwort.
- Bliebe bei der Annahme eines Suffixes als Restwort lediglich ein Morphem übrig, das formal nur Präfix sein kann, so muss das Kernwort vorliegen, da sich sonst eine nur aus Affixen bestehende kernwortlose Bildung ergäbe. So kann es sich bei *bære* und *mæze* in *ge-bære*, *ge-mæze* nur um Kernwörter handeln, da *ge* wegen des Schwa formal nur Affix sein kann (daher nicht → *pl_sA, sondern → plwV|A). Bei Lexemen wie *un-mæze* müsste das Script dagegen wissen, dass *un* nur Präfix und nicht auch Kernwort sein kann.
- Allgemeiner kann die vorstehende Regel so gefasst werden: Unmittelbar links von einem Suffix kann kein Präfix und unmittelbar rechts von einem Präfix kann kein Suffix stehen. So kann *über* in *hōch-über-ig* nur Kernwort sein, daher nicht → *pl_sA, sondern → plwAD|sA oder wAlwAD|sA, falls *hōch* als Kernwort und nicht als gradatives Präfix (Klein, Solms und Wegera 2009: § A24) gewertet würde.

Es bleiben Fälle, die durch diese Regeln nicht abgedeckt sind. So ist bei *ant-vogel* ‚Ente(nvogel)‘ und *ant-vuoz* ‚Entenfuß‘ formal nicht entscheidbar, ob *ant* Präfix oder Erstglied eines Kompositums ist. Soweit sie bekannt sind, können Kernwörter in solchen Bildungen vorab durch Großschreibung blockiert werden (da das Script nur kleingeschriebene Affixe kennt), z. B. *ant-vuoz* → *ANT-vuoz*, das dann automatisch als wX+wX analysiert würde.

3.2.4. Erweiterung der Bonner Lemmaliste

Die Bonner mhd. Lemmaliste wird durch ein vorbereitendes Script um zwei Felder erweitert. Das erste Feld füllt das Script mit einer Form des Lexems (Lexem2), in der die „Altsuffixe“ abgetrennt sind. Das zweite Feld ist für das Kernwort vorgesehen und wird provisorisch mit der Ausgangsform des Lexems gefüllt, die bei Simplicia wie *hûs* oder *blint* ja bereits mit dem Kernwort identisch ist.

Tab. 4: Erweiterte Bonner Lemmaliste.

	Mittelhochdeutsch		Lexer	Althochdeutsch
Lexem, POS	Lexem2	Kernwort	...	Lexem, POS Kernwort
<i>ab-grünt-lich</i> , adj	<i>ab-grünt-lich</i>	<i>ab-grünt-lich</i>	...	–
<i>runs(e)-âder(e)</i> , f	<i>runs-(e)-âder-(e)</i>	<i>runs(e)-âder(e)</i>	...	
<i>bëte-gültig</i> , adj	<i>bët-e-gült-ig</i>	<i>bëte-gültig</i>	...	

3.2.5. Konstruktion der mhd. Strukturformel und Kernwortextraktion

Aus den Segmenten von Lexem2 erzeugt das Hauptsript mittels der aus der Affixliste bezogenen Informationen Strukturformeln, die gegenüber den Strukturformeln Spletts (s. Splett, Ahd.Wb. I,1, XXXf.) erheblich vereinfacht, teils auch modifiziert sind:

- Auf eine Klammerung, mit der Splett die Abfolge der einzelnen Wortbildungsschritte anzeigt, wird verzichtet, da die Strukturanalyse, die der Klammerung zugrunde liegt, automatisch kaum zuverlässig durchzuführen ist. Stattdessen werden Morphemgrenzen durch „|“ angezeigt.
- „+“, das in Spletts Formeln Zusammenrückungsfugen anzeigt, bezeichnet auch Kompositionsfugen.
- Die Wortartkennzeichnung durch A = Adjektiv, S = Substantiv, V = Verb, AD = Adverb unterbleibt bei Kernwörtern grundsätzlich: Statt wS, wS usw. steht stets wX. Bei Suffixen steht das Wortartkennzeichen nur in eindeutigen Fällen (z. B. *-ig* sA), nicht aber z. B. bei *-e* und *-en*.
- Von Präfixen (p) werden abweichend von Splett Verbpartikel (vp) unterschieden.

Die vereinfachte Strukturformel wird durch das Hauptsript aus Lexem2 erzeugt. Zunächst werden die Affixe sukzessive durch ihre Struktursymbole ersetzt. Bei dem danach verbleibenden Lexemsegment muss es sich um das Kernwort oder eine Variante des Kernworts handeln, bei Komposita entsprechend um zwei Kernwörter; z. B. („KW“ steht für Kernwort(variante)):

- ab-grünt-lich*: *ab-* → pl, *-lich* → lsA, verbleibend KW *grünt-* → wX; ergibt die Strukturformel: plwXlsA
- bëte-gültig*: *-ig* → lsA, *-e* → ls, verbleibend KW *bët* → wX, KW *gült-* → wX; ergibt die Strukturformel: wXls+wXlsA
- un-ab-ge-zèl(e)t*: *un-* → pl, *ab-* → pl, *ge-* → pl, *-(e)t* → |Vpart2|A; verbleibend KW *zèl*; ergibt die Strukturformel: pl|plwX|Vpart2|A

Tab. 5: Beispiele aus der um Strukturformel und Kernwort erweiterten Lemmaliste.

Lexem, POS	Lexem2	Strukturformel	Kernwort
<i>ab-grünt-lich</i> , adj	<i>ab-grünt-lich</i>	plwXlsA	grünt
<i>bëte-gültig</i> , adj	<i>bët-e-gült-ig</i>	wXls+wXlsA	bët+gült
<i>un-ab-ge-zèl(e)t</i> , adj_part	<i>un-ab-ge-zèl-(e)t</i>	plplpwXIVpart2lA	zèl

3.2.6. Konstruktion des umlautbereinigten Kernworts

Sodann erzeugt das Script ein weiteres Feld für das „umlautbereinigte“ Kernwort. Das bedeutet, dass ein ahd. *i*-Umlaut im Kernwort rückgängig gemacht wird, z. B. *ü* → *u*, *üe* → *uo*, *öu* → *ou*. Bei Kernwörtern, die nicht vom ahd. *i*-Umlaut betroffen sind,³ wird das unveränderte Kernwort als umlautbereinigte Form genommen. Nicht berücksichtigt wird bei diesem Schritt der germ. *i*- und *a*-Umlaut, der erst später einbezogen wird (s. 3.2.7). Enthält das Lexem als Kompositum mehr als ein Kernwort, so wird für jedes Kernwort ein weiterer Datensatz hinzugefügt, der im letzten Feld nur die umlautbereinigte Form dieses Kernworts enthält (s. *bëte-gültig* in Tab. 6).

Tab. 6: Beispiele aus der um das umlautbereinigte Kernwort erweiterten Lemmaliste.

Lexem, POS	Lexem2	Strukturformel	Kernwort	umlautbereinigtes Kernwort
<i>ab-grünt-lich</i> , adj	<i>ab-grünt-lich</i>	plwXlsA	<i>grünt</i>	grunt
<i>bëte-gültig</i> , adj	<i>bët-e-gült-ig</i>	wXls+wXlsA	<i>bët+gült</i>	bët
<i>bëte-gültig</i> , adj	<i>bët-e-gült-ig</i>	wXls+wXlsA	<i>bët+gült</i>	gult
<i>un-ab-ge-zèl(e)t</i> , adj_part	<i>un-ab-ge-zèl-(e)t</i>	plplpwXIVpart2lA	<i>zèl</i>	zal

3.2.7. Weitere Alternanzbereinigung: Ablaut, Grammatischer Wechsel, Geminatio

Während bei *grunt* und *zal* mit der Umlautbereinigung auch formal bereits das Kernwort der Wortfamilie extrahiert ist, unterscheidet sich *gult* noch durch Ablaut vom Kernwort *gëlt* seiner Wortfamilie GËLTEN.

Wie lässt sich dieses Problem lösen? Eine globale „Ablautbereinigung“ würde jeden Vokal soweit wie möglich einer Ablautreihe zuordnen und dann

3 Bei mhd. *iu* ist eine Umlautbereinigung nicht sinnvoll, da sonst z. B. auch *liut* (in *liut-e* ‚Leute‘) zu *lüt* umgeformt und *liute* damit der Wortfamilie LÛT zugeschlagen würde.

eine „ablautebereinigte“ Form des Kernworts mit dem grundstufigen Vokal erzeugen, z. B. Kernwort *bôt* → Ablautreihe 2b → ablautebereinigtes Kernwort *bièt*. Dies würde jedoch zu einer Menge von Fehlzuordnungen und untauglichen Kernwörtern führen; so würde *brôt* ‚Brot‘ der Ablautreihe 2b zugeordnet und dazu die abwegige Kernwortform *brièt* erzeugt; entsprechend z. B. *bluot* ‚Blut‘/‚Blüte‘ → Ablautreihe 6 → ablautebereinigtes Kernwort *blat*, wodurch sowohl *bluot* ‚Blüte‘ als auch *bluot* ‚Blut‘ der Wortfamilie BLAT (oder gar *BLATEN!) von *blat* ‚Blatt‘ zugeordnet würden.

Daher empfiehlt sich ein anderer Weg: Ausgangspunkt ist hierbei, dass die Hauptlieferanten solcher ablautebedingter Kernwortvarianten die Wortfamilien starker Verben sind. Deshalb werden durch ein eigenes Script zunächst möglichst alle erwartbaren Kernwortvarianten erzeugt, die bedingt sind durch

- Ablaut,
- Gramm. Wechsel (z. B. *ziehen* – *zugen*),
- Geminatio u. expressive Verschärfung (z. B. *zücken* < **tukkjan* < **tuggjan*)
- germ. *i*-Umlaut (z. B. *siht* < **seht-i-* zu *sëhen*)⁴
- germ. *a*-Umlaut (z. B. *ge-bot* < **ga-bud-a-* zu *bièten*)

Unter „Alternanzbereinigung“ sei die Zuordnung aller so entstandenen Kernwortvarianten zum Kernwort der jeweiligen Wortfamilie zusammengefasst, das grundstufig und frei von Gramm. Wechsel, Geminatio und germ. *i*-Umlaut ist. So wird z. B. zu umlautebereinigtem *zuck* (und ebenso auch zu *zôh*, *zug*, *zog*, *ziug*, *zock*) das alternanzbereinigte Kernwort *zieh* 2b erzeugt. Die Variantenliste zu den mhd. starken Verben (und Präterito-Präsentien) enthält aktuell ca. 1740 Varianten. Wünschenswert wäre es, auch die Kernwort-Varianten der Präteritopräsentien in die Liste aufzunehmen. Die Liste kann verbessert werden durch Veränderung des sie erzeugenden Scripts, durch manuelles Löschen oder Hinzufügen von Einträgen ohne Veränderung des Quellcodes.

4 Hierher auch *i-e*-Alternanzen, die sich anders erklären, wie z. B. *stich* mit altem schwundstufigem *i* zu „ablauteentgleistem“ *stëchen* (Seebold 1970: 467f.); bei *j*-Präsentien wie *sitzen* und *bitten* werden wegen der umgelauteten Grundformen *sitz* und *bitt* als alternanzbereinigte Kernwörter angesetzt.

Tab. 7: Ausschnitt aus der Kernwort-Variantenliste (KW = Kernwort der Wortfamilie).

Variante	KW	Reihe	Variante	KW	Reihe	Variante	KW	Reihe
back	back	6	bar	bër	4	bièg	bièg	2a
bâg	bâg	7	bâr	bër	4	biug	bièg	2a
bann	bann	7	bër	bër	4	bock	bièg	2a
balg	bëlg	3b	bir	bër	4	bog	bièg	2a
bëlg	bëlg	3b	bor	bër	4	boug	bièg	2a
bolg	bëlg	3b	bur	bër	4	buck	bièg	2a
bulg	bëlg	3b	barg	bërg	3b	bug	bièg	2a
ball	bëll	3b	bërg	bërg	3b	bugg	bièg	2a
bëll	bëll	3b	borg	bërg	3b	bièt	bièt	2b
boll	bëll	3b	burg	bërg	3b	bot	bièt	2b
bull	bëll	3b				bôt	bièt	2b
						but	bièt	2b

Diese Variantenliste liest das Hauptscript als externe Hilfsdatei ein und prüft, ob die jeweilige umlautbereinigte Kernwort-Variante in der linken Spalte der Liste vorkommt. Wenn ja, wird ihr das alternanzbereinigte Wortfamilien-Kernwort der rechten Spalte zugeordnet. Mit diesem Kernwort und seiner Ablautreihe werden zwei weitere Felder der Lemmaliste gefüllt. Kommt die umlautbereinigte Kernwort-Variante nicht in der Variantenliste vor, so ist sie zugleich auch die alternanzbereinigte Form und erhält die Ablautreihenanzahl 0, das heißt: sie ist synchron mittelhochdeutsch keiner der Wortfamilien starker Verben zugehörig (vgl. z. B. *tag* von *vièr-tègig* ‚viertägig‘ in Tab. 8).

Tab. 8: Beispiele aus der um das alternanzbereinigte Kernwort erweiterten Lemmaliste.

Lexem, POS	Lexem 2	Strukturformel	KW	umlaut-ber. KW	alternanz-ber. KW	Reihe
<i>bëte-gültig</i> , adj	<i>bët-e-gült-ig</i>	wXls+wXlsA	<i>bët+gült</i>	<i>gult</i>	<i>gëlt</i>	3b
<i>be-sëtzære</i> , m	<i>be-sëtz-ære</i>	plwXlsS	<i>sëtz</i>	<i>satz</i>	<i>sitz</i>	5b
<i>viir(e)-ge-zoèhe</i> , n	<i>viir(e)-ge-zoèh-e</i>	plplwXls	<i>zoèh</i>	<i>zôh</i>	<i>zièh</i>	2b
<i>vièr-tègig</i> , adj	<i>vièr-tèg-ig</i>	wX+wXlsA	<i>tèg</i>	<i>tag</i>	<i>tag</i>	0

Im letzten Schritt werden die erzeugten Datensätze der Lemmaliste nach dem alternanzbereinigten Kernwort als erstem und der Ablautreihe als zweitem Sortierschlüssel sortiert. Dadurch werden im günstigen Fall alle Mitglieder einer Wortfamilie zusammengeführt. Als dritter und vierter Sortierschlüssel empfehlen sich das umlautbereinigte Kernwort und das Ausgangslexem.

Tab. 9: Ausschnitt aus der sortierten erweiterten Lemmaliste (Anfang der Wortfamilie DWINGEN ‚zwingen‘).

Lexem, POS	Lexem 2	Strukturformel	KW	umlaut-ber. KW	alternanz-ber. KW	Reihe
<i>be-dwang</i> , m	<i>be-dwang</i>	plwX	<i>dwang</i>	<i>dwang</i>	<i>dwing</i>	3a
<i>be-dwangsals</i> , n	<i>be-dwang-sals</i>	plwXlsS	<i>dwang</i>	<i>dwang</i>	<i>dwing</i>	3a
<i>be-dwènge</i> , n	<i>be-dwèng-e</i>	plwXlS	<i>dwèng</i>	<i>dwang</i>	<i>dwing</i>	3a
<i>be-dwèngen</i> , swv	<i>be-dwèng-en</i>	plwXlV	<i>dwèng</i>	<i>dwang</i>	<i>dwing</i>	3a
<i>be-dwèngēt</i> , part adj	<i>be-dwèng-et</i>	plwXlVpart2lA	<i>dwèng</i>	<i>dwang</i>	<i>dwing</i>	3a
<i>be-dwèng-lich</i> , adj	<i>be-dwèng-lich</i>	plwXlsA	<i>dwèng</i>	<i>dwang</i>	<i>dwing</i>	3a
<i>be-dwèngnisse</i> , f	<i>be-dwèng-nisse</i>	plwXlsS	<i>dwèng</i>	<i>dwang</i>	<i>dwing</i>	3a
<i>dwang</i> , m	<i>dwang</i>	wX	<i>dwang</i>	<i>dwang</i>	<i>dwing</i>	3a

4. Fehlerdiagnose und -korrektur

Bei automatisch erzeugten Ergebnissen ist eine gründliche Fehlerdiagnose in aller Regel unerlässlich – sowohl für eine anschließende Verbesserung des Verfahrens als auch für die notwendige Korrektur seiner Ergebnisse! Bei dem hier vorgestellten Verfahren gibt es zwei Haupttypen von Fehlern:

1) Lexeme sind ihrer Wortfamilie nicht zugeordnet worden

Aufgrund von Sonderformen, die von den beteiligten Scripten nicht erfasst sind oder in der Kernwort-Variantenliste (s. 3.2.7) nicht berücksichtigt sind, oder aus anderen Gründen haben Mitglieder einer Wortfamilie nicht das richtige alternanzbereinigte Kernwort erhalten. Dieser Fehlertyp dürfte in zahlreichen Einzelfällen vorliegen, fällt aber bezogen auf die einzelnen Wortfamilien – zumal die großen – quantitativ meist wohl weniger ins Gewicht. So fehlen bei der Wortfamilie ZIEHEN nur 5 (1,1%) von 452 Fällen, bei SITZEN nur einer von 347 Fällen, bei SNĪDEN nur 5 (3,2%) von 158 Fällen (s. Tab. 10). Die Gründe sind verschieden: In einigen Fällen hat das Script Suffixe im Wortinneren nicht abgetrennt, so *-ig-* in *un-ge-zūhtig-et*, *zūhtig-ære*, *zūhtig-unge* und *-el-* in *snitzel-ære*. Fälle wie *all-zoges*, *-zuges* ließen sich durch eine auf Adverbien beschränkte Abtrennung von *-es* beheben, die daher substantivische Lexeme wie *ackes* ‚Axt‘ oder *nickes* ‚Nix, Krokodil‘ nicht betraf. In *ge-sēz(ze)* hätten vor der Suffixabtrennung die Klammern getilgt werden müssen. Für *snitt-el* und *vrôn-(e)-snitt-ære* fehlt in der Kernwort-Variantenliste der Eintrag des geminierten *snitt* (*snitt* – *snid* 1a). Bei *snitterære* und *zwèi-snīdrig* statt der gewöhnlichen Formen *snitære* und *zwèi-snīdig* handelt es sich um vereinzelte Sonderformen.

Vielfach liegt der Grund der Fehlzuordnung aber auch in der Gestalt der Lemmata. So ist nicht selten dasselbe Lexem als Kompositumsglied oder in abgeleiteten Bildungen abweichend von der Form des Simplex angesetzt, z. B.

iuwel(e) ‚Eule‘, aber *naht-*, *stein-*, *turn-iule*. Auch an nicht standardkonformen Lemmaansätzen, die gelegentlich unkorrigiert in die Bonner Lemmaliste gelangt sind, kann die Zuordnung scheitern; zwei Beispiele: Lexers Ansatz des präfigierten *ge-efern* (1, 777) stimmt nicht zu dem des Simplex *avern*, *ävern*, *äfern* ‚wiederholen‘ (1, 106); im Findebuch-Ansatz *achselspange* (Gärtner u.a. 1992: 5) verhindert *-chs-* statt *-hs-* die Zuordnung zur Wortfamilie AHSE.

Tab. 10: Falsch zugeordnete Lexeme der Wortfamilien ZIEHEN, SITZEN und SNĪDEN.

Lexem, POS	Lexem 2	Strukturformel	KW	umlaut-ber. KW	alternanz-ber. KW	Reihe
<i>all-zoges</i> , adv	<i>all-zoges</i>	wX+wX	<i>all+zoges</i>	<i>zoges</i>	<i>zoges</i>	0
<i>all-zuges</i> , adv	<i>all-zuges</i>	wX+wX	<i>all+zuges</i>	<i>zuges</i>	<i>zuges</i>	0
<i>un-ge-zühtiget</i> , part adj	<i>un-ge-zühtig-et</i>	plplwXlVpart2lA	<i>zühtig</i>	<i>zuchtig</i>	<i>zuchtig</i>	0
<i>zühtigære</i> , m	<i>zühtig-ære</i>	wXlsS	<i>zühtig</i>	<i>zuchtig</i>	<i>zuchtig</i>	0
<i>zühtigunge</i> , f	<i>zühtig-unge</i>	wXlsS	<i>zühtig</i>	<i>zuchtig</i>	<i>zuchtig</i>	0
<i>ge-sëz(ze)</i> , n	<i>ge-sëz(ze)</i>	plwX	<i>sëz(ze)</i>	<i>sëz(ze)</i>	<i>sëz(ze)</i>	0
<i>zwei-snidrig</i> , adj	<i>zwei-snidr-ig</i>	wX+wXlsA	<i>zwei+snidr</i>	<i>snidr</i>	<i>snidr</i>	0
<i>snittel</i> , n	<i>snitt-el</i>	wXlsS	<i>snitt</i>	<i>snitt</i>	<i>snitt</i>	0
<i>vrôn(e)-snittære</i> , m	<i>vrôn(e)-snitt-ære</i>	wXls+wXlsS	<i>vrôn+snitt</i>	<i>snitt</i>	<i>snitt</i>	0
<i>snitterære</i> , m	<i>snitter-ære</i>	wXlsS	<i>snitter</i>	<i>snitter</i>	<i>snitter</i>	0
<i>snitzelære</i> , m	<i>snitzel-ære</i>	wXlsS	<i>snitzel</i>	<i>snitzel</i>	<i>snitzel</i>	0

2) Lexeme einer Wortfamilie sind einer anderen Wortfamilie zugeordnet worden

Nicht selten erhalten Lexeme zweier Wortfamilien dasselbe alternanzbereinigte Kennwort. Dadurch kommt es zu einer partiellen oder kompletten Fusionierung verschiedener Wortfamilien. Dieses Defizit ist gravierender als die zuvor besprochenen Zuordnungsfehler bei einzelnen Lexemen. Zudem ist es in der Regel weder durch Verbesserungen des automatischen Verfahrens noch durch nachträgliches Ergänzen der Kernwort-Variantenliste zu beheben.

So enthält die Wortfamilie des starken Verbs *WINTEN* 3a Mitglieder mit den Ablautformen *wint*, *want* und *wunt*, denen in der Kernwort-Variantenliste das alternanzbereinigte Kernwort *wint* zugeordnet ist. Daher werden der Wortfamilie *WINTEN* im Rahmen dieses Verfahrens zwangsläufig auch die Mitglieder der Wortfamilien *wint* ‚Wind‘, *wint* ‚Windhund‘ (wendischer Hund) und *wunt* ‚wund‘ einverleibt und müssen bei der späteren manuellen Korrektur als eigene Wortfamilien abgetrennt werden (s. unten).

Ein Extrembeispiel ist die Wortfamilie von md. *ver-mëchen* ‚aufhalten, hindern, schwächen‘, die eigentlich nur aus maximal zwei Lexemen bestehen dürfte: aus dem nur in einigen mitteldeutschen, vor allem mittelfränkischen Quellen belegten *ver-mëchen* und seiner erschlossenen Basis *mëchen*. Dieser

„Kleinstfamilie“ wird über die als Ablautvariante gewertete Kernwortform *mach*⁵ die gesamte Wortfamilie MACHEN zugeschlagen, außerdem über *mich* auch *michelen* ‚vergrößern‘: Von 225 Lexemen sind damit 223 (99%) falsch zugeordnet!

Tab. 11: Beispiele von Fehlzuordnungen zur Wortfamilie MËCHEN.

Lexem, POS	Lexem 2	Strukturformel	KW	umlaut-ber. KW	alternanz-ber. KW	Reihe
<i>ab-got-machære</i> , m	<i>ab-got-mach-ære</i>	plwX+wXlsS	<i>got+mach</i>	<i>mach</i>	<i>mëch</i>	5
<i>all-ge-mëch-liche</i> , adv	<i>all-ge-mëch-liche</i>	wX+plwXlsAD	<i>all+mëch</i>	<i>mach</i>	<i>mëch</i>	5
...	
<i>mëchen</i> , swv	<i>mëch-en</i>	wXIV	<i>mëch</i>	<i>mëch</i>	<i>mëch</i>	5
<i>ver-mëchen</i> , stv5	<i>ver-mëch-en</i>	plwXIV	<i>mëch</i>	<i>mëch</i>	<i>mëch</i>	5
<i>michel(e)n</i> , swv	<i>mich-el(e)n</i>	wXlsV	<i>mich</i>	<i>mich</i>	<i>mëch</i>	5

Im Rahmen des vorgestellten Verfahrens lässt sich dieses Problem am leichtesten manuell dadurch lösen, dass die Einträge „*mach – mëch* 5“ und „*mich – mëch* 5“ aus der Kernwort-Variantenliste gestrichen werden, s. Tab. 12.

Tab. 12: Korrektur der Fehlzuordnungen zur Wortfamilie MËCHEN.

Lexem, POS	Lexem 2	Strukturformel	KW	umlaut-ber. KW	alternanz-ber. KW	Reihe
<i>ab-got-machære</i> , m	<i>ab-got-mach-ære</i>	plwX+wXlsS	<i>got+mach</i>	<i>mach</i>	<i>mach</i>	0
<i>all-ge-mëch-liche</i> , adv	<i>all-ge-mëch-liche</i>	wX+plwXlsAD	<i>all+mëch</i>	<i>mach</i>	<i>mach</i>	0
...	
<i>mëchen</i> , swv	<i>mëch-en</i>	wXIV	<i>mëch</i>	<i>mëch</i>	<i>mëch</i>	5
<i>ver-mëchen</i> , stv5	<i>ver-mëch-en</i>	plwXIV	<i>mëch</i>	<i>mëch</i>	<i>mëch</i>	5
<i>michel(e)n</i> , swv	<i>mich-el(e)n</i>	wXlsV	<i>mich</i>	<i>mich</i>	<i>mich</i>	0

Schon die vorstehenden Beispiele haben gezeigt: Die primäre Sortierung nach dem alternanzbereinigten Kernwort als dem Kandidaten für das Wortfamilie-Kernwort ergibt ein Angebot an möglichen Wortfamilien, das einer gründlichen manuellen Korrektur und Nachbearbeitung bedarf. Allgemein besteht diese Nachbearbeitung aus

- der Ergänzung fehlender Wortfamilienangehöriger, die in der sortierten Liste blockartig, aber auch vereinzelt und versprengt vorkommen können;
- der Trennung fälschlich ganz oder partiell fusionierter Wortfamilien.

5 Außerdem stützt sich die Annahme, dass (*ver*)*mëchen* starkes Verb sei, offenbar auf nur einen Beleg des Partizip II *vermëchen* in der hessisch-thüringischen ‚Kaiserchronik‘-Handschrift Heidelberg, Cpg 361: *Der lip ift mir uirmechen. Ich nemac nicht mer gesprechin*, 919.

Die einzelnen Korrekturschritte müssen auf die besonderen Gegebenheiten der fusionierten Wortfamilien abgestellt sein. Dies sei am Beispiel der Wortfamilie WINTEN (mit den inkorporierten Wortfamilien WINT und WUNT) erläutert:

- Der Datensatzblock mit dem alternanzbereinigten Kernwort *wint* wird intern nach dem umlautbereinigten Kernwort sortiert (falls dies nicht schon zuvor der dritte Sortierschlüssel war). Dadurch ergeben sich drei Blöcke mit umlautbereinigtem *want*, *wintund* und *wunt*.
- Der Block mit umlautbereinigtem *want* dürfte komplett zur Wortfamilie WINTEN ‚winden‘ gehören, der *wunt*-Block zu WINTEN und WUNT ‚wund‘ und der *wint*-Block zu WINTEN und WINT ‚Wind‘. Die Lexeme der beiden letzten Blöcke müssen daher manuell aufgeteilt und ihren Wortfamilien zugewiesen werden. Bei den zu WINTEN gehörigen Lexemen ist keine Änderung nötig. Bei den zu WUNT gehörigen Lexemen müssen alternanzbereinigtes Kernwort und Ablautreihe von „*wint* 3a“ zu „*wunt* 0“ geändert werden; bei den zu WINT gehörigen ist lediglich die Reihe in „0“ zu ändern.
- Bei der manuellen Auftrennung kann man sich wortfamilienspezifische Besonderheiten zunutze machen. Zu WINT ‚Wind‘ beispielsweise zählen nahezu alle Lexeme mit *-wint*, m (= stm.) als Kompositionsgrundwort und mit *wint-* als Kompositionserstglied (im Ausgangslexem der Lemmaliste); das ist die Hauptmasse der zu dieser Wortfamilie gehörigen Wörter.
- Ähnlich verhält es sich im *wunt*-Block: Fast alle Lexeme mit (*-wunt*, adj-*wunte*, f oder mit *wunte-*, *wunten-* als Erstglied zählen zur Wortfamilie WUNT.
- Bei einer abschließenden Durchsicht sind alle bislang noch falsch zugeordneten Fällen entsprechend zu korrigieren.
- Anschließend ist neu zu sortieren mit dem alternanzbereinigtem Kernwort als erstem, der Ablautreihe als zweitem und dem umlautbereinigtem Kernwort als drittem Sortierschlüssel.
- Dadurch sind nun jeweils zusammensortiert: 1. die Wortfamilie WINTEN (*wint* 3a) mit 482 Mitgliedern, 2. die Wortfamilie WINT (*wint* 0) mit 75 Mitgliedern und 3. die Wortfamilie WUNT (*wunt* 0) mit 58 Mitgliedern.

Die Beispiele verdeutlichen, dass der benötigte Korrekturaufwand erheblich sein wird. Da sich die vorbereitenden Schritte (s. 3.2) in Grenzen halten und das Weitere bis hin zu der nach den Wortfamilienkernwörtern sortierten Liste weitestgehend automatisch abläuft, beginnt bei dem hier vorgestellten Verfahren mit der manuellen Korrektur die eigentliche Arbeit. Der Hauptaufwand wird dabei nicht in der Kontrolle der sehr großen Wortfamilien wie WINTEN bestehen, deren Zahl überschaubar ist. Ein größeres Problem dürfte

die Masse von 5635 Einzeleinträgen⁶ sein, die in der aktuellen sortierten Liste den 4849 potentiellen Wortfamilien mit wenigstens zwei Mitgliedern gegenüberstehen. Zum Vergleich: In Spletts Ahd. Wörterbuch kommen auf ca. 1400 Wortfamilien nur ca. 720 Einzeleinträge (Splett 1993 I,1: LV–LXXVI; I,2: 1209–1240). Bei einer ähnlichen Proportion sollten im Mittelhochdeutschen auf die potentiellen 4849 Wortfamilien nur etwa 2500 Einzelträge kommen. Auch wenn der Anteil der Einzeleinträge, also der familiär isolierten Lexeme, im Mittelhochdeutschen schon wegen des höheren Lehnwortanteils⁷ deutlich größer sein dürfte als im Althochdeutschen, muss wohl damit gerechnet werden, dass eine hohe Zahl von isolierten Lexemen aus unterschiedlichen Gründen (s. oben 4.1) ein falsches (das heißt, nicht zielführendes) alternanzbereinigtes Kernwort erhalten hat. Dies zu korrigieren, ist eine mühselige und zeitraubende Aufgabe. Es ist aber zu betonen, dass sich diese Arbeit auch bei jeder anderweitigen umfassenden Wortfamilienuntersuchung des mhd. Wortschatzes nicht umgehen lässt. Bei dem hier vorgestellten Verfahren wird die Korrektur immerhin dadurch erleichtert, dass das Zusammengehörige in der sortierten Liste in der Regel nahe beieinander steht. So wird *èckern* ‚Buchecker, Eichel‘ in *buoch-èckern*⁸ über umlautbereinigtes *ackern* in der sortierten Liste zum Nachbarn von *ackeran*⁹ ‚Buchecker, Eichel‘ und kann so leicht mit ihm zu einer nur aus diesen beiden Lexemen bestehenden Wortfamilie zusammengeführt werden.

5. Ausblick

Nach Abschluss der Korrekturarbeiten lässt sich der Beitrag der mhd. Zeit zur Strukturgeschichte des deutschen Wortschatzes genauer untersuchen. So fragt sich beispielsweise: Wie verändert sich global gesehen der Bestand an Wortfamilien vom Althochdeutschen zum Mittelhochdeutschen und weiter zum Neuhochdeutschen? Welche Wortfamilien sterben wann aus und welche entstehen wann neu? Bleibt dabei der Anteil verbaler und nominaler Kernwörter gleich oder wandelt auch er sich und wenn ja, in welcher Weise?

6 Das sind Lexeme, deren alternanzbereinigtes Kernwort in der Liste nur einmal vorkommt.

7 Nach Stichproben scheint etwa die Hälfte der isolierten Lexeme aus Lehnwörtern zu bestehen.

8 Von Lexer 1, 386 nur in Diefenbachs Glossarium aus einem Vokabular des 15. Jh.s belegt, daher im MWB nicht aufgenommen.

9 In MWB 1, 96 nach einer Sonderform der Urkunde Corpus Nr. 679 (Kloster Andlau/ Elsass 1284) 91,46, 92,9f. kaum zu Recht als *ackeram* statt *ackeran* angesetzt, vgl. Kluge/Seebold (2002: 226 s.v. Ecker). Zum Suffixablaut *-ana/-ina-* in germ. **akr-ana-* vgl. Much (1909: 215f.); daher ist *eckern* (< **akr-ina-*) unter Wortbildungsaspekt ein besonderes Lexem und nicht nur Variante von *ackeran*.

Beschränkt man sich zunächst auf einzelne (schon hinreichend korrigierte) Wortfamilien, so lassen sich deren Schrumpfs- und Wachstumsprozesse bereits vorab untersuchen. Als Beispiele dafür mögen abschließend die Wortfamilien ahd. BIOTAN / mhd. BIETEN und ahd. KIOSAN / mhd. KIESEN ‚wählen, prüfen‘ dienen. Vom Althochdeutschen her gesehen setzen sich etwa 53% der Lexeme der Wortfamilie BIOTAN im Mittelhochdeutschen fort; das ist etwas mehr als der Durchschnitt von 47% auch im Mittelhochdeutschen belegten ahd. Lexemen. Die Wortfamilie KIOSAN entspricht mit ca. 48% dagegen dem Durchschnitt. Aus mittelhochdeutscher Warte stellt sich der Vergleich so dar: Nur 17% der mhd. Lexeme haben ahd. Vorläufer, 83% kommen neu hinzu. Davon weicht die Wortfamilie BIETEN mit 13% aus dem Althochdeutschen übernommener und 87% neuer Lexeme nur wenig ab. Bei KIESEN ist dagegen der Anteil schon ahd. belegter Lexeme mit 26% doppelt so hoch wie bei BIETEN und der Anteil neuer Lexeme mit 76% entsprechend deutlich niedriger. Diese Zahlenverhältnisse weisen möglicherweise darauf hin, dass die Wortfamilie BIETEN lebenskräftiger ist als KIESEN: BIETEN hat vom Althochdeutschen zum Mittelhochdeutschen prozentual weniger Lexeme verloren als KIESEN und in mhd. Zeit deutlich mehr Lexeme hinzugewonnen. Bei KIESEN deutet sich in dem unterdurchschnittlichen Lexemzuwachs in mhd. Zeit vielleicht schon ein Vitalitätsverlust an, der auf den fast kompletten Untergang dieser Wortfamilie im Neuhochdeutschen vorausweisen könnte.

Literatur

- Ganslmayer, Christine. 2012. *Adjektivderivation in der Urkundensprache des 13. Jahrhunderts. Eine historisch-synchrone Untersuchung anhand der ältesten deutschsprachigen Originalurkunden* (Studia Linguistica Germanica 97). Berlin & Boston: de Gruyter.
- Gärtner, Kurt, Christoph Gerhardt, Jürgen Jährling, Ralf Plate, Walter Röll & Erika Timm. 1992. *Findebuch zum mittelhochdeutschen Wortschatz, mit einem rückläufigen Index*. Stuttgart: Hirzel.
- Herbers, Birgit. 2002. *Verbale Präfigierung im Mittelhochdeutschen. Eine semantisch-funktionale Korpusanalyse* (Studien zur Mittelhochdeutschen Grammatik 1). Tübingen: Niemeyer.
- Klein, Thomas, Hans-Joachim Solms & Klaus-Peter Wegera (eds.). 2009. *Mittelhochdeutsche Grammatik. Teil III: Wortbildung*. Tübingen: Niemeyer.
- Kluge, Friedrich. 2002. *Etymologisches Wörterbuch der deutschen Sprache*. Bearb. von Elmar Seebold. 14. Aufl. Berlin & New York: de Gruyter.
- Leipold, Aletta. 2006. *Verbableitung im Mittelhochdeutschen. Eine synchron-funktionale Analyse der Motivationsbeziehungen suffixaler Verbwortbildungen*. Tübingen: Niemeyer.
- Much, Rudolf. 1909. Got. haiþnô. *Zeitschrift für deutsche Wortforschung* 11. 211–218.

- MWB = Gärtner, Kurt, Klaus Grubmüller & Karl Stackmann (eds.). 2006ff. *Mittelhochdeutsches Wörterbuch*. Bd. 1: *a – êvrouwe* (2013). Bd. 2: Doppellieferung 1/2: *êvüegerinne – geværlich* (2014); Doppellieferung 3/4: *gevaterne – hanken* (2017). Stuttgart: Hirzel.
- Ring, Uli. 2008. *Substantivderivation in der Urkundensprache des 13. Jahrhunderts. Eine historisch-synchrone Untersuchung anhand der ältesten deutschsprachigen Originalurkunden*. Berlin & New York: de Gruyter.
- Seebold, Elmar. 1970. *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben* (Janua linguarum. Series practica 85). Den Haag & Paris: Mouton.
- Splett, Jochen. 1993. *Althochdeutsches Wörterbuch. Analyse der Wortfamilienstrukturen des Althochdeutschen, zugleich Grundlegung einer zukünftigen Strukturgeschichte des deutschen Wortschatzes*. 3 Bde. Berlin & New York: de Gruyter.
- Splett, Jochen. 2009. *Deutsches Wortfamilienwörterbuch. Analyse der Wortfamilienstrukturen der deutschen Gegenwartssprache, zugleich Grundlegung einer zukünftigen Strukturgeschichte des deutschen Wortschatzes*. 18 Bde. Berlin & New York: de Gruyter.

Prof. em. Dr. Thomas Klein
 Rheinische Friedrich-Wilhelms-Universität Bonn
 Institut für Germanistik, Vergleichende Literatur- und Kulturwissenschaft
 Am Hofgarten 22
 D-53113 Bonn
 thomas.klein@uni-bonn.de

Lisa Dücker

Die Getrennt- und Zusammenschreibung von Komposita in frnhd. Hexenverhörprotokollen

Abstract: This paper investigates the spelling of compound nouns in a corpus comprised of Early New High German protocols of witch trials from the 16th and 17th century. Previous studies on the spelling of compound nouns in printed texts have found that scribes increasingly write compound nouns as one word during the 16th century. However, this paper will show that there is still much variation in handwritten texts from that time. The study focusses on identifying factors that lead scribes to write compound nouns either as one word or two, such as linking elements and the use of upper case letters. I will argue that while there is more variation in the spelling of compound nouns in the handwritten corpus than in printed texts, there still is a strong tendency to line up the boundaries of the graphemic and syntactic words.

1. Einleitung

Die Zusammen- und Getrenntschreibung von Komposita in der Geschichte des Deutschen wurde in der Vergangenheit hauptsächlich an Drucken untersucht (vgl. u.a. Solling 2012, 2016; Kopf 2017; Schindler 1999; Pavlov 1972, 1983, 2004). Zu handschriftlichen Texten liegen hingegen vereinzelte Studien vor: So analysiert V. Moser (1982) zwei literarische Handschriften Fischarts. Haubold (1914) vergleicht wiederum handschriftliche Manuskripte und Bieldrucke von Martin Luther und stellt dabei unter anderem Folgendes fest: „Über die Schreibung der zusammengesetzten Worte ist zu sagen, daß sie im Druck genau so schwankend ist wie in der Handschrift“ (Haubold 1914: 24). Wie genau sich aber die Regularitäten bei der Worttrennung von Komposita in Handschriften und Drucken unterscheiden, darüber gibt es bisher keine gesicherten Ergebnisse.

Im Fokus des vorliegenden Artikels steht deshalb, ob die Kriterien, die für die Spatiensetzung in gedruckten, meist lektorierten Texten erarbeitet wurden, auch auf Handschriften zutreffen, und welche Unterschiede es gibt. Dies wird anhand eines Korpus von Hexenverhörprotokollen aus dem SiGS-Projekt¹

1 Das SiGS-Korpus ist im Rahmen des DFG-finanzierten Projekts zur „Entwicklung der satzinternen Großschreibung“ (kurz: SiGS), das in der ersten Phase an zwei Universitäten angesiedelt war (Uni Hamburg und Uni Münster), entstanden (SZ 280/2-1 und KO 909/12-1: Münsteraner Gruppe (Klaus-Michael Köpcke, Mark Schutzzeichel), Hamburger Gruppe (Fabian Barteld, Renata Szczepaniak). AnnotatorInnen: Annemarie Bischoff, Lisa Dücker, Julia Hübner, Johanna Legrum, Katja Politt, Nicolai Pudimat, Eleonore Schmitt, Annika Vieregge und Nicholas

untersucht. Diese Texte können als (semi-) spontan gelten und sind als Teil von Gerichtsakten für einen mittleren Grad von Öffentlichkeit bestimmt (vgl. Topalović 2003; Szczepaniak & Barteld 2016: 45).

In Kap. 2 wird zunächst der Reanalyseprozess von der Genitivphrase zum uneigentlichen Kompositum nachvollzogen und in Kap. 3 der Forschungsstand zur Schreibung von Komposita im Frnhd. zusammengetragen. Auf Grundlage des Konzepts des graphematischen Worts nach Fuhrhop (2008) werden in Kap. 4 alle Substantivkomposita aus 56 Hexenverhörprotokollen aus der Zeit von 1570 bis 1665 ausgewertet. Dabei wird der Zusammenhang zwischen Worttrennung und Majuskelsetzung (Kap. 4.6) ebenso beleuchtet wie der Einfluss von Fugenelementen (Kap. 4.5) und der Frequenz (Kap. 4.7).

2. Substantivkomposita im Frühneuhochdeutschen

Das Wortbildungsmuster der Komposition gewinnt im Frnhd. immer mehr an Bedeutung (vgl. Solling 2012: 37). Zu dem aus dem Germanischen ererbten Muster der sog. eigentlichen Komposition tritt nun verstärkt die sog. uneigentliche hinzu. Eigentliche Komposita werden gebildet, indem zwei Substantive mit Hilfe eines (fakultativen) Stammbildungselements verbunden werden (ahd. *sper-a-scaft*, *tag-a-lôn*). Dieses Wortbildungsmuster ist im Frnhd. nicht mehr produktiv; die Stammbildungselemente sind größtenteils ausgefallen oder zu Schwa abgeschwächt (frnhd. *speer-Ø-schaft*, *tag-e-lohn*, vgl. Szczepaniak 2016: 329).

Uneigentliche Komposita sind hingegen diachron betrachtet aus Genitivphrasen entstanden. Demske (1999, 2001) sieht die Entstehung der Determiniererphrase als Ursache für deren Wandel zum Substantivkompositum: Da die vorangestellten Genitivattribute den neu entstandenen Determiniererslot in der Phrase nicht ausfüllen können, setzt ein Stellungswandel ein. Genitivattribute werden zunehmend nachgestellt und die verbliebenen vorangestellten Attribute werden als Kompositionserstglieder reanalysiert:

[[*der Brücken*] *Zoll*] > [*der [Brücken Zoll]*] > [*des [Brücken Zolls]*] > [*des [Brückenzolls]*]

Durch diesen Entstehungsprozess der Komposita aus Genitivphrasen werden an der Grenze zwischen den beteiligten Substantiven ehemalige Flexionsendungen als Fugenelemente konserviert (*Brücke-n-zoll*, *Teufel-s-tanz*; vgl. Nübling & Szczepaniak 2011: 53–55).

Wieling). Die zweite Phase ist an der Universität Hamburg (Lisa Dücker, Stefan Hartmann, Renata Szczepaniak) angesiedelt (SZ 280/2-3). Ich möchte allen an dem Projekt Beteiligten für ihre Arbeit danken. Mein Dank gilt außerdem den anonymen Gutachtern, die mit hilfreichen Hinweisen zur Verbesserung dieses Artikels beigetragen haben.

Im Zuge dieses Prozesses ändern sich die Kongruenzverhältnisse innerhalb der Nominalphrase: Der Determinierer kongruiert nun nicht mehr mit dem ehemaligen Genitivattribut (*der Brücken*), sondern mit dem zweiten Substantiv (*des Zolls*), das als Kopf des Kompositums analysiert wird. Auch auf graphematischer Ebene schlägt sich die Reanalyse nieder, indem die ursprünglich voneinander getrennt geschriebenen Substantive vermehrt auch ohne Spatium vorkommen. Pavlov (2004) konstatiert für frnhd. Texte gar, dass dieses graphematische Kriterium für die Abgrenzung von Genitivkonstruktionen und Komposita hinreichend ist (ähnlich auch Nitta 1987; Solms 1999). Er nimmt an, dass die Schreibung als ein oder zwei graphematische Wörter ein direkter Ausdruck dessen ist, was ein frnhd. Schreiber als ein oder zwei Wörter *empfindet* (Pavlov 2004: 106):

Unserer Ansicht nach handelt es sich bei den Schreibungen wie *lant hërre* vs. *lanthërre* oder *shâf stîge* vs. *shâfstîge* um Wortgruppe auf der einen und Kompositum auf der anderen Seite, trotzdem die beiden Konstruktionsweisen im gegebenen Fall semantisch und funktional, als lexematische Einheiten, identisch sind. (Pavlov 2004: 106)

Wie Szczepaniak & Barteld (2016: 56) und Barteld, Szczepaniak & Zinsmeister (2014: 254–255) anhand der handschriftlichen Hexenverhörprotokolle zeigen, die auch der vorliegenden Untersuchung zugrunde liegen (s.u. Kap. 4.1), kann keineswegs eine 1:1-Zuordnung zwischen graphematischen und syntaktischen Wörtern für das Frnhd. gezogen werden. Sie verweisen darauf, dass diese Uneinheitlichkeit der Grenzen von syntaktischem und graphematischem Wort nicht nur bei Komposita, sondern auch bei Präposition-Artikel-Klisen und weiteren Phänomenen zu beobachten ist (vgl. <auffm Teufelß dantz>, Alme 1630²).

Auch wenn die Zusammen- oder Getrenntschreibung für die vorliegende Studie nicht als hinreichendes Kriterium für die Analyse als Kompositum oder Genitivphrase angesehen werden kann, ist die Schreibung ohne Spatium zweifelsohne ein wichtiger Schritt auf dem Weg von der Phrase hin zu der Analyse als ein nicht nur syntaktisches, sondern auch graphematisches Wort (vgl. Fuhrhop 2008). Im nächsten Kapitel soll deshalb der aktuelle Forschungsstand zur Entwicklung der Worttrennung bei Komposita und mögliche Einflussfaktoren zusammengetragen werden.

Die Unterteilung in eigentliche und uneigentliche Komposita ist nicht unproblematisch (vgl. Solling 2012: 34–36). Bei starken Feminina, die im Genitiv keine Endung zeigen, kann bspw. keine morphologisch basierte Unterscheidung zwischen Kompositionserstglied und Genitivobjekt getroffen

2 Bei Beispielen wird mit einer Chiffre aus Entstehungsort und -zeit nach Macha et al. (2005) auf den entsprechenden Korpus text referiert.

werden. In der vorliegenden Untersuchung wird deshalb in der Auswertung mit der Unterteilung in verfügte und unverfügte Komposita gearbeitet. Die Termini eigentliche und uneigentliche Komposita werden jedoch im Theorie- teil zum Vergleich mit anderen Studien herangezogen.

3. Frnhd. Komposita als graphematische Wörter

Ein graphematisches Wort ist laut Fuhrhop (2008: 193) eine ununterbrochene Graphemkette, die zwischen zwei Spatien steht. Prototypische graphematische Wörter erlauben zudem nur Großschreibung am Wortanfang (vgl. Fuhrhop 2008: 194), aber prinzipiell können sie Wortzeichen wie Bindestriche oder Apostrophe enthalten (z. B. *M'gladbach*, *Software-Entwickler*). Auch die Binnenmajuskelschreibung bei Komposita (mit oder ohne Bindestrich) führt laut Fuhrhop (2008: 225) nicht dazu, dass diese als zwei graphematische Wörter interpretiert werden: „Eine Schreibung wie *BahnCard* erscheint zwar einerseits als künstliche Schreibung, kann aber quasi als Zwischenstufe zwischen einfachen Komposita (*Bahnkarte*) und Bindestrichkomposita gesehen werden“ (Fuhrhop 2008: 225; Hervorhebungen im Original).

In Dücker & Szczepaniak (2017) wurde ein Modell mit verschiedenen Gra- den der graphematischen Wortwertigkeit von Komposita erarbeitet (Tabelle 1).

Tab. 1: *Graphematische Wortwertigkeit von Komposita (nach Dücker & Szcze- paniak 2017: 49).*

zwei graphematische Wörter			ein graphematisches Wort		
mit Binnenmajuskel (und fakultativer Initialmajuskel)	ohne Majuskel	mit Initialmajuskel	mit Binnenmajuskel (und fakultativer Initialmajuskel)	ohne Majuskel	mit Initialmajuskel
<i>Teufels Tanz</i> <i>teufels Tanz</i>	<i>teufels</i> <i>tanz</i>	<i>Teufels tanz</i>	<i>TeufelsTanz</i> <i>teufelsTanz</i>	<i>teufelstanz</i>	<i>Teufelstanz</i>

Alle Komposita auf der rechten Seite der Darstellung können jeweils als ein graphematisches Wort gelten. Obwohl alle drei Unterkategorien mit dem syntaktischen Wort übereinstimmen, gibt es dennoch Abstufungen: Das graphematische Wort ohne Majuskel stellt im Deutschen das unmarkierte graphematische Wort dar. Eine Majuskelsetzung betont den Anfang eines graphematischen Worts. Eine Initialmajuskel hebt somit ein Kompositum besonders hervor, eine Binnenmajuskel wie in *TeufelsTanz* betont hingegen die einzelnen Bestandteile der Zusammensetzung und verschlechtert die Wahrnehmung als *ein* zusammengehöriges graphematisches und syntaktisches Wort.

Die Schreibung als zwei graphematische Wörter, die auf der linken Seite der Darstellung zu sehen ist, erschwert die Analyse als ein syntaktisches Wort.

Das gilt insbesondere dann, wenn die zweite Konstituente mit einer Majuskel beginnt. Großschreibung beim Bestimmungsglied hebt hingegen auch hier den Beginn des syntaktischen Worts besonders hervor.

Auch wenn Schreibungen auf der linken Seite der Tabelle mit frühen Entwicklungsstadien der Substantivkomposition assoziiert werden und die unmarkierte Schreibung im Nhd. diejenige ist, die ganz rechts steht, verläuft die graphematische Entwicklung nicht so gradlinig für alle Substantivkomposita, wie eine solche Darstellung vermuten lassen könnte. Stattdessen zeigen sich große Unterschiede zwischen den eigentlichen und uneigentlichen Komposita.

Noch für das 14. und die erste Hälfte des 15. Jhs. wird in bisherigen Studien ein hoher Anteil an Zusammenschreibung von eigentlichen Komposita dokumentiert.³ Ab dem Ende des 15. Jahrhunderts gilt dies jedoch nicht mehr für alle Substantivkomposita gleichermaßen: Pavlov (1972, 1983) zeigt für den Zeitraum von 1470–1530, dass eigentliche Komposita 81,5% Zusammenschreibung (Token) aufweisen. Die nun häufiger auftretenden uneigentlichen Komposita, entstanden aus Phrasen, werden noch häufiger getrenntgeschrieben; in Pavlovs (1983: 33–34) Daten werden sie nur in 33,4% der Belege zusammengeschrieben. Um 1700 finden sich wiederum insgesamt knapp 3% Getrenntschreibung, die mehrheitlich auf uneigentliche Komposita entfallen (Pavlov 1983: 105–106, 2009: 64–65).

Die Arbeiten von Solling (2012, 2016), der gedruckte Predigten aus den Jahren 1550–1710 untersucht, zeigen ein ähnliches Bild. Wie in Tab. 2 dargestellt, steigt bei den eigentlichen Komposita der Anteil der getrenntgeschriebenen Belege nie über 6%, während bis zu 45% der uneigentlichen Komposita mit *Spatium* geschrieben werden. Ab Zeitschnitt V (1660) kommt Getrenntschreibung bei allen Komposita nur noch selten vor (<7% pro Zeitschnitt), was auf den starken Anstieg der Bindestrichschreibung zurückzuführen ist.

3 Für einen Überblick über die Worttrennung vom Ahd. an vgl. Solling 2012: 73–98.

Tab. 2: Worttrennung bei *eigentlichen* und *uneigentlichen* Komposita 1550–1710 nach Solling (2016: 296).

	Eigentliche Komposita			Uneigentliche Komposita		
	zusammen	mit Bindestrich	getrennt	zusammen	mit Bindestrich	getrennt
Zeitschnitt I (1550)	94% (280)	0	6% (19)	55% (74)	0	45% (60)
Zeitschnitt II (1570)	97% (1046)	0	3% (35)	69% (346)	0	31% (152)
Zeitschnitt III (1600)	96% (603)	<0,5% (1)	4% (26)	79% (209)	0	21% (55)
Zeitschnitt IV (1620)	96,5% (631)	<0,5% (1)	3,5% (22)	88% (354)	0	12% (47)
Zeitschnitt V (1660)	85% (544)	13% (82)	2% (12)	53,5% (219)	40,5% (165)	6% (24)
Zeitschnitt VI (1710)	46% (382)	53,5% (444)	0,5% (4)	17,5% (101)	78,5% (448)	4% (22)

Sowohl bei *eigentlichen* als auch bei *uneigentlichen* Komposita sinkt somit der Anteil der Getrennschreibung ab der zweiten Hälfte des 16. Jhs. stark, *uneigentliche* Komposita zeigen über den gesamten Zeitraum hinweg jedoch wesentlich höhere Anteile von Spatiensetzung als die *eigentlichen*.

Kopf (2017: 197) sieht in ihrer Studie, die sich auf ein Korpus von kirchlichen Texten und Sachtexten aus dem Zeitraum von 1500 bis 1710 stützt, eine noch deutlichere Abnahme der Worttrennung bei verfügbaren Komposita: 1500 werden 81% aller Komposita mit Fugenelement getrenntgeschrieben, 1710 sind es nur noch 8%. Bei den unverfügbaren Komposita sinkt der Anteil an Getrennschreibung im gleichen Zeitraum von 22% auf 2% (vgl. Kopf 2017: 197).

Während nach relativ konstanter Zusammenschreibung bis zur Mitte des 15. Jhs. die Getrennschreibung von *eigentlichen* Komposita um 1500 Werte von knapp 20% erreicht, sinkt der Anteil im 16. und 17. Jh. bei beiden Klassen von Substantivkomposita so sehr, dass um 1700 kaum noch getrenntgeschriebene Komposita vorkommen. *Eigentliche* und *uneigentliche* Komposita nähern sich also in ihrer graphischen Gestalt einander immer mehr an. Dabei weisen *uneigentliche* Komposita in den bisherigen Studien durchweg einen höheren Anteil an Getrennschreibung auf.

Sowohl Pavlov (2004: 108) als auch Schindler (1999: 318) und Solling (2016: 300) machen das Aufkommen der *uneigentlichen* Komposita für die vorübergehende Zunahme der Getrennschreibung bei den *eigentlichen* Komposita verantwortlich. Pavlov (2004: 108) sieht die Gründe dafür in einer „gegenseitige[n] funktionale[n] Annäherung“ der beiden Kompositionstypen. Diese

führte einerseits zur breiteren Übernahme des getrennten Schriftbildes der Fügungen mit dem vorangestellten flexionslosen Attribut, andererseits aber kamen die genitivischen Komplexe funktional auf gleiche Linie mit den ‚Stammkomposita‘, was ihre Zusammenschreibung stimulierte. (Pavlov 2004: 108)

Auch Solling (2016) erklärt die Schwankungen bei der Worttrennung von Substantivkomposita in gedruckten Predigttexten mit wechselseitigen Analogiebeziehungen zwischen eigentlichen und uneigentlichen Komposita. Dass die uneigentlichen Komposita am Ende seines Untersuchungszeitraums nur noch 4% Getrenntschreibung aufweisen, wertet Solling (2016: 299) als ein Zeichen dafür, dass diese nun als Wortbildungsmuster etabliert sind und (wie eigentliche Komposita) zunehmend als ein syntaktisches Wort interpretiert werden:

In the first part of the period covered by the study [um 1550, L.D.], the artificial compound nouns [uneigentliche Komposita, L.D.] adapted to match the genuine ones [eigentliche Komposita, L.D.], but the opposite tendency can be observed over the last two thirds of the 17th century. (Solling 2016: 300)

Neben dem Auftreten von Fugenelementen ist die Majuskelsetzung beim Kopf entscheidend für die Getrennt- oder Zusammenschreibung von Komposita. Zwar kommen in den Daten von Solling alle Kombinationen von Getrennt-, Zusammen- und Bindestrichschreibung mit Groß- oder Kleinschreibung des Kopfs vor, doch werden Komposita mit kleingeschriebenem Kopf beinahe durchweg zusammengeschrieben, „während Substantivkomposita, in denen eine Binnenmajuskel vorhanden ist, dagegen in den früheren Zeitschnitten überwiegend getrennt geschrieben und in den späteren Zeitschnitten vorwiegend mit einem Bindestrich in der Fuge auftreten“ (Solling 2012: 179).

Weidman (1941: 96) stellt zudem für den von ihm untersuchten Codex Manesse fest, dass frequente Lemmata häufiger zusammengeschrieben werden als weniger frequente. Die Ergebnisse von Solling (2012: 267) zeigen hingegen, dass in den Predigttexten frequente Komposita nicht durchweg häufiger Zusammenschreibung aufweisen. Stattdessen schwanken Komposita, die in seinem Korpus häufig vorkommen, nicht so sehr in Bezug auf die Worttrennung wie weniger frequente Komposita, sondern zeigen entweder überwiegend Getrennt- oder überwiegend Zusammenschreibung. Diese beiden Beobachtungen können als Hinweise auf eine stärkere Konventionalisierung („Lexikalisierung“) der Schreibung frequenter Lemmata interpretiert werden. Im Gegensatz dazu werden Hapax Legomena häufiger getrennt als zusammengeschrieben (vgl. Solling 2012: 101, Pavlov 1983: 28–29).

Zur Zusammen- und Getrenntschreibung von Substantivkomposita speziell in Handschriften gibt es bisher nur wenige empirische Untersuchungen: V. Moser (1982: 418) stellt für die von ihm untersuchten Handschriftenbruchstücke fest, dass am Ende des 16. Jhs. „zusammenschreibung [sic] das

gewöhnliche“ ist und Getrenntschreibung bei eigentlichen wie uneigentlichen Komposita nur „vereinzelt“ belegt ist. Haubold (1914) findet in seiner Studie keinen signifikanten Unterschied in der Getrenntschreibung von Komposita zwischen den Drucken und Handschriften Martin Luthers. Folglich wäre davon auszugehen, dass in einem Korpus von handschriftlichen Texten aus dem Zeitraum vom letzten Drittel des 16. Jhs. bis zur Mitte des 17. Jhs. überwiegend mit Zusammenschreibung von Substantivkomposita zu rechnen ist.

4. Korpusstudie in frnhd. Hexenverhörprotokollen

Die Analyse der Schreibung von Komposita in Hexenverhörprotokollen zeigt insgesamt ein wesentlich uneinheitlicheres Bild, als auf Grundlage der bisherigen Forschung zu erwarten wäre. In Kap. 4.1. wird das SiGS-Korpus vorgestellt und in Kap. 4.2 und 4.3 dargelegt, wie Substantivkomposita sowie Getrennt- und Zusammenschreibung für die Korpusanalyse definiert bzw. operationalisiert wurden. Kap. 4.4 bietet einen Überblick über die Zusammen- und Getrenntschreibung der Komposita, bevor anschließend Faktoren für die Worttrennung überprüft werden, darunter das Auftreten von Fugenelementen (4.5), die Majuskelsetzung bei den Kompositionsgliedern (4.6) und Lexikalisierung häufiger Lemmata (4.7).

4.1. Das SiGS-Korpus

Abb. 1: SiGS-Gesamtkorpus.



Im Gegensatz zu bisherigen Analysen von Substantivkomposita wird die vorliegende Untersuchung in einem Korpus aus 56 handschriftlichen Hexenverhörprotokollen, die von Macha et al. (2005) ediert wurden (vgl. Abb. 1), durchgeführt. Die Texte stammen aus der Zeit zwischen 1570 und 1665 und sind nach Macha et al. (2005) in sechs Regionen eingeteilt: Nordwestdeutschland (Region I), Nordostdeutschland (Region II), Mittelwestdeutschland (Region III), Mittelostdeutschland (Region IV), Südwestdeutschland (Region V) und Südostdeutschland (Region VI). Für die Untersuchung der zeitlichen Entwicklung wurde eine Einteilung in Zeitschnitte von jeweils zehn Jahren gewählt (1570–1579, 1580–1589 etc.).⁴

Es handelt sich bei den Texten um Verhörmit- und -abschriften, die in der Regel von einem anonymen, aber geübten Schreiber verfasst wurden (vgl. Szczepaniak & Barteld 2016: 47–48).⁵ Sie enthalten zwischen 550 und 2848 graphische Token und sind in der Kanzleikursive abgefasst. Das Korpus wurde mithilfe des Programms GATE (Cunningham et al. 2014) linguistisch aufbereitet und annotiert. Für diese Studie wurden stichprobenartig Faksimiles der handschriftlichen Akten zur Überprüfung der Transkription herangezogen.

4.2. Operationalisierung der Substantivkomposita

In die Untersuchung gehen alle Komposita ein, die in ihrer Bedeutung transparent sind und im Frnhd. eindeutig aus zwei Substantiven bestehen, was anhand des *Deutschen Wörterbuchs* der Gebrüder Grimm überprüft wurde. So werden Belege mit Erstgliedern, die potentiell einen Verbstamm darstellen (bspw. *Bademagd* oder *Spielmann*) und stark lexikalisierte Belege wie *Herzog* oder *Sonntag* ausgeschlossen (anders: Solling 2012). Wortbestandteile wie *-werk* in *Handwerk*, die im Untersuchungszeitraum zwar als eigenständige Substantive vorkommen, aber bereits dabei sind, sich zu Affixoiden zu entwickeln, werden als Kompositionsbestandteile gewertet.⁶ Zudem werden ausschließlich Appellativa betrachtet: Zusammengesetzte Eigennamen (*Rheinwiesen*, *Rosenfeld*) werden von der Untersuchung ausgeschlossen. Zusätzlich werden Belege ausgeschlossen, bei denen aus der Transkription hervorgeht,

4 Dieses Vorgehen führt zu Unterschieden in der Größe der einzelnen Teilkorpora. So sind für die Hochzeit der Hexenverfolgung in den 1620er-Jahren mehr Texte bei Macha et al. (2005) ediert als aus den Zeitschnitten zu Beginn und am Ende des Untersuchungszeitraums.

5 Für Informationen über den Entstehungsprozess einzelner Texte s. Macha et al. (2005), die zu jedem transkribierten Protokoll, wo möglich, Informationen über den Status des Dokuments als Mit- oder Abschrift und die Anzahl und Identität der Schreiber vor jedem Transkript angeben.

6 Das gilt allerdings nicht für *-tum*, *-schaft* und *-heit*, die bereits im Mhd. ihren Status als freie Lexeme eingebüßt haben (vgl. Klein, Solms & Wegera 2009: 158).

dass ein Großteil des Worts bzw. eines Kompositionsglieds aus aufgelösten Abkürzungen besteht, bspw. <gg> für *Goldgulden*. Erstglieder koordinierter Komposita werden ebenfalls nicht aufgenommen; in Beispielen wie *Rueben oder Pflanzgärtl* (München 1600) wird nur der zweite Teil (in diesem Fall *Pflanzgärtl*) in die Untersuchung miteinbezogen.

Die Unterscheidung zwischen Substantivkomposita und Phrasen mit vorangestelltem Genitivattribut erfolgt im Rahmen dieser Untersuchung nach dem morphologischen Kriterium der Artikelkongruenz (vgl. Dücker & Szczepaniak 2017). Die Schreibung als ein oder zwei Wörter hat hingegen keinen Einfluss auf ihre Interpretation als Kompositum. Wie in Kap. 2 beschrieben, unterscheiden sich Genitivphrasen und Substantivkomposita dadurch, dass der Determinierer bei einer Genitivphrase mit dem Attribut (vgl. Bsp. (1)) und bei einem Kompositum mit dessen Kopf kongruiert (vgl. Bsp. (2)).⁷

- (1) Genitivkonstruktion: *solchem Pauren Auch Ain Rosß Inns Theüfels Nammen Zue Todt geschlagenn* (Rosenfeld 1603)
- (2) Kompositum: *Seie sie bei ¼ Jahr auf einem hexen tanz beim Salzstadel gewesen* (Memmingen 1665)

Wie in Kap. 2 erwähnt wurde, wird für die Auswertung keine Unterteilung in eigentliche und uneigentliche Komposita vorgenommen. Stattdessen wird zwischen verfügten und unverfügten Komposita unterschieden. Ein Fugenelement wird hier definiert als das Wortmaterial, das zwischen Bestimmungsglied und Kopf eines Kompositums steht und nicht zum Nominativ Singular des Bestimmungsglieds gerechnet werden kann (vgl. Kürschner & Szczepaniak 2013: 1). Dabei ist es unerheblich, ob es sich um ein aus einem ehemaligen Genitivflexiv entstandenes Fugenelement handelt (*gericht+s+personen*, Gaugrehweiler 1610) oder um den erstarrten Kompositionsvokal eines eigentlichen Kompositums (*hertz+e+leidt*, Perleberg 1588).⁸

4.3. Operationalisierung der Getrennt- und Zusammenschreibung

Die Operationalisierung von Getrennt- und Zusammenschreibung erfolgt wegen der Besonderheiten von handschriftlichen Texten nicht nach den Kriterien, die Solling (2012) auf Grundlage von Saenger (1997) anwendet. Dass bisherige Studien zur Worttrennung bei Komposita sich auf gedruckte Texte

7 Für das Vorgehen in Fällen mit ambiger Determiniererkongruenz vgl. Dücker & Szczepaniak (2017: 38–39). Belege, die auch mit dem dort vorgestellten semantischen Kriterium nach Solling (2011, 2012) nicht eindeutig erfasst werden können, werden von der Untersuchung ausgeschlossen. Es handelt sich dabei um 24 Belege.

8 Solling (2012: 35–36) geht bei der Einteilung anders vor; er zählt Belege vom Typ *Herzeleid* zu den eigentlichen Komposita.

konzentriert haben, liegt auch daran, dass gängige Methoden zur Spatienerkennung hier einfacher anwendbar sind: „The decisive factor here [bei der Wahl von Drucken als Untersuchungskorpus, L.D.] was that it is easier to determine whether a word is written as one word or two in printed texts than in manuscripts“ (Solling 2016: 291).

Für gedruckte Texte kann nach Saenger (1997: 27) ganz allgemein gelten, dass ein Abstand zwischen zwei Buchstaben, der 1,5-mal so groß ist wie „the distance between the minum stroke of the ‚lowercase‘ letters *u*, *m*, or *n*“, eine Grenze zwischen zwei Wörtern konstituiert. Diese recht mechanische Vorgehensweise ist für die handschriftlichen Mitschriften, aus denen das hier untersuchte Korpus besteht, allerdings nicht anwendbar. Busch & Fleischer (2015: 566) nennen das Spatium in handschriftlichen Texten einen „skalare[n] Begriff“ und weisen darauf hin, dass es für Handschriften „keineswegs eine Definition, ab welchem Abstand ein worttrennendes Spatium vorliegt“, gibt (Busch & Fleischer 2015: 566). Zur Bestimmung von Spatien setzen sie den horizontalen Abstand zwischen zwei Graphien mit der vertikalen Höhe der umgebenden Graphien ohne Ober- und Unterlänge ins Verhältnis: Entspricht der horizontale Abstand mindestens der Hälfte der vertikalen Höhe, sprechen sie von einem „deutliche[n] Spatium“ (Busch & Fleischer 2015: 568).

Die vorliegende Untersuchung stützt sich vorrangig auf die Transkription von Macha et al. (2005). Die Edition liefert aber leider keine Hinweise darauf, wie die Unterscheidung zwischen Getrennt- und Zusammenschreibung von Wörtern gemacht wurde. In der vorliegenden Untersuchung wird deshalb eine konservative Operationalisierung der Getrennt- und Zusammenschreibung von Komposita vorgenommen: Prinzipiell wird von Getrenntschreibung der Kompositionsglieder ausgegangen. Zwei Wörter gelten nur dann als zusammengeschieden, wenn sie in der Handschrift durch eine physische Linie miteinander verbunden sind, da nur in solchen Fällen eindeutig davon ausgegangen werden kann, dass der Schreiber Zusammenschreibung intendiert hat. Dazu kommen einige Regeln für besondere Fälle: Enthält das Kompositum bspw. auch innerhalb einer Konstituente Buchstaben, die nicht miteinander verbunden sind, so gilt das Kompositum als getrenntgeschrieben, wenn der Abstand zwischen den Buchstaben an der Grenze zwischen den Konstituenten größer ist als der im Inneren der Konstituente.

17 der Transkriptionen von Macha et al. (2005) wurden auf Grundlage dieser Kriterien mit Faksimiles der Handschriften verglichen. Da die Transkription nur in sehr wenigen Fällen von den hier vorgestellten Maßstäben abweicht, werden die Texte von Macha et al. (2005) als authentische Darstellung der Zusammen- und Getrenntschreibung von Komposita in frnhd. Handschriften angesehen. So wird es möglich, erstmals auch handschriftliche historische Texte in Bezug auf die Getrennt- und Zusammenschreibung von

Komposita zu untersuchen. Gedruckte Texte gehen von der Konzeption bis zur Drucksetzung durch viele verschiedene Hände und unterliegen oftmals Veränderungen durch Lektorat und Drucker. Die Untersuchung handschriftlicher Texte verspricht entsprechend einen direkteren Einblick in spontane Schriftlichkeit und die „online-Graphematik“ der Schreibenden (vgl. Szczepaniak & Barteld 2016: 45).

4.4. Zusammen- und Getrenntschreibung im SiGS-Korpus

Anhand der in Kap. 4.3 genannten Kriterien wurden insgesamt 880 Komposita ermittelt, die sich auf 472 unterschiedliche Types verteilen. Dies entspricht einem Anteil von 5,1% an allen im Korpus vorhandenen Substantiven (n=17152). In den einzelnen Texten sind Substantivkomposita unterschiedlich oft vertreten: In Helmstedt 1580 ist bspw. nur ein einziges Kompositum belegt (*kerkhoue*), was einem Anteil von 0,6% an allen in dem Protokoll enthaltenen Substantiven entspricht. In Dillenburg 1631 sind hingegen 31 Komposita verzeichnet, die 14,4% aller Substantive ausmachen.

Das Verhältnis zwischen getrennt- und zusammengeschriebenen Komposita ist wesentlich ausgeglichener als in den Daten von Solling (2012): Von den 880 im Korpus befindlichen Substantivkomposita werden 529 (60,1%) aller Belege zusammengeschrieben. Darunter befinden sich 46 Komposita, die an der Kompositionsfuge am Zeilenende mit einem Bindestrich getrennt werden und ein Beleg mit Bindestrichschreibung innerhalb der Zeile (*Sondern Ihr daumen und Schien-schrauben Angelegt worden*, Rhens 1629). Von den 351 (39,9%) getrenntgeschriebenen Belegen werden 55 an ihrer Kompositionsfuge am Zeilenende getrennt (vgl. Tab. 3).⁹

9 Die am Zeilenende getrennten Komposita mit und ohne Bindestrich werden hier nicht als eigenständige Gruppen aufgefasst, sondern im weiteren Verlauf der Untersuchung werden sie jeweils als Untergruppe der zusammen- bzw. getrennt geschriebenen Komposita betrachtet. Dieses Vorgehen wird dadurch begründet, dass sich die am Zeilenende mit Bindestrich getrennten Komposita in Bezug auf die hier untersuchten Eigenschaften genauso verhalten wie die zusammengeschriebenen Belege. Auch die Komposita mit Trennung innerhalb der Zeile und am Zeilenende ohne Bindestrich verhalten sich gleich. Die Aufteilung in Trennung am Zeilenende mit und ohne Bindestrich ist deshalb notwendig, da nur bei Bindestrichsetzung ein klarer Wille zur Zusammenschreibung erkennbar ist.

Tab. 3: *Getrennt- und Zusammenschreibung im Gesamtkorpus.*

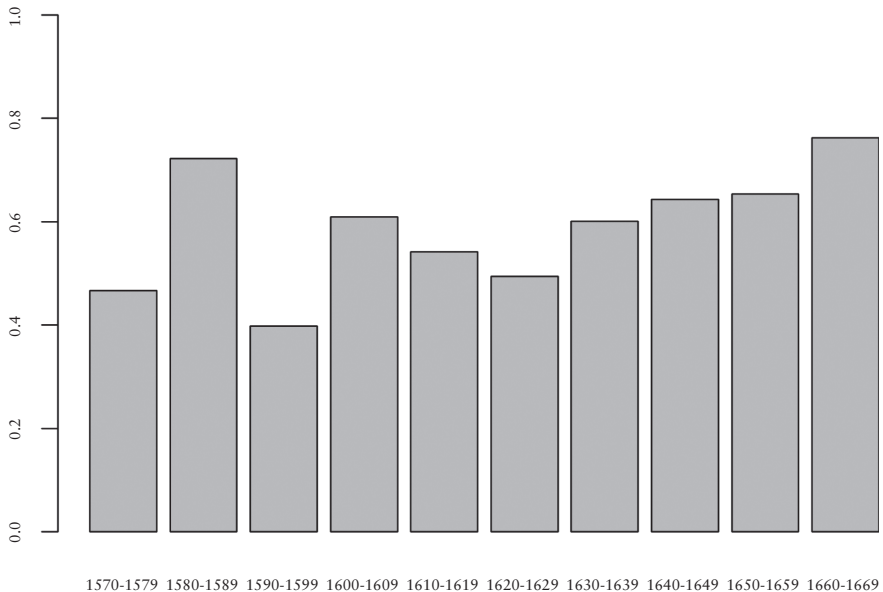
Getrennt- und Zusammenschreibung	Absoluter Wert	Prozentwert	Beispiel
Getrennt	351	39,9%	<i>Ross koth</i> (Baden-Baden 1628)
davon am Zeilenende	55		<i>Donnerstag Abendt</i> (Werl 1630)
zusammen	529	60,1%	<i>stutenbrei</i> (Alme 1630)
davon mit Bindestrich am Zeilenende	46		<i>hosenbandt</i> (Schweinfurt 1616)
davon mit Bindestrich innerhalb der Zeile	1		<i>Schien-schrauben</i> (Rhens 1629)
Gesamt	880		

Diese Zahlen sind weit von Sollings (2012) Ergebnissen entfernt, dessen Daten für die Zeitschnitte 1550 bis 1620 stets mehr als 80% Zusammenschreibung aufweisen (vgl. Solling 2012: 121); in den handschriftlichen Texten herrscht deutlich mehr Getrenntschreibung vor (60,1%). Wie Abbildung 2 zeigt, wird diese Marke für keinen der hier untersuchten Zeitschnitte erreicht: Zwei von drei Zeitschnitten vor 1600 weisen weniger als die Hälfte zusammengeschriebene Komposita auf; der Höchstwert liegt im letzten Zeitschnitt bei 76,2%.

Über den gesamten Zeitraum hinweg ist allerdings ein Anstieg der Zusammenschreibung zu beobachten (Kendall's $\tau = 0,6$).¹⁰

10 Der hohe Anteil an Zusammenschreibung im zweiten Zeitschnitt ist auf den starken Einfluss des Schreibers des Protokolls Perleberg 1588 zurückzuführen: Bis auf einen Beleg stammen alle in diesem Zeitraum erfassten Komposita aus diesem Text. In die Berechnung von Kendall's τ sind alle Belege aus dem Zeitschnitt eingegangen.

Abb. 2: Anteil der Zusammenschreibung pro Zeitschnitt ($n=880$, Kendall's $\tau = 0,6$, $p=0,01667$).



Bei der räumlichen Verteilung ergeben sich in Bezug auf den Anteil der Zusammenschreibung keine signifikanten Unterschiede zwischen den Regionen (vgl. Tab. 4). Der höchste Anteil an Zusammenschreibung findet sich im westlichen Mitteldeutschland (64,6%) und östlichen Norddeutschland (64,5%), während der niedrigste Anteil an zusammengesetzten Komposita im östlichen Mitteldeutschland gemessen wird (47,8%). Es ist kein eindeutiger Ausbreitungsverlauf der Zusammenschreibung von Substantivkomposita erkennbar, was sich mit den Ergebnissen von Solling (2012: 160) und Pavlov (1983: 12) deckt.

Tab. 4: Anteil der Zusammenschreibung pro Dialektgebiet ($\chi^2=8,1207$, $p=0,1497$).

	Westdeutschland	Ostdeutschland	gesamt
Norddeutschland	94/164 (57,3%)	49/76 (64,5%)	143/240 (59,6%)
Mitteldeutschland	153/237 (64,6%)	33/69 (47,8%)	186/306 (60,1%)
Süddeutschland	76/121 (62,8%)	124/213 (58,2%)	200/334 (59,9%)
gesamt	323/522 (61,9%)	206/358 (57,5%)	

In fünf von sechs Regionen wird die Mehrzahl der Komposita zusammengeschieden. Im Gegensatz zu dem in Dücker & Szczepaniak (2017) untersuchten SiGS-Kernkorpus, das 18 Protokolle umfasst, ist nach der Auswertung

des SiGS-Gesamtkorpus mit 56 Texten kein eindeutiges Ost-West-Gefälle mehr erkennbar; die westdeutschen Protokolle weisen (anders als im Kernkorpus) mehr Zusammenschreibung auf als die ostdeutschen Texte. Dies ist mit der Heterogenität innerhalb der Dialekträume zu erklären, auf die auch Schutzeichel & Szczepaniak (2015) in Bezug auf den regionalen Einfluss der Substantivgroßschreibung hinweisen. Wichtiger als die Region ist für die Zusammenschreibung möglicherweise der Idiolekt oder auch der Einfluss einzelner Schreibschulen und Kanzleien, was sich allerdings aufgrund der fehlenden Daten zur Identität und Ausbildung der jeweiligen Schreiber derzeit nicht nachweisen lässt (vgl. Fußnote 5).

4.5. Zusammenhang zwischen Worttrennung und Verfugung

Von den 880 Substantivkomposita, die in den Hexenverhörprotokollen vorkommen, weisen 255 ein Fugenelement auf (29,0%). Beinahe drei Viertel aller Belege sind somit unverfugt (71,0%). Im SiGS-Korpus sind dabei sechs unterschiedliche Fugenelemente belegt:

- (1) *-e-*: *Schwineborsten* (Flensburg 1608)
- (2) *-s-*: *weibsbildt* (Bamberg 1628)
- (3) *-es-*: *kindes händtel* (Münüchen 1600)
- (4) *-n-*: *katzenbirn* (Lemberg 1630)
- (5) *-en-*: *Narren werk* (Mergentheim 1629)
- (6) *-er-*: *Kinder tauff* (Ahaus 1608)

Die Hexenverhörprotokolle lassen keine zeitliche Entwicklung bezüglich des Verhältnisses von unverfugten und verfugten Komposita erkennen (vgl. Tab. 5). Der Anteil der Komposita mit Nullfuge¹¹ liegt über den gesamten Zeitraum bei mindestens zwei Dritteln.

Tab. 5: Zeitliche Entwicklung des Anteils von unverfugten und verfugten Komposita (Kendall's $\tau = 0,06667$).

Zeitschnitt	Anzahl der Texte pro Zeitschnitt	Anzahl aller Wörter pro Zeitschnitt	unverfugte Komposita	verfugte Komposita	gesamt
1570–1579	1	1127	11	4	15
1580–1589	2	2206	17	1	18
1590–1599	8	11366	69	39	108
1600–1609	4	7571	48	21	69

11 Darunter werden hier neben Komposita, in denen die beiden Wortstämme zusammengefügt werden (*Gaißbokh*, Bamberg 1628), auch Komposita mit Reduktionsfuge (*Schulmeister*, Georgenthal 1597) verstanden.

Zeitschnitt	Anzahl der Texte pro Zeitschnitt	Anzahl aller Wörter pro Zeitschnitt	unverfugte Komposita	verfugte Komposita	gesamt
1610–1619	10	12229	108	25	133
1620–1629	13	21303	166	71	327
1630–1639	11	17297	121	62	183
1640–1649	5	10104	52	18	70
1650–1659	1	2287	17	9	26
1660–1669	1	1463	16	5	21
gesamt	56	86953	625	255	

Solling (2012: 295) und Kopf (2017: 197) weisen darauf hin, dass der Status als uneigentliches oder eigentliches Kompositum bzw. das Auftreten eines Fugenelements einen großen Einfluss auf die Zusammen- oder Getrenntschreibung eines Kompositums hat. Die Daten aus den Hexenverhörprotokollen stützen diese Befunde (vgl. Tab. 6).

Tab. 6: Worttrennung und Verfugung ($\chi^2=102,73$, $p<0,0001$).

	Zusammenschreibung	Getrenntschreibung	gesamt
Fugenelement	86 (33,6%)	169 (66,3%)	255 (100%)
kein Fugenelement	443 (70,9%)	182 (20,7%)	625 (100%)
Gesamt	529	351	

Verfugte und unverfugte Belege zeigen in den untersuchten Daten ein komplementäres Bild: Während Komposita mit Fugenelement in nur einem Drittel aller Fälle zusammengeschrieben werden, zeigen Belege mit Nullfuge über 70% Zusammenschreibung.

Die Aufschlüsselung in Komposita mit und ohne Fugenelemente in Abb. 3 und 4 zeigt deutlich, wie sehr die Verfugung die Worttrennung der Komposita beeinflusst: Unverfugte Komposita wie *beinschraube* (Göttingen 1649) weisen bereits zu Beginn des Untersuchungszeitraums einen Anteil von über 60% Zusammenschreibung auf, der in der Folge nur einmal unterschritten wird. Bis auf den Zeitschnitt 1600–1609 (49,3%) weisen die unverfugten Komposita stets überwiegend Zusammenschreibung auf. Im letzten Zeitschnitt werden 87,5% aller Komposita mit Nullfuge zusammengeschrieben. Auch wenn der Anteil der Zusammenschreibung nicht kontinuierlich ansteigt, so ist über den gesamten Untersuchungszeitraum hinweg doch ein Trend hin zu mehr Zusammenschreibung zu beobachten (Kendall's $\tau = 0,64$).

Abb. 3: Zeitliche Entwicklung der Zusammenschreibung bei unverfügten Komposita (Kendall's $\tau=0,64$, $p=0,009148$).

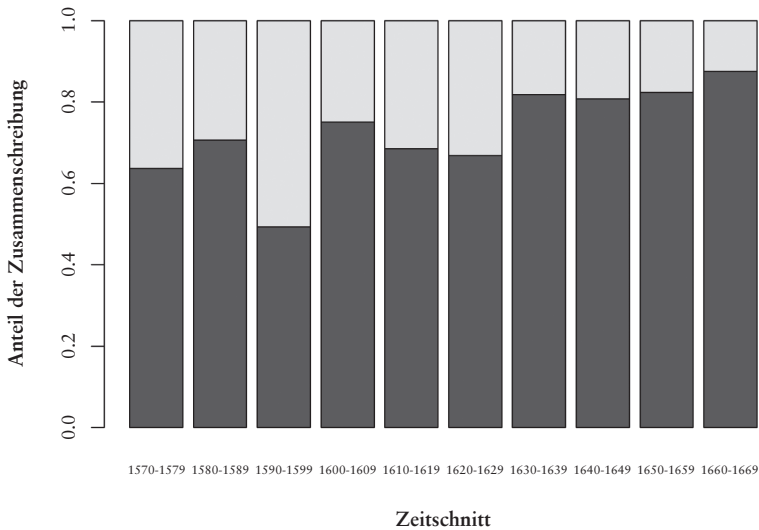
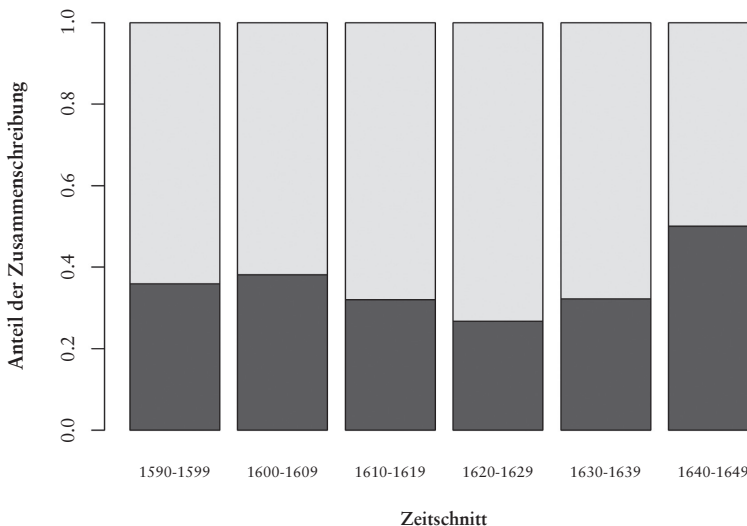


Abb. 4: Zeitliche Entwicklung der Zusammenschreibung bei verfügbaren Komposita (Kendall's $\tau=0,0667$, $p=1$)¹².



12 Für die graphische Darstellung der Entwicklung der Worttrennung bei verfügbaren Komposita wurden nur die Zeitschnitte ausgewählt, für die mindestens zehn Belege vorliegen. Die zwei frühesten und die zwei spätesten Zeitschnitte sind entsprechend nicht in der Graphik abgebildet.

Auch bei verfügten Komposita wie *teufels hure* (Leipzig 1640) ist am Ende des Untersuchungszeitraums mehr Zusammenschreibung zu beobachten als zu Beginn, jedoch ist hier keine eindeutige Aufwärtskurve auszumachen, sondern vielmehr ein Auf und Ab, bei dem die zusammengeschrriebenen Komposita jedoch nie in der Mehrzahl sind (Kendall's $\tau = 0,0667$). Im Zeitschnitt 1640–1649 stehen neun zusammengeschrriebene neun getrennten Komposita gegenüber; in allen anderen Zeitschnitten überwiegt die Getrenntschrreibung. Der Anteil der Zusammenschreibung ist bei den Komposita mit Fugenelement also deutlich geringer als in den gleichen Zeitschnitten bei den unverfügten Komposita.

In Bezug auf die Verfügung lässt sich somit festhalten, dass unverfügte Komposita in späten Zeitschnitten besonders zur Zusammenschreibung tendieren, während Zusammensetzungen mit Fugenelementen häufiger getrenntgeschrieben werden, unabhängig vom Entstehungsjahr des entsprechenden Textes.

Angesichts der Tatsache, dass sich das Verhältnis zwischen verfügten und unverfügten Komposita im Zeitraum von 1570 bis 1669 nicht wesentlich verändert und die unverfügten Komposita so eindeutig in der Mehrzahl sind, stellt sich außerdem die Frage, ob die verfügten Komposita im 15. Jh. überhaupt schon frequent genug waren, um einen analogischen Sog auf die unverfügten auszuüben und zu mehr Getrenntschrreibung bei den unverfügten Komposita zu führen, wie in der bisherigen Forschungsliteratur angenommen wird.

4.6. Zusammenhang zwischen Worttrennung und Majuskelsetzung

Für den Status als graphematisches Wort ist neben der Zusammen- oder Getrenntschrreibung auch die Majuskelsetzung von großer Bedeutung (vgl. Kap. 3, Tab. 1). Im SiGS-Korpus sind alle möglichen Kombinationen aus Majuskelsetzung bei Bestimmungsglied oder Kopf und Getrennt- oder Zusammenschreibung des Kompositums belegt (vgl. Tab. 7).

Tab. 7: Majuskelsetzung und Worttrennung (n=608).¹³

Majuskelsetzung bei Bestimmungsglied und Kopf	Zusammenschreibung	Getrennschreibung	gesamt
Groß + klein	219	96	315
klein + klein	110	87	197
Groß + Groß	23	58	81
klein + Groß	3	12	15
Gesamt	355	253	

Am häufigsten ist mit 219 Belegen jedoch die zusammengeschiedene Kombination aus majuskelhaltigem Erstglied und kleingeschriebenem Kopf (*Melthau*, Ellingen 1590), doch auch durchgängige Kleinschreibung (*gerichtsheffen*, Gaugrehweiler 1610) ist mit 197 Belegen häufig vertreten. Großschreibung des Kopfs kommt überwiegend in Kombination mit einem majuskelhaltigen Erstglied vor (*Garten Thuer*, Ostrau 1628). Nur dreimal sind hingegen Komposita belegt, die ein kleingeschriebenes Erstglied und einen großgeschriebenen Kopf nicht mit einem Spatium trennen (z. B. *kunstGeygerin*, Leipzig 1640).

Es gibt keinen signifikanten Zusammenhang zwischen der Worttrennung und der Großschreibung des Bestimmungsglieds ($\chi^2=3,1518$, $p=0,07584$); vielmehr wird der Anfangsrand des Erstglieds unabhängig von der Worttrennung bevorzugt mit einer Majuskel markiert, was die Wortwertigkeit der gesamten Konstruktion hervorhebt. Der Zusammenhang zwischen der Worttrennung und der Großschreibung des Kopfs ist hingegen hochsignifikant ($p<0,0001$): Nur 26 von 96 Komposita mit majuskelhaltigem Kopf werden zusammengeschieden (*AmbtMan*, Dillenburg 1631). Wird der Kopf kleingeschrieben, überwiegt stattdessen die Zusammenschreibung (329 von 512 Belegen), wodurch die Wortwertigkeit auch auf graphematischer Ebene betont wird. Diese Werte bestätigen die Ergebnisse, die Dücker & Szczepaniak (2017: 47) bereits für das SiGS-Kernkorpus gezeigt haben. Lediglich der Anteil an Binnenmajuskelschreibungen ist im Gesamtkorpus wesentlich höher (15,8%) als bei der Auswertung von 18 Texten (7,9%).

13 Prinzipiell ist bei der in den Hexenverhörprotokollen verwendeten Kanzleikursive eine Unterscheidung zwischen Majuskeln und Minuskeln möglich. Macha et al. (2005: XXIII) stellten bei der Transkription jedoch Probleme bei der Unterscheidung von <v>/<V>, <h>/<H> und <z>/<Z> fest. Aus diesem Grund sind alle Komposita, deren Bestimmungsglied oder Kopf mit einem dieser drei Grapheme beginnt, von der Untersuchung in Bezug auf die Majuskelsetzung ausgeschlossen, wodurch sich die Gesamtzahl hier auf 608 reduziert.

Majuskelsetzungen, die die graphematische Wortwertigkeit des Kompositums beeinträchtigen können, werden also im Allgemeinen vermieden.¹⁴ Weitere Gemeinsamkeiten zwischen Komposita mit majuskelhaltigem Kopf lassen sich nicht erkennen: Sie stammen aus allen Dialektgebieten und sind in den meisten Zeitschnitten vertreten. Auch in Bezug auf Lexik, Semantik oder die Stellung am Zeilenanfang lässt sich kein Einfluss auf die Großschreibung des Zweitglieds erkennen.

4.7. Das Lemma als Faktor

Als weiterer Faktor für die Getrennt- und Zusammenschreibung von Komposita wird das Lemma in den Blick genommen. Dabei wird untersucht, ob bestimmte Lemmata zu der Schreibung mit oder ohne Spatium tendieren und ob es sich dabei um besonders frequente Komposita handelt.

Mithilfe des R-Pakets *Collostructions* (Flach 2017, Version 0.1.0) wird eine distinktive Kollexemanalyse durchgeführt (vgl. Gries & Stefanowitsch 2004). Dabei werden zwei Konstruktionen miteinander verglichen und berechnet, in welcher der Konstruktionen (hier Getrenntschreibung und Zusammenschreibung) ein bestimmtes Lemma häufiger auftritt, als bei einer zufälligen Verteilung zu erwarten wäre. Als Maß der Assoziationsstärke wird in dieser Untersuchung die Log-Likelihood Ratio verwendet. Tab. 8 enthält die zehn Types mit der höchsten Log-Likelihood Ratio für Zusammenschreibung, während Tab. 9 die zehn Types mit der größten Affinität zur Getrenntschreibung auflistet. Die Daten sind jeweils absteigend nach der Stärke der Assoziation angeordnet.

Tab. 8: Distinktive Kollexemanalyse, Assoziation: Zusammenschreibung.

Type	Tokenanzahl Getrennt- schreibung	Tokenanzahl Zusammen- schreibung	Stärke der Assoziation	Signifikanzniveau
<i>hausfrau</i>	2	45	18.37306	*****
<i>amtman</i>	0	18	9.27824	*****
<i>rathaus</i>	1	16	6.21995	*****
<i>steinmetzer</i>	0	10	5.12359	*****

14 Von den 26 Belegen mit großgeschriebenem Kopf bei Zusammenschreibung der Kompositionsglieder entfallen allein neun auf die Berufsbezeichnung *SteinMetzer* (Dillenburg 1631). Dieser Begriff wird stets gebraucht, um einen bestimmten Mann zu bezeichnen. Es handelt sich dabei somit möglicherweise um eine konventionalisierte Schreibung (vgl. Kap. 4.2) oder einen Eigennamen, worauf auch Macha et al. (2005: 192) hinweisen, die den Titel der Gerichtsakte als „Inquisitio contra Hansen Holttschenhäwer sonsten Steinmetzger genant zu Dillenburgk“ wiedergeben.

Type	Tokenanzahl Getrennt- schreibung	Tokenanzahl Zusammen- schreibung	Stärke der Assoziation	Signifikanzniveau
<i>ehebruch</i>	0	9	4.60778	****
<i>bürgermeister</i>	1	12	4.39068	****
<i>burgmann</i>	1	9	3.0759	***
<i>schafmeister</i>	0	6	3.06497	***
<i>kirchtag</i>	0	6	3.06497	***
<i>ehemann</i>	1	8	2.65286	**

Tab. 9: Distinktive Kollexemanalyse, Assoziation: Getrenntschreibung.

Type	Tokenanzahl Getrennt- schreibung	Tokenanzahl Zusammen- schreibung	Stärke der Assoziation	Signifikanzniveau
<i>teufelssalbe</i>	7	0	6.47023	*****
<i>nachtzeit</i>	6	0	5.54069	*****
<i>mannsperson</i>	4	0	3.68687	***
<i>teufelspulver</i>	4	0	3.68687	***
<i>dienstverwalter</i>	3	0	2.76256	**
<i>gerichtsperson</i>	3	0	2.76256	**
<i>leinentuch</i>	3	0	2.76256	**
<i>reichstaler</i>	3	0	2.76256	**
<i>teufelskind</i>	3	0	2.76256	**
<i>verdachtspunkt</i>	3	0	2.76256	**

Belege für *hausfrau* und *amtman*n kommen demnach überzufällig häufig mit Zusammenschreibung vor, *teufelssalbe* und *nachtzeit* bevorzugen hingegen die Schreibung mit Spatium. Es zeigt sich, dass die stark zur Zusammenschreibung tendierenden Komposita insgesamt häufiger im Korpus vertreten sind als solche, die zur Getrenntschreibung tendieren, deren Frequenz relativ niedrig ist. Nur die zwei Types, die am stärksten zur Getrenntschreibung tendieren, sind häufiger als fünfmal im Korpus belegt. Entsprechend werden für die Assoziation mit Zusammenschreibung häufiger höhere Signifikanzniveaus erreicht. Bei einem Type wie *dienstverwalter*, der mit nur drei Token belegt ist, fällt ein Beleg mit Zusammenschreibung wesentlich mehr ins Gewicht als ein Beleg mit Spatium bei *hausfrau*, das auch bei Getrenntschreibung bei zwei von 47 Token die stärkste Assoziation zur Zusammenschreibung zeigt. So lässt sich auch der auffällige Umstand erklären, dass diverse der mit Zusammenschreibung assoziierten Types auch in getrennter Schreibweise vorliegen, die am deutlichsten zur Getrenntschreibung tendierenden Types aber ausschließlich mit Spatium geschrieben werden.

Unter den Komposita, die zur Zusammenschreibung tendieren, befinden sich nur solche, deren Nennform laut *Deutschem Wörterbuch* kein Fugenelement aufweisen und die entsprechend nicht als uneigentliche Komposita gelten. Mit Getrenntschreibung sind hingegen sowohl verfugte (*teufelskind*) als auch unverfugte Komposita (*dienstverwalter*) assoziiert. Die verfügten Komposita sind mit acht Lemmata allerdings dominant. Hier zeigt sich erneut der in Kap. 3.3 beobachtete Zusammenhang zwischen Verfugung und Worttrennung. Mit Blick auf einen möglichen analogischen Sog in Bezug auf die Getrenntschreibung, wie er in der Forschungsliteratur vorgeschlagen wurde, muss aber auch hier auf die niedrigen Tokenzahlen bei den verfügten Komposita verwiesen werden.

Ein Urteil über den Zusammenhang zwischen Tokenzahl und Tendenz zur Zusammenschreibung, wie ihn Pavlov (1983) formuliert, kann anhand dieser Daten nicht gefällt werden, da Types, die ein ausgeglichenes Verhältnis zwischen Zusammen- und Getrenntschreibung zeigen, mit der distinktiven Kollexemanalyse nicht erfasst werden.

Ein Blick auf den Anteil der Zusammenschreibung bei den zehn häufigsten Komposita im SiGS-Korpus (vgl. Tab. 10) zeigt aber deutlich, dass bei neun von zehn Lemmata die Zusammenschreibung bei weitem überwiegt.

Tab. 10: Anteil der Zusammenschreibung bei den frequentesten Komposita.

Lemma	Anteil der Zusammenschreibung
<i>hausfrau</i>	45/47
<i>amtman</i>	18/18
<i>rathaus</i>	16/17
<i>bürgermeister</i>	12/13
<i>burgmann</i>	9/10
<i>kuhhirte</i>	5/10
<i>steinmetzer</i>	10/10
<i>ehebruch</i>	9/9
<i>ehemann</i>	8/9
<i>teufelstanz</i>	4/9

Von den häufigsten Komposita sind nur *kuhhirte* und *teufelstanz* nicht in Tab. 8 enthalten. Sie sind die einzigen frequenten Komposita, die nicht stark zu Zusammenschreibung tendieren, sondern ein ausgewogenes Verhältnis zwischen Getrennt- und Zusammenschreibung aufweisen. *Teufelstanz* ist zudem das einzige verfugte Kompositum, das mehr als acht Token im SiGS-Korpus aufweist. Anders als bei Solling (2012: 267), in dessen gedruckten Predigttexten frequente Komposita je nach Zeitschnitt entweder mit Bindestrich, mit Spatium oder zusammengeschrieben werden, zeigen die Hexenverhörprotokolle eine

klare Tendenz dafür, häufig vorkommende Komposita zusammenzuschreiben. Genau wie bei Solling (2012) ist diese Beobachtung nicht an einzelne Texte gebunden, sondern auch bei Lemmata wie *hausfrau* zu sehen, die in Texten aus allen Regionen und Zeitschnitten belegt sind.

Um einer möglichen Lexikalisierung der Komposita nachzuspüren, wurden die zehn häufigsten Komposita im *Deutschen Wörterbuch*, dem mhd. Wörterbuch von Lexer sowie im ahd. Wörterbuch von Schützeichel nachgeschlagen. Wo möglich, wurde das ahd. Wörterbuch von Karg-Gasterstädt & Frings herangezogen.¹⁵ Alle in Tabelle 10 enthaltenen Zusammensetzungen sind im DWB belegt und nur *teufelstanz* hat keinen Eintrag bei Lexer. Im ahd. Wörterbuch sind wiederum nur *amtman* (ahd. *ambabtman*) und *steinmetzer* (ahd. *steinmezzo*) aufgelistet. Damit steht einerseits fest, dass die im SiGS-Korpus frequenten Komposita überwiegend schon in mhd. Zeit vorkommen. Ob dies andererseits aber ausreicht, um daraus eine Lexikalisierung als Grund für die Zusammenschreibung abzuleiten, ist fraglich, zumal fünf der am stärksten mit Zusammenschreibung assoziierten Komposita auch mit Spatium vorkommen. Für das SiGS-Korpus lässt sich somit festhalten, dass die frequentesten Komposita überwiegend unverfugte sind und normalerweise zusammengeschieden werden und dass Komposita, die durchgehend getrenntgeschrieben werden, im Vergleich weniger frequent sind. Inwiefern eine mögliche Lexikalisierung von frequenten Lemmata einen Einfluss auf die Worttrennung v.a. bei eigentlichen Komposita hat, muss nach wie vor als Desiderat angesehen werden.

5. Fazit: Die zunehmende graphematische Wortwertigkeit der frnhd. Komposita

Die Untersuchung von frnhd. handschriftlichen Hexenverhörprotokollen hat gezeigt, dass unterschiedliche Faktoren dafür verantwortlich sind, ob ein Kompositum mit oder ohne Spatium geschrieben wird. Neben einer zeitlichen Entwicklung hin zu mehr Zusammenschreibung hat das Auftreten von Fugenelementen eine blockierende Wirkung und führt eher zu Getrenntschreibung. Auch ist die Majuskelsatzung beim Kopf des Kompositums ein guter Indikator für Getrenntschreibung. Ob das Bestimmungsglied groß oder klein geschrieben wird, ist hingegen nicht von Bedeutung. Auch eine hohe Tokenfrequenz wirkt sich begünstigend auf die Zusammenschreibung von Substantivkomposita aus.

15 Das *Deutsche Wörterbuch*, das mhd. Wörterbuch von Lexer und das ahd. Wörterbuch von Karg-Gasterstädt & Frings wurden über www.woerterbuchnetz.de eingesehen (gesehen am 8. Juni 2017).

Die höchste Wahrscheinlichkeit für eine Zusammenschreibung der Konstituenten ist gegeben, wenn das Kompositum in einem Text aus der Mitte des 17. Jhs. häufig auftaucht, es nicht verfugt ist und der Kopf nicht mit Majuskel geschrieben wird. Ein Spatium ist hingegen in solchen Fällen wahrscheinlich, wenn ein verfugtes Kompositum am Ende des 16. Jhs. nur selten vorkommt und einen majuskelhaltigen Kopf aufweist.

Somit sind in den handschriftlichen Hexenverhörprotokollen dieselben Faktoren einflussreich, die bereits für die Spatiensetzung in Drucken herausgearbeitet wurden. Der Anteil der Getrenntschreibung ist aber wesentlich höher als in gedruckten Texten aus der gleichen Zeit.

In Kap. 3 wurde eine graphische Darstellung von verschiedenen Stufen der graphematischen Wortwertigkeit frnhd. Komposita vorgestellt. Tab. 11 zeigt, wie sich die Belege aus dem SiGS-Gesamtkorpus in dieses Schema eingliedern.

Tab. 11: Die graphematische Wortwertigkeit der im Korpus belegten Komposita.

zwei graphematische Wörter			ein graphematisches Wort		
mit Binnenmajuskel (und fakultativer Initialmajuskel)	ohne Majuskel	mit Initialmajuskel	mit Binnenmajuskel (und fakultativer Initialmajuskel)	ohne Majuskel	mit Initialmajuskel
<i>Teufels Tanz</i> <i>teufels Tanz</i>	<i>teufels</i> <i>tanz</i>	<i>Teufels tanz</i>	<i>TeufelsTanz</i> <i>teufelsTanz</i>	<i>teufelstanz</i>	<i>Teufelstanz</i>
70 Belege	87 Belege	96 Belege	26 Belege	110 Belege	219 Belege

Deutlich ist zu sehen, dass die frnhd. Schreiber alle ihnen zur Verfügung stehenden Möglichkeiten in Bezug auf Worttrennung und Majuskelsetzung ausnutzen. Sowohl bei den Komposita mit als auch ohne Spatium nimmt die Belegzahl von links nach rechts ständig zu, d.h. dass immer mehr Substantivkomposita eine Gestalt mit möglichst hoher graphematischer Wortwertigkeit annehmen. Auch insgesamt gesehen nimmt die Anzahl der Belege zu, je höher die Wortwertigkeit eines Kompositums einzuschätzen ist. Die Ausnahme bilden dabei diejenigen Komposita, die trotz Zusammenschreibung einen majuskelhaltigen Kopf aufweisen, was die Interpretation als ein syntaktisches Wort erschwert. Somit zeigt sich bei aller Varianz eine klare Tendenz hin zu einer graphematischen Gestalt, die im Einklang mit dem Status als syntaktisches Wort ist.

Literatur

Althochdeutsches Wörterbuch. Auf Grund der von Elias v. Steinmeyer hinterlassenen Sammlungen im Auftrag der Sächsischen Akademie der Wissenschaften zu Leipzig.

- Bearbeitet und herausgegeben von Elisabeth Karg-Gasterstädt und Theodor Frings. Leipzig: 1952–2015ff.
<http://awb.saw-leipzig.de> (accessed 08 June 2017).
- Barteld, Fabian, Renata Szczepaniak & Heike Zinsmeister. 2014. The definition of tokens in relation to words and annotation tasks. In V. Henrich, E. Hinrichs, D. de Kok, P. Osenova & A. Przepiórkowski (eds.), *Thirteenth International Workshop on Treebanks and Linguistic Theories (TLT13)*. December 12–13, 2014, Tübingen, Germany. 250–257.
<http://tlt13.sfs.uni-tuebingen.de/tlt13-proceedings.pdf> (letzter Aufruf 13.06.2017).
- Busch, Nathanael & Jürg Fleischer. 2015. Zusammenschreibung im Althochdeutschen und Altniederdeutschen bis ca. 975. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 137(4). 563–598.
- Cunningham, Hamish, Diana Maynard & Kalina Bontcheva. 2014. *GATE Developer*. General Architecture for Text Engineering. Version 8.3. Sheffield.
- Demske, Ulrike. 1999. Case Compounds in the History of German. In Matthias Butt & Nanna Fuhrhop (eds.), *Variation und Stabilität in der Wortstruktur: Untersuchungen zu Entwicklung, Erwerb und Varietäten des Deutschen und anderer Sprachen*, 150–176. Hildesheim [u.a.]: Olms.
- Demske, Ulrike. 2001. *Merkmale und Relationen. Diachrone Studien zur Nominalphrase des Deutschen*. Berlin & New York: de Gruyter.
- Deutsches Wörterbuch* von Jacob und Wilhelm Grimm. 16 Bde. in 32 Teilbänden. Leipzig 1854–1961. Quellenverzeichnis Leipzig 1971.
<http://woerterbuchnetz.de/dwb> (accessed 08 June 2017).
- Dücker, Lisa & Renata Szczepaniak. 2017. „Auffm Teuffelßdantz haben sie auff knotten korffen linen gedantzet“. Die graphematische Markierung von Komposition in den Hexenverhörprotokollen aus dem 16./17. Jh. In Florentine Oehme, Hans Ulrich Schmid & Franziska Spranger (eds.), *Wörter. Wortbildung, Lexikologie und Lexikographie, Etymologie* (Jahrbuch für Sprachgeschichte Band 8), 30–51. Berlin & Boston: de Gruyter.
- Flach, Susanne. 2017. *Collostructions: An R implementation for the family of collostructional methods*. R package version 0.1.0, www.bit.ly/sflach (accessed 14 December 2017).
- Fuhrhop, Nanna. 2008. Das graphematische Wort (im Deutschen). Eine erste Annäherung. *Zeitschrift für Sprachwissenschaft* 27. 189–228.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on “alternations”. *International Journal of Corpus Linguistics* 9(1). 97–129.
- Haubold, Fritz. 1914. *Untersuchung über das Verhältnis der Originaldrucke der Wittenberger Hauptdrucker Lutherscher Schriften: Grunenberg, Lother, Döring-Cranach und Lufft zu Luthers Druckmanuskripten*. Zugl. Diss., Univ. Jena, 1914. Borna-Leipzig: Noske.
- Klein, Thomas, Hans-Joachim Solms & Klaus-Peter Wegera. 2009. *Mittelhochdeutsche Grammatik. Band 3: Wortbildung*. Tübingen: Niemeyer.

- Kopf, Kristin. 2017. Fugenelement und Bindestrich in der Compositions-Fuge. In Nanna Fuhrhop, Renata Szczepaniak & Karsten Schmidt (eds.), *Sichtbare und hörbare Morphologie*, 177–204. Berlin & New York: de Gruyter.
- Kopf, Kristin. in press. From genitive suffix to linking element. A corpus study on the genesis and productivity of a new compounding pattern in (Early) New High German. In Horst Simon, Christian Zimmer & Tanja Ackermann (eds.), *Germanic Genitives*, 2–25. Amsterdam & Philadelphia: Benjamins.
- Kürschner, Sebastian & Renata Szczepaniak. 2013. Linking elements – origin, change, and functionalization. *Special issue of Morphology* 23(1). 1–6.
- Lexner, Matthias. 1872–1878/2011. *Mittelhochdeutsches Handwörterbuch*. 3 Bände. Leipzig: Hirzel. <http://woerterbuchnetz.de/Lexer/> (accessed 08 June 2017).
- Macha, Jürgen, Elvira Topalović, Iris Hille, Uta Nolting & Anja Wilke (eds.). 2005. *Deutsche Kanzleisprache in Hexenverhörprotokollen der Frühen Neuzeit*. 2 Bände. Berlin & New York: de Gruyter.
- Moser, Virgil. 1982. *Schriften zum Frühneuhochdeutschen*. Herausgegeben von Hugo Stopp, mit einem Register erstellt von Angelika Schmitt unter Mitw. von Maria Walch. 2 Bände. Heidelberg: Winter.
- Nitta, Haruo. 1987. Zur Erforschung der ‚uneigentlichen‘ Zusammensetzungen im Frühneuhochdeutschen. *Zeitschrift für deutsche Philologie* 106. 400–416.
- Nübling, Damaris & Renata Szczepaniak. 2011. Merkmal(s?)analyse, Seminar(s?)arbeit und Essen(s?)ausgabe: Zweifelsfälle der Verfungung als Indikatoren für Sprachwandel. *Zeitschrift für Sprachwissenschaft* 30(1). 45–73.
- Pavlov, Vladimir M. 1972. *Die substantivische Zusammensetzung im Deutschen als syntaktisches Problem*. München: Hueber.
- Pavlov, Vladimir M. 1983. *Zur Ausbildung der Norm der deutschen Literatursprache im Bereich der Wortbildung (1470–1730). Von der Wortgruppe zur substantivischen Zusammensetzung*. Berlin: Akademie-Verlag.
- Pavlov, Vladimir M. 2004. Zur Entwicklung der substantivischen Zusammensetzung im Frühneuhochdeutschen. In Klaus J. Mattheier & Haruo Nitta (eds.), *Sprachwandel und Gesellschaftswandel – Wurzeln des heutigen Deutsch: Studien des deutsch-japanischen Arbeitskreises für Frühneuhochdeutschforschung*, 99–119. München: Iudicium.
- Pavlov, Vladimir M. 2009. *Deutsche Wortbildung im Spannungsfeld zwischen Lexikon und Syntax. Synchronie und Diachronie*. Frankfurt a.M. [u. a.]: Peter Lang.
- Saenger, Paul. 1997. *Space between words: the origins of silent reading*. Stanford: Stanford University Press.
- Scherer, Carmen. 2012. Vom „Reisezentrum“ zum „Reise Zentrum“: Variation in der Schreibung von N+N-Komposita. In Livio Gaeta (ed.), *Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte*, 57–81. Berlin & Boston: de Gruyter.
- Schindler, Wolfgang. 1999. Bindestrich-Komposita im Frühneuhochdeutschen. In Wolfgang Schindler & Jürgen Untermann (eds.), *Grippe, Kamm und Eulenspiegel. Festschrift für Elmar Seebold zum 65. Geburtstag*, 313–329. Berlin & Boston: de Gruyter.

- Schutzzeichel, Marc & Renata Szczepaniak. 2015. Die Durchsetzung der satzinternen Großschreibung in Norddeutschland am Beispiel der Hexenverhörprotokolle. In Markus Hundt & Alexander Lasch (eds.), *Deutsch im Norden. Varietäten des norddeutschen Raums*, Bd. 1, 151–167. Berlin & Boston: de Gruyter.
- Schützeichel, Rudolf. 2012. *Althochdeutsches Wörterbuch*. 7., durchgesehene und verbesserte Auflage. Berlin & Boston: de Gruyter.
- Solling, Daniel. 2011. Zur Problematik der Unterscheidung zwischen pränominalen Genitivattribut und getrennt geschriebenem Substantivkompositum im Frühneuhochdeutschen. In Jörg Riecke (ed.), *Historische Semantik*, 294–311. Berlin & Boston: de Gruyter.
- Solling, Daniel. 2012. *Zur Getrennt-, Zusammen- und Bindestrichschreibung von Substantivkomposita im Deutschen (1550–1710)*. Zugl. Diss. Univ. Uppsala, Schweden, 2012. Uppsala: Univ.
- Solling, Daniel. 2016. Compound nouns in German (1550–1710): open, closed and hyphenated forms. In Paola Cotticelli-Kurras und Alfredo Rizza (eds.), *Variation within and among writing systems. Concepts and methods in the analysis of ancient written documents*, 291–302. Wiesbaden: Reichert.
- Solms, Hans-Joachim. 1999. Der Gebrauch uneigentlicher Substantivkomposita im Mittel- und Frühneuhochdeutschen als Indikator kultureller Veränderung. In Andreas Gardt, Ulrike Haß-Zumkehr & Thorsten Roelcke Thorsten (eds.), *Sprachgeschichte als Kulturgeschichte*, 225–246. Berlin & Boston: de Gruyter.
- Szczepaniak, Renata. 2016. Is the development of linking elements in German a case of exaptation? In Muriel Norde & Freek van de Velde (eds.), *Exaptation and Language Change*, 317–340. Amsterdam & Philadelphia: Benjamins.
- Szczepaniak, Renata & Fabian Barteld. 2016. Hexenverhörprotokolle als sprachhistorisches Korpus. In Sarah Kwekkeboom & Sandra Waldenberger (eds.), *Perspektivwechsel oder: Die Wiederentdeckung der Philologie. Bd. 1: Sprachdaten und Grundlagenforschung in der Historischen Linguistik*, 43–70. Berlin: Erich Schmidt Verlag.
- Topalović, Elvira. 2003. *Sprachwahl – Textsorte – Dialogstruktur: Zu Verhörprotokollen aus Hexenprozessen des 17. Jahrhunderts*. Trier: WTV.
- Weidman, Robert H. 1941. The orthographic conflation of nominal compounds in mhg based on a study of the Manesse manuscript. In Arno Schirokauer & Wolfgang Paulsen (eds.), *Corona: Studies in celebration of the eightieth birthday of Samuel Singer, professor emeritus, University of Berne, Switzerland*, 89–99. Durham, North Carolina: Duke Univ. Press.

Lisa Dücker
 Otto-Friedrich-Universität Bamberg
 Lehrstuhl für Deutsche Sprachwissenschaft
 Hornthalstraße 2
 96045 Bamberg
 lisa.duecker@uni-bamberg.de

Kristin Kopf

The role of syntax in the productivity of German N+N compounds. A diachronic corpus study¹

Abstract: This paper studies the morphological productivity of German N+N compounding patterns from a diachronic perspective. It argues that the productivity of compounds increases due to syntactic influence from genitive constructions (“improper compounds”) in Early New High German. Both quantitative and qualitative productivity measures are adapted from derivational morphology and tested on compound data from the Mainz Corpus of (Early) New High German (1500–1710).

1. Introduction

While N+N compounding in present-day German is subject to very few restrictions, the situation was different in Old High German (OHG, 500–1050 AD) and Middle High German (MHG, 1050–1350 AD). Obviously, the productivity of compounding has changed between MHG and today. This development warrants a more detailed investigation.

The paper is structured as follows. I will first review some fundamental aspects of compounding in modern German and its history, and describe my data sources and analysis (Section 2). I will then turn to productivity in word-formation which can be measured in different ways: First, it can be calculated from the numbers of types, tokens and hapax legomena. This is what I will refer to as quantitative measures. In its most basic sense, productivity is the ability of a word-formation pattern to form new units (operationalized as potential productivity; cf. Baayen 2009: 902–904) and to actually do so (operationalized as realized productivity; cf. Baayen 2009: 901–902). The question of to what extent productivity in word-formation can be measured has been hotly debated, albeit almost exclusively in the context of derivation. Compounding is considered to follow similar principles (e.g. “all the discussion here should equally apply to compounds”; Bauer 2005: 316), although I am unaware of case studies. I will apply quantitative productivity measures to corpus data from the second half of Early New High German (ENHG, 1350–1650) and from the beginning New High German (NHG, since 1650) in Section 3, and discuss their benefits and problems. Second, a word-formation pattern can lose semantic, morphological or phonological

1 I'd like to thank the two anonymous reviewers, Susanne Flach and Celeste Brennecka for their helpful comments.

restrictions, thereby opening up to formations that were unavailable before. While this is reflected in the quantitative measures, it warrants an in-depth discussion (Section 4). I will refer to this approach as qualitative, although the loss of restrictions will also be measured in numbers – strictly speaking, it is also largely quantitative, but applied to selected phenomena within compounding, not to compounding as such. I will show that the two perspectives complement each other and that it would not be wise to forego one for the other. In Section 5, I will combine the results and observations, offering an explanation for the considerable productivity of compounding in present-day German. I will show that it is likely rooted in changes in ENHG compounding patterns that were facilitated by syntactic constructions.

2. Fundamentals

2.1. Compounds in New High German

Modern German shows a strong affinity for compounding as a means to expand its lexicon. Munske (2009: 227) considers compounding “the greatest laboratory of ‘integration’ for lexical elements of different origin into new word structures” (own translation).² 83% of all new word-formation products in Harlass & Vater’s (1974: 91) corpus data are compounds – compared to single-digit percentages for derivation (9%), simplexes (3%) and abbreviations (2%).³ They conclude: “The extension of the lexicon thus relies mostly on compounding.”⁴

However, in present-day German compounding, input mostly consists of nouns and adjectives for both constituents, with N+N compounds being by far the most common type (Ortner et al. 1991: 37; Pümpel-Mader et al. 1992: 19). Many other types often subsumed under compounding are in fact cases of univerbation with a change in the part of speech: A combination of two prepositions used as an adverb, such as *durchlaus* ‘definitely (verb. throughlout)’, should therefore not be analyzed as compounding.⁵

2 „[D]as größte Labor der ‚Integration‘ lexikalischer Elemente verschiedener Herkunft in neuen Wortkonstruktionen.“

3 The material also contains phrasal constructions like *sowjetische Besatzungszone* ‘Soviet occupation zone’, i.e. mostly names in the form of an NP. These were not included here. In total, the corpus comprises 20,000 types from newspapers published in the early 1960s. The types denote concepts/objects that were either not known before 1945 or had not been recorded in dictionaries before. All types were recorded on punch cards and analyzed automatically.

4 „Eine Erweiterung des Wortbestandes findet also vor allem durch Komposition statt.“

5 A few notes on formal conventions: The internal structure of compounds is marked by <|> instead of <-> to avoid confusion with actual spelling, e.g.

Not only are compounds a productive means of word formation in German, the process also lacks phonological, morphological or semantic restrictions apart from preference for certain parts of speech. The widely cited *Donaul dampfschiff|fahrts|gesellschaft|skapitän* ‘captain of the Danube Steamboat Shipping Company’⁶ is no anecdotal exception: The length of a compound is merely restricted by the speaker’s working memory, much as syntactic units are. The longest compound (by letters) that appears at least 40 times in the Duden corpus consists of seven nouns, many of them containing derivational affixes.⁷ As long as there is a communicative need to further modify a concept, speakers form compounds (cf. Meineke 1991: 41). Even when presented with compounds that do not seem to make much sense (i.e. *Brücke|n|brücke* ,bridge bridge’, Günther 1981), they are willing to construct possible contexts (a general presupposition of existence, see Klos 2011). The main semantic relation between the two constituents of a German compound is determinative, i.e. the first constituent modifies the second, which forms the semantic core. Other semantic relations do exist but will only be discussed in passing.⁸

Anfahrts|beschreibung ‘directions (verb. approach|description)’, standard spelling: <Anfahrtsbeschreibung>. Linking elements are always separated and marked as LE. Historical corpus data often contains hyphenation in the original (using <=> or <->). In these cases, <|> is only used to additionally segment linking elements, while the original hyphen is kept, e.g. *Infektion|s|zeiten* ‘times of infection (verb. infection|times)’. Likewise, if a compound is not written as one orthographic word, the original space is preserved, e.g. *Vermanungs|red* ‘admonishing speech (verb. admonishment|speech)’. If internal compound structure is relevant, the embedded structure is enclosed in square brackets, e.g. [*Weih|brauch*]essig ‘vinegar of frankincense (verb. consecrate|smoke|vinegar)’. If a compound is not glossed, its English translation is accompanied by a verbatim equivalent in parentheses where the two differ, this is marked by ‘verb.’. Glosses follow the Leipzig Glossing Rules.

6 Verb. Danube|steam|ship|drive|company|captain.

7 *Grundstück|s|verkehrs|genehmigung|s|zuständigkeit|s|übertragung|s|verordnung* ‘Regulation on the delegation of authority concerning land conveyance permissions (verb. ground|piece|traffic|permission|jurisdiction|transfer|decree)’, see <http://www.duden.de/sprachwissen/sprachratgeber/die-laengsten-woerter-im-dudenkorpus> (8/9/2017).

8 Compounding terminology is notoriously heterogeneous (for an overview, see Scalise & Bisetto 2009). I follow Olsen (2001), distinguishing determinative, copulative and possessive compounds. Determinative compounds subsume both synthetic compounds like *Arbeit|nehmer* ‘employee (verb. work|taker)’ and root compounds like *Arbeit|s|zeit* ‘working time (verb. work|time)’. The choice is motivated by the different genesis of these types: Only determinative compounds could arise from syntactic constructions in ENHG (see below) while copulative and possessive compounds were always combined directly by a word-formation pattern (cf. Kopf 2018b: 5–8).

2.2. Compounding types: Etymology and form

N+N compounds can be divided into two groups from a synchronic and a diachronic perspective, respectively. Synchronically, we distinguish compounds with linking elements (LE, “linked compounds”) and compounds without linking elements (“unlinked compounds”). Diachronically, there are proper compounds (formed after the IE pattern) and improper compounds (reanalyzed from genitive phrases). These four groups are not independent, but there is no straightforward correlation between linked/improper compounds and unlinked/proper compounds, respectively. We will now consider all four types in turn.

IE nominal morphology consists of three elements: The root combines with a thematic vowel (stem-building element, SBE)⁹ to form a stem that can then be combined with a second noun or adjective (for compounds, c.f. 1a), a derivational suffix or an inflectional suffix (Krahe & Meid 1967). Only a few so-called root nouns act as stems without further additions (c.f. 1b).

(1) Gothic

- a. marlilsaiws ‘lake’
sealsBellake
- b. manlleika ‘figure/image of a man’
manlimage

This pattern has been called “proper compounding” (“eigentliche Komposition”) by Grimm (1826: 407): It is older than improper compounding and its compounds are formed by a morphological process. In OHG and MHG, most thematic vowels in compounds lose their systematicity, and are then lost completely (cf. Gröger 1911; Demske 2001: 297; Nübling & Szczepaniak 2013: 69–70). The only residuals can be found in a small group of short-stemmed nouns, e.g. MHD *taglellôn* ‘daily wage (verb. day|LE|wage)’.¹⁰ Thus, the new default case is a direct combination of roots without any intervening elements.

Improper compounds are a later addition to the system. They develop from univerbation of prenominal genitive attributes (cf. Pavlov 1983; Nitta 1987; Demske 1999, Nübling & Szczepaniak 2008; 2013; Solling 2012; Kopf 2017; 2018a; 2018b). Isolated cases are attested early on in many IE languages (Krahe & Meid 1967: 17–18):

9 Or a combination of vowel and consonant; these are, however, irrelevant for German compounds (cf. Nübling & Szczepaniak 2013, Szczepaniak 2016).

10 Vowels following former long-stemmed nouns must, on the other hand, have originated in improper compounds (e.g. *Schwein e futter* ‘pig LE food’, cf. Szczepaniak 2016: 329).

- (2) a. Greek *dióslkouroi*
 Zeus.GEN-son.PL
 b. Latin *aquaeductus*
 water.GEN-pipeline
 c. Sanskrit *jás|pati-*
 house/family.GEN-lord-
 d. Gothic *baurgslwaddjus*
 castle-GEN-wall

However, the pattern only becomes more common in ENHG. Nouns with overt genitive markers carry these with them when they undergo univerbation, hence giving rise to compounds with linking elements. This can be seen in (3) for linking-*s* (univerbation from a to b) and in (4) for linking-*n* (Kopf 2018a):

- (3) ENHG: genesis of linking-*s*
 a. *wegen [des Leibls] Erbeln >*
 because the.GEN.SG body(M)-GEN.SG heir(M)-GEN.SG
 b. *wegen des Leibls erbeln*
 because the.GEN.SG body-LE-heir(M)-GEN.SG
 ‘because of the heir of the body’
- (4) ENHG: genesis of linking-*n*
 a. *[der Sonneln] Schein >*
 the.GEN.SG sun(F)-GEN.SG shine(M).NOM.SG
 ‘the shine of the sun’
 b. *der Sonnelnschein*
 the.NOM.SG sun-LE-shine(M).NOM.SG
 ‘the sunshine’

The results resemble compounds, but did not arise by compounding. At some point, they are reanalyzed as part of the word-formation system. Linking elements are then productively inserted into new compounds that do not go back to genitive constructions. Such cases are still called “improper compounds” by Grimm, even though they are “properly” (i.e. directly) compounded.

Nouns unmarked for genitive case in the singular – i.e. feminine nouns from the strong *i*-declension (5a) – and strong masculine/neuter nouns without plural markers (5b) are formally identical with the old compounding type (for a detailed heuristic cf. Kopf 2018b, in preparation). It is thus impossible to tell how they came about:

(5) ENHG

			< <i>Nacht + Zeit</i>			
a. <i>Nacht:</i>	<i>Nachtzeit</i>	<	<i>in</i>	<i>der</i>	<i>Nacht</i>	<i>Zeit</i>
night (F)	night-time		in	the	night:GEN.SG	time
			< <i>Engel + Schar</i>			
b. <i>Engel:</i>	<i>Engellschar</i>	<	<i>mit</i>	<i>der</i>	<i>Engel</i>	<i>Schar</i>
angel (M)	angel-group		with	the	angel:GEN.PL	group

The distinction of proper and improper compounds thus poses two problems for empirical studies: First, in many cases without a linking element, it is impossible to tell if the compound is “proper” or not. Second, in many cases with a linking element, it is impossible to tell if the linking element is due to univerbation or part of a productive word-formation pattern.

In studies of present-day German, the division between proper and improper compounds is irrelevant: The direct connection with syntactic structures has been lost due to postposing of genitive attributes and the spread of determiner use (especially articles, cf. Pavlov 1983, Demske 2001, Kopf 2018b). Hence, no new univerbations occur. All compounds are now understood to belong to word-formation, even if their individual histories might not bear this out. Linked compounds amount for 27.2% to 42.4% of all compound types (Wellmann et al. 1974: 365, Kürschner 2003) and 29.5% of compound tokens (Donalies & Bubenhofer 2011: 32). The distribution of linking elements has been widely discussed (e.g. Fuhrhop 1996, Aronoff & Fuhrhop 2002, Kürschner 2003, Nübling & Szczepaniak 2008, Kopf 2018b), and although interested laypersons often concentrate on dubious cases, there are large areas in which the use of a linking element can be predicted with high certainty.

Almost all linked compounds go back to former improper compounds or were formed analogically, in line with their pattern. Only very few cases still contain an old linking vowel (like *Tagelohn*, see Tab. 1, upper row). Most unlinked compounds go back to former proper compounds and the pattern they provided. There is, however, a sizeable number of potential exceptions: In cases in which the genitive singular or plural did not have a suffix in MHG/ENHG (such as *(der) Nacht* ‘night’s (GEN SG)’ or *(der) Engel* ‘angels’ (GEN PL)’), we cannot tell whether compounds containing them are due to reanalysis/analogy or the old compounding pattern (see Tab. 1, lower row).

Tab. 1: Difficulties when distinguishing proper and improper compounds by linking elements.

	proper compounds	improper compounds
linked compounds	<i>Tagelohn</i> (< OHG <i>taglallōn</i> ‘daily wage (verb. day LE wage’)	from univerbation: <i>Tagelzeit</i> ‘time of day (verb. day LE time)’ (< OHG <i>zuo tagles zīt</i> ‘at day GEN time’) from compounding: <i>Religion s=Anfechtungen</i> ‘contestations of religion (verb. religion LE contestation PL)’
unlinked compounds	<i>Nacht ø zeit</i> ‘nighttime’ (< OHG <i>zuo naht zīt</i> ‘at night. GEN time’ OR OHG <i>naht ø zīt</i> ‘night-time’) ¹¹	

2.3. Corpus & data analysis

Most of the empirical data for this paper was collected using the following corpora:

time	corpus	composition	tokens	annotation	data extracted	description
1500–1710	Mainz Corpus of (E)NHG	8 periods, 2 genres, 5 regions	320,000	none	all Genitive constructions and NN compounds with common nouns	Kopf (2018b)
1650–1800	GerManC (subcorpora SCIE, SERM, NEWS)	3 periods, 3 genres, 5 regions	90,000	POS	all NN compounds (manual check of all tokens POS-tagged NN with 6+ characters)	Durrell et al. (2007)
1850, 1900	excerpts from Mannheim Historical Newspapers	2 periods, 1 genre	60,000	none	all NN compounds (manual check of all tokens with initial upper case and 6+ characters)	http://hdl.handle.net/10932/00-01B8-AE41-41A4-DC01-5

11 The empty set symbol (∅) is used where no linking element occurs to mark the contrast. It does, however, not assume some kind of invisible element.

The Mainz Corpus of (E)NHG, based on Bergmann & Nerius (1998) but heavily modified for present purposes (cf. Kopf 2018b), has been analyzed most thoroughly. It contains 8 periods, 2 genres (non-fiction and theological texts) and 5 regions (East Upper German, West Upper German, North Upper German, East Middle German, West Middle German). The GerManC sub-corpora SCIE and SERM (scientific texts and sermons) were chosen because they are similar in text composition. From 1650 to 1900, newspaper texts of the same size can also be compared (3 periods of NEWS in GerManC and 2 added periods from the Mannheim Corpus of Historical Newspapers). Each corpus was balanced for token number per period and, where applicable, genre and region.

Quantitative productivity measures were used for the Mainz Corpus, qualitative data was analyzed for all three corpora. In one case, additional material from the German Text Archive (DTA) was used; in another case, the Corpus of the MHG Grammar (MiGraKo) was consulted.

Identification of compounds in historical sources poses serious problems, especially in the Mainz Corpus, as they partially overlap with syntactic constructions (cf. Pavlov 1983, Nitta 1987, Reagan 1981, Solling 2011, 2012). Due to morphosyntactic changes in ENHG (e.g. positional change of the genitive attribute, obligatorization of definite and indefinite article), this is no longer the case. Today, morphosyntactic, phonological and graphematic properties clearly distinguish compounds from phrases. In Kopf (2018b, in preparation), I develop a detailed heuristic based on Pavlov (1983) for distinguishing unambiguous compounds, unambiguous genitive constructions and three types of ambiguous intermediate constructions. One of these ambiguous constructions, the so-called “framing constructions” (Rahmenkonstruktionen), is often taken for compounding (e.g. Kopf in print) as it has a determiner or modifier that precedes the whole construction and agrees with the second noun (see 4a). I have included this construction when measuring quantitative productivity. As shown in Kopf (2018b: 243–252), there is, however, little difference with a scenario in which they are excluded.

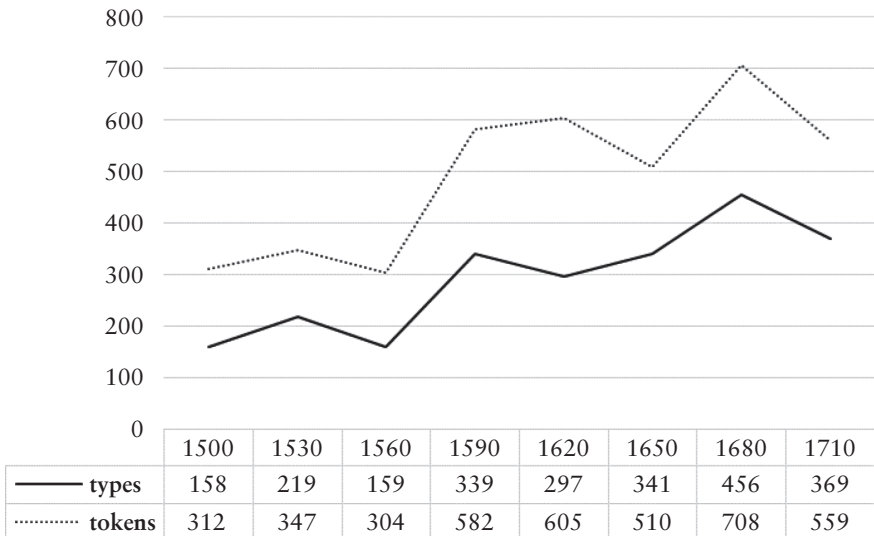
3. Quantitative measures of productivity

3.1. Compounding in general

It is indisputable that compounding has become more and more frequent when comparing OHG and NHG inventory and usage. While numbers for types and tokens are lacking in earlier periods, they have been determined for the Mainz Corpus, so part of the gap between OHG and NHG can now be filled with data. This allows us to apply established measures of productivity and compare the outcome with findings from a more quantitative approach.

The numbers for tokens and types, given in Fig. 1, show a strong increase from 1590 onwards: Compounds were used to a greater extent than before, and people used more different compounds than previously – e.g. compounds that were not possible in OHG or MHG due to morphological restrictions (see Section 4). The type frequency corresponds to the realized productivity of the pattern,¹² i.e. it shows how productive the pattern has been in the past (Baayen 2009: 901–902). As it is rather low for the years 1500 to 1560, its productivity must have expanded immensely shortly before 1590.

Fig. 1: *Types* ($n=2,338$) and *tokens* ($n=3,927$) of compounds in the Mainz Corpus. The increase is statistically highly significant for both types and tokens (Kendall's Tau, $\tau=1$; $T=28$; $p<0.001$).



3.2. Compounding patterns: linked vs. unlinked compounds

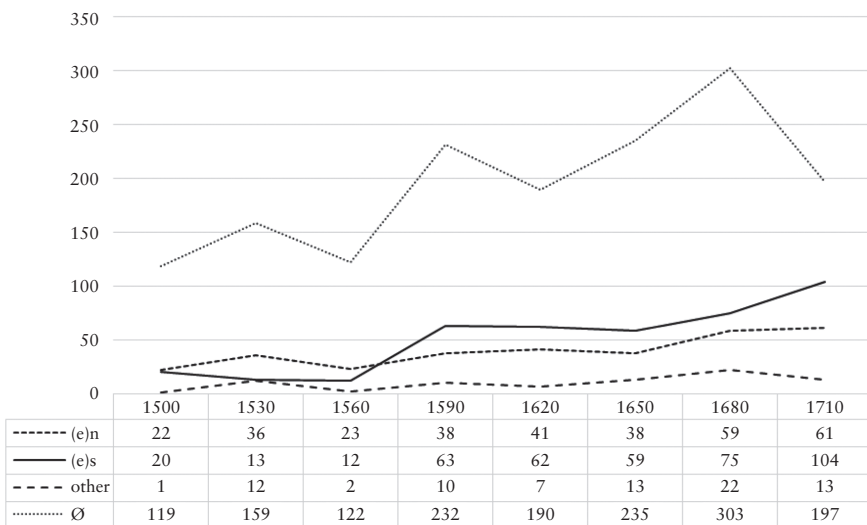
As has been discussed in Section 2.2, most compounds with linking elements (apart from a few exceptions with linking-*e*) point either to reanalysis from genitive phrases or to a new, purely morphological pattern of compounding. Compounds without linking elements, on the other hand, do not point to a particular genesis (but see Kopf 2018b: 228–230). Changes in their productivity cannot be clearly attributed to either univerbation or compounding. I will therefore focus on linked compounds for the remainder of this section.

12 In a corpus of invariant size, realized productivity is simply the number of types (while usually it would be the number of types divided by the number of text words).

3.2.1. Realized productivity

The realized productivity of compounds in general has been shown to increase over time (3.1). If we look at different compounding patterns in turn, large differences arise: The unlinked pattern increases, but shows a great variation, probably due to the heterogeneous genesis of its members. The linked patterns increase in a more straightforward way ((*e*)*n*, (*e*)*s*, other in Fig. 2).

Fig. 2: Realized productivity by compounding pattern. The increase is statistically significant for the three major patterns (Kendall's Tau, (*e*)*s*: $\tau=0.571$; $T=22$; $p<0.05$), (*e*)*n*: $\tau=0.815$; $z=2.764$;¹³ $p<0.01$, no LE: $\tau=0.643$; $T=23$; $p<0.05$) (Kopf 2018b: 245).



As for the overall development, compounds with linking-(*e*)*s* show a drastic increase in 1590, and again in 1710. We cannot tell for sure from the numbers if this is due to reanalysis or due to the establishment of a new, purely morphological pattern. I would, however, argue for the latter, as most constructions that form potential input for reanalysis have disappeared by this time (due to advanced positional change of genitive attributes, see Kopf 2018: 86–110).

3.2.2. Potential productivity

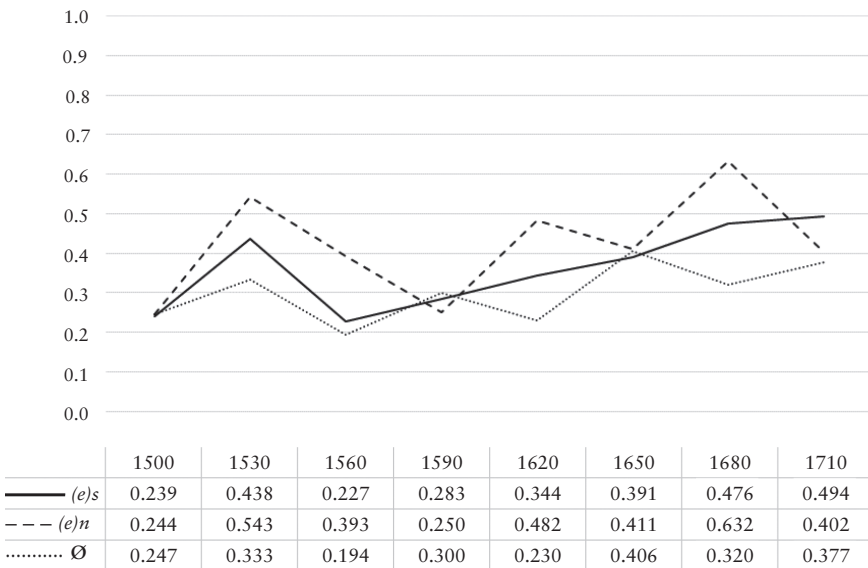
Potential productivity is a measure that is supposed to determine not how productive a pattern has been in the past, but what potential it has to form new words in the present (Baayen 2009: 901–902). This makes a lot of

13 T could not be calculated due to multiple identical values.

sense when looking at derivational morphology: If a suffix appears on many words, but morphological or semantic restrictions hinder its further spread, the pattern has done all it can already. The existing types will be used, but as they are firmly established, they will surely occur several times in a large collection of texts, not just once. If, however, the pattern can still be expanded (e.g. by shedding restrictions), speakers will coin words that are not (yet) part of the lexicon, and those will appear seldom at first. Potential productivity is therefore estimated by dividing hapax legomena by tokens of the pattern under investigation. The reasoning goes as follows: If a word appears only once in a good-sized corpus, it is very likely a new formation. We can then determine the ratio of such new formations to all formations. The closer the resulting number is to 1, the higher the potential productivity.

By drawing on hapax legomena, potential productivity hinges on corpus size: The smaller the corpus, the higher the number of compounds that appear only once, and the higher the probability that these are not actually new formations. The relationship between hapax legomena, corpus size and number of tokens relevant to the pattern is not linear (see Kopf 2018b: 246–250). While the Mainz Corpus accounts for invariable corpus size, comparison between patterns of different frequency is notoriously difficult. It has been shown that the overall number of compounds increases over time (3.1), thus lowering the probability of hapax legomena in later periods. When separating the groups of linking elements as done here, the problem increases: As unlinked compounds show much higher numbers than linked compounds (and thus lower probabilities for hapax legomena), their development cannot be compared to each other. This renders the concept of potential productivity highly problematic.

Fig. 3: Potential productivity of different compounding patterns in the Mainz Corpus.



The findings are therefore limited to determining if there is an increase of productivity in one of the patterns. This is not the case; Kendall's Tau yields no significant results for any of them. This seems peculiar when compared to realized productivity: How can the potential of a pattern be unchanged if the number of different types it produces does in fact increase? The most likely explanation is that the corpus is not suited for measuring potential productivity: The number of hapax legomena is artificially high for two reasons: First, the corpus size is rather small (40,000 tokens per period). Second, the compounds are not evenly distributed over time. The number of compound tokens is low in the first three periods (between 300 and 350) and then almost doubles in 1590 (see Fig. 1). This means that the number of hapax legomena will be higher than expected from 1500 to 1560 compared to the following years. That might lead to the impression that potential productivity does not change over time. We cannot tell for sure whether realized or potential productivity is a better indicator of what is happening in the corpus, which is why it is instructive to have a closer look at qualitative measures of productivity.

4. Qualitative measures of productivity: Loss of restrictions, extension to new contexts

4.1. General remarks

The use of new word-formation patterns can be limited by pragmatic or structural restrictions (e.g. Rainer 2005, Schmid 2011: 115–119). The main prerequisite for the use of a pattern is of course the need to designate a concept or object. This need may not arise if the referent is non-existent (and not imaginable, e.g. *carpet opener*), if the pattern would make use of obvious properties (**eyed man*) or if there is no “nameable” concept (Schmid 2011: 116). Preexisting synonyms¹⁴ or homonyms (**liver* ‘living person’) might also be hindering an otherwise regular word-formation process. All these “restrictions” pertain to individual words; they do not hinder a word-formation pattern as such, and thus they tell us more about human imagination than about word-formation (see also Schmid 2011: 116, who calls them “not very helpful”).¹⁵

The case is different for semantic, morphological or phonological restrictions: A pattern might only be applicable to certain parts of speech (i.e. present-day German *-er* denoting agentive nouns can only be used with a verbal base) or to certain parts of the lexicon (i.e. native vs. non-native bases,¹⁶ complex vs. simple bases). It might be hindered when certain sound patterns would arise

14 Much-used examples like *Stehler* (*steht-* ‘steal’ + *-er* ‘agentive suffix’) which is supposedly blocked by *Dieb* ‘thief’ lose much of their strength by a diachronic perspective: When *er*-formations arose in OHG, they met with quite a number of synonyms, e.g. *wartil* (*-il*) ‘warden’ or *becko* (*-o*) ‘baker’. Still, they were finally replaced by *Wärter* and *Bäcker*. Cases like *Dieb* are marginal at best and hold no influence over the word-formation pattern as such; they only hinder individual words. Kempf (2016) shows partially synonymous adjectival suffixes (e.g. *-lich* and *-bar*) competing for bases which lead to functional differentiation (via temporarily coexisting forms).

15 It is also not true that speakers would generally feel no need to coin such terms: They might want to promote a new tool used to pry a carpet away from the floor to remove it (*carpet opener*), or they might want to distinguish a person employed with unfolding a rug from a roll from other workers. Children might speak about the only doll left that still has its eyes as *eyed man* and a science-fiction author may combine a seemingly disparate array of actions and objects into a single verb to denote something aliens do.

16 This restriction is considered etymological by Plag (1999: 58). I would argue that – at least for German – the synchronic structure of the base is the deciding factor. Of course suffixes like *-ier*, *-ize* (restricted to latinate bases) can also be found with roots that do not look like loans, but in fact are (e.g. *kassieren* ‘to cash in’ < *Kasse* ‘cash register’, which looks like native *Gasse* ‘alley’). However, most of these cases are remnants from earlier times – *kassieren* hails from the

(i.e. German diminutive suffix *-chen* which doesn't combine with bases ending in a palatal fricative – in these cases, synonymous *-lein* is preferred, so *Büchlein* 'small book' instead of *Büchchen*) or favored when a certain prosody arises.

4.2. What to measure in compounding

While most derivatives and compounds show a binary structure, the functions of the two morphemes are not identical. In derivational morphology, the analysis of restrictions is usually framed in one of two perspectives: Either the base accepts only certain affixes, or the affix selects only certain bases. This is clearly not the case in compounding. Apart from the fact that there should be a semantically motivated reason to combine the two elements, the specific constituents do not restrict each other. However, there are restrictions on the whole compounding pattern – or at least there were in earlier periods of German. In OHG, compounds with a verbal stem as first constituent are not yet in use (for their development, see Gröger 1911), tripartite compounds are uncommon and mostly found in translations and glosses (Carr 1939:197), and derivationally complex first constituents are almost nonexistent. The default case for the first constituent is a monosyllabic stem (Wilmanns 1896: 388).¹⁷ The case is different for second constituents: They show no restrictions in older periods (Wilmanns 1896: 388). Even though it cannot be said that the second constituent formally restricts the first, OHG speakers were clearly hesitant to make use of morphologically complex elements while NHG speakers are not. In this section, I will focus on the change the compounding pattern underwent in ENHG. First, the use of morphologically complex input over time will be analyzed with data from grammars and the Mainz Corpus. Then, a second aspect must be considered: Most OHG and MHG compounds adhere to the old compounding pattern, although there are a few cases of linking elements from genitive reanalysis (Grimm 1826: 409). From ENHG onwards, the new compounding pattern plays an important role. I will show that there is an overarching explanation for both observations.

4.2.1. Number of roots: tripartite compounds

Tripartite compounds with three nominal elements are not attested in Gothic or other minor Germanic sources (Carr 1939: 197). Even if “particle compounds” like Gothic [*fralbauhta*]*boka* ‘deeds of sale’ or *ufar*[*biminkunda*] ‘heavenly’

17th century when its base still was an obvious loan from Italian (*Cassa*). It seems more adequate to classify the input restriction for *-ier* as structural.

17 There are different opinions on how common compounding was in Germanic (Gröger 1911: 1, Salus 1963: 47, Carr 1939), but we can assume tentatively that it was similar to early OHG.

are included, cases are rare and the embedded compounds are “no longer felt as compounds” (Wilmanns 1896: 388).¹⁸

In early OHG, tripartite compounds are still rare (Carr 1939: 197), although more numerous than in Gothic (Wilmanns 1896: 388). They are mostly found as glosses and often contain a lexicalized constituent. In many cases, the second constituent is undergoing grammaticalization (6d, ending in today’s derivational suffix *-schaft* ‘-ship’).

- (6) a. [*poob|stap*]*zila* ‘row of letters’
book-staff-row
(Samanunga glossary, Vienna, Cod. 162)
- b. [*uuerolt|reht*]*uuisoln* ‘scholars of the law of the world’
world-right-scholar-PL
(Muspilli, Clm 14098)
- c. [*hasel|nuze*]*chermeln* ‘hazelnut kernel’
hazel-nut-pit-DAT.SG
(Notker: Martianus Capella, St Gall, Cod. 872)
- d. [*heril|ganoz*]*scaf* ‘military division’
army-comrade-form
(Monsee fragments, Munich, Clm 14747)

Starting from the later OHG period, tripartite compounds are “not uncommon” (Carr 1939: 197), although most seem to be a result of translation.¹⁹ Language users rely on them to solve a communicative problem. Their usage is telling: As compounding is almost nonexistent in Latin, tripartite compounds are a Germanic structure through and through. Even though sparingly used in indigenous texts, the pattern is available. Such structures are most likely not (yet) a part of the general compounding pattern, but this pattern’s extension comes naturally when expressing complex concepts.

We can assume that the use of the pattern rose gradually, though not steeply from OHG to MHG and then ENHG. Reliable data is available from 1500 onwards: Compounds with three roots occur from the beginning in the Mainz corpus, albeit at a low level (Kopf 2018b: 268–270);²⁰ They account for 2.6% of all compound types and are used even less (1.8% of all compound

18 I count such cases towards prefixation (see below).

19 “The whole type is, in fact, mainly restricted to glosses, and is not found in the OHG poetry, although there is one example from the L[ow]G[erman] Heliand.” (Carr 1939: 198)

20 The data includes compounds with three roots. The embedded compound may have one of the structures V+N, A+N or N+N.

tokens). Almost half of the types contain a lexicalized constituent that is already attested in OHG (e.g. [*Weih|rauch*]essig ‘vinegar of frankincense (verb. consecratelsmoke|vinegar)’). The GerManC data (1650–1800) shows no difference in types (2.7%) but a higher usage frequency (2.9%), which is mostly due to the inclusion of newspaper texts (they contain 93 of 155 tripartite tokens).

In present-day German, compounds with three or more lexical roots are more common, but not widely used: While accounting for 11.8% of Ortner et al.’s (1991: 13) nominal compound types (X+N), their usage varies between 1% and 8% in newspaper texts. If we consider only new formations instead of inventory, tripartite compounds play a bigger role: In an early corpus study on German neologisms, Harlass & Vater (1974: 94) show 36% of all compounds to be made up of more than two lexical roots. This points to a stronger productivity of the pattern, probably genre-related.

4.2.2. *Derivationally complex constituents*

This group contains cases of suffixation (e.g. *Beziehung|s|ratgeber* ‘relationship advisor’), prefixation (e.g. *Umweltschutz* ‘environment|protection’), but see below) and circumfixation (*Geschmeidelkunst* ‘jewelry (verb. jewelry|art)’). Such derivationally complex first constituents occur regularly in present-day German. They are, however, uncommon in OHG, which limits itself to complex second constituents (Wilmanns 1896: 388).

Only when a **suffix** has become intransparent in OHG (e.g. in *kuning* ‘king’, *jugund* ‘youth’), its word-formation product may be used at the beginning of a compound. Transparent formations in *-āri*, *-ida*, *-unga*, *-nissi* cannot be found in OHG compounds (Tatian, Otfrid, Notker, cf. Wilmanns 1896: 388).²¹ These restrictions still hold for MHG, first constituents in *-er(e)* (< OHG *-āri*) being the only exception (Carr 1939: 222):

- (7) a. *dörper|diēt* ‘farmers (verb. farmer|people)’
- b. *jeger|meister* ‘hunter (verb. hunter|master)’
- c. *rihter|stuel* ‘bench (verb. judge|chair)’

The group denotes agent nouns, as opposed to most other OHG/MHG suffixes, which bring about abstract or collective semantics. This makes *er*-derivations especially common in possessive constructions and therefore in many compounds. These nouns are structurally identical with simplex

21 Carr (1939: 219) names two early exceptions, *hamalunc-stat* ‘Skull Hill’ and *gesmīde-ziereda* ‘jewel ornaments’ (from the Germanic circumfix *ga-X-ja*), but concedes that the former is likely due to OE influence Carr (1939: 221).

nouns in *-er* (e.g. MHG *kevere* ‘bug’, *hamer* ‘hammer’), which might have helped as well.

The situation is very different in the (E)NHG data of the Mainz Corpus. I considered all first constituents ending in suffixes or suffix-like endings of Latin origin (“exogenous suffixes”, cf. Fleischer & Barz 2012: 239) and mapped their first appearance in a compound (Fig. 4). While first appearance in a corpus should not be mistaken as first use or even first attested use, the relative chronology is still instructive.

Fig. 4: First appearance of a suffixed first constituent in the Mainz Corpus (Kopf 2018b: 261). Suffixes that appeared less than ten times on either a first constituent of a compound or a genitive attribute were not considered. Real suffixes are set in boldface to set them apart from suffix-like endings.

1500	1530	1560	1590	1620	1650	1680	1710
<i>-er</i> (<i>Burgermeister</i> ‘mayor’)							
<i>-ent/ant</i> (<i>conuent bruder</i> ‘friar’, <i>Proviand heuser</i> ‘storehouse for food’)							
				<i>-ur</i> (<i>Fractur=Feder</i> ‘quill for black letter’)			
				<i>-ung</i> (<i>Vermanungs red</i> ‘admonishing speech’)			
				<i>-ion</i> (<i>Infection s=zeiten</i> ‘times of infection’)			
				<i>-nis</i> (<i>Vermächtnu ßsachen</i> ‘inheritance’)			
				<i>-tät</i> (<i>Majestät Briefen</i> ‘Letters of Majesty’)			
				<i>-ei</i> (<i>Cantzley buch</i> ‘registry’)			
				<i>-lein</i> (<i>Mährlein schreiber</i> ‘writer of (fairy) tales’)			
					<i>-schaft</i> (<i>Vormundschaft s Zweck</i> ‘purpose of a guardianship’)		
							<i>-at</i> (<i>Canonicat=Stift</i> ‘monastery chapter’)
							<i>-enz</i> (<i>Residentz=Stadt</i> ‘royal seat’)
1500	1530	1560	1590	1620	1650	1680	1710

The data suggest that the restriction on derivationally complex first constituents started to crumble in the late 16th century: Derivatives that look much like simplexes occur from the beginning (with *-er* even going back to MHG). However, suffixes that are formally set apart first appear in 1590 and grow in number. The first suffixes to be included are phonologically reanalyzed due to their vocalic onset, and thus merged with the base. Suffixes with consonantal onset that form their own phonological word appear from 1620 on. Some suffixes, while present outside of compounds in the Mainz Corpus (e.g. *-tum*, *-heit/keit*), do not appear yet. A closer look at the most frequent of these suffixes, *-ung*, a

cognate of engl. *-ing*, adds to the finding that complex first constituents only became part of the German compounding pattern in ENHG. Grimm (1826: 937) remarks on Martin Luther’s German bible translation (first half of the 16th century):

As in earlier use, [Luther] never compounds the frequent formations in *-ung*. He says neither *nahrung-sorge*, *versönung-geld* nor *nahrungs-sorge*, *versönungs-geld*, but *sorgen der nahrung* [‘worries about food/sustenance’] Luke 21, 34, *geld der versönung* [‘atonement money’] Exod. 30, 16. [...] or if he wants to compound, he uses a verb, e.g. *versön-tag* [‘day of atonement’] Levit. 23, 28. (own translation and formatting)²²

This ties in nicely with the corpus data reported above, and more specifically with the usage of *-ung* in first constituents of compounds: The first use in the Mainz corpus dates from 1590. An increase can only be seen in 1680/1710 (11 tokens, 10 different *ung*-types, cf. Kopf 2018b: 258–259). One could suspect that growing numbers of *ung*-compounds simply reflect growing productivity of *ung*-derivation. However, this is not the case: The productivity of *-ung* remains diachronically stable in the Mainz Corpus (Hartmann 2016: 171). Growing use in compounds therefore shows loss of morphological restrictions. We will come back to the fact that Grimm’s (1826: 937) potential compound suggestions show both unlinked and linked compounding (*nahrung-sorge* vs. *nahrung-s-sorge*) – this is no coincidence.

The topic of nominal **prefixes** is somewhat complicated: Although they could be mistaken for prefixes on first sight, preposition-like elements in many complex words are actually due to conversion (8a) or derivation of verbal stems (8b):

- (8) a. *stēn* ‘stand’ > *fir+stantan* ‘realize, understand’
 conversion: *firstant* ‘understanding, intellect’
 b. *waltan* ‘govern’ > *anawalt(an)* ‘to be powerful’
 derivation: *anawalt+o* ‘administrator, ruler’
- (9) a. *lust* ‘desire, pleasure’ > *un+lust* ‘dislike, displeasure’
 b. *richten* ‘to judge’ > *ge+riht+i* ‘court of law’

22 German original: „die häufigen bildungen mit -ung setzt er [Luther, KK], gleich der früheren sprache, [...] nie zusammen. Er sagt weder *nahrung-sorge*, *versönung-geld* noch *nahrungs-sorge*, *versönungs-geld*, sondern *sorgen der nahrung* Luc. 21, 34, *geld der versönung* Exod. 30, 16. [...] oder wenn er componieren will, thut ers verbal, z. B. *versön-tag* Levit. 23, 28.“

Nouns that have carried over verbal prefixes or particles are rare as first constituents in OHG (Wilmanns 1896: 139). Nouns with true nominal prefixes do not occur (9). This also holds true for derivatives with a *gi-X-i* circumfix (9b),²³ the single exception being *gesmīdelziereda* ‘jewel ornaments’ (Carr 1939: 219). A spot-check in MiGraKo yields no clear examples of MHG compounds with former verbal prefixes or particles.²⁴ A few examples with circumfixed forms (e.g. [*gelrihte*]hūs ‘courthouse’, [*gelburt*]muoter ‘birth mother’) are attested in MHG dictionaries. In the (E)NHG corpus data, the pattern is present, independent of its origin: 33 tokens (9 different first constituents) are based on former prefixed verb stems (e.g. 10a)²⁵, and 61 tokens (11 different first constituents) bear a nominal prefix (e.g. 10b) – albeit almost all of them examples of the former circumfix.

- (10) a. *beyschlaff weiber* ‘common-law wives’ (1557), *Ablaßkrämer* ‘seller of indulgences’ (1588), *Geburtszeit* ‘time of birth’ (1617), *Anfangslbüchstaben* ‘initial letters’ (1648), *Andachts=Blick* ‘devoted gaze’ (1684), *Besitz=Nehmung* ‘appropriation, seizure’ (1708)
- b. *gericht|ß|costung* ‘court fees’ (1507), *Gesang-Buch* ‘hymn-book’ (1706), *Gebüt=Folge* ‘succession line’ (1708), *Gebetts=Puncten* ‘parts of a prayer’ (1708), *Unglück|slstiffterin* ‘(female) causer of misfortune’ (1708)

The pattern is ubiquitous in present-day German. It was probably furthered by the growing usage frequency of stem conversions in the 17th and 18th century and some new back-formations, both due to prescriptivist abhorrence of *ung*-derivatives (Hartmann 2016: 172).

4.2.3. Growing complexity: summary

Overall, the analysis of restrictions on structural complexity of compound constituents yields the following picture:

23 The closing part of the circumfix is later dropped in many cases, so the form resembles a prefix *Ge-* (OHG *girihti* > NHG *Gericht* ‘court of law’).

24 However, cases which might be syntactic phrases do occur: *antlaz(e) tag* ‘Maundy Thursday’, *vor dem auffert. tag* ‘before (the) Ascension Day’.

25 For a potential connection with simplex stem conversions, see Kopf (2018b: Kap. 11.7.3).

Tab. 2: Restrictions on morphological complexity of compound constituents in OHG, MHG, and NHG.

	OHG	MHG	NHG
N1 = compound	✓ (mostly in glosses)	✓	✓
N2 = compound	✓ (but rare)	✓	✓
N1 = suffixed	✗ no transparent/ productive suffixes possible	✗ no transparent/ productive suffixes possible (exception: <i>-er(e)</i>)	✓
N1 = prefixed	✗	✗ (only very few cases of <i>ge-</i>)	✓
N2 = prefixed/suffixed	✓	✓	✓

Tripartite compounds have gradually expanded, but they were always possible. The increase in both types and tokens is most likely due to the growing need for a more specialized vocabulary as the use of German becomes more common in all kinds of written texts. The case is different for compounds that contain a derivative: While derivatives have always been used as second constituents, their use as first constituents only came about in ENHG and has expanded since. The input restriction on derivatives as first constituents in OHG and MHG was rather abruptly dropped in ENHG. Specialization in vocabulary might also play a role, but due to the speed of the change, it is unlikely to be the only reason. I have also shown that the change is mostly independent from increasing productivity of the respective derivational pattern. I will now offer a different explanation for the change. It hinges on the observation that the loss of this restriction coincides with the genesis of a new compounding pattern derived from syntax. I assume that in this case, the correlation is due to a causal relation between the two.

4.2.4. Compounding patterns

Productivity of the new compounding type must be further differentiated. The pure increase in linked compounds cannot tell us whether these are cases of univerbation stemming from genitive constructions or cases of direct word-formation. While quantitative productivity can be computed for the former, this is not truly productivity in a morphological sense: The numbers do not prove the existence of a compounding pattern. They simply show lexical growth in a special part of the vocabulary. There are, however, a few systematic cases that can be taken as proof of an independent word-formation pattern; they could theoretically be used to determine the terminus ante quem for morphological productivity. All are either mismatches between semantics

and form of the compounding stem form (first constituent and LE, cf. Fuhrhop 1998), or formal mismatches between inflection and word-formation. While they are often cited in literature (e.g. Demske 2001, Meibauer et al. 2015), their usefulness in actual corpus texts has not been tested before.

There are two kinds of mismatches between semantics and form: The compounding stem form in (11) matches a genitive in the original paradigm (cf. Kopf 2018b: 140–141).

(11) *Bischofslskonferenz* ‘episcopal conference (verb. bishop-LE-conference)’

However, its number is semantically wrong: If *Bischofslskonferenz* were a case of univertation, the preceding genitive construction would have been *Bischofsls Konferenz* ‘bishop’s conference’. This is impossible, as the referent is a meeting of several bishops. Speakers using a genitive construction to convey this sense would have said (*der*) *Bischöfe Konferenz* ‘(the.GEN) bishops’ conference’, resulting in a hypothetical univertation *Bischöflekonferenz*, not *Bischofslskonferenz*. The mismatch between form and semantics shows that the linking-*s* in *Bischofslskonferenz* must be due to a word-formation process. The phenomenon is not very frequent, and no such cases are attested in the Mainz Corpus. A non-exhaustive search in DTA,²⁶ however, yields the following examples for my time period:

- (12) a. 1614: *häringslsfang* ‘catch of herrings (verb. herring|LE|catch)’
 b. 1626: *Wolffslsjagt* ‘wolf hunt (verb. wolf|LE|hunt)’
 c. 1627: *Diebslshauffen* ‘pile of thieves (verb. thief|LE|pile)’
 d. 1640: *Strömlingsl-fang* ‘catch of Baltic herrings (verb. Baltic herring|LE|catch)’
 e. 1659: *Mannleslshauffe* ‘pile of men (verb. man|LE|pile)’
 f. 1660: *Feindes-Hauffen* ‘pile of enemies (verb. enemy|LE|pile)’
 g. 1682: *Hunds-Zucht* ‘dog breeding (verb. dog|LE|breeding)’

All these cases show linking elements that do not fall in line with the semantics of the first constituent: A *Diebslshauffen* is a pile of thieves, but contains linking-*s*, which is derived from a genitive singular. It is therefore impossible that the compound is a case of univertation from a genitive construction; a plural (e.g. *der Dieble Hauffen*) would have been used. The linking-*s* must be part of a word-formation process. We can thus conclude that a productive

²⁶ These data were collected by Andreas Klein (Mainz), mostly by searching for collective second constituents like *Haufen* ‘pile’ preceded by a linking element. Only the first attestation for each first constituent are taken into account.

compounding pattern with linking elements is already in place by the early 17th century.

The second mismatch between semantics and form is based on the case relation between the two constituents. It is usually argued that copulative compounds cannot be based on a genitive construction because a genitive relation between the two constituents is not conceivable: A genitive relation is one of subordination while a copulative compound shows coordination (e.g. Demske 2001: 311). If copulative compounds like NHG *Hosenrock* ‘pant skirt, pair of culottes (verb. trouser|LE|skirt)’ contain linking elements, this must therefore be due to morphological productivity. However, almost all historical examples are problematic. Demske’s (2001: 311) cases like *Baur|s|man* ‘farmer (verb. farmer|LE|man)’ should not be used because the relation is one between hyponym and hypernym. As is shown in Kopf (2018b: 178–180), parallel genitive constructions do actually exist at the time. In the only linked example found in the Mainz Corpus, *Schalk|s|narren* ‘fools (verb. joker|LE|fool|PL)’, the first constituent might as well be based on the abstract meaning of *Schalk*, ‘wag-gishness’.²⁷

The last and most useful case to be considered here is that of the so-called “non-paradigmatic linking-s” as in (13):²⁸

- (13) *Vormundschaft|s|Zweck* ‘purpose of a guardianship (verb. guardianship|LE|purpose)’

Compounds with feminine first constituents using linking-s are non-paradigmatic because the s-genitive is (and was) exclusively used for neutral und masculine nouns. These cases can therefore not be based on univerbation; the linking element must have been included when a regular compounding pattern was applied. Non-paradigmatic linking-s in NHG is mostly found with derivatives in *-heit/keit*, *-ion*, *-(i)tät*, *-sal*, *-schaft* and *-ung* and there are

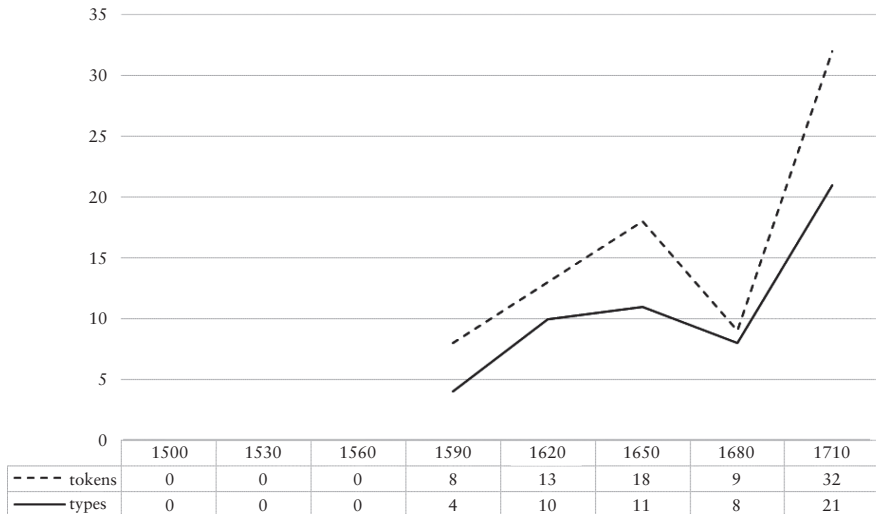
27 For the sake of completion, I will shortly point to another possible indicator of morphological productivity that is irrelevant for my corpus data. These are first constituents which only appear after all possible reanalysis contexts disappeared due to the positional change of the genitive attribute, i.e. in later NHG, usually loans. One of the (very few) examples is *Trainings|s|anzug* ‘tracksuit (verb. training|LE|suit)’. As univerbation may still be ongoing during the period of my investigation, linked compounds with recent loans can by definition not be used in my Corpus data to determine productivity. They are, however, useful for addressing present-day productivity questions (see Kopf 2018b: Kap. 10.9.2).

28 Other cases of non-paradigmatic linking elements came about because the simplex changed its inflectional paradigm while the compound kept the old, fossilized element. They can, naturally, not be used to determine productivity (see also Kopf 2018b: Kap. 5.1.1).

almost no cases in which these suffixes are not followed by linking-*s*. On the other hand, only 4.3% of all non-paradigmatically linked compound types (n=443) from the Wortwarte Corpus are simplexes (Kopf 2018b: Kap.12.1). Their first constituents form a small, closed group. The non-paradigmatic linking-*s* is thus closely tied to certain suffixes or lexemes; when it is used in new compounds, it is usually kept. However, it does not spread to new first constituents. This shows, unsurprisingly, that linked compounds in NHG follow a morphological pattern, although it points to low productivity of the linked pattern.

It is more instructive to analyze the phenomenon in the Mainz Corpus when it was still new and spreading. The first examples can be found in 1590, and the number of different types (i.e. the realized productivity of the pattern) increases strongly in 1710 (Fig. 5):

Fig. 5: Non-paradigmatic linking-*s* in the Mainz Corpus (tokens=80, types=54).



The non-paradigmatic linking-*s* is closely connected to a number of suffixes and lexemes from the beginning (see below), but still shows some variation with unlinked forms. How and why this changed will be addressed in the next section.

Linking-*(e)n*, the only other linking element of any consequence, shows almost no innovative use (but see Klein 2015). It is firmly tied to the respective inflectional paradigm. It is the default linking element with mixed feminines and weak masculines – the latter consisting almost exclusively of a small group of animate nouns (for details, see Kopf 2018b: 43). Linking-*(e)n* is therefore not relevant to determine the productivity of the linked compounding pattern.

We can conclude that a morphologically productive compounding pattern with linking elements was present from 1590 in the Mainz Corpus. (For an overview of earlier attestations in other sources, see Kopf 2018b: Kap. 12.2.) This is mostly shown by the non-paradigmatic linking-*s*, but supported by the occurrence of semantically “wrong” compounding stem forms by the beginning of the 17th century.

5. Compounding productivity: An explanation

Thus far, the following has been shown:

1. Transparent derivatives were not used as first constituents in OHG and MHG compounds but are attested in ENHG compounds, where they increase from 1590. This happened rather abruptly, compared with the slow extension of tripartite compounds that were attested from the beginning. Derivatives are completely unremarkable in NHG compounds.
2. In ENHG, a new compounding pattern with linking elements arose due to univerbation and reanalysis of prenominal genitive constructions. This pattern is still in use today, while its source, univerbation, has run dry due to the positional change of the genitive attribute in (E)NHG.
3. One of these linking elements, *-s-*, was extended to non-paradigmatic use in combination with feminine nouns. This can be taken as proof of a word-formation pattern independent from syntax. The non-paradigmatic linking-*s* is strongly associated with a number of suffixes in NHG.

In addition, it is important to note that NHG paradigmatic linking-*s* (used with masculine and neuter nouns) also shows a remarkable affinity for complex first constituents: Apart from a few suffixes (*-tum*, *-ling*), it also combines with many prefixed nouns, although there is some variation (Kopf 2018b: 367, see also Augst 1975: 113–120). In total, 81.3% of prefixed masculine and neuter compound types with a monosyllabic basis tested in a comprehensive corpus study showed linking-*s* (Kopf 2017, Kopf 2018b: Kap. 14.2.2).²⁹ The Mainz Corpus contains only 26 prefixed cases for all grammatical genders; however, 19 of these use linking-*s* (73%).

These observations lead to a plausible explanation for the increased productivity of compounding in general, which is mostly due to loss of restrictions: The traditional pattern did not include derivatives as first constituents.

²⁹ Feminine nouns, for which *-s-* is non-paradigmatic, still used it in 53.8% of the prefixed cases.

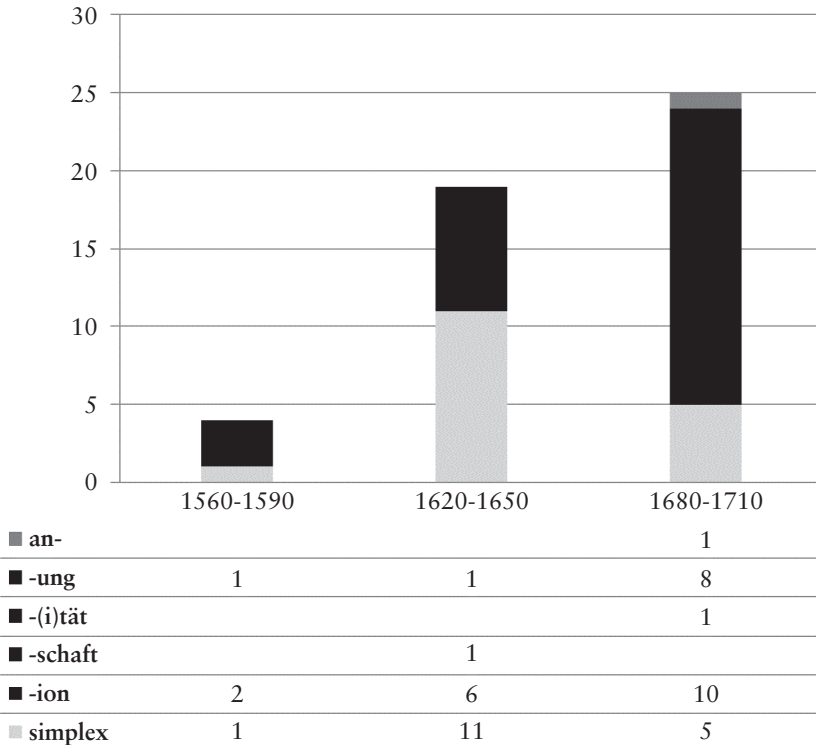
When speakers started to reanalyze genitive constructions as compounds in ENHG,³⁰ this changed. As genitive constructions, being syntactic units, did not show formal restrictions for their genitive attribute, univerbation “smuggled” structures into morphology that had not been used up to this point, especially derivatives. The morphological restriction on derivatives in compounds fell.

In many cases, these derivatives brought a former genitive suffix with them, which then became a paradigmatic linking element. In most cases, this was *-s-*. Of course, this not only happened with derivatives but also with simplexes, e.g. *des leibls erben* > *die Leiblslerben*. The important difference is that morphologically simple first constituents were already possible in the old, unlinked compounding pattern. The linking-*s* therefore appeared much more regularly with derivatives than with simplexes. This quickly turned the linking-*s* into a marker for internal complexity: Where it appeared, a complex first constituent was very likely. This might have been a useful segmentation aid for speakers that were now confronted with much more complicated compounds.

This marker was quickly applied to compounds with derivative feminine first constituents. As most of them did not show an inflectional suffix in the genitive singular, complex constituents would have looked like old compounds. A closer look at the internal structure of feminine first constituents in the Mainz Corpus (Fig. 6) shows, however, that they make up a considerable part of the non-paradigmatic cases from the beginning.

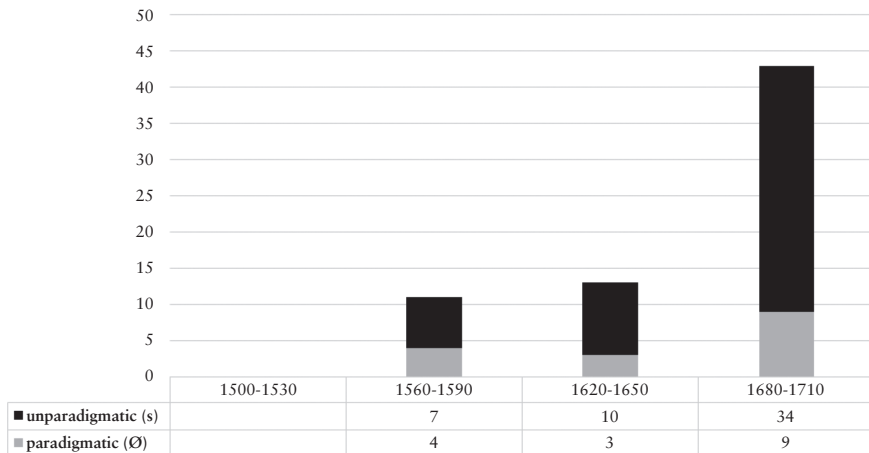
30 Several explanations have been brought forward as to why univerbation becomes so common in ENHG, e.g. by Pavlov (1983) and Demske (2001). See Kopf (2018b: Kap. 8.6) for an evaluation and some new aspects.

Fig. 6: First attestations of compounds with non-paradigmatic linking-*s* by morphological complexity in the Mainz Corpus ($n=48$). For ease of comparison, all suffixed first constituents have been assigned the same color; individual numbers can be seen in the table.



The non-paradigmatic linking-*s* is not used as invariably as today when encountering suffixes. However, Fig. 7 shows a remarkably early affinity: Non-paradigmatic cases dominate from the beginning for first constituents in *-ung*, *-(i)tät*, *-schaft* and *-ion*, while cases without linking-*s* persist for the whole period of investigation, but to a much smaller degree.

Fig. 7: *Compounding patterns of first constituents in -ung, -(i)tät, -schaft and -ion in the Mainz Corpus (tokens, n=67), the other suffixes are not yet attested. Data combined from Kopf (2018b: 313).*



Thus, it seems reasonable to assume that the spread of the non-paradigmatic linking-s was motivated by the unconscious will to mark derivatives in compounds, segmenting them for listeners and readers (on (E)NHG compound segmentation, see also Kopf 2017). The linking element is an even better signal here than with paradigmatic cases; as it appears only in compounds, confusion with syntactic structures is completely impossible.

The spread of the non-paradigmatic linking-s shows a temporary functionalization of linking-s in general as a complexity marker (for factors that might have helped it along additionally, see Kopf 2018b: Kap. 12.3). I consider this as a case of exaptation (Kopf 2018b: Kap. 14.4). When the new compounding pattern was firmly integrated into word-formation and the use of derivatives was no longer restricted, the linking-s was no longer needed. The exaptation process was interrupted and linking-s stopped expanding to new lexemes. It fossilized in fixed compounding stem forms.

6. Conclusion

The present paper has shown why and how productivity in German compounds can and should be measured: The status quo would hardly raise any interest in the topic, as compounds are obviously the means of word-formation in NHG. However, OHG and MHG show a different picture. Compounds were formed and used, but the morphological complexity of the first constituent was heavily restricted, especially in the case of derivatives. As conditions have obviously changed since then, this warrants a detailed investigation.

Quantitative measurements of productivity that are regularly employed in derivational morphology have been extended to compounding. Realized productivity shows a heavy increase in overall compounding as well as in patterns differentiated by linking pattern. Potential productivity shows no significant change, but it is shown that this is most likely due to corpus size and fluctuations in token numbers. It has become clear that the quantitative measures, while yielding interesting results, cannot stand for themselves and can hardly be used to generate any further hypotheses. This can be done much more elegantly by a qualitatively motivated approach. Comparing OHG and NHG yields morphological complexity as a suspect. It can in fact be shown that derivationally complex first constituents became possible and quickly more frequent in the 16th century. Tripartite compounds, on the other hand, have always existed. Their usage frequency rose constantly over the centuries, but is still only moderately high today. Concentrating on derivatives, it can be shown that their rise in compounds coincides with univertation and reanalysis of compounds from genitive constructions. Syntax was a back door that enabled derivatives to enter morphology, removing a long-standing restriction on first constituents. As masculine and neuter derivatives usually brought along a former genitive-s, the new, more complex pattern was overtly marked. The linking-s was partly functionalized as a marker of internal derivational complexity, even spreading to feminine nouns where it was non-paradigmatic. It was, however, only needed as a temporary crutch: When the new compounding type was firmly integrated into word-formation, the spread of the linking-s was halted. Since the beginning of the 20th century, very few isolated cases of analogy are attested (e.g. *Trainings|hose* ‘tracksuit (verb. training|E|suit)’).

References

- Aronoff, Mark & Nanna Fuhrhop. 2002. Restricting suffix combinations in German and English: Closing suffixes and the monosuffix Constraint. *Natural Language & Linguistic Theory* 20. 451–490.
- Augst, Gerhard. 1975. *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache*. Tübingen: Narr.
- Baayen, R. H. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics* (Handbücher zur Sprach- und Kommunikationswissenschaft 2), 900–919. Berlin & New York: de Gruyter.
- Bauer, Laurie. 2005. Productivity: Theories. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation* (Studies in natural language and linguistic theory 64), 315–334. Dordrecht: Springer.
- Bergmann, Rolf & Dieter Nerius. 1998. *Die Entwicklung der Großschreibung im Deutschen von 1500 bis 1700*. Heidelberg: Universitätsverlag Winter.
- Carr, Charles T. 1939. *Nominal compounds in Germanic*. London: Oxford University Press.

- Demske, Ulrike. 1999. Case Compounds in the history of German. In Matthias Butt & Nanna Fuhrhop (eds.), *Variation und Stabilität in der Wortstruktur. Untersuchungen zur Entwicklung, Erwerb und Varietäten des Deutschen und anderer Sprachen*, 150–176. Hildesheim: Olms.
- Demske, Ulrike. 2001. *Merkmale und Relationen. Diachrone Studien zur Nominalphrase des Deutschen*. Berlin & New York: de Gruyter.
- Donalies, Elke & Noah Bubenhofer. 2011. *Tagtraum, Tageslicht, Tagedieb: Ein korpuslinguistisches Experiment zu variierenden Wortformen und Fugenelementen in zusammengesetzten Substantiven* (Arbeitspapiere und Materialien zur deutschen Sprache 42). Mannheim: Institut für Dt. Sprache.
- Durrell, Martin, Astrid Ensslin & Paul Bennett. 2007. The GerManC project. *Sprache und Datenverarbeitung* 31. 71–80.
- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache* (de Gruyter Studium), 4th edn. Tübingen: Niemeyer.
- Fuhrhop, Nanna. 1996. Fugenelemente. In Ewald Lang & Gisela Zifonun (eds.), *Deutsch – typologisch*, 525–550. Berlin & New York: de Gruyter.
- Fuhrhop, Nanna. 1998. *Grenzfälle morphologischer Einheiten*. Tübingen: Stauffenburg.
- Grimm, Jacob. 1826. *Deutsche Grammatik. Zweiter Theil* (2). Göttingen: Dieterich.
- Gröger, Otto. 1911. *Die althochdeutsche und altsächsische Kompositionsfuge. Mit Verzeichnis der althochdeutschen und altsächsischen Composita*. Zürich: Zürcher und Furrer.
- Günther, Hartmut. 1981. N+N: Untersuchungen zur Produktivität eines deutschen Wortbildungstyps. In Leonhard Lipka & Hartmut Günther (eds.), *Wortbildung*, 258–280. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Harlass, Gertrude & Heinz Vater. 1974. *Zum aktuellen deutschen Wortschatz*. Tübingen: Narr.
- Hartmann, Stefan. 2016. *Wortbildungswandel. Eine diachrone Studie zu deutschen Nominalisierungsmustern*. Berlin & New York: de Gruyter.
- Kempf, Luise. 2016. *Adjektivsuffixe in Konkurrenz. Wortbildungswandel vom Frühneuhochdeutschen zum Neuhochdeutschen* (Studia Linguistica Germanica 126). Berlin & New York: de Gruyter.
- Klein, Andreas. 2015. *Instrumentenklänge und Temperamentszustände – Fälle zwischen -en und -s: Zur unparadigmischen en-Fuge in Lehnwortkomposita: Genese und Distribution*. Mainz: Johannes Gutenberg-Universität Hausarbeit.
- Klos, Verena. 2011. *Komposition und Kompositionalität: Möglichkeiten und Grenzen der semantischen Dekodierung von Substantivkomposita* (Reihe Germanistische Linguistik 292). Zugl.: Siegen, Univ., Diss., 2009. Berlin & New York: de Gruyter.
- Kopf, Kristin. in preparation. Erkennung von Komposita im Frühneuhochdeutschen. In Christian Schwarz & Christine Ganslmayer (eds.), *Historische Wortbildung. Theorie - Methoden - Perspektiven*. Hildesheim: Olms.
- Kopf, Kristin. 2017. Fugenelement und Bindestrich in der Compositions-Fuge. Zur Herausbildung phonologischer und graphematischer Grenzmarkierungen in (früh)neuhochdeutschen N+N-Komposita. In Renata Szczepaniak, Nanna Fuhrhop &

- Karsten Schmidt (eds.), *Sichtbare und hörbare Morphologie*, 177–204. Berlin & New York: de Gruyter.
- Kopf, Kristin. 2018a. From genitive suffix to linking element. A corpus study on the genesis and productivity of a new compounding pattern in (Early) New High German. In Horst J. Simon, Christian Zimmer & Tanja Ackermann (eds.), *Germanic Genitives*. Amsterdam & Philadelphia: Benjamins.
- Kopf, Kristin. 2018b. *Fugenelemente diachron. Eine Korpusuntersuchung zu Entstehung und Ausbreitung der verfügenden N+N-Komposita*. Berlin & Boston: de Gruyter.
- Krahe, Hans & Wolfgang Meid. 1967. *Germanische Sprachwissenschaft* (Sammlung Göschen 1218). Berlin: de Gruyter.
- Kürschner, Sebastian. 2003. *Von Volk-s-musik und Sport-Ø-geist im Lemming-Ø-land – af folk-e-musik og sport-s-ånd i lemming-e-landet: Fugenelemente im Deutschen und Dänischen – eine kontrastive Studie zu einem Grenzfall der Morphologie*. Freiburg: Albert-Ludwigs-Universität Magisterarbeit.
- Meibauer, Jörg, Ulrike Demske, Jochen Geilfuß-Wolfgang, Jürgen Pafel, Karl-Heinz Ramers, Monika Rothweiler & Markus Steinbach. 2015. *Einführung in die germanistische Linguistik*, 3rd edn. Stuttgart: Metzler.
- Meineke, Eckhard. 1991. Springlebendige Tradition. Kern und Grenzen des Kompositums. *Sprachwissenschaft* 16(1). 27–88.
- Munske, Horst H. 2009. Was sind eigentlich ‚hybride‘ Wortbildungen? In Peter O. Müller (ed.), *Studien zur Fremdwortbildung* (Germanistische Linguistik 197/198), 223–260. Hildesheim: Olms.
- Nitta, Haruo. 1987. Zur Forschung der ‚uneigentlichen‘ Zusammensetzung im Frühneuhochdeutschen. *Zeitschrift für deutsche Philologie* 106(3). 400–416.
- Nübling, Damaris & Renata Szczepaniak. 2008. On the way from morphology to phonology: German linking elements and the role of the phonological word. *Morphology* 18. 1–25.
- Nübling, Damaris & Renata Szczepaniak. 2013. Linking elements in German. Origin, Change, Functionalization. *Morphology* 23(1). 67–89.
- Olsen, Susan. 2001. Copulative compounds: A closer look at the interface between syntax and morphology. In Geert Booij, Jaap van Marle, Stephen Anderson, Mark Aronoff, Mark Baker, Laurie Bauer, Rudie Botha, Joan Bybee, Andrew Carstairs-McCarthy, Greville Corbett, Wolfgang Dressler, Martin Haspelmath, Jack Hoeksema, Rochelle Lieber, Peter Matthews, Franz Rainer, Sergio Scalise, Henk Schultink & Andrew Spencer (eds.), *Yearbook of Morphology 2000* (Yearbook of Morphology), 279–320. Dordrecht: Springer Netherlands.
- Ortner, Lorelies, Elgin Müller-Bollhagen, Hanspeter Ortner, Hans Wellmann, Maria Pümpel-Mader & Hildegard Gärtner. 1991. *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache. Eine Bestandsaufnahme des Instituts für deutsche Sprache, Forschungsstelle Innsbruck* (4 Substantivkomposita (Komposita und kompositionsähnliche Strukturen 1)). Berlin & New York: de Gruyter.

- Pavlov, Vladimir M. 1983. *Von der Wortgruppe zur substantivischen Zusammensetzung* (Zur Ausbildung der Norm der deutschen Literatursprache (1470–1730) 4). Berlin: Akademie-Verlag.
- Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation* (Topics in English Linguistics [TiEL] 28). Berlin & New York: de Gruyter.
- Pümpel-Mader, Maria, Elsbeth Gassner-Koch & Hans Wellmann. 1992. *Adjektivkomposita und Partizipialbildungen: Komposita und kompositionsähnliche Strukturen 2* (Sprache der Gegenwart 80). Berlin & New York: de Gruyter.
- Rainer, Franz. 2005. Constraints on productivity. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation* (Studies in natural language and linguistic theory 64), 335–352. Dordrecht: Springer.
- Reagan, Sherman C. 1981. *Compound nouns in the Luther Bible and some of its printed predecessors*. Madison, Wisconsin: University of Wisconsin-Madison dissertation.
- Salus, P. H. 1963. *Nominal composition in the Indo-European languages*. New York Dissertation.
- Scalise, Sergio & Antonietta Bisetto. 2009. The Classification of Compounds. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Compounding*, 49–82. Oxford: Oxford University Press.
- Schmid, Hans-Jörg. 2011. *English morphology and word-formation: An introduction* (Grundlagen der Anglistik und Amerikanistik 25), 2nd edn. Berlin: Schmidt.
- Solling, Daniel. 2011. Zur Problematik der Unterscheidung zwischen pränominalem Genitivattribut und getrennt geschriebenem Substantivkompositum im Frühneuhochdeutschen. In Hans U. Schmid & Arne Ziegler (eds.), *Historische Semantik* (Jahrbuch für germanistische Sprachgeschichte 2), 294–311. Berlin & Boston: de Gruyter.
- Solling, Daniel. 2012. *Zur Getrennt-, Zusammen- und Bindestrichschreibung von Substantivkomposita im Deutschen (1550–1710)* (Acta Universitatis Upsaliensis: Studia Germanistica Upsaliensia 57). Zugl.: Uppsala, Univ., Diss., 2012.
- Szczepaniak, Renata. 2016. Is the development of linking elements in German a case of exaptation? In Muriel Norde & Freek van de Velde (eds.), *Exaptation and language change* (Current Issues In Linguistic Theory Volume 336), 317–340. Amsterdam, Philadelphia: John Benjamins.
- Wellmann, Hans, Nikolaus Reindli & Annemarie Fahrmeier. 1974. Zur morphologischen Regelung der Substantivkomposition im heutigen Deutschen. *Zeitschrift für deutsche Philologie* 93. 358–378.
- Wilmanns, Wilhelm. 1896. *Deutsche Grammatik. Gotisch, Alt-, Mittel- und Neuhochdeutsch. Bd. 2: Wortbildung*. Straßburg: Trübner.

Dr. Kristin Kopf
 Westfälische Wilhelms-Universität Münster
 Germanistisches Institut
 Schlossplatz 34
 48143 Münster
 kristin.kopf@uni-muenster.de

Luise Kempf & Matthias Eitelmann

Von *diutisk* zu *dynamisch*, von *englisc* zu *anything-ish*. *-is(c)h* kontrastiv diachron

Abstract: German *-isch* and English *-ish* share a common Germanic origin, which is evidenced by striking similarities concerning the derivation of ethnic adjectives (*englisch/English*) or property-denoting adjectives (*kindisch/childish*). However, after an initial period of parallel characteristics, the two languages display drastic changes, with English developing an approximative sense when attached to adjectival bases (e.g. *greenish*) and expanding to a wide range of other word categories, while German *-isch* develops multiple functions and also comes to firmly occupy a morphological niche with non-native bases. The paper sheds light on the evolving divergence between German and English by presenting results from two diachronic corpus-based studies. Additionally, explanations with respect to the typological parameter of ‘Boundary Permeability’ are provided.

1. Einleitung

Im Deutschen wie im Englischen steht das Derivationsuffix *-isch* bzw. *-ish* als gleichermaßen produktives und kreatives Wortbildungselement zur Verfügung. Davon zeugen exemplarisch gegenwartssprachliche Neubelege bzw. Hapax legomena, wie sie sich etwa aus der *Wortwarte* (siehe (1)) bzw. dem *Corpus of American Soap Operas* (siehe (2)) extrahieren lassen:

- (1) *nutrigenomisch, pugliesisch, hypodisch, biosystematisch, cardobarbarisch, trophologisch, anamorphotisch, genografisch, spagyrisch, aurisch, demokraturisch, fallibilistisch, fidelistisch, fernsehtopographisch, präwikipedianisch*
- (2) *superhero-ish, love-ish, down-under-ish, Thansgivingish, responsible-ish, 94-ish, happily-ish, déjà-vu-ish, big-brother-ish, chocolate-blondish, caveman-esque-ish, demi-ish, semi-normal-ish, there-ish, done-ish*

Auffällig bei den deutschen *isch*-Derivaten ist, dass ihre Basen strukturell den Status von Fremdwörtern haben – so dürfte einem Außenseiter ohne einen ausreichenden fachwissenschaftlichen Hintergrund nicht auf Anhieb klar sein, was mit *nutrigenomisch* oder *hypodisch* gemeint ist. Der Fremdwortcharakter resultiert aus dem Umstand, dass die Basen allesamt fremden Ursprungs sind, von nominalen Basen lateinischer Herkunft wie bei *Aura* – *aurisch* über komplexere Kontaminationen wie bei *Demokratur* – *demokraturisch* bis hin zu Eigennamenbasen wie bei *Fidel (Castro)* – *fidelistisch*. Eine Nähe zum akademischen Register ist bei all diesen Neubildungen leicht erkennbar, auch bei dem zusätzlich präfigierten *präwikipedianisch*.

Dem gegenüber stehen englische Hapaxbildungen, die auf den ersten Blick eine größere Vielfalt bezeugen, da sie bei weitem nicht auf denominaler oder deonymischer Ableitungen fremder Herkunft beschränkt sind. Die Basen sind weit heterogener, und bei der Mehrzahl dieser Hapaxe liegt nahe, dass sie in formalen, schriftbasierten Registern eher selten bis gar nicht vorkommen. Die Semantik so mancher dieser *-ish*-Bildungen weicht von derjenigen der deutschen eklatant ab, da das Suffix neben einer vergleichenden eine approximative Bedeutung innezuhaben vermag: so z. B. *responsible-ish* ‚quasi verantwortungsvoll‘, *94-ish* ‚um 94 herum‘, *there-ish* ‚ungefähr dort‘ oder *done-ish* ‚vollbracht, aber irgendwie nicht wirklich‘. Die umständliche Wiedergabe dieser englischen *-ish*-Hapaxe im Deutschen zeigt direkt, dass die Funktionsspektren der Suffixe in den beiden Sprachen nur bedingt überlappen.

Die beobachtbare Divergenz zwischen dem deutschen *-isch* und dem englischen *-ish* ist umso verwunderlicher, als beide auf einen gemeinsamen Ursprung zurückgehen: das germanische Suffix **-isko-*. Tabelle 1 skizziert die formalen Entwicklungen, die das Suffix in beiden Sprachen bis heute durchlaufen hat.

Tab. 1: Formale Entwicklung von dt. *-isch*, engl. *-ish*.¹

germ. <i>*-isko-</i>	
ahd. <i>-isc, -isk, -isg, ...</i>	ae. <i>-isc, -isce, ...</i>
mhd. <i>-isch, -esch, ...</i>	me. <i>-ish, -isshe, ...</i>
(f)nhd. <i>-isch</i>	(f)ne. <i>-ish</i>

Während formal große Ähnlichkeit besteht, haben sich dt. *-isch* und engl. *-ish* im Laufe ihrer jeweiligen Sprachgeschichten funktional stark auseinanderentwickelt. Ursprünglich hatte das Suffix die Funktion, Adjektive der Zugehörigkeit und Herkunft abzuleiten, z. B. altnordisch *engliskr* ‚englisch‘. Davon zeugen einige aus verschiedenen altgermanischen Sprachen belegte Bildungen, vgl. die Beispiele in Tabelle 2. Allerdings weisen auch einige altsprachliche Lexeme idiomatische Bedeutungen auf. Dabei handelt es sich um Eigenschaften, die der entsprechenden Zugehörigkeit bzw. Herkunft zugeschrieben werden sowie nicht selten auch deren pejorative Evaluation z. B. an. *heimskr* ‚einfältig, dumm, töricht‘, vgl. Tabelle 3.

1 Die Formvarianten sind folgenden Quellen entnommen: Goetze (1899: 470), Klein (1994: 385), Ganslmayer (2012: 493), Bosworth (2010), Middle English Dictionary, OED online (Eintrag *-ish*). Die Epochenabkürzungen stehen für: Althochdeutsch (500/750–1050), Mittelhochdeutsch (1050–1350), Frühneuhochdeutsch (1350–1650); Altenglisch (450–1100), Mittelenglisch (1100–1500), Frühneuenglisch (1500–1750).

Tab. 2: Frühe -isch-Adjektive der Zugehörigkeit und Herkunft².

Ahd.	Ae.	An.	Got.	Bedeutung
<i>mennisc</i>	<i>mennisc</i>	<i>men(n)skr</i>	<i>mannisks</i>	‚menschlich‘
<i>heidanisc</i>	<i>hæðenisc</i>			‚heidnisch‘
<i>judeisc</i>	<i>Judéisc</i>		<i>judaiwisks</i>	‚jüdisch‘
	<i>Englisc</i>	<i>engliskr</i>		‚englisch‘

Tab. 3: Frühe idiomatisierte -isch-Adjektive.

Ahd.	Ae.	An.	Got.	Etymon
<i>himilisc</i> ‚himmlisch, göttlich‘		<i>himneskr</i> ‚himmlisch‘		germ. * <i>hemila-</i> ‚Himmel‘
<i>diutisk</i> ‚national, deutsch‘	<i>þéodisc</i> ‚national‘		<i>þiudiskô</i> ‚heidnisch‘	germ. * <i>þéudô</i> ‚Volk‘
<i>heimisc, -k</i> ‚heimisch; ungebildet u.a.‘		<i>heimskr</i> ‚dumm, töricht, einfältig‘		germ. * <i>haimaz</i> ‚Heim‘
	<i>cildisc</i> ‚kindlich, kindhaft‘			ae. <i>cild</i> ‚Kind‘, vgl. got. <i>kilþei</i> ‚Uterus‘
		<i>bernskr</i> ‚kindisch‘	<i>barnisks</i> ‚kindlich, kindisch‘	germ. * <i>barna-</i> ‚Kind‘

Die metonymische Erweiterung von ‚zu X gehörig‘ zu ‚in der (typischen) Art von X, wie X‘ ist also – soweit man aus den belegten Bildungen schließen kann – in allen altgermanischen Sprachen zumindest im Ansatz vorhanden. Jedoch stellt sich die kontrastiv-diachrone Frage, was von diesem Stadium zu den heutigen stark auseinanderklaffenden Funktionsspektren führt und inwiefern diese Ausdifferenzierung aus der mehr oder weniger produktiven Eroberung stets neuer Kontexte und Funktionen resultiert. Mit anderen Worten: Wie genau kommt es nach einem gemeinsamen Ursprung, repräsentiert durch die jeweiligen Nationalitätsbezeichnungen *diutisk* und *englisc*, zu solch

2 Die Daten in Tabelle 2 und 3 sind aus folgenden Quellen zusammengestellt: Bosworth (2010), Goetze (1899: 470), Klein (1994: 385), Köbler (2014a, b, c), OED online (Eintrag *-ish*), Pfeifer (1993). ‚An.‘ = Altnordisch, ‚Got.‘ = Gotisch.

einer Auseinanderentwicklung, die sich mit *dynamisch* und (so der Eindruck) *anything-ish*³ veranschaulichen lässt?

Um dieser Frage empirisch nachzugehen, werden wir zunächst getrennt die Entwicklungspfade der beiden Suffixe in den jeweiligen Sprachen nachzeichnen, jeweils eingeleitet durch einen kurzen Abriss relevanter Forschungsliteratur und methodologische Vorbemerkungen zu den einzelsprachlichen Fallstudien. Die Einzelstudien, die in dem vorliegenden Aufsatz vorgestellt werden, beruhen auf unterschiedlichen methodologischen Vorgehensweisen, die dadurch bedingt sind, dass für das Deutsche und das Englische zum einen unterschiedlich zusammengesetzte Korpora und zum anderen divergierende Forschungsstände vorliegen. Daher sind die beiden Einzelstudien nicht komplett analog aufgebaut, berücksichtigen aber analoge Analyse Kriterien (z. B. Komplexität der Ableitungsbasen). Beide sind als langzeitdiachrone Studien angelegt und erlauben aufgrund überlappender Korpuszeiträume sowie einer gemeinsamen Schnittmenge der untersuchten Genres (i.e. narrative Prosa) eine prinzipiell vergleichbare Einschätzung der Entwicklungspfade vom Germanischen bis heute. Die in den beiden Einzelstudien getroffenen Beobachtungen werden in einer kontrastiven Synopse zusammengeführt, die Ähnlichkeiten und Unterschiede in den jeweiligen Karrieren der Wortbildungselemente auf den Punkt bringt. Im Anschluss daran werden die kontrastiven Beobachtungen in einen größeren theoretischen Kontext eingeordnet, indem die beobachtbaren gravierenden Divergenzen mit Bezug auf Bergs (2014) typologischen Parameter der *Boundary Permeability* diskutiert werden. Auf diese Weise soll ein substantieller Beitrag geleistet werden, das bislang noch offene Forschungsdesiderat einer empirisch fundierten diachron-kontrastiven Analyse von dt. *-isch*/engl. *-ish* aufzuarbeiten.

2. Diachrone Entwicklungspfade

2.1. Diachrone Entwicklung von dt. *-isch*

2.1.1. Kurzer Abriss relevanter Forschungsliteratur

Die deutsche *-isch*-Derivation ist bislang überwiegend synchron erforscht. Der gegenwartssprachliche Stand wird zum einen in allgemeinen Wortbildungslehren (Motsch 2004, Fleischer und Barz 2012, u.a.m.) und Korpusstudien (Kühnhold, Putzer und Wellmann 1978) behandelt, zum anderen in zwei Monographien (Schlaefter 1977, Eichinger 1982). Historisch-synchrone

3 Einen Beleg für *anything-ish* findet man in der Synopse eines Kinderbuchs von Peter H. Reynolds (2005), das den bezeichnenden Titel *Isb* trägt: „Ramon loves to draw, especially when he learns that a drawing doesn't have to look exactly like anything. It can simply look anything-ish“.

Korpusuntersuchungen mit diachronen Ausblicken liefern Thomas (2002) und Ganslmayer (2012) (in der Adjektivderivation im Nürnberger Fnhd. um 1500 bzw. zur Urkundensprache des 13. Jh.) sowie Klein, Solms und Wegera (2009) (zur mhd. Wortbildung insgesamt). An diachronen Betrachtungen liegen Kleins (1994) Analyse alt- und mittelhochdeutscher *-isch*-Derivate, Goetzes (1899) Beitrag mit Schwerpunkt v.a. auf pejorativen Ableitungen sowie Sugarewas (1974) Studie zu deonymischer Derivation vor. Diesen Schwerpunkt verfolgt auch Hornbruch (1996) in ihrer Studie zu ahd. und mhd. Deonomastika. Die vorliegende Untersuchung knüpft an die Untersuchungen in Kempf (2016a, zum Wandel der Adjektivderivation vom Fnhd. zum Nhd.), Kempf (2016b, zur Modellierung polysemer Ableitungsmuster) und Kempf (2017, zur deonymischen *-isch*-Derivation im Fnhd. und Nhd.) an, wobei sie die kontrastiv relevanten Aspekte (z. B. Basiskategorien) fokussiert und ausbaut.

2.1.2. *Methodologische Vorbemerkungen*

Die folgende Beschreibung des Entwicklungspfades der deutschen *-isch*-Derivation stützt sich auf eine diachrone Hauptstudie und zwei kleinere synchrone Nebenstudien. Die diachrone Studie arbeitet mit der Datenbank, die auch der Untersuchung in Kempf (2016a) zugrundeliegt. Diese Datenbank wurde erstellt durch Extrahierung aller mit Suffixen abgeleiteten Adjektive aus dem Bonner Fnhd.-Korpus (1350–1700) sowie Teilen des German Manchester Korpus (GerManC, 1650–1800). Beide Korpora sind in 50-Jahres-Perioden gegliedert, wobei die Periode 1650–1700 von beiden Korpora abgedeckt ist. Um je Zeitschnitt ein etwa konstantes Korpusvolumen zu erzielen, wurden nur vier der acht Textsorten des GerManC-Korpus benutzt. Dabei wurden die Textsorten mit dem Ziel einer möglichst hohen Vergleichbarkeit zum Bonner Fnhd.-Korpus gewählt: die Textsorten Predigt (da hoher Anteil religiöser Texte im Bonner Korpus), Zeitung (korrespondiert am ehesten mit „Chronikalischer und Berichtstext“), narrative Prosa (Entsprechung zu Belletristik) sowie Naturwissenschaftliche Sachtexte (Entsprechung zu „Fachprosa“; für genauere Angaben zu Korpus und Datenbank s. Kempf 2016a: 102–113). Beide Korpora zusammengenommen umfassen ca. 890.000 Textwörter, was ca. 24.000 Tokens/1.885 Types suffigierter Adjektive ergab. Auf die *-isch*-Derivation entfielen dabei 2.744 Tokens/517 Types (vgl. Tabelle 4).

Tab. 4: *Quantitative Werte zur diachronen Korpusstudie des Deutschen.*

Periode	Bonner FnhdC				GerManC			Gesamt
	1350–1400	1450–1500	1550–1600	1650–1700	1650–1700	1700–1750	1750–1800	
Korpusgröße	123.952	138.314	125.163	127.525	125.318	126.172	123.848	890.292
alle suffigierten Adjektive, Tokens	2.189	2.447	3.201	4.317	3.698	4.052	4.108	24.012
alle suffigierten Adjektive, Types	335	407	517	689	726	794	904	1.885
<i>-isch</i> Tokens	103	124	315	491	538	587	586	2.744
<i>-isch</i> Types	15	24	84	140	175	208	218	517

Die diachrone Korpusstudie wurde um zwei kleinere synchrone Studien ergänzt: Zum einen wurde eine Stichprobenanalyse aus DWDS-Korpusdaten des 20. Jh. durchgeführt. Hierbei wurden 100 zufällige *-isch*-Types sowie 100 zufällige *-isch*-Hapax legomena untersucht.⁴ Zum anderen wurde eine Sammlung möglichst rezenter *-isch*-Neubildungen aus verschiedenen Quellen zusammengestellt,⁵ um Einblicke in die aktuellste Wortbildung zu ermöglichen.

Die Analyse der Basiskategorien und des funktionalen Wandels stützt sich außerdem auf die Ergebnisse aus Schlaefer (1977), Kühnhold, Putzer und Wellmann (1978), Klein (1994), Thomas (2002), Klein, Solms und Wegera (2009) sowie Ganslmayer (2012). Entsprechend orientiert sich die Bezeichnung der Wortbildungsfunktionen (z. B. ‚possessiv–ornativ‘ = ‚mit X‘ wie in *tückisch*) an den in diesen Werken verwendeten Klassifikationssystemen (die im Wesentlichen übereinstimmen oder sich ineinander überführen lassen; für einen vergleichenden Überblick s. Ganslmayer 2012: 1124–1127).

4 Beide Zufallssamples entstammen der Treffermenge folgender Abfrage: „\$p=/ADJ./ with /.*.[fr]isch(e(slnlmlr)?)?\$/“, ddc2 Kernkorpus 20/Version 1.1, aus der 5000 Zufallstreffer exportiert wurden (12.05.2017).

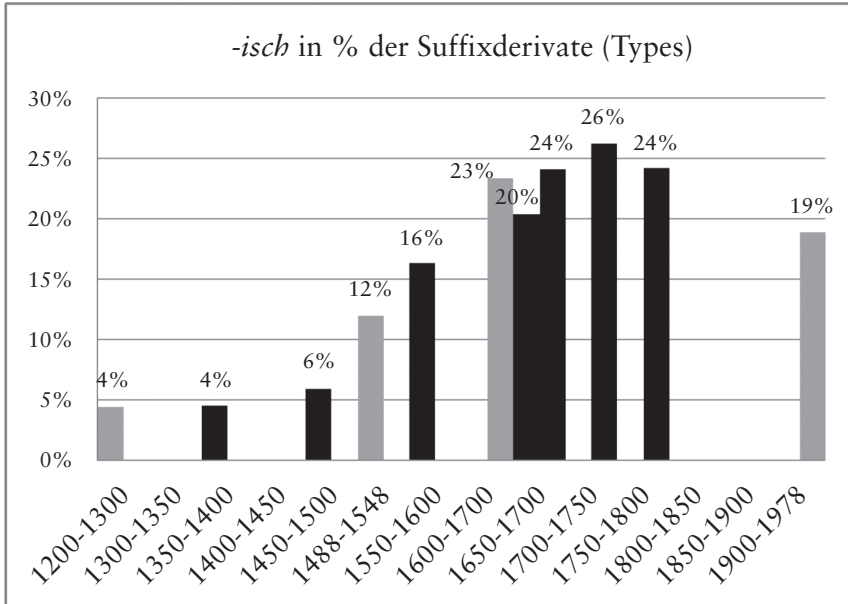
5 Die Sammlung umfasst 18 Neubildungen aus dem späten 20. sowie frühen 21. Jh. Sie wurde – unter gründlichem Aussortieren dessen, wofür sich doch ältere Belege finden ließen (Google-Suche, Google NGram Korpus u.a.m.) – anhand folgender Quellen zusammengestellt: Motsch (2004: 198, 221, 247), Daten der Neologismen-Plattform *Wortwarte* (<http://www.wortwarte.de/>) sowie der Wortwarte zugrundeliegende ungefilterte Rohdaten, die uns der Betreiber L. Lemnitzer freundlicherweise zur Verfügung gestellt hat und die wir in Zufallsausschnitten ausgewertet haben.

2.1.3. Entwicklungspfad von dt. -isch

Die *-isch*-Derivation im Deutschen zeigt eine bemerkenswerte Produktivitätsentwicklung: Vom 16. bis zum 18. Jh. hat das Muster eine immense Produktivitätszunahme erfahren. Dies wird in Kempf (2016a: 248–253) durch den Vergleich mehrerer Produktivitätsmaße⁶ ermittelt, die in dieser Tendenz übereinstimmen. Abbildung 1 zeigt als ein Beispiel unter den Produktivitätsmaßen den Verlauf, wie er sich im Spiegel des Typenanteils darstellt – d.h. des Anteils, den die *-isch*-Derivate eines Zeitschnitts an den Types aller suffigierten Adjektive (*-lich*, *-ig* u.a.m.) dieses Zeitschnitts halten. Dieses Maß zeigt nur die realisierte Produktivität und muss immer mit Rücksicht auf die Entwicklungen der anderen Muster interpretiert werden. Der Vorteil des Maßes liegt in der guten Vergleichbarkeit mit anderen Studien.⁷ Die Ergebnisse aus anderen Studien sind in die Grafik integriert (graue Balken). Die schwarzen Balken zeigen die Ergebnisse der eigenen Korpusuntersuchung (Bonner Fnhd.- und GerManC-Korpus). Deutlich zu erkennen ist die Produktivitätszunahme vom 16. bis 18. Jh. Die anschließende Stagnation kann nicht als gesichert gelten, da Vergleichswerte aus dem 19. bis 21. Jh. noch fehlen. Allerdings ist ein gewisser Rückgang nicht unwahrscheinlich, da er sich schon im letzten Untersuchungszeitraum (1750–1800) andeutet und mit dem Rückgang der Ableitung aus Städte- (*Maynzisch*) und Personennamen (*Brömsisch* < *Brömsen*, vgl. Kempf 2017) sowie dem Rückgang der gräko-romanischen Entlehnungen (s.u.) in Verbindung gebracht werden kann.

6 Types, Hapax Legomena, Neubelege (jeweils relativ zu den Werten aller suffigierten Adjektive jedes Zeitschnitts), sowie die Produktivität im engeren Sinne, d.h. Hapax Legomena/Tokens (vgl. Baayen 2009: 906).

7 Den Anteil von *-isch* an allen suffigierten Types kann man leicht den vorliegenden Studien entnehmen, die jeweils das Derivationssystem eines Zeitschnitts untersuchen. Für Hapaxe oder gar Neubelege ist dies i.d.R. nicht möglich. Anders ist die Situation auch im Hinblick auf das Englische, wo umfassende empirische Studien für verschiedene Zeiträume noch fehlen (vgl. Trips 2014: 385–386).

Abb. 1: *-isch*-Derivate in% der suffigierten Adjektive (Types), 1200–1978⁸.

8 Quellen sowie absolute Zahlen, graue Balken: 1200–1300: Ganslmayer (2012: 35), 26 (587) Types; 1488–1548: Thomas (2002: 99–485), 67 (561) Types (neu berechnet, vgl. Kempf 2016a: 102–104); 1600–1700: Schulz (2002, 2007: 193), 356 (1526) Types; 1900–1978: Kühnhold, Putzer und Wellmann (1978: 100–117, 266–421, 444), 1387 (7189) Types; schwarze Balken: Kempf (2016a: 322f), 1350–1400: 15 (335) Types; 1450–1500: 24 (407) Types; 1550–1660: 84 (517) Types; 1650–1700: 140 (689) Types bzw. 175 (726) Types; 1700–1750: 208 (794) Types; 1750–1800: 218 (904) Types. Man beachte, dass die X-Achse nicht strikt linear verläuft, sondern nur die untersuchten Zeiträume chronologisch anordnet. Zu Details s. Kempf (2016a: 249).

Ein/e anonyme/r GutachterIn fragt zu Recht nach der Vergleichbarkeit der Korpora. Hierzu sei Folgendes angemerkt: Alle Korpora enthalten die vergleichbaren Textsorten Urkunden, Berichte, Chroniken oder Zeitungen. Darüber hinaus sind in allen Korpora außer Ganslmayer (2012) weitere Textsorten wie Predigten, wissenschaftliche Sachtexte oder narrative Prosa zu variierenden Anteilen enthalten. Die Vergleichbarkeit ist somit nicht vollkommen (was insbesondere die Differenzen im 17. Jh. erklären kann, s. Kempf 2016a: 249), aber doch akzeptabel gemessen an dem für die deutsche Sprachgeschichte Möglichen. Der Eindruck der (bei reflektierter Interpretation gegebenen) Vergleichbarkeit verstärkt sich dadurch, dass auch für die anderen hochfrequenten Suffixe (*-lich* und *-ig*) eine ähnlich gute Passung der benachbarten Balken auftritt wie bei *-isch* (s. Kempf 2016a: 128, 208). Zum Zusammenhang zwischen Textsorte und Derivation s. ausführlich Kempf (im Erscheinen).

Um den Produktivitätsverlauf der deutschen *-isch*-Derivation zu verstehen, ist ein Blick auf die Entwicklung der beteiligten Basiskategorien aufschlussreich. Daher werden im Folgenden zunächst die Derivationsbasen untersucht, und zwar im Hinblick auf ihre Wortarten, ihren Status als Eigennamen und/oder Fremdwörter sowie ihre Komplexität. Anschließend wird die Entwicklung der Wortbildungsfunktionen dargestellt.

Was die Wortart der Derivationsbasen betrifft, so lässt sich durch die gesamte deutsche Sprachgeschichte eine starke Dominanz substantivischer Basen feststellen. Beide mhd. Korpusstudien verzeichnen einen fast ausschließlichen Anteil substantivischer Basen (bei Ganslmayer 2012: 494 sind alle 26 Types desubstantivisch, bei Klein, Solms und Wegera 2009: 304 sind es 59 von 60 Types). Für das Nürnberger Fnhd. ermittelt Thomas (2002: 322) einen desubstantivischen Anteil von ca. 87% (aus 67 Types). Im Nhd. schließlich liegt der Wert für substantivische Basen bei über 94% (neben ca. 5% verbalen und 0,36% adjektivischen Basen Kühnhold, Putzer und Wellmann 1978: 110–111). Auf etwas andere Werte kommt Schlaefter (1977: 187), der nur 73%, i.e. 320 seiner 439 Types als desubstantivisch und weitere 4,8% als deadjektivisch analysiert. In seiner Studie wird besonders deutlich, wie schwierig die Basiswortart zu bestimmen ist, denn einige Bildungen sind doppelt motiviert (z. B. *grüblerisch* zu *Grübler* oder mit Suffixvariante *-erisch* zu *grübeln*), und sehr viele sind aus einem gebundenen und/oder komplexen fremdsprachigen Stamm abgeleitet (z. B. *sem-ant-isch*, *lexik-al-isch*, 1977: 47–48).

Aus diesen Beobachtungen lässt sich schon erahnen: Anders als im Englischen steht der Produktivitätszuwachs im Deutschen nicht mit einer starken Expansion der Basiswortarten in Verbindung. Während Verben und Adjektive nur einen bescheidenen Zuwachs als Basiskategorien erfahren, spielen Adverbien und Numeralien in der deutschen *-isch*-Derivation überhaupt keine Rolle. Allerdings konnte sich die deutsche *-isch*-Derivation zwei einträgliche Quellen neuer Basislexeme erschließen – nämlich Fremdwörter und Eigennamen (*apostolisch*; *platonisch*; vgl. Wilmanns ²1899: 475, Schlaefter 1977: 198, Fleischer ⁵1982: 264, Kühnhold 1985: 1614, Klein 1994: 383, 394, Fleischer & Barz ⁴2012: 340, Kempf 2017).

Die Affinität zu onymischen Basen – die ja u.a. auch im (Alt)englischen auftritt – besteht schon mit der ursprünglichen (gemein-germanischen) Funktion, Adjektive der Zugehörigkeit und Herkunft zu bilden (vgl. Wilmanns ²1899: 471). Sie setzt sich im Ahd. fort mit Derivaten wie *arabisc*, *crèhhisc*, *sirisc*, *spanisc* (Goetze 1899: 471, Hornbruch 1996: 127–238). Eigennamen machen dort schon ca. 51% (Klein 1994) bzw. 54% (Decke 1993) der Basen aus.⁹

9 Decke (1993) und Klein (1994) analysieren Daten aus ahd. Wörterbüchern. Im Einzelnen nennen sie folgende Werte: Decke (1993): fremdes Onym 42; fremd

Klein (1994: 382) stellt fest, dass die *-isch*-Ableitung aus onymischen Basen im Deutschen „seit ahd. Zeit kontinuierlich produktiv geblieben“ ist, und spätere Untersuchungen geben ihm recht: In Ganslmayers (2012: 494–495) mhd. Korpusdaten sind 22 von 26 Types aus Eigennamen abgeleitet. Auch in der mhd. Wortbildung von Klein, Solms und Wegera (2009) sowie im fnhd. Korpus von Thomas (2002: 286, 301, 313) treten Eigennamen als Basis hervor, wengleich etwas schwächer. Die vorliegende Korpusuntersuchung erlaubt, den Anteil der Onyme und Fremdwörter über einen längeren Zeitraum zu quantifizieren. Tabellen 5 und 6 zeigen die absoluten und prozentualen Werte der *-isch*-Types jedes Zeitschnitts, Tabellen 7 und 8 das Entsprechende für die neubelegten Types jedes Zeitschnitts.¹⁰

Tab. 5: *Basistypen aller -isch-Types 1350–1800, absolut.*¹¹

Types absolut	1350–1400	1450–1500	1550–1600	1650–1700	1650–1700*	1700–1750	1750–1800	gesamt
fremd, Onym	3	9	27	50	59	80	45	172
fremd, nicht-Onym	0	1	11	29	37	39	91	136
nativ, Onym	5	5	22	35	51	61	50	128
nativ, nicht-Onym	7	9	23	23	25	28	32	74
Summe	15	24	83	137	172	208	218	510

Tab. 6: *Basistypen aller -isch-Types 1350–1800, prozentual.*

Types in%	1350–1400	1450–1500	1550–1600	1650–1700	1650–1700*	1700–1750	1750–1800	gesamt
fremd, Onym	20%	38%	33%	36%	34%	38%	21%	34%
fremd, nicht-Onym	0%	4%	13%	21%	22%	19%	42%	27%
nativ, Onym	33%	21%	27%	26%	30%	29%	23%	25%

nicht-Onym 9; natives Onym 13; nativ nicht-Onym 37; Summe 101. Klein (1994): Onym 43; fremd nicht-Onym 9; nativ nicht-Onym 33; Summe 85.

10 Als neubelegt wurden für jede Epoche diejenigen Types gezählt, die innerhalb des Untersuchungskorpus in einem gegebenen Zeitraum erstmalig auftreten (so dass im ersten Zeitschnitt alle Types als „neubelegt“ gelten). In einem begrenzten Korpus sind Neubelege selbstverständlich nicht mit Neubildungen gleichzusetzen. Allerdings lässt sich die Wortbildung eines Zeitraums auf Grundlage der Neubelege besser einschätzen als anhand aller Types, da dann das „Rauschen“ durch alte, tradierte Bildungen zumindest reduziert ist. Zu einer Methode, die mit exakterer Datierung einzelner Derivate arbeitet s. Kempf und Hartmann (im Erscheinen).

11 Als Beispiele für die vier Typen seien folgende Derivate des letzten Zeitschnitts genannt: fremd, Onym: *Crusiusisch*; fremd, nicht-Onym: *atmosphärisch*; nativ, Onym: *Tannenbergisches*; nativ, nicht-Onym: *rächerisch*.

Types in%	1350–1400	1450–1500	1550–1600	1650–1700	1650–1700*	1700–1750	1750–1800	gesamt
nativ, nicht-Onym	47%	38%	28%	17%	15%	13%	15%	15%
Summe	100%	100%	100%	100%	100%	100%	100%	100%
Summe fremd	20%	42%	46%	58%	56%	57%	62%	60%
Summe Onym	53%	58%	59%	62%	64%	68%	44%	59%

Tab. 7: Basistypen neubelegter -isch-Types 1350–1800, absolut.

Neubelege absolut	1350–1400	1450–1500	1550–1600	1650–1700	1650–1700*	1700–1750	1750–1800	gesamt
fremd, Onym	3	6	24	40	35	43	21	172
fremd, nicht-Onym	0	1	11	24	21	21	58	136
nativ, Onym	5	1	16	21	33	31	21	128
nativ, nicht-Onym	7	5	18	8	12	11	13	74
Summe	15	13	69	93	101	106	113	510

Tab. 8: Basistypen neubelegter -isch-Types 1350–1800, prozentual.

Neubelege in%	1350–1400	1450–1500	1550–1600	1650–1700	1650–1700*	1700–1750	1750–1800	gesamt
fremd, Onym	20%	46%	35%	43%	35%	41%	19%	34%
fremd, nicht-Onym	0%	8%	16%	26%	21%	20%	51%	27%
nativ, Onym	33%	8%	23%	23%	33%	29%	19%	25%
nativ, nicht-Onym	47%	38%	26%	9%	12%	10%	12%	15%
Summe	100%	100%	100%	100%	100%	100%	100%	100%
Summe fremd	20%	54%	51%	69%	55%	60%	70%	60%
Summe Onym	53%	54%	58%	66%	67%	70%	37%	59%

An den Werten lässt sich erkennen, dass sowohl Eigennamen als auch Fremdwörter langfristig eine tragende Rolle in der deutschen -isch-Derivation gespielt haben. Besonders aus den Werten der Neubelege wird die zunehmende Bedeutung beider Basistypen ersichtlich: Eigennamen haben sich von anfangs schon 53% auf später 70% gesteigert und sacken lediglich im letzten Zeitraum auf 37% ab; dies lässt sich, wie gesagt, mit dem Rückgang der Ableitung aus Städte- und Familiennamen in Verbindung bringen (vgl. Kempf 2017). Die DWDS-Stichprobe des 20. Jh. scheint diesen neueren Abwärtstrend zu bestätigen: Hier entfallen von 100 Hapax legomena nur noch 25 auf deonymische Bildungen (davon 22 fremd, z. B. *kuwaitisch*, 3 nativ, z. B. *sorbisch*). Schlaefer (1977: 186–187) ermittelt für das mittlere 20. Jh. nur

noch 27% Deonymika unter seinen Types. Eigennamen haben damit also die Führungsrolle an Fremdwörter abgegeben. Im Untersuchungskorpus haben sich Fremdwörter von anfangs 20% auf 70% der Neubelege ‚hochgearbeitet‘; unter den 100 Hapaxen des 20. Jh. finden sich sogar 93 fremdstämmige (z. B. *anthroposophisch*).¹²

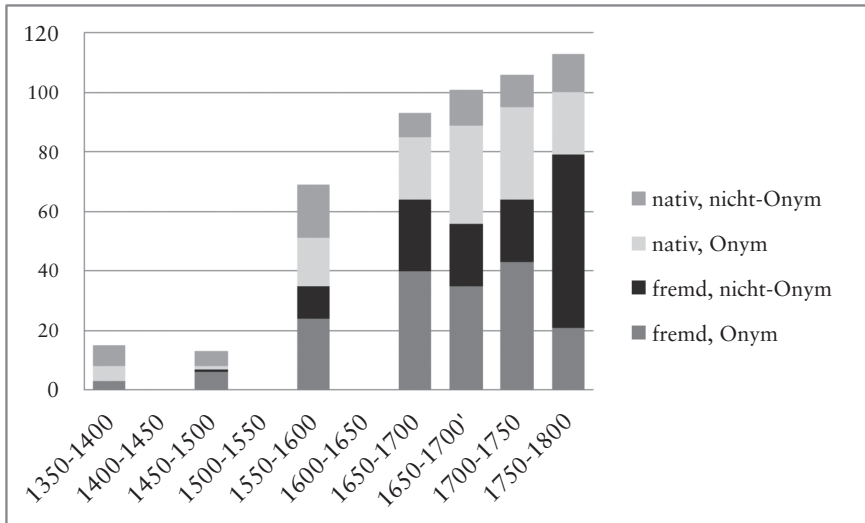
Wie lassen sich die starke Kopplung zwischen *-isch* und Fremdbasen und der große Erfolg dieses Musters erklären? Klein (1994: 383) nimmt an, dass die Ableitung aus Fremdwörtern durch das deonymische Muster „hervorgerufen oder wenigstens gefördert worden sein [dürfte], da die herkunftsbezeichnenden [ahd.] *-isk*-Adjektive großenteils von fremdsprachlichen Eigennamen abgeleitet waren.“ Tatsächlich ist es sehr plausibel, dass ein solcher Übergang vom deonymischen zum fremdbasierten Muster schon im Ahd. angelegt war: Die zahlreichen fremden Onyme hätten gute Vorlagen für analogische Nachbildungen auf der Basis formaler Ähnlichkeit geliefert (*platonisch* – *kanonisch*, beide 16. Jh.). Auch hätten deonymisierte Derivate den Übergang gefördert, d.h. Derivate, die nicht mehr primär auf den Namenträger referieren, sondern auf ein durch ihn bezeichnetes Konzept (z. B. ahd. *dorisch* ‚dorisch [singen]‘, nicht *‚den Dorern zugehörig‘, Decke 1993: 145; *salomonisch* ‚weise‘, 16. Jh.). Mit solchen Derivaten wird die Wortbildungsfunktion von ‚X zugehörig‘ zu ‚in der Art von X‘, ‚wie X‘ – d.h. zu ‚vergleichenden‘ Bildungen – ausgeweitet. Auch dies würde Ableitungen aus Appellativen begünstigen, da hier oft andere Wortbildungsfunktionen als ‚Zugehörigkeit‘ intendiert sind, z. B. *poetisch* ‚dichterisch‘ (i.e. ‚wie ein Dichter, auf Art und Weise eines Dichters‘, 15. Jh.).

Allerdings ist die Kategorie der fremden nicht-onymischen Stämme im Ahd. noch nicht sehr stark etabliert (nur 9 von 101 Types bei Decke 1993). Auch im eigenen Untersuchungskorpus setzt die Zunahme dieser Kategorie erst relativ spät ein (noch 0 Types/1 Type in den ersten zwei Zeiträumen, s.o.). Diese zunächst verzögerte und dann starke Expansion zu nicht-onymischen Fremdbasen ist in engem Zusammenhang zu sehen mit dem Sprachkontakt, der sich mit der humanistisch-lateinischen Entlehnungswelle im 16. Jh. und der französischen im 17. und 18. Jh. intensiviert. Entlehnungen spielen nicht nur als Basen eine wichtige Rolle (*Gravität* > *gravitatisch*, *Satan* > *sataniſch*, *Exemplar* > *exemplariſch*, 16. Jh., Decke 1993: 111), sondern auch in Form von Adjektiven, deren lateinische Endung durch *-isch* ersetzt wird, z. B. „*heroisch*, *infernaliſch*, *katholiſch* im 16. Jh. < lat. *heroicus*, *infernalis*, *catholicus*“

12 Der Einteilung fremd–nativ liegt ein strukturell-perzeptueller Fremdheitsbegriff (vgl. Kempf 2010 u.a.), kein historisch-etymologisierender zugrunde. Als fremd gilt demnach, was z. B. fremde Segmente oder Phonotaktik, nicht-initialen oder beweglichen Akzent, Mehrfüßigkeit, Vollvokale in Nebensilben etc. aufweist.

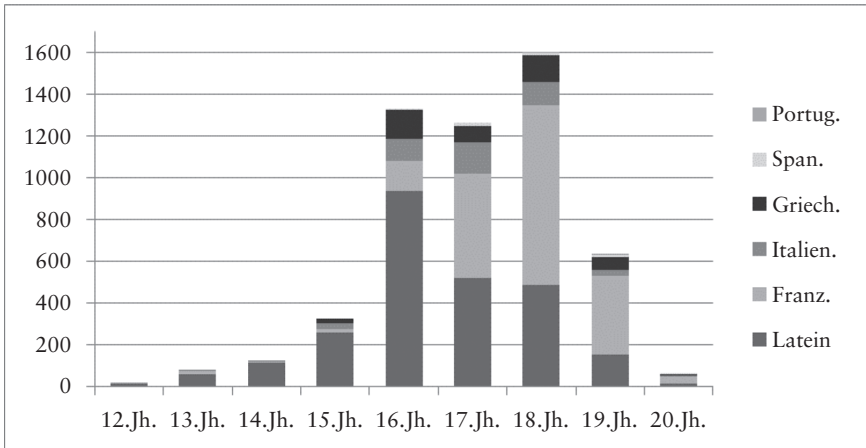
(Decke 1993: 111).¹³ Der Zusammenhang zwischen Entlehnungswellen und -isch-Derivation lässt sich gut nachvollziehen, wenn man den absoluten Verlauf der Neubelege mit dem der Entlehnungen vergleicht – s. Abbildung 2, die die Werte aus Tabelle 7 visualisiert, und Abbildung 3, in der der Verlauf der gräko-romanischen Entlehnungen ins Deutsche dargestellt ist. Abbildung 2 lässt deutliche Sprünge der fremdbasierten -isch-Ableitung im 16. und 17. Jh. sowie einen weiteren Zuwachs im 18. Jh. erkennen. Ähnlich sieht der in Abbildung 3 gezeigte Befund für gräko-romanische Entlehnungen aus, der das 16., 17. und 18. Jh. als Hochphasen ausweist. Der Zusammenhang zwischen Entlehnung und -isch-Produktivität ist umso bemerkenswerter, als andere Muster nicht oder nur in deutlich geringerem Umfang von dieser Ressource profitieren und entsprechend keinen parallelen Produktivitätsschub erfahren. So weisen z. B. unter den -lich-Derivaten nur 0,01% im ersten Untersuchungsabschnitt (1350–1400) und 0,02% im letzten (1750–1800) fremde Basislexeme auf.

Abb. 2: Basistypen neubelegter -isch-Derivate.



13 Vgl. auch Wellmann (1997: 71) zu *romantisch*: „Derivat zu *Roman*, französisch *roman*, wobei – wie das -t- verrät – französisch *romantesque* als direktes Vorbild gedient hat.“

Abb. 3: Gräko-romanische Entlehnungen ins Deutsche nach Best (2001: 11–14).



Die durch die Entlehnungswellen gestärkte Tendenz zu fremden Basiswörtern scheint sich zu einem stabilen Subschema der *-isch*-Derivation verfestigt zu haben, denn auch nach Abklingen der gräko-romanischen Entlehnungen bleibt die Affinität zu Fremdbasen bestehen. In Schlaefers (1977: 186) Korpus des mittleren 20. Jh. basieren 245 von 439 Types (i.e. 55%) auf Fremdwörtern (ähnlich Kühnhold, Putzer und Wellmann 1978: 292, 300–301, doch ohne konkrete Zahlen), in der DWDS-Stichprobe sind es 93 von 100 Hapaxen (davon 71 nicht-deonymische, z. B. *hymnisch*). In der Sammlung rezenter Neubildungen (s.u. (7)–(11)) basieren 12 von 18 Types auf fremdsprachlichem Material (z. B. *kreationistisch*).

Diese Affinität der deutschen *-isch*-Derivation zu fremdsprachlichen Basen kann teilweise im Sinne der von Lindsay und Aronoff (2013) beschriebenen funktionalen Nischen begriffen werden: Bei ganz oder teilweise synonymen Mustern kann es zum Überleben des schwächeren Musters in funktionalen (z. B. phonologisch oder morphologisch definierten) Nischen kommen. An partiell synonymen Konkurrenz wäre hier vor allem an die *-lich*-Derivation zu denken (Klein 1994, Kempf 2016b u.a.) – wobei allerdings nicht klar ist, ob *-isch* tatsächlich als schwächeres Muster gelten kann. Vielmehr könnten sich beide (etwa gleichstarken) Muster komplementäre Nischen erschlossen haben. Wichtig ist hierbei nach Lindsay und Aronoff (2013), dass eine entsprechende Subdomäne für die Sprecher klar erkennbar und vorhersagbar sein muss, um als Nische für das betreffende Muster zu funktionieren. Dieses Kriterium der Vorhersagbarkeit trifft für den deutschen Fremdwortschatz zu, da er sich verhältnismäßig scharf vom nativen abgrenzt (anders als im Englischen, das bekanntlich über einen stark hybridisierten germanisch-romanischen

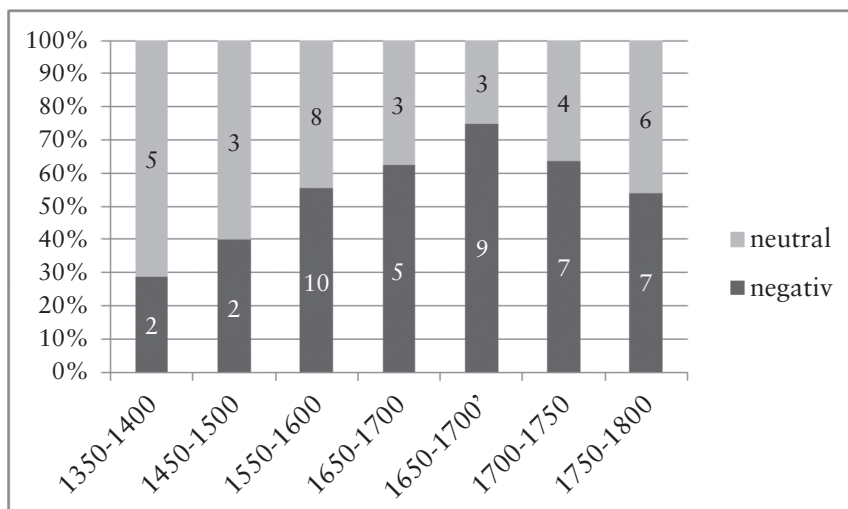
Wortschatz verfügt¹⁴). Im Deutschen heben sich insbesondere die für *-isch* relevanten gräko-romanischen Fremdexeme durch Merkmale wie fremde Phonotaktik (*mnemotechnisch*), nicht-initialem oder beweglichem Akzent (*Ásthma – asthmätisch*) und vollvokalige Nebensilben hervor. Darüber hinaus sind viele der Fremdbasen durch charakteristische Endungen gekennzeichnet. Auch wenn diese im Deutschen nicht zwingend transparent sind, tragen sie dennoch aufgrund ihres reihenweisen Vorkommens zu charakteristischen Nischen bei (allen voran *-ist-isch*, gefolgt von *-at-iv-isch*, s. Schlaefer 1977: 184–186; s. außerdem Lindsay und Aronoff zu ähnlichen Phänomenen im Englischen (*-olog-ic-al*)).

Die Kategorie der nativen nicht-onymischen Basen, auf die im Ahd. noch 37% der Types entfallen (Decke 1993) ist unter den Neubelegten des Untersuchungszeitraums am geringsten vertreten und teilt auch nicht die steigende Tendenz der anderen drei Kategorien – im Gegenteil, sie sackt von anfangs 47% auf rund 10% der Neubelege ab (vgl. Tabelle 8). Hier herrscht die stärkste Konkurrenz durch andere Suffigierungsmuster, vor allem die *-lich*-Suffigierung. Im Ahd. und Mhd. bestehen noch viele synonyme Dubletten, z. B. *esellich – eselisch, vihelich – vihisch, wíplich – wíbisich, hêrlich – hêrisch, tôrlich – tôrisch, tiufellich – tiuwelisch, geloubic – gelöübisich* (Klein, Solms und Wegera 2009: 310, Lexer 1872–1878, Pfeifer 1993, Goetze 1899: 492, vgl. auch Klein 1994: 395). In der Konkurrenz um native nicht-onymische Basen zeigt *-isch* den Ansatz, eine pragmatisch definierte Subdomäne als Nische zu besetzen – nämlich die der negativ konnotierten Basen und Kontexte (*verbrecherisch, mörderisch*; Goetze 1899: 464–514, Fleischer & Barz 2012: 315). Seit dem Fnhd. lässt sich beobachten, dass bei Dubletten zu negativ konnotierten Basen die *-isch*-Variante gewinnt, z. B. *teuflisch* gegen *teufellich*, *höhnisch* gegen *höhnlich*, *neidisch* gegen *neidig* (alle vor 1600, s. Kempf 2016: 273–288). *-isch*-Bildungen des späten Fnhd. sind zunehmend aus negativ konnotierten Basen abgeleitet (z. B. *verführerisch, tyrannisch, ketzerisch*, im Korpus ab ca. 1550) oder stehen in pejorativen Kontexten, vgl. *der alte Pábstische grewel* ‚Greuel-taten des Papsts‘ (Mitte 16. Jh., Thomas 2002: 307). Die Übergänge zwischen negativ konnotierter Basis und pejorativem Kontext sind fließend. Daher sind in Abbildung 4 alle entsprechenden Fälle unter „negativ“ zusammengefasst. Die Gruppe „neutral“ umfasst Derivate wie *einheimisch, oberländisch, fürstentumsch, südisch, kupferisch*. Berücksichtigt werden nur die Ableitungen aus nativen nicht-onymischen Basen, da die Affinität zu negativen Basen bzw. Kontexten beim übrigen Material kaum von Bedeutung ist. Durch diesen Filter sind die absoluten Werte sehr niedrig, doch es zeigt sich immerhin ein bis 1700

14 Vgl. Kap. 3.2, das diese Abgrenzungstendenz in Bergs (2014) typologisches Konzept der *Boundary Permeability* einbettet.

konstant steigender Anteil der „Negativ-Affinität“, die sich ja auch im o.g. Dubletten-Abbau widerspiegelt.

Abb. 4: Anteil negativ konnotierter Derivate unter den *-isch*-Ableitungen aus nativen nicht-onymischen Lexemen (neubelegte Typen).



Die Frage, wie es zur Affinität zwischen *-isch* und negativen Basen/Kontexten kam, liegt noch teilweise im Dunkeln (s. aber Goetze 1899 und Kempf 2016a: 273–288). Denkbar wäre Folgendes: In der seit dem Mhd. bestehenden Konkurrenzsituation konnte womöglich schon ein kleiner Faktor genügen, um das – in diesem Kontext unterlegene – *-isch*-Muster in eine Nische zu drängen. Als ein solcher Faktor käme etwa der immer wieder auftretende lexikalische Wandel einzelner *-isch*-Derivate in Frage, wie wir ihn schon oben bei verschiedenen altgermanischen Sprachen gesehen haben, bei dem aus der Funktion ‚Zugehörigkeit‘ via Stereotypisierung eine ‚Eigenschaft‘ wird, z. B. an *heimskr* ‚einfältig‘ (Henzen 1965: 200), ahd. *dorfisc* u.a. ‚dumm‘ (Wilmanns 1899: 472), *flämisch* u.a. ‚grob, plump‘ (Goetze 1899: 471). Ein weiterer möglicher Faktor liegt in der Polemik der Reformationszeit. Hier werden zum einen (mit *-isch* bezeichnete) Gruppen- oder Personenzugehörigkeiten als negativ inszeniert (vgl. *ketzerisch*, *päpstisch*, *teuflisch*, *abergläubisch* (im Korpus ab ca. 1550), *antichristisch*, *abgöttisch*, *papistisch* (im Korpus ab ca. 1650)). Zum anderen werden ursprünglich neutrale Tieradjektive zur Diffamierung durch Vergleich eingesetzt, vgl. *D. Fabers vnd dergleichen lügenhafttge, lästerliche, eselische schrift* (Luther nach Goetze 1899:495), *äffisch* ‚nährisch‘ bei Luther, Fischart, Alberus, 16. Jh. (DWB, Goetze 1899:495). Außerdem treten in den Korpusbelegen ab dieser Zeit zunehmend Ableitungen des folgenden Typs auf:

- (3) *kriegerisch* (15. Jh.)
- (4) *verführerisch, bettlerisch, ketzerisch* (ab dem 16. Jh.)
- (5) *räuberisch, verräterisch, mörderisch, einbilderisch, verleumderisch, buhlerisch* u.a. (ab dem 17. Jh.)
- (6) *verschwenderisch, betrügerisch, rächerisch, knauserisch* u.a. (18. Jh.).

Die *-erisch*-Bildungen scheinen sich diachron sukzessive von entsprechenden *-er*-Nomina Agentis (*Krieger, Bettler, Mörder*) zu lösen, so dass die Derivate eher als deverbale ‚aktivische‘ Ableitungen (*sich einbilden, knausern*) aufzufassen sind. Für gegenwartssprachliche Bildungen stellen Fleischer und Barz (⁴2012: 339–340) einen (semantischen) Bezug auf Verben oder Verbalabstrakta fest, die sich im Übrigen nicht mehr auf negativ konnotierte Begriffe beschränken (*feldforscherisch, gestalterisch* neben *selbstzerstörerisch, kriecherisch*). Diese Entwicklung kann als formaler Verstärkungsprozess gedeutet werden, mit dem *-isch* (dann als Variante *-erisch*) sich in Ansätzen die verbalen Basen als neue Ressource erschließt.

Anhand der Sammlung rezenter Neubildungen kann man erahnen, auf welche Weise sich *-isch* trotz Rückgang der relevanten Entlehnungen und trotz partiellen Verlusts der onymischen Basen seine Produktivität (auf einem gewissen Level) erhält. Bei der Derivation aus nativen Basen ist das *-erisch*-Schema deutlich vertreten, die pejorative Nische höchstens bei *schuldenmacherisch* und *langschläferisch*, je nach Kontext.

- (7) Derivation aus nativer Basis: *schuldenmacherisch, langschläferisch, fledermäusisch, funkisch, boxerisch, bläserisch*
- (8) Lehnübertragung: *retrostylich*
- (9) (fremde) Kontamination: *demokraturisch* (zu *Demokratie/Diktatur*)
- (10) Analogie (Fremdname): *fidelistisch* (zu *sozialistisch*)
- (11) Derivation aus fremder Basis: *aurisch, hypodisch, spagyrisch, anamorphotisch, kreationistisch, fallibilistisch*; dabei deonymisch: *pugliesisch* (zu *Puglia*); dabei kombinierende Ableitung:¹⁵ *präwikipedianisch, postkatastrophisch*

Im Bereich der stark vertretenen Fremdwortbildung funktioniert das Anhalten der Produktivität wohl über verschiedene Wege, z. B. über erst jüngst frequenter gewordene Basen (*Aura*), z.T. Fachtermini (*Spagyrik*). Außerdem zeigt sich insgesamt eine starke Tendenz zur komplexen Wortbildung: Die Mehrheit der Bildungen weist eine komplexe Basis oder eine komplexe Affixkonstellation auf (*-ist-isch; prä-/post-X-isch*). Dass die Komplexität der *-isch*-Derivation diachron sukzessive angestiegen ist, zeigen die in Tabelle 9 zusammengestellten Ergebnisse. Damit weist die deutsche *-isch*-Derivation

15 Zum Konzept der kombinierenden Derivation s. Kempf und Hartmann (im Erscheinen).

eine partielle Parallele zum englischen Pendant auf, wobei aber dort noch weitaus komplexere Basen möglich sind (z. B. *light-at-the-end-of-the-tunnel-ish*, s.u.), während die deutsche *-isch*-Derivation nur selten (und nur einfache) Phrasen ableitet (z. B. *vorexilisch* < *vor dem Exil*).

Tab. 9: *Simplizische vs. komplexe -isch-Derivation diachron.*

Fnhd.:	
um 1500, Daten aus Thomas (2002: 286–324), eigene Analyse:	
56 Simplizia (85%) Bsp. <i>päpstisch, närrisch, irdisch</i>	10 komplexe Basen (15%) Bsp. <i>abgöttisch, widersinnisch</i>
Nhd.:	
Types 1750–1800, eigene Korpusanalyse:	
154 Simplizia (70%) Bsp.: <i>arabisch, lydisch, Wolfisch, höhnisch, knechtisch, magellanisch, ätherisch</i>	65 komplexe Basen (30%) Bsp. <i>metaphysisch, philosophisch, topographisch, Churpfälzisch, kraftmännisch, Holländisch, ausländisch, heimtückisch</i>
Schlaefer (1977: 184), Mitte 20. Jh.:	
247 Simplizia (56%) Bsp.: <i>schulisch, kanadisch</i>	192 komplexe Basen (44%) Bsp.: <i>künstlerisch, lexikalisch, sauertöpfisch</i>
Zufallssample 100 Types aus DWDS, 20. Jh.:	
48 Simplizia Bsp.: <i>drastisch, modisch, thrakisch, diebisch, sächsisch, stoisch, Moldauisch, kubisch, honduranisch, schlesisch</i>	52 komplexe Basen Bsp.: <i>pyrotechnisch, logopädisch, misstrauisch, westeuropäisch, mittelständisch</i>
Zufallssample 100 Hapaxe aus DWDS, 20. Jh.:	
36 Simplizia Bsp.: <i>balsamisch, botanisch, iranisch, sorbisch, ständisch</i>	64 komplexe Basen (oder komplexe Derivation) Bsp.: <i>anthroposophisch, bioklimatisch, präsynaptisch, vorexilisch</i>

Die Komplexitätszunahme ist einerseits sehr deutlich, andererseits kein Alleinstellungsmerkmal der *-isch*-Derivation. Sie tritt auch bei der *-ig*-Derivation auf (Kempf 2016a: 247) und – dies ist für *-isch* gravierender – beim Konkurrenten *-lich*: Hier sind bereits 1750–1800 fast zwei Drittel der Neubelege komplex, und dieser Trend setzt sich bei rezenten Neubildungen fort (z. B. *erstinstanzlich, nachbarschaftsstaatlich, globaldörflich*, Kempf 2016a: 197–199). Konkurrenz besteht nicht nur in der für *-isch* zentralen Zugehörigkeits- und Herkunftsfunktion (ahd. *himilisc* – *himillih*, *kuningisc* – *kuninglih*, Klein 1994: 395), sondern auch u.a. in der daraus hervorgegangenen

,vergleichenden‘ Funktion (s. z. B. die ursprünglich synonymen Bildungen *kindisch* – *kindlich*). Dieses Konkurrenzgeflecht umfassend darzustellen würde zu weit führen, s. aber Kempf (2016a) und (2016b). Stattdessen sei die funktionale Entwicklung der *-isch*-Derivation grob skizziert, mit dem Fokus auf den kontrastiv relevanten Aspekten.

In der Entwicklung des deutschen *-isch*-Suffixes lässt sich eine zunehmende funktionale Auffächerung gegenüber der ursprünglich klaren Funktion der Zugehörigkeits- und Herkunftsbezeichnung beobachten:¹⁶ Im Mhd. liegt der Anteil dieser beiden Funktionsklassen zusammengenommen noch zwischen ca. 65% und 79% der *-isch*-Derivate, während er im Nürnberger Fnhd. um 1500 auf ca. 46% abgesunken ist. Im Nhd. des 20. Jh. erreichen Zugehörigkeit und Herkunft zusammen nur noch rund 36% der Types.¹⁷ Immerhin sind dabei die Adjektive der Zugehörigkeit (mit gut 27%, z. B. *hanseatische (Froh-natur)*, Kühnhold, Putzer und Wellmann 1978: 262) noch die mit Abstand frequenteste Gruppe innerhalb der *-isch*-Derivation. Die Verbindung zu den germanischen Ursprüngen ist also noch erkennbar (viel deutlicher als beim englischen *-ish*, das zwar noch nationalitätsbezeichnende Derivate tradiert (*English, Swedish*), dieses Muster aber nicht mehr aktiv nutzt). Trotzdem ist die nhd. *-isch*-Derivation als semantisch diffus zu bezeichnen. Sie umfasst in Kühnhold, Putzer und Wellmanns (1978: 110–111) Studie des Nhd. ganze dreizehn Funktionsklassen, dagegen in den Analysen des Mhd. (Klein, Solms und Wegera 2009: 303, Ganslmayers 2012: 497) nur fünf bzw. sechs – trotz vergleichbarer Funktionsklasseneinteilung. Tabelle 10 zeigt nur die vier frequentesten Funktionsklassen der (f)nhd. *-isch*-Derivation.

Tab. 10: Die frequentesten Funktionsklassen der (f)nhd. *-isch*-Derivation¹⁸.

Funktion	Bsp. fnhd.	Fnhd. semas.	Nhd. semas.	Fnhd. abs.	Nhd. abs.	Fnhd. onomas.	Nhd. onomas.
Zugehörigkeit	<i>hellisch</i> ‚der Hölle zugehörig‘	38,69%	27,39%	33	376	61,90%	62,56%
vergleichend	<i>pewrisch</i> ‚wie ein Bauer‘	11,90%	18,43%	10	253	45,80%	24,66%

16 Vgl. Ganslmayer (2012: 509–511) sowie Kempf (2016a: 253–257).

17 Die Prozentangaben entstammen Klein, Solms und Wegera (2009: 303) sowie Ganslmayer (2012: 497) für das Mhd.; Thomas (2002: 313–324) für das Nürnberger Fnhd. (Prozentsätze neu berechnet, vgl. Kempf 2016a: 102–104); Kühnhold, Putzer und Wellmann (1978: 259–419) für das Nhd.

18 „semas[iologisch]“ = Anteile der Funktionsklassen an der *-isch*-Derivation, „onomas[iologisch]“ = Anteil der *-isch*-Derivation an den Funktionsklassen. Die Werte sind errechnet nach Thomas (2002: 286–319) sowie Kühnhold, Putzer und Wellmann (1978: 259–419); für eine vollständige Auflistung der Funktionsklassen s. Kempf (2016a: 255).

Funktion	Bsp. fnhd.	Fnhd.	Nhd.	Fnhd.	Nhd.	Fnhd.	Nhd.
		semas.	semas.	abs.	abs.	onomas.	onomas.
determinierend/ referentiell	<i>hüntisch</i> (<i>geschlecht</i>) ,Hunderasse‘	8,33%	12,16%	7	167	35,90%	37,78%
Herkunft	<i>mosisch</i> ‚von Moses ausgehend‘	7,74%	9,10%	7	125	41,94%	40,72%

Nach den Zugehörigkeitsbezeichnungen steht die ‚vergleichende‘ Funktion an zweiter Stelle (mit knapp 12% gut 18% der fnhd./nhd. *-isch*-Types), die vermutlich über die Zuschreibung typischer Eigenschaften aus der Zugehörigkeitsfunktion hervorgegangen ist (vgl. z. B. mhd. *hellisch düsternisse* (Klein, Solms und Wegera 2009: 309) ‚zur Hölle gehörige Dunkelheit‘ oder ‚Dunkelheit wie in der Hölle‘). Beim englischen *-ish* ist diese Funktion viel stärker ausgeprägt als beim deutschen Pendant und kann als Ausgangspunkt zur Entwicklung einer ‚approximativen‘ Funktion gesehen werden (z. B. *greenish* ‚(ähnlich) wie grün‘ > ‚ein bisschen grün/grünlich‘). Diese kann das deutsche *-isch* nicht realisieren, da die Nische seit dem frühen Nhd. durch die *-lich*-Derivation eingenommen wird (z. B. *grünlich*, *ältlich*, *dicklich*, auch rezente Okkasionalismen wie *komischlich*, *kaputtlich*, Kempf 2016a: 198).

Die dritthäufigste Klasse bildet die ‚determinierende‘ (das Bezugsnomen wird, ähnlich wie bei Komposita, determiniert) bzw. ‚referentielle‘ Funktion (z. B. *schulische Leistungen* ‚Leistungen in Bezug auf die Schule‘). Sie leitet keine Qualitäts-, sondern Relationsadjektive ab, die v.a. dazu dienen, ein nominales Konzept für den attributiven Gebrauch verfügbar zu machen.¹⁹ Diese Funktion kann sich aus der Affinität zu fremden Basen – und dem Bedarf, diese attributiv einzusetzen, entwickelt haben (vgl. z. B. *akrobatische Geschicklichkeit*, Kühnhold, Putzer und Wellmann 1978: 287). Ebenso können Fremdbasen eine Brücke zur ‚possessiv-ornativen‘ Funktion (*melancholisches Achselzucken* ‚mit Melancholie‘, Kühnhold, Putzer und Wellmann 1978: 300) und vielen anderen Funktionsklassen dargestellt haben. Die ‚aktivische‘ Funktion kann sich via Nomina Agentis auf *-er* entwickelt haben (vgl. o. (3)–(6) sowie das Beispiel *anklägerischer (Zeitungsartikel)*, das als *-erisch*-Derivat mit Bedeutung ‚Xend‘ analysiert werden kann).

Die Beschreibung der Funktionsklassen könnte noch weiter fortgesetzt und vertieft werden, doch dürfte ein wichtiger Aspekt deutlich geworden sein: Das deutsche *-isch*-Suffix zeichnet sich kaum durch ein klares Bedeutungsprofil aus – anders als manche stärker autosemantischen Suffixe wie etwa *-los* (‚ohne X‘, i.e. ‚privativ‘) oder *-bar* (‚kann ge-X-t werden‘, i.e. ‚passivisch-potentiell‘). Die Funktionen entstammen bei *-isch* weniger dem Suffix selbst als vielmehr

19 Vgl. Dornseiff (1921: 193), Hotzenköcherle (1968), Frevel und Knobloch (2005).

der Basis und teilweise auch dem Kontext (vgl. Kempf 2016b). Was *-isch* von anderen Suffixen unterscheidet und ihm doch ein eigenes Profil verleiht, ist wie oben gezeigt seine Affinität zu fremden und onymischen Basen sowie (schwächer ausgeprägt) zu negativ konnotierten Basen und Kontexten. Der eingangs dargestellte Produktivitätsverlauf kann in Verbindung mit diesen Basisklassen und den für sie relevanten soziohistorischen Entwicklungen – nämlich dem Sprachkontakt einerseits und dem Diskurs der Reformationszeit andererseits – begriffen werden.

2.2. Diachrone Entwicklung von engl. *-ish*

2.2.1. Kurzer Abriss relevanter Forschungsliteratur

Ähnlich zur deutschen *-isch*-Derivation ist auch das englische *-ish* bisher großenteils aus einer synchronen Perspektive behandelt worden, so vor allem in allgemeinen Wortbildungslehren, die eine Beschreibung des Wortbildungsmusters und des Funktionsspektrums leisten (z. B. Marchand 1969, Bauer 1983, Plag 1999, 2003, Bauer et al. 2013). Gelegentlich finden sich Aussagen zu Wortbildungsbeschränkungen im Sinne von phonologischem, morphologischem, semantischem und etymologischem *blocking* (z. B. postuliert in Malkiel 1977, Morris 1998, Sugawara 2012), die jedoch einer umfassenderen empirischen Überprüfung ausgesetzt werden müssten, zumal sie allesamt durch mitunter zahlreiche Einzelbelege widerlegt werden. Eine Reihe von Studien beleuchtet die Produktivität von *-ish* im Gesamtbild der zur Verfügung stehenden Wortbildungselemente und attestiert *-ish* jeweils einen der vorderen Plätze in verschiedenen Produktivitätsrankings (Baayen & Lieber 1991, Hay & Baayen 2003, Plag 2006), wobei eine Differenzierung zwischen den unterschiedlichen Basiskategorien, an die sich *-ish* anhängen kann, fehlt. Sprachhistorische Analysen von *-ish* konzentrieren sich eher auf bestimmte Sprachperioden (so Mendaza 2015 für das Altenglische, Ciszek 2012 sowie Dalton-Puffer 1996 für das Mittelenglische), ohne jedoch einen breiten langzeitdiachronen Blickwinkel einzunehmen. Von sprachwandeltheoretischem Interesse ist das sich im Gegenwartsendlichen aus der approximativen Funktion herausbildende freie *ish*, das als Repräsentant diverser Sprachwandelprozesse herangezogen wird: so Lexikalisierung (Szymanek 2005), Anti-Grammatikalisierung (Kuzmack 2007), Degrammatikalisierung im Sinne von *debonding* (Norde 2011) oder Konstruktionalisierung (Traugott und Trousdale 2013). Die hier vorgestellte Studie hat vor dem Hintergrund der bisherigen Forschungsliteratur den Anspruch, das diachrone Desiderat einer empirisch fundierten Untersuchung von englisch *-ish* aufzuarbeiten, und basiert im Wesentlichen auf den Ergebnissen der Studie von Eitelmann, Haugland und Haumann (in Vorb. a), die im Detail die Ausbreitung des englischen *-ish*-Suffixes über verschiedene Basiskategorien vom Altenglischen

bis zum Gegenwartsgenesis analysiert. Um einen kontrastiven Vergleich mit dem deutschen Gegenstück zu erlauben, wird besonderes Augenmerk auf kontrastiv relevante Aspekte gelegt, die somit in einem neuen Zusammenhang betrachtet werden.

2.2.2. Methodologische Vorbemerkungen

Der nachfolgenden Untersuchung liegen die in Tabelle 11 aufgelisteten historischen Korpora zugrunde. Insgesamt decken die allesamt dem britischen Englischen zugehörigen Korpora einen Zeitraum vom Altenglischen bis zum späten 20. Jh. ab, wobei für die Perioden ab dem Frühneuenglischen das Genre narrative Prosa weitestgehend konstant gehalten wurde. Als etwas problematisch ist die Korpuslage im Mittelenglischen zu bezeichnen, für die – in Ermangelung einer geeigneteren Textgrundlage – über den Umweg des *Middle English Dictionary* auf Prosa- und Verssprache zurückgegriffen wurde.

Tab. 11: Übersicht über die herangezogenen englischen Korpora.²⁰

	Korpora	Zeitraum	Wortanzahl
Altenglisch	Dictionary of Old English Corpus	ca. 600–1150	3 Mio.
Mittelenglisch	Penn-Helsinki-Parsed Corpus of Middle English	1150–1500	1,2 Mio.
	Corpus of Middle English Prose and Verse via Middle English Dictionary	1100–1500	3,7 Mio.
FNE (16. & 17. Jhd.)	Early English Prose Fiction Eighteenth Century Fiction Part 1	*1460–1682	10 Mio.
		*1660–1699	5 Mio.
18. Jhd.	Eighteenth Century Fiction Part 2 Nineteenth Century Fiction Part 1	*1700–1752	5 Mio.
		*1728–1799	12 Mio.
19. Jhd.	Nineteenth Century Fiction Part 2	*1800–1869	27 Mio.
20. Jhd.	British National Corpus (BNC), wridom1	^p 1960–1993	19 Mio.
Summe			85,9 Mio.

Aus den oben gelisteten Korpora wurden sämtliche *-ish*-Derivate extrahiert und bezüglich ihrer Basiskategorie (Nomen, Eigennamen, Adjektiv etc.) sowie nach Komplexität der Basis klassifiziert. Auf diese Weise wurden Token- und Type-Frequenzen ermittelt sowie Hapax legomena identifiziert. Anders als bei der deutschen Korpusanalyse wurde die *-ish*-Derivation nicht im Zusammenhang anderer adjektivischer Wortbildungsmuster betrachtet; nicht nur wäre ein solches Unterfangen angesichts der Korpusgrößen mit einem durch die erforderliche manuelle Klassifikation extremen Zeitaufwand verbunden, der

²⁰ In der Tabelle steht * für die Geburtsdaten der Autoren, ^p für Publikationsdaten.

Versuch einer klaren Kategorisierung sämtlicher Derivationsuffixe wäre auch aufgrund der beobachtbaren funktionalen Heterogenität englischer Wortbildungselemente im Allgemeinen problematisch.²¹ Ebenfalls ist zu bemerken, dass bei der englischen Korpusanalyse alle nationalitätsbezeichnenden *-ish*-Formationen herausgenommen wurden, da der verhältnismäßig übergroße Anteil an Tokens das Bild verzerrt hätte. Gerechtfertigt ist dies auch durch den Umstand, dass diese Funktion im Englischen nur noch eine geringe bis gar keine Rolle mehr spielt in Bezug auf das Vermögen, neue nationalitätsbezeichnende Adjektive zu generieren. Im gegenwartssprachlichen Englisch wird die aus dem Germanischen ererbte ethnische Funktion durch rivalisierende romanische Suffixe (*-ese*, *-(i)an*) oder Nullmorphem ausgedrückt, weswegen *-ish* für Herkunftsbezeichnungen nicht mehr zur Verfügung zu stehen scheint.²²

2.2.3. *Entwicklungspfad von engl. -ish*

Zweifelsohne steht das Derivationsuffix *-ish* als ein produktives Wortbildungsmuster über alle Sprachperioden des Englischen hinweg zur Verfügung, wie die Ergebnisse aus Eitelmann, Haugland und Haumann (in Vorb. a), zusammengefasst in Tabelle 12, belegen. Aufgrund der stark schwankenden Korpusgrößen sind die Produktivitätsmaße zwar nur bedingt diachron vergleichbar. Auffällig ist jedoch, dass alle Werte auf eine gesteigerte Produktivität im 20. Jh. verweisen: Der P-Wert des 20. Jh. liegt mehr als doppelt so hoch wie der des Zeitschnitts 1500–1700 (bei dem ein grob ähnlicher Token-Wert im Nenner steht und somit ein Vergleich der jeweiligen P-Werte durchaus zulässig ist); dasselbe gilt für die Type/Token-Ratio. Auch der Index Hapax/Types zeigt für das 20. Jh. eine Steigerung gegenüber dem 19. Jh. (bei dem ein ähnlicher Type-Wert im Nenner steht).

21 Hinzu kommt, wie bereits angemerkt, dass anders als im Deutschen keine umfassenden korpusbasierten Studien zu den gesamten Derivationsystemen einzelner Epochen vorliegen.

22 Dies lässt sich exemplarisch anhand fiktiver Nationalitätsbezeichnungen konkretisieren: Wenn jemand die Kunstsprache *klingtonisch* spricht, könnte man im Englischen gerade nicht von *Klintonish* sprechen, da dies eben nur eine dem Klingtonischen ähnliche, aber nicht identische Sprache meinen würde (diese wird als *Klinton* mit Nullmorphem bezeichnet). Gleichzeitig illustriert das Beispiel kontrastive Unterschiede: Im Deutschen steht *-isch* als Herkunftsbezeichnung nach wie vor zur Verfügung. Tatsächlich gibt es bei einer Google-Suche für „speaks Klintonish“ keine Belege, für „speak Klintonish“ nur wenige und eher zweifelhafte, während „speak(s) Klinton“ die deutlich etablierte Variante darstellt.

Tab. 12: Übersicht über Type-Token-Frequenzen sowie Anzahl der Hapax legomena (Eitelmann, Haugland und Haumann in Vorb. a).

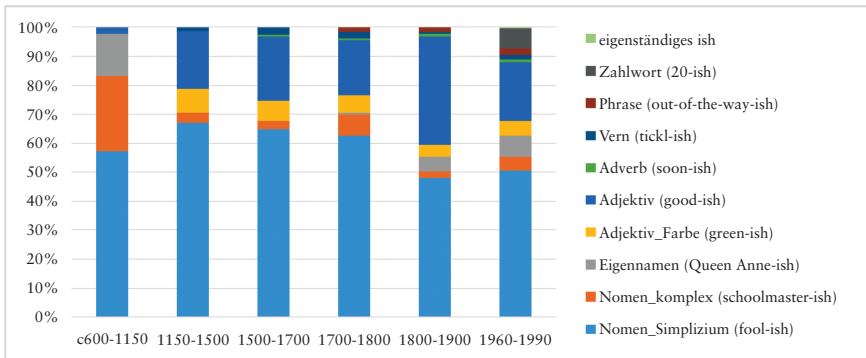
	c600–1150	1150–1500	1500–1700	1700–1800	1800–1900	1960–1993
Korpusgröße Mio. Wörter	3	4,9	15	17	27	19
-ish Token	632	511	3236	2337	5535	3046
-ish Typen	42	142	154	136	252	259
Type/ Token-Ratio	0,0664	0,2779	0,0475	0,0581	0,0455	0,085
-ish Hapaxe	21	69	55	57	101	110
P-Wert (= Hapaxe/ Tokens)	0,0332	0,1350	0,017	0,0244	0,0182	0,0361
Hapax/ Type-Ratio	0,50	0,49	0,36	0,42	0,40	0,42

Angesichts dieser beobachtbaren Produktivitätszunahme stellt sich die Frage, wodurch diese bedingt ist. Wie wir in der Einzelstudie zum Deutschen gesehen haben, ist die anhaltende Beliebtheit des *-isch*-Suffixes dadurch bedingt, dass es mit der ‚Eroberung‘ der Fremdbasen im Konkurrenzgefüge mit anderen Derivationssuffixen eine neue, stabile Funktionsnische besetzt hat. Im Englischen tragen zwei wesentliche Faktoren zur Produktivitätsexplosion bei: Zum einen wird das Funktionsspektrum erweitert, indem zu der Zugehörigkeitsfunktion, die das Englische mit dem Deutschen teilt, die approximative Funktion hinzutritt, die annähernde Bedeutung ausdrückt. Zum anderen kommt es zu einer beträchtlichen Erweiterung der Basiskategorien, womit das Deutsche und das Englische letzten Endes unterschiedliche Pfade einschlagen. So weisen Deutsch und Englisch einerseits mit nominalen Basen (*fool-ish*), Eigennamenbasen (*Queen Anne-ish*) und verbalen Basen (*tickl-ish*) eine Überlappung in den zur Verfügung stehenden Wortkategorien auf; andererseits sind adjektivische Basen (*green-ish*, *happy-ish*) und Numeralia (*20-ish*) exklusiv dem Englischen vorbehalten²³. Ebenso können die Basen von variierender Komplexität sein: So finden sich, wiederum ähnlich zum Deutschen, neben Simplizia (*childish*) auch komplexere Wortbildungen wie Komposita (*schoolmaster-ish*), aber auch vom Deutschen abweichend, ganze Phrasen, die den Skopus für die

23 Um Approximation zu Numeralia auszudrücken, wird im Deutschen v.a. auf syntaktische Konstruktionen zurückgegriffen, vgl. die z.T. umgangssprachlichen Konstruktionen (*so(etwa) um acht (rum)* (Uhrzeit); (*so(etwa) um die vierzig* (Alter); (*so(etwa) dreihvier, um und bei zehn* (Anzahl; schriftsprachlich: *ca. X, rund X*); u.a.m.

-ish-Affigierung bilden können (*light-at-the-end-of-the-tunnel-ish*)²⁴. Wie sich eben diese Erweiterung der Basiskategorien vom Altenglischen bis zum Gegenwartsendenglischen des späten 20. Jahrhunderts entwickelt, zeigt Abbildung 5, die die diachrone Verteilung aller -ish-Derivationstypen, differenziert nach Basiskategorien, darstellt.

Abb. 5: Entwicklung der nach Basiskategorien differenzierten -ish-Typen (Eitelmann, Haugland und Haumann in Vorb. a).



Werfen wir zunächst einen Blick auf den ersten Balken, der die -ish-Derivation im Altenglischen repräsentiert. Am stärksten vertreten ist die Basiskategorie der Simplizia (*eorlisc*), mit komplexen Kompositionsbasen (z. B. *godspelisc*, *ofersæwisc*) und Eigennamenbasen (*Davidisc*) auf dem zweiten und dritten Platz. Bezeichnenderweise findet sich in den altenglischen Daten mit *gæglisc* (von *gægol*, ‚freudig‘) ein erstes deadjektivisches Hapax, wobei -isc hier primär zur Stärkung des adjektivischen Charakters beitragen dürfte als tatsächlich ein Vorbote all jener deadjektivischen Bildungen aus dem Mittelenglischen zu sein, die approximative Bedeutung innehaben. Die im Altenglischen vertretenen Basiskategorien sind somit diejenigen, die nach wie vor im Deutschen von Relevanz sind, und auch in den nachfolgenden Sprachperioden des Englischen spielen die Basiskategorien, die die Ähnlichkeit zum Deutschen unterstreichen, eine ungebrochen große Rolle. So zeigt sich deutlich, dass simplizische Substantive über alle Zeiträume hinweg die am meisten vertretene Basiskategorie bilden; nur in den letzten beiden Perioden geben sie, auch bedingt durch die zunehmende Vielfalt, an die anderen Basiskategorien ab. Der Anteil an komplexen Nominalbasen ist verglichen mit den Simplizia geringfügiger und

24 Desubstantivisch-vergleichende und phrasale Ableitungen werden im Deutschen v.a. durch die -mäßig-Derivation realisiert (vgl. *schulbuchmäßig*, *limousinenmäßig*, *Licht-am-Ende-des-Tunnels-mäßig* [<http://0weil-ichs-kann0.blogspot.de/2012/03/>], zuletzt aufgerufen 02.02.2018]).

wechselhaft.²⁵ Auch deonymische Basen sind in allen Zeiträumen vertreten (wenn auch im Mittenglischen und 16./17. Jh. nur schwach) und befinden sich gerade im jüngsten Zeitraum wieder im Aufschwung. Eine weitere Parallele zum Deutschen liegt in einer beträchtlichen Anzahl von denominalen *-ish*-Derivaten, die im Zuge eines Lexikalisierungsprozesses pejorative Bedeutung innehaben und die sich insbesondere im Mittenglischen herausgebildet haben (vgl. z. B. *childish*, das seit dem Mittenglischen nicht mehr ‚kindlich‘ wie noch im Altenglischen bedeutet, oder auch *sheepish*, *wolfish*, *foolish*, etc.). Allerdings scheint es sich im Englischen eher um ein nicht weiter verfolgtes Zwischenstadium zu handeln, da in späteren Sprachstufen keine besondere Affinität zu negativen Basen festzustellen ist und deshalb das englische Suffix anders als sein deutsches Gegenstück nicht latent negativ assoziiert ist. Ein Grund hierfür ist sicherlich die seit dem Mittenglischen sich entwickelnde approximative Bedeutung, jedoch wäre diese Annahme noch empirisch genauer zu überprüfen.

Eklatante Unterschiede zum Deutschen beginnen sich erst ab dem Mittenglischen herauszubilden, die sich in den nachfolgenden Sprachperioden nach und nach weiter festigen. So kommen im Mittenglischen adjektivische Basen hinzu, zunächst prominent vertreten mit den Farbadjektiven, die jedoch recht schnell ausgeschöpft sind, und überspringend auf alle Arten von Adjektiven. Mit den adjektivischen Basen entwickelt sich dann auch die Funktion, die dem Deutschen fremd ist, nämlich die approximative Funktion: *greenish* ‚vage grün, grünlich‘. Gleichzeitig kommt es aufgrund dieser Funktionserweiterung zu zahlreichen Dublettenbildungen; so bedeutet *waterish*, wie es im Mittenglischen belegt ist, ursprünglich ‚wässrig‘ mit vergleichendem *-ish*, während *waterish* in späteren Belegen auch approximativ verwendet werden kann im Sinne von ‚vage an Wasser erinnernd, aber eben ungleich Wasser‘. Beobachtbar ist in diesem Zusammenhang eine zunehmende Subjektivierung (sensu Traugott 1989). Was onymische Basen angeht, so ist die alte Zugehörigkeitsfunktion verdrängt worden durch die approximative Funktion, d.h. *das Shakespeare'sche Werk* ließe sich nicht mit *Shakespeareish* übersetzen, weil dies eben kein von Shakespeare geschaffenes Werk bezeichnen würde (sondern eher ‚shakespeare-artig‘).²⁶ Somit können wir, beginnend mit dem Mittenglischen, von zwei mehr oder weniger getrennten Funktionen ausgehen, die

25 Wie ein Reviewer angemerkt hat, wäre es interessant zu untersuchen, ob der beobachtbare Rückgang an komplexen Kompositionsbasen auch damit zusammenhängt, dass das Wortbildungsmuster der Komposition relativ im Vergleich zu romanischstämmiger Wortbildung vom Alt- zum Mittenglischen generell an Bedeutung verliert.

26 Übernommen wird die Zugehörigkeitsfunktion von Suffixen wie *-ean* (*Shakespearean*), oder sie wird durch periphrastische Konstruktionen ausgedrückt.

-ish ausübt: zum einen die semantisch neutrale vergleichende Funktion, die bei Neubildungen vor allem in attributiver Stellung zur Geltung kommt und die die originäre gemeingermanische Zugehörigkeitsfunktion fortsetzt, und zum anderen die approximative Funktion, die bei gegenwartssprachlichen Neubildungen vor allem mit prosodisch besonders markierter Aussprache vorkommt, indem -ish mit Glottisverschluss sonderbetont wird.²⁷

Die Extension auf weitere Basiskategorien über die herkömmlichen nominalen und onymischen Basen sowie die neu hinzugetretenen adjektivischen hinaus setzt sich beobachtbar über die nachfolgenden Jahrhunderte fort, wobei verbale und adverbale Basen jedoch durch die Reihe hinweg spärlich belegt sind. Erste phrasale Basen tauchen im 18. Jh. mit *old maidish* oder *not know how-ish* auf und bezeugen eine neue Art von Komplexität, die sich durch den erweiterten Skopus ergibt; im Falle von *old maidish* z. B. hängt sich -ish nicht etwa in erster Linie an *maid* an, sondern an die gesamte Nominalphrase: [*old maid*]_{NP}-ish. Das 20. Jh. ist schließlich insofern von Innovation geprägt, als -ish erstmals an Numeralia angehängt wird und die approximative Funktion somit eine neue Stärkung erfährt. Überdies sind Numeralia die erste Basiskategorie, bei der durch die -ish-Affigierung bemerkenswerterweise kein Wortartenwechsel induziert wird; affigierte Numeralia werden im Großen und Ganzen weiterhin als eben solche Numeralia im syntaktischen Zusammenhang verwendet (z. B. *fortyish-year-old women* [COCA]).

Begünstigt durch die Erweiterung auf phrasale Basen und die damit einhergehende allmähliche Klitisierung von -ish, kommt es schließlich im 20. Jahrhundert zum Aufkommen von *ish* als freies Lexem, wie in diesem Beleg von 1990 aus dem BNC:

You must try to remember that some people are normal. Ish.

Es ist bezeichnend, dass freies *ish*, das besonders im gesprochenen Englisch stark vertreten zu sein scheint, nachgeschoben werden oder gar einen vorangegangenen Redebeitrag modifizieren kann wie in folgendem Befund aus dem *Corpus of Contemporary American English*:

- (15) GIFFORD: All right, one of them is cheap...
 KOTB: So – inexpensive.
 GIFFORD:... er, ish. (COCA_SPOKEN, 2011)

²⁷ Traugott und Trousdale (2013: 233–237) konstruktionsgrammatische Behandlung von -ish nimmt hierfür zwei Makroschemata an: vergleichendes (oder in ihrer Terminologie ‚associative‘) *ish*₁ und approximatives *ish*₂, hervorgegangen durch den Prozess der Konstruktionalisierung.

Solche Verwendungsweisen, in denen *ish* als sprachliche Hecke fungiert, können dabei auch als Ausprägung eines gerade im Englischen beobachtbaren Demokratisierungsprozesses verstanden werden. Unter *democratization* (Fairclough 1992, Farrelly & Seoane 2012) wird gemeinhin die Vermeidung overter Ausdrucksmittel gefasst, die Machtasymmetrien zwischen Sprechern anzeigen, und die Präferenz gesichtswahrender Strategien. Gerade approximatives *-ish* steht in diesem Zusammenhang als Option zur Verfügung, im Nachhinein das Sprecher-*commitment* zu reduzieren und eine gesichtswahrende Korrektur vorzunehmen. In dieser Hinsicht kann man von einem Pragmatikalisierungsprozess sprechen, den klitisiertes und freies *-ish* hin zu einem Diskursmarker durchläuft.

Neben diesen Entwicklungspfaden, die zu einer Auseinanderentwicklung von dt. *-ish* und engl. *-ish* geführt haben, zeigt sich ein weiterer Unterschied darin, dass das englische Pendant keine auffallende Affinität zu einer bestimmten Art von Basis hinsichtlich deren Fremdstatus aufweist. Seit frühesten Zeiten sind hybride Wortbildungen belegt, wie z. B. altenglisch *puerisc* (von lat. *puer*, ‚Junge‘) oder umgelautetes *milisc* (von lat. *melus* ‚Honig‘, im Sinne von ‚süß‘, Klein 1994); ebenso kommt es bei Entlehnungen wie vom lat./gr. *mechanikós/mechanicus* zu einem Suffixwechsel, indem *-isc* das gräko-romanische Derivationsaffix ersetzt: vgl. altengl. *mechanisc*. Wie Dalton-Puffer (1996) in ihrer Studie zum Mittelenglischen zeigt, hängen sich native Suffixe wie *-ish* weiterhin an Fremdbasen an, also auch nachdem ab dem Mittelenglischen das englische System der Derivationsuffixe regen Input durch neue französische und lateinische Affixe erfährt. Dies bestätigt sich auch in unseren eigenen Korpusdaten, die eine große Menge an hybriden Wortbildungen aufweisen (z. B. *foolish*, *beauish*, *aguish*, *brutish*, *nourish*, *decentish*, etc.). Selbst die Annahme einer etymologischen Einschränkung auf native Basen, derzufolge germanisch-stämmiges *-ish* und romantisches *-ous* miteinander rivalisieren und sich hinsichtlich des zugrundeliegenden Etymons auseinanderdividieren lassen (Baayen & Lieber 1991: 826), lässt sich so nicht halten, da sich in den Daten auffallende Dublettenbildungen wie etwa *vapourish* neben *vapourous* finden. Im Gegensatz zur Situation beim deutschen *-isch* stellen die Hybridbildungen des englischen *-ish* aber keine Spezialisierung dar, insbesondere da auch andere native Suffixe Fremdbasen ableiten bzw. ein gegenüber dem Fremdheitsstatus indifferentes Derivationsverhalten zeigen (vgl. die nativen adjektivischen Suffixe *-y* und *-ly*, die ebenfalls zahlreich mit Fremdwortbasen vorkommen, z. B. *flowery*, *cushiony*, *arty*, *faulty*; *cowardly*, *masterly*, *scholarly*, *mannerly*, etc. (Beispiele aus Dixon 2014: 228–231)). Dass es im Englischen nicht zu einem ähnlichen ‚Nischenverhalten‘ wie im Deutschen kommt, ist sicherlich dem Umstand geschuldet, dass generell Fremdbasen im Englischen weit besser

integriert und oft, aufgrund phonologischer und morphologischer Anpassungsprozesse, als solche gar nicht mehr erkennbar sind.²⁸

Summa summarum lässt sich eine Zunahme in zweierlei Hinsicht feststellen. Einerseits ist insofern eine Zunahme an Variabilität zu beobachten, als die *-ish*-Typen zunehmend heterogen werden aufgrund stets neu hinzukommender Basiskategorien, die wiederum von unterschiedlicher Komplexität sind: nominale, onymische, adjektivische, verbale, adverbiale, numerale bis hin zu phrasalen. Andererseits ist eine Zunahme an Kreativität im Sinne des Vermögens, Wortbildungsmuster auf unberechenbare und in einer nicht regelhaften Weise anzuwenden (sensu Lyons 1977: 549), zu verzeichnen: So sind späte *-ish*-Ableitungen aus z. B. Phrasen, Zahlwörtern, Pronomen etc. zwar nicht hochproduktiv, weisen jedoch in den ad hoc-Bildungen ein hohes Maß an kreativer Innovation auf, da sie die vermeintlichen Regeln der Wortbildung austesten und auch sprengen.²⁹

3. Kontrastive Auswertung

3.1. Empirische Synopse

Im Folgenden sollen die wichtigsten Entwicklungsstränge beider Sprachen synoptisch zusammengefasst werden. Angesichts des gemeinsamen germanischen Ursprungs interessiert besonders die Frage, wo und ab wann divergierende Pfade eingeschlagen wurden.

Ausgehend von der gemeinermanischen Zugehörigkeits- und Herkunftsfunktion findet in beiden Sprachen ein Übergang zur Bezeichnung typischer Eigenschaften statt, die jemand oder etwas aufweist, der/das zu X gehörig ist oder aus X stammt. Die Funktion „typische Eigenschaft“ weitet sich in beiden Sprachen auf eine allgemeine Vergleichsfunktion aus (,wie (ein) X‘), bei der keine Zugehörigkeit mehr bestehen muss. Im Zusammenhang mit

28 Die Behauptung eines nicht beobachtbaren Nischenverhaltens lässt sich natürlich nicht pauschal auf alle Affixe übertragen; so zeigt z. B. romanischstämmiges *-ity* in der Konkurrenz zu *-ness* ganz deutlich eine Präferenz zu romanischstämmigen Basen (vgl. Lindsay 2012: 195).

29 Diese kreative Innovation wird gerade in noch jüngeren Daten ersichtlich: gegenwartssprachlichen Daten des Amerikanischen Englisch im späten 20., frühen 21. Jh., die ein noch weiter facettiertes Bild von *-ish* entwerfen, wie die Studie von Eitelmann, Haugland und Haumann (in Vorb. b) gezeigt hat. Bemerkenswert sind innovative Vorkommnisse von *ish* als freiem Lexem in den Daten, so z. B. in attributiver Stellung: „She opened her ish eyes“, bezogen auf eine nicht eindeutig bestimmbare Augenfarbe. Des Weiteren vermag sich *-ish* an gebundene Basen und Pronomen anzuhängen – was insofern innovativ und regelsprengend ist, als *-ish* interessanterweise bei den Pronomen keinen Wortartenwechsel bewirkt wie in folgendem Beleg: „Feeling responsible-ish for everybody-ish“ [COCA, NEWS].

stereotypischen Zuschreibungen und Vergleichen entstehen in beiden Sprachen reihenhafte Gruppen pejorativer *-isch*-Derivate (*childish*, *kindisch*). Dies führt in keiner von beiden Sprachen zu einer inhärenten pejorativen Funktion von *-is(c)h*, bewirkt aber im Deutschen v.a. in der Frühmoderne eine starke Affinität zu negativ bewerteten Basiswörtern (*Ketzer*, *Neid*).

Im Englischen ergibt sich aus der allgemeinen Vergleichsfunktion die approximative Funktion. Dies geht mit einer Ausweitung der Basiswortarten einher, wobei zunächst (im Mittelenglischen, vgl. Abbildung 5) Adjektive eine prominente Rolle spielen (*greenish* ‚wie grün, ein bisschen grün‘). Beide Funktionen, die vergleichende wie die approximative, bleiben bis zum Gegenwartenglischen produktiv. Weiterhin dehnen sich die Basisklassen aus, nämlich auf weitere Wortarten (Numeralia, Pronomina) und komplexe Basen (bis hin zu Phrasen). Die approximative Funktion ist so stark entwickelt (vor allem im gesprochensprachlich-expressiven Gebrauch), dass sie die Zugehörigkeits- und Herkunftsfunktion verdrängt. Ganz anders verhält sich das im Deutschen, wo sich aus der ursprünglichen Funktion der Zugehörigkeits- und Herkunftsbezeichnung eine starke Affinität zu Eigennamen und über diesen Umweg auch zu fremdwörtlichen Basen entwickelt hat. Beide Basisklassen sind heute noch aktiv. Das Versiegen der gräko-romanischen Entlehnungen wird durch komplexe Derivation (*postkatastrophisch*) und Neubildungen (*Wikipedia*) teilweise, wenn auch nicht ganz aufgewogen; ähnliches gilt für Eigennamen, wo neue abzuleitende Namen (*Fidel*) den Verlust der Städtenamen (*Mainzisch* > *Mainzer*) teilweise wettmachen. Aufgrund der fremdwörtlichen und onymischen Basen ist die rezente Produktivität im schriftsprachlichen Bereich zu verorten, und zwar vor allem in Wissenschafts- und Zeitungstexten.³⁰

Neben der Zugehörigkeits- und Herkunftsfunktion ist im Gegenwartsteutschen v.a. die vergleichende Funktion, daneben auch u.a. die ‚aktivi-sche‘ Funktion (*schuldenmacherisch*) aktiv (für Genaueres s. Kempf 2016a: 253–257). Hier ist eine latente Ausweitung auf Basisverben zu beobachten (die im Englischen nur eine marginale Rolle spielen), insgesamt findet aber keine ansatzweise vergleichbare Basisklassenausweitung statt: Adjektive sowie komplexere Phrasen fehlen gänzlich. Die deutsche *-isch*-Derivation hatte, bedingt durch die lateinische und französische Entlehnungswelle, im 16. bis 18. Jh. Hochkonjunktur, während die englische im 20. und 21. einen immensen Produktivitätsschub erlebt. Hier wäre beim approximativen Schema

30 Dies ergeben für das 20. Jh. sowohl eine DWDS-Abfrage (<http://eins.dwds.de/>, 11.5.2017, Abfrageformel „\$p=/ADJ./ with /.*.[^fr]isch(e(s|nl|mlr)?):?\$/“) als auch Schlaefers (1977: 177–184) Studie. Von 1650 bis 1800 dominieren Zeitungen, während v.a. Predigten das Suffix meiden (s. Kempf im Ersch.). Für das Mhd. beschreibt Klein (1994: 393) das Genre Fachliteratur als führend in der *-isch*-Verwendung.

zu prüfen, ob es sich wirklich noch um derivationelle Produktivität handelt oder ob (und wie weit) sich die Konstruktion schon zu einem epistemischen Marker des Zweifels (dubitativ) entwickelt hat.

Bevor wir im nächsten Kapitel die empirischen Beobachtungen einer theoretisch-typologischen Einordnung unterziehen, seien die wichtigsten Ergebnisse in einer vergleichenden Übersicht zusammengefasst (Tabelle 13).

Tab. 13: *Entwicklungsmerkmale der deutschen vs. englischen -is(c)h-Derivation im Überblick.*³¹

Deutsch	Englisch
Produktivitätsverlauf:	
vom Ahd. bis zur Gegenwart produktiv, dabei Hochkonjunktur ca. 16.–18. Jh.	vom Ae. bis zur Gegenwart produktiv, dabei starker Produktivitätsschub seit dem 20. Jh.
Derivationsbasen:	
Basiswortarten:	
Substantive durchgehend äußerst dominant, doch Bestimmung bei Doppelmotivation oder gebundenen Stämmen schwierig; geringer Zuwachs bei Verben im Nhd.	Substantive anfangs dominant, ab dem Mittelengl. Ausdehnung von Adjektiven (bis hin zu 42% im 19. Jh.); sukzessive Ausdehnung auf weitere Wortarten (Adverbien, Verben, Zahlwörter, Pronomina; komplexe Phrasen) bis hin zu insgesamt ca. 12% im 20. Jh.
Onymische Basen:	
stark dominanter Typus vom Ahd. bis ins frühe Nhd. bis hin zu 68% (1700–1750), dann Rückgang auf 27% im 20. Jh.	nur im Ae. auffällig mit 14%, danach erst im 19. und 20. Jh. wieder nennenswert (5% bzw. 7%)
Nicht-native Basen:	
wichtiger Produktivitätsfaktor; seit dem Ahd. belegt, (spätestens) seit dem Fnhd. kontinuierliche Zunahme (bis hin zu 62%, 1750–1800); mindestens 55% im 20. Jh., in Neubildungen deutlich mehr	Ableitung aus Fremdbasen seit dem Ae. durchgehend vertreten, doch kein Spezifikum der -isb-Derivation, da fremder und nativer Wortschatz weniger separiert als im Deutschen

31 Alle Prozentangaben beziehen sich auf Types.

Deutsch	Englisch
Basiskomplexität:	
seit dem Fnhd. kontinuierliche Zunahme der Komplexität, doch dies generelle Tendenz der dt. Wortbildung; keine <i>-isch</i> -Ableitung komplexer Phrasen	wichtiger Produktivitätsfaktor; komplexe Substantive durchgehend belegt; de-phrasale Derivation seit 1700 steigend (bis hin zu 7% im 20. Jh.), dabei auch hochkomplexe Phrasen; Merkmal der Degrammatikalisierung/ Verselbständigung vom Suffix zum Klitikon und mitunter freiem <i>ish</i>
Negativ konnotierte Basen:	
im Fnhd. zunehmende Affinität zwischen <i>-isch</i> und negativ besetzten Basen oder Kontexten; v.a. im Bereich nativer Basen (dort Nische gegenüber Konkurrenzsuffixen); negative Bedeutung des Musters nur bei Dubletten (<i>kindisch</i> – <i>kindlich</i> , <i>dörfisch</i> – <i>dörflich</i>)	im Zuge der ‚vergleichenden‘ Funktion im Mittelenglischen reihenweise Bildungen mit negativ besetzter Gesamtbedeutung (<i>foolish</i> , <i>sheepish</i>), doch keine weitere Spezialisierung oder produktive Nischenbildung
Wortbildungsfunktionen:	
ursprüngliche Funktion: Bildung von Zugehörigkeits- und Herkunftsadjektiven, dann auch zur Bezeichnung typischer Eigenschaften; davon ausgehend ‚vergleichende‘ Funktion; zwischen dem Mhd. und der Gegenwart starke Polyfunktionalisierung; v.a. durch Fremdbasen Ausweitung auf weitere, v.a. Relationsadjektive bildende Funktionen wie ‚determinierend/ referentiell‘, ‚gleichsetzend‘, ‚korrespondierend‘ u.a.m.; gelegentlich deverbale: ‚aktivisch‘	ursprüngliche Funktion: Bildung von Zugehörigkeits- und Herkunftsadjektiven, dann auch zur Bezeichnung typischer Eigenschaften; davon ausgehend ‚vergleichende‘ Funktion; dann Ausdehnung auf approximative Funktion (über Farb- und später andere Adjektive); dadurch Verdrängung der Zugehörigkeits-/Herkunftsfunktion; approximative Funktion Grundlage rezenter Produktivität/Basisklassenausweitung/Verselbständigung

3.2. Typologische Faktoren

Die beobachtbaren Unterschiede zwischen dem Deutschen und dem Englischen, die oben synoptisch zusammengefasst wurden, werfen die Frage auf, inwiefern diese durch typologische Eigenschaften des Deutschen und Englischen bedingt sind. Einen hilfreichen Erklärungsansatz bietet hierbei Bergs (2014) Kriterium der *boundary permeability*. Dieser typologische Parameter bezieht sich auf eine Grenzdurchlässigkeit in dreierlei Hinsicht:

- Erstens ist Grenzdurchlässigkeit auf den Grad der Überschneidung von linguistischen Einheiten, also Wortkategorien, bezogen (d.h. wie stark lassen sich Wortkategorien voneinander abgrenzen?);
- zweitens lässt sich Grenzdurchlässigkeit im Sinne einer Überlappung von Eigenschaften verstehen (gemeint im Sinne eines mehr oder weniger eindeutigen *form-function-mapping*); und
- drittens bezieht sich Grenzdurchlässigkeit auf die Interaktion zwischen diversen Sprachbetrachtungsebenen, d.h. von Interesse ist, ob es zu gegenseitigen Beeinflussungen zwischen beispielsweise Morphologie und Phonologie oder Morphologie und Syntax kommt.

Für das Deutsche stellt Berg (2014) eine deutliche Tendenz zum Grenzen wahren Typus fest. Im gegenwärtigen Kontext sind dabei vor allem die folgenden Beobachtungen relevant: Im adjektivischen Bereich wird heute eine eindeutigere Wortartmarkierung realisiert als noch im Fnhd. So kann das Suffix *-lich* heute nur noch Adjektive ableiten und nicht mehr wie noch im Fnhd. daneben auch Adverbien (vgl. Relikte wie *bitterlich weinen*, aber *bittere Tränen*). Nhd. Suffixe sind i.d.R. wortarteindeutig, daher wären Ausweitungen wie etwa *everybodyish*_{PRON} nicht möglich. Auch ist es z. B. nicht (mehr) möglich, Substantivstämme mit unklarem morphosyntaktischen Status in die attributive Position zu setzen – vgl. fnhd. Konstruktionen des Typs *eisen kette*, *leinen hemd*, bei denen nicht klar ist, ob es sich um Nominalkomposita oder *-e(r)n*-Derivate (vgl. *golden*) handelt (Sattler 1992: 230). Auf den Bereich der Zugehörigkeits- und Herkunftsfunktion übertragen, bedeutet dies, dass entsprechende Personenbezeichnungen auf jeden Fall morphologisch markiert werden müssen, um in die attributive Position gesetzt werden zu können. Bloße Voranstellung konvertierter oder wortartambiger Elemente wie im Englischen (*heathen customs*, *Muslim clergyman*) ist nicht möglich – entsprechend kommt die *-isch*-Derivation hier häufig zum Einsatz (*heidnische Sitten*, *muslimischer Geistlicher*).

Ein weiterer relevanter Aspekt liegt darin, dass Namen und Fremdwörter als Sonderbereiche des Wortschatzes abgegrenzt werden, indem sie ein grammatisches Sonderverhalten zeigen. Für Namen gilt das auch im Englischen, vgl. z. B. deren Großschreibung oder das auf Namen spezialisierte Ableitungsmuster *-ian*. Im Deutschen sind in der Entwicklung der Substantivgroßschreibung neben den Namen auch Fremdwörter dem übrigen Wortschatz vorangeschritten (Bergmann und Nerius 1998: 873). Fremdwörter werden auch durch Apostrophgebrauch stärker hervorgehoben als native Wörter (*Tempo's*, Scherer 2010, Nübling 2014). Vor allem aber stehen fremdsprachige Elemente in der Wortbildung in einem partiell autonomen Teilsystem (Bergmann 1998, Kempf 2010). Während in der Komposition hybride Bildungen möglich sind, können fremde Suffixe so gut wie gar nicht mit nativen

Elementen kombiniert werden. Im Kontext dieser partiellen Autonomie lässt sich die *-isch*-Derivation als systemgerecht begreifen, da sie mit ihrer Spezialisierung auf Fremdmaterial dazu beiträgt, dieses vom übrigen Wortschatz abzugrenzen. Damit entspricht sie auch der von Berg (2014) festgestellten typologischen Tendenz des Deutschen zum Grenzen wahrenden Sprachtypus.

Das Englische lässt sich hingegen als ‚soft boundary language‘ charakterisieren, wie Berg (2014) anhand von 14 Fallstudien ausführt. So verweist beispielsweise der im Englischen verhältnismäßig hohe Anteil an Konversionen auf diffuse Grenzen zwischen Wortkategorien, während im Deutschen die Überführung eines Wortes in eine andere Wortart in der Regel morphologisch markiert werden muss (Berg 2014: 495–496). Auch der hohe Anteil an ambitransitiven oder semireflexiven Verben im Englischen zeugt von einer enormen Grenzdurchlässigkeit: Kanonisch intransitive Verben wie *walk* oder *sleep* können durchaus transitiv verwendet werden (*walk the dog home, sleep the hangover away*), und das Weglassen des Reflexivpronomens ist bei einer Reihe von Verben, die gemeinhin an sich selbst verübte Tätigkeiten bezeichnen (insbesondere sogenannte *grooming verbs* wie *dress*), längst üblich. Im Deutschen wiederum ist solch eine Zugehörigkeit zu verschiedenen Subkategorisierungsschemata nicht möglich (Berg 2014: 496–498). Dass die Grenzen zwischen Morphologie und Syntax im Englischen weicher sind, lässt sich anhand des Possessivmarkers veranschaulichen: Während im Deutschen das *s*-Suffix an Simplizia und Eigennamen angehängt werden kann, jedoch bei zunehmender Komplexität vermieden wird, kann im Englischen das *s*-Suffix auch an Nominalphrasen mit interner Präpositionalphrase (*the Queen of England's throne*) oder gar integriertem Relativsatz (*the man I know's hat*) herantreten (Berg 2014: 513–515).

In diesem Zusammenhang bildet *-ish* weitere Evidenz für *soft boundaries*. Hierfür spricht zum einen die Tatsache, dass *-ish* an unterschiedlichste Basen herantreten kann: sowohl was die Basiskategorie an sich als auch die Komplexität angeht. Die Eigenschaft von *-ish*, sich nicht auf bestimmte Basen spezialisiert zu haben und keine Präferenz für ausschließlich native Basen oder Fremdbasen aufzuweisen, kann in dieser Hinsicht als Verletzung von Aronoffs (1976) *Unitary Base Hypothesis* angesehen werden, da sich *-ish* eben nicht als ideales Affix verhält, das präferiert an eine Art von Basis herantritt und – zumindest was die Numeralia und die ersten Anzeichen einer weiteren Ausweitung auf z. B. Pronomen angeht – einen klaren Wortartenwechsel induziert. Zum anderen zeigt die Ausweitung auf phrasale Basen, dass es zu einer ähnlichen Interaktion zwischen Syntax und Morphologie kommt wie im beschriebenen Fall des Possessivmarkers. Auch die funktionale Heterogenität zwischen Vergleich und Approximation, die oftmals zu einem hohen Maß an Ambiguität beiträgt, zeugt von einer Grenzdurchlässigkeit, da ein eindeutiges *form-function-mapping* gesprengt wird.

4. Schluss

Wie die vorangegangene empirische diachron-kontrastive Studie von dt./engl. -is(c)h gezeigt hat, lassen sich eklatant divergierende Entwicklungen feststellen – und dies trotz des gemeingermanischen Ursprungs. Zu Buche schlagen hier vor allem die folgenden Kontraste, die die Besonderheiten der jeweiligen Sprache als solche wahrzunehmen erlauben:

- Spezialisierung von dt. -isch auf Fremdbasen vs. zunehmende Basendiversifizierung von engl. -ish
- dt. Persistenz der Zugehörigkeitsfunktion vs. engl. Koexistenz von vergleichender und approximativer Funktion
- kanonisches Suffixverhalten von dt. -isch (Integration mit der Basis zu einem phonologischen Wort, z.T. mit Umlautung) vs. zunehmende Klitisierung von engl. -ish aufgrund komplex-phrasaler Basenausweitung im Zuge von Skopus-Erweiterung

Während die Divergenzen für die Beschreibung der Einzelsprachen von besonderem Interesse sind, liefern die Konvergenzen gleichzeitig einen Erkenntniswert im Hinblick auf sprachübergreifende Prinzipien des Wortbildungswandels. Der vorliegende Fall erlaubt die folgenden Überlegungen: Im Gegensatz zur germanisch noch klar umrissenen Wortbildungsfunktion hat sich die -is(c)h-Derivation in beiden Sprachen über die Jahrhunderte breit aufgefächert: Während im Deutschen im Rahmen der Schwerpunkte Fremdwort-, Eigennamen und negativ besetzter Basen eine starke Diversifizierung der Wortbildungsfunktionen zu beobachten ist, findet im Englischen im Rahmen der beiden funktionalen Makroschemata Vergleich und Approximation eine außerordentliche Erweiterung der möglichen Basisklassen statt. Es wurden also unterschiedliche Pfade eingeschlagen, aber in beiden Fällen lässt sich eine Ausweitung feststellen. Dies scheint ein wiederkehrendes Prinzip in der Entwicklung von Derivationsaffixen zu sein, das sich z. B. darin zeigt, dass jüngere Affixe wie -los/-less tendenziell funktional klarer sind, während ältere (wie -lich, -ig/-y und -isch/-ish) zur Polyfunktionalität bzw. semantischer Ausbleichung neigen. Ob bei Derivationsaffixen generell von Grammatikalisierung gesprochen werden kann, diskutiert Munske (2002). Es sollte allerdings auch darauf hingewiesen werden, dass nicht alle Affixe eine funktionale Ausweitung erfahren müssen (z. B. hat die -bar-Derivation im Lauf ihrer Geschichte ihr funktionales Profil geschärft, s. Flury 1964).

Ein weiterer allgemein wortbildungsrelevanter Aspekt liegt darin, dass die vorliegende Studie die immense Bedeutung unterstreicht, die dem Faktor ‚Vorrat potentiell geeigneter Basiswörter‘ zukommt (vgl. den Begriff *pool of new words* bei Lindsay und Aronoff 2013). Sind native Simplizia (einer Wortart X) erschöpft, bleiben als mögliche weitere Domänen andere Wortarten,

komplexe Basen (bis hin zur phrasalen Ebene) sowie Fremdwörter. Die englische *-ish*-Derivation hat vor allem die ersten beiden Strategien genutzt (wobei Fremdwörter keineswegs tabu sind, nur weniger abgegrenzt als im Deutschen), die deutsche v.a. die letzte Strategie, doch auch zunehmend die komplexe Derivation (auch beides kombiniert: *präwikipedianisch*, *postkatastrophisch*).

Zweifelsohne ist die vorliegende Studie nur ein erster Schritt in der Aufarbeitung eines diachron-kontrastiven Desiderats in der Wortbildungsforschung. Wünschenswert wäre in zukünftigen Arbeiten, eine verstärkte Betrachtung von engl. *-ish* innerhalb des Systems der Derivationsuffixe des Englischen anzustreben, ähnlich wie dies bereits für das Deutsche geleistet wurde (s. die unter 2.1.1 genannten entsprechenden Arbeiten). Interessant wäre etwa, die Konkurrenzbeziehungen von *-ish* zu funktional äquivalenten Suffixen wie beispielsweise *-y* (vornehmlich assoziativ) oder *-esque* (approximativ) empirisch auszuloten und dies wiederum um eine kontrastive Betrachtung zu ergänzen.

Ebenso von hohem Interesse wäre eine Blickerweiterung auf andere Sprachen und Varietäten der Germania, in denen gemeingermanisches **-isko-*fortbesteht und potentiell eigene Entwicklungspfade eingeschlagen haben mag. Vereinzelt liegen Studien zum Norwegischen (Nilssen und Kinn 2017) sowie zum Niederländischen (Heynderickx und van Marle 1994) vor, jedoch läge ein großes Potential darin, solche und weitere Studien zu einer übereinzelsprachlichen Betrachtung zusammenzubringen und dadurch zu überprüfen, inwiefern Variation und Wandel in der Wortbildung durch universelle Tendenzen bedingt sind oder aber durch andere, für die jeweils betroffenen Einzelsprachen charakteristischen Ausprägungen beeinflusst werden.

Referenzen

I. Korpora

- Bonner Frühneuhochdeutsch-Korpus*. Online verfügbar unter <http://www.korpora.org/Fnhd/>.
- British National Corpus* (BNC) 1995 BNC Consortium/Oxford University Computing Services.
- Corpus of Middle English Prose and Verse / Dictionary of Middle English* [CME/MED] (2006).
- Davies, Mark. (2001–2012) *Corpus of American Soap Operas*. Online verfügbar unter <http://corpus.byu.edu/soap/>.
- Davies, Mark. (1990–2017) *Corpus of Contemporary American English* [COCA]. Online verfügbar unter <http://corpus.byu.edu/coca/>.
- Dictionary of Old English Corpus* [DOE] (2000) Centre for Medieval Studies, University of Toronto.
- DWDS. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart. Online verfügbar unter <https://www.dwds.de/>.

- Early English Prose Fiction* (1997) Chadwyck-Healey, Cambridge.
Eighteenth-Century Fiction (1996) Chadwyck-Healey, Cambridge.
GerManC-Korpus. Online verfügbar unter <http://www.ota.ox.ac.uk/desc/2544/>.
Nineteenth-Century Fiction (1999–2000) Chadwyck-Healey, Cambridge.
Wortwarte. Online verfügbar unter <http://www.wortwarte.de/>.

II. Sekundärliteratur

- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge: MIT Press.
- Baayen, R. Harald & Rochelle Lieber. 1991. Productivity and English Word-Formation: A Corpus-Based Study. *Linguistics* 29. 801–843.
- Baayen, R. Harald. 2009. Corpus linguistics in morphology. Morphological productivity. In Herbert E. Wiegand, Merja Kytö & Anke Lüdeling (Hgg.), *Corpus Linguistics*, 899–919. Berlin: De Gruyter.
- Bauer, Laurie. 1983. *English word formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2015. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.
- Berg, Thomas. 2014. Boundary Permeability: A Parameter for Linguistic Typology. *Linguistic Typology* 18(3). 489–531.
- Bergmann, Rolf. 1998. Autonomie und Isonomie der beiden Wortbildungssysteme im Deutschen. *Sprachwissenschaft* 23(2). 167–183.
- Bergmann, Rolf & Dieter Nerius. 1998. *Die Entwicklung der Großschreibung im Deutschen*. 2 Bde. Heidelberg: Winter.
- Best, Karl-Heinz. 2001. Wo kommen die deutschen Fremdwörter her?. *Göttinger Beiträge zur Sprachwissenschaft* 5. 7–20.
- Bosworth, Joseph. 2010. *An Anglo-Saxon Dictionary Online*. <http://www.bosworth-toller.com> (05.10.2016).
- Ciszek, Ewa. 2012. The Middle English Suffix *-ish*: Reasons for Decline in Productivity. *Studia Anglica Posnaniensia* 47(2–3). 27–39.
- Dalton-Puffer, Christiane. 1996. *The French Influence on Middle English Morphology. A Corpus-Based Study of Derivation*. Berlin/New York: Mouton De Gruyter.
- Decke, Angela. 1993. Wortbildung im Dienst der Integration fremdsprachlicher Morpheme (am Beispiel der Adjektive auf *-isc* im Althochdeutschen). In Hans Wellmann (Hg.), *Synchrone und diachrone Aspekte der Wortbildung im Deutschen*, 105–112. Heidelberg: Winter.
- Dixon, Robert M. W. 2014. *Making new words: Morphological derivation in English*. Oxford: Oxford University Press.
- Dornseiff, Franz. 1964 [1921]. Das Zugehörigkeitsadjektiv und das Fremdwort. [Germanisch-Romanische Monatsschrift 9. 193–200.] Reprinted in Franz Dornseiff: *Kleine Schriften*, vol. 1, 221–234. Leipzig: Koehler & Amelang.
- Eichinger, Ludwig M. 1982. *Syntaktische Transpositionen und semantische Derivation. Die Adjektive auf -isch im heutigen Deutsch*. Tübingen: Niemeyer.

- Eitelmann, Matthias, Kari Haugland & Dagmar Haumann. In Vorb. a. From *englisc* to *what-ish*. A Corpus-based Investigation of *-ish* Derivation in the History of English.
- Eitelmann, Matthias, Kari Haugland & Dagmar Haumann. In Vorb. b. *Teletubbyish paintings, youngerish sisters and ish eyes: A Corpus-based Investigation of the Extravagance of English -ish-Derivatives*.
- Fairclough, Norman. 1992. *Discourse and Social Change*. Cambridge: Polity.
- Farrelly, Michael & Elena Seoane. 2012. Democratization. In Terttu Nevalainen & Elizabeth Closs Traugott (Hgg.) *The Oxford Handbook of the History of English*, 392–401. Oxford: Oxford University Press.
- Fleischer, Wolfgang. 1982. *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*. Berlin/Boston: De Gruyter.
- Flury, Robert. 1964. *Struktur- und Bedeutungsgeschichte des Adjektiv-Suffixes -bar*. Winterthur, Keller, Ph.D. Dissertation Zürich.
- Frevel, Claudia & Clemens Knobloch. 2005. Das Relationsadjektiv. In Clemens Knobloch & Burkhard Schaefer (eds.), *Wortarten und Grammatikalisierung: Perspektiven in System und Erwerb*, 151–175. Berlin: De Gruyter.
- Ganslmayer, Christine. 2012. *Adjektivderivation in der Urkundensprache des 13. Jahrhunderts: Eine historisch-synchrone Untersuchung anhand der ältesten deutschsprachigen Originalurkunden*. Berlin: De Gruyter.
- Goetze, Alfred. 1899. Zur Geschichte der Adjectiva auf *-isch*. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 24. 464–522.
- Grimm, Jacob & Wilhelm Grimm (1854–1971) *Deutsches Wörterbuch [DWB]*. Online abrufbar unter <http://dwb.uni-trier.de/de/> [09.02.2018].
- Hay, Jennifer B. & R. Harald Baayen. 2003. Phonotactics, Parsing and Productivity. *Rivista di Linguistica* 15(1). 99–130.
- Henzen, Walter. 1965. *Deutsche Wortbildung*. Tübingen: Niemeyer.
- Heynderickx, Priscilla & Jaap van Marle. 1994. Over het hybride karakter van *-isch*: Op de grens van inheems en uitheems. *Spektator* 23. 229–243.
- Hornbruch, Heike. 1996. *Deonomastika: Adjektivbildungen auf der Basis von Eigennamen in der älteren Überlieferung des Deutschen*. Göttingen: Vandenhoeck & Rupprecht.
- Hotzenköcherle, Rudolf. 1968. Gegenwartsprobleme im deutschen Adjektivsystem. *Neuphilologische Mitteilungen* 69. 1–28.
- Kempf, Luise. 2010. Warum die Unterscheidung fremd–nativ in der deutschen Wortbildung nicht obsolet ist. In Carmen Scherer & Anke Holler (Hgg.), *Strategien der Integration und Isolation nicht-nativer Einheiten und Strukturen*, 123–142. Berlin/New York: De Gruyter.
- Kempf, Luise. 2016a. *Adjektivsuffixe in Konkurrenz. Wortbildungswandel vom Frühneuhochdeutschen zum Neuhochdeutschen*. Berlin/Boston: De Gruyter.

- Kempf, Luise. 2016b. Modeling polyfunctional word formation patterns: A Construction Morphology account of adjectival derivation in the history of German. *SKASE Journal of Theoretical Linguistics* 13(2). 140–163.
- Kempf, Luise. 2017. *Engländisch, Hamburgisch, Lutherisch* – Degrees of onymicity reflected in the history of German *isch*-derivation. *Folia Linguistica* 51(2), 391–417. <https://doi.org/10.1515/flin-2017-0014>.
- Kempf, Luise. Im Erscheinen. Genre influence on word formation (change): A case study on German adjectival derivation. In Richard Jason Whitt (Hg.): *Diachronic Corpora, Genre and Language Change*. Amsterdam: John Benjamins.
- Kempf, Luise und Stefan Hartmann. Im Erscheinen. Schema unification and morphological productivity: A diachronic perspective. In Geert Booij (Hg.): *The construction of words. Advances in Construction Morphology* (Studies in Morphology). Dordrecht: Springer.
- Klein, Thomas. 1994. Althochdeutsch *diutisc* und die Adjektive auf *-isk* im Alt- und Mittelhochdeutschen. In Uecker, Heiko (Hg.), *Studien zum Altgermanischen. Festschrift für Heinrich Beck*. Berlin/New York: De Gruyter, 381–410.
- Klein, Thomas, Hans-Joachim Solms & Klaus-Peter Wegera, Klaus-Peter (Hgg.). 2009. *Mittelhochdeutsche Grammatik. Wortbildung (Teilbd. 3)*. Tübingen: Niemeyer.
- Köbler, Gerhard. 2014a. *Althochdeutsches Wörterbuch*. <http://www.koeblergerhard.de/ahdwbhin.html> (05.10.2016).
- Köbler, Gerhard. 2014b. *Gotisches Wörterbuch*. <http://www.koeblergerhard.de/gotwbhin.html> (05.10.2016).
- Köbler, Gerhard. 2014c. *Altnordisches Wörterbuch*. <http://www.koeblergerhard.de/gotwbhin.html> (05.10.2016).
- Kühnhold, Ingeburg. 1985. Wortbildung des Neuhochdeutschen seit dem 17. Jahrhundert. In Werner Besch, Oskar Reichmann & Stefan Sonderegger (Hgg.), *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung, 1614–1622*. Berlin/New York: De Gruyter.
- Kühnhold, Ingeburg, Oskar Putzer & Hans Wellmann (Hgg.). 1978. *Deutsche Wortbildung. Typen und Tendenzen in der Gegenwartssprache. Teil 3: Das Adjektiv*. Eine Bestandsaufnahme des Instituts für Deutsche Sprache, Forschungsstelle Innsbruck. Berlin: De Gruyter. (=Sprache der Gegenwart 43).
- Kuzmack, Stefanie. 2007. *Isb*: A New Case of Antigrammaticalization?. Paper presented at the meeting of the Linguistic Society of America (LSA). Anaheim, January 4–7.
- Lexer, Matthias. 1872–1878. *Mittelhochdeutsches Handwörterbuch*. 3 Bde. <http://woerterbuchnetz.de/Lexer/>.
- Lindsay, Mark. 2012. Rival suffixes: synonymy, competition, and the emergence of productivity. In Angela Ralli, Geert Booij, Sergio Scalise & Athanasios Karasimos (Hgg.), *Morphology and the Architecture of Grammar. Proceedings of the 8th International Morphology Meeting*, 192–203. Patras: University of Patras.
- Lindsay, Mark & Mark Aronoff. 2013. Natural selection in self-organizing morphological systems. In Fabio Montermini, Gilles Boyé & Jesse Tseng (Hgg.), *Selected Proceedings of the 7th Décembrettes*. München: LINCOM Europa.

- Lyons, John. 1977. *Semantics*. Vol. 1. Cambridge: Cambridge University Press.
- Mair, Christian. 1997. Parallel Corpora: A Real Time Approach to the Study of Language Change in Progress. In Magnus Ljung (Hg.) *Corpus-Based Studies in English*, 195–209. Amsterdam: Rodopi.
- Malkiel, Yakov. 1977. Why *Ap-ish* but *Worm-y?*. In Paul J. Hopper (Hg.), *Studies in Descriptive and Historical Linguistics. Festschrift for Winfred P. Lehmann*, 341–364. Amsterdam: John Benjamins.
- Marchand, Hans. 1969. *The Categories and Types of Present-Day English Word-Formation. A Synchronic-Diachronic Approach*. 2nd ed. München: Beck.
- Mendoza, Raquel Mateo. 2015. Matching Productivity Indexes and Diachronic Evolution: The Old English Affixes *ful-*, *-isc*, *-cund*, and *-ful*. *Canadian Journal of Linguistics* 60(1). 1–24.
- Morris, Lori. 1998. A Toughish Problem: The Meaning of *-ish*. *LACUS Forum* 24. 207–215.
- Motsch, Wolfgang. 2004. *Deutsche Wortbildung in Grundzügen*. Berlin: De Gruyter.
- Munske, Horst Haider. 2002. Wortbildungswandel. In Mechthild Habermann, Peter Müller & Horst Haider Munske (Hgg.), *Historische Wortbildung des Deutschen*, 23–40. Tübingen: Niemeyer.
- Nilssen, Signe & Torodd Kinn. 2017. A chameleon abroad: English *-ish* and *ish* used in Norwegian. *Maal og Minne* 2017. 123–145.
- Norde, Muriel. 2011. Degrammaticalization. In Heiko Narrog & Bernd Heine (Hgg.), *The Oxford Handbook of Grammaticalization*, 475–487. Oxford: Oxford University Press.
- Nübling, Damaris. 2014. Sprachverfall? Sprachliche Evolution am Beispiel des diachronen Funktionszuwachses des Apostrophs im Deutschen. In Albrecht Plewnia & Andreas Witt (Hgg.), *Sprachverfall? Dynamik – Wandel – Variation*, 99–126. (Institut für Deutsche Sprache Jahrbuch 2013). Berlin & Boston: De Gruyter.
- Oxford University Press. 2016. *OED Online*. <http://www.oed.com/> (05.10.2016).
- Pfeifer, Wolfgang. 1993. *Etymologisches Wörterbuch des Deutschen*. [Ergänzte Online-Version]. <http://www.dwds.de/>.
- Plag, Ingo. 1999. *Morphological Productivity. Structural Constraints in English Derivation*. Berlin/New York: Mouton De Gruyter.
- Plag, Ingo. 2006. Productivity. In Bas Aarts & April MacMahon (Hgg.), *The Handbook of English Linguistics*, 537–556. Malden/Oxford: Wiley-Blackwell.
- Sattler, Lutz. 1992. Zur Verwendung von Wortgruppen mit adjektivischem Attribut und Komposita in der deutschen Literatursprache (1570–1730). In Joachim Schildt (Hg.), *Aspekte des Sprachwandels in der deutschen Literatursprache 1570–1730*, 227–271. Berlin: Akademie-Verlag.
- Scherer, Carmen. 2010. Das Deutsche und die dräuenden Apostrophe. Zur Verbreitung von 's im Gegenwartssprache. *Zeitschrift für germanistische Linguistik* 38. 1–24.
- Schlaefel, Michael. 1977. *Die Adjektive auf -isch in der deutschen Gegenwartssprache*. Heidelberg: Winter.

- Schulz, Matthias. 2002. Wortbildung in Wörterbüchern und Texten des 17. Jahrhunderts. In Mechthild Habermann, Peter Müller & Horst Haider Munske (Hgg.), *Historische Wortbildung des Deutschen*, 269–288. Tübingen: Niemeyer.
- Schulz, Matthias. 2007. *Deutscher Wortschatz im 17. Jahrhundert. Methodologische Studien zu Korpus­theorie, Lexikologie und Lexikographie von historischem Wortschatz*. Tübingen: Niemeyer.
- Sugarewa, Tekla. 1974. Adjektivderivate zu Eigennamen und ihre Konkurrenz mit Substantivkomposita und syntaktischen Wortverbindungen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 94. 199–256.
- Sugawara, Ayaka. 2012. Semantics of English Suffix *-ish*. Paper Presented at CLS 48, April 20.
- Szymanek, Bogdan. 2005. The Latest Trends in English Word-Formation. In Pavol Štekauer & Rochelle Lieber (Hgg.), *Handbook of Word-formation*, 429–448. Dordrecht: Springer.
- Thomas, Barbara. 2002. *Adjektivderivation im Nürnberger Frühneuhochdeutsch um 1500: Eine historisch-synchrone Analyse anhand von Texten Albrecht Dürers, Veit Dietrichs und Heinrich Deichslers*. Berlin: De Gruyter.
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65. 31–55.
- Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
- Trips, Carola. 2014. Derivation and historical change. Chapter 22. In Rochelle Lieber & Pavol Štekauer (Hgg.), *The Oxford handbook of derivational morphology* (Oxford handbooks in linguistics), 384–406. Oxford: Oxford University Press.
- Wellmann, Hans. 1997. Wortbildung im Sprachwandel. In Rainer Wimmer & Franz J. Berens (Hgg.), *Wortbildung und Phraseologie*, 65–87. Tübingen: Narr.
- Wilmanns, Wilhelm. 1899. *Deutsche Grammatik: Gotisch, Alt-, Mittel-, und Neuhochdeutsch, Zweite Abteilung: Wortbildung*. 2. Aufl. Berlin: De Gruyter.
- Winkler, Gertraud. 1995. *Die Wortbildung mit -lich im Alt-, Mittel- und Frühneuhochdeutschen*. Heidelberg: Winter.

Dr. Luise Kempf
 Johannes Gutenberg-Universität Mainz
 Dt. Phil.: Historische Sprachwissenschaft
 Jakob-Welder-Weg 18
 D-55128 Mainz
 E-Mail: kempf@uni-mainz.de

Dr. Matthias Eitelmann
Johannes Gutenberg-Universität Mainz
Department of English and Linguistics
Jakob-Welder-Weg 18
D-55128 Mainz
E-Mail: eitelman@uni-mainz.de

Eva Bütthe-Scheider

Verwandtschaft oder Entlehnung? Die Beziehung der ripuarischen Kollektiva mit *ge-* + *-ze* zu den mittelniederländischen mit *ge-* + *-te*

Abstract: A typical characteristic of Central German dialects, especially of the Ripuarian dialect, is that it has collective nouns with *ge-* + *-ze* (cf. *gesteinze*) besides those with *ge-* + *-e* (cf. *gesteine*) corresponding to Dutch *gesteente* and *gestene*. A relationship between *ge-* + *-ze* and *ge-* + *-te* has been assumed for a long time. A corpus-based comparison is given in order to explain the genesis of these different formation types (*ge-* + *-e*, *ge-* + *-ze*, *ge-* + *-te*) and their relations. It seems likely that earlier Dutch formations influenced their Ripuarian counterparts. Rarely, the circumfix *ge-* + *-te* also occurs in Ripuarian texts and may be autochthone. One main result is that the suffix *-ze* in Ripuarian restores the collective formation in the circumfix *ge-* + *-e* when it was destroyed by the *e*-apokope. This is a rare instance where an element of word formation is replaced by another one in order to neutralize the isolation effect of sound change.

1. Einleitung

Eine Besonderheit der mitteldeutschen (md.) Wortbildung zeigt sich in dem vermehrten Auftreten von *ge-* + *-ze* zur Bildung von Kollektiva. Solche Bildungen treten ab der späteren mittelhochdeutschen (mhd.) Zeit besonders im Westmitteldeutschen (Wmd.) in Erscheinung, und dort vor allem im Ripuarischen (Rip.), um das es im Folgenden vornehmlich gehen wird. Das Rip. zählt neben dem Moselfränkischen (Mslfrk.) zur mittelfränkischen (mfrk.) Dialektgruppe, die ihrerseits Teil des Wmd. ist. Das Rip. ist durch eine Zwischenstellung zwischen den hochdeutschen (hd.) und niederdeutschen (nd.)/niederländischen (nl.) Dialekten gekennzeichnet, die alle Ebenen der Sprache betreffen, so auch die Wortbildung, um die es im Folgenden gehen wird.

Bis zum 16. Jh. ist eine Zunahme der Mitglieder der *ge-* + *-ze*-Klasse im Rip. beobachtbar. Sie setzen sich größtenteils bis in die modernen Mundarten hinein fort. Das Rheinische Wörterbuch (RhWb) verzeichnet eine Vielzahl solcher Bildungen (besonders für das Rip. s.u.), aber auch in Wörterbüchern zu anderen md. Mundarten finden sich Belege für solche Bildungen. Sie lassen sich bis in die rezenten Mundarten hinein nachweisen (vgl. für das Neuripuarische (Nrip.) Münch [1904] 1970: 115; RhWb; Wrede 1958: 265–319) und nicht nur im Rip., aus dem Luxemburgischen seien beispielhaft etwa *Gekraids* ‘Gewürz’ (LLU, Bd. 1, Sp. 171), *Geštèng(s)* ‘Gestein’ und *Geštills* n. ‘Gestühl’ (WLM, Bd. 1, Sp. 143) genannt.

Ein Zusammenhang der *ge-* + *-ze*-Bildungen mit den ihnen vergleichbaren nl. und nd. *ge-* + *-tel/-de*-Bildungen ist seit der frühesten Forschung immer wieder hergestellt worden. Die Herkunft von md. *-ze* ist bislang nicht abschließend geklärt. Es wurde bisher schon häufiger vermutet, dass md. *-ze* als eine verhochdeutsche Übersetzung von nl./nd. *-del/-te* entstanden sein könnte. Eine systematische vergleichende Bestandsaufnahme der verschiedenen Bildungstypen ist bislang noch nicht erfolgt. Durch die inzwischen gut zugängliche und aufgearbeitete Quellenlage des Rip. und Mittelniederländischen (Mnl.) ist ein solches Unterfangen derzeit schon möglich.¹ Die Frage ist, ob sich durch einen kontrastiven Blickwinkel weitere Erkenntnisse zur Genese der rip. *ge-* + *-ze*-Bildungen gewinnen lassen. Zunächst wird eine kurze Übersicht über den bisherigen Forschungsstand unter 2.1 gegeben, danach wird das Vorgehen unter 2.2. beschrieben, und die verwendeten Quellen werden vorgestellt. Die Ergebnisse des Vergleichs der rip. *ge-* + *-ze*-Bildungen, die in der mhd. und frühneuhochdeutschen (frnhd.) Zeit in Erscheinung treten, mit entsprechenden Bildungen im Mnl. werden in Kap. 3.1 beschrieben. Im Anschluss wird versucht, eine Einordnung und Erklärung selten bezeugter *ge-* + *-te*-Bildungen im Rip. zu finden (s. Kap. 3.2). In Kap. 4 werden alle Ergebnisse der Analyse einer Schlussbetrachtung zugeführt und zusammengefasst.

2. Vorüberlegungen

2.1. Forschungsstand

Die Erforschung der *ge-* + *-ze*-Bildungen reicht in das 19. Jh. zurück: Bereits Jacob Grimm waren diese Bildungen bekannt, die er unter den starken Neutra auf germanisch (germ.) **-ati* anführt (vgl. Grimm [1878] 1989: 214; vgl. weiterhin auch Grimm [1890] 1989: 526). Die älteste umfänglichere Sammlung mit Belegen dieses Bildungstyps von Bech (1865, 1869 und 1877) zeigt, was seitdem immer wieder zu *ge-* + *ze*-Bildungen bestätigt worden ist, und sich bisher auch durch das Kölner Material weiterhin bestätigen lässt:

Sie entsprechen fast durchweg den niederdeutschen und besonders den niederländischen Wörtern auf *-ete*, *-te*, ja sind ihnen zum Theil wohl erst nachgebildet. Die Sprache bedient sich ihrer hauptsächlich da, wo sie eine Menge, eine Masse, ein Durcheinander bezeichnen will. (Bech 1865: 395)

1 Ein Einbezug des Mittelniederdeutschen (Mnd.) wurde hier noch nicht vorgenommen. Das derzeit im Entstehen befindliche „Referenzkorpus Mittelniederdeutsch/Niederrheinisch (1200–1650)“ (ReN) ist noch nicht abgeschlossen, so dass derzeit nur ein Eindruck anhand einiger Texte gewonnen werden könnte. Nach Abschluss des ReN wird die Untersuchung unter Einschluss der mnd. Sprachdenkmäler fortgeführt werden.

Sowohl zu den nd./nl. Belegen als auch zu den hd. passt eine Herleitung von *-tel/-ze* aus germ. **-it̥ia* statt germ. **-iþ̥ia* (vgl. Bütthe/Wich-Reif 2012: 274), falls es sich um ein altes Suffix handeln sollte. Diese Herleitung wird auch von nl. Seite für die meisten Bildungen angenommen, die im Mnl. mit *ge-* + *-te* gebildet erscheinen (vgl. EWN s.v. *ge-te* Bd. 2, S. 181; Pijnenburg 1976: 29). Daneben gibt es auch vereinzelt im Mnl. als *ge-* + *-te*-Bildung erscheinende Bildungen, die auf germ. **-iþ̥ia* zurückgehen (s.u. Anm. 9).

Im Nrip. überwiegt die Anzahl der deverbalen Bildungen. Diese haben alle eine iterative Bedeutung, z. B. nrip. „*jabomels* Anhängsel, welche baumeln“ (Münch [1904] 1970: 115; s. auch die Vielzahl der deverbalen Bildungen, die Wrede 1958: 265–319 verzeichnet) und sind erst später entstanden als die desubstantivischen Bildungen (vgl. Henzen 1967: 137), was sich auch an ihrer Lautung zeigen ließe. Sie haben nämlich einige Lautwandel nicht vollzogen, die die desubstantivischen hingegen aufweisen. Als Beispiel sei an dieser Stelle etwa auf den fehlenden Umlaut gegenüber der Basis hingewiesen, der den alten desubstantivischen Bildungen als germ. **ȝa*-stämmigen Bildungen lautgesetzlich zukommt. Man vergleiche etwa die folgenden nrip. deverbalen und desubstantivischen Bildungen, die samt Bedeutungsangaben Wrede (1958: 272–274) entnommen sind: *Jeboldersch* ‘Gepolter, Getöse’ zu *boldere*, *Jebruddels* ‘Fehlstiche beim Sticken; Verworrenes, Durcheinander’ zu *bruddele*, *Jebbabbels* ‘kindliches Geplauder; Geklatsch’ zu *babbele* gegenüber *Jeböd(t)s* ‘Einfassung eines Kleidungsstückes durch Band, Borte’ zu *Bood* ‘Borte’, *Jeblöts* ‘das gesamte Blut im Körper als Lebenssaft’ zu *Blut*, *Jeblöms* ‘Vielerlei Blumen’ zu *Blume*. Die deverbalen Bildungen finden im Folgenden in diesem Beitrag keine Berücksichtigung.

Bemerkenswert ist an den hd. *ge-* + *-ze*-Bildungen, dass sie nach bisherigem Forschungsstand erst im 14. Jh. aufkommen und in der Regel die älteren *ge-* + *-e*-Bildungen (< germ. **ga-* + *-ȝa*) neben sich haben, die sie allmählich ersetzen. Gleiches könnte auch für die niederländischen Pendanten gelten: Nicht selten verzeichnet das MNW oder das VMNW eine *ge-* + *-(e)*- neben einer *ge-* + *-te*-Bildung. Aus dem VMNW seien folgende bedeutungsgleiche Bildungen beispielhaft genannt: *ghescoe* neben *ghescoeite* ‘Menge von Schuhen’, *gheberghe* neben *gheberchte* ‘Gebirge’, *ghebeene* neben *ghebeente* ‘Gebein’.

In den ‚Kölner Akten‘ (Stein [1893] 1993 und [1895] 1993), einer großen Quellensammlung aus dem 14. und 15. Jh., hat sich das allmähliche Aufkommen dieser Bildungen auch im Rip. zeigen lassen (zu einzelnen Beispielen vgl. Bütthe/Wich-Reif 2012: 276ff.).

Bei den Wortbildungen mit *ge-* + *-e* ist *ge-* + *-e* diachron als Zirkumfix einzuordnen (vgl. auch Henzen 1967: 137)², wie es schon Wilmanns (1899: 242) formulierte:

Aus dem copulativen Sinn, der der Vorsilbe *ga-* zukommt, konnte sich leicht ein collectiver ergeben, wie wir denn in der älteren Sprache auch solche Collectiva ohne *-ja* finden [...]. Ebenso aber konnte er sich leicht aus dem Begriff der Zugehörigkeit entwickeln, den das Suffix ausdrückt: *givugili* das zum Vogel gehörige, das Vogelartige, alle Vögel. Die beiden Elemente fördern sich also gegenseitig.

Falls *ge-* + *-e* immer durch *ge-* + *-ze/-te* ersetzt worden sein sollte, wäre diese Entwicklung adäquat wohl zunächst als Suffixersatz³ im Zirkumfix zu beschreiben, das anschließend in seiner neuen Gestalt produktiv wird. Dies sei vorab schon bemerkt.

In der Auswertung wird u.a. dann auch die Frage verfolgt, wie das Entstehungsverhältnis dieser Konkurrenzbildungen im Rip. und Nl. der mhd. und frnhd. Zeit zueinander jeweils zu denken ist. Außerdem wird der Versuch unternommen, auf breiter Quellengrundlage (s. Kap. 2.2) den hier referierten Forschungsstand zu überprüfen, das heißt der Frage nachzugehen, in welchem zeitlichen Verhältnis die nl. *ge-* + *-te-* und rip. *ge-* + *-ze-*-Bildungen zueinander in Erscheinung treten und ob sie andere Bildungen voraussetzen.

2.2. Vorgehen und Untersuchungsmaterial

Um das zeitliche Verhältnis der nl. *ge-* + *-te-* und rip. *ge-* + *-ze-*-Bildungen zueinander zu bestimmen, wurde zunächst nach rip. Erstbelegen gesucht. Dabei stellte sich die Frage, ob die Bildungen, die bei Henzen (1965: 137) nach Wilmanns (1899: 365, § 274.4) und Kluge (1899: 36, §71 und 72–73, §144) als mhd. (und dort vorwiegend rip.) beschrieben worden sind, hier einbezogen werden sollten oder nicht, zumal es sich dabei auch um Bildungen handelt, die im vorhandenen Quellenmaterial zum Rip. (s. u.) zum Teil nicht in Erscheinung treten. Da sich bei diesen Bildungen Datierungs- und Lokalisierungsprobleme ergaben, wurden sie außer Acht gelassen. Am Beispiel von mhd. *gewürmze*

2 Aus synchroner Perspektive wird *ge-* + *(-e)* in Klein, Solms und Wegera (2009: 54, § S 44) im Anschluss an Brendel (et al. 1997: 46) als *Präfigierung* bezeichnet. Bei Brendel (et al. 1997: 46) wird jedoch den „20% nicht-apokopierten Belegen“ Rechnung getragen, „indem [...] neben dem Präfix *ge-* ein Zirkumfix *ge-e-*“ angesetzt wird. Gegen dieses Vorgehen zweier Ansätze ließe sich einwenden, dass es nicht strikt synchron ist, durch die Apokopierung wird ein diachroner Lautwandel in die Betrachtung eingeschlossen und zur Kategorisierung genutzt. Da diese Untersuchung diachron ausgerichtet ist, folge ich der o. g. sprachhistorischen Kategorisierung.

3 Ich verwende hier und im Folgenden den Terminus *Suffix* für das Hinterglied des Zirkumfixes.

‘Menge von kriechenden Tieren’, die bei Wilmanns und Kluge genannt sind, soll dies verdeutlicht werden. Bei Wilmanns heißt es z. B.: „Diese neutrale Endung *-eze* wird in der mhd. Zeit, namentlich in Ripuarien, gern gebraucht [...]. Zu Substantiven gehören z. B. *gebeinze*, [...] *gedierze*, [...] *gevogelze*, [...]“ (Wilmanns 1899: 365). Eine Datierung und genaue Lokalisierung der einzelnen Bildungen ist damit nicht gegeben. Das Gleiche gilt auch für Kluge (1899: 36, §71). Dort wird auf die Belegsammlungen von Bech (1865, 1869 und 1877) verwiesen (s. u.), die Belege aus ganz verschiedenen Zeiten und md. Gegenden enthalten. Für die Bildung *gewürmze* fand sich bei Bech (1869: 432) z. B. ein Verweis auf Vilmar (1868: 461–462), der in seinem *Hessischem Idiotikon* die Bildung *gewürmze* unter dem Stichwort *Wurm* nennt. Eine Datierung der Bildung oder ein historischer Beleg findet sich dort aber nicht, so dass das Alter der Bildung ungewiss ist. Zudem ist nicht klar, ob es sie auch im Rip. schon früher gegeben hat. Die Bildungen, die in der beschriebenen Weise genannt werden, finden hier daher keine Berücksichtigung, wenn sich für sie kein historischer rip. Beleg finden ließ, auch wenn die Anzahl der Bildungen dadurch insgesamt geringer ist und eine Bildung auch einmal fehlen kann, weil sie zufällig nicht im betrachteten Material vorhanden ist.

Den Ausgangspunkt für die Untersuchung bildeten die schon in anderen Zusammenhängen betrachteten ‚Kölner Akten‘ (Stein [1893] 1993 und [1895] 1993) und die in ihnen bezeugten *ge-* + *-ze*-Bildungen. Sodann wurde geprüft, ob sich für das Rip. noch frühere Erstbelege für die jeweilige *ge-* + *-ze*-Bildung finden lassen. Dazu wurde mein Korpus zur *e*-Apokope durchsucht (s. Bütthe-Scheider 2017: 38–97), ferner die mfrk. Quellen aus dem Korpus der Mittelhochdeutschen Grammatik (MiGraKo)⁴ und die Wörterbücher ¹DWB⁵, Lexer, MWB, BMZ, FWB, DRW und das RhWb, das selten auch historische Belege bietet. Dabei zeigte sich – das sei bereits hier erwähnt –, dass die Belege in den ‚Kölner Akten‘ (Stein [1893] 1993 und [1895] 1993) jeweils die am frühesten bezeugten sind. Weitere *ge-* + *-ze*-Bildungen aus rip. Quellen, die in den zuvor genannten Korpora und Wörterbüchern bei diesem Arbeitsschritt gefunden wurden, werden auch in die Betrachtung eingeschlossen, da der Vergleich der rip. *ge-* + *-ze*-Bildungen mit den nl. *ge-* + *-te*-Bildungen umso aussagekräftiger wird, je mehr Bildungen in die Betrachtung eingeschlossen werden können.

Die Schriften Hermann Weinsbergs aus dem 16. Jh. sind für Fragen des Wortschatzes sehr ergiebig, weil Weinsberg über Vielerlei – besonders auch Alltägliches – schreibt und sich somit Wörter in seinen Schriften finden, die

4 S. <https://www.linguistics.rub.de/rem/> (verfügbar am 30.06.2017).

5 Die o. g. Belegsammlungen von Bech (1865, 1869 und 1877) wurden im ¹DWB berücksichtigt, so dass sie auf diesem Wege auch hier Berücksichtigung gefunden haben. Der räumliche Schwerpunkt dieser Sammlungen liegt nicht auf dem Rip.

in sonstigen Quellen mangels eines Anwendungsbezuges nicht bezeugt sind. Dies zeigte sich auch bei den hier behandelten Bildungen. Deshalb finden seine Aufzeichnungen als ein sehr spätes Zeugnis auch Eingang in die Betrachtung. Allerdings finden nur diejenigen Textteile seiner Schriften Berücksichtigung, die im Rahmen des DFG-geförderten Projekts „Digitale Erfassung sowie historische und sprachgeschichtliche Auswertung der autobiographischen Aufzeichnungen des Kölner Bürgers Hermann Weinsberg (1518–1597)“⁶ diplomatisch erfasst wurden, um nach Entsprechungen für nl. *ge-* + *te-*-Bildungen zu suchen, die in den vorgängigen rip. Quellen nicht zu finden waren. Ergänzend wurde das Glossar zur Auswahl-edition von Höhlbaum ([1886] 2000, [1887] 2000) und Lau ([1897] 2000, [1898] 2000) zu Rate gezogen. Durch den Einbezug dieser Quelle konnte die rip. Beleg-sammlung um ein paar weitere relevante Bildungen ergänzt werden.

Für das Mnl. ist eine sehr große Anzahl von Quellen durch die Cd-rom Middelnerlands elektronisch verfügbar. Sie enthält die bekannten großen Korpora wie das Corpus Gys-seling. Für die Zeit bis 1300 enthält die Cd-rom die gesamte mnl. Überlieferung und zusätzlich 336 weitere mnl. Quellen, darunter Vers- und Prosatexte.

Bei der Suche nach mnl. Erstbelegen konnte bei *ge-* + *te-*-Bildungen, die es im Nnl. noch gibt, auf das EWN zurückgegriffen werden, das die mnl. Erstbelege verzeichnet. Das EWN stützt seine Angaben auch auf die Quellen, die die Cd-rom Middelnerlands enthält und die auch dem MNW und teils dem VMNW zugrunde liegen. Durch die digital aufbereiteten Quellen konnte im EWN gegenüber dem MNW in einigen Fällen ein früherer Erstbeleg gefunden werden, so dass es häufiger eine zeitliche Differenz zwischen dem Erstbeleg im MNW-Artikel und dem EWN gibt.⁷

Bei einigen alten *ja-*-stämmigen Kollektiva lässt sich beobachten, dass die kollektive Bedeutung in manchen Fällen einer sächlich konkreten gewichen ist (vgl. Henzen 1967: 137, der u. a. *Gemüse* und *Gesimse* als Beispiele für diesen Wandel anführt). Die gleiche Entwicklung betrifft auch einige *ge-* + *-ze-* und *ge-* + *-te-*-Kollektiva. Sofern die hier ausgewerteten Bildungen keine kollektive Bedeutung haben, wurde vor ihrer Aufnahme in die Auswertung geprüft, ob es sich bei ihnen nach herrschender Lehrmeinung in den etymologischen

6 Vgl. <http://www.weinsberg.uni-bonn.de/> (verfügbar am 27.06.2017).

7 Z. B. datiert der früheste Beleg in MNW für mnl. *gewelfte* ‘Gewölbe’ (verzeichnet unter dem Lemma *gewelf*) mit der Schreibung *ghewulfft* auf das Jahr 1488. In einer späteren Alphabetstrecke findet sich mit *ghewulfte* noch ein früherer Beleg aus dem Jahr 1369 im Kontext zu einem Beleg für das Lemma *orbare* ‘(finanzieller) Nutzen, Gewinn’. Nach dem MNW wäre *gewelfte* auf das Jahr 1488 zu datieren, nach dem EWN aber auf das Jahr 1369. In diesem und allen vergleichbaren Fällen wurde die Datierung aus dem EWN gewählt.

Wörterbüchern und den Darstellungen zur historischen Morphologie im Deutschen und Niederländischen sicher um ursprüngliche Kollektiva handelt oder mit hinreichender Sicherheit gehandelt haben kann. Ein Beispiel für letzteren Fall wäre etwa die Bildung mhd. *ge-steiger-el-ze* ‘Gerüst, Tribüne’ mit seiner mnl. Entsprechung *steiger*, die u.a. ‘Leiter’, ‘Gerüst (für Maurerarbeiten)’ und ‘Treppe’ bedeutet (vgl. MNW s.v. *steiger*). Diese Bildungen wurden in die Betrachtung eingeschlossen, da sowohl ein Gerüst als auch eine Leiter eine Vielzahl von Stufen enthält, so dass eine kollektive Grundbedeutung anzunehmen ist. Dieses Beispiel ist zugleich geeignet, eine weitere Besonderheit des Mnl. vorab zu erläutern, die im ausgewerteten Material nur die zuvor genannte Bildung mnl. *steiger* betrifft, darüber hinaus eine sonst häufiger im Mnl. anzutreffende Erscheinung ist: Gelegentlich fehlt im Mnl. das *ge-* bzw. es finden sich Bildungen mit *ge-* + *-te/-e*, die teils auch mit *-te*, aber ohne *ge-* erscheinen. Aus dem MNW kann hier beispielhaft *gebeente* neben *beente* mit gleicher Bedeutung ‘Gebein’ angeführt werden.⁸

Eine weitere Besonderheit des Mnl. betrifft den Wechsel von *<t>* und *<d>* für sowohl germ. **t* als auch germ. **þ*, für den es ganz verschiedene Gründe gibt (vgl. van Loey 1968: 109–115, §§ 114–115), die hier nicht im Einzelnen referiert werden. In diesem Zusammenhang sind nur die wechselnden Dentschreibungen von Bedeutung, die die Fortsetzer des hier vornehmlich behandelten Suffixes germ. **-itja* und seiner Variante germ. **-iþja* betreffen. Für letztere wäre als Dentschreibung lautgesetzlich stets *<d>* zu erwarten, es findet sich aber auch *<t>*. Diese Fälle erklären sich nach van Loey (1968: 110) wie folgt: „In *begeerte*, *lengte*, *veete*, *scaemte* naast *-de* staan we voor suffixverwisseling“ (van Loey 1968: 110) ‘Die Erklärung für *-te* neben *-de* [d.h. bei Bildungen mit ursprünglichem *d*], die sich bei den Bildungen mnl. *begeerte* ‘Begierde’, *lengte* ‘Länge’, *veete* ‘Fehde’, *scaemte* ‘Scham’ zeigt, ist Suffixverwechslung’. Da es in dieser Untersuchung um die Frage geht, ob eine nl. Bildung eine Vorbildwirkung für eine rip. *ge-* + *-ze*-Bildung gehabt haben kann, werden auch die mnl. Bildungen des Typs germ. **ga-* + *-iþja* berücksichtigt, deren Lemma zum Teil im MNW auf *-de* auslautet⁹, wenn für das

-
- 8 Bei mnl. *steiger* ‘Anlegeplatz für Schiffe; Leiter; Gerüst (für Maurerarbeiten); Treppe’ wird es sich wohl um einen ähnlichen Fall wie *beente* ‘Gebein’ handeln. Zwar gibt es für mnl. *steiger* in diesem Falle innerhalb des Mnl. keinen Beleg, der auch *ge-* enthält, dafür aber eine Konkurrenzbildung mit *-te* (s. Kap. 3.1), wie sie auch für andere Kollektiva im Mnl. typisch ist; darüber hinaus hat mnl. *steiger* eine in Teilen semantisch übereinstimmende Entsprechung mit mhd. *ge-steiger-el-ze* im Rip.
- 9 Zudem wird das Lemma älterer **ga-* + *-iþja*-Bildungen in VMNW und MNW in einigen Fällen auch mit *-te* angesetzt, wenn sich ein Suffixwandel vollzogen hatte, z. B. bei *gebuurte* ‘Umgebung, umliegendes Land’ (s. MNW); zu diesem Lemma gibt das VMNW die Information: „suffix *-te* < ouder *-de*“ ‘Suffix *-te* < [aus] älterem

betreffende Lemma auch Belegschreibungen mit <-te> tradiert worden sind. Das ist z. B. der Fall bei mnl. *gemaelde* ‘Gemälde’ im MNW.

Der nachfolgende Vergleich beinhaltet keine Bildungen, die im Mnl. ausdrucksseitig nicht zweifellos interpretiert werden können, weil das Suffix im Mnl. nicht sicher bestimmt werden kann. Das betrifft alle Wörter mit dentalem Stammauslaut, z. B. bei der Bildung mhd. *ge-kriute* ‘Menge von Kräutern, Gras’, die im Rip. irgendwann durch -ze erweitert wurde und sich im Nrip. als *ge-* + -ze-Bildung fortsetzt (vgl. RhWb Bd. 4, Sp. 1438). Ob an diesen dentalen Stammauslaut noch ein dentalhaltiges Suffix angeschlossen wurde, das mit diesem verschmolzen ist, ließe sich für eine entsprechende mnl. *ge-* + -te-Bildung – in diesem Fall gibt es sie nicht, in anderen Fällen aber schon – nicht mit Gewissheit sagen.

3. Empirische Untersuchung

3.1. Auswertungsergebnisse

Beim Vergleich der im Rip. in Erscheinung tretenden *ge-* + -ze-Bildungen zeigte sich, dass nicht in jedem Fall ein nl. Pendant mit *ge-* + -(e)/-te vorhanden ist. Für die im Rip. bezeugten Bildungen mhd. *gebærze* (Kollektivum zu mhd. *bære* ‘Sänfte, (Toten-)Bahre’) und *gesimeze* ‘Gesims’ existiert keine parallele Bildung im Nl., wobei mhd. *gesimeze* als Kollektivum zu mhd. *simez* ‘Sims, Gesims’ keine spezifische rip. Kollektivbildung mit -ze ist, so dass *gesimeze* hier nicht mitgezählt wird. Nur für eine Bildung lässt sich also keine nl. Entsprechung finden, für die übrigen 9 Bildungen sind mnl. Parallelbildungen vorhanden.

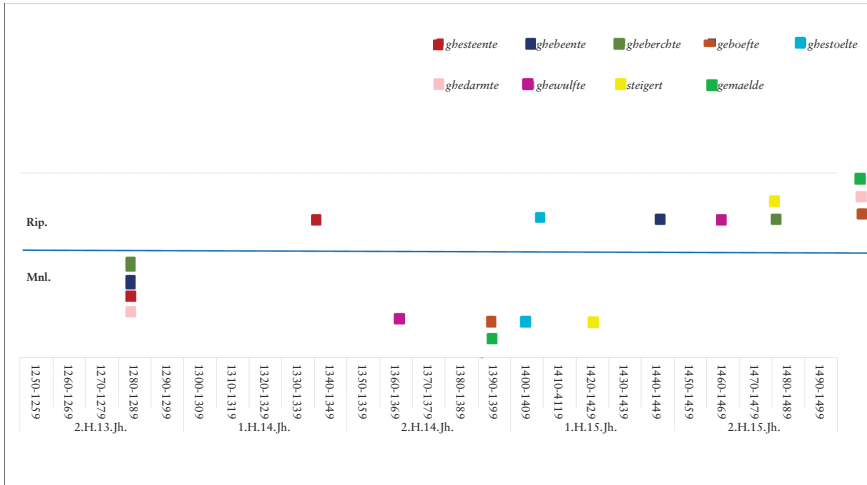
Der Vergleich der Erstbelege dieser Bildungen im Rip. und Mnl. ist in Abb. 1 dargestellt; das heißt, dass Abb. 1 nicht alle im Mnl. existenten mnl. *ge-* + -te-Bildungen enthält, sondern nur diejenigen, die in den betrachteten rip. Quellen bis 1500 und bei Weinsberg eine Parallele haben. Die einander entsprechenden Bildungen sind durch ein farblich jeweils gleiches Quadrat dargestellt. In der Legende wurde der Kürze halber nur die mnl. Form angegeben, die immer *ge-* + -te lautet¹⁰; die rip. Bildung ist in all jenen Vergleichsfällen immer eine *ge-* + -ze-Bildung. Im Einzelnen handelt es sich um die Bildungen mnl. *ghesteente* ‘Gestein’, *ghebeente* ‘Gebein’, *gheberchte* ‘Gebirge’, *geboefte* ‘Menge von Schurken’ (Ableitung von mnl. *boef* u.a. ‘Schurke’ ~ mhd. *buobe* ‘Knabe; zuchtloser Mensch’), *ghestoelte* ‘Gestühl’, *ghedarmte* ‘Gedärm’, *ghewulfte* ‘Gewölbe’, *steigert* u.a. ‘Leiter’, *gemaelde* ‘Gemälde’. Ist ein Quadrat weit rechts außerhalb der Zeitachse eingezeichnet, bedeutet dies, dass es sich

-de’ (s.v. *ghebuurte*). Dieser Fall betrifft jedoch keines der mnl. Lemmata, die in der Auswertung behandelt werden.

10 Eine Ausnahme bildet hier nur *steigert*, s. dazu Anm. 8.

um eine erst spät bei Weinsberg bezeugte Bildung handelt. Es handelt sich um eine vereinfachte Darstellung, die einen groben zeitlichen Überblick vermitteln soll, die exakten Datierungen sind der später folgenden Tab. 1 zu entnehmen.

Abb. 1: Erstbelege der am frühesten im Rip. bezeugten *ge-* + *-ze-*Bildungen im Vergleich zu den ihnen entsprechenden *nl. ge-* + *-te-*Bildungen.



Das Ergebnis der Gegenüberstellung in Abb. 1 lässt sich auf einen einfachen Nenner bringen: Wenn einer *rip. ge-* + *-ze-*Bildung eine *mn. ge-* + *-te-*Bildung entspricht, datiert der *mn.* Erstbeleg früher, häufig sogar mit großem zeitlichem Abstand. Durch den hier durchgeführten Vergleich bestätigt sich somit die Einschätzung, die schon in der älteren Forschung zu diesem Thema vertreten wurde (s.o. in Kap. 2.1).

Nun könnte das Ergebnis zu den Datierungen und den bezeugten Bildungen vielleicht dem Zufall der Überlieferung geschuldet sein. In einem weiteren Schritt wäre daher zu fragen, ob es die später bezeugten *rip. ge-* + *-ze-*Bildungen, denen im *Mnl. ge-* + *-te-*Bildungen entsprechen, in den Quellen ab dem 13. Jh. auch als *mhd. ge-* + *-e-*Bildungen (< germ. **ga-* + *-ja*) gegeben hat, die im *Hd.* sonst üblicher sind und weiter verbreitet waren (vgl. Wilmanns 1899: 242–244). Hierzu wurden die Erstbelege für die Vergleichsbildungen mit *mhd. ge-* + *-e-* zu den in Abb. 1 genannten *rip. ge-* + *-ze-*Bildungen ermittelt und in diesem Zuge auch die *mn. ge-* + *-e-*Bildungen mit ihren Erstbelegen. Das Ergebnis dieses Vergleichs ist in der folgenden Tabelle 1 dargestellt, die ein weiteres Lexem aus einer sekundären Quelle einschließt, für das ein Beleg über das *RhWb* erschlossen werden konnte.

Tab. 1: Das Verhältnis der *rip. ge- + -ze-* und *ge- + -e-*-Bildungen untereinander und im Vergleich zu den *mnl.* Bildungen.

Quelle	Lemma mnl.	Datierung	Lemma mhd.	Quelle und Datierung
VMNW	ghebeene	1200	gebeine	2. Drittel 12. Jh. (RBib)
VMNW	ghebeente	1285	gebeinze	1448 (Kölner Akten)
VMNW	gheberghe	1240	gebirge	1. V. 13. Jh. (RhTun)
VMNW	gheberchte	1285	gebirgze	2. H. 15. Jh. (Kais)
-	-	-	gebüebe	Lexer + MWB + FWB
MNW	geboefte	ca. 1390	gebüebze	2. H. 16. Jh. (Weinsberg, LD)
-	-	-	gederme	2. H. 16. Jh. (Weinsberg, LD)
VMNW	ghedarmte	1285	gedermze	2. H. 16. Jh. (Weinsberg, LD) + <i>nrip.</i> vorhanden
-	-	-	gemæle	Lexer + MWB + FWB
MNW	<i>gemaelte</i> s.v. <i>gemaelde</i> ¹¹	1391–1400	gemælze	2. H. 16. Jh. (Weinsberg) [nach MWB schon 2. H. 14. Jh. in Pil (V. 7998)]
-	-	-	-	-
MNW	geraemte ¹²	1340–1379 ¹³	geremze	RhWb (Bd. 7, Sp. 49): „ <i>gereymptze</i> Köln 1352“ (ed. Ennen (1878: 391))
MNW	gestene	15. Jh.	gesteine	2. H. 13. Jh. (PLilie)
VMNW	ghesteente	1285	gesteinze	1341 (Kölner Akten)
EWN	steiger	1270	gesteiger	1410 (Kölner Akten)
MNW	<i>steigert/</i> <i>steigaert</i> s.v. steiger	1451–1500	gesteigerze	1482 (Kölner Akten)
-	-	-	gestüele	DRW (Bd. 4: 593) + FWB (Bd. 6, Sp. 1667): 1401 (ed. Grimm 1840: 667)

11 Zu der nicht existenten Form *mnl. gema(e)l(e)* ließe sich anmerken, dass *mnl. gemael* mit anderer Bedeutung ‘Gemahlenes’ als Ableitung von *mnl. malen* ‘zerreiben; zerkleinern; mit einer Mühle zermahlen’ schon existiert (s. MNW). Ob manche *nl.* Bildungen durch andere, schon vorhandene Bildungen blockiert worden sind, bedürfte noch einer eigenen Untersuchung.

12 Kollektivum zu *mnl. raem* ‘Umrandung, Einrahmung, Rahmen’, *mhd.* entspricht *rame* mit ähnlicher Bedeutung.

13 Der Text datiert wohl früher: „Dl. I, blz. 3–68: Liber Albus — De oudste codificatie van het stadrecht van vóór 1340 [...] Hs.: Utrecht, Sted. arch., geschreven in 1340 en vervolgd tot 1379.“ (http://gtb.inl.nl/iWDB/search?wdb=MNWBRONNEN&actie=article&cid=s1144_sub000, verfügbar am 23.06.2017).

Quelle	Lemma mnl.	Datierung	Lemma mhd.	Quelle und Datierung
MNW	gestoelte	1408	gestüelze	1410 (Kölner Akten)
EWN	gewelf	1369	ge-welbe	1396 (Kölner Akten)
EWN	<i>gewulft</i> s.v. gewelf	1370	ge-welbze	1467 (Kölner Akten)

Bei den heller schattierten Einträgen in der Spalte für das mhd. Lemma (s. Tab. 1) gibt es eine *ge-* + *-e*-Bildung jeweils im Mhd., aber nur Belege aus anderen hd. Regionen. In diesen Fällen ist es durchaus wahrscheinlich, dass es das Wort als *ge-* + *-e*-Bildung auch im Rip. gegeben hat, es aber zufällig nicht in einer älteren rip. Quelle bezeugt ist. Bei einigen der betrachteten Bildungen in Tab. 1 gibt es allerdings im Rip. zu den *ge-* + *-ze*-Bildungen früher bezeugte *ge-* + *-e*-Bildungen. Auf die knapp überwiegende Mehrheit der in Tab. 1 betrachteten Bildungen trifft die Aussage zu, dass es zu der betreffenden rip. *ge-* + *-ze*-Bildung eine ältere *ge-* + *-e*-Bildung gibt¹⁴, so dass sich der Eindruck eines Suffixersatzes im Rip. einstellt. Dieser Eindruck wird durch den Umstand verstärkt, dass sich Belege dafür finden lassen, dass einer *ge-* + *-ze*-Bildung eine *ge-* + *-e*-Bildung in einer früheren Bearbeitungsstufe des gleichen Textes entspricht. Auf einen solchen Fall wurde schon in Büthe/Wich-Reif (2012: 278) mit einem Beleg für mhd. *gebirge* in der Sächsischen Weltchronik (SW) der rip. Hs. mgq 284 (fol. 4va,14) hingewiesen, die den Vorlagentext für den SW-Teil der Kölner Prosa-Kaiserchronik (Kais) bildet, in der sich im parallelen Text ein Beleg für mhd. *gebirgze* (fol. 1ra,24) findet.

Für die Ausgangsfrage bedeutet das Ergebnis in Tab. 1, dass *ge-* + *-ze*-Bildungen im Rip. bereits viel früher hätten in Erscheinung treten können – das gilt besonders für die im 13. Jh. schon bezeugten Pendanten mit *ge-* + *-e*. Dieser Befund legt die Vermutung nahe, dass es sie in dieser Zeit noch nicht gab und sie erst im 14. Jh. entstanden sind, so dass also der Zeitpunkt ihres Auftretens nicht weit vom Zeitpunkt ihrer Entstehung entfernt liegen wird.

14 In Klein, Solms und Wegera (2009: § 42–48) wird zur sprachräumlichen Verteilung der Belege für *ge-* + *-e*-Kollektiva nicht explizit etwas zum Mfrk. gesagt, die Aussage, dass die „zeitliche und sprachräumliche Verteilung sowie die Verteilung auf Textsorten [...] unauffällig“ sei (Klein, Solms und Wegera 2009: 56, § S 48), deutet darauf hin, dass sich die 487 Tokens für die 40 bezeugten Types (Klein, Solms und Wegera 2009: 53, § S 42) etwa gleichmäßig über das Korpus der Mittelhochdeutschen Grammatik (= MiGraKo) verteilen. Eine eigene Korpussuche ergab für die mfrk. Texte des 13. Jh.s aus MiGraKo insgesamt 26 Tokens für die 8 unzweifelhaften Types mhd. *ge-birge*, *ge-kleide* ‘Kleidung’, *ge-kriute*, *ge-müese* (Kollektivum zu mhd. *muos* ‘breiige Speise, Essen’), *ge-steine*, *ge-stirne* ‘Menge von Sternen’ (Kollektivum zu mhd. *stërne*), *ge-würme* und *ge-zimber*.

Ein Unterschied zwischen dem Rip. und dem Mnl. zeigt sich im Verhältnis von nl. *-te* und *-e* – das in der Regel schon apokopiert wurde, so dass die mnl. Lemmata in vielen Fällen nicht mehr auf *-e* auslauten; dessen ungeachtet schreibe ich *-e*, um diese Bildungen von jenen auf *-te* zu unterscheiden – im Vergleich zu rip. *-ze* und *-e*: In vier Fällen gibt es im gesamten Nl. kein Pendant auf *-e* oder ehemals *-e* zu einer Bildung mit *-te*. Da vier solcher Fälle nicht sehr aussagekräftig sind, wurden weitere *ge-* + *-e* und *ge-* + *-te*-Bildungen aus den mnl. Wörterbüchern exzerpiert (s. Tab. 2). Die rip. oder sonstigen hd. Bildungen sind in Tab. 2 ergänzt, sofern dies möglich war. Das Lemma erscheint in der Spalte „Lemma mhd.“ stets in einer helleren Graustufe, wenn kein historischer rip. Beleg für das Lemma gefunden werden konnte.

Tab. 2: Das Verhältnis weiterer mnl. *ge-* + *-te-* und *ge-* + *-e-*-Bildungen untereinander und im Vergleich zu den rip. Bildungen.

Quelle	Lemma mnl.	Datierung	Lemma mhd.	Quelle und Datierung
-	-	-	ge-böume	Lexer, MWB
VMNW	gheboomte ¹⁵	1285	?	
VMNW	ghescoe	1240	geschüehze	Lexer, MWB (s.v. <i>geschuohe</i>)
VMNW	ghescoeite	1291–1300	geschüehze	RhWb: nrip.
-	-	-	gestirne	2. H. 13. Jh. (VLilie)
MNW	gesternte	1301–1325	gestirnze	Bech (1865: 397)
-	-	-	gewürme	2. H. 13. Jh. (VLilie)
MNW	gewormte	1390–1410?	gewürnze	Bech (1869: 432) ¹⁶ : im rezenten Hessischen vorhanden + nrip. (s. Münch [1904] 1970: 115)
-	-	-	getiere	¹ DWB, aber nicht rip.
MNW	gedierte ¹⁷	1399	getierze	¹ DWB: md. 1479, nrip. vorhanden (s. Münch [1904] 1970: 115)
-	-	-	gevocele	Lexer, FWB: ostmitteldeutsch 1338; aber nicht rip.

15 ‘Menge von Bäumen’; Kollektivum zu mnl. *boom* ‘Baum’ (~ mhd. *boum*).

16 Mit Verweis auf Vilmar (1868: 461–462).

17 ‘Menge aller vierfüßigen Tiere’.

Quelle	Lemma mnl.	Datierung	Lemma mhd.	Quelle und Datierung
EWN	ghevogelte ¹⁸	1287	gevogelze	ja, laut Henzen (1965: 139), nrip. vorhanden (s. Münch [1904] 1970: 115)
MNW	getimmer ¹⁹	1343–1371	ge-zimbere	1328 (Kölner Akten)
MNW	getimmerte	1301–1325	?	

Tab. 2 zeigt, dass es weitere Beispiele für den Fall gibt, dass einer nl. *ge-* + *-te*-Bildung keine ältere *ge-* + *-e*-Bildung entspricht: Dies betrifft 5 von 7 betrachteten Bildungen. Sofern es zu den betrachteten Bildungen ein rip. Pendant gibt, handelt es sich dabei stets um eine *ge-* + *-e*-Bildung. Der Vergleich der Bildungen in Tab. 1–2 lässt daher die folgenden Schlussfolgerungen zu:

1. im Rip. setzen *ge-* + *-ze*-Bildungen überwiegend ältere *ge-* + *-e*-Bildungen voraus, so dass
2. die Beschreibung als Suffixersatz im Zirkumfix adäquat sein dürfte. Im Vergleich zum Rip. zeigt sich für das Nl., dass
3. nl. *ge-* + *-te* viel früher als rip. *ge-* + *-ze* in Erscheinung tritt und viele nl. *ge-* + *-te*-Bildungen ein sehr hohes Alter haben,
4. die nl. *ge-* + *-te*-Bildungen zahlreicher sind als die rip. *ge-* + *-ze*-Bildungen,
5. den nl. *ge-* + *-te*-Bildungen – anders als im Rip. – vielfach keine *ge-* + *-e*-Bildungen vorausgehen.

Eine Vorbildwirkung der nl. *ge-* + *-te*-Bildungen für die rip. *ge-* + *-ze*-Bildungen – wie sie bisher schon erwogen wurde – ist daher gut möglich.

Im Vergleich zu den nl. und rip. Bildungen lässt sich noch aus systematischer Perspektive ergänzen: Nicht jedes mnl. *ge-* + *-e*-Kollektivum erscheint im Mnl. auch als *ge-* + *-te*-Bildung. Gleiches gilt auch für das Rip. (s. die Beispiele in Tab. 3).

18 ‘Menge aller Vögel’.

19 ‘Wohnung, Gebäude’; Kollektivum zu mnl. *timmer* ‘Zimmer’ (~ mhd. *zimber*).

Tab. 3: Beispiele für Kollektiva mit *ge-* + *-e*, die es nicht auch als *ge-* + *-te/ge-* + *-ze*-Bildung gibt.

Quelle	Lemma mnl.	Datierung	Lemma mhd.	Quelle und Datierung
VMNW	ghestubbe ²⁰	1240	gestüppe	Lexer, ¹ DWB, [RhWb: nrip. <i>Stübb</i> ²¹]
-	-	-	-	-
VMNW	ghegare ²²	1285	geger	14.-15. Jh. (Kölner Akten)
-	-	-	-	-

Bei der Beurteilung des Nebeneinanders von nl. *ge-* + *-te* und *ge-* + *-e* gehörte zu einer abschließenden Bewertung noch die Klärung der Frage, ob es eine sprachräumliche Präferenz der Varianten gibt. Ersten Beobachtungen zufolge, die an dem wenigen mnl. Sprachmaterial gemacht werden konnten, das hier einbezogen wurde, stammen die *ge-* + *-te*-Bildungen meist aus Flandern, Brabant oder Holland; in diesen Gebieten finden sich andererseits aber auch *ge-* + *-e*-Bildungen. Dieser Punkt bedürfte also noch einer eigenen Untersuchung.

Ein letzter Unterschied zwischen den nl. *ge-* + *-te*-Bildungen und den rip. *ge-* + *-ze*-Bildungen, der die Frage nach der Herkunft der *ge-* + *-ze*-Bildungen zwar nicht berührt, aber in systematisch-vergleichender Perspektive erwähnenswert ist, betrifft die weitere Genese dieser Bildungen, der schon in EWN (Bd. 2, S. 181) vermerkt wurde:

Het huidige verschil tussen Duits *ge-e* (met *i*-umlaut) en Nederlands *ge-te* lijkt dus niets meer te zijn dan het gevolg van generalisatie, die in het Duits en Nederlands toevallig net andersom uitpakte. ‘Der heutige Unterschied zwischen deutsch *ge-e* (mit *i*-Umlaut) und niederländisch *ge-te* scheint nichts anderes zu sein als die Folge einer Verallgemeinerung, die im Deutschen und Niederländischen zufällig anders herum verlief’.

Die *ge-* + *-ze*-Bildungen haben im Hd. keinen Eingang in die Standardsprache gefunden – was wohl durch die vergleichsweise geringe sprachlandschaftliche

20 ‘Staub’ (zu mnl. *stof* ‘Staub’).

21 Nrip. *Ge-stub(s)* (s. RhWb, Bd. 8, Sp. 901) ist eine Ableitung vom Lehnwort *stuben* < spätlat. **extupare*, *-fare* (s. RhWb, Bd. 8, Sp. 900) und gehört somit nicht hierher. Nrip. *Gestübbels* mit der Bedeutung „grober Staub, Abfall von gesägtem Holz, von Gemüse udgl., Holzsplitter, ganz kleines Brennholz“ (RhWb: Bd. 8, Sp. 893) ist keine direkte *ge-* + *-ze*-Bildung zu mhd. *gestüppe*, vielleicht aber weiträumig mit zu den nrip. deverbale Bildungen des Typs nrip. *jabomels* ‘Anhängsel, welche baumeln’ zu zählen.

22 ‘Kleidung; Waffenrüstung’.

Verbreitung dieser Bildungen sowie den geringen Einfluss der wmd. Dialekte auf die Standardsprache zu erklären ist –, im Nl. aber schon. Bei der Beseitigung der Varianten der konkurrierenden Stammbildungen ist im Nl. die Wahl häufiger – aber nicht immer – auf *ge-* + *-te* gefallen und die entsprechende Bildung mit *ge-* + *-e* ist in diesen Fällen untergegangen. So kennt das rezente Niederländische standardsprachlich beispielsweise weder *gesteen*, *geberg* noch *gebeen*, sondern nur *gesteente*, *gebergte* und *gebeente*.

Eine abschließende Frage betrifft nun noch den eingangs genannten Aspekt, ob für die rip. *ge-* + *-ze*-Bildungen auch eine authochtone Entstehung denkbar ist. In diesem Zusammenhang soll der Blick auf einige Belege mit *ge-* + *-te* statt *ge-* + *-ze* gelenkt werden, die auch im Rip. vereinzelt bezeugt sind. Sie werden im Folgenden Abschnitt 3.2 behandelt, in dem es auch um die Frage gehen wird, ob diese Bildungen als Entlehnungen oder authochtone Bildungen einzustufen sind.

3.2. Rip. *ge-* + *-te*-Bildungen?

Etwa zeitgleich zum Auftreten der ersten *ge-* + *-ze*-Bildungen im Rip. lassen sich auch vereinzelte Belege für *ge-* + *-te* finden. Der früheste Beleg dieser Art stammt im untersuchten Material aus dem *Buch der Minne* aus MiGrako²³ (s. (1)).

- (1) *An deme nuffchelen sint dru dinc. Dat golt. inde dat **ge=steinte**. īde dat meisterliche werc* (BuMi 62v,14–16) ‘An der Mantelschnalle sind drei Dinge zu bemerken: das Gold und das **Gestein** sowie das meisterhafte Werk [d.h. die meisterhafte, kunstfertige Verfertigung selbst]’.

Mhd. *gesteine* mit *ge-* + *-te* ist in BuMi nur an dieser einen Stelle bezeugt, weitere *ge-* + *-ze*-Bildungen oder *ge-* + *-te*-Bildungen sind dort nicht vorhanden, was die Deutung dieser Stelle erschwert. Wie sich *ge-* + *-te* in BuMi erklärt, ist aufgrund der Textgeschichte und Überlieferung nicht zweifelsfrei zu klären. Grundsätzlich lässt sich zur Überlieferung von BuMi sagen: „Der Traktat ist frühzeitig nach dem niederländischen Nordwesten gewandert und ist dort breiter überliefert als im Deutschen.“ (Ruh 1993: 233). BuMi datiert in die 1. Hälfte des 14. Jh.s und vergleicht sich schreibsprachlich dem *Rheinischen Marienlob* und der *Lilie* aus dem 13. Jh. (vgl. Honemann 1989: 1063, Klein 1982: 8), für die z. T. der gleiche Verfasser angenommen wurde (vgl. Ruh 1993: 243f.). Außer bei *t* in *ge=steinte* sind in BuMi weitere Lautverschiebungsausnahmen anzutreffen, die auf nördlichere Überlieferungszusammenhänge hindeuten könnten, zumal der Erzählstoff für BuMi wohl aus

23 Referenzkorpus Mittelhochdeutsch (= RefM), Nr. M404 (<https://www.linguistics.rub.de/rem/>, verfügbar am 24.05.2017).

dem Französischen stammt und über das Mnl. vermittelt wurde.²⁴ Einige der Lautverschiebungsausnahmen, die in BuMi vorkommen, sind allerdings auch im Mfrk. bodenständig und möglich, z. B. gibt es bei Komparativen auch Belege mit unverschobenem germ. *t wie etwa <aller fütifte> ‘aller süßeste’ (BuMi 078v,13.05) neben Belegen mit Lautverschiebungsfrikativ <aller füzifte> (BuMi 064v,14.05) (vgl. hierzu schon Klein 1982: 8 mit weiteren Beispielen und Verweis auf Lerchner 1971: 212f., 262). Schreibsprachlich zeigt BuMi nach Klein (1982: 8) insgesamt Züge des nördlichen Mfrk., die außer im Rip. eventuell auch im nördlichen Mslfrk. angetroffen werden könnten. Die meisten LV-Ausnahmen, die in BuMi bezeugt sind, sind heute noch im Rip. gebräuchlich, aber nicht alle. Nach Lerchner (1971: 224ff.) und RhWb (Bd. 7, Sp. 663) wäre etwa <faken> ‘Sachen’ (BuMi 067v,09.06) mit germ. *-k- für mhd. *sachen* (Dat. Pl.) nicht zu den rezent im Rip. auftretenden LV-Ausnahmen zu zählen. Dennoch muss man sich angesichts dieses Belegs für die Form

24 „Ons hs., de Mnl. vertaling van een Mhd. tekst laat zien hoe Duitsland niet slechts door bemiddeling van Brabant de Franse mystieke werken importeerde, maar deze ook in nieuwe, of althans vernieuwde gedaante teruggaf.“ (Willeumier-Schalij 1946: XVII) ‘Unsere Handschrift – die mnl. Übersetzung eines mhd. Textes – zeigt, dass Deutschland nicht nur durch Vermittlung Brabants die französischen mystischen Werke importierte, sondern diese auch in neuer oder zumindest erneuerter Gestalt wieder zurückgab’. Die mnl. Textzeugen werden als Übersetzungen des hier behandelten Textzeugen BuMi angesehen: „De Middelnederlandse vertaler is zeer nauwgezet en met overleg te werk gegaan en wel zo nauwkeurig, dat wij bijna van een letterlijke vertaling kunnen spreken [...] Het Duits heeft zijn sporen [...] vooral in de spelling achtergelaten.“ (Willeumier-Schalij 1946: LXXVIII) ‘Der mnl. Übersetzer ist sehr genau und mit Bedacht vorgegangen und sogar so präzise, dass wir fast von einer wörtlichen Übersetzung sprechen können [...] Das Deutsche hat seine Spuren vor allem in den Schreibungen hinterlassen’. Allerdings heißt es mit Bezug zur mnl. Übersetzung auch: „Zeer zelden geeft de Middelnederlandse vertaling meér dan de tekst uit hs. P [= BuMi, Anm. d. Verf.], maar het is niet uit te maken of dit uit het voorbeeld is overgenomen, dan wel of het eigen toevoegingen van de Middelnederlandsen copïist zijn.“ (LXXIX) ‘Sehr selten enthält die mnl. Übersetzung mehr als den Text aus Hs. P [= BuMi], aber es ist nicht zu entscheiden, ob diese aus der Vorlage übernommen wurden oder ob es eigene Zusätze des mnl. Kopisten sind’. Mit Bezug zu einigen Stellen, die im mnl. Text abweichend vom Hd. lateinischsprachig gesetzt wurden, vermutet Willeumier-Schalij, dass der mnl. Bearbeiter eine weitere Vorlage zur Verfügung hatte, in der bereits diese lat. Wörter standen (vgl. Willeumier-Schalij ebd.). Ob diese Vorlage mnl. gewesen sein soll, bleibt offen. Es heißt nur mit Bezug zu BuMi: „De Duitse afschrijver had óf puriteinse neigingen óf hij copieerde met het oog op een minder ontwikkeld publiek, dat geen Latijn kende.“ (ebd.) ‘Der deutsche Abschreiber hatte entweder eine sprachpuristische Neigung oder er kopierte für ein weniger gebildetes Publikum, das kein Latein verstand’.

gesteinte in BuMi fragen²⁵, ob *ge-* + *-te* auch im Rip. vorkam und ob sich weitere Hinweise diesbezüglich finden lassen.

In den Kölner Akten gibt es weitere Belege für Bildungen mit *ge-* + *-tel-de* statt *ge-* + *-ze*, auf die schon in Bütthe und Wich-Reif (2012: 277) hingewiesen wurde, einer dieser Belege ist unter (2) genannt sowie weitere Belegstellen mit Varianten unter (2)-(4).

- (2) *Van den slusselen van dem gewulfe, daynne de privilegien sint. Vort solen dye drij slussele van dem gewulfde, [...] bevoylen sijn drijn birven mannen, die unse heren vanme raide dartzo dünckent nützlich syn* (Stein ([1893] 1993: 216,10–15) ‘Von den Schlüsseln des **Archivs**, worin sich die Privilegien befinden. Weiterhin sollen die drei Schlüssel von dem **Archiv** [...] drei tüchtigen Männern befohlen sein, die unsere Herren vom Rad dazu für tauglich befinden’.
- (3) *Van offenongen des gewulfs. Unse heren v. r. haint verdragen ind bevoylen den gewulffheren, dat sij myt dem doctoir ind prothonotorem bynnen desem nyesten jaire, so vill yn moegelijk sall syn, in dat gewulffs soelen gaen, dat zo reformyeren ind zo besseren ind die privilegien [...] ind schriften zo betzeychenen ind in kisten ind laden zo legen ind zo verwaeren [...]* (Stein [1895] 1993: 569,24–30 u.ö.) ‘Von Öffnungen des **Archivs**. Unsere Herren vom Rad haben übereinkommend festgesetzt und den Gewölbeherren befohlen, dass sie mit dem Doktor und Proto notar innerhalb des nächsten Jahres, so oft [es] ihnen möglich sein wird, in das **Archiv** gehen sollen, um dieses zu reformieren [i.e. umzugestalten] und auszubessern und die Privilegien, [...] und Schriften zu bezeichnen und in Kisten und Laden zu legen und zu verwahren’; [...] *dat dan zwene van yn mit dem doctoir oder prothonotaren in dat gewulfs gaen, upind zosliessen moegen, solange dese reformacie duyren sall, [...]* (Stein [1895] 1993: 570,05-07) ‘dass dann zwei von ihnen mit dem Doktor und Protonotar in das **Archiv** gehen, auf- und zuschließen dürfen, so lange diese Umgestaltung dauern wird’.
- (4) *Item sal he mit den gewulffmeistern in dat gewulffe gaen [...], so wanne des noit geburt [...]* (Stein [1895] 1993: 211, 20–22)²⁶ ‘Weiterhin soll er mit den Gewölbeameistern in das **Archiv** gehen [...], wann immer es nötig ist’.

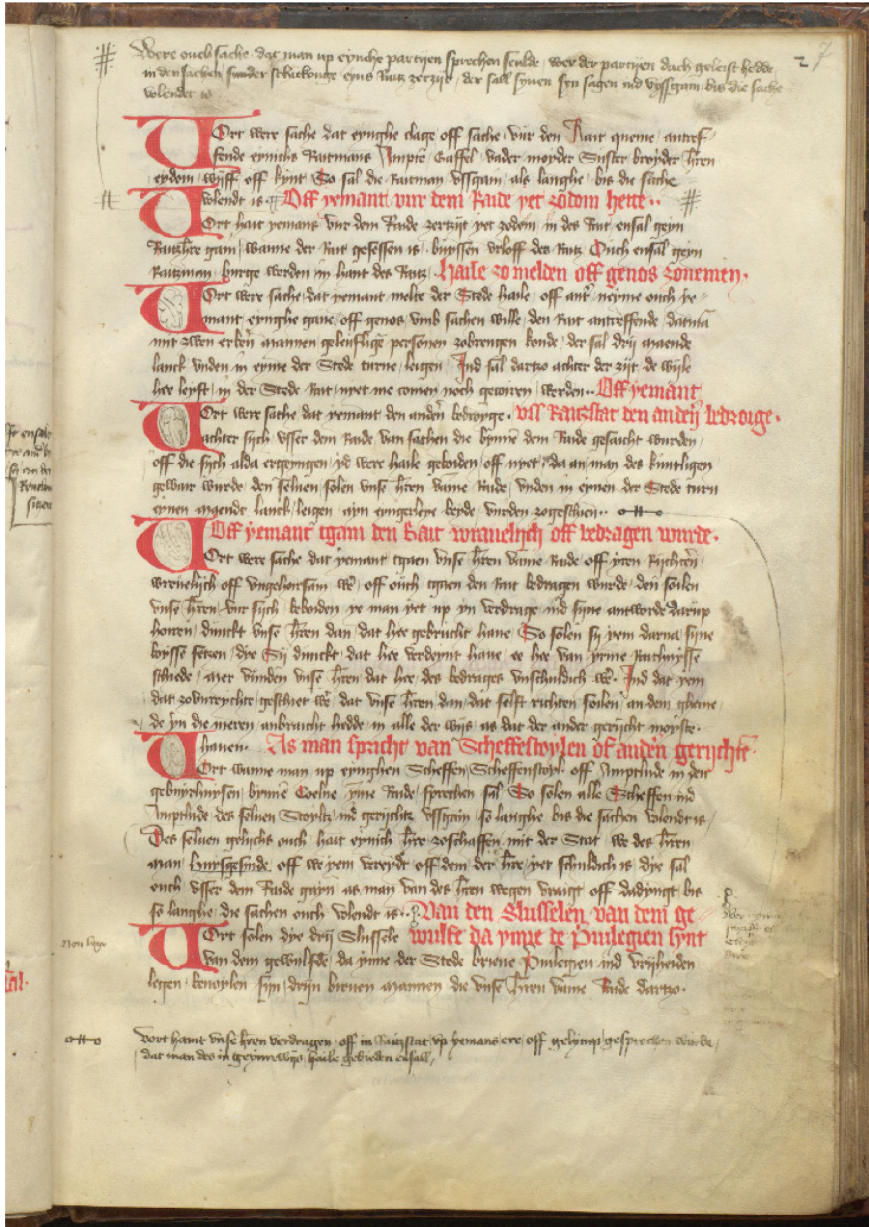
25 Auch in der Koelhoffschen Chronik scheint es *gesteinte* neben *gesteinze* zu geben (vgl. Cardauns 1876: 330,30; 417,30; 419,11; 456,13). Wie sich das Nebeneinander dort erklärt, bedürfte noch weiterer Untersuchung.

26 Ausschnitt aus: „Vertrag mit Heinrich Vront über seine Thätigkeit als Rath der Stadt Köln 1410“ (Stein [1895] 1993: 211); geschrieben vom städtischen Schreiber Emund von Eilsich (vgl. ebd.).

Neben den unter (2)–(4) genannten Belegen gibt es noch weitere Belege bei Stein ([1893] 1993 und [1895] 1993), die das Nebeneinander von Varianten mit *-s*, *-(e)* und *-t/d(e)* für mhd. *gewelbe* bezeugen. Die gewählten Belege in (2)–(4) illustrieren die genannte ausdrucksseitige Varianz und zeigen überdies, dass mit *gewulf(s/d/(e))* jeweils auf den gleichen Ort referiert wird, nämlich das städtische Archiv, in dem Urkunden aufbewahrt wurden.²⁷ Die Schreibungen ohne *-s* und *-t/-d* sind bei allen Belegen für das Lemma mhd. *gewelbe* am häufigsten: Belegschreibungen mit *-s* kommen zu 23% vor, Belegschreibungen mit *-t/-d* zu 6% bei 31 Belegen insgesamt. Dies zeigt, dass auch *ge-* + *-te* neben *ge-* + *-ze* als Variante in Köln bekannt war, von der selten auch Gebrauch gemacht wurde, es sei denn, dass hier ein nl. oder nd. Schreiber am Werk war.

Besonders zu Beginn der volkssprachigen Schriftlichkeit in Köln gibt es schreibsprachlich niederdeutsche Spuren (vgl. Klein und Peters 2015). Aus diesem Grunde wurde die Hs. eingesehen, die (2) überliefert, um weitere Passagen zu betrachten, die dieser Schreiber geschrieben hat. Ausschnitt (2) stammt aus dem 1. Eidbuch des 15. Jhs. (HASTK 30/V V 8) und wurde von einem namentlich „unbekannten Kanzleischreiber[]“ geschrieben, „der vorwiegend in den letzten Jahren des 14. Jahrhunderts und im ersten Jahrzehnt des 15. Jahrhunderts in der Kanzlei thätig ist“ (Stein ([1893] 1993: LXII). Die Passagen, die er geschrieben hat, enthalten mehrfach Ergänzungen vom bekannten Kanzlisten Heinrich Vront (vgl. Stein [1893] 1993: LXII, CXXXII) und auch von Emund von Eilsich, so auch auf fol. 7r (s. Abb. 2).

27 Bei Weinsberg findet sich kein Beleg für dieses Archiv, sondern nur für mhd. *gewelbe* mit anderer Bedeutung. Lediglich ein Beleg für mhd. *gewelbe-hërre* findet sich mit der Schreibung ›*Gewolffs her*› (LS 321v,61–62) mit fehlendem Bindestrich am Zeilenende. Hier ist jedoch nicht erkennbar, ob ›s› für *-ze* steht oder ein Fugenelement ist.

Abb. 2: Historisches Archiv der Stadt Köln, Best. 30/V (Verfassung), V8, fol. 8r²⁸.

28 Für die Abbildungserlaubnis, die mir am 24.04.2017 erteilt wurde, danke ich Dr. Max Plassmann (Historisches Archiv der Stadt Köln).

Daher wird man zwar sagen können, dass die Passagen, die der unbekannte Schreiber schrieb, von den städtischen Schreibern nochmals gelesen und bei Bedarf inhaltlich ergänzt wurden. Für den Eintrag (2) bedeutet dies aber leider nicht, dass Eilsich und Vront die Schreibung *gewelfde* nicht für korrekturbedürftig hielten, denn es finden sich keine inhaltlichen Korrekturen dieser Schreiber, wie die Randeinträge verdeutlichen (s. Abb. 2).

Die schreibsprachliche Analyse von Folio 6r–8v im 1. Eidbuch des 15. Jh.s, die von besagtem Schreiber geschrieben wurden, ergab allerdings keine schreibsprachlichen Merkmale, die zwingend auf einen nd. oder nl. Schreiber hindeuten.²⁹ Viele der auffälligen Merkmale finden sich auch sonst in anderen etwa zeitgleichen Texten aus Köln, die Stein ([1893] 1993 und [1895] 1993) ediert hat. Merkmale, die sich sonst nicht finden, sind nicht zahlreich genug, um sichere Rückschlüsse auf die Herkunft des Schreibers zu ziehen.³⁰ Ob aus dem Beleg *gewelfde* geschlossen werden darf, dass *ge-* + *-t/d(e)* auch in Köln möglich war und nicht als fremd empfunden wurde, lässt sich daher nicht mit völliger Gewissheit sagen, es wäre aber durchaus denkbar.

Schließlich soll noch auf ein Lexem mit *ge-* + *-te* im Rip. eingegangen werden, das mangels eines Anwendungsbezugs in den spätmittelalterlichen Quellen der Kölner Stadtverwaltung kaum in Erscheinung tritt, aber zum Zeitpunkt seiner Bezeugung die Spuren verschiedener Lautwandel zeigt, aufgrund derer es als eine sehr alte Bildung bestimmt werden kann. Es handelt sich hierbei um das Wort, das im Nhd. als *Gehöfte* (fort-)besteht.

Die im Rip. – nach Ausweis des RhWb (Bd. 3, Sp. 866) anzutreffenden dialektalen Lautungen – zeugen davon, dass das Wort den Wandel *-ft* > *-ht* mitgemacht hat und den später (etwa im 13./14. Jh.) erfolgten finalen *t*-Ausfall (vgl. Bütthe-Scheider 2017: 326–327). Besonders der Wandel von *-ft* > *-ht* lässt auf ein sehr hohes Alter der Bildung schließen, da der Lautwandel sich im Rip. wie auch in einigen nl. – auch den an das Rip. angrenzenden – Dialekten sehr früh vollzogen haben muss (vgl. Schützeichel 1974: 233–280; Klein 2000: 15). Am Niederrhein gibt es z. B. den heutigen Namen *Süchteln*

29 Zu nd. und nl. Merkmalen in rip. Rechtstexten – besonders der Frühzeit im 13. Jh. – vgl. Klein und Peters 2015.

30 Ein schreiberspezifisches Merkmal dieses Schreibers ist, dass er für kurze und ehemals lange *i*-Laute die Schreibung *ij* verwendet; Belege mit dieser Schreibung für kurze *i*-Laute wären z. B. *dijske* ‘oft’ (6v,03), *syndenckijch* ‘erinnerlich’ (7v,15) und *moegijch ind mechtijch* (eine Paarformel, etwa: ‘Kraft und Macht’) (8v,30 in einer Abschrift des Verbundbriefs). Hierbei handelt es sich nicht um eine typische Schreibgewohnheit in Köln. In einem Abschnitt aus der sog. Kölner Prosa-Kaiserchronik (Kais), die schreibsprachlich im Westripuarischen (wrip.) zu verorten ist (vgl. Bütthe-Scheider 2016), finden sich vergleichbare Schreibungen für *i* in den Nebensilben. Dieses Kriterium allein ist jedoch nicht ausreichend, um wrip. Schreibeeinfluss oder eine wrip. Herkunft des Schreibers anzunehmen.

schon im Jahr 1159 mit der Schreibung <Suchtele> statt vormals <Suftete> (vgl. Schützeichel 1974: 248).

Laut Grimm in ¹DWB soll Kölnisch *gehüchte* erst später lautlich und inhaltlich an *hov* m. ‘Hof’ angebunden worden sein, ursprünglich habe es sich um eine Ableitung von mhd. *huobe* ‘Stück Land von einem gewissen Maße’ gehandelt³¹, zu dem auch *Gehöfer* eine Ableitung ist (vgl. ¹DWB s.v. *Gehöfer*), die u. a. im nhd. Familiennamen *Hübner* noch fortlebt. Angesichts der angetroffenen Vokale im Nrip. kann eine lautliche Herleitung aus *huob-* als möglich gelten. Aufgrund des Phonemzusammenfalls von mhd. *üe* und *ü* vor *-ht* kann aus der rezenten Lautung allein kein Rückschluss auf *üe* oder *ü* gezogen werden, er ist aber auch obsolet.³²

Mangels eines Anwendungsbezuges tritt dieses Wort, wie erwähnt, in den Kölner Akten (ed. Stein ([1893] 1993 und [1895] 1993) nicht in Erscheinung, ist aber bei Hermann Weinsberg häufiger bezeugt, ein Beleg ist unter (5) genannt.

31 „so wird aber auch jenes cölnische gehufte vielmehr zu hufe, hube gehören, wie schon der vocal angibt, also gehüfte, das ganze das ein huber, gehuber, gehover (s. gehöfer 1, c) zu verwalten hat, hauptsächlich der grund und boden, der denn auch in jener bestimmung des pachtvertrags wahrscheinlich gemeint ist, denn sie steht mitten in andern bestimmungen über die behandlung der hoflant mit aussaat und düngen (mirgelen)“ (¹DWB, Bd. 4.1, Sp. 2492).

32 Vor *-ht* ist die Reihe mhd. *ie – üe – uo* im Rip. mit der Reihe mhd. *î – û – û* zusammengefallen (vgl. Wiesinger 1979 II: 57); mhd. *û* und *û* haben sich wiederum gleich entwickelt wie mhd. *u* und *ü* vor *-ht* (vgl. Müller 1912: 51): Im Landkölnischen hat sich *ûə* für mhd. *û* entwickelt und entsprechend *ÿə* für *û*; im Stadtkölnischen wird langes geschlossenes [o] und [ö] angetroffen (vgl. Müller 1912: 51). Diese Lautungen sind bei den meisten Formen anzutreffen, die das RhWb für das Stadt- und Landkölnische verzeichnet, sofern *-h-* nicht ausgefallen ist. Wenn das Wort so alt wäre, dass es eine Ableitung von **huv-* ‚Hof‘ darstellte, bevor der Vokal durch den germ. *a*-Umlaut *> -o-* gesenkt wurde (vgl. mhd. *holz* ‘Holz’– *gehülze* ‘Gehölz’ oder *hülzîn* ‘hölzern’, *gold* ‘Gold’– *güldîn* ‘golden’ mit *u > ü* infolge des althochdeutschen Sekundärumlautes), wäre eine alternative Herleitung gegeben, die zum gleichen lautlichen Resultat im Nrip. geführt hätte. Es ist auch möglich, dass es beide Bildungen gab – Entsprechendes wird auch für das Mnl. angenommen (vgl. MNW „nl. wb. III, 896, z. B. *eensame ghehuchte*“). Sie wären später lautlich zusammengefallen.

- (5) *Daß hoiffrecht ligt zû Swartzem, helt drittenhalben morgen vngeferlich, Eß mach myn Erb künfftich zû glück vnd habseligkeit gerathen vnd darüff oder vff eyn ander stuck landes eyn **gehüchte** hauß hoff schüren vnd stell setzen [...]* (Weinsberg LD, 500r,03–06) ‘Die Hofstätte liegt zu Schwarzheim, umfasst ungefähr drei halbe Morgen. Sie möge meinem Erben künftig zu Glück und Habe reichen; und darauf oder auf ein anderes Stück Land [kann er] **Wohnstätten**, Haus, Hof, Scheunen und Ställe setzen’³³.

Dieses Lexem nhd. *gehöfte* gibt ein weiteres Indiz, dass es *ge-* + *-te* auch im Rip. gegeben haben könnte.

Wie bereits oben erwähnt wurde, gibt es im Rip. bekanntlich eine Reihe von Lautverschiebungsausnahmen, die auch *-te* in *ge-* + *-te* betroffen haben könnten (vgl. Lerchner 1971: 231–234 mit ähnlichen Beispielen). Diese LV-Ausnahmen treten in den historischen Schriftzeugnissen nicht geregelt in Erscheinung. Heinrichs (1961) hat Hinweise dafür gefunden, dass sie immer schon dagewesen sind, aber eher der gesprochenen Sprache der unteren Bevölkerungsschichten angehörten. Das Ergebnis von Heinrichs Untersuchung lautet (1961: 151):

Ich habe sehr viele Beispiele gesammelt, die m. E. dartun, daß noch im späteren Mittelalter im Kölner Raum in der Grundschrift weitgehend unverschobene Formen gebraucht wurden, daß aber anderseits gewisse bodenständige Lautentwicklungen dem Eindringen verschobener Formen Vorschub leisteten.

Es wäre vor diesem Hintergrund gut möglich, dass *ge-* + *-te* im Rip. immer schon vorhanden war – das wäre auch eine Erklärung für das gelegentliche Auftreten neben *ge-* + *-ze* – und ab dem 14. Jh. dann verhochdeutsch wurde nach dem Muster, das weitere Fälle im Rip. betraf, wo ein Wort mit lautverschobenem und nicht lautverschobenem Laut im Gebrauch war – vielleicht diastratisch unterschiedlich distribuiert. Die Rolle des Nl. kann in diesem Falle darin bestanden haben, als einer Nachbarsprache, die von *ge-* + *-te* regen Gebrauch machte, den Ersatz von *ge-* + *-e* durch *ge-* + *-ze* anzuregen. Es wäre vor dem Hintergrund der von Möller (1998) ermittelten geringeren empfängerorientierten Anpassung der Schreibsprache zum Nl. und Nd. hin im Vergleich zum südlicheren Hd. bei eher mehr Unterschieden zwischen dem Rip. und Nl./Nd. als zwischen dem Rip. und südlichen Hd. (vgl. Möller 1998: 279) und der ermittelten Höherbewertung des südlicheren

33 Ob für *gehüchte* die Bedeutung ‘Weiler’ anzusetzen ist, die sich auch im Mnl. findet (vgl. MNW s.v. *gehocht* ‘Weiler’, ‘(kleines) Dorf’, ‘zueinander gehörende Gebäude’) oder eher nur die kollektivische von mhd. *hov* als ‘zum Hof gehörende Bauwerke’, ist aus dem Kontext nicht zu entscheiden.

Hd. im 15. Jh. (vgl. Möller 1998: 284) plausibel anzunehmen, dass das Rip. sich durch das hd. klingende *ge-* + *-ze* vom nördlicheren Niederland absetzen wollte (und vielleicht auch von Teilen der eigenen Bevölkerung (s. o.)); *ge-* + *-te* muss aber selbst nicht aus dem Nl. entlehnt worden sein, wie die Beispiele weiter oben gezeigt haben.

Vom Nl. ausgehend gäbe es noch eine weitere Erklärungsmöglichkeit für *-ze*, das auch als *<-s(e)>* in Erscheinung tritt (vgl. Bütthe und Wich-Reif 2012: 275, 281–282): Besonders im westlichen Flämischen und im Holländischen kommt ein *s*-Plural vor, der auch schon im 13. Jh. bezeugt ist (vgl. Marynissen 1996: 406–407), dort vornehmlich bei Maskulina auf *-er* vorkam, daneben aber auch schon in geringerem Umfang bei Tier- und Sachbezeichnungen auf *-er* verwendet wurde, darunter z. B. auch für *steiger* (vgl. Marynissen 1996: 210). Beispiele für die Nebenform auf *-t* in Kombination mit *s*-Plural gibt es auch, vgl. z. B. „Item en salmen gheen vecken (*afval van vogels?*) noch vuylnisse van beesten schudden van der groter bruggen in die havene noch van gheenen *steegaerts* noch bruggen, O. K. v. Dordr. 60, 215, Holland, 1451–1500“ ‘Außerdem soll man keinen Abfall von Vögeln(?) noch Exkreme von Tieren von der großen Brücke in die Häfen schütten noch von keinen Anlegeplätzen noch von Brücken’ in MNW (s.v. *steiger* I). Beispiele wie diese dürften jedoch nicht so zahlreich gewesen sein, denn Bildungen auf *-er* gibt es unter den Kollektiva nicht viele. Zudem dürften Kollektiva, die keine konkrete Bedeutung angenommen hatten, nicht im Plural verwendet worden sein. Daher ist eher nicht anzunehmen, dass Fälle wie der zuvor genannte eine musterbildende Wirkung gehabt haben können.

4. Schlussbetrachtung und Zusammenfassung

Im vorliegenden Beitrag konnten die zentralen Ergebnisse der bisherigen Forschung zu den rip. *ge-* + *-ze*-Bildungen anhand von Quellenuntersuchungen bestätigt werden. In der Tat existieren für die allermeisten rip. *ge-* + *-ze*-Bildungen nl. Pendanten auf *ge-* + *-te*, die einen Einfluss auf das Aufkommen und die Zunahme von *ge-* + *-ze*-Bildungen im Rip. gehabt haben können. Im Vergleich sind *ge-* + *-te*-Bildungen im Nl. deutlich zahlreicher als *ge-* + *-ze*-Bildungen im Rip., und sie sind bereits im 13. Jh. mit vielen Bildungen bezeugt, während die rip. *ge-* + *-ze*-Bildungen erst im 14. Jh. in Erscheinung treten. Im 13. Jh. kommen die Bildungen, die im Nl. zu dieser Zeit schon als *ge-* + *-te*-Bildungen existieren, im Rip. mit der älteren Bildungsweise *ge-* + *-e* vor, die gelegentlich auch im Nl. anzutreffen sind, dort aber nicht so zahlreich in Erscheinung treten wie im Hd. Die Betrachtung einiger *ge-* + *-tel-de*-Bildungen, die vereinzelt auch in rip. Quellen auftreten, deutet darauf hin, dass *ge-* + *-te* auch im Rip. als Lautverschiebungsausnahme autochthon vorhanden

gewesen sein könnte (s. 3.2). Daher muss es sich um kein aus dem Nl. übernommenes Suffix handeln.

Ähnliche Beispiele, die es auch in den Belegsammlungen von Bech (1865, 1869 und 1877) gibt, stammten dort meist aus nd. Quellen. Das FWB verzeichnet jedoch Beispiele für *ge-* + *-te(-de)* auch aus anderen eindeutig hochdeutschen Quellen, z. B. aus einer Straßburger Quelle der Beleg *«gestielter»* ‘Chorgestühl’ (Nom. Pl./Gen. Pl.) unter dem Lemma *gestüile*, Bedeutung 2 (Bd. 6, Sp. 1666). Demnach wäre es möglich, dass *ge-* + *-te(-de)* großräumiger verbreitet war. Dieser Punkt bedürfte noch weiterer Untersuchung.

Für das Rip. stellt sich die Entwicklung nach bisherigen Erkenntnissen wie folgt dar: *-te* aus *ge-* + *-te* hat in wohl verhochdeutscher Gestalt *-e* in *ge-* + *-e* ersetzt und damit zu einer erneuten ausdrucksseitigen Kennzeichnung der Kollektiva dieses Bildungstyps beigetragen. Den Auslöser zu diesem Wandel kann im Grunde nur die *e*-Apokope gegeben haben, durch die die Kollektiva einen Teil ihrer Kennzeichnung einbüßten, denn Neutra mit *ge-* und ohne finales Schwa gab es auch sonst schon. Mit der Entwicklung von *ge-* + *-e* > *ge-* + *-ze* im Rip. ist also ein seltenes Beispiel für den Fall gegeben, dass eine durch Lautwandel verursachte Klassenauflösung einer etymologisch zusammengehörigen Gruppe (nach Hermann Paul auch *Isolation* genannt) durch Suffixersatz verhindert wird.

Quellen und Literatur

1. Ungedruckte Quellen

1. Eidbuch 15. Jh. = Köln, Historisches Archiv der Stadt Köln, Best. 30/V (Verfassung), V8.
- BuMi = «Buch der Minne»/«Rede von den 15 Graden», Prag, Klosterbibl. in Strahov, Cod. DG IV 17, aus MiGraKo (= ReM M404).
- Kais = Kölner Prosa-Kaiserchronik, München, Bayerische Staatsbibl., cgm 691, 17ra,15–278vb; SW-Teil, 1ra–17ra,14.
- Koelhoffsche Chronik = s. Cardauns 1876.
- Kölner Akten = s. Stein [1893] 1993 und [1895] 1993.
- MiGraKo = *Korpus der Mittelhochdeutschen Grammatik*.
<https://www.linguistics.rub.de/rem/> (verfügbar am 28.06.2017), die O. K. v. Dordr. = s. Fruin 1876.
- Pil = *Pilgerfahrt des träumenden Mönchs*. Köln, Historisches Archiv der Stadt Köln, Best. 7004 (GB 4°), 223 (= Hs. C).
- PLilie = Die Lilie [Prosateil], Wiesbaden, Hessische Landesbibl., Hs. 68, 3r,1–26v,4, aus MiGraKo (= ReM M327).
- RBib = Mittelfränkische Reimbibel (A), [a] Halle (Saale), Universitäts- und Landesbibl., 6 Blätter (2 Doppelblätter und zwei einzelne Blätter); [b] Moskau, Bibl.

- der Lomonossow-Universität, 4 Blätter (2 Doppelblätter) aus MiGraKo (= ReM M199A).
- RhTun = Niederrheinischer Tundalus, Krakau, Bibl. Jagiellońska, Berol. mgq 624, 3 Bll., aus MiGraKo (= ReM M232).
- SW = Sächsische Weltchronik, Berlin, Staatsbibl., mgq 284, 1ra-53va.
- VLilie = Die Lilie [Versteil], Wiesbaden, Landesbibl., Hs. 68, 26v,4-115r,5, aus MiGraKo (= ReM M354).
- Weinsberg LS = *Liber Senectutis*, Köln, Historisches Archiv der Stadt Köln, C + D, Best. 7030, Nr. 50.
- Weinsberg LD = *Liber Decrepitvdinis*, Köln, Historisches Archiv der Stadt Köln, C + D, Best. 7030, Nr. 51.

2. Literatur

- Bech, Fedor. 1865. Kleine Beiträge. 1. *Gebeinze, geborgze, gebuscheze* u. s. w. *Germania* 10. 395–398.
- Bech, Fedor. 1869. Wortformen auf *-eze*. Nachtrag zu *Germania* 10, 395–398. *Germania* 14. 431–432.
- Bech, Fedor. 1877. Wortformen auf *-eze*. Nachtrag zu *Germania* X, 395–398 und XIV, 431 folg. *Germania* 22. 290–293.
- BMZ = *Mittelhochdeutsches Wörterbuch. Mit Benutzung des Nachlasses von Georg Friedrich Benecke ausgearbeitet von Wilhelm Müller und Friedrich Zarncke.* 1854–1866. 3 Bde. Leipzig: Hirzel.
- Brendel, Bettina, Regina Frisch, Stephan Moser & Norbert Richar Wolf. 1997. *Wort- und Begriffsbildung in frühneuhochdeutscher Wissensbildung. Substantivische Affixbildung* (Wissensliteratur im Mittelalter 226). Wiesbaden: Reichert.
- Büthe, Eva & Claudia Wich-Reif. 2012. Das „Historische Rheinische Wörterbuch“ als Addendum zum „Rheinischen Wörterbuch“. Die *ge-* + *-zel-s*-Kollektivbildungen. *Rheinische Vierteljahrsblätter* 76. 268–283.
- Büthe-Scheider, Eva. 2016. Zur schreibsprachlichen Lokalisierung der sogenannten ‚Kölner Prosa-Kaiserchronik‘ (Cgm 691). *Rheinische Vierteljahrsblätter* 80. 1–23.
- Büthe-Scheider, Eva. [vor. Sept.] 2017. *Die e-Apokope im Ripuarischen. Eine korpuslinguistische Untersuchung spätmittelhochdeutscher und frühneuhochdeutscher Quellen* (Studia Linguistica Germanica 130). Berlin & New York: de Gruyter.
- Cardauns, Heinrich (ed.). 1876. Die Cronica van der hilliger stat van Coellen. 1499 (Koelhoffsche Chronik). In *Die Chroniken der niederrheinischen Städte. Cöln. Zweiter Band. Auf Veranlassung und mit Unterstützung seiner Majestaet des Königs von Bayern Maximilian II herausgegeben durch die historische Commission bei der königl. Academie der Wissenschaften* (Chroniken deutscher Städte 13). Leipzig: Hirzel. 263–640.
- Cd-rom Middelnederlands* = Instituut voor Nederlandse Taal [vormals: Instituut voor Nederlandse Lexicologie] (ed.). [1998] 2017. *Cd-rom Middelnederlands. Woordenboek en teksten*. Leiden: Sdu/Standdaard.

- DRW = *Deutsches Rechtswörterbuch. Wörterbuch der älteren deutschen Rechtssprache. Herausgegeben von der Heidelberger Akademie der Wissenschaften.* 1914 ff. Weimar: Hermann Böhlau Nachfolger.
<http://drw-www.adw.uni-heidelberg.de/drw-cgi/zeige> (verfügbar am: 01.07.2017).
- ¹DWB = *Deutsches Wörterbuch von Jacob und Wilhelm Grimm.* 1854–1960. 16 Bde. Quellenverzeichnis 1971. Leipzig: Hirzel.
http://woerterbuchnetz.de/cgi-bin/WBNetz/wbgui_py?sigle=DWB (verfügbar am: 01.07.2017).
- EWN = Philippa, Marlies, Frans Debrabandere, Arend Quak, Tanneke Schoonheim & Nicoline van der Sijs. 2003–2009. *Etymologisch woordenboek van het Nederlands.* 4 Bde. Amsterdam: University Press.
<http://www.etymologiebank.nl/> (verfügbar am: 01.07.2017).
- FWB = Goebel, Ulrich, Anja Lobenstein-Reichmann & Oskar Reichmann (eds.). 1989ff. *Frühneuhochdeutsches Wörterbuch.* Begr. v. Robert R. Anderson, Ulrich Goebel & Oskar Reichmann. Bd. 1ff. Berlin & New York: de Gruyter.
<https://fwb-online.de/> (verfügbar am: 01.07.2017).
- Grimm, Jacob. 1840. *Weisthümer. Gesammelt von Jacob Grimm. Mitherausgegeben von Ernst Dronke und Heinrich Beyer.* Bd. 2. Göttingen: Dietherichsche Buchhandlung.
- Grimm, Jacob. [1878] 1989. *Deutsche Grammatik. Zweiter Theil. Neuer vermehrter Abdruck besorgt durch Wilhelm Scherer.* 2. Aufl., verm. Abdruck. [Berlin: F. Dümmers] Hildesheim, Zürich & New York: Olms.
- Grimm, Jacob. [1890] 1989 *Deutsche Grammatik. Dritter Theil. Neuer vermehrter Abdruck besorgt durch Gustav Roethe und Edward Schröder.* 2. Aufl., verm. Abdruck. [Gütersloh: Bertelsmann] Hildesheim, Zürich & New York: Olms.
- Heinrichs, Heinrich Matthias. 1961. ‚Wye grois dan dyn andait eff andacht is...‘. Überlegungen zur Frage der sprachlichen Grundsicht im Mittelalter. *Zeitschrift für Mundartforschung* 28. 98–153.
- Henzen, Walter. 1965. *Deutsche Wortbildung.* 3., durchges. und erg. Aufl. Tübingen: Niemeyer.
- Höhlbaum, Konstantin (Bearb.). [1886] 2000. *Das Buch Weinsberg. Kölner Denkwürdigkeiten aus dem 16. Jahrhundert. Erster Band* (Publikationen der Gesellschaft für Rheinische Geschichtskunde III). [Leipzig: Dörr] Düsseldorf: Droste.
- Höhlbaum, Konstantin (Bearb.). [1887] 2000. *Das Buch Weinsberg. Kölner Denkwürdigkeiten aus dem 16. Jahrhundert. Zweiter Band* (Publikationen der Gesellschaft für Rheinische Geschichtskunde IV). [Leipzig: Dörr] Düsseldorf: Droste.
- Honemann, Volker. 1989. Rede von den fünfzehn Graden. In Kurt Ruh, *Die deutsche Literatur des Mittelalters. Verfasserlexikon. Begründet von Wolfgang Stammer, fortgeführt von Karl Langosch.* 2., völlig neu bearbeitete Aufl. unter Mitarbeit zahlreicher Fachgelehrter. Bd. 7: ‚Oberdeutscher Servatius‘ – Reuchart von Salzburg, 1061–1065. Berlin & New York: de Gruyter.
- Klein, Thomas. 1982. *Untersuchungen zu den mitteldeutschen Literatursprachen des 12. und 13. Jahrhunderts.* Habil. (Masch.) Bonn.

- Klein, Thomas. 2000. Rheinische und westfälische Sprachgeschichte bis 1300. In Jürgen Macha, Elmar Neuss & Robert Peters (eds.), *Rheinisch-Westfälische Sprachgeschichte* (Niederdeutsche Studien 46). 3–48.
- Klein, Thomas, Hans-Joachim Solms & Klaus-Peter Wegera (eds.). 2009. *Mittelhochdeutsche Grammatik. Teil III. Wortbildung*. Tübingen: Niemeyer.
- Klein, Thomas & Robert Peters. 2015. Niederdeutsche Schreiber, Gottfried Hagen und die Anfänge der deutschsprachigen Urkunde in Köln. In Anna Karin, Silvia Ulivi & Claudia Wich-Reif (eds.), *Regiolekt, Funktiolekt, Idiolekt: Die Stadt und ihre Sprachen. Akten der 31. Tagung des Internationalen Arbeitskreises Historische Stadtsprachenforschung. Bonn, 29. September - 02. Oktober 2013*. Bonn: University Press. 25–66.
- Kluge, Friedrich. 1899. *Nominale Stammbildungslehre der altgermanischen Dialekte* (Sammlung kurzer Grammatiken germanischer Dialekte, Ergänzungsreihe: 1. Nominale Stammbildungslehre der altgermanischen Dialekte). Halle: Niemeyer.
- Lau, Friedrich (Bearb.). [1897] 2000. *Das Buch Weinsberg. Kölner Denkwürdigkeiten aus dem 16. Jahrhundert. Dritter Band* (Publikationen der Gesellschaft für Rheinische Geschichtskunde XVI). [Bonn: Hanstein] Düsseldorf: Droste.
- Lau, Friedrich (Bearb.). [1898] 2000. *Das Buch Weinsberg. Kölner Denkwürdigkeiten aus dem 16. Jahrhundert. Vierter Band* (Publikationen der Gesellschaft für Rheinische Geschichtskunde XVI). [Bonn: Hanstein] Düsseldorf: Droste.
- Lerchner, Gotthard. 1971. *Zur II. Lautverschiebung im Rheinisch-Westmitteldeutschen. Diachronische und diatopische Untersuchungen* (Mitteldeutsche Studien 30). Halle (Saale): Niemeyer.
- Lexer, Matthias. 1872–1878. *Mittelhochdeutsches Handwörterbuch*. 3 Bde. Leipzig: Hirzel.
- LLU = Gangler, Jean-François 1847. *Lexicon der Luxemburger Umgangssprache (wie sie in und um Luxemburg gesprochen wird) mit hochdeutscher und französischer Uebersetzung und Erklärung, verbunden von J. F. Gangler*. Luxemburg: Hoffmann. http://engelmann.uni.lu:8080/portal/WBB2009/LLU/wbgui_py?lemid=UA00001 (verfügbar am: 01.07.2017).
- Loey, Adolphe van. 1968. *Middelnerlandse spraakkunst. I. Vormleer*. 4. Aufl. Groningen: Wolters.
- Marynissen, Ann. 1996. *De flexie van het substantief in het 13de-eeuwse ambtelijke Middelnerlands. Een taalgeografische studie* (Studies op het gebied van de Nederlandse Taalkunde 2). Zugl. Leuven, Univ., Diss., Leuven: Peeters.
- Münch, Ferdinand. [1904] 1970. *Grammatik der ripuarisch-fränkischen Mundart*. [Bonn: Bouvier] Wiesbaden: Martin Sändig oHG.
- MWB = *Mittelhochdeutsches Wörterbuch. Im Auftrag der Akademie der Wissenschaften und der Literatur Mainz und der Akademie der Wissenschaften zu Göttingen herausgegeben von Kurt Gärtner, Klaus Grubmüller und Karl Stackmann*. 2006ff. Stuttgart: Hirzel. <http://www.mhdwb-online.de/wb.php> (verfügbar am: 01.07.2017).
- MNW = Verwijs, Eelco und Jacob Verdam: *Middelnerlandsch Woordenboek*. 10 Dln. 's-Gravenhage, 1885–1952. Dl. 9 voltooid door F.A. Stoett. Dl.

- 10 (Tekstcritiek van J. Verdam en bouwstoffen) 1e gedeelte (A-F) door W. de Vreese; 2e gedeelte (G-Z) door G.I. Liefstinck. Ook opgenomen op: Cd-rom Middelnederlands. Woordenboek en teksten. Den Haag, Antwerpen 1998. <http://gtb.inl.nl/?owner=MNW> (verfügbar am 01.07.2017).
- Möller, Robert. 1998. *Regionale Schreibsprachen im überregionalen Schriftverkehr. Empfängerorientierung in den Briefen des Kölner Rates im 15. Jahrhundert* (Rheinisches Archiv 139). Köln, Weimar, Wien: Böhlau.
- Müller, Wilhelm. 1912. *Untersuchungen zum Vokalismus der stadt- und landkölnischen Mundart*. Zugl.: Bonn, Univ., Diss. Bonn: Hauptmann'sche Buchdruckerei.
- Pijnenburg, Willy J.J. 1976. Eeuwsel. *Naamkunde* 8. 1–53.
- Ruh, Kurt. 1993. *Geschichte der abendländischen Mystik. Bd. 2: Frauenmystik und Franziskanische Mystik der Frühzeit*. München: Beck.
- RhWb = *Rheinisches Wörterbuch. Im Auftrag der Preußischen Akademie der Wissenschaften, der Gesellschaft für Rheinische Geschichtskunde und des Provinzialverbandes der Rheinprovinz auf Grund der von Johannes Franck begonnenen, von allen Kreisen des Rheinischen Volkes unterstützten Sammlung bearbeitet und herausgegeben von Josef Müller, Heinrich Dittmaier, Rudolf Schützeichel und Mattias Zender. 1928–1971. 9 Bände*. Bonn: Fritz Klopp, Berlin: Erika Klopp. http://woerterbuchnetz.de/cgi-bin/WBNetz/wbgui_py?sigle=RhWB (verfügbar am 01.07.2017).
- Schützeichel, Rudolf 1974. *Mundart, Urkundensprache und Schriftsprache. Studien zur rheinischen Sprachgeschichte* (Rheinisches Archiv 54). 2., stark erweiterte Aufl. mit 39 Karten. Bonn: Röhrscheid.
- Stein, Walther (ed.). [1893] 1993. *Akten zur Geschichte und Verfassung der Stadt Köln im 14. und 15. Jahrhundert* (Publikationen der Gesellschaft für Rheinische Geschichtskunde 10). Bd. 1. [Bonn: Behrendt] Düsseldorf: Droste.
- Stein, Walther (ed.). [1895] 1993. *Akten zur Geschichte und Verfassung der Stadt Köln im 14. und 15. Jahrhundert* (Publikationen der Gesellschaft für Rheinische Geschichtskunde 10). Bd. 2. [Bonn: Behrendt] Düsseldorf: Droste.
- Wiesinger, Peter. 1970. *Phonetisch-phonologische Untersuchungen zur Vokalentwicklung in den deutschen Dialekten*. Bd. I: *Die Langvokale im Hochdeutschen*. Bd. II: *Die Diphthonge im Hochdeutschen*. (Studia Linguistica Germanica 2/1-2/2). Berlin: de Gruyter.
- Wilmanns, Wilhelm. 1899. *Deutsche Grammatik. Gotisch, Alt-, Mittel- und Neuhochdeutsch*. Bd. 2: *Zweite Abteilung: Wortbildung*. 2. Aufl. Strassburg: Trübner.
- Willeumier-Schalij, Johanna Marie. 1946. *Dat boec der Minnen, die Rede von den 15 Graden*. Zugl.: Leiden, Univ., Diss., Leiden: Brill.
- WLM = 1906. *Wörterbuch der luxemburgischen Mundart*. Luxemburg: M. Huss. 1906. http://engelmann.uni.lu:8080/portal/WBB2009/WLM/wbgui_py?lemid=MA00001 (verfügbar am: 01.07.2017).
- Wrede, Adam. 1958. *Neuer Kölner Sprachschatz. Zweiter Band K-R*. Köln: Greven Verlag.
- Vilmar, August Friedrich Christian. 1868. *Idiotikon von Kurhessen*. Bd. 1. Marburg, Leipzig: Elwert'sche Universitäts-Buchhandlung.

VMNW = Pijnenburg, Willy J.J., Karina H. van Dalen-Oskam, Katrien A.C. Depuydt & Tanneke H. Schoonheim (eds.). 2001. *Vroegmiddelnederlands Woordenboek. Woordenboek van het Nederlands van de dertiende eeuw in hoofdzaak op basis van het Corpus-Gysseling*. 4 Bde. Leiden: Gopher Publishers.
<http://gtb.inl.nl/?owner=VMNW> (verfügbar am 01.07.2017).

Dr. Eva Bütthe-Scheider
Rheinische Friedrich-Wilhelms-Universität Bonn
Institut für Germanistik, Vergleichende Literatur- und Kulturwissenschaft
Arbeitsstelle Rheinische Sprachforschung
Am Hofgarten 22
D-53113 Bonn
ebuschei@uni-bonn.de

Sören Stumpf

Textsortenorientierte Wortbildungsforschung. Desiderate, Perspektiven und Beispielanalysen

Abstract: This paper provides an overview of the connection between word formation and text type linguistics. Following a brief outline of the current state of research, desiderata and weaknesses of previous research as well as perspectives of a text type oriented research on word formation will be introduced. Here, I advocate a stronger inclusion of oral (with regard to the medium) and conceptually spoken text types (cf. Koch/Oesterreicher 1985). The focus is on the analysis of word formations within the text type of battle rap, which can be classified as oral and conceptually spoken. The analysis gives an insight into my habilitation project outlined in the essay and shows how this project can be realized.

1. Einleitung

Die Verbindung der beiden linguistischen Teildisziplinen der Wortbildung und der Textlinguistik stellt keineswegs einen neuen Forschungsansatz dar, liegen die ersten wegweisenden Ansätze doch bereits seit den 1970er und 1980er Jahren vor (siehe unter anderem Agricola 1969; Koch 1972; Schröder 1978, 1983, 1985; von Polenz 1980 und Wildgen 1982). Im Zentrum dieser innerdisziplinären Vernetzung stehen insbesondere die Fragen, inwiefern Wortbildungen am Aufbau von Texten beteiligt sind und inwieweit sie als textsortenspezifische und -distinktive Charakteristika angesehen werden können. Trotz dieser recht langen Forschungsgeschichte ist es verwunderlich, dass auch heute noch in den meisten Einführungswerken und Gesamtdarstellungen zur Wortbildung eine systemlinguistische, an der *langue* orientierte Sichtweise dominiert (vgl. Elsen/Michel 2007: 1). So bilden in der Regel die verschiedenen Wortbildungseinheiten und -arten den Hauptschwerpunkt. Verwendungsweisen und Funktionen von Wortbildungen im konkreten Sprachgebrauch (sei es in Texten oder Gesprächen) werden kaum thematisiert. Eine Ausnahme ist sicherlich das Überblickswerk von Fleischer/Barz (2012), das dem Themenbereich „Wortbildung und Text“ ein eigenes Kapitel widmet. In anderen Einführungen wie Eichinger (2000a), Donalies (2005, 2011), Motsch (2004), Erben (2006), Lohde (2006) und Elsen (2014) sucht man jedoch vergeblich nach textlinguistischen Betrachtungsweisen zur Wortbildung. Im Gegensatz dazu lässt sich nach einem Blick in einzelne Qualifikationsschriften, Zeitschriften- und Sammelbandaufsätze konstatieren, dass von einer generellen Vernachlässigung der Erforschung des Zusammenhangs von Wortbildung und

Text keine Rede sein kann – so wie es in den Einführungswerken zunächst den Anschein erweckt.

Ziel des Artikels ist es, einen Überblick über den Forschungsstand zu Wortbildung und Text(bildung) zu geben sowie Forschungsdesiderate und -perspektiven aufzuzeigen. Die theoretischen Ausführungen werden mithilfe von Beispielanalysen verdeutlicht. Insgesamt wird zum einen für eine umfassendere Auseinandersetzung mit den Beziehungen zwischen Wortbildung und Text und zum anderen für einen stärkeren Einbezug medial mündlicher und konzeptionell gesprochener Texte plädiert, da hierzu „bislang vergleichsweise wenig differenzierte Untersuchungen [existieren]“ (Elsen/Michel 2011: 34).

Nach einem Forschungsüberblick (Kapitel 2) werden Desiderate aufgezeigt, die sich aus den bisherigen Schwächen und Mängeln der Forschung ableiten (Kapitel 3). Kapitel 4 stellt daran anschließend Perspektiven einer textsortenorientierten Wortbildungsforschung vor, wobei in diesem Kontext genauer auf das Projekt „Wortbildung in Mündlichkeit und Schriftlichkeit“ eingegangen wird. Anhand der Textsorte Battle-Rap wird schließlich exemplarisch eine textsortenorientierte Wortbildungsanalyse durchgeführt (Kapitel 5). Im Fokus stehen dabei insbesondere qualitative Merkmale der Wortbildungen, indem deren Funktionen innerhalb dieser Textsorte mithilfe von authentischen Textbelegen herausgearbeitet werden.

2. Forschungsstand: Wortbildung und Textlinguistik

Die Beziehungen zwischen Wortbildung¹ und Text zeigen sich darin, „dass einerseits Wortbildungsprodukte Bausteine des Textes sind und dass andererseits Texte die Rahmenbedingungen setzen für die Bildung von Wörtern sowie für ihren Gebrauch“ (Barz/Schröder/Hämmer/Poethe 2007: 60). In der bisherigen Forschung stehen dabei vor allem die textkonstitutiven und textdistinktiven Eigenschaften von Wortbildungen im Fokus. Textkonstitutiv heißt, dass Wortbildungen „am Aufbau von Texten beteiligt sind“ (Fleischer/Barz 2012: 26), und textdistinktiv, dass „sie Einzeltexte voneinander

1 Unter „Wortbildung“ werden im vorliegenden Aufsatz sowohl alle morphologisch komplexen Einzelausdrücke im Text jenseits der Flexionsformen als auch die Verfahren zur Bildung neuer Lexeme verstanden (vgl. Fleischer/Barz 2012: 1f.). Barz/Schröder/Hämmer/Poethe (2007: 13) verwenden im ersten Fall den Terminus „Wortbildungsprodukt“. Im zweiten Fall kann von „Wortbildungstypen/-arten“ bzw. „Wortbildungsverfahren“ gesprochen werden. Im Folgenden wird an Stellen, an denen diese terminologische Unterscheidung eine Rolle spielt, auf sie zurückgegriffen.

zu unterscheiden oder Textsorten gegeneinander abzugrenzen helfen, sie aber auch musterhaft kennzeichnen können“ (ebd.).²

Studien zur **textkonstitutiven Funktion** (z. B. Schröder 1978, 1983, 1985; Wildgen 1982; Dederding 1983; Gataullin 1990; Eichinger 1995; Wolf 1996; Schlienz 2004 und Kauffer 2006) versuchen, das kohäsionsbildende Potenzial von Wortbildungen herauszuarbeiten.³ So zeigt sich, dass in einem Text Wortbildungsprodukte mit demselben Grundmorphem einer Isotopiekette angehören können (vgl. Erben 1995: 546) und sie somit „die inhaltsseitige Kohärenz eines Textes ausdrucksseitig unterstützen“ (Barz/Schröder/Hämmer/Poethe 2007: 60). Dies soll im Folgenden anhand eines Zeitungskommentars aus der Frankfurter Allgemeinen Zeitung (FAZ) verdeutlicht werden (siehe Beispiel 1 im Anhang).

Beispiel 1 zeigt recht eindrucksvoll, „dass alle Wortbildungen eines Textes aufgrund rekurrenter Grundmorpheme ein sachorientiertes Umfeld lexikalisch hervortreten lassen“ (Fleischer/Barz 2012: 28). Allein anhand der Wortbildungsketten ist ersichtlich, um was es in dem Zeitungskommentar geht (siehe Tabelle 1). Thematisch lässt er sich anhand der wiederkehrenden Grundmorpheme bzw. lexikalischen Einheiten in den Bereich Wirtschaft und (Sozial-)Politik einordnen (z. B. *-beitrag-*, *-rente-*, *-sozial-*, *-arbeit-*, *-kasse-*, *-minister-* und *-finanz-*). Dabei sind die einzelnen Isotopieketten miteinander vernetzt, indem bestimmte Einheiten gemeinsam innerhalb von Wortbildungen auftreten (z. B. *Sozialbeiträge*, *Rentenkasse* und *Finanzminister*). Die Wortbildungen erstrecken sich über die drei Wortarten Substantiv (*Pflegeleistungen*), Verb (*kassieren*) und Adjektiv (*finanziell*), wobei Substantive eindeutig überwiegen. Bezüglich der Wortbildungstypen nehmen Determinativkomposita (z. B. *Wahlkampf*, *Sozialsystem* und *Gesundheitsminister*) den Löwenanteil ein. Vereinzelt werden im Beispieltext auch Derivate gebildet (z. B. *erarbeiten*).

2 Für weiterführende Informationen sei auf die Forschungsüberblicke in Matussek (1994: Kap. 1); Peschel (2002: Kap. 3); Schlienz (2004: Kap. 2) und Fleischer/Barz (2012: Kap. 1.4) verwiesen.

3 Anzumerken ist, dass auch Arbeiten existieren, die sich speziell mit dem Vorkommen und den Funktionen von Wortneubildungen/Okkasionalismen innerhalb von Texten beschäftigen (siehe unter anderem Matussek 1994; Michel 1997 und Peschel 1998, 2002).

Tab. 1: Isotopieketten im FAZ-Zeitungskommentar „Sozial kurzsichtig“.

-sozial-	Sozialabgaben Sozialleistungen Sozialsystem Sozialausgaben Soziale Marktwirtschaft Sozialpolitik	-rente-	Renten- und Arbeitslosenversicherung Rentenkasse Mütterrente Rente mit 63
		-pflege-	Pflegeversicherung Pflegeleistungen
-beitrag-	Beitragsatz/-sätze Zusatzbeitrag/-beiträge Beitragsprung Beitragszahler Sozialbeiträge	-leist-	Pflegeleistungen Leistungen Sozialleistungen
		-wirtschaft-	Wirtschaftsministerium Marktwirtschaft
-kasse-	Arbeitslosenkasse Krankenkasse Rentenkasse kassieren	-arbeit-	Arbeitslosenkasse erarbeitet Arbeitslosenversicherung Arbeitnehmer
-wahl-	Wahlkampf Wahljahr	-gesund-	Gesundheitsminister Gesundheitsfonds
-steuer-	Steuersenkung steuerfinanzierte Gegensteuern ⁴	-steig-	steigen steigende Anstieg
-prozent-	Prozent Prozentpunkte	-zusatz-	Zusatzbeitrag/-beiträge Zusatzeinnahmen
-polit-	Politik politisch Sozialpolitik	-finanz-	Finanzminister steuerfinanzierten finanzieller
-kurz-	kurzsichtig kurzerhand	-wettbewerb-	Wettbewerb Wettbewerbsfähigkeit
-minister-	Finanzminister Gesundheitsminister Wirtschaftsministerium	-weit-	Ausweitung ausweiten Weitsicht
-nehm-	Zusatzeinnahmen Unternehmen Arbeitnehmer zunehmend	-schutz-	Schutz Klimaschutz

- 4 Bei *Gegensteuern* handelt es sich um eine stilistische Besonderheit. Während in den Wortbildungen *Steuersenkung* und *steuerfinanzierte* das Substantiv *Steuer* ‚Abgabe an den Staat‘ auftritt, liegt der Konversion *Gegensteuern* das homonyme Verb *steuern* ‚das Steuer eines Fahrzeugs bedienen‘ zugrunde. Die Vernetzung liegt strenggenommen also nur ausdrucksseitig vor. Inhaltlich wird dennoch mittels eines Wortspiels eine Verbindung hergestellt. So kann das Lexem aufgrund des Kontextes auch auf das Substantiv *Steuer* bezogen werden; in dem Sinne, dass der Staat mit seiner *Steuerpolitik gegen* mögliche Bedrohungen der Sozialen Marktwirtschaft vorgeht, er also sozusagen *gegensteuert*.

Die textkonstitutive Funktion von Wortbildungen kommt selbstverständlich auch in medial mündlichen/konzeptionell gesprochenen Texten zum Vorschein. In der bisherigen Forschung finden sich hierfür jedoch keine Beispiele. Im Folgenden sollen daher die durch Wortbildung erzeugten Isotopieketten in einem Interview aufgezeigt werden. Hierbei handelt es sich um ein Gespräch, das zwischen dem Moderator Claus Lufen (ARD) und dem Kugelstoßer David Storl kurz nach dessen Silbermedaillengewinn bei der Leichtathletik-WM in Peking 2015 stattgefunden hat (siehe Beispiel 2 im Anhang).

Tabelle 2 fasst diejenigen Wortbildungen zusammen, die Anteil an der Kohäsion haben. Im vorliegenden Interview sind es vor allem die Grundmorpheme *-train-*, *-stoß-* und *-techn-*, die beim Aufbau von Isotopieketten in Form von Wortbildungen fungieren. Besonders hervorzuheben ist das dreimalige Aufgreifen der Partikel *rein-* in den Verben *reinkommen*, *reinstürzen* und *reinbringen*, die ebenfalls als kohäsionsbildend angesehen werden kann (vgl. Schlienz 2004: 296–306). Augenscheinlich ist aber auch, dass im Gegensatz zum konzeptionell geschriebenen Kommentar in dem vorliegenden Interview aus quantitativer Sicht nicht so viele verschiedene Grundmorpheme und daraus aufbauend Wortbildungsprodukte auftreten. Stattdessen werden häufiger die gleichen Lexeme wiederaufgegriffen, um somit ebenfalls Kohäsion herzustellen (beispielsweise findet sich *gehen* sieben Mal, *Wettkampf/-kämpfe* fünf Mal und *fehlen* sowie *Titel* drei Mal in dem Gespräch wieder).

Tab. 2: Isotopieketten im Interview mit David Storl.

<i>-train-</i>	<i>Training Trainer trainieren</i>	<i>-stoß-</i>	<i>stoßen Standstöße Stöße</i>	<i>-rein-</i>	<i>reinkommen reinstürzen reinbringen</i>	<i>-geh-</i>	<i>gehen umgehen (gutgehen)</i>
<i>-meter-</i>	<i>Meter Zentimeter</i>	<i>-arbeit-</i>	<i>gearbeitet verarbeiten</i>	<i>-techn-</i>	<i>technisch Technik</i>	<i>-tag-</i>	<i>Tagen Dienstag</i>
<i>-spann-</i>	<i>Spannung Vorspannung</i>						

Steht die **textdistinktive Funktion** im Vordergrund, so erforscht man, inwiefern Wortbildungsphänomene als stilbildende Potenzen für den Einzeltext eingesetzt werden können. Darüber hinaus trägt die Untersuchung von Wortbildungsverfahren dazu bei, eine Textsorte zu beschreiben und sie von anderen Textsorten zu differenzieren (vgl. von Polenz 1980: 178 und Barz/Schröder 2001: 185). Ausgehend von der Annahme, dass Wortbildungen einen Anteil an der Textsortentypik haben können, sollten bei einer mehrdimensionalen/holistischen Textsortenanalyse, wie sie mittlerweile Standard innerhalb der Textsortenlinguistik ist, Wortbildungen mitberücksichtigt werden, wenn es

um die Untersuchung lexikalischer Charakteristika geht.⁵ Sichtet man die Forschungsliteratur, so lassen sich nicht wenige Studien anführen, die die Typik von Wortbildungen in bestimmten Textsorten fokussieren. Meist sind dies kleinere Analysen im Rahmen von Aufsätzen, die jeweils eine Textsorte in den Mittelpunkt stellen und dabei insbesondere zu Erkenntnissen bezüglich des quantitativen Vorkommens der unterschiedlichen Wortbildungsarten gelangen. So existieren beispielsweise Arbeiten zu Wortbildungen bzw. Wortneubildungen in

- literarischen Texten (z. B. phantastische Kinderliteratur, Elsen 2011; Science-Fiction-Literatur, Ortner 1985 und Siebold 2000),
- Werbetexten (z. B. Werbeanzeigen, Krieg 2005 und Meliss 2008; Hörfunkwerbung, Greule/Reimann 2011),
- Hörfunknachrichten (siehe Schröder 2005) und Presstexten (siehe Fleischer/Barz 2012: Kap. 1.4.2.2.2) sowie
- Gesprächen (z. B. Sprechstundengespräche, Elsen/Michel 2010; Familiengespräche, Schröder 1992; (telefonische) Beratungsgespräche, Weber 2014).

3. Forschungsdesiderate

Wie der Überblick zeigt, ist der Themenkomplex „Wortbildung und Text“ bereits aus verschiedenen Blickwinkeln beleuchtet worden. Nichtsdestotrotz drängen sich bei der Sichtung der Forschungsliteratur einige Schwachstellen bzw. größere Desiderate auf:

- 1) Die Studien, die sich mit den Beziehungen zwischen Wortbildung und Text beschäftigen, bestehen meist nur aus eng gesteckten Einzelanalysen. „[E]ine geschlossene Darstellung steht noch aus“ (Fleischer/Barz 2012: 26). Dabei mangelt es vor allem an Korpusauswertungen, denen eine breitere Materialsammlung zugrunde liegt und die für allgemeinere Aussagen dienlich sein können (vgl. Stein 2016: 240).
- 2) Es stehen häufig Texte im Mittelpunkt, die keine authentische Kommunikation widerspiegeln bzw. die in stark fachbezogenen Bereichen anzusiedeln sind (also literarische Texte wie z. B. Eichinger 2000b, Siebold 2000 und Elsen 2011; Werbetexte wie z. B. Krieg 2005; Fachsprache wie z. B. Zhang 1990, Poethe 2000 und Crestani 2010). Textsorten, die der *face-to-face*-Kommunikation zuzuordnen sind und „Sprache-in-Interaktion“ (Imo 2013: 1) zeigen bzw. die nicht an eine bestimmte

5 Siehe hierzu das Anforderungsprofil für die Beschreibung von Texten/Textsorten in Stein (2004: 191).

Fachsprache gebunden sind, finden bislang kaum Beachtung (vgl. Elsen/Michel 2011: 6).

- 3) Es werden textkonstitutive und -distinktive Funktionen zum größten Teil nur in medial schriftlichen und konzeptionell geschriebenen Texten untersucht (vgl. Elsen/Michel 2007: 9):⁶

„Im Gegensatz zur Textlinguistik liegen nur wenige empirische Untersuchungen zur Wortbildung in der gesprochenen Sprache vor [...]“ (Elsen/Michel 2011: 5)

Diese Forschungslücke macht sich auch in dem Standardwerk zum Gesprochenen Deutsch von Schwitalla (2012) bemerkbar, in dem das Kapitel zur Wortbildung noch nicht einmal eine Seite umfasst.

- 4) Die meisten Studien gehen nicht über einzelne Textsorten hinaus. Die Erforschung, welche Rolle Wortbildungen in sogenannten Textkomplexen/„Puzzletexten“ und Textsortenkettten/-netzen (siehe Schröder 2000; Püschel 1997 und Janich 2009) bzw. in transtextuellen Diskursen (siehe Busse/Teubert 1994) einnehmen, kann weitgehend als Desiderat angesehen werden. So fordert Schröder (2002: 110) in einem programmatischen Aufsatz,

„dass die Wortbildungsforschung künftig verstärkt ihr Augenmerk auf textübergreifende Wortbildungserscheinungen im Textkomplex und auch im Diskurs [...] zu richten hat.“

Hervorzuheben ist hierbei ein Artikel von Schröder (2000), in dem sie aufzeigt, dass auch in sogenannten Textkomplexen⁷

Wortbildungsprodukte mit gleichem Kernwort Isotopieketten [bilden], die onomasiologische Begriffsstrukturen bezeichnen und die, weitergeführt im Textkomplex, Informationen lexikalisch miteinander vernetzen. (Schröder 2000: 399)

Innerhalb der Diskurslinguistik werden „Phänomene der Wortbildung, also morphologische Aspekte von Diskursen“ (Spitzmüller/Warnke 2011: 138), durchaus als beschreibungsrelevant erachtet.⁸ Dennoch lassen sich nur vereinzelt diskurslinguistische Arbeiten finden, die Wortbildungen intensiver

6 Siehe in Ansätzen unter anderem Gersbach/Graf (1984); Schröder (1992); Iluk (1993); Pankow (1993) und Elsen/Michel (2010).

7 Ein Textkomplex besteht nach Schröder (2000: 389) „aus mehreren Texten zu einem gemeinsamen Thema“, die untereinander durch „referentielle Intertextualität“ verbunden sind und die sich „in ihrer Intention und in ihrer sprachlichen Ausformung“ voneinander unterscheiden.

8 Spitzmüller/Warnke (2011: 138, 201) ordnen Wortbildungen innerhalb ihres diskurslinguistischen Mehr-Ebenen-Analysemodells der wortorientierten Ebene zu.

und systematisch berücksichtigen. So untersucht Schmidt-Brücken (2015: Kap. 5.3) sogenannte generische Wortbildungen mit den Suffixen *-tum* (z. B. *Deutschtum*), *-schaft* (z. B. *Beamten-schaft*), *-ie* (z. B. *Sozialdemokratie*), *-heit* (z. B. *Christenheit*) und *-al* (z. B. *Großkapital*) in kolonialen Diskursen. Tereick (2016) zeigt, dass innerhalb des Klimawandel-Diskurses das Morphem „*Klima* als Wortbildungselement in allerlei Komposita [dient] [...]“. Insgesamt enthält ihr crossmediales Korpus über 1200 Wortbildungen mit *Klima*, die in ihrer Vielfalt allesamt im Anhang der Arbeit eingesehen werden können (vgl. Tereick 2016: 371–376). Diese erstrecken sich unter anderem über verschiedene Wortarten (Nomen: *Klimaabkommen*, Adjektiv: *klimaunfreundlich*, Verb: *klimaretten*), wobei *Klima* nicht nur als Erst-, sondern auch als Zweitglied verwendet wird (z. B. *Arktis-Klima* und *Weltklima*). Römer (2017) verdeutlicht in seiner Arbeit zu Wirtschaftskrisen mit einer Zusammenstellung von Wortbildungen, in denen das Wort *Krise* als Determinatum fungiert, welche Vorstellungen innerhalb der sogenannten Ölkrise von 1973/74 vorherrschen. Unter den zehn häufigsten Wortbildungen finden sich unter anderem *Energiekrise*, *Ölkrise*, *Versorgungskrise*, *Nahost-Krise/Nahostkrise* und *Wirtschaftskrise* (vgl. Römer 2017: 190). Spitzmüller/Warneke (2011: 145) betonen explizit, dass insbesondere Okkasionalismen – und das sind die meisten Belege, die in Tereick (2016) und Römer (2017) aufgelistet sind, –

als Wortbildungsprodukte von besonderem Interesse für die Diskurslinguistik [sind], weil sie Einstellungen konstituieren, Ironie markieren usw., mithin kontextuelle und pragmatische Dimensionen von Bedeutung kennzeichnen.

Insgesamt lässt sich allerdings festhalten, dass die diskurslinguistische Wortbildungsforschung insgesamt noch in den Kinderschuhen steckt und die traditionellen Beschreibungseinheiten wie Schlagwörter, Metaphern oder Topoi/Argumentationsmuster noch im Mittelpunkt stehen.

- 5) Auch der Zusammenhang zwischen Wortbildung und Sprachkritik ist nur vereinzelt herausgearbeitet worden (siehe Donalies 2003). Hierbei ist zu fragen, welche Rolle Wortbildungen als Mittel der Bewertung innerhalb bestimmter gesellschaftlicher Gruppen spielen (z. B. im Zusammenhang mit dem zum Unwort des Jahres 2015 gewählten Kompositum *Gutmensch* und dessen Varianten wie *Gutmenschentum* und *Gutbürger* oder auch im Hinblick auf den sprachkritischen Diskurs über Wortbildungen mit *Asyl-* wie *Asylkritik*, *Asylkritiker*, *Asylgegner* und *Asyldebatte*, siehe Niehr 1996, 2014: 57–62).

4. Perspektiven einer textsortenorientierten Wortbildungsforschung: Das Projekt „Wortbildung in Mündlichkeit und Schriftlichkeit“

Ausgehend von den herausgearbeiteten Desideraten und Mängeln der bisherigen Forschung lassen sich Perspektiven aufzeigen, die für eine Erforschung der Beziehungen zwischen Wortbildung und Text berücksichtigt werden sollten und die die Ausgangsbasis meines sich im Entstehen begriffenen Habilitationsprojekts mit dem Titel „Wortbildung in Mündlichkeit und Schriftlichkeit“ bilden (siehe Stumpf in Vorbereitung). Das Projekt zielt auf eine geschlossene theoretische wie empirische Gesamtdarstellung der Beziehungen zwischen Wortbildung und Text ab, wobei die Untersuchung über einzelne Texte hinausgeht und systematisch sowohl Einzeltexte und Textkomplexe/„Puzzletexte“ als auch Textsorten, Textsortenketten/-netze und Diskurse mitberücksichtigt. Die Auswertungen basieren auf einer Materialauswahl unterschiedlicher Textsorten, da „nur breitangelegte, an verschiedensten Textsorten orientierte Untersuchungen [...] zeigen [können], welchen Stellenwert die einzelnen Wortbildungstypen tatsächlich haben“ (Elsen/Michel 2007: 9).

Die ausgewählten Textsorten decken das gesamte Spektrum des Nähe-Distanz-Modells von Koch/Oesterreicher (1985, 2008)⁹ ab. Dass ein solches Vorgehen unerlässlich ist, um die Zusammenhänge zwischen Wortbildung und mündlichen sowie schriftlichen Texten aufzudecken, heben Elsen/Michel (2011: 4) deutlich hervor. Die Textsortenauswahl erstreckt sich daher von medial schriftlichen/konzeptionell geschriebenen (z. B. Zeitungsberichte) und medial mündlichen/konzeptionell geschriebenen (z. B. Berichte in Fernsehnachrichten) bis hin zu medial schriftlichen/konzeptionell gesprochenen (z. B. WhatsApp-Freizeitchat) und medial mündlichen/konzeptionell gesprochenen Textsorten (z. B. Alltagsgespräche). Dabei finden auch die konzeptionellen Zwischenstufen systematisch Berücksichtigung (z. B. durch Textsorten wie Wikipedia-Löschdiskussionen und Hochschulprüfungsgespräche). Bei der Korpuserstellung kann zum Teil auf bereits existierende Materialsammlungen zurückgegriffen werden. So sind in der „Datenbank für Gesprochenes Deutsch“¹⁰ des Instituts für Deutsche Sprache (IDS) Aufnahmen und Transkriptionen verschiedener Gesprächssorten vorhanden. Zudem stehen WhatsApp-Kommunikationsverläufe zur Verfügung, die im Projekt „What’s Up Deutschland“¹¹ gesammelt werden.

9 Zu neuesten Ansätzen rund um das Nähe-Distanz-Modell siehe Feilke/Hennig (2016).

10 https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome (Stand 17.12.2017).

11 https://www2.uni-mannheim.de/1/presse_uni_medien/pressemitteilungen/2014/November/Whats%20up%20Deutschland/ (Stand 17.12.2017).

Das Habilitationsvorhaben ist durch eine computergestützte Vorgehensweise gekennzeichnet. Die Korpusanalyse erfolgt mithilfe einer digitalen Forschungsumgebung (FuD)¹², die es ermöglicht, Daten zu sammeln, dokumentieren, klassifizieren und quantifizieren.¹³ Dabei werden die einzelnen Wortbildungen einer detaillierten Analyse unterzogen, die verschiedene Sprachbeschreibungsebenen in den Blick nimmt. Als Hintergrundfolie dienen die Modellierungsschritte, die Fleischer/Barz (2012: 73) „[f]ür die Zuordnung einer Wortbildung zu einem Modell“ anführen. Dabei werden strukturelle, morphologische und semantische Merkmale berücksichtigt (z. B. morphologischer Status der unmittelbaren Konstituenten, Wortbildungsart und funktional-semantische Klasse). Diese werden ergänzt durch weitere Analysekategorien wie Anzahl der enthaltenen Morpheme oder Vorhandensein von fremdsprachlichen Elementen.

Mithilfe der sich über das gesamte Kontinuum von Nähe und Distanz erstreckenden Textsorten kann folgenden zentralen Fragestellungen nachgegangen werden:

- 1) Vergleich von Wortbildungen in medial mündlicher und schriftlicher Sprache: Lassen sich Frequenzen und Spezifika feststellen, die nur im Mündlichen oder im Schriftlichen auftreten? (→ vertikale Perspektive auf das Koch/Oesterreicher-Modell)¹⁴
- 2) Vergleich von Wortbildungen in konzeptionell gesprochener und geschriebener Sprache: Wie verändert sich die Wortbildungsaktivität und -typik auf dem Kontinuum zwischen Nähe und Distanz? (→ horizontale Perspektive auf das Koch/Oesterreicher-Modell)

5. Beispielanalyse: Wortbildungen in der Textsorte Battle-Rap

5.1. Materialgrundlage, Untersuchungsziele und Vorgehen

Im Folgenden werden die Typik sowie die Funktionen von Wortbildungen innerhalb der Textsorte Battle-Rap vorgestellt. Die Auswertungen basieren

12 <http://fud.uni-trier.de/de/> (Stand 17.12.2017).

13 Das Textmaterial und die dazugehörigen Annotationen werden in Zusammenarbeit mit dem „Trier Center for Digital Humanities / Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften“ digital aufbereitet, archiviert und somit auch nachhaltig verfügbar gemacht: <http://kompetenzzentrum.uni-trier.de/de/> (Stand 19.12.2017).

14 Hierbei muss es sich nicht unbedingt nur um absolute Gegensätze handeln. Sollten sich zwischen Mündlichkeit und Schriftlichkeit deutliche Asymmetrien abzeichnen, wäre dies sicherlich auch eine interessante Beobachtung.

auf 37 Battle-Rap-Runden des sogenannten *JuliensBlogBattle* (JBB)¹⁵, welches das meistgesehene deutschsprachige Online-Battle-Rap-Turnier ist.¹⁶ Die Runden stammen aus den Jahren 2013, 2014 und 2015. Sie sind nach GAT 2¹⁷ transkribiert und umfassen eine Gesamtlänge von circa drei Stunden. Insbesondere werden die Texte aus den Viertelfinal-, Halbfinal- und Finalrunden berücksichtigt, da davon auszugehen ist, dass diese den Battle-Rap am besten repräsentieren.¹⁸

Die Battle-Rap-Texte werden im Hinblick auf die enthaltenen Wortbildungskonstruktionen untersucht. Zum einen wird aufgezeigt, inwiefern Wortbildungen in Battle-Raps als prototypische Textsortenmerkmale einen Beitrag dazu leisten, die „Textsorte zu charakterisieren und sie von anderen Textsorten zu unterscheiden“ (Fleischer/Barz 2012: 30). Hierbei werden diejenigen Wortbildungsverfahren und -besonderheiten zusammengefasst, die besonders häufig in den zugrunde gelegten Texten auftreten. Zum anderen werden die Funktionen von Wortbildungen innerhalb der Texte herausgearbeitet, wobei auch von textsortentypischen Funktionen gesprochen werden kann. Mit Battle-Raps steht dabei eine Textsorte im Mittelpunkt, die medial mündlich ist und auf konzeptioneller Ebene in Richtung kommunikativer Nähe tendiert.¹⁹ Um einen Einblick in die Merkmale der Textsorte Battle-Rap zu erhalten, steht der Beschreibung der Wortbildungstypik eine Charakterisierung der Textsorte mithilfe eines holistischen Beschreibungsansatzes in Anlehnung an Stein (2004) voran. Die Analyse basiert auf dem oben erwähnten Textkorpus (siehe hierzu auch Stumpf/Kämmer 2017: 176–179).

15 Kanal und Battle-Format wurden 2012 gegründet. Seit 2012 gibt es ein jährliches JBB-Turnier (mit Ausnahme des Jahres 2016, in dem Sewering ein Musikturnier veranstaltete, das sich nicht nur auf Rap-Battles konzentrierte).

16 So hat der YouTube-Kanal des Turniers 1,45 Millionen Abonnenten und 680 Millionen Aufrufe: <http://socialblade.com/youtube/user/juliensblogbattle> (Stand 19.12.2017).

17 Die Transkripte basieren auf den Konventionen für Minimaltranskripte (siehe Selting et al. 2009). Da mit der Wortbildung ein lexikalisches Phänomen im Mittelpunkt steht, werden bei der Transkription spezifischere prosodische Informationen nicht berücksichtigt. Lediglich der Fokusakzent wird vermerkt, da er für die im Battle-Rap typischen Reimstrukturen von Bedeutung ist.

18 Diese Runden werden fokussiert, da typische und repräsentative Exemplare von Battle-Rap-Texten und somit der Textsorte untersucht werden sollen. Es ist davon auszugehen, dass die Rezipienten den Textproduzenten zum Gewinner einer Runde küren, der für sie die Idee des Battle-Raps am besten umsetzt.

19 Es ist anzumerken, dass die vorliegende Untersuchung zunächst einmal eine Einzelanalyse darstellt. Die Ergebnisse sollen aber im Habilitationsprojekt mit den Ergebnissen zu anderen Textsorten verglichen werden.

5.2. Die Textsorte Battle-Rap: holistische Kurzcharakterisierung

Kommunikativ-pragmatische Charakteristika: Die Textsorte Battle-Rap kann in dem Handlungs- und Kommunikationsbereich der Musikszene verortet werden. Sie ist dort fester Bestandteil der Hip-Hop-Kultur. Die hier untersuchten Battle-Raps werden im Rahmen von sogenannten Online-Battle-Rap-Turnieren ausgetragen, die in Deutschland in den letzten Jahren sehr an Beliebtheit gewonnen haben und millionenfache Klickzahlen generieren.

Das für die Untersuchung ausgewählte *JuliensBlogBattle* (JBB)²⁰ ist ein Battle-Rap-Turnier, das auf dem gleichnamigen YouTube-Kanal von Julien Sewering stattfindet. Das System der Online-Battles dieses Kanals ist so aufgebaut, dass von den Teilnehmern Qualifikationsvideos eingesendet werden, aus denen der Kanalhaber die seiner Meinung nach besten veröffentlicht. Ein User-Voting entscheidet daraufhin, wer am Wettbewerb teilnehmen darf. In den einzelnen Runden treten einander zugeteilte Rapper gegeneinander an, indem sie jeweils ein Video einsenden, in dem sie den Gegner „battlen“. Diese Videos werden wiederum veröffentlicht und von Abonnenten sowie dem Turniergründer bewertet. Die Bewertungskategorien, nach denen sich entscheidet, wer eine Runde weiterkommt, haben sich im Laufe der Jahre zum Teil stark verändert. Kriterien, die jedoch immer eine Rolle spielen, sind das User-Voting sowie die Qualität des Videos und der Punchlines.²¹

Durch den Wettkampfcharakter des Battle-Rap-Turniers liegt ein Hauptziel der Textproduzenten darin, die Gunst der sie bewertenden Rezipienten zu gewinnen. Die Produzenten sind darauf aus, die Rezipienten davon zu überzeugen, sie zum Gewinner der jeweiligen Runde zu wählen. Die Battle-Rap-Texte sind demnach durch eine starke Appellfunktion geprägt. Es finden sich zwar auch informierende Anteile, wenn Rapper etwas über ihre eigene Person oder über die ihres Gegenübers berichten, die Informationsfunktion dient dabei aber immer der übergeordneten Appellfunktion. Auch eine poetische oder ästhetische Funktion ist in den Texten deutlich zu erkennen, da die kreative Nutzung und Umformung von Sprache eines der Hauptmerkmale des Battle-Raps ist.

Textstruktur: Die Kohärenz der Texte zeigt sich in einer Vielfalt und Dichte der Verknüpfung zwischen den einzelnen Sätzen, wobei explizite Wiederaufnahmen dominieren. *Das* zentrale Thema der Battle-Rap-Texte lässt sich wie folgt paraphrasieren: „Unterlegenheit des Gegners gegenüber dem Textproduzenten“ (Stumpf/Kämmer 2017: 178). Es geht also darum, selbst seine

20 <https://www.youtube.com/user/JuliensBlogBattle> (Stand 19.12.2017).

21 Es ist offensichtlich, dass es sich bei der Qualität des Videos und der Punchlines um recht subjektive Kriterien handelt. Es ist daher keine Seltenheit, dass die Entscheidungen des Turniergründers Sewering in der Community zuweilen auf größere Kritik stoßen.

(Rap-)Stärken aufzuzeigen (Boasten) und gleichzeitig den Kontrahenten in ein schlechtes Licht zu rücken (Dissen). Dieses Thema wird primär argumentativ entfaltet. So lassen sich immer wiederkehrende Argumentationsmuster finden (vgl. Stumpf/Kämmer 2017: 185–189). Zum Teil spielt auch die narrative Thementafelung eine Rolle (sogenannter *Realtalk*, d. h. Geschichten über die Vergangenheit des Kontrahenten oder die eigene). Diese dient aber immer der übergeordneten Argumentation. Der Sprechakt des Dissens wird meist in Form sogenannter Punchlines (von engl.: *punch* (Faustschlag) und *line* (Zeile, Vers)) umgesetzt. Punchlines sind essentieller Bestandteil des Battle-Raps und werden als „Zeilen verstanden, die Pointen generieren“ (Wolbring 2015: 342).

Äußere und sprachliche Charakteristika: Betrachtet man die materielle Textgestalt, so kann festgehalten werden, dass es sich bei den hier untersuchten Battle-Rap-Runden um eine multimodale Textsorte handelt, da diese in Musikvideos präsentiert werden.²² Die Texte haben in der Regel eine Länge von circa 500 bis 1000 Wörtern, Videos eine Länge von drei bis sieben Minuten.

Hinsichtlich des Textsortenstils steht die Sprache des Raps der gesprochenen Alltagssprache nahe, wobei mehr oder weniger verschiedene Substandard-Elemente in die Texte miteinfließen (z. B. Drogensprache, Fäkalsprache, Jugendsprache) (vgl. Androutsopoulos 2003: 120 und Deppermann/Riecke 2006: 165). Bezüglich der Syntax zeigt sich, dass die Sätze innerhalb der untersuchten Battle-Rap-Texte für gewöhnlich keine hohe Komplexität aufweisen. Im Bereich der Lexik finden sich viele Anglizismen (*Skills, Image*) und Substantive mit pejorativer lexikalischer Bedeutung (*Nutte, Hurensohn*). Signifikante Wortfelder sind die der Sexualität (*ficken, Blowjob*) und der Gewalt (*morden, zerfleischen*). Weitere sprachlich-stilistische Besonderheiten sind vor allem Wortspiele jeder Art.

Einzeltextübergreifende Merkmale: Die Battle-Rap-Texte zeichnen sich durch zahlreiche intertextuelle Bezüge aus.²³ Diese zeigen sich darin, dass auf vorherige Runden des Turniers oder auch auf Filme und Fernsehserien Bezug genommen wird. Der Bezug auf vorherige Runden dient in fast allen Fällen dazu, die vom Gegner angebrachten Argumente zu entkräften oder ihn durch seine Vorgeschichte im Battle negativ darzustellen. Im Rahmen der intertextuellen Verweise auf die Bereiche Film und Fernsehen wird meist auf einzelne fiktive Personen zurückgegriffen, wobei vorteilhaftere oder stärkere Charaktere

22 In den Musikvideos sind meist die Rapper selbst zu sehen sowie typische Statussymbole des (Gangster-)Raps wie Schmuck, Autos, Waffen, Drogen und Frauen.

23 Unter Intertextualität werden Anspielungen auf andere Texte verstanden (vgl. Fix 2008: 31). Der Begriff der Intertextualität muss im Hinblick auf Battle-Rap-Texte jedoch etwas weiter gefasst werden, da sich die Verweise oft auf Bereiche beziehen, die nicht dem prototypischen Textbegriff entsprechen (z. B. Filme).

vom jeweiligen Produzenten für sich beansprucht werden und der Gegner im Sinne des Dissens mit unvoreilhafteren Charakteren verglichen wird (so finden sich beispielsweise Anspielungen aus den Filmen/Serien *Matrix*, *Herr der Ringe*, *Star Wars* und *Die Simpsons*). Als Mustervariation/Musterdurchbrechung können (ungewöhnliche) Gastauftritte anderer Rapper, die Länge der Runden, Selbstironie sowie (extreme) Übertreibungen angeführt werden.

5.3. Überblick über textsortentypische Wortbildungen des Battle-Raps

Betrachtet man die Wortbildungen innerhalb von Battle-Rap-Texten, so lassen sich folgende Besonderheiten feststellen, die für die Textsorte charakteristisch erscheinen:

- 1) **Polymorphemische Konstruktionen:** In Bezug auf die Länge und Komplexität der Wortbildungen zeigt sich, dass polymorphemische Nominal-Komposita besonders typisch für Battle-Rap-Texte sind. So finden sich im Korpus zahlreiche Wortbildungsprodukte, die sich aus drei (z. B. *Großstadtecken*, *Kandiszuckerstücke*, *Paarungstrieb*, *Abschlussball*, *Strichcodezeichen*, *Sehvermögen* usw.) oder sogar aus mehr als drei Morphemen zusammensetzen (z. B. *Waldbrandrettungskräfte*, *Gokartstreckenzielgeraden*, *Gayclubbefürworter*, *Halloweenverkleidung*, *Krabbenkokainverteiler* usw.).
- 2) **Okkasionelle Wortbildungen:** Auffällig ist auch der hohe Anteil okkasioneller Wortbildungen. Sie werden im Rahmen des jeweiligen Themas der Texte konstruiert und sind in den meisten Fällen auch nur mithilfe des Kontextes sowie aufgrund von Hintergrundwissen verstehbar. Als Beispiele können *Veggieburgerflow*, *Schachbrettquadrate*, *Mittelpunktneurosen*, *Kindergartenpunchlines* und *Hurensohnuniform* angeführt werden. Wie zu erkennen ist, handelt es sich bei diesen auch meist um komplexe Konstruktionen, die aus mindestens drei Morphemen bestehen.
- 3) **Hybridbildungen:** Die Textsorte ist zudem stark von Hybridbildungen gekennzeichnet, die sich aus deutschen und englischen lexikalischen Einheiten zusammensetzen, fast ausschließlich als Okkasionalismen zu interpretieren sind und häufig auch aus mehr als zwei Morphemen bestehen (z. B. *Fischkutterboys*, *Gasthook*, *Schwammconnection*, *Fakegeschichten*, *famegeil*, *Cocklecker*, *Gammelslut*, *Imagewechsel* usw.). Die Hybridbildungen reihen sich ein in die generelle Tendenz von Rap-Texten, stark auf die englische Sprache zurückzugreifen, was wohl vor allem darauf zurückzuführen ist, dass Rap seinen Ursprung in den USA hat und zahlreiche Begrifflichkeiten des Hip-Hops internationalen bzw. universellen Charakter besitzen.
- 4) **Kontaminationen:** Textsortentypisch ist ferner das relativ häufige Vorkommen der für das Standarddeutsche eher peripheren Wortbildungsart

der Kontamination, „bei der Teile meist zweier Wörter zu einem neuen verschmelzen, das dann Bedeutungsaspekte beider Wörter besitzt“ (Elsen 2014: 31). Kontaminationen sind im Deutschen selten und fast ausschließlich okkasionell. Man findet sie in der Regel in Texten, die ein starkes sprachlich-kreatives Moment besitzen (z. B. literarisch-poetische Texte, Werbetexte, Kabarett-Texte, journalistische Texte usw.) (vgl. Elsen 2008: 115–117 und Müller/Friedrich 2011: 74). Auch in Battle-Rap-Texten werden sie vor allem als sprachspielerisches und humorvolles Element eingesetzt (siehe Kapitel 5.4). Als Beispiele lassen sich *Spongearogunz* (Verschmelzung aus den beiden Rappernamen *SpongeBOZZ* und *Punch Arogunz*), *Kinguin* (Verschmelzung aus dem Anglizismus *King* und *Pinguin*) sowie *Tunefisch* (Verschmelzung aus dem Rappernamen *Tune* und *Thunfisch*) anführen.

Insgesamt lässt sich festhalten, dass vor allem okkasionelle Nominal-Komposita typisch für die Textsorte Battle-Rap sind, die in der Regel aus mindestens drei Morphemen bestehen und darüber hinaus nicht selten englischsprachige Lexeme enthalten. Diese wortbildungsspezifischen Besonderheiten ergeben sich dabei direkt aus dem Wettbewerbscharakter der Textsorte. So ist die Hauptintention der Textproduzenten, die Rezipienten von ihrem sprachlichen Können beispielsweise in Form von ungewöhnlichen und sprachlich-kreativen Gelegenheitsbildungen zu überzeugen, um sich dadurch ihren Kontrahenten gegenüber als bessere Rapper darzustellen (vgl. Wolbring 2015: 341f.).²⁴

5.4. Funktionen von Wortbildungen im Battle-Rap

Aus funktionaler Perspektive lassen sich drei Aspekte anführen, die in Bezug auf den Gebrauch von Wortbildungen in Battle-Rap-Texten eine Rolle spielen:

- 1) **Informationsverdichtung:** Die Funktion der Informationsverdichtung gilt generell für jede Wortbildung (vgl. Erben 2006: 25). Darunter ist zu verstehen, dass mithilfe von Wortbildungen das, was syntaktisch gesehen auf mehrere Elemente verteilt werden müsste, in einer einzigen Einheit zusammengefasst werden kann (vgl. Naumann 2000: 27). In Battle-Rap-Texten kommt diese Funktion jedoch in besonderer Weise zum Tragen, was daran liegt, dass die Textproduzenten auf einen relativ engen Raum begrenzt sind. Zum einen müssen die Rapper beachten, dass ihr Song nicht den zeitlichen Rahmen sprengt, und zum anderen sind sie auch innerhalb ihrer Texte

²⁴ Zukünftige Untersuchungen sollten der Frage nachgehen, ob die Textsortenspezifik der Wortbildungsverfahren im Battle-Rap auch in anderen Textsorten festzustellen ist. Da Battle-Rap als eine Form der *oral literature* angesehen werden kann, wäre unter anderem ein Vergleich mit weiteren Textsorten aus diesem Bereich lohnenswert.

auf einzelne Verse beschränkt. Ziel der Rapper ist es daher, möglichst viel Inhalt bzw. viele Informationen in ihrem Text bzw. in den Versen unterzubringen. Um dies zu erreichen, greifen sie auf (komplexe) Wortbildungskonstruktionen zurück. So verwendet beispielsweise der Rapper Laskah im folgenden Textbeleg das Kompositum *Fishboneline* statt der syntaktischen Umschreibung *deine Punchline, in der du die Fishbone-Kleidung von Gary Washington schlecht machst*²⁵ (siehe Beispiel 3)²⁶:

(3) **JBB 2014 – Halbfinale – Laskah vs. Aytee**

070 LK: und alexander man du DACHtest deine **fishbo-**
neline zerstört gary
071 NEIN
072 denn vor paar jahren liefst du selber rum
in sir benny MILEs
073 dicke BLING bling in den ohren
074 und hoch gegelte HAAre
075 aytee du warst nur ein KOLlegah für arme

- 2) **Reimstruktur:** Eine angesehene Fertigkeit innerhalb des Battle-Raps ist es, am Ende von Verszeilen nicht nur eine Silbe aufeinander zu reimen, sondern mehrsilbige Reime zu kreieren. Es zeugt von künstlerischem Können, wenn Takt, Betonung der Wörter und Silbenanzahl zusammenpassen und klangliche Ähnlichkeiten zwischen aufeinanderfolgenden Verszeilen gegeben sind (beispielsweise in der Form, dass die Vokale in den Reimen gleich sind). Besondere „Skills“ zeigt daher derjenige, der möglichst lange (aber kontextuell sinnvolle) Mehrfachreime bilden kann (vgl. Wolbring 2015: 322f.). Diese Mehrfachreime können sich durch die syntaktisch sinnvolle Reihenfolge der Wörter in einem Vers, aber auch durch Wortbildungen ergeben, die wiederum oft nur für diesen Zweck konstruiert werden. Wortbildungen dienen im Battle-Rap demnach häufig gegen Ende von Versen der mehrsilbigen Reimstruktur.²⁷ Dabei können sich zwei Wortbildungen aufeinander reimen wie in Beispiel 4, in dem mit *Koalabär* und *Oralverkehr* ein viersilbiger Reim vorhanden ist. Es können aber auch eine Wortbildung und eine aus mehreren Wörtern zusammengesetzte

25 Im Viertelfinale standen sich die beiden Rapper Gary Washington und Aytee gegenüber. In diesem Battle bezog sich Aytee unter anderem in einem Vers auf den Kleidungsstil seines Kontrahenten.

26 In den nachfolgenden Transkriptionsbeispielen werden die Wortbildungen, auf die näher eingegangen wird, (fett) hervorgehoben.

27 Dass solche mehrsilbigen Reime aufgrund ihrer Komplexität häufig nicht vollkommen rein sind, liegt auf der Hand. Durch die jeweilige Betonung und den Rhythmus gelingt es guten Rappern aber, auch etwas unreinere Silben geschickt aneinander zu reihen.

Wortgruppe, innerhalb derer auch Wortbildungen auftreten können, als Ausgangseinheiten für den Reim fungieren wie in Beispiel 5, in dem mit *Gio macht Ferien* und *Mikrobakterien* ein sechssilbiger Reim zu finden ist. In der Regel wird innerhalb solcher mehrsilbigen Reime auch der Hauptakzent jeweils auf der gleichen Silbe gesetzt (in Beispiel 4 auf der zweiten und in Beispiel 5 auf der vierten Silbe):

(4) JBB 2015 – Viertelfinale – Jay Jiggy vs. EnteTainment

014 JJ: du penner ich besuche deine clique und gehe
los auf den

koAlabär

015 und breche ihm den kiefer mit rigorosem

oRALverkehr

(5) JBB 2013 – Finale (HR) – SpongeBOZZ vs. Gio

072 SB: jeder normale mensch geht arbeiten aber
!GI:!o

073 **macht !FE:!rien**

074 wir zahlen steuern doch er liegt uns nur
auf der tasche wie **MIkrobak!TE:!ri:en**

Eine besondere Art ist die Verwendung von mehrsilbigen Reimen in mehreren Versen nacheinander. Diese Reimketten können sich teilweise über recht lange Textabschnitte erstrecken. Im folgenden Beispiel gebraucht der Rapper EnteTainment in fünf aufeinanderfolgenden Versen achtsilbige Reime, die er zum einen aus mehreren Wörtern (*Großstadtecken Zivilklagen, Flohmarktdecken liegend Bares, Strothmann exen niederschlagend*) und zum anderen aus komplexeren Wortbildungen (*Kokapäckchendiwaage, Gokartstreckenzielgeraden*) erzeugt. Zu beachten ist hierbei auch der Hauptakzent, der innerhalb der Reimketten zunächst auf der fünften Silbe (*zi, di*) und in den darauffolgenden drei Versen auf der ersten Silbe liegt (*floh, stroth, go*):

(6) JBB 2014 – Achtelfinale – EnteTainment vs. Vocal

046 ET: ich sammel in **großstadtecken ZIvilklagen**

047 mach mit der **kokapäckchenDIGIwaage**

048 du nur auf **FLOHmarktdecken liegend bares**

049 ich komm dich nach dem **STROTHmann exen
niederschlagen**

050 und erreiche die nächste runde binnen se-
kunden wie

GOKartstreckenzielgeraden

- 3) **Kreativer Sprachgebrauch/Wortspiele:** Die Battle-Rap-Texte zeichnen sich durch eine hohe Anzahl an Wortspielen im weitesten Sinne aus (vgl. Winter-Froemel 2009: 1429). Sie stellen somit „sprachlich ausgesprochen

kreative Produkte“ (Ehrhardt 2005: 33) dar, innerhalb derer „die deutsche Sprache auf geistreiche, originelle und/oder intelligente Weise“ (ebd.) verwendet wird. Die wortspielerischen Elemente dienen in erster Linie als Ausdruck sprachlicher Kreativität der Rapper und somit als Beitrag zur poetischen Funktion der Battle-Rap-Texte (vgl. Wolbring 2015: 341f.). Das Spektrum der Typen an Wortspielen ist dabei äußerst weit (z. B. modifizierte Phraseme, *wie*-Vergleiche, Homophonien usw.), wobei unabhängig vom konkreten Verfahren meistens das Spielen mit Ambiguitäten im Mittelpunkt steht (vgl. Winter-Froemel 2013: 131–141). Aus der Perspektive der Wortbildung sind es insbesondere sprachspielerische, ungewöhnliche, gegebenenfalls bestimmte Wortbildungsmuster widersprechende sowie okkasionelle Wortbildungen, mit denen der Rapper seine Wortgewandtheit und seine (sprachliche) Überlegenheit gegenüber seinem Gegner demonstriert. So wird – wie oben bereits erwähnt – nicht selten auf die durchaus kreative Wortbildungsart der Kontamination zurückgegriffen. In Beispiel 7 werden der Rappename des Gegners (*Gio*) und das aus dem Französischen entlehnte Wort *Guillotine* miteinander verschmolzen:

(7) JBB 2014 – Halbfinale – Diverse vs. Gio

043 DV: dennoch lass ich KÖPfe rollen im battle
044 hiernach nennt man mich die **gioTine**

In Beispiel 8 wird mit einer Homophonie gespielt. *Hammer* innerhalb der Wortbildung *Hammerlines* lässt sich einerseits als Affix(oid) (steigernd im Sinne von ‚sehr gut, super‘) und andererseits als freies lexikalisches Grundmorphem (mit der Bedeutung ‚Werkzeug zum Schlagen‘) interpretieren. Die zweite Bedeutung wird zudem durch die Lexeme *Butterflies* und *Axt* hervorgerufen bzw. unterstützt. Für diese Lesart ist aber auch Hintergrundwissen notwendig. Denn sie spielt auf das Image des Gegners an, dessen Markenzeichen es ist, in seinen Videos den Gegnern nicht nur sprachlich, sondern auch physisch mittels diverser Werkzeuge wie Axt, Säge und Hammer Schaden zuzufügen (sogenannter Splatter-Battle-Rap).

(8) JBB 2015 – Halbfinale – Casa vs. EnteTainment

129 CS: jetzt kommst du wieder mit paar **HAMmerlines**
130 in denen du mich brutal mit deinen BUTter-
flies und axt
zerfleischst

Zwei weitere kreative Wortbildungen finden sich in Beispiel 9. Die beiden Komposita *Schwanzflossenbros* und *Fischkutterboys* sind eine Anspielung auf den tatsächlichen Namen der gegnerischen Rap-Crew, die sich *Lachsneckengang* nennt. Analog zum Ursprungsnamen werden die Wortbildungseinheiten sowohl im Hinblick auf ihren Inhalt als auch in

Bezug auf ihre sprachliche Herkunft eins zu eins ersetzt (dt. *Lachsnacken* → *Schwanzflossen* und *Fischkutter*; engl. *Gang* → *Bros* und *Boys*).

(9) JBB 2015 – Finale – EnteTainment vs. Neo

- 072 ET: ne schwule BOYgroup deren richtung nich
läuft
073 nennt euch die **SCHWANZflossenbros**
074 oder **FISCHKutterbo:ys**

Ein letztes Beispiel zeigt das Einsetzen von Wortbildungen im Rahmen des sogenannten Spittens. Dies ist eine Rap-Technik, bei der sich aufeinanderfolgende Wörter reimen und gleich bzw. ähnlich geschrieben/betont werden, jedoch unterschiedliche Bedeutungen besitzen (vgl. Wolbring 2015: 324f.). Ein Spit befindet sich in Textbeispiel 10. In 022 ist das Konvertat *Schwerwiegende* ein Substantiv und *folgen* ein Verb, in 024 stellt *schwerwiegende* ein Adjektivattribut zu dem substantivierten *Folgen* dar. Die Homophonie wird zudem durch die gleiche Akzentsetzung begünstigt, die in beiden Versen auf der ersten Silbe liegt (*schwer*):

(10) JBB 2015 – Finale – Neo vs. EnteTainment

- 022 Neo: die runde hat für dich **!SCHWER!wiegende
folgen**
023 du verkaufst dein merchandise nur in ÜBer-
größen
024 weil dir nur **!SCHWER!wiegende folgen**

5.5. Sprachkritische Perspektive

Die Verwendung von bestimmten (okkasionellen) Wortbildungen wird zum Teil entweder von den gegnerischen Rappern oder vom Turniergründer Se-wering, der die Runden bewertet, kritisiert. Dies ist vor allem dann der Fall, wenn es sich bei diesen um sogenannte Zweckreime handelt. Als Zweckreime bezeichnet man im Rap „auffällig ungebräuchliche Formulierungen, die dem Anschein nach nur des Reimes wegen eingesetzt werden“ (Wolbring 2015: 327). Hierbei werden häufig morphologisch recht komplexe Wortbildungen verwendet, die sich zwar auf vorangehende Verse reimen, die aber nichtssagend, inhaltlich unpassend, unlogisch oder schlicht redundant erscheinen.²⁸ Ein solcher Zweckreim, der in der Analyse auch konkret zu Punktabzügen

28 Dabei ist zu betonen, dass es in bestimmten Fällen sicherlich nur schwer möglich ist, Zweckreime analytisch wohlbegründet zu bestimmen. So scheint die Frage nach zweckgereimten Zeilen auch von der Fantasie oder dem Vorwissen des jeweiligen Rezipienten abhängig zu sein.

geführt hat,²⁹ findet sich in Beispiel 12 mit der Wortbildung *stubenreiner*, die lediglich für Reimzwecke in den Vers eingebaut wurde (als Reim auf *du dir deine*), im Grunde aber aus inhaltlicher Perspektive wenig Sinn ergibt bzw. überflüssig ist (warum sollte es in diesem Kontext von Bedeutung sein, dass der Suchhund stubenrein ist?):

(12) JBB 2014 – Viertelfinale – Gary Washington vs. Aytee

105 GW: dein crackkonsum muss nich mal ein **stuben-**
reiner SUCHhhund
 wittern
 106 digga durch mein klunker könntest du dir
 deine ZUKunft sichern

Wer solche Zweckreime konstruiert, kann sich der Kritik seines Gegners sicher sein. So prangert in Beispiel 13 der Rapper SpongeBOZZ die *hingezweckten Aussagen* seines Gegenübers an und verweist darauf, dass dieser für Reimzwecke Wörter (in Form von Wortbildungen) erfindet.

(13) JBB 2013 – Finale (RR) – SpongeBOZZ vs. Gio

031 SB: ich schenk dir von dem battlecontestgeld ein
 deutsches LExikon
 032 denn außer einer **hingezweckten AUSSage**
 bleibt nix
 033 du **erfindest WÖRter** gio
 034 **hauptsache REIMT sich**

Als sprachkritisch kann auch folgende Auseinandersetzung betrachtet werden, die sich innerhalb des JBB 2014 über mehrere Runden erstreckt. Ausgangspunkt ist ein Vers im Halbfinale des Turniers, in dem der Rapper Laskah sein Gegenüber als *rechtsradikalisch* bezeichnet (siehe Beispiel 14). Dass diese Wortbildung nicht der Norm entspricht (siehe STEIN 2007), ist Laskahs Finalgegner nicht entgangen, weshalb dieser auf recht humorvolle Art und Weise darauf Bezug nimmt. Er bildet analog zu *rechtsradikalisch* die ebenfalls normwidrige Konstruktion *unintelligentisch* (siehe Beispiel 15). Laskah reagiert in der Rückrunde des Finales auf diese Unterstellung und versucht den Wortbildungsverstoß zu rechtfertigen, indem er behauptet, dieser sei volle Absicht gewesen (siehe Beispiel 16):

(14) JBB 2014 – Halbfinale – Laskah vs. Aytee

077 LK: keiner checkte dass ne eightyEIGHT in dei-
 nem namen is
 078 willst du uns damit sagen dass du **RECHTS-**
radikalisch bist

²⁹ Siehe <https://www.youtube.com/watch?v=hhHcHspWPSM#t=2m53s> (Stand 19.12.2017).

(15) JBB 2014 – Finale (HR) – Gio vs. Laskah

- 024 Gio: dass durch die ganze kifferei auch heute
noch dein kopf
geFICKT is
- 025 denn du hundesohn erKENNST nicht
- 026 aytee **rechtsradikalisch** zu nennen is echt
unintelligentisch

(16) JBB 2014 – Finale (RR) – Laskah vs. Gio

- 161 LK: pass ma auf die line in der ich aytee
RECHTSradikalisch nenn
- 163 macht sinn weil ich sage dass ich auch
ZWECKreime haben kann

6. Zusammenfassung

Der Blick auf die Forschung zu Wortbildung und Text hat gezeigt, dass es insbesondere die textkonstitutiven und -distinktiven Eigenschaften von Wortbildungen sind, die bislang untersucht wurden. Textkonstitutiv sind Wortbildungen, da sie Anteil an der Textkohäsion haben, indem Grundmorpheme in verschiedenen Wortbildungskonstruktionen innerhalb des Textes wieder aufgegriffen werden. Die daraus hervorgehenden Isotopieketten haben einen entscheidenden Anteil an der Textkohärenz. Wie anhand der Beispielanalysen zu sehen ist, gilt die textkonstitutive Funktion sowohl für schriftliche als auch für mündliche Texte. Textdistinktiv heißt demgegenüber, dass Wortbildungen die sprachliche Spezifik von Textsorten prägen können. Dies wurde im vorliegenden Artikel exemplarisch anhand der Textsorte Battle-Rap verdeutlicht.

Ausgehend vom Forschungsstand lassen sich diverse Desiderate ableiten, die an dieser Stelle nochmals stichpunktartig zusammengefasst werden:

- Die meisten Studien stellen eng gesteckte Einzelanalysen dar. Eine umfassende Darstellung der Zusammenhänge zwischen Wortbildung und Text auf der Grundlage breiterer Materialsammlungen steht noch aus.
- Bislang orientieren sich die Auswertungen stark an Texten, die keiner authentischen Kommunikationssituation entspringen (z. B. literarische Texte und Werbetexte).
- Texte, die medial mündlich und/oder konzeptionell gesprochen sind, werden weitgehend vernachlässigt. Die bisherigen Untersuchungen – insbesondere diejenigen zur textkonstitutiven Funktion von Wortbildungen – basieren fast ausschließlich auf Schrifttexten.
- Es existieren nur wenige Arbeiten, die über den Einzeltext oder die einzelne Textsorte hinausblicken. Analysen zum Vorkommen und zur Funktion von Wortbildungen in Textkomplexen, Textsortennetzen und Diskursen sind die Ausnahme.

- Zuletzt fehlt auch eine stärkere Verbindung von Wortbildung und Sprachkritik aus text- und diskurslinguistischer Perspektive.

Im Rahmen meines Habilitationsprojekts „Wortbildung in Mündlichkeit und Schriftlichkeit“ werden die aufgezeigten Probleme überwunden, indem auf der Grundlage korpusbasierter Auswertungen ein Überblick über die vielfältigen Beziehungen zwischen Wortbildung und Text angestrebt wird (siehe STUMPF in Vorbereitung).

Literatur

- Agricola, Eberhard. 1969. *Semantische Relationen im Text und im System*. Halle: Niemeyer.
- Androutsopoulos, Jannis K. 2003. HipHop und Sprache: Vertikale Intertextualität und die drei Sphären der Popkultur. In Jannis Androutsopoulos (ed.), *HipHop: globale Kultur – lokale Praktiken* (Cultural studies 3), 111–136. Bielefeld: Transcript.
- Barz, Irmhild & Marianne Schröder. 2001. Grundzüge der Wortbildung. In Wolfgang Fleischer, Gerhard Helbig & Gotthard Lerchner (eds.), *Kleine Enzyklopädie Deutsche Sprache*, 178–217. Frankfurt am Main: Peter Lang.
- Barz, Irmhild, Marianne Schröder, Karin Hämmer & Hannelore Poethe. 2007. *Wortbildung – praktisch und integrativ. Ein Arbeitsbuch*. 4. überarbeitete Auflage (Leipziger Skripten 2). Frankfurt am Main: Peter Lang.
- Busse, Dietrich & Wolfgang Teubert. 1994. Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In Dietrich Busse, Fritz Hermanns & Wolfgang Teubert (eds.), *Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik*, 10–28. Opladen: Westdeutscher Verlag.
- Crestani, Valentina. 2010. *Wortbildung und Wirtschaftssprachen – Vergleich deutscher und italienischer Texte*. Bern: Peter Lang.
- Dederding, Hans-Martin. 1983. Wortbildung und Text. Zur Textfunktion (TF) von Nominalkomposita (NK). *Zeitschrift für germanistische Linguistik* 11. 49–64.
- Deppermann, Arnulf & Andrea Riecke. 2006. Krieg der Worte – Boasten und Dissen im HipHop-Battle. In Birgit Richard & Klaus Neumann-Braun (eds.), *Ich-Armeen. Täuschen – Tarnen – Drill*, 157–165. München: Wilhelm Fink.
- Donalies, Elke. 2003. „Gebt endlich die Wortbildung frei!“ Über unsinnige und sinnige Kritik an der Wortbildung. *Sprachreport* 19. 26–32.
- Donalies, Elke. 2005. *Die Wortbildung des Deutschen. Ein Überblick*. 2. überarbeitete Auflage (Studien zur deutschen Sprache 27). Tübingen: Narr.
- Donalies, Elke. 2011. *Basiswissen Deutsche Wortbildung*. 2., überarbeitete Auflage. Tübingen: Francke.
- Ehrhardt, Claus. 2005. „Jetzt mal ohne Spaß“. Zur diskursbildenden Potenz von Phraseologismen in deutschen Raps. *Der Deutschunterricht* 57. 33–43.
- Eichinger, Ludwig M. 1995. Wegweiser durch Textwelten. Wozu komplexe Substantive gut sind. In René Métrich & Marcel Vuillaume (eds.), *Rand und Band. Abgrenzung*

- und Verknüpfung als Grundtendenzen des Deutschen* (Eurogermanistik 7), 169–182. Tübingen: Narr.
- Eichinger, Ludwig, M. 2000a. *Deutsche Wortbildung. Eine Einführung* (Narr-Studienbücher). Tübingen: Narr.
- Eichinger, Ludwig M. 2000b. Verstehen und Spaß haben. Wortbildung im literarischen Text. In Irmhild Barz, Marianne Schröder & Ulla Fix (eds.), *Praxis- und Integrationsfelder der Wortbildungsforschung* (Sprache – Literatur und Geschichte 18), 145–158. Heidelberg: Winter.
- Elsen, Hilke. 2008. Kontamination im Randbereich der deutschen Grammatik. *Deutsche Sprache* 36. 114–126.
- Elsen, Hilke. 2011. Das besondere Funktionsspektrum der Wort(neu)bildung in der phantastischen Kinderliteratur. In Hilke Elsen & Sascha Michel (eds.), *Wortbildung im Deutschen zwischen Sprachsystem und Sprachgebrauch. Perspektiven – Analysen – Anwendungen* (Perspektiven germanistischer Linguistik 5), 211–223. Stuttgart: Ibidem.
- Elsen, Hilke. 2014. *Grundzüge der Morphologie des Deutschen. 2.*, aktualisierte Auflage (De Gruyter Studium). Berlin & Boston: de Gruyter.
- Elsen, Hilke & Sascha Michel. 2007. Wortbildung im Sprachgebrauch. Desiderate und Perspektive einer etablierten Forschungsrichtung. *Muttersprache* 117. 1–16.
- Elsen, Hilke & Sascha Michel. 2010. Wortbildung in Sprechstundengesprächen an der Hochschule. In Nina Hinrichs & Annika Limburg (eds.), *Gedankenstriche – Reflexionen über Sprache als Ressource* (Stauffenburg Festschriften), 33–45. Tübingen: Stauffenburg.
- Elsen, Hilke & Sascha Michel. 2011. Wortbildung im Spannungsfeld zwischen Sprachsystem und Sprachgebrauch. Zur Einführung in diesen Band. In Hilke Elsen & Sascha Michel (Hgg.), *Wortbildung im Deutschen zwischen Sprachsystem und Sprachgebrauch. Perspektiven – Analysen – Anwendungen* (Perspektiven germanistischer Linguistik 5), 1–15. Stuttgart: Ibidem.
- Erben, Johannes. 1995. Wortbildung und Textbildung. In Heidrun Popp (ed.), *Deutsch als Fremdsprache. An den Quellen eines Faches. Festschrift für Gerhard Helbig zum 65. Geburtstag*, 545–552. München: Iudicium.
- Erben, Johannes. 2006. *Einführung in die deutsche Wortbildungslehre. 5.*, durchgesehene und ergänzte Auflage. Berlin: Erich Schmidt.
- Feilke, Helmuth & Mathilde Hennig (eds.). 2016. *Zur Karriere von „Nähe und Distanz“. Rezeption und Diskussion des Koch-Oesterreicher-Modells* (Reihe Germanistische Linguistik 306). Berlin & Boston: de Gruyter.
- Fix, Ulla. 2008. *Texte und Textsorten – sprachliche, kommunikative und kulturelle Phänomene*. (Sprachwissenschaft 5). Berlin: Frank & Timme.
- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache. 4.*, völlig neu bearbeitete Auflage (De Gruyter Studium). Berlin & Boston: de Gruyter.
- Gataullin, Ravil. 1990. Zu wortbildenden Textpotenzen. *Deutsch als Fremdsprache* 27. 240–245.

- Gersbach, Bernhard & Graf, Rainer. 1984. *Wortbildung in gesprochener Sprache*. 2 Bände (Idiomatica 12/13). Tübingen: Niemeyer
- Greule, Albrecht & Sandra Reimann. 2011. Von *Doppel- und Normaltuben* und dem *Anti-Enzym BX* – Wortbildung in der Hörfunkwerbung. In Hilke Elsen & Sascha Michel (eds.), *Wortbildung im Deutschen zwischen Sprachsystem und Sprachgebrauch. Perspektiven – Analysen – Anwendungen* (Perspektiven germanistischer Linguistik 5), 225–263. Stuttgart: Ibidem.
- Iluk, Jan. 1993. Zur Frequenz der Wortbildungsstrukturen in der gesprochenen Alltagssprache. Dargestellt am Beispiel der deutschen Privativa. In Günther Richter (ed.), *Methodische Grundfragen der Erforschung gesprochener Sprache* (Arbeiten zur Sprachanalyse 16), 91–97. Frankfurt am Main: Peter Lang.
- Imo, Wolfgang (2013): *Sprache in Interaktion. Analysemethoden und Untersuchungsfelder*. (Linguistik – Impulse und Tendenzen 49). Berlin & Boston: de Gruyter.
- Janich, Nina. 2009. Zur Analyse von Textsorten-in-Vernetzung. Eine Modelldiskussion an einem Fallbeispiel aus der Unternehmenskommunikation. *LAUD 734* (Series A: General & Theoretical Papers). 1–26.
- Kauffer, Maurice. 2006. Wie wird ein Text aufgebaut? Zur Rolle der Makrokomposita und der Kompositionsnetze. In Jean François Marillier, Martine Dalmás & Irmtraud Behr (eds.), *Text und Sinn. Studien zur Textsyntax und Deixis im Deutschen und Französischen. Festschrift für Marcel Vuillaume zum 60. Geburtstag* (Eurogermanistik 23), 49–66. Tübingen: Stauffenburg.
- Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36. 15–43.
- Koch, Peter & Wulf Oesterreicher. 2008. Mündlichkeit und Schriftlichkeit von Texten. In Nina Janich (ed.), *Textlinguistik. 15 Einführungen* (Narr-Studienbücher), 199–215. Tübingen: Narr.
- Koch, Sabine. 1972. Semantische Relationen in sprachlichen Texten (Topikrelationen). *Automatische Sprachübersetzung II*. Berlin (DDR). 129–158.
- Krieg, Ulrike. 2005. *Wortbildungsstrategien in der Werbung. Zur Funktion und Struktur von Wortneubildungen in Printanzeigen* (Beiträge zur germanistischen Sprachwissenschaft 18). Hamburg: Buske.
- Lohde, Michael. 2006. *Wortbildung des modernen Deutschen: Ein Lehr- und Übungsbuch* (Narr-Studienbücher). Tübingen: Narr.
- Matussek, Magdalena. 1994. *Wortneubildung im Text* (Beiträge zur germanistischen Sprachwissenschaft 7). Hamburg: Buske.
- Meliss, Meike. 2008. Wortbildungsprozesse in der Anzeigenwerbung für technische Produkte im Vergleich: Deutsch-Spanisch. In Ludwig M. Eichinger, Meike Meliss, Vázquez Domínguez & José María (eds.), *Wortbildung heute. Tendenzen und Kontraste in der deutschen Gegenwartssprache* (Studien zur deutschen Sprache 44), 231–256. Tübingen: Narr.
- Michel, Georg. 1997. Okkasionalismen und Textstruktur. In Irmhild Barz & Marianne Schröder (eds.), *Nominationsforschung im Deutschen. Festschrift für Wolfgang Fleischer zum 75. Geburtstag*, 337–344. Frankfurt am Main: Peter Lang.

- Motsch, Wolfgang. 2004. *Deutsche Wortbildung in Grundzügen*. 2., überarbeitete Auflage (Germanistische Arbeitshefte 4). Berlin: de Gruyter.
- Müller, Peter O. & Cornelia Friedrich. 2011. Kontamination. In Hilke Elsen & Sascha Michel (eds.), *Wortbildung im Deutschen zwischen Sprachsystem und Sprachgebrauch. Perspektiven – Analysen – Anwendungen* (Perspektiven germanistischer Linguistik 5), 73–107. Stuttgart: Ibidem.
- Naumann, Bernd. 2000. *Einführung in die Wortbildungslehre des Deutschen*. 3., neubearbeitete Auflage. Tübingen: Niemeyer.
- Niehr, Thomas. 1996. „Der Spiegel“ und die Asylanten. Über eine angebliche Wende in der Wortwahl eines deutschen Nachrichtenmagazins. *Sprache und Literatur in Wissenschaft und Unterricht* 78. 84–92.
- Niehr, Thomas. 2014. *Einführung in die linguistische Diskursanalyse*. Darmstadt: WBG.
- Ortner, Lorelies. 1985. Wortbildungs- und Satzbildungsmittel zum Ausdruck von Metaphern und Vergleichen in Science-Fiction-Texten. Oder: Von *wurstförmigen Raumkrümmern* und *Wesen wie Ameisenigel*. In Erwin Koller & Hans Moser (eds.), *Studien zur deutschen Grammatik. Festschrift für Johannes Erben zum 60. Geburtstag* (Innsbrucker Beiträge zur Kulturwissenschaft: Germanistische Reihe 25), 255–275. Innsbruck.
- Pankow, Marion. 1993. Einige methodische Überlegungen zur Erforschung von Wortbildungsprodukten in gesprochener Sprache. In Günther Richter (ed.), *Methodische Grundfragen der Erforschung gesprochener Sprache* (Arbeiten zur Sprachanalyse 16), 86–90. Frankfurt am Main: Peter Lang.
- Peschel, Corinna. 1998. Von Milliardenjungfern, Luthertöttern und Sperminatoren. Zu einer text(sorten)-spezifischen Interpretation von Wortneubildungen. *Zielsprache Deutsch* 29. 121–128.
- Peschel, Corinna. 2002. *Zum Zusammenhang von Wortneubildung und Textkonstitution* (Reihe Germanistische Linguistik 237). Tübingen: Niemeyer.
- Poethe, Hannelore. 2000. Fachsprachliche Aspekte der Wortbildung. Die Leistung der Wortbildung für Fachsprache und Fachtext. In Irmhild Barz, Marianne Schröder & Ulla Fix (eds.), *Praxis- und Integrationsfelder der Wortbildungsforschung* (Sprache – Literatur und Geschichte 18), 199–218. Heidelberg: Winter.
- Polenz, Peter von. 1980. *Wortbildung*. In Hans Peter Althaus, Helmut Henne & Herbert Ernst Wiegand (eds.), *Lexikon der germanistischen Linguistik*. 2., vollständig neu bearbeitete und erweiterte Auflage, 169–180. Tübingen: Niemeyer.
- Püschel, Ulrich. 1997. „Puzzle-Texte“ – Bemerkungen zum Textbegriff. In Gerd Antos & Heike Tietz (eds.), *Die Zukunft der Textlinguistik. Traditionen, Transformationen, Trends* (Reihe germanistische Linguistik 188), 27–41. Tübingen: Niemeyer.
- Römer, David. 2017. *Wirtschaftskrisen. Eine linguistische Diskursgeschichte* (Sprache und Wissen 26). Berlin & Boston: de Gruyter.
- Schlienz, Michael. 2004. *Wortbildung und Text. Eine Untersuchung textverknüpfender Wortbildungselemente* (Erlanger Studien 130). Erlangen: Palm & Enke.
- Schmidt-Brücken, Daniel. 2015. *Verallgemeinerung im Diskurs. Generische Wissensindizierung in kolonialem Sprachgebrauch* (Diskursmuster / Discourse Patterns 9). Berlin, München & Boston: de Gruyter.

- Schröder, Marianne. 1978. Über textverflechtende Wortbildungselemente. *Deutsch als Fremdsprache* 15. 85–92.
- Schröder, Marianne. 1983. Zum Anteil von Wortbildungskonstruktionen an der Konstitution von Texten. *Beiträge zur Erforschung der deutschen Sprache* 3. 108–118.
- Schröder, Marianne. 1985. Überlegungen zur textorientierten Wortbildungsforschung. In Wolfgang Fleischer (ed.), *Textbezogene Nominationsforschung. Studien zur deutschen Gegenwartssprache* (LS/ZISW. A 123), 69–94. Berlin: Akademie d. Wissenschaften der DDR.
- Schröder, Marianne. 1992. Wortbildung in Familiengesprächen. In Rudolf Grosse, Gotthard Lerchner & Marianne Schröder (eds.), *Beiträge zur Phraseologie – Wortbildung – Lexikologie. Festschrift für Wolfgang Fleischer zum 70. Geburtstag*, 93–99. Frankfurt am Main: Peter Lang.
- Schröder, Marianne. 2000. Wortbildung in Textkomplexen. In: Irmhild Barz, Ulla Fix, Marianne Schröder & Georg Schuppener (eds.), *Sprachgeschichte als Textsortengeschichte. Festschrift zum 65. Geburtstag von Gotthard Lerchner*, 385–403. Frankfurt am Main: Peter Lang.
- Schröder, Marianne. 2002. Brauchen wir in der Wortbildungslehre einen textlinguistischen Beschreibungsansatz, der sich an neuen Textphänomenen orientiert? In Ulla Fix (ed.), *Brauchen wir einen neuen Textbegriff?* (Forum angewandte Linguistik 40), 107–111. Frankfurt am Main: Peter Lang.
- Schröder, Marianne. 2005. Wortbildung in der Hörfunknachricht. In: Ulla Fix, Gotthard Lerchner, Marianne Schröder & Hans Wellmann (eds.), *Zwischen Lexikon und Text. Lexikalische, stilistische und textlinguistische Aspekte* (Abhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Philologisch-Historische Klasse 78 (4)), 247–262. Stuttgart & Leipzig: Verlag der Sächsischen Akademie der Wissenschaften.
- Schwitalla, Johannes. 2012. *Gesprochenes Deutsch. Eine Einführung*. 4., neu bearbeitete und erweiterte Auflage (Grundlagen der Germanistik 33). Berlin: Erich Schmidt.
- Selting, Margret et al. 2009. Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10. 353–402.
- Siebold, Oliver. 2000. *Wort – Genre – Text. Wortbildung in der Science Fiction*. Tübingen: Narr.
- Spitzmüller, Jürgen & Ingo H. Warnke. 2011. *Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse*. Berlin & Boston: de Gruyter.
- Stein, Stephan. 2004. Texte, Textsorten und Textvernetzung. Über den Nutzen der Textlinguistik (nicht nur) für die Fremdsprachendidaktik. In Heinz-Helmut Lüger & Rainer Rothenhäusler (eds.), *Linguistik für die Fremdsprache Deutsch* (Beiträge zur Fremdsprachenvermittlung: Sonderheft 7), 171–222. Landau: Verlag Empirische Pädagogik.
- Stein, Stephan. 2007. Wortbildungsfehler? Zur Akzeptabilität von Wortbildungsprodukten aus wortbildungstheoretischer und kommunikationspraktischer Perspektive. *Wirkendes Wort* 57. 459–485.
- Stein, Stephan. 2016. Wortbildung und Wortbildungssemantik. In Inge Pohl & Winfried Ulrich (eds.), *Wortschatzarbeit* (Deutschunterricht in Theorie und Praxis 7), 225–255. Baltmannsweiler: Schneider-Verlag Hohengehren.

- Stumpf, Sören. In Vorbereitung. *Wortbildung in Mündlichkeit und Schriftlichkeit. Beziehungen zwischen Wortbildung und Textbildung im geschriebenen und gesprochenen Deutsch*. Universität Trier Habilitationsschrift.
- Stumpf, Sören & Viola Kämmer. 2017. Battle-Rap – Holistische Textsortenanalyse und soziolinguistische Verortung. *Muttersprache* 127, 173–196.
- Tereick, Jana. 2016. *Klimawandel im Diskurs. Multimodale Diskursanalyse crossmedialer Korpora* (Diskursmuster 13). Berlin: de Gruyter.
- Weber, Tilo. 2014. Funktion und Bedeutung von Wortneubildungen in telefonischen Beratungsgesprächen. In Sascha Michel & József Tódt (eds.), *Wortbildungssemantik zwischen Langue und Parole. Semantische Produktions- und Verarbeitungsprozesse komplexer Wörter* (Perspektiven germanistischer Linguistik 10), 205–226. Stuttgart: Ibidem.
- Wildgen, Wolfgang. 1982. Makroprozesse bei der Verwendung nominaler Ad-hoc-Komposita im Deutschen. *Deutsche Sprache* 10. 237–257.
- Winter-Froemel, Esme. 2009. Wortspiel. In Gert Ueding (ed.), *Historisches Wörterbuch der Rhetorik*. Band 6, 1429–1443. Tübingen: Niemeyer.
- Winter-Froemel, Esme. 2013. Ambiguität im Sprachgebrauch und im Sprachwandel: Parameter der Analyse diskurs- und systembezogener Fakten. *Zeitschrift für französische Sprache und Literatur* 123. 130–170.
- Wolbring, Fabian. 2015. *Die Poetik des deutschsprachigen Rap* (Westwärts Studien zur Popkultur 2). Göttingen: V&R Unipress.
- Wolf, Norbert Richard. 1996. Wortbildung und Text. *Sprachwissenschaft* 21. 241–261.
- Zhang, Dingxian. 1990. *Komplexe lexikalische Einheiten in Fachsprachen. Eine Untersuchung am Beispiel der Fachsprachen der Umformtechnik und der Fertigungstechnik* (Sammlung Gross 42). Heidelberg: Gross.

Anhang

(1) Kommentar, FAZ³⁰

Beitragssätze

Sozial kurzsichtig

Die Sozialabgaben steigen weiter, die Arbeitslosenkasse häuft Reserven an. Warum gibt sie das Geld nicht an die zurück, die es erarbeitet haben? (02.01.2017, von Heike Göbel)

Was der Finanzminister zum neuen Jahr gibt, nimmt der Gesundheitsminister. Die leichte Steuersenkung wird zumindest teilweise aufgezehrt durch den höheren Beitragssatz zur Pflegeversicherung und den höheren Zusatzbeitrag, den jetzt so manche Krankenkasse erbittet. Ein großes Thema sind die steigenden Sozialabgaben bislang nicht, auch weil die große Koalition ihr Vorgehen geschickt tarnt. Minister Gröhe hat den Beitragssprung um ein halbes Prozent, den die Ausweitung der Pflegeleistungen um ein sattes Viertel mit sich bringt, in zwei Schritte aufgeteilt.

30 Die Wortbildungen und Simplicia, die Teil einer (Wortbildungs-)Isotopiekette sind, werden durch Kursivierung und Unterstreichung hervorgehoben.

Ins Wahljahr fällt mit 0,2 Prozentpunkten nur noch der kleinere. Und den kräftigeren Anstieg der Zusatzbeiträge haben Union und SPD verhindert, indem sie den Kassen kurzerhand aus der Reserve des Gesundheitsfonds 1,5 Milliarden Euro zuschusterten. Dass der Beitragssatz der Renten- und Arbeitslosenversicherung stabil bleibt, ist kein Trost.

Die Arbeitslosenkasse häuft Reserven an, allein 2016 hat sie den Beschäftigten und ihren Unternehmen 4,9 Milliarden Euro zu viel abgeknöpft. Statt das Geld an die zurückzugeben, die es erarbeitet haben, denkt die Politik über neue Leistungen nach. Auch in der Rentenkasse läge der Beitragssatz niedriger, hätte die Koalition die dank guter Konjunktur unerwartet erzielten Zusatzeinnahmen nicht in Mütterrente und Rente mit 63 gesteckt.

Die Wettbewerbsfähigkeit der Unternehmen interessiert kaum

Für ihre Politik, die kollektiven Sozialleistungen mit Hilfe der Beitragszahler beherzt auszuweiten, bekommt die Regierung in Umfragen viel Lob. Beifall finden auch jene, die immer noch neue Lücken im Sozialsystem ausmachen. Ein Wahlkampf um höhere Renten zeichnet sich ab. Hingegen hat der Ökonomenrat beim Wirtschaftsministerium nur ein Achselzucken kassiert für seine eindringliche Warnung, dass die Sozialbeiträge nach 2030 von heute rund 40 Prozent auf mehr als 50 Prozent des Bruttolohns schießen werden. Da auch die steuerfinanzierten Sozialausgaben zulegen, bliebe dann von dem, was Bürger und Unternehmen verdienen, nicht mehr viel zur privaten Verfügung, ob für Konsum oder Investition. Was das für die Motivation der Arbeitnehmer und die Dynamik und Wettbewerbsfähigkeit der Unternehmen bedeutet, interessiert kaum.

Politische Weitsicht und frühes Gegensteuern verlangt eine Mehrheit der Bürger bloß im Klimaschutz, nicht zum Schutz der Sozialen Marktwirtschaft. Das Bewusstsein, dass es finanzieller Freiheit bedarf, um im Wettbewerb erfolgreich etwas zu wagen, fehlt zunehmend. Doch sozial bleibt die Marktwirtschaft nur, wenn die Sozialpolitik die Ansprüche zu zügeln lernt.³¹

(2) Interview, Storl³²

001 CL: david ich würde eigentlich gerne äh gratuLIERen zu
ner
silbermedaille
002 aber sie sehen gar nicht aus als würden sie sich
darüber
FREUen
003 kann das SEIN
004 (-)
005 DS: ja=ich denke mal das kommt jetzt NOCH
006 (-) (ich) war jetzt überHAUPT nicht mein wettkampf
007 bin überHAUPT nicht **reingekommen**

31 <http://www.faz.net/aktuell/wirtschaft/wirtschaftspolitik/beitragssaetze-sozial-kurzfristig-14601917.html> (Stand 18.01.2017).

32 Die nach GAT 2 angefertigte Transkription (siehe SELTING et al. 2009) wurde mir freundlicherweise von Anne Klee zur Verfügung gestellt, wofür ich mich an dieser Stelle recht herzlich bedanken möchte. Die Wortbildungen, die einer Isotopiekette angehören, sind (fett) hervorgehoben.

008 hat mich selber n bisschen geÄrgert
009 (-) dass es nich ganz so (.) so flüssig lief wie im
TRAINing
010 die lockerheit hat n bisschen gefEHLT
011 und (.) ja war ein verKRAMPFter wettkampf
012 und äh ja bin ich noch FROH
013 dass es noch zu platz ZWEI gereicht hat
014 muss man so SAgen
015 CL: sie haben VORher selber gesagt
016 für GOLD muss man in peking wahrscheinlich
[zweiundzwanzig] **meter stoßen**
017 DS: [((nickt))]
018 CL: genau DIE hätten gereicht
019 dann wären es nämlich SIEben **zentimeter** mehr gewesen
020 ähm (.) WARum geht das nich
021 können sie uns erKLÄren
022 wie wie wie was das für ein gefÜHL is
023 dass man nich in den WETtkampf **reinkommt**
024 (.) weiß man das schon VORher
025 DS: nein VORher weißtes nich
026 also DRAUßen wars noch ganz gut
027 draußen macht man immer son paar **STANDstöße**
028 um son bisschen ne **vorspannung** AUFzubauen
029 und dann (.) WEISS ich nicht
030 hab ich irgendwie aufm weg ins stadion son bisschen
meine Linie verlorn
031 hab (im) **technisch** UNSauber **gearbeitet**
032 hat man in ganz ungewöhnlichen **STÖßen** und im zwanzig
vierundvierzig
033 (.) hat man geSEHen
034 das is ja wenn_n bisschen aufregung daZU kommt
035 dann **stürzt** man die **stöße REIN**
036 und da fehlt dann so der der DRUCK ausm bein
037 und ja wenn das dann FEHLT dann
038 (.) isses SCHWIERig dann irgendwie wieder ne linie
zu
finden
039 CL: is ja beMERkenswert
040 dass ein zweimaliger weltmeister tatsächlich dann
auch
noch nervös werden [KANN]
041 DS: [((schmunzelt))]
042 CL: zwischendurch haben sie sich immer wieder die WAd
gehalten
043 waren beim PHYSiotherapeuten
044 beim **TRAIner**
045 hat das n EINfluss gehabt
046 weil äh_äh körperlich defizite spielen ja SCHON im-
mer ne
rolle

047 auch in sachen KNIE bei ihnen
 048 (-)
 049 DS: ja=is halt IMmer so ne sache=ne
 050 das is (.) man geht ans LIimit bei jedem wettkampf
 051 man baut so viel **spannung** AUF
 052 das is ne (.) JA
 053 (.) dass es entweder GUT geht
 054 oder NICH gut geht
 055 jetzt hat ich hier n bisschen proBLEme in der wade
 056 das is aber (.) haben wir alles in_n GRIFF bekommen
 057 und hat WIRKlich hinten raus daran gelegen
 058 dass die **TECHnik** n bisschen gehadert hat
 059 CL: sie haben VORher gesagt
 060 ich bin nur an Titeln äh interessiert
 061 reKORde kommen vielleicht nebenbei
 062 ähm sind sie da fahren sie jetzt richtig SAUER nach
 hause
 063 weil KEIN titel dabei rausgekommen is
 064 DS: ja KLAR
 065 ich wollt meinen titel verTEIdigen
 066 und es war auch jetzt keine LEISTung
 067 die ich nicht hätte **STOßen** können
 068 sone zweiundzwanzig hab ich mir eigentlich ZUgetraut
 hier
 069 war auch gestern im **training** nochmal ja (-) SEHR gut
 070 und hab viel SELBSTbewusstsein getankt
 071 aber das war halt dann SCHWIERig hier umzusetzen
 072 und ja wie gesagt (.) hab n bisschen die LInie
 verloren=ne
 073 CL: wie gehen sie jetzt damit UM
 074 was MACHen sie in den nächsten tagen
 075 DS: joa ich bin jetzt MORgen noch in peking
 076 hab morgen noch die SIEgerehrung
 077 und dann gehts am DIENStag nach hause
 078 und und erstmal alles so n bisschen **verARbeiten**
 079 [bisschen] **trainIERen** wieder
 080 CL: [(nickt))]
 081 DS: bisschen struktUR **reinbringen**
 082 und dann gehts zu zu den nächsten WETTkämpfen
 083 CL: alles GUTE dabei danke
 084 DS: DANke

Dr. Sören Stumpf
 Universität Trier
 FB II – Germanistik – GermanistischeLinguistik
 Universitätsring 15
 D-54296 Trier
 E-Mail: stumpf@uni-trier.de

Joshgun Sirajzade

Korpusbasierte Untersuchung der Wortbildungsaffixe im Luxemburgischen. Technische Herausforderungen und linguistische Analyse am Beispiel der Produktivität

Abstract: The project WBLUX (Wortbildung des moselfränkisch-luxemburgischen Raumes) at the University of Luxembourg aims at the investigation of Luxembourgish word formation through different text sorts and genres. In order to achieve this goal the compilation of an annotated corpus is needed. This article gives an example for benefits of using a corpus with annotations like parts of speech, lemmata and word formation affixes in the analysis of productivity of some selected word formation affixes of Luxembourgish. Then it describes how one can achieve such a corpus from a technical point of view. This includes the choice of corpus format, of a database platform and the designing of programs needed for the annotation process of word formation itself. This article also suggests new corpus linguistic approaches for research of word formation like analyzing the usage of word formation bases in the entire corpus or performing context analysis in order to determine semantical functions of each suffix.

1. Einleitung

Korpuslinguistische Methoden sind nicht nur ein fester Bestandteil der heutigen linguistischen Forschung, sondern auch eine notwendige Voraussetzung für die zuverlässige, empirische und nachhaltige Untersuchung einer Sprache. Somit sind sie auch nicht aus der Wortbildungsforschung des Luxemburgischen wegzudenken. Sie erlauben sowohl eine strukturierte Organisation der Daten dank Technologien wie Datenbanken, die zum Auffinden und Speichern von gesuchten Sprachphänomenen sehr hilfreich sein können, als auch die schnellere Durchführung von formalen und stochastischen Analysen. Dabei kann man außer der korpuslinguistischen Theorie selbst auch auf die Erkenntnisse der Informatik sowie Linguistik und insbesondere der einzelnen benachbarten interdisziplinären Fachrichtungen wie Computerlinguistik, Quantitative Linguistik sowie Digital Humanities zurückgreifen. Die Erforschung der luxemburgischen Sprache, die oft als jüngste germanische Sprache bezeichnet wird, hat inzwischen eine Tradition. Eine Übersicht findet sich bei Moulin (2004). Dennoch sind grammatische Beschreibungen und Analysen des Luxemburgischen im Vergleich zu Untersuchungen der benachbarten Sprachen immer noch rar (Gilles 2006). Besonders fällt dies im Bereich der korpuslinguistischen Analysen auf (Sirajzade 2013: 5). Obwohl

bereits elektronische Ressourcen und Datenbanken für die Untersuchung des Luxemburgischen existieren (Sirajzade 2013: 5), gibt es hier einen Bedarf nach einem nationalen Sprachkorpus, das etwa vergleichbar wäre mit dem British National Korpus¹ oder dem Deutsche Referenzkorpus², um einige von vielen Beispielen weltweit zu nennen. Solch ein Korpus für das Luxemburgische würde eine repräsentative Untersuchung ermöglichen und die Ergebnisse der linguistischen Analysen vergleichbar machen. Das Projekt „Die Wortbildung des moselfränkisch-luxemburgischen Raumes (WBLUX)“ am Institut für luxemburgische Sprache und Literatur der Universität Luxemburg hat sich zum Ziel gesetzt, die luxemburgische Wortbildung flächendeckend über verschiedene Stile und Textsorten zu untersuchen. Zu diesem Zweck mussten technische Hürden überwunden werden, wie die Aufbereitung des Korpus aus bestehenden elektronischen Texten sowie die Suche und Annotation der Wortbildungsaffixe. Dieser Umstand eröffnete zur gleichen Zeit eine Möglichkeit, die aktuellsten Standards und Technologien in der Korpuslinguistik zu untersuchen und einzusetzen. Anschließend mussten neuste Erkenntnisse und Methoden aus der Wortbildungsforschung elaboriert und angewandt werden. Dieser Aufsatz fungiert als ein Bericht, wie auf der einen Seite bereits existierende elektronische Texte des Luxemburgischen zu einem annotierten Korpus zur Untersuchung der Wortbildung aufbereitet werden können. Auf der anderen Seite werden statistische Möglichkeiten vorgestellt und diskutiert, mit deren Hilfe die Wortbildung in einem Korpus untersucht werden kann – konkret eine formale Analyse der Produktivität der ausgewählten Wortbildungssuffixe des Adjektivs und Substantivs im Luxemburgischen. Es soll gezeigt werden, wie sich die Entwicklungstendenzen der Wortbildungssuffixe im Luxemburgischen darstellen. Dabei werden Kategorien wie die realisierte und expandierende Produktivität diskutiert und konkrete Beispiele für die Konkurrenz der einzelnen analysierten Suffixe beschrieben. Diese Untersuchungen fanden im Rahmen des WBLUX Projekts an der Universität Luxemburg statt.

1.1. Forschungsstand

Die bisherige Forschungsarbeit, die für die Anwendung von korpuslinguistischen Methoden in der Wortbildungsforschung relevant sein kann, lässt sich in zwei Gruppen unterteilen: Die Erkenntnisse hinsichtlich der allgemeinen Algorithmen und Werkzeuge in der Linguistik einerseits, und andererseits die Erkenntnisse der Wortbildungsforschung selbst. Eine gute Übersicht darüber findet sich im HSK-Band ‚Word-formation‘ (Müller 2015: 2333–2372), in dem die Werkzeuge und Algorithmen aufgrund des zugrunde liegenden

1 <http://www.natcorp.ox.ac.uk/>, zuletzt geprüft am 29.11.2017.

2 <http://www1.ids-mannheim.de/kl/projekte/korpora/>, zuletzt geprüft am 29.11.2017.

empirischen Materials in drei einzelnen Aufsätzen behandelt werden: 1. Dictionaries (Belentschikow 2015); 2. Corpora (Heid 2015) und 3. Internet (Dal und Namer 2015). Den Nutzen der Wörterbücher sehen die Autoren in den Informationen zur Wortbildung, die diese zur Verfügung stellen, zum einen in den einzelnen Wortartikeln der einsprachigen Wörterbücher. Zum anderen existieren spezialisierte Wörterbücher, insbesondere die sogenannten „morpheme and word-formation dictionaries“, die bereits eine Liste von Affixen mit ausführlichen Informationen zu Wortbildung enthalten (Belentschikow 2015: 2345). Hinsichtlich der Wortbildung des Luxemburgischen findet sich leider kein solches Wörterbuch. Eine besondere Rolle spielt im HSK-Band der Aufsatz ‚Corpora‘ (Heid 2015). Hier werden nicht nur die wichtigsten Eigenschaften eines Korpus wie Repräsentativität, Zusammenstellung und Größe diskutiert, sondern es werden auch wichtige Strategien in der Wortbildungsforschung wie die Verwendung der vorhandenen Annotationen (Part-of-speech (POS), Lemmata etc.) bei der Suche nach Wortbildungselementen, deren automatisierte Zerlegung, die Festlegung der Produktivität durch Frequenzanalysen und die Untersuchung der Funktionsklassen durch Kontexte aufgezeigt (Heid 2015: 2354–2371). Einiges an Arbeit wurde auch bereits geleistet, um Werkzeuge zu entwickeln, die eine automatisierte morphologische Zerlegung und damit auch die Zerlegung der Wortbildungselemente gewährleisten können. Was die in der Korpuslinguistik etablierten Werkzeuge angeht, existiert bereits ein POS-Tagger für das Luxemburgische (Sirajzade 2012: 264–280). Darüber hinaus existieren andere korpuslinguistische Werkzeuge, die zwar nicht explizit für das Luxemburgische implementiert sind, jedoch stochastische Mittel verwenden und im Prinzip sprachunabhängig sind. Eines davon ist Morfessor³, hauptsächlich entwickelt für das Finnische, das eine relativ komplexe morphologische Struktur aufweist (Creutz und Lagus 2002; Kohonen et al. 2010). Eine andere Herangehensweise ist die semiautomatische morphologische Zerlegung, die meist auf einem Sprachmodell basiert. Getestet am Material des Arabischen ergab diese Herangehensweise einen Korrektheitsgrad von 97% (Lee et al. 2003: 399–406). Einen deutlich klassischeren Weg geht das Projekt MorphoDiTa⁴, das hauptsächlich aus einem POS-Tagger und einem morphologischen Wörterbuch besteht. Die beiden Teile des Programms können trainiert werden (Straková et al. 2014). Es gibt bereits für das Tschechische und Englische trainierte Modelle, die man verwenden kann. Jedoch betreffen alle diese Arbeiten die gesamte Morphologie und beschäftigen sich nicht zielgerichtet mit der Wortbildung. Bisher wurde viel Forschung hinsichtlich der morphologischen Produktivität und ihrer quantitativen Modellierung betrieben (Baayen 2009; Pustynnikov und

3 mittlerweile in der Version 2.0. <http://morpho.aalto.fi/projects/morpho/>, zuletzt geprüft am 29.11.2017

4 <http://ufal.mff.cuni.cz/morphodita>, zuletzt geprüft am 29.11.2017

Schneider-Wiejowski 2010; Schneider-Wiejowski 2011). Eine Beschreibung der Untersuchungsgeschichte der luxemburgischen Wortbildung findet sich bei Sirajzade (2013: 57–60). Außer den hier ausgeführten klassischen Grammatiken wäre die Arbeit von Lulling (2002) zu nennen, der die Kreativität der luxemburgischen Wortbildung anhand eines selbst kompilierten Korpus untersucht.

1.2. Das Korpus

Die Korpusgrundlage für die Untersuchung der luxemburgischen Wortbildung besteht aus zwei Teilkorpora: 1. Ein ausgewogenes, repräsentatives Kernkorpus, bestehend aus ca. einer Mio. Token. 2. Ein Erweiterungskorpus für weitere, meist stochastische Analysen, bestehend aus ca. 19 Mio. Token. Insgesamt bestehen beide Teilkorpora aus knapp unter 20 Mio. Token. Hierzu gehören Parlamentsreden (Chambre des Députés) mit ca. 10 Mio. Token in 293 Dokumenten, RTL News (Radio Télé Luxembourg) Nachrichten mit ca. 5 Mio. Token in 109 Dokumenten, luxemburgische Literatur vertreten mit ca. 2 Mio. Token in 54 Dokumenten sowie Interviews der Uni Luxemburg mit ca. 0,5 Mio. Token in 49 Dokumenten.

2. Statistische Analyse der Produktivität

2.1. Festlegung der Produktivität

Die Produktivität der Wortbildungselemente ist eine der wichtigsten Eigenschaften und zeichnet den Charakter und die Dynamik einer Sprache aus. So thematisiert Heid (2015: 2361) die Produktivität im HSK-Band „Word-formation“ ausführlich und beruft sich bei der mathematischen Beschreibung des Problems hauptsächlich auf die Arbeiten von Baayen. Baayen (2009) versucht, das Phänomen der Produktivität theoretisch und mathematisch zu formulieren, allerdings im Hinblick auf die gesamte Morphologie. In einfachster Form wird die Produktivität einer morphologischen Kategorie C geschätzt mit der Typezahl $V(C, N)$, wobei N die Gesamtanzahl der Tokens in einem Korpus ist (Baayen 2009: 902). Baayen nennt dies die realisierte Produktivität. Durch die Miteinbeziehung der Types einer Kategorie spiegelt diese Maßgröße deren Möglichkeit zur lexikalischen Vielfalt wieder (Sirajzade 2013: 50). Laut Baayen ist die realisierte Produktivität der aktuelle Zustand in einer Sprache, sie repräsentiert jedoch nicht die Tendenzen einer Sprache. Dafür ist die expandierende Produktivität („expanding productivity“) verantwortlich. Sie bezieht Hapaxlegomena einer Kategorie in die Berechnung ein, also die Anzahl der Wörter $V(1, C, N)$ in einer Kategorie C, die in einem Korpus von der Größe N nur einmal auftauchen. Wenn $V(1, N)$ die totale Anzahl der Hapaxlegomena in einem Korpus darstellt, dann schätzt das Verhältnis $P^* = V(1, C, N) / V(1, N)$ den Beitrag einer Kategorie zur Erweiterung des Lexikons einer Sprache. Eine

weitere Formel von Baayen, genannt die potenzielle Produktivität $P^* = V(1, C, N) / N(C)$, setzt die Hapaxlegomena ins Verhältnis zu der Gesamtanzahl der Tokens in einem Korpus. Die potenzielle Produktivität beschreibt das Wachstum des Lexikons einer grammatischen Kategorie (Baayen 2009: 902).

2.2. Produktivität bei der Adjektivbildung

Während *-esch* als adjektivbildendes Suffix 4 273 Mal in 737 Types im Kernkorpus (1 000 000 Tokens) vorkommt, lässt sich *-eg* als adjektivbildendes Suffix in 4 341 Tokens bestehend aus 489 Types finden. Im Verhältnis zur Gesamtanzahl der Tokens im Kernkorpus ergibt sich, dass die realisierte Produktivität von *-eg* $489/1\,000\,000 \approx 0,000489$ beträgt und bei *-esch* $737/1\,000\,000 \approx 0,000737$. Zum Suffix *-eg* gibt es 211 Hapaxlegomena im Kernkorpus, d.h. die expandierende Produktivität für *-eg* beträgt hier $211/16\,372 \approx 0,0129$, während sie bei *-esch* mit 346 Hapaxlegomena $346/16\,372 \approx 0,0211$ beträgt. Laut beiden Messungen ist *-esch* bei der Adjektivbildung im Luxemburgischen demnach produktiver als *-eg*, obwohl die Tokenanzahl bei *-eg* höher ist. Das wird offensichtlicher, je mehr Daten man in die Analyse miteinbezieht. Im Gesamtkorpus gibt es ca. 50 000 Tokens zu *-eg*, zu *-esch* lediglich ca. 45 000. Das bedeutet, dass die mit *-eg* gebildeten Adjektive in der sprachlichen Realisierung häufiger verwendet werden, aber *-esch* viel mehr lexikalische Vielfalt bei den Basen besitzt, mit denen es Verbindungen eingeht.

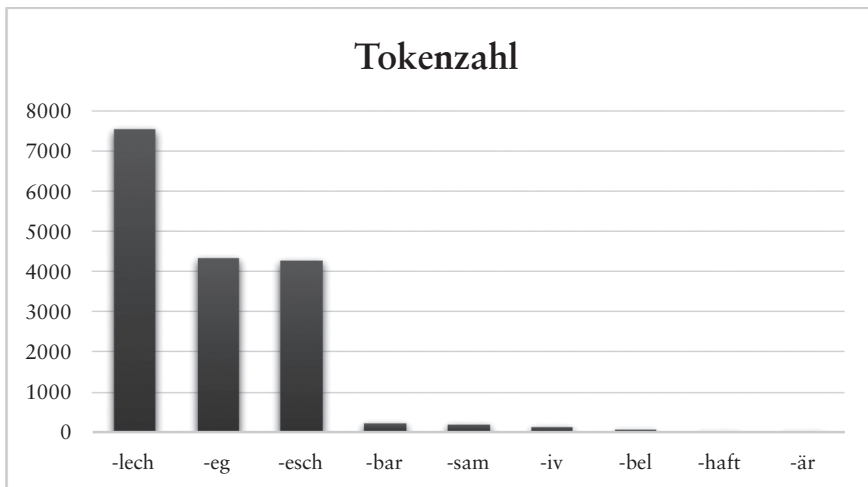
Betrachtet man die Zahl der Types bei dem Suffix *-lech*, das tokenmäßig häufigste adjektivbildende Suffix im Kernkorpus (7 555), so lässt sich feststellen, dass das Suffix *-lech* auch hinter *-esch* steht (556). Die Zahl der Hapaxlegomena ist mit 211 auch relativ niedrig. So kommt man auf eine realisierte Produktivität von $556/1\,000\,000 \approx 0,000556$, wobei die expandierende Produktivität dabei $211/16\,372 \approx 0,0129$ beträgt, ähnlich wie bei dem Suffix *-eg*. In der folgenden Tabelle finden sich die Token-, Type- und Hapaxlegomenon-Frequenzen der häufigsten adjektivbildenden Suffixe.

Tab. 1: Adjektivbildende Suffixe im Luxemburgischen.

Suffix	POS	Tokenzahl	Typezahl	Hapaxlegomena
<i>-lech</i>	ADJ	7555	556	211
<i>-eg</i>	ADJ	4341	489	211
<i>-esch</i>	ADJ	4273	737	346
<i>-bar</i>	ADJ	231	92	51
<i>-sam</i>	ADJ	183	18	10
<i>-iv</i>	ADJ	115	12	0
<i>-bel</i>	ADJ	59	28	5
<i>-haft</i>	ADJ	52	22	11
<i>-är</i>	ADJ	2	1	0

Außer diesen drei produktivsten adjektivbildenden Suffixen, kann man eine andere Gruppe mit geringerer Tokenanzahl unterscheiden. Zu dieser Gruppe gehören *-bar*, *-sam*, *-iv*, *-bel*, *-haft*. Zu *-bar* gibt es im Kernkorpus 232 Tokens und 94 Types, wovon 51 Hapaxlegomena sind. Bei dem Suffix *-sam*, welches im Kernkorpus 183 Mal in 18 Types vorkommt, gibt es nur 10 Hapaxlegomena. Das Suffix *-iv* verhält sich ähnlich; für 115 Tokens gibt es 12 Types, wobei es hier keine Hapaxlegomena gibt. Das Element *-bel* wie bei *honorabel* kommt als Token 59 Mal und als Type 28 Mal vor. Hapaxlegomena dazu gibt es nur in fünf Fällen. Zu *-haft* lassen sich 52 Tokens, 22 Types und 11 Hapaxlegomena finden, was zu folgenden Ergebnissen führt: Die realisierte Produktivität für das Suffix *-bar* beträgt $94/1\ 000\ 000 \approx 0,000\ 094$, für *-sam* $18/1\ 000\ 000 \approx 0,000\ 018$, für *-iv* $12/1\ 000\ 000 \approx 0,000\ 012$, für *-bel* $28/1\ 000\ 000 \approx 0,000\ 028$ und für *-haft* $22/1\ 000\ 000 \approx 0,000\ 022$. Also haben in dieser Gruppe die Suffixe *-bar* und *-bel* die höchsten Werte für die realisierte Produktivität. Die expandierende Produktivität für *-bar* ist mit $92/16\ 372 \approx 0,0056$ jedoch deutlich höher als bei *-bel* mit $5/16\ 372 \approx 0,00031$.

Abb. 1: Tokenzahl der adjektivbildenden Suffixe im Luxemburgischen.



Hapaxlegomena sind eine wichtige Eigenschaft eines Korpus, sie werden neben und im Gegensatz zu der Token-Type-Ratio, auch in der sogenannten Hapaxlegomena-Token-Ratio verwendet (Ali und Hussein 2014). Pustyl'nikov und Schneider-Wiejowski (2010: 13) schlagen in ihrer Untersuchung der Wortbildungssuffixe *-nis*, *-ung*, *-er* und *-heit/-keit* im Deutschen vor, die komplette Distribution der Types eines Wortbildungselements zu berücksichtigen, allerdings erhalten sie bessere Ergebnisse mit dem Verfahren, das Hapaxlegomena berücksichtigt. Die Anzahl der Hapaxlegomena in einem

Korpus ist jedoch von dessen Größe abhängig. Im Gesamtkorpus (ca. 20 Mio. Tokens) ist zu *-esch* 1370 Hapaxlegomena zu finden, zu *-eg* nur 736, was den Unterschied in der Produktivität noch deutlicher macht, da die expandierende Produktivität von *-esch* $1\ 370 / 85\ 424 \approx 0,016$ deutlich höher ist als bei *-eg* $736 / 85\ 424 \approx 0,0086$.

Bei der Analyse der Hapaxlegomena im Falle von *-esch* und *-eg* lässt sich erkennen, dass *-esch* eher Bildungen mit Fremdwörtern romanisch oder gar lateinischen Ursprungs eingeht als *-eg*, z. B. *hygienesch* ‚hygienisch‘, *mikrobiologesch* ‚mikrobiologisch‘, *organisatoresch* ‚organisatorisch‘, *olympesch* ‚olympisch‘, *diabolesch* ‚diabolisch‘, *didaktesch* ‚didaktisch‘, *dynamesch* ‚dynamisch‘, *pädagogesch* ‚pädagogisch‘, *thematesch* ‚thematisch‘, *philosophesch* ‚philosophisch‘, *statistitesch* ‚statistisch‘, *ekonometresch* ‚ökonomisch‘, *logistes* ‚logistisch‘ etc. Es gibt aber auch ursprünglich luxemburgische Wörter, die mit *-esch* gebildet werden können, wie z. B. *ënnerierdesch* ‚unterirdisch‘, *stiermesch* ‚stürmisch‘, *auslännesch* ‚ausländisch‘. Das Suffix kann außerdem nahezu mit allen Länder-, Orts- und Flurnamen kombiniert werden: *dänesch* ‚dänisch‘, *irlännesch* ‚irländisch‘, *australesch* ‚australisch‘, *éisträiches* ‚österreichisch‘, *südeuropäesch* ‚südeuropäisch‘ etc. Das zeigt die Dynamik und das Potenzial des Suffixes. *-eg* hingegen bildet mehrheitlich Wörter von luxemburgischen bzw. germanischen Stämmen: *stéchalteg* ‚stichhaltig‘, *réckfällleg* ‚rückfällig‘, *zweetrangeg* ‚zweitrangig‘, *gemengnützeg* ‚gemeinnützig‘, *zäitwëlleg* ‚zeitweilig‘, *eekeleg* ‚eklig‘, *flësseg* ‚flüssig‘, *gehéiereg* ‚gehörig‘. Das Suffix *-eg* ist nicht offen für Wörter romanischen bzw. lateinischen Ursprungs, ist dennoch eines der produktivsten adjektivbildenden Suffixe im Luxemburgischen. Die Tatsache, dass man *-eg* im Korpus in Wörtern mit griechisch-lateinischen Ursprung findet, wie z. B. *drasteg*, welches im Korpus zwei Mal registriert wurde (RTL_News_National_06.2 und Parlament 2003–2004_09), ist auf einen Rechtschreibfehler zurückzuführen. Die richtige Schreibweise ist *drastesch* ‚drastisch‘.

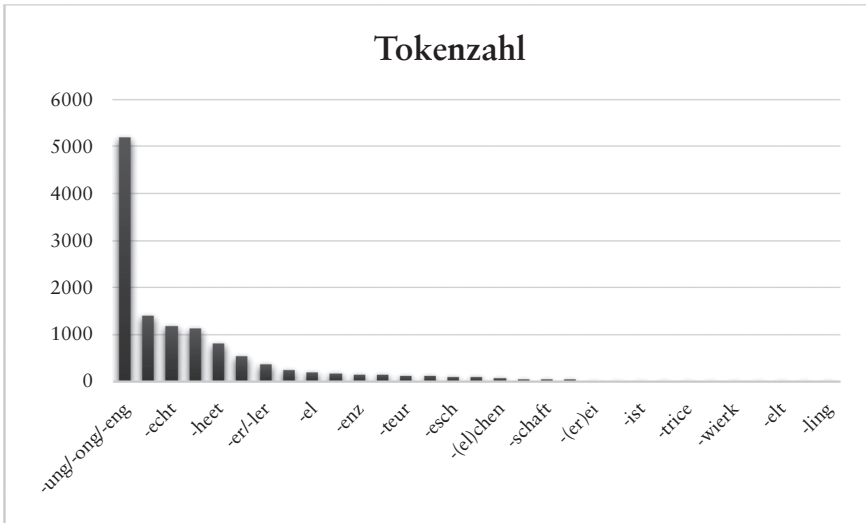
Was die Adjektive angeht, die mit *-bar* gebildet werden, ist deutlich der Einfluss des benachbarten Deutschen zu sehen, obwohl es aufgrund der Tatsache, dass beide Sprachen verwandt sind, schwer ist, diesen Einfluss ganz genau in Zahlen zu fassen. Es gibt im Kernkorpus Wörter wie *absehbar*, *wunderbar*, *furchtbar*, *fruchtbar* die deutlich deutschen Ursprungs sind, jedoch auch andere Wörter wie *bemierkbar* ‚bemerktbar‘, *ëmsetzbar* ‚umsetzbar‘, *virstellbar* ‚vorstellbar‘, *notzbar* ‚nutzbar‘, deren Ursprung strittig sein könnte (gebildet nach dem deutschen Muster) und Wörter, die sehr Luxemburgisch erscheinen wie *cumuléierbar* ‚kumulierbar‘, *uklobar* ‚anklagbar‘, *drobar* ‚tragbar‘. Auf jeden Fall scheint das Suffix im Luxemburgischen in der zweiten Gruppe der Suffixe mit weniger Tokens am produktivsten zu sein, aufgrund der lexikalischen Vielfalt, die man auf den ersten Blick feststellen kann. Zahlen aus dem Korpus und die Berechnungen der realisierten und expandierenden Produktivität bestätigen

dies nochmals. *-bar* konkurriert jedoch mit dem Suffix *-bel*, das laut Produktivitätsmessungen zwar weniger produktiv ist als *-bar*, dennoch kommt es hinter *-bar* und ist produktiver als alle übrigen Suffixe in der zweiten Gruppe. *-bel* ist in Adjektiven romanischen Ursprungs zu finden, wie z. B. *applicabel* ‚anwendbar‘, *recevabel* ‚zulässig‘, *flexibel* ‚flexibel‘, *indispensabel* ‚indispensabel‘, *formidabel* ‚großartig‘ etc. Man könnte meinen, beide Suffixe haben ihre eigene Nische, führt man allerdings eine Basisanalyse durch, die als Verfahren später diskutiert wird, dann stellt man fest, dass es im Korpus Basen gibt, um die diese zwei Suffixe im Luxemburgischen konkurrieren, z. B. *realisabel* vs. *realiséierbar*. Beide Wörter sind im Kernkorpus ein Hapaxlegomenon, „...datt de lëtzebuergeschen Effort muss realisabel sinn...“ (Parlament 2003–2004_26) und „...déis sech fir eng friddlech Koexistenz, mä eng réaliséierbar Koexistenz...“ (Parlament 2003–2004_17), Jedoch findet sich der erste Type 6 Mal und der zweite 18 Mal im Erweiterungskorpus. Somit wird deutlich, dass *realiséierbar* die häufigere Form der beiden Types ist. *-haft* verhält sich wie *-bar*; es gibt hier Adjektive, die definitiv aus dem Deutschen kommen, z. B. *mangelhaft*, *lückenhaft*, *gönnerhaft*. Auf der anderen Seite findet man das Suffix bei den luxemburgischen Basen, wie *feelerhaft* ‚fehlerhaft‘, *meeschterhaft* ‚meisterhaft‘, *eeschthhaft* ‚ernsthaft‘, *niewelhaft* ‚nebelhaft‘. Allerdings ist *-haft* nicht so produktiv wie *-bar* und erreicht auch nicht seine lexikalische Vielfalt an Basen. Dasselbe gilt auch für *-sam* in Lexemen deutschen Ursprungs z. B. *gemeinsam*, *gewaltsam* gegen luxemburgische Lexeme *spuersam* ‚sparsam‘, *opmierksam* ‚aufmerksam‘.

2.3. Produktivität bei der Substantivbildung

Substantive besitzen unter den Wortklassen die meisten Wortbildungselemente. Das hängt damit zusammen, dass Substantive Dinge aus der objektiven Realität benennen und die meisten Types und damit die umfangreichste lexikalische Vielfalt in der Sprache besitzen. Das Luxemburgische ist in diesem Fall keine Ausnahme (Sirajzade 2013: 58). Das Diagramm 2 und die dazu gehörige Tabelle 2 zeigt, dass das Luxemburgische 6 bzw. 7 Suffixe kennt, die sehr häufig Substantive bilden, je nachdem, ob man *-heet* und *-(eg)keet* als ein oder zwei Suffixe betrachtet. Diese sind *-ungl/-ongl/-eng*, *-ioun*, *-echt*, *-(eg)keet*, *-heet*, *-age* und *-er/-ler*. Die zweite Gruppe, bestehend aus den Suffixen *-el*, *-ement*, *-nes/nis*, *-enz*, *-ment*, *-teur*, *-esch*, *-in*, und *-turl/-dur*, bezieht, was die Häufigkeit betrifft, eine mittlere Position. Die Suffixe der dritten Gruppe *-mus*, *-schaft*, *-eur*, *-(el)chen*, *-(er)ei*, *-tuml/-tem*, *-ist*, *-ert* und *-trice* kommen in der luxemburgischen Sprache seltener vor. Die vierte Gruppe besteht aus den sehr seltenen Suffixen *-(e)s*, *-wierk*, *-wiesen*, *-elt*, *-inne* und *-ling*, die im Kernkorpus in einer Anzahl unter 10 Tokens vorkommen.

Abb. 2: Tokenanzahl der substantivbildenden Suffixe im Luxemburgischen.



Das Suffix *-ung* scheint im Luxemburgischen mit Abstand das geläufigste Suffix für die Substantivbildung zu sein. Sowohl die realisierte Produktivität von $1\,068/1\,000\,000 \approx 0,00107$ als auch die expandierende Produktivität von $491/16372 \approx 0,029$ ist am höchsten unter allen substantivbildenden Suffixen. Gefolgt wird dieses Suffix von *-ioun*, das zwar unter den substantivbildenden Suffixen nach *-ung* mit 1 404 und 407 die höchste Anzahl von Tokens und Types hat, aber mit 82 weniger Hapaxlegomena als das Suffix *-(eg)keet* besitzt (noch weniger, wenn man *-(eg)keet* und *-heet* als die Formen eines Suffixes sehen würde). Daher beträgt die realisierte Produktivität von *-ioun* $407/1\,000\,000 \approx 0,000407$ und die expandierende Produktivität $82/16372 \approx 0,005$. Somit hat *-ung*, statistisch gesehen keine ernstesten Konkurrenten unter den substantivbildenden Suffixen in der luxemburgischen Sprache. Das Suffix *-ioun* konkurriert vielmehr mit dem germanischen Suffix *-keet/-heet*. Alleine *-keet* besitzt eine expandierende Produktivität von $123/16372 \approx 0,0075$, also eine höhere als *-ioun*. Indes ist die realisierte Produktivität bei *-ioun* höher als bei *-keet*. Dieser Unterschied fällt natürlich kleiner aus, wenn man *-keet* und *-heet* als ein Suffix betrachtet, woraus sich eine Typeanzahl von 311 ergäbe. Trotzdem erreicht diese Zahl nicht die Anzahl der Types bei *-ioun* mit 407. Laut der Anzahl der Tokens befindet sich zwischen diesen beiden das Suffix *-echt*, allerdings mit deutlich weniger Types und Hapaxlegomena. Interessant ist bei diesem typisch luxemburgischen Suffix, dass es eine erstaunlich hohe Anzahl an Tokens hat. Diese Gruppe von Suffixen endet mit *-age* und *-er* mit einer realisierten Produktivität von jeweils $64/1\,000\,000 \approx 0,00\,0064$

und $110/1\ 000\ 000 \approx 0,00\ 011$ und einer expandierenden Produktivität von jeweils $24/16\ 372 \approx 0,00146$ und $22/16\ 372 \approx 0,00124$.

Tab. 2: *Substantivbildende Suffixe im Luxemburgischen.*

<i>Suffix</i>	<i>POS</i>	<i>Tokenzahl</i>	<i>Typezahl</i>	<i>Hapaxlegomena</i>
<i>-ung/-ongl/-eng</i>	N	5187	1068	491
<i>-ioun</i>	N	1404	407	82
<i>-echt</i>	N	1175	113	44
<i>-(eg)keet</i>	N	1137	222	123
<i>-heet</i>	N	817	89	40
<i>-age</i>	N	541	64	24
<i>-er/-ler</i>	N	354	110	22
<i>-nes/-nis</i>	N	236	73	38
<i>-el</i>	N	193	56	16
<i>-ement</i>	N	172	52	9
<i>-enz</i>	N	145	44	5
<i>-ment</i>	N	134	13	1
<i>-teur</i>	N	124	31	4
<i>-tur/-dur</i>	N	124	46	11
<i>-esch</i>	N	105	15	6
<i>-in</i>	N	100	36	22
<i>-(el)chen</i>	N	61	32	20
<i>-mus</i>	N	40	13	4
<i>-schaft</i>	N	38	11	4
<i>-eur</i>	N	37	16	4
<i>-(er)ei</i>	N	25	11	2
<i>-tum/-tem</i>	N	18	5	1
<i>-ist</i>	N	17	12	5
<i>-ert</i>	N	13	7	2
<i>-trice</i>	N	12	3	1
<i>-(e)s</i>	N	7	4	1
<i>-wierk</i>	N	4	2	0
<i>-wiesen</i>	N	4	3	0
<i>-elt</i>	N	1	1	0
<i>-inne</i>	N	1	1	1
<i>-ling</i>	N	1	1	0

Schaut man sich die lexikalische Vielfalt an, die die erste Gruppe der substantivbildenden Suffixe des Luxemburgischen haben, so sieht man, dass *-ung/-ongl/-eng* im Kernkorpus von nahezu über 1 000 Verbtupes Substantive

bilden kann. Größtenteils sind es luxemburgische Substantive wie *Fuerschung* ‚Forschung‘, *Belaaschtung* ‚Belastung‘, *Finanzéierung* ‚Finanzierung‘, *Ofmaachung* ‚Abmachung‘, *Uweisung* ‚Anweisung‘, *Eenegung* ‚Einigung‘, *Kloerstellung* ‚Klarstellung‘, *Opschwong* ‚Aufschwung‘ etc. Es finden sich aber auch Lehnwörter aus dem Deutschen, wie *Bevölkerung*, *Bekämpfung* oder *Handlung*. Es scheint, dass die deutschen Lehnwörter die Produktivität des Suffixes statistisch erhöhen, denn viele von diesen Lehnwörtern können im Luxemburgischen durch andere Wörter ersetzt werden, z. B. mit Lehnwörtern aus dem Französischen. Trotzdem ändert diese Tatsache nichts an der ersten Position des Suffixes bei der Produktivität. Mit einer lexikalischen Vielfalt und Flexibilität zeichnet sich das Suffix *-ioun* aus. Dieses Suffix ist eher bei den Lexemen aus dem Lateinischen zu finden und konkurriert sowohl mit *-ung*, als auch am Rande mit *-heet/-keet*, z. B. *Realisatioun* ‚Realisation‘ vs. *Realiséierung* ‚Realisierung‘, dabei ist die erste Form häufiger im Kernkorpus zu finden als die zweite, 29 vs. 5 Tokens, *Globalisatioun* ‚Globalisierung‘ vs. *Globaliséierung* ‚Globalisierung‘, diesmal mit der häufigeren Verwendung der zweiten Form, 3 vs. 9 Tokens, und *Regularisatioun/Regulatioun* ‚Regulierung‘ vs. *Reguléierung* ‚Regulierung‘, wieder zu Gunsten der ersteren Form, also 8 vs. 3 Tokens im Kernkorpus. Die zahlentechnische Konkurrenz von *-heet/-keet* und *-ioun* bestätigt sich semantisch gesehen nicht und findet daher nur am Rande statt. Beide Suffixe scheinen ihre eigenen Nischen zu haben und konkurrieren daher bezüglich der Basen nur innerhalb des Erweiterungskorpus in drei Fällen: *Absolutheet* ‚Absolutheit‘ – *Absolutioun* ‚Absolution‘, *Korrektheet* ‚Korrektheit‘ – *Korrektioun* ‚Korrektur‘, *Resolutheet* ‚Resolutheit‘ – *Resolutioun* ‚Beschluss‘. Hierbei handelt es sich eher um eine formale Konkurrenz und der Unterschied in der Semantik der Wörter lässt es zu, dass diese Formen parallel existieren können, z. B. ist *Korrektheet* ein (idealer) Zustand im Sinne von „fehlerfrei“, *Korrektioun* ist der Vorgang, der dazu führt.

Eine besondere Rolle spielt in der luxemburgischen Substantivbildung das Suffix *-echt*. Im Kernkorpus findet sich dieses Suffix in 113 Types, davon sind 44 Hapaxlegomena. Allerdings ist ein Großteil der Lexeme hier nicht dem Suffix selber zu verdanken, sondern der Tatsache, dass das Wort *Aarbecht* ‚Arbeit‘ im Luxemburgischen sehr kompositionsfreudig ist. *Schweessaarbecht* ‚Schweißarbeit‘, *Stroossenaarbecht* ‚Straßenarbeit‘, *Ënnerhaltsaarbecht* ‚Unterhaltsarbeit‘, *Botzaarbecht* ‚Putzarbeit‘, *Interimsaarbecht* ‚Interimsarbeit‘, *Gesamtaarbecht* ‚Gesamtarbeit‘ etc. Darüber hinaus gibt es tatsächlich Substantive im Luxemburgischen, die mit diesem Suffix gebildet werden, z. B. *Wourecht* ‚Wahrheit‘, *Klorecht* ‚Klarheit‘, *Heemecht* ‚Heimat‘, *Deierecht* ‚Teuerung, Preisanstieg‘, *Gewunnecht* ‚Gewohnheit‘ (auch Komposita wie *Iessgewunnecht* ‚Essgewohnheit‘). Hier ist eine klare Paradigmabildung anhand der bestehenden Analogie zu sehen. *-er* ist ein klassisch luxemburgisches

Suffix zur Bildung von Substantiven aus Verben mit mehreren Funktionsklassen, auf die hier nicht weiter eingegangen wird.

Ein interessanter Fall im Luxemburgischen ist die Konkurrenz von *-esch* und *-in* zur Bildung von Feminina aus Maskulina, besonders aus Wörtern, meist Verben, die auch eine Verbindung mit *-er* eingehen können. Das erstere als substantivbildendes Suffix kommt im Kernkorpus 105 Mal in 15 Types vor, davon sind 6 Hapaxlegomena. Dicht gefolgt wird dieses Suffix von *-in*, 100 Mal mit 36 Types und 22 Hapaxlegomena. Laut der Anzahl der Tokens haben die Suffixe gleiche Häufigkeit, allerdings legt letztere Gruppe mit doppelt so vielen Types und fast viermal mehr Hapaxlegomena eine höhere Produktivität an den Tag: $15/1\ 000\ 000 \approx 0,000\ 015$ vs. $36/1\ 000\ 000 \approx 0,000\ 036$ bei der realisierten Produktivität und $6/16\ 372 \approx 0,000\ 36$ vs. $22/16\ 372 \approx 0,0013$. Schaut man sich *-in* jedoch näher an, so findet man viele Wörter aus dem Deutschen, z. B. *Sekretärin*, *Beamtin*, *Schülerin*, aber auch viele luxemburgische Wörter, die das Suffix besitzen, wie z. B. *Kolleegin* ‚Kollegin‘, *Burgermeeschterin* ‚Bürgermeisterin‘, *Wielerin* ‚Wählerin‘. Die Konkurrenz der Suffixe um die Basen zeigt sich relativ deutlich in Beispielen wie *Meeschtesch* ‚Meisterin‘ vs. *Meschterin* ‚Meisterin‘, *Riednesch* ‚Rednerin‘ vs. *Riednerin* ‚Rednerin‘ oder *Ministesesch* ‚Ministerin‘ vs. *Ministerin* ‚Ministerin‘. Es gibt aber auch Basen, die nur mit einem der beiden Suffixe eine Verbindung eingehen, z. B. *Wielerin* ‚Wählerin‘, *Lëtzebuurgerin* ‚Luxemburgerin‘, *Nopesch* ‚Nachbarin‘. Bei den ersten zwei Wörtern besteht eine höhere Gefahr der Verwechslung mit dem adjektivbildenden Suffix *-esch*.

3. Technische Realisierung

3.1. Infrastruktur

Da im Institut für die Luxemburgische Sprache und Literatur bereits eine Sammlung von elektronischen Fließtexten in der luxemburgischen Sprache vorlag, war es naheliegend, die Organisation der Daten in XML zu halten. Es wurde hier absichtlich auf eine relationale Datenbank verzichtet. Stattdessen wurde Text Encoding Initiative⁵ (TEI) verwendet, wobei ein wichtiger Bestandteil der Arbeit bei der Überführung der Textdaten in XML die Erstellung von TEI-Headern war. TEI stellt Richtlinien für die Strukturierung von XML-Dateien und für die Benennung der XML-Elemente zur Verfügung, um Texte, die in den Geisteswissenschaften als Forschungsobjekte dienen, zu strukturieren und zu kodieren. Der Vorteil der Verwendung des TEI liegt einerseits in der Tatsache, dass für viele Textsorten jeweils ein bereits gut überlegtes Schema verwendet werden kann. Andererseits sind auf eine solche

5 <http://www.tei-c.org/index.xml>, zuletzt geprüft am 29.11.2017

Art und Weise kodierte Daten nicht nur leichter zu lesen für Forscher, die diesen Standard bereits kennen, vor allem dadurch, dass TEI englischsprachige Bezeichnungen für XML-Elemente verwendet, sondern können auch bereits dafür programmierte fertige Anwendungen mit diesen Daten umgehen. Im Untersuchungskorpus wurden zu einzelnen Textdateien Metadaten erstellt, um eben diese internationale Transparenz zu gewährleisten. Die Informationen aus dem Header wurden auch in die Arbeit von anderen Skripten und Programmen integriert, sodass in jedem Schritt der Verarbeitung genau ersichtlich war, aus welchen Daten die gefundenen Belege kommen. In den Metadaten wurden z. B. Informationen bezüglich des Titels, Autors, Pseudonyms und Veröffentlichungsdatums etc. festgehalten. Folgender Auszug aus dem Korpus ist ein Beispiel für die Metadaten:

```
<TEI>
<teiHeader xml:lang="deu">
<fileDesc>
<titleStmt>
<title type="main">D'Kerfegsblo'm</title>
<title type="sub">Én Geschicht a'us dem ale Letzeborger Volleksliewen an der Mu-
selspröch</title>
<title type="short">D'Kerfegsblo'm 1</title>
<author> <forename>Adolf</forename> <nameLink/> <surname>Berens</surname>
<addName type="pseudonym">Ale Mann</addName> </author>
<editor> <forename/> <surname/> </editor>
</titleStmt>
<publicationStmt>
<publisher/>
<pubPlace/>
<date>1921</date>
</publicationStmt>
<sourceDesc>
<bibl>Berens, Adolf: D'Kerfegsblo'm. Én Geschicht a'us dem ale Letzeborger
Volleksliewen an der Muselspröch. vun ém ale Mann. Grevenmacher 1921.</bibl>
<file>Berens_D'Kerfegsblom 1.pdf</file>
</sourceDesc>
</fileDesc>
```

Für die Speicherung der XML-Daten wurde die XML-Datenbank eXist-db (Version 2.2) eingesetzt⁶. Eine XML-Datenbank erlaubt nicht nur das zentrale Speichern von XML-Daten und einen Zugriff über das Internet, sondern auch eine elegante Möglichkeit zur Abfrage der Daten mittels X-Query, eines W3-Standards für die Abfrage von XML-Daten (Meier 2003, S. 171). Als

6 <http://exist-db.org/>, zuletzt geprüft am 29.11.2017

eine native XML-Datenbank stellt eXist-db darüber hinaus zusätzliche Tools zur Verfügung, beispielsweise eine graphische Oberfläche für Browser und/oder ein eigenständiges Java-Admin-Tool, um die XML-Daten übers Netz anzuzeigen und abzufragen. Eine andere Alternative zu eXist-db stellt die XML-Datenbank BaseX dar, die ähnliche Funktionen zur XML-Abfrage zur Verfügung stellt⁷. XML-Datenbanken genießen als NoSQL-Datenbanken den Vorteil, beliebig erweiterbar zu sein, und müssen keiner rigiden einheitlichen Struktur folgen. Allerdings brauchen sie mehr Platz zur Speicherung der Daten, im Falle von eXist-db 2,7 mal mehr Speicherplatz als es beispielsweise bei MySQL der Fall ist (Freire, Sergio Miranda et al. 2016). Die Reaktionszeiten sind von der Art der Abfragen und der zurückgelieferten Datenmengen abhängig, allerdings hat BaseX laut der letzten Studien die schnellste Reaktionszeit unter den XML-Datenbanken (Freire, Sergio Miranda et al. 2016). Der Grund für eine geringere Laufzeitgeschwindigkeit im Vergleich zu anderen Datenbankarten liegt in der Natur des XML-Formats – vor allem wegen der voll ausgeschriebenen öffnenden und schließenden Elementen-Syntax. Als Format scheint JSON bis zu 30% schlanker zu sein als XML, jedoch sind hier besonders große Mengen von Daten sehr schlecht menschenlesbar.

3.2. Algorithmen und Heuristiken zur automatischen Analyse der Wortbildung

3.2.1. Automatische Suche nach Wortbildungsauffixen und Probleme

Die Ambiguität von Sprachzeichen zeigt sich auch in der Wortbildung. Es gibt im Luxemburgischen viele Wortbildungselemente, die homonym oder Homografen sind. Das Morphem *-esch* beispielsweise, kann sowohl der Bildung weiblicher Substantive wie *Nop-er+esch* > *Nopesch* ‚Nachbarin‘ oder *Ried-er+esch* > *Riednesch* ‚Rednerin‘ als auch der Bildung von Adjektiven wie *afrikanesch* ‚afrikanisch‘, *technesch* ‚technisch‘ oder *evangesch* ‚evangelisch‘ dienen. Darüber hinaus gibt es viele Zeichenketten bei der Suche nach Wortbildungselementen, die je nach Kontext eine gänzlich andere, z. B. morphologische Funktion erfüllen, oder fester Bestandteil eines Wortstammes sind. Je kleiner das gesuchte Sprachzeichen, z. B. bei Wortbildungssuffixen *-el*, *-eg*, *-er* oder *-ei*, desto höher ist die Ambiguität. Das Element *-eg* beispielsweise, bildet Adjektive wie *dreckeg* ‚dreckig‘, *eekeleg* ‚eklig‘ oder *heefeg* ‚häufig‘, jedoch findet sich diese Zeichenkette auch in Substantiven wie *Mëtteg* ‚Mittag‘ oder *Deeg* ‚Tage‘ und somit in anderer Funktion. Daraus lässt sich schlussfolgern, dass eine formale Suche nach Zeichenketten oder regulären Ausdrücken sehr viele falsche Kandidaten liefern kann, die gesichtet werden müssen. Einen Teil davon kann man automatisch ausschließen, wenn man in die Suche die

7 <http://basex.org/>, zuletzt geprüft am 29.11.2017

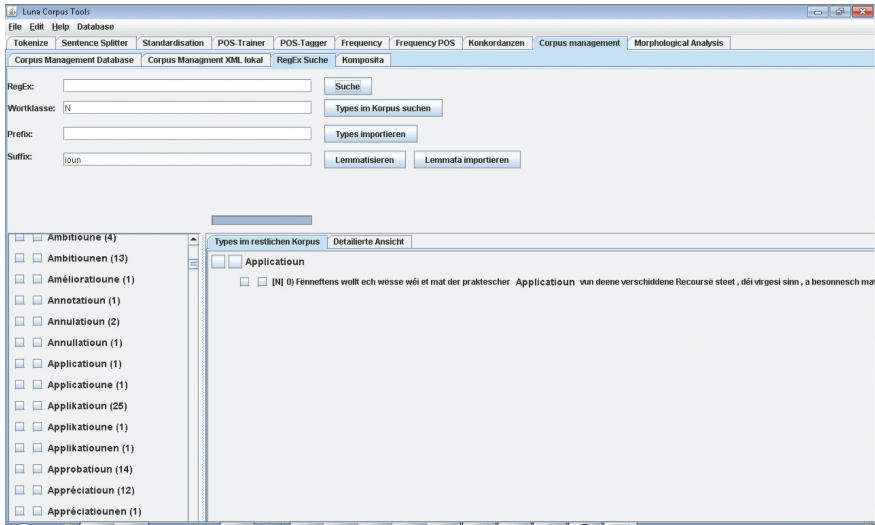
Annotation der Wortklassen miteinbezieht. Eine weitere Schwierigkeit bei der Suche nach Wortbildungssuffixen ist die Tatsache, dass diese sich sowohl im Luxemburgischen, wie auch in anderen Sprachen morphologisch gesehen näher bei dem Stamm befinden und meist vor den grammatischen Flexions-elementen vorkommen. Hinzu kommen im Luxemburgischen die N-Tilgung und die uneinheitliche Schreibung, die später diskutiert werden sollen. So hilft es nicht, die formale Suche mit einem Wortende einzuschränken. Im Falle der Zeichenkette *esch*, die im Gesamtkorpus als adjektivbildendes Suffix ca. 45 000 und als substantivbildendes Suffix ca. 1 200 Mal vorkommt, findet eine formale Suche nur 32 000 Wörter, die mit der Zeichenkette *-esch* enden. Denn die restlichen Belege sind flektierte Formen, bei Substantiven Pluralendung und bei Adjektiven die jeweiligen Angleichungen an den Kasus des beschriebenen Substantivs. Das Suffix *-ei* wie bei *Roserei* ‚Ärger, Zorn‘ oder *Molkerei* ‚Molkerei‘ zu finden hat sich als einfacher erwiesen als andere Suffixe. Es gibt im Kernkorpus über 92 000 Tokens, die mit dieser Zeichenkette enden. Beschränkt man die Suche auf die minimale Größe von vier Zeichen, bekommt man nur ca. 4 000, denn viele von diesen Tokens sind Funktionswörter wie *hei* ‚hier‘. Von diesen 4 000 sind aber fast die Hälfte Adjektive wie *schei* ‚scheu‘, *trei* ‚treu‘, *fonkelnei* ‚funkelneu‘ und Adverbien wie *elei* ‚hier‘ und damit uninteressant für die eigentliche Suche. Der Rest sind Substantive, aber sogar diese brauchen eine weitere Überprüfung, denn unter den Funden sind auch Wörter wie *Tierkei* ‚Türkei‘, die Ländernamen sind und keineswegs ein Produkt der Wortbildung. Anders verhält es sich bei dem Suffix *-eg*. Über 46 000 Tokens im Gesamtkorpus enden mit dieser Zeichenkette. Berücksichtigt man alle möglichen Flexionsendungen, steigt die Zahl der Funde auf ca. 62 000, wobei davon über 12 000 Substantive sind, die aus der Suche ausgeschlossen werden können. Zusammenfassend kann gesagt werden, dass eine bloße Suche nach Zeichenketten in einem Korpus nicht immer zu gewünschten Ergebnissen führt. Über die Regulären Ausdrücke hinaus zeigt sich hier die Verwendung der vorhandenen POS-Annotationen als eine gute Strategie. Falsche Kandidaten sind überall; auch dort, wo man sie nicht erwarten würde, z. B. bei Eigennamen wie *Akbar* bei der Suche nach *-bar* in Adjektiven wie *bemierkbar* ‚bermerkbar‘, *ëmsetzbar* ‚umsetzbar‘, oder *Ensembel* ‚Gesamtheit‘ und *Qualitéitslabel* ‚Qualitätslabel‘ bei der Suche nach *-bel* in Adjektiven wie *räsonnabel* ‚vernünftig‘ oder *akzeptabel* ‚akzeptabel‘. Diese falschen Funde können hervorragend vermieden werden, wenn in den Suchmechanismus POS-Annotationen miteinbezogen werden. Auf diese Art und Weise wird das Programm die Zeichenkette „bar“ oder „bel“ nur bei den Wörtern suchen, die als Adjektiv identifiziert wurden. Substantive wie *Ensembel* oder *Qualitéitslabel* können somit aus der Suche ausgeschlossen werden, obwohl sie rein formal mit der gesuchten Zeichenkette enden.

3.2.2. Lemmatisierung

Bei der korpusbasierten Arbeit mit Wortbildungsaffixen ist eine Lemmatisierung notwendig. Wie vorher im Abschnitt „Automatische Suche nach Wortbildungsaffixen“ beschrieben, kommen viele Wörter mit Wortbildungsaffixen im Korpus in flektierter Form vor. Viele Korpora in der englischen, deutschen oder französischen Sprache haben heutzutage bereits Annotationen mit Lemmata zu den jeweiligen Tokens, die hervorragend für diese Aufgabe verwendet werden können (Müller 2015: 2357). Allerdings gibt es für das Luxemburgische weder ein solch annotiertes Korpus noch verfügbare Werkzeuge, um solche Annotationen zu erzeugen. Um diese Hürde zu überwinden, stünden zwei Möglichkeiten zur Verfügung: Einen Lemmatisierer für das Luxemburgische zu implementieren oder nur die Belege zu lemmatisieren, also die Tokens, die ein Wortbildungsaffix besitzen. Im vorliegenden Fall wurde sich für die zweite Variante entschieden, zumal einerseits eine Implementierung eines Lemmatisierers für das Luxemburgische eine eigenständige Aufgabe ist und sehr lange dauern würde, andererseits besteht das Risiko, dass sich Fehler in die Analyse einschleichen. Wenn man jedoch nur die Belege lemmatisiert und dafür stochastische Methoden verwenden möchte, dann steht man vor dem bekannten „sparse data problem“. Also bleibt hier ein regelbasiertes, semi-automatisches und nicht trainierbares Verfahren übrig. Diese Belege müssen ohnehin gesichtet werden, sodass entschieden wurde, diesen Schritt damit zu kombinieren.

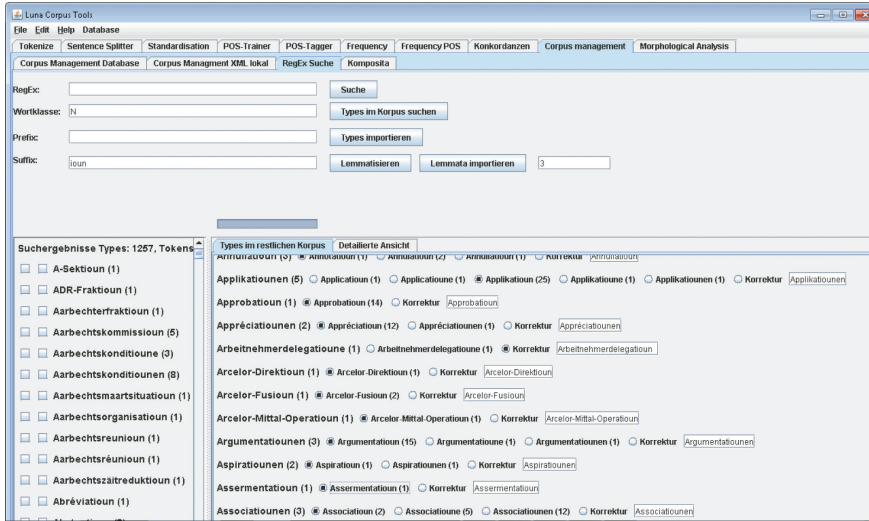
Das Verfahren, das angewendet wurde, besteht darin, dass man die Belege alphabetisch sortiert und diejenigen, die sich ähnlich sind, bündelt, ein Verfahren, das schon bei der Programmierung der ersten Lemmatisierer verwendet wurde (Klein und Rath 1971). Für die Bündelung der Belege wurde die Levenshtein-Distanz eingesetzt. Ein zusätzlicher Vorteil dieses Vorgehens besteht darin, dass man auf diese Art und Weise auch Wörter mit unterschiedlicher Schreibung bündeln kann. Wie zuvor erwähnt, besitzt das Luxemburgische mehr uneinheitliche Schreibung als die benachbarten Sprachen, obwohl schon seit 1975 eine offizielle Orthographie existiert.

Abb. 3: Überarbeitung der annotierten Belege in LuNa.



Als Beispiel für die Bündelung der flektierten Formen zu einem Lemma kann man z. B. die Types *Applicatioun*(1), *Applicatioune*(1), *Applikatioun*(25), *Applikatioune*(1) und *Applikatiounen*(1) betrachten. Hier wurde das Wort mit *c* einmal im Singular und einmal im Plural + *n*-Tilgung geschrieben und drei Mal mit *k*, einmal im Singular und zweimal im Plural ohne und mit *n*-Tilgung. In Klammern wird die Anzahl der jeweiligen Tokens in einem Type angegeben. Alle diese Tokens müssen auf ein Lemma zurückgeführt werden. Als Default bietet das Programm die häufigste Form als Lemma an, die auch mit einer Liste der möglichen Lemmata im Luxemburgischen automatisch verglichen wird. Dies muss dennoch von einem menschlichen Bearbeiter bestätigt werden, um automatisch generierte Fehler zu vermeiden. Wenn passende Lemmata nicht gefunden werden, kann der menschliche Bearbeiter das Lemma per Hand eingeben. Mit so einer Bestätigung, die meist nur einen Klick erfordert, werden in diesem Beispiel bereits 29 Wörter lemmatisiert und in der Datenbank gespeichert. Die Levenshtein-Distanz, benannt nach dem russischen Mathematiker Wladimir Iossifowitsch Lewenstein, der das Modell entwickelt hat, wird schon seit einer Weile in der linguistischen Forschung erfolgreich eingesetzt (List 2014: 145). Um die flektierten Formen eines Lemmas zu bündeln, reicht meistens eine Distanz von 3 oder 4 aus.

Abb. 4: Lemmatisierung der Belege in LuNa.



3.2.3. Basis- und Kontextanalyse

Die bisher beschriebenen Ansätze gehen meist von den Affixen aus. In Baayens Formeln werden Hapaxlegomena berücksichtigt, um vor allem die Neubildungen, die in der Sprache noch nicht häufig auftreten, in die Produktivität miteinzubeziehen. Hapaxlegomena können aber auch fossilisierte Belege beinhalten, die eher das Gegenteil darstellen. Aus diesem Grund ist es sinnvoll, eine Analyse der Basen der Wortbildung durchzuführen, die man auch größtenteils automatisieren kann. Solch eine Analyse würde nicht nur bezüglich der Verwendung der Basen ohne das dazugehörige Suffix Klarheit schaffen, sondern auch helfen, die Semantik und die Funktionsklasse der Suffixe besser zu bestimmen. Die Vorgehensweise ist einfach: Nachdem die Belege identifiziert und lemmatisiert bzw. annotiert wurden, kann man automatisch die Basen aus den Lemmata ausschneiden und im Korpus nach ihrer sonstigen Verwendung suchen. Eine solche Analyse wurde von Sirajzade (2012:87) mit den Verben durchgeführt. Ob die Ergebnisse solch einer Analyse mathematisch formuliert oder gar in die Produktivitätsanalyse miteinbezogen werden können, bleibt noch zu untersuchen.

Eine Kontextanalyse ist unabdingbar, wenn man bestimmen möchte, welche Funktionen Wortbildungsaffixe haben. Wie so häufig, kann auch hier Software den Linguisten nicht komplett ersetzen. Allerdings können die Tools die linguistischen Forschungen unterstützen; einerseits um die Arbeit schneller voranzutreiben, andererseits können hier Statistiken sehr hilfreich sein, um Beobachtungen zu untermauern oder zu widerlegen. Hier kann die Software

ein Fenster aus dem Kontext der Belege schneiden und die Wörter, die darin vorkommen, nach ihrer Häufigkeit sortieren. Zur algorithmischen Kontextanalyse passt auch die berühmte Feststellung von Firth: „You shall know a word by the company it keeps!“ (Evert 2004: 15), die in vielen linguistischen Applikationen im Bereich der Semantik zitiert wurde.

4. Fazit und Ausblick

Korpuslinguistische Methoden können einen wichtigen Beitrag zur linguistischen Forschung allgemein, und so auch zur Wortbildungsforschung im Speziellen leisten. Programmatisches Vorgehen kann viele Probleme lösen: Die Suche nach Wortbildungsaffixen, die Bestimmung ihrer Häufigkeiten, ihre Verteilung in einem Korpus, Festlegung ihrer Produktivität durch grundlegende statistische Analyse, Analyse des Kontextes der Belege und der Basen im sonstigen Sprachgebrauch. Jedoch entwickeln sich diese Methoden beständig weiter. Die Entwicklungen heutzutage, besonders im Bereich der Korpora, sind gewaltig. Vorhandene Annotationen erlauben tiefergehende Suche und Analysen. Es gibt mittlerweile nicht nur viele fertige Werkzeuge, die man verwenden kann, sondern die Algorithmen und Heuristiken, auf denen diese Werkzeuge basieren, werden ständig weiterentwickelt. Bei der Speicherung der Daten empfiehlt sich XML. Damit erzeugt man zwar mehr Daten als mit z. B. JSON oder wenn man die Daten in einer MySQL-Datenbank speichern würde, jedoch sind sie besser menschenlesbar. Durch TEI verwendet man einen internationalen Standard – diese Richtlinie gibt den geisteswissenschaftlichen Daten nicht nur eine gut überlegte Struktur, sondern ermöglicht es auch, dass diese Daten international verständlich und austauschbar sind. Die genaue Dokumentation der TEI-Versionen (P5) ist außerdem eine Voraussetzung für die Nachhaltigkeit. Setzt man TEI-XML in einem Projekt mit mehreren Mitarbeitern ein, so ist es empfehlenswert eine XML-Datenbank zu verwenden, um die Daten zentral zu verwalten.

Die Wortbildung ist ein fester Bestandteil der linguistischen Forschung und dafür zuständig, dass das Lexikon einer Sprache wächst und erneuert wird. Das Luxemburgische, das sich zwischen germanischem und romanischem Sprachraum befindet, zeigt hier eine besondere Dynamik. Wie festgestellt wurde, besitzt das Luxemburgische einerseits viele Affixe, die über Lehnwörter in die Wortbildung gelangt sind, andererseits gibt es Affixe, die sehr typisch für die luxemburgische Sprache sind und die sich in benachbarten Sprachen nicht wiederfinden lassen. Bei der Bildung von Adjektiven stellen *-lech*, *-eg*, *-esch* die produktivsten Suffixe dar. Die Suffixe *-bar* und *-bel* konkurrieren um einige gemeinsame Basen. Jedoch stellt sich *-bar* statistisch gesehen als die produktivere Variante dar, denn es kann zusätzlich zu den germanischen auch mit romanischen Basen eine Verbindung eingehen. Das Umgekehrte ist bei *-bel* nicht der Fall, dieses Suffix verbindet sich meist mit Basen romanischen

Ursprungs. Die Affixe, die mit Lehnwörtern in das Luxemburgische gelangen, können die Präsenz dieser zwar stärken, aber ändern die Wortbildung prinzipiell nicht sonderlich. Das Gesagte gilt auch für die Bildung der Substantive; auch hier gibt es Affixe, die über Lehnwörter Einzug in die Sprache gehalten haben. Das romanische *-ioun* trifft auf germanische Konkurrenten in Gestalt von *-heet/-keet* und *-ung*. Während die Konkurrenz mit *-heet/-keet* nur formal stattfindet und ein leichter Unterschied in der Semantik der Bildungen beiden Formen eine gleichzeitige Existenz erlaubt, stellt *-ung* einen ernsthaften Konkurrenten für *-ioun* dar und zeigt sich produktiver. Die Produktivität von *-ung* wird zusätzlich durch die Lehnwörter aus dem Deutschen gestärkt. Ein Großteil der Affixe ist jedoch als ursprünglich luxemburgisch zu bezeichnen. Sie sind moselfränkischen bzw. germanischen Ursprungs und entwickeln sich unabhängig vom benachbarten Deutschen, wie beispielsweise die Analyse der substantivbildenden Suffixe *-echt* oder *-esch* zeigt.

Literatur

- Ali, Sundus Muhsin & Khalid Shakir Hussein. 2014. The Comparative Power of Type/Token and Hapax legomena/Type Ratios. A Corpus-based Study of Authorial Differentiation. *Advances in Language and Literary Studies* 5(3). 112–119.
- Baayen, R. Harald. 2009. Corpus linguistics in morphology: morphological productivity. In Anke Lüdeling (ed.), *Corpus Linguistics. An International Handbook*. Volume 2 (Handbücher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science (HSK), 29(2)). 899–919. Berlin & Boston: de Gruyter.
- Belentschikow, Renate. 2015. Dictionaries. In Peter O. Müller (ed.), *Word-Formation. An International Handbook of the Languages of Europe* (Handbücher zur Sprach- und Kommunikationswissenschaft/ Handbooks of Linguistics and Communication Science (HSK), 40(3)). 2333–2354. Berlin & Boston: de Gruyter.
- Creutz, Mathias & Krista Lagus. 2002. Unsupervised Discovery of Morphemes. *Proceedings of the ACL-02 Workshop on Morphological and Phonological Learning* 6. 21–30. Stroudsburg, PA, USA: Association for Computational Linguistics (MPL '02). <http://dx.doi.org/10.3115/1118647.1118650>.
- Dal, Georgette & Fiammetta Namer. 2015. Internet. In Peter O. Müller (ed.), *Word-Formation. An International Handbook of the Languages of Europe* (Handbücher zur Sprach- und Kommunikationswissenschaft/ Handbooks of Linguistics and Communication Science (HSK), 40(3)). 2372–2386. Berlin & Boston: de Gruyter.
- Evert, Stefan. 2004. *The statistics of word cooccurrences. Word pairs and collocations*. Stuttgart, Univ., Diss. <http://elib.uni-stuttgart.de/opus/volltexte/2005/2371>.
- Freire, Sergio Miranda, Douglas Teodoro, Fang Wei-Kleiner, Eric Sundvall, Daniel Karlsson & Patrick Lambrix. 2016. Comparing the Performance of NoSQL Approaches for Managing Archetype-Based Electronic Health Record Data (PLoS ONE 11(3): e0150069). <https://doi.org/10.1371/journal.pone.0150069>.

- Gilles, Peter. 2006. Dialektausgleich im Luxemburgischen. In Claudine Moulin & Nübling, Damaris (eds.), *Perspektiven einer linguistischen Luxemburgistik*. 1–28. Heidelberg: Universitätsverlag Winter.
- Heid, Ulrich. 2015. Corpora. In Peter O. Müller (ed.), *Word-Formation. An International Handbook of the Languages of Europe* (Handbücher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science (HSK), 40(3)). 2354–2372. Berlin & Boston: de Gruyter.
- Klein, Wolfgang & Rainer Rath. 1971. *Automatische Lemmatisierung. Ein Bericht*. Saarbrücken: Universität des Saarlandes.
- Kohonen, Oskar, Sami Virpioja & Krista Lagus. 2010. Semi-supervised Learning of Concatenative Morphology. *Proceedings of the 11th Meeting of the ACL Special Interest Group on Computational Morphology and Phonology*. 78–86. Stroudsburg, PA, USA: Association for Computational Linguistics (SIGMORPHON '10). <http://dl.acm.org/citation.cfm?id=1870478.1870488>.
- Lee, Young-Suk, Kishore Papineni, Salim Roukos, Oussama Emam & Hany Hassan. 2003. Language Model Based Arabic Word Segmentation. *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics* 1. 399–406. Stroudsburg, PA, USA: Association for Computational Linguistics (ACL '03). <http://dx.doi.org/10.3115/1075096.1075147>.
- List, Johann-Mattis. 2014. *Sequence comparison in historical linguistics*. Düsseldorf: Düsseldorf University Press. <http://sequencecomparison.github.io/>.
- Lulling, Jérôme. 2002. *La créativité lexicale dans la langue luxembourgeoise*. Montpellier: Th. doctorat. Etudes germaniques.
- Meier, Wolfgang. 2003: eXist. An Open Source Native XML Database. In Akmal B. Chaudhri, Mario Jeckle, Erhard Rahm & Rainer Unland (eds.), *Web, Web-Services, and Database Systems: NODE 2002 Web- and Database-Related Workshops Erfurt, Germany, October 7–10, 2002 Revised Papers*. 169–183. Berlin & Heidelberg: Springer Berlin Heidelberg. http://dx.doi.org/10.1007/3-540-36560-5_13.
- Moulin, Claudine. 2004: Lëtzebuergesch, Université a Recherche. Lëtzebuergesch: Quo vadis? Actes du cycle de conférences. 107–119. Mamer: Melusina Conseil.
- Müller, Peter O. (ed.) 2015. *Word-Formation. An International Handbook of the Languages of Europe* (Handbücher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science (HSK), 40(3)). Berlin & Boston: de Gruyter.
- Pustyl'nikov, Olga & Karina Schneider-Wiejowski. 2010. Measuring Morphological Productivity. *Studies in Quantitative Linguistics* 5. 1–9.
- Schneider-Wiejowski, Karina. 2011. *Produktivität in der deutschen Derivationsmorphologie*. Bielefeld: Universität Bielefeld Dissertation.
- Sirajzade, Joshgun. 2013. *Das luxemburgischsprachige Oeuvre von Michel Rodange (1827–1876). Editionsphilologische und korpuslinguistische Analyse*. Trier: Universität Trier. <http://ubt.opus.hbz-nrw.de/volltexte/2015/914/pdf/Sirajzade2013.pdf>.

Straková, Jana, Milan Straka, Milan & Jan Hajič. 2014. Open-Source Tools for Morphology, Lemmatization, POS Tagging and Named Entity Recognition. *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 13–18. Baltimore, Maryland, Juni 2014. Association for Computational Linguistics.

Joshgun Sirajzade
University of Luxembourg
Faculty of Science, Technology and Communication
Computer Science and Communication
Maison du Nombre
6, Avenue de la Fonte
L-4364 Esch-sur-Alzette
joshgun.sirajzade@uni.lu

Rezensionen / Reviews

Pius ten Hacken

**Carola Trips & Jaklin Kornfilt (eds.) (2017),
*Further Investigations into the Nature of
Phrasal Compounding*, Berlin:
Language Science Press**

Phrasal compounding is a phenomenon illustrated by *slept all day look*. Prototypical examples are determinative compounds with a nominal head and a phrasal non-head. They raise interesting questions about the interaction of syntax and morphology and have been discussed in this context by Botha (1981) for Afrikaans and Lieber (1992) for English. Also in German and Turkish, they have received ample attention. This volume has as its main purpose to extend the range of languages for which phrasal compounds are discussed. It consists of a brief introduction (chapter 1), six chapters devoted to individual languages, and a final chapter with a more general outlook. The use of *further* in the title is perhaps surprising, in particular because the volume under review is the first of a new series. It is motivated by the fact that the papers are from “the second workshop on phrasal compounding”, held in Mannheim in 2015. In this review, I will first present and discuss each chapter, then consider some general points about the volume.

Chapter 2 by Kristín Bjarnadóttir is devoted to Icelandic. The author collected 900 phrasal compounds from corpora and from a large morphological database. She starts by outlining compounding in Icelandic in general. Icelandic has very long compounds formed by the recursive application of compounding. A nominal non-head can be a stem or a genitive noun. In the latter case, the ending is a proper genitive and not a linking element as in German. There are two types of phrasal compound, the traditional type and an innovative type. The former is not stylistically marked, the latter is informal. The range of structures is more restricted in the traditional type.

This chapter gives a good overview of Icelandic compounding. It raises many questions of delimitation, which are not really addressed. It remains unclear how compounds that look like phrases are distinguished from the corresponding phrases. Also the distinction between the two types of phrasal compounding is not described in a way that could lead to a clear delimitation. It seems a rather intuitive, pretheoretical distinction, but for claims of the kind made in the chapter, this does not seem sufficient to me.

Chapter 3 by Bogdan Szymanek focuses on Polish but also looks at other Slavic languages. It starts with an outline of compounding in Polish. Szymanek adopts phonological criteria to delimit compounds. In this definition

of compounding, most English compounds translate as syntactic phrases in Polish. Although phrases can be input to derivation, there are no phrasal compounds in Polish. This conclusion can be generalized to Slavic languages with the exception of Bulgarian.

This chapter shows that Polish differs from Germanic languages in the area of compounding. The question of how this difference is interpreted depends on which properties of compounding are taken as a criterion. Szymanek chooses phonology as central. This is a coherent perspective, but, as argued in ten Hacken (2013), it neglects the significant similarities in the onomasiological use of certain types of what Szymanek calls phrases in Polish to compounds in English. The question of whether Polish has phrasal compounds in such an onomasiological perspective of compounding is not addressed here.

Chapter 4 by Alexandra Bagasheva is about Bulgarian. Bagasheva argues that a new type of phrasal compound has gained currency in Bulgarian in certain registers or genres. Analysing specific sections of the magazine *Cosmopolitan*, she found phrasal compounds that cannot all be analysed as borrowings and calques, because in some cases there is no English counterpart. They can be left-headed or right-headed. Bagasheva assumes that they are evidence for a new construction schema. Whether this schema is borrowed or emerged as an extension of the existing N+N compounding schema is not easy to establish.

The chapter raises some interesting questions as to the nature of word formation as a component of a language. The use of phrasal compounds in a tightly delimited context of communication suggests that they are bound to a particular register. Speakers of Bulgarian using this register may then have the word formation rule (or construction schema) in their linguistic competence as well as the information about the restricted use of the rule in their pragmatic competence. The question of the origin of the rule can in my view not be answered at the level of the language, but only for individual speakers.

Chapter 5 by Kathrin Hein reports on a corpus-based study of phrasal compounds in German. Hein extracted a set of 1576 phrasal compounds from the newspaper component of the *Deutsche Referenz-Korpus*. Adopting a constructional model, she then classified these compounds in a bottom-up fashion. The main criterion for distinguishing fine-grained classes is the semantic relation between the head and the non-head. For higher-level classes, also form-based criteria were used. The result is an inheritance hierarchy of constructions.

The main problem with this chapter is that the author tries to cover too much ground in a short space. In many cases, she refers to her PhD thesis for a more detailed discussion of definitions and classifications. This is a work of over 500 pages. Of course, it is difficult to summarize a 500-page work in less than 8000 words, but when for a classification only the labels of the

classes are given, this is not sufficient to understand the argument that is made with the classification.

Chapter 6 by Kunio Nishiyama is on Japanese. In Japanese, phrasal compounds of the type discussed by Lieber (1992) are translated as phrases, e.g. noun phrases with a postposition or with the clitic *teki* ('like'). Because of rules of accent placement, it is possible to identify a different class of phrasal compounds. A minimal pair combines *doitu* ('Germany'), *bungaku* ('literature') and *kyookai* ('association') into a compound meaning 'association for German literature' or a phrase meaning 'German association of literature'. This contrast can be accounted for in Distributed Morphology by the distinction between real compounds and noun incorporation, which involve different morpheme combination rules. In this analysis, the level of *Word Plus*, proposed by Kageyama (2001) is no longer necessary.

Compared to the other chapters, this chapter is different on at least two counts. It is the only chapter on a non-European language and the only chapter adopting a Distributed Morphology framework. This double specificity creates a larger need for explanation of basic assumptions. The identification of compounds in European languages raises very different questions to their identification in Japanese. Much of the argument in this chapter seems to be directed to other Japanese researchers who will be familiar with the literature referred to here.

Chapter 7 by Metin Bağrıaçık, Aslı Göksel and Angela Ralli treats the Greek dialect of Pharasa. Pharasa is a place in Anatolia which had a Greek-language population until 1923, when this population was moved to Northern Greece after the Greek-Turkish war. The dialect was influenced by Turkish and Armenian when the speakers lived in Anatolia and by the Modern Greek standard after the relocation. Whereas almost all Modern Greek dialects have a compounding construction with a compound marker *-o-*, Phrasiot Greek has a compounding pattern with the first constituent in the genitive, which as such has a striking similarity to Turkish compounding. However, as opposed to Turkish, no phrasal compounding is found. This may be because in Phrasiot Greek, the compound marker is on the non-head, whereas in Turkish, it is on the head. This hypothesis is supported by Khalkha, a Turkic language with a compound marker on the non-head and no phrasal compounds.

This chapter is significantly longer than the others, over 12,000 words as against around 8,000 words for the others. It is also remarkable in the sense that only section 5 (c. 1,500 words) is devoted to the analysis of phrasal compounds. The earlier sections give a detailed overview of compounding in Phrasiot Greek, other "Hellenic" dialects and Turkish. Each assumption or distinction that is made in the analysis is explained with well-chosen examples. In this way, the chapter is interesting for the general description of

compounding and contextualizes the position of phrasal compounds in an exemplary way.

Chapter 8 by Jürgen Pafel has a different orientation. It is based on data from a range of languages that are not described systematically but used for a more general consideration of phrasal compounding. Pafel distinguishes four types of phrasal compound by means of the features [\pm well-formed] and [\pm quotative]. Quotations are reanalysed as nouns, so that quotative phrasal compounds are N+N compounds. Non-quotative phrasal compounds are distinguished as involving well-formed phrases (e.g. *over-the-fence gossip*) or non-well-formed phrases (e.g. German *Vor-Nobelpreis-Ära*, ‘before-Nobel prize era’). Then he distinguishes three ways of accounting for phrasal compounds, merge, insertion and conversion. Merge leads to [XP Y]_Y structures, which is only adequate for non-quotative, well-formed phrasal compounds. Insertion is also not sufficient for all types. Conversion, can account for all types of phrasal compound and provides the best mechanism for doing so.

With its emphasis on broader theoretical questions, this chapter could be taken as a kind of conclusion. However, it does not refer to the material in the preceding chapters and some of the assumptions are hardly compatible with them. In the classification of phrasal compounds, Pafel gives Italian *carta di credito* (‘card of credit’, i.e. credit card) and *cambialvalute* (‘change currencies’, i.e. money changer) as examples of the non-quotative types. However, in these cases the entire compound corresponds to the phrase, whereas in more prototypical phrasal compounds, the phrase is the non-head. Prepositional constructions of the type *carta di credito* are explicitly excluded from compounding in other chapters. Another problem I see is the criterion for [\pm quotative]. Here, Pafel contrasts the sentences in (1).

- (1) a. Jeder hat die Ob-ich-glücklich-bin-Frage beantwortet.
 ‘Everyone has the whether-I-happy-am-question answered’
 b. Jeder hat die “Bin ich glücklich?”-Frage beantwortet.
 ‘Everyone has the “am I happy?” question answered’

According to Pafel (p. 247), *ich* refers to the speaker in the non-quotative (1a), ‘Everyone replied to the question whether I am happy’, whereas in the quotative (1b) it refers to each individual in the scope of *jeder*, ‘Everyone replied to the question whether he/she is happy’. To me, both sentences are ambiguous between both readings with a strong pragmatic preference for the reading Pafel ascribes to (1b).

A general question that arises in the discussion of phrasal compounding is how to determine the boundary between phrasal compounds and phrases. Where this question is addressed explicitly, most chapters adopt phonological criteria, especially stress assignment. Interestingly, the editors state in their

introduction that they consider the non-separability of head and non-head the most reliable, crosslinguistically valid criterion for compounds (p. 6). In later chapters, this criterion is hardly mentioned if at all. In ten Hacken (2013), I argue for a semantic criterion, based on the way compounds receive their meaning. A certain correlation between theoretical outlook and compounding criteria can be expected. In a theoretical approach in which the generation of forms is the focus, phonological criteria receive a stronger weight. In an onomasiological approach, which highlights the naming function of compounds, semantics is more important. It is therefore not surprising that chapter 6, with its Distributed Morphology outlook, uses a phonological definition of compound. Chapter 4 on Bulgarian, which mentions Štekauer's (1998) onomasiological approach, makes more use of semantic considerations.

As a catalogue of overviews of compounding in different languages and phrasal compounding in particular, the editors produced a useful volume. The theoretical background of most chapters is fairly homogeneous. Apart from chapter 6 (Japanese), all chapters adopt or are compatible with a view of morphology based on constructions. Especially chapters 2 (Icelandic), 3 (Polish) and 7 (Greek) give well-documented general overviews of compounding that can be used more widely.

There are some rather unfortunate errors in the volume, some of which should normally have been caught in the editing process. Thus, Trips & Kornfilt (p. 2) mention Dutch examples, but they are Afrikaans. Bjarnadóttir (p. 20) has German *liebessbrief* without a capital. Nishiyama (p. 170) has "Chomsky's (2001) 2001 conjecture". Bağrıaçık et al. (p. 209) have "Jaspersen" for "Jespersen". Pafel (p. 240) uses "in the same vain". In some chapters, also the English could have benefited from a more careful proofreading, with some superfluous or missing articles and misplaced adverbs. Hein uses *substantive* instead of *noun*, which is not the normal English terminology in theoretical linguistics. Whereas these problems do not seriously affect the quality of the chapters, the fact that many non-English examples in chapters 4 and 5 are presented without translations restricts their use to readers who know Bulgarian and German, respectively.

In general, one can say that the volume is more valuable as a collection of individual papers than as a coherent overview of phrasal compounding. The editorial introduction is very brief. The part before the summaries of the chapters is just over 2000 words. The summaries in the introduction are very uneven, ranging in length from 68 words for Bagasheva's chapter 4 to 500 words for Bağrıaçık et al.'s chapter 7. The brevity of the introduction is not compensated for by a proper conclusion. Although Pafel's final chapter has a more comparative perspective, he does not refer to the earlier chapters. Perhaps this lack of attention for the editorial finishing of the volume is connected to the mode of publication. The default presentation of the volume

is as a freely downloadable PDF file or a separate file for each chapter. This has the important advantage that interested readers can download individual chapters, which may have been the typical use of the volume intended by the editors. In any case, in its electronic form this volume is great value for (no) money. A hard copy can also be ordered through Amazon.

References

- Botha, Rudolf P. 1981. A base rule theory of Afrikaans synthetic compounds. In Moortgat, Michael, van der Hulst, Harry & Hoekstra, Teun (eds.), *The scope of lexical rules*, 1–77. Dordrecht: Foris.
- ten Hacken, Pius. 2013. Compounds in English, in French, in Polish, and in General. *SKASE Journal of Theoretical Linguistics* 10. 97–113.
- Kageyama, Taro. 2001. Word Plus: The intersection of words and phrases. In Jeroen van de Weijer & Tetsuo Nishihara (eds.), *Issues in Japanese phonology and morphology*, 245–276. Berlin & New York: de Gruyter.
- Lieber, Rochelle. 1992. *Deconstructing Morphology: Word Formation in Syntactic Theory*. Chicago: University of Chicago Press.
- Štekauer, Pavol. 1998. *An Onomasiological Theory of English Word-Formation*. Amsterdam: Benjamins.

Prof. Dr. Pius ten Hacken
Leopold-Franzens-Universität Innsbruck
Institut für Translationswissenschaften
Herzog-Siegmond-Ufer 15
A-6020 Innsbruck
pius.ten-hacken@uibk.ac.at

Stefanie Stricker

**Nicolaus Janos Raag: Substantivkomposita
und Sinnggebung im Kontext
frühmittelalterlicher Wissensvermittlung.
Eine kulturanalytisch-linguistische
Untersuchung zur Wortbildung bei Notker III.
Studia Germanistica Upsaliensia 60. 229
S. Uppsala: Acta Universitatis Upsaliensis.
ISBN 978-91-554-9614-2.**

Die hier zu besprechende Dissertation des schwedischen Germanisten Nicolaus Janos Raag ist an der Universität Uppsala (Schweden) unter der Betreuung von Dessislava Stoeva-Holm entstanden. Die Arbeit will zeigen, welche Rolle Substantivkomposita im Rahmen der Wissensvermittlung und des Kulturtransfers spielen, wie deutsche Komposita lateinisch vermittelte Inhalte in die eigene frühmittelalterliche klösterliche Kultur der Rezipienten integrieren und welche Veränderungen sie dabei erfahren. Für die Untersuchung wurden Substantivkomposita ausgewählt, da sie prädestiniert für die Benennung von bisher Unbekanntem sind und durch ihre binäre Struktur die Fähigkeit besitzen, zwei Größen zueinander in Beziehung zu setzen, wobei die semantische Relation zwischen den Konstituenten eines Kompositums prinzipiell offen und auf morphologischer Ebene nicht ausgedrückt ist.

Komposita kommt zweifellos die Funktion der Benennung zu, wodurch sie geeignet sind, neue Bezeichnungen für Begriffe zu schaffen. Entsprechend tragen sie gerade in der frühen schriftlichen Überlieferung des Deutschen zur Ausdifferenzierung des Wortschatzes bei. Gleichzeitig verfügen sie über eine Typisierungs- bzw. Spezifizierungsfunktion, indem sie die Einordnung neuer Tatbestände in den Kontext bekannter Tatbestände ermöglichen. Für die Wissensvermittlung ist schließlich von Vorteil, dass Komposita durch ihre Prägnanz leicht memorierbar sind. In diesen Eigenschaften von Substantivkomposita liegt für den Verfasser das Potential, sie für die Wissensvermittlung nutzbar zu machen und somit auch für die eigene Fragestellung auszuwerten.

Angesichts dieser Eigenschaften scheint verwunderlich, dass es kaum umfassendere Untersuchungen gibt, die sich der Komposition im Althochdeutschen widmen (zum Forschungsstand Meineke 2007: 233; Raag 2016 mit weiterer Literatur), insbesondere nicht solche, die nicht die morphologischen Eigenschaften des Kompositums, sondern funktionale Fragestellungen in den

Blick nehmen. Diese Aufgabe nimmt sich N. J. Raag vor. Konkret konzentriert sich die Arbeit auf die Anfänge der deutschsprachigen Schriftlichkeit und der frühmittelalterlichen Wissensvermittlung. Als Untersuchungskorpus dienen sämtliche althochdeutsches Wortgut enthaltende Schriften Notkers III. von St. Gallen (um 950-28.6.1022) unter Einschluss der Glossen seiner Psalter-Bearbeitung. Ein wichtiger gemeinsamer Nenner der Schriften wird in ihrer Intention der frühmittelalterlichen Wissensvermittlung gesehen. Notker ist die große Lehrerpersönlichkeit am Ausgang des Althochdeutschen; seine Texte sind primär im schulischen Kontext der *septem artes liberales* zu verorten und auf den frühmittelalterlichen Unterricht und die Wissensvermittlung angelegt. Damit sind die Substantivkomposita aus Notkers volkssprachigem Schrifttum zweifellos ein idealer Ausgangspunkt für eine Untersuchung, die Strategien der Wissensvermittlung im historischen Sprachgebrauch auf die Spur kommen will. N. J. Raag will also zeigen, wie der Lehrer und Übersetzer Notker das Substantivkomposita innewohnende Potential kreativ nutzt, wenn er Begriffe aus dem Lateinischen ins Deutsche übersetzt, für die es bisher noch keine Bezeichnungen im Deutschen gab. Die Substantivkomposition soll hier folglich als eine mögliche sprachliche Strategie betrachtet werden, die dem frühmittelalterlichen Übersetzer zur Verfügung steht, wenn er lateinische Texte in die deutsche Sprache überträgt.

Methodisch sind für Untersuchungen historischen Sprachmaterials besondere Gegebenheiten zu berücksichtigen, die N. J. Raag auch umsichtig reflektiert. So ist weder der Sprecher noch der Rezipient klar greifbar. Die mittelalterliche Kommunikationssituation kann allein aus der sprachlichen Äußerung rekonstruiert werden. Die Substantivkomposita müssen daher in ihrem Kontext, also in ihrem konkreten textlichen Zusammenhang, sodann in ihren intratextuellen und intertextuellen Beziehungen sowie schließlich mit ihrer außersprachlichen Einbettung in das kulturelle Umfeld der frühmittelalterlichen Klosterschule beziehungsweise der Unterrichtssituation betrachtet werden. Es sind also möglichst alle Bedingungen aufzudecken, unter denen die Komposita als sprachliche Äußerungen produziert wurden.

Für dieses Vorhaben gliedert N. J. Raag seine Arbeit in sieben Kapitel. Die umfangreiche Einleitung enthält neben der Fragestellung auch eine Charakteristik der Wissensvermittlung und Bildung im Frühmittelalter sowie eine Vorstellung des im Mittelpunkt stehenden Wortbildungstyps der Komposition als grundlegendem Verfahren des Wortschatzausbaus (S. 13–44). Kapitel 2 liefert den theoretischen Rahmen, wobei insbesondere das Verhältnis von Sprache, Wissen und Kultur behandelt wird sowie kulturanalytische Ansätze in der Kompositionsforschung (so von Pasques 2003a und b zu Substantivkomposita bei Notker III., Solms 1999 zu Substantivkomposita im Frühneuhochdeutschen und Lobenstein-Reichmann 2004 zu Substantivkomposita bei Luther) vorgestellt werden. Sodann werden Substantivkomposita hinsichtlich

ihrer Funktion der Benennung neuer Begriffe sowie ihrer Verortung in der bekannten Welt (S. 45–71) charakterisiert. Kapitel 3 (S. 72–132) erläutert die Materialgrundlage, also Notkers Übersetzungen kanonischer Schulschriften und weitere Einzelwerke Notkers, sowie das Verfahren der Materialerhebung, das zu 864 Tokens (die Zahl der Types wird nicht genannt; das Verzeichnis auf S. 205–210 listet 116 Komposita auf) geführt hat. Diese stehen einem Gesamtwortschatz Notkers gegenüber, der in der Literatur mit etwa 8.000 Lexemen angegeben wird. Aus den aus Notkers Schrifttum aufgenommenen Wortbildungen (es handelt sich durchgehend um Determinativkomposita) hat N. J. Raag eine qualitative Auswahl getroffen, nämlich einerseits von Neubildungen, andererseits von solchen Bildungen, die gegenüber ihrer lateinischen Übersetzungsvorlage Abweichungen von einer einfachen Glied-für-Glied-Übersetzung zeigen. Es folgen Erläuterungen zur Bestimmung althochdeutscher Substantivkomposita sowie die Definition von Neubildungen, wobei die Probleme, die historische Korpora bei der Bestimmung einer Bildung als neu oder als einmalig, als usuell oder als produktiv mit sich bringen, ausführlich dargelegt werden. Für die Beurteilung der Wahrscheinlichkeit einer Neubildung werden fünf nachvollziehbare Kriterien (S. 111f.) entwickelt. Es folgen Darlegungen zu Substantivkomposita im Verhältnis zu ihrer lateinischen Vorlage (S. 112–125) sowie ein Exkurs zur Frage des Lexemstatus althochdeutscher Bildungen (S. 126–132). Kapitel 4 erläutert die methodischen Grundlegungen (S. 133–138) und Kapitel 5, das Herzstück der Arbeit, liefert die Ergebnisse der Analyse (S. 139–195), auf die noch genauer einzugehen ist. In Kapitel 6 schließen sich in knapper Form Überlegungen zu anderen Wortbildungsmodellen an [S. 196–199; mit Berücksichtigung von *prūtestôl* ‚Brautstuhl, Hochzeitsstuhl‘ (Np 483,15); *brūte stūol* ‚Brautstuhl, Hochzeitsstuhl, Hochzeitsbett(?)‘ (Nc 112, 1)]. Eine konzentriert verfasste Zusammenfassung (Kapitel 7, S. 200–204) beschließt die Dissertation, der dann noch ein Verzeichnis der 116 erhobenen Komposita (S. 205–210) und das Literaturverzeichnis (S. 211–229) folgen.

Für die Arbeit grundlegend ist die Unterscheidung von Kompositum und Nominalphrase. Als Substantivkomposition wird ausschließlich Stammkomposition verstanden (die sogenannte eigentliche Komposition). Dagegen wird die uneigentliche Komposition, bei der das Erstelement ein ehemaliges Flexionselement aufweist, noch nicht als Wortbildung betrachtet. Begründet wird diese Entscheidung durch die Genese der uneigentlichen Komposition aus Nominalphrasen mit Genitivattribut. Im Althochdeutschen war der attributive Genitiv im Normalfall noch vorangestellt. In dem Maße, wie sich die Nachstellung des Genitivattributs im Laufe der Sprachgeschichte durchzusetzen begann, wurden lexikalisierte Nominalphrasen zunehmend als zusammengesetzte Wörter und nicht mehr als Phrasen aufgefasst. Es handelt sich also da, wo das erste Element eine Genitivendung zeigt, um eine Nominalphrase

und nicht um ein Kompositum. Für die Materialauswahl ist damit ein streng formales Kriterium zugrunde gelegt worden, wonach alles, was der Form nach Genitivattribut sein kann, auch als Genitivattribut gewertet wird. Grundsätzlich werden nur Bildungen mit einem Wortstamm im Erstglied als Komposita betrachtet. Es werden also zwei unterschiedliche Modelle angenommen, einmal das Wortbildungsmodell der (Stamm-)Komposition und einmal ein syntaktisches Modell der Nominalphrasen.

Für die sich anschließende Bestimmung von Neubildungen stellt N. J. Raag fünf Kriterien auf, die in abnehmender Wahrscheinlichkeit Indikator für das Vorliegen einer Neubildung sein können. Es sind dies (S. 111–112): 1. Die Bildung ist ein Hapax legomenon; 2. Die Bildung übersetzt ein lateinisches Lemma, das an anderer Stelle anders übersetzt wird; 3. Die Bildung ist zwar nur im Notker-Korpus belegt, allerdings in höherer Frequenz; 4. Die Bildung ist außer bei Notker im Althochdeutschen nicht belegt, findet sich aber im Mittelhochdeutschen oder in rezenten Mundarten; 5. Die Bildung findet sich auch außerhalb des Notker-Korpus im Althochdeutschen. Die für die Substantivkomposition wie die Neubildung streng formalen Kriterien der Materialauswahl sind nachvollziehbar und gut operationalisierbar.

Für die Bedeutungsermittlung wird das von der Erlanger Forschergruppe anhand des frühneuhochdeutschen Dürer-Korpus entwickelte Verfahren des Aufbaus einer „Suppletivkompetenz“ (Müller 1993: 55–66; Habermann 1994: 57–70) nutzbar gemacht, das sich in der Forschung zur historischen Wortbildung für die Bedeutungsermittlung bewährt und etabliert hat. Die Suppletivkompetenz wird durch textinterne Paraphrasen, aufgrund von Wörterbuchparaphrasen oder nur aufgrund einer extralingualen semantischen Analyse aufgebaut. Sie trägt dazu bei, die gegenwartssprachliche Verstehenskompetenz zu ergänzen und damit die historische Distanz zu überbrücken. Die Notwendigkeit einer semantischen wie morphologischen Suppletivkompetenz gilt für das Althochdeutsche noch stärker als für das Frühneuhochdeutsche, da mit der größeren zeitlichen Entfernung auch die Möglichkeiten der Anwendbarkeit der eigenen gegenwartssprachlichen Kompetenz des Forschers abnehmen. N. J. Raag legt für seine Untersuchung fest, dass noch eine synchron motivierte Wortbildung anzunehmen ist, wenn eine Konstituente im Althochdeutschen außerhalb des Notker-Korpus oder im Mittelhochdeutschen belegt ist. Dieses Vorgehen ist schon wegen der geringeren Überlieferung des Althochdeutschen gegenüber dem Frühneuhochdeutschen nachvollziehbar.

Seine Analysemethode charakterisiert N. J. Raag als Verbindung eines morphosemantischen und eines hermeneutischen Zugriffs, indem neben dem unmittelbaren Kotext der außersprachliche kulturelle Kontext in die semantische Analyse einbezogen wird:

Die Bestimmung der aktualisierten Bedeutungen der einzelnen Substantivkomposita in ihren Belegstellen ist der Ausgangspunkt für die hier durchgeführte

qualitative kulturanalytisch-linguistische Untersuchung auf der Suche nach Musterhaftigkeiten, denen kulturelle Signifikanz zukommt. Für die semantische Analyse eines Kompositums heißt dies, dass man über die bloße syntaktische bzw. morphosemantische Auflösung des Kompositums in der Wortbildungsparaphrase hinaus auch und vor allem dessen Einbettung in den Kontext und den kulturellen Kontext geleistet werden muss. (N.J. Raag: 137)

Zentral für den gewählten Ansatz ist also, dass die morphosemantische Analyse von Substantivkomposita als Wortbildungsprodukten mit hermeneutischen Deutungsverfahren im kulturellen Kontext der historischen Kommunikationssituation, in der sie geäußert wurden, kombiniert werden. Die Substantivkomposita werden einerseits als komplexe Wortbildungsprodukte begriffen, bei denen die einzelnen Glieder zueinander, aber auch zur lateinischen Vorlage in Beziehung gesetzt werden, andererseits werden sie aber auch als sprachliche Äußerungen in einem bestimmten kulturellen Kontext begriffen und gefragt, welche Konsequenzen ihre strukturellen Eigenschaften auf kultureller Ebene haben.

Die Analyse zeigt, dass viele der Substantivkomposita auf Notkers Didaktisierungsbestrebungen zurückgeführt werden können, insbesondere die Neubildungen. Gleichzeitig wird Notkers Übersetzungswerk auch als ein Produkt des Kulturtransfers sichtbar, bei dem die zu vermittelnden Begriffe durch die Übersetzung in die eigene Kultur eingepasst werden, was N. J. Raag mit dem Begriff der Rekontextualisierung fasst (Raag 2016, 98). Diese Rekontextualisierung wird dann besonders deutlich, wenn Notker die Anpassung an die eigene Kultur als mehr oder weniger bewusstes Mittel zur Didaktisierung anwendet. Durch die wiederholte Verwendung bewährter Sprachmittel entstehen sprachliche Muster, von denen N. J. Raag drei als zentral herausarbeiten kann:

1. die Explikation von Inhalten. Die Substantivkomposita ermöglichen es, bestimmte Inhalte, die im lateinischen Ausgangstext nur implizit enthalten sind, in der volkssprachigen Wiedergabe explizit zu machen. So tragen sie auch zur Verdeutlichung der Textintention bei. Als Beispiele für metaphorische Deutungen, die nicht als Mittel, sondern als Inhalt der Wissensvermittlung dienen, erläutert N. J. Raag folgende Komposita: (1) *pérg-fúgeli* ‚Bergvöglein‘ (Np 34,15f.; S. 142f.); (2) *fínstir-land* ‚Finsterland, finsternes Land, Land der Finsternis: Ägypten‘ (Npgl 299,3; S. 144f.); für Wertungen, die durch die Komposition zum Ausdruck gebracht werden, diskutiert N. J. Raag: (3) *súnder-man* ‚Pharisäer‘ (Npgl 167,18); *súndir-lébo* ‚Pharisäer‘ (Npgl 347,13); *súndirguôt* ‚Pharisäer‘ (Npgl 234,23; S. 145–153); (4) *abGot-pilde* ‚Götzenbild‘ (Np 284,25); *crábeuuergh* ‚Schnitzwerk‘ (Np 397,23); *gráftpilide* ‚Schnitzbild‘ (Np 358,14; S. 153–156); (5) *uuilsálda* ‚Geschick, Glück (das nur begrenzte Zeit währt)‘ (Nb 39, 15; 217,10; 226,29; Nc 19,10; 20,1;

- 23,20; 25,10; 77,15); *uuil(e)-uuéndigi* ‚Veränderlichkeit innerhalb der kurz begrenzten Zeit (Zufall, Geschick, Schicksal)‘ (Nb 32,26; 47,15); *trúgesâlda* ‚falsches (trügerisches) Glück‘ (Nb 30,2); *trúgetíeuel* ‚Trugteufel, betrügerischer Dämon‘ (Nb 44,30); *zuífelsâlda* ‚schwankende Glücksverhältnisse‘ (Nc 26,16; S. 156–164).
2. die Zusammenfassung komplexerer Inhalte, also ein der Explikation entgegengesetztes Verfahren. Komposita leisten eine Univerbierung von Inhalten, die im Lateinischen durch komplexere Syntagmen ausgedrückt werden. Damit dienen sie gleichzeitig einer leichteren Memorierbarkeit des entsprechenden Sachverhalts. Aufgezeigt wird dieses Muster durch die Bildungen (6) *erdþíba* ‚Erdbeben‘ (Np 302,13f.; 165f.); (7) *érd-rât* ‚Ertrag der Erde; Nahrungsmittel, die die Erde hervorbringt (vgl. Vorrat)‘ (Nc 73,6; S. 166–168); (8) *âbent-chûoli* ‚Abendkühle‘ (Nb 31,15) und (9) *blíg-skîmo* ‚Lichtschein, Strahl‘ (Nc 58,6f.; S. 168f.).
 3. die erklärenden Übersetzungen. Hier werden solche Komposita betrachtet, die als erklärende oder didaktisierende Übersetzungen von lateinischen Fachbegriffen gesehen werden können. Das Muster wird verdeutlicht anhand der Bildungen (10) *síben bûohlíste* ‚die sieben Buchkünste = die Sieben Freien Künste‘ (Nb 54,23); *bûohlíst* ‚schriftliche Gelehrsamkeit und Bildung‘ (Nc 124,17; S. 170–172); (11) *óbeslíhtí* ‚Fläche‘ (Nk 41,5); *uêldslíhtí* ‚Fläche‘ (Nk 43,10; S. 173–176); (12) *crúzze-mélo* ‚Grüzmehl, Manna‘ (Nppl 279,12f.); *hímel-brôt* ‚Himmelsbrot, Manna‘ (Np 279,13); *Engelo brôt* ‚Brot der Engel‘, NP mit Genitivattribut (Np 279,14; S. 176f.).

Ein Bereich, in dem erklärende Übersetzungen massiert vorkommen, betrifft die Wiedergabe der Götter der griechisch-römischen Mythologie, die vor allem im Martianus Capella begegnen. Diese Gottheiten konnte Notker kaum als bekannt voraussetzen und so wählt er an vielen Stellen in seiner Übersetzung statt des Götternamens ein Kompositum oder eine komplexe Konstruktion, die die Eigenschaft oder Funktion der jeweiligen Gottheit ausweist. Das gilt für (13) *frámspúotesâre pales* ‚der Glücksbringer Pales‘ (Nc 53,4; S. 178f.); (14) *fúotergót* ‚Pales (Feldgottheit)‘ (Nc 53,1; S. 180); (15) *héllegóten* Pl. ‚unterirdische Götter, Götter der Unterwelt‘ (Nb 180,5); *héllegót* ‚Gott der Unterwelt: Pluto‘ (Nc 8,14f.; S. 181f.); (16) *méregót* ‚Gott des Meeres: Portunus, Neptunus‘ (Nc 8,15; 53,16; S. 182f.); (17) *sáng-cútten* ‚Göttin des Gesangs: Muse, Kamene/lat. Camena‘ (Nc 3,11); *sáng-cúten* ‚Göttin des Gesangs: Kalliope/lat. Caliopea‘ (Nc5,5f.); *sángcútenna* Pl. ‚Göttinnen des Gesangs: Musen/lat. Musae‘ (Nc 38,115. 183–185); (18) *uuíngot* ‚Kriegsgott: Gradivus (= Mars)‘ (Nc 8,15; S. 185); (19) *uuíngot* ‚Weingott: Bacchus‘ (Nc 4,19); *uuíngot* ‚Weingott: Liber‘ (Nc 53,4; S. 185–187); (20) *érdrónua* ‚Erdherrin/Mutter Erde: Cybele‘ (Nc 146,18; S. 187f.); (21) *búrghalto* ‚Beschützer der Stadt: Mars Quirinus‘ (Nc 53,2; S. 188); (22) *chórngéba*

,Korngeberin: Ceres‘ (Nc 52,20; 76,9f.; S. 188); (23) *érdcot* ‚Erdgott: Tellurus‘ (Nc 52,20); *érd-kóta* Pl. ‚Erdgötter, irdische Götter: Heroen‘ (Nc 141,12; S. 189f.); (24) *zuifelgót* ‚Gott der Ungewissheit/des Zufalls: Fors‘ (Nc 52,2; S. 190f.); (25) *hímel-iuno* ‚himmlische Juno: Ops (vielleicht als Mutter der himmlischen Götter)‘ (Nc 54,5; S. 191f.); (26) *líut-frómuu* ‚Schutzherrin des Volkes: Populonia (= Juno)‘ (Nc 134,17); (27) *flôz-kéba* ‚Fluvonia, Juno als für die Monatsblutung zuständige Ehegöttin‘ (Nc 134,2f.); (28) *cúrtilfléga* ‚Hüterin des Brautgürtels: Cinxia, Juno als Ehegöttin‘ (Nc 134,9); (29) *hém-brínga* ‚die Nach-Hause-Geleitende (Juno als Ehegöttin, die die Braut in das Haus des Bräutigams führt)‘ (Nc 134,8f.); (30) *hífûoga* ‚Ehestifterin: Juno‘ (Nc 134,15); (31) *sâmo-géba* ‚Samengeberin: Saticena (= Juno)‘ (Nc 134,14; S. 192f.); (32) *hígót* ‚Ehegott: Hymenaeus‘ (Nc 3,1); *hímachare* ‚Ehestifter: Hymenaeus‘ (Nc 4,9; S. 193f.); (33) *uuérh-méistera* ‚Werkmeisterin: Pallas Athene, die Göttin aller Künste und Gewerbe‘ (Nc 56,15f.; S. 194); (34) *fúoro gébo* ‚Triptolemus, Erfinder des Ackerbaus; Speisegeber‘ (Nc 156,15f.); *héllo gót* ‚Pluto, Gott der Unterwelt‘ (Nb 181,7); *héllo iouis* ‚Veiovis, Pluto oder Orcus; Gott der Unterwelt‘ (Nc 54, 7); *héllo uuárt* ‚Vedius .i. Orcus; Höllenhüter, Gott der Unterwelt‘ (Nc 130,19); *uuéido gúten* ‚Göttin der Jagd (Diana)‘ (Nc 146,13; S. 195).

Die einzelnen Analysen verdeutlichen, dass die Komposita nur richtig verstanden werden, wenn sowohl ihr unmittelbarer Kontext als auch weitere Informationen aus dem Kontext theologischen Denkens im frühmittelalterlichen Kloster berücksichtigt werden. Für das Verständnis der Komposita werden von N. J. Raag auch intertextuelle und intratextuelle Referenzen nutzbar gemacht. Insgesamt spiegeln Notkers Bildungen das Bemühen des Lehrers und Übersetzers um die adäquate, erklärende Wiedergabe des im lateinischen Text Gemeinten wider.

Die Arbeit leistet eine nachvollziehbare und gelungene Analyse der Substantivkomposita im Werk des bedeutenden Klosterlehrers Notker von St. Gallen. Durch die kulturanalytische Betrachtung der Komposita mit Blick auf ihre Rolle bei der Vermittlung neuer Wahrnehmungs- und Deutungsmuster wurde eine ganzheitliche Perspektive auf Komposita in ihrem pragmatischen und kommunikativen Verwendungszusammenhang eröffnet, die über eine bloße morphologische Betrachtung hinausgeht, sie aber nicht außer Acht lässt. Es wird sichtbar, wie Notker zum Vermittler der zunächst nur lateinisch fassbaren Kultur wird, indem er den lateinischen Text kommentierend, verdeutlichend oder zusammenfassend ins Deutsche überträgt, ihn dabei in der Zielsprache und in der Zielkultur verständlich macht und auf diese Weise gleichzeitig zu einer Wortschatz- wie Funktionserweiterung der deutschen Sprache beiträgt. Die Ergebnisse der Analyse demonstrieren die Tragfähigkeit der Methode, die aber zumindest an einer Stelle auch kritikwürdig ist. So ist nicht nachvollziehbar, dass in der Arbeit (wie in weiteren Aufsätzen

des Verfassers zu diesem Thema) das Leipziger Althochdeutsche Wörterbuch gänzlich unbeachtet bleibt. Dieses hätte (für die bereits erschienenen Alphabetstrecken) nicht allein eine Erleichterung für die Analyse dargestellt, sondern ein unverzichtbares Korrektiv. Das Wörterbuch bietet zu den behandelten Komposita mehrere, nicht selten sogar alle Notker-Stellen mit einer grammatischen Bestimmung und einer Bedeutungsangabe und – wo nötig – auch einem Kommentar, den kein anderes Wörterbuch bietet, jedenfalls nicht in dieser Ausführlichkeit und Exzellenz. So zeigen sich dort auch Varianten (z. B. bei *gráftpilide* S. 156, *bûohlist* S. 170–172, *érdfróuuvn* S. 187, *búrghalto* S. 188, *chórngéba* S. 88, *flôz-kéba* S. 192, *hîfûoga* S. 192, *fûoro gébo* S. 195) oder hilfreiche Hinweise (z. B. bei *érd-râtes* S. 167, *blîg-skîmo* S. 168f., *uêldslîhtî* S. 173–175, *frámsspûotesâre pales* S. 178, *fûoter-gót* S. 180, *érdcot* S. 189, *hîmel-iuno* S. 191, *hîgot* S. 193), insbesondere hinsichtlich der Bedeutungsangaben. Zudem wäre es zweckmäßig gewesen, die Frage der Zusammen- und Getrennschreibung bzw. Bindestrichschreibung für die eigene Darstellung zu erläutern, zumal, wenn eine Varianz zwischen der Schreibung im Ansatz des Kompositums und im zitierten Textauszug begegnet (bei *brûte stûol* S. 196 und anderen Fällen). Dieser Aspekt ist schon deshalb nicht trivial, da die Bindestrichschreibung als reine Verständnishilfe eingesetzt werden kann (so bei King in seiner Ausgabe Notker der Deutsche 1972–2009, Band 5, S. XVIII), aber auch eine morphologisch-syntaktische Entscheidung anzeigen kann. Der Bindestrich wird auch nicht in der Legende des Verzeichnisses der im Text genannten und behandelten Komposita (S. 205–210) thematisiert, wo Getrennschreibungen, Zusammenschreibungen und Bindestrichschreibungen nebeneinander begegnen. Ungeachtet dieser kleineren Monita trägt das Buch zweifellos zu unserer Kenntnis über Sprache und Leistung des großen Klosterlehrers Notker von St. Gallen am Ende der althochdeutschen Zeit bei. Es liefert – insbesondere in Zusammenschau mit Studien anderer Zeitschnitte – einen wertvollen Baustein für eine diachrone Beschreibung des Systems der Substantivkomposition im Deutschen.

Literatur

- Althochdeutsches Wörterbuch*. Auf Grund der von Elias von Steinmeyer hinterlassenen Sammlungen im Auftrag der Sächsischen Akademie der Wissenschaften zu Leipzig. Bearbeitet und hg. von Elisabeth Karg-Gasterstädt und Theodor Frings. I–VI. VII. Lieferungen. Berlin 1952–2017 [weitere Bände in Bearbeitung].
- Habermann, Mechthild. 1994. *Verbale Wortbildung um 1500. Eine historisch-synchrone Untersuchung anhand von Texten Albrecht Dürers, Heinrich Deichslers und Veit Dietrichs* (Wortbildung des Nürnberger Frühneuhochdeutsch 2). Berlin & New York. de Gruyter.
- Lobenstein-Reichmann, Anja. 2004. Wortbildung bei Martin Luther: Systematische, semantische und kommunikative Aspekte. In Klaus J. Mattheier & Haruo Nitta

- (eds.), *Sprachwandel und Gesellschaftswandel – Wurzeln des heutigen Deutsch. Studien des deutsch-japanischen Arbeitskreises für Frühneuhochdeutschforschung*, 69–97. München: iudicium.
- Meineke, Eckhard. 2007. Bindungssemantik der Substantivkomposita im Althochdeutschen, Altsächsischen und Altniederfränkischen: Überlegungen zu einem Forschungsvorhaben. In Claudia Wich-Reif (ed.), *Strukturen und Funktionen in Gegenwart und Geschichte: Festschrift für Franz Simmler zum 65. Geburtstag*, 231–256. Berlin: Weidler.
- Müller, Peter O. 1993. *Substantiv-Derivation in den Schriften Albrecht Dürers. Ein Beitrag zur Methodik historisch-synchroner Wortbildungsanalysen* (Wortbildung des Nürnberger Frühneuhochdeutsch 1). Berlin & New York: de Gruyter.
- Notker der Deutsche. 1972–2009. *Die Werke Notkers des Deutschen*: Neue Ausgabe. Hg. von James C. King/Petrus W. Tax. Begründet von Edward H. Sehr/Taylor Starck. 10+8 Bde. (Altdeutsche Textbibliothek 73–75, 80–81, 84, 87, 91, 93–94, 98, 100–101, 109, 117, 120–122). Tübingen: Niemeyer.
- Pasques, Delphine. 2003a. ‘Ausdrucks- und Auslösfunktionen’ der Nominalkomposita in Notkers Psalter. *Sprachwissenschaft* 28. 95–110.
- Pasque, Delphine. 2003b. La composition nominale dans le Psautier de Notker (XI^e siècle): Modèles et fonctions. Paris: Université Paris-Sorbonne 4 Dissertation.
- Raag, Nicolaus Janos. 2016. Wissensvermittlung durch Substantivkomposita im frühmittelalterlichen Kloster. In Elke Hentschel (eds.), *Wortbildung im Deutschen: Aktuelle Perspektiven*, 89–110. Tübingen: Narr Francke Attempto.
- Solms, Hans-Joachim. 1999. Der Gebrauch uneigentlicher Substantivkomposita im Mittel- und Frühneuhochdeutschen als Indikator kultureller Veränderung. In Andreas Gardt, Ulrike Haß-Zumkehr & Thorsten Roelcke (eds.), *Sprachgeschichte als Kulturgeschichte* (Studia Linguistica Germanica 54), 225–246. Berlin & New York: de Gruyter.

Stefanie Stricker
 Universität Bamberg
 Deutsche Sprachwissenschaft
 Hornthalstraße 2
 D-96047 Bamberg
 stefanie.stricker@uni-bamberg.de

Varia / Miscellaneous

Conference report

“Morphology Days”, Louvain-la-Neuve (Belgium), November 22–24 2017 (Edwige Dugas)

Morphology Days is a (nearly) biennial international meeting which deals with morphology within different frameworks and in various perspectives. Previous editions of this conference have taken place in Leuven (2015), Leeuwarden (2013), Leiden (2012), Nijmegen (2011), Luik (2009) and Amsterdam (2007). While the first editions of the conference were mainly addressed to researchers working on morphology in the Netherlands and in Belgium, the last editions – including this one – included international contributions. The programme and the book of abstract is available at the conference’s homepage at <https://morphologydays2017.wordpress.com/program/>. Organized by Philippe Hiligsmann, Kristel Van Goethem, Nikos Koutsoukos and Isa Hendriks from the Université Catholique de Louvain, and Laurent Raiser from the Université de Liège, this edition of *Morphology Days* hosted more than 30 researchers, among which 3 plenary speakers, coming not only from Belgium but also from France, Germany, Greece, Hungary, Italy, Japan, the Netherlands, Spain, the United Kingdom and the United States. Although both inflection and derivation (affixation) were dealt with in the talks, this conference report will only address the studies on derivation.

The *Morphology Days* conference gathered postgraduate students and young researchers as well as more “expert” ones, thus enabling interesting discussions and fruitful feedback. The conference presented a wide range of topics, approaches and methods: general and special languages, diachrony and synchrony, composition and affixation (including zero-affixation, conversion), lexematic morphology, construction morphology, distributed morphology, relational morphology, collocation analysis, etc. This blending of different ways to study derivational processes undoubtedly provided very interesting perspectives for further research. Besides, several languages were under study; although English represented the lion’s share of the presentations, Dutch, French, German, Greek, but also Japanese were among the languages examined in the talks. All in all, this *Morphology Days* meeting was particularly successful and all participants are looking forward the next edition of the conference.

Second language acquisition was at the center Hélène Giraudo’s and Isa Hendriks’ talks. Hélène Giraudo (Université de Toulouse) opened the conference with a plenary talk on morphological processing and the role of morphological

structure of complex and derived words. On the basis of data coming from second language acquisition and visual word perception, Giraudo chose to take a construction morphology perspective (Booij, 2010) rather than a decomposition one, to show that systematic form-meaning correspondences are crucial for language and reading acquisition. This question of the link between form and meaning was actually a common thread linking most talks of conference.

The talk of Isa Hendrikx (Université catholique de Louvain) also dealt with language acquisition, in a usage-based Construction Morphology perspective. The author was interested in the impact of Content and Language Integrated Learning (CLIL) in the acquisition of intensifying constructions (e.g. *blood-heet*, lit ‘blood-hot’, very hot) in French-speaking L2 learners of Dutch and she investigated the longitudinal developments in the learners’ acquisition of these constructions. The diversity of the constructions available to express intensification (phrasal constructions, compounding, etc.) was assumed to complicate the acquisition process and to involve a competition between L1 and L2 constructions (Ellis & Cadierno 2009). The data used for this study was a corpus of written productions in the form of fictional e-mails on the subject of a party or holidays. Using colostruational analysis to determine the degree of attraction or repulsion of a lexeme to an intensifying construction (e.g. Stefanowitsch & Gries 2003), the author showed, among other things, that CLIL and non-CLIL learners do not use the same constructions with the same frequency.

The link between form and meaning was further examined by Serena Dal Maso (Università di Verona) and Sabrina Piccinin (CLLE, Université de Toulouse), who investigated the consistency of three morphological series in Italian: words with suffixes *-mento* (e.g. *trasferimento* ‘transfer’), *-ino* (e.g. *tavolino* ‘small table’), and *-olo* (e.g. *fruttivendolo* ‘greengrocer’). The analysis of corpus data revealed that these three suffixes display various degrees of consistency, which was measured with respect to several parameters. A first group of parameters were related to the consistency of the word ending series: (i) the proportion between the number of truly affixed and pseudo-affixed words containing the same ending, (ii) the proportion between token frequency counts of affixed and pseudo-affixed words containing the same word ending, (iii) the number of meanings carried by a given suffix, (iv) the occurrence or non-occurrence of semantic bleaching, (v) possible subpatterns due to cumulation of affixes, (vi) the stability in the prosodic structure. A second group of parameters dealt with the stability of the functional features of the base: (vii) the lexical category of the base, (viii) the stability of the gender of the base.

Two groups of authors addressed the diachronic evolution of language. First, Arne Dhondt, Timothy Coleman and Johan De Caluwe (Ghent University) examined Dutch particle verbs such as *nalezen* ‘proofread’ from a

Construction Morphology perspective. These verbs have been analyzed by Booij as instantiations of constructional schemas with a lexically specified particle slot and an open verb slot, and an idiosyncratic meaning which differs from the meaning of the adposition (in that case, *na*) from which it is diachronically derived. Drawing on data from the Modern Dutch period (i.e. from the 17th century), the authors sought to describe the emergence of such particle verbs schemas with the particle *op* ‘up’ as a case study. They showed that the meaning of the particle, ‘to use to its full extent by V-ing’ started out in semantic contexts of consumption before spreading to other contexts.

Diachrony was also the focus of Asimakis Fliatouras’ (Democritus University of Thrace) talk, where Greek language was used as a case study to (re-)examine the grammaticalization parameters traditionally described in the literature (a.o., Lehmann 1982/1995), namely: (i) phonological erosion, (ii) de- or resemanticization, (iii) de- or transcategorialization, (iv) extension, (v) paradigmaticization. On the basis of evidence from Ancient Greek, the author proposed to add another parameter, namely morphological analysis. This parameter plays a role in the case of verbal stems as second constituents in exocentric compound adjectives which are detached from the compounds absorbing the external adjectival zero suffix and which become suffixes (e.g. *-o:d-is* ‘who smells’ > *-ó:d-is* ‘diminutive’), or adverbs which have been re-analyzed as prefixes (e.g. *hupó-lefkos* ‘white-like’). Asimakis Fliatouras also showed that the parameters playing a role in morphological grammaticalization are not of equal weight and he underlined the fact that the parameters are closely interrelated.

Jenny Audring (Leiden University) examined the structural ambiguity of complex words in Dutch, English and German. She showed that in some cases, complex words match two or more derivational patterns instead of only one, as in Dutch *selfie*, which can be analyzed as $[_N[_{prt} self]-ie]$ or as $[_N[sel]-fie]$. Indeed, Dutch has developed a productive suffix *-fie* ‘photo of oneself with X/while doing X’, which gives to the English loan *selfie* a synchronically ambiguous structure. To account for structurally ambiguous words as reflecting multiple lexical relations, the author proposed a declarative and relational model of morphology (Jackendoff & Audring, *fc.*), based on some of the premises of Construction Morphology (Booij 2010).

Edward Göbbel (University of Wuppertal) studied *-ie* hypocoristics (*comfy*, *fattie*), a phenomenon which is at the interface of derivational (e.g. *leftiness*, *leftie friendly*) and inflectional morphology (e.g. *comfier*), in the framework of Distributed Morphology. It was shown that most *-ie* hypocoristics are based on noun phrases and that, as a consequence, *-ie* must be analyzed as a phrasal affix. Distributed Morphology allowed the author to demonstrate that *-ie* is a head in the extension of a noun or adjective and that the phrasal hypocoristics are derived by ellipsis of the noun and cliticisation of *-ie* onto the adjective.

Camiel Hamans (Universiteit van Amsterdam) studied the phenomenon of blends in English. Whereas certain authors seem to call into question the fact that blending is a regular morphological process (Bauer 2003), other emphasize the regularity of the phenomenon (e.g. Kelly 1998, Gries 2004). The author provided evidence in favour of these analyses; the regularity of this phenomenon lies in the fact that the second source word is crucial as there is a faithfulness constraint on its phonological and metrical properties which is actually stronger than the constraints aiming at preserving the properties of the first source word.

Katrin Hein (Institute for the German Language IDS) investigated the productivity of German compounds (e.g. *papstviolett* ‘pope purple’, *Höhenangst*, ‘fear of heights’) in lexematic morphology. While the productivity of affixation processes has received much attention in the literature, that of composition is a relatively understudied domain. Taking as a starting point the observation that semantically similar words display different tendencies to occur as head words in compounds (Fleischer & Barz 2012), the author examined groups of compounds with similar head words. She also evaluated the respective roles of the different factors of productivity, such as the morpho-phonological, morpho-syntactic and semantic properties of the immediate constituents, and focused particularly on the characteristics of the head words when they occur in isolation and how it affects the productivity of the compound.

In another talk, Katrin Hein (Institute for the German Language IDS) together with Adrien Barbaresi (Berlin-Brandenburg Academy of Sciences and Austrian Academy of Sciences) discussed the detection and characterization of German phrasal compounds (PCs, e.g. *Man-muss-doch-über-alles-reden-können-Credo*, ‘one-should-be-able-to-talk-about-everything-motto’) in large web corpora, with a method based on word segmentation and data-driven learning. German PCs come in two flavours, with or without hyphen, and the authors proposed the hypothesis that the former are less expressive than the latter. It was also shown that on the web certain lexemes have a bigger probability to appear as head words of PCs, such as nouns denoting human beings, as is the case of PCs found in newspapers (Hein 2015, Trips 2016), but that PCs also show different characteristics in the two types of corpora. This study also underlined the existence of borderline cases between phrasal compounding and phrasal derivation.

Kazuya Nishimaki’s (University of Tsukuba) presentation dealt with a somewhat neglected word-formation phenomenon, that of conversion. He examined the case of nouns and their so-called ren’yō verbal form (e.g. *basiri* ‘a run’ / *basiru* ‘to run’). Although they are generally treated as converted nouns, the author proposed an analysis based on a comparison with English and showed that these forms are similar to English gerundive nominals. As English gerundive nominals, which transform verb phrases into nouns (e.g.

to keep a house / house-keeping), they inherit their argument structure from their underlying verb phrase. As for *-ing*, the only function of the morpheme *-i* is to change the category of the base.

Finally, let us mention two groups of authors who took a different perspective on word-formation processes: Andreas Nolda, who underlined the necessity to make a distinction between the phonological and graphemic forms of words, and Edwige Dugas and Jérôme Michaud, who applied the framework of Construction Morphology to the domain of specialized languages.

Andreas Nolda (University of Szeged) examined the understudied domain of German cardinals such as *sechzehn* ‘sixteen’ to demonstrate that phonological and graphematic forms have their own characteristics and must be examined separately. As a matter of fact, while the written form of German cardinals can be analyzed as a case of compounding, their spoken form should rather be seen as an instance of suffixation. Evidence for such an analysis comes from the fact that German cardinals are phonetically realized in a reduced way, a phenomenon which is typical of accentless affixes. The study was couched within the framework of Pattern-and-Restriction Theory (Nolda 2012), traditionally used for the description of the spoken characteristics of word-formation and which it adapted to describe its written characteristics as well. This theory also enabled the author to present data and hypotheses which are computationally implementable.

Edwige Dugas (Université de Lille) and Jérôme Michaud (Uppsala Universitet) applied the theoretical framework of Construction Morphology to the study of specialized languages, with the patterns [super-X] and [s-X] of the vocabulary of supersymmetry in particle physics as a case study. They showed that it is possible to represent these two constructions in a “constructicon” (Jurafsky 1991) that is common to other similar constructions of the general language, in particular the [super-X] construction which forms lexemes with a meaning of ‘being above’ or with an evaluative meaning.

References

- Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Ellis, Nick & Teresa Cadierno. 2009. Constructing a Second Language. Introduction to the Special Section. *Annual Review of Cognitive Linguistics* 7. 111–139.
- Fleischer, Wolfgang & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*. Berlin & Boston: de Gruyter.
- Hein, Katrin. 2015. *Phrasenkomposita im Deutschen. Empirische Untersuchung und konstruktionsgrammatische Modellierung*. Tübingen: Narr.
- Jackendoff, Ray & Jenny Audring. in preparation. *The Texture of the Mental Lexicon*. Oxford: OUP.

- Jurafsky, D. 1991. *An on-line computational model of human sentence interpretation: A theory of the representation and use of linguistic knowledge*. Technical report, DTIC Document.
- Lehmann, Christian. 1982/1995. *Thoughts on Grammaticalization*. München: Lincom Europa.
- Nolda, Andreas. 2012. *Konversion im Deutschen – Muster und Beschränkungen: Mit einem Grundriss einer allgemeinen Theorie der Wortbildung*. Habilitation thesis, Humboldt-Universität zu Berlin.
- Stefanowitsch, Anatol. & Stefan Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.
- Trips, Carola. 2016. An analysis of phrasal compounds in the model of Parallel Architecture. In Pius ten Hacken (ed.), *The semantics of compounding*, 153–177. Cambridge: Cambridge University Press.

Edwige Dugas, Ph.D.
Université de Lille 3 – Bât. B4
Laboratoire STL ‘Savoirs, Textes, Langage’
Rue du Barreau – BP 60149
F-59653 Villeneuve d’Ascq Cedex
edwige.dugas@gmail.com

Tagungsankündigung / Conference Announcement

18th International Morphology Meeting May 10–13, 2018; Budapest, Hungary

Organizing committee: Ferenc Kiefer, András Márton Baló, Huba Bartos, Péter Rebrus

The 18th International Morphology Meeting is organized by the Research Institute for Linguistics, Hungarian Academy of Sciences, in collaboration with the Department of Theoretical Linguistics and the Department of English Linguistics of Eötvös Loránd University (ELTE). The meeting will be held in Budapest, Hungary, May 10–13, 2018.

Invited were papers on subjects including but not limited to the main theme of the conference:

Paradigms in inflection and word formation synchronically and diachronically

Registration begins: 1 March 2018

Invited speakers:

Olivier Bonami (Université Paris Diderot)

Farrell Ackerman (UCSD)

Marilyn Vihman (York University)

Workshops:

Workshop 1: Models and methods in morphology (James P. Blevins)

Workshop 2: The learnability of complex constructions from a cross-linguistic perspective (Marcel Schlechtweg)

Workshop 3: Morphological aspects of Uralic and Turkic languages (Darya Kavitskaya, Piibi-Kai Kivik, Öner Özçelik, Anne Tamm)

Contact: imm18@nytud.mta.hu

Tagungsankündigung / Conference Announcement

Word-Formation Theories III & Typology and Universals in Word-Formation IV June 27–30, 2018; Košice, Slovakia

Main organizers:

Livia Körtvélyessy and Pavol Štekauer, P.J. Šafárik University, Košice, Slovakia

Call for Papers:

Another edition of a twin conference concentrates on two areas of research: word-formation theories and word-formation typology/universals. Papers discussing any aspect of these general areas are most welcome.

Plenary speakers (in alphabetic order):

Balthasar Bickel, University of Zurich

Dirk Geeraerts, KU Leuven

Nicola Grandi, University of Bologna

Bernd Heine, University of Cologne

Susan Olsen, Humboldt University, Berlin

Frans Plank, Konstanz University

Special event:

A workshop 'Revisiting paradigms in word-formation' will be organized by Aleksandra Bagasheva and Jesus Fernandez Dominguez.

Deadlines:

Submission of abstracts: 31 March, 2018

Notification of acceptance: 10 April, 2018

Submission of a registration form: 30 April, 2018

Abstracts (max. 200 words) should be emailed to livia.kortvelyessy@upjs.sk

Tagungsankündigung / Conference Announcement

Morphosyntactic misfits: Clitics, particles, and non-canonical affixes in the languages of the Americas 23–25 July 2018; Pontificia Universidad Católica del Perú (PUCP), Lima

Workshop languages: English, Spanish and Portuguese

Keynote speaker: Balthasar Bickel

Convenors: Roberto Zariquiey (PUCP Lima) and Rik van Gijn (University of Zurich)

In the traditional, dichotomic division of coding systems in language between morphology and syntax, linguistic units are ideally either morphological or syntactic in nature. This perspective on linguistic organization has generated a vast literature in which linguistic units are classified as one or the other based on checklists (see e.g. Dixon & Aikhenvald 2002, Haspelmath 2011 for discussion). The main idea expressed in those checklists is that relations between units in a morphological construction are tight and rigid, whereas relations between units in a syntactic construction are loose and flexible.

It has been recognized by many scholars for a long time (e.g. Dixon 1977, Zwicky 1977) that the dichotomic approach is problematic in that there are many elements that do not fit neatly into one of the two categories. This is mainly due to the fact that rigidity and tightness are measured across different dimensions (Bickel 2007). A prototypical morphological element, for instance, is often described as having the following characteristics, that span a variety of dimensions.

- lexically restricted host selection
- phonological interaction with host (prosodic, morphophonological)
- morphological interaction with host
- rigid positioning with respect to its host
- non-manipulable by syntax

Because of this multi-dimensionality, even if we can define prototypical or idealized morphological and syntactic units, elements can deviate from these idealizations in many ways (see e.g. Anderson 2006, Spencer & Luís 2014, Van Gijn & Zúñiga 2016, Bickel & Zúñiga 2017). This has led to considerable terminological confusion, in which elements with similar behavior are

classified differently and elements with different behavior are classified in the same way. Testimony to the terminological confusion are, furthermore, the many different terms that have been proposed for these “morphosyntactic misfits”, which include simple clitics, special clitics, phrasal clitics, phrasal affixes, non-cohering affixes, Wackernagel clitics, Wackernagel affixes, clause-final particles, clause-initial particles, Wackernagel particles, etc.

For the Americanist descriptive tradition, with its many (poly)synthetic languages this problem is particularly relevant, and likely to frustrate fruitful morphological comparison across languages. At the same time, South American languages can be highly informative to shaping our ideas of the possible variation within this group of so-called morphosyntactic misfits (see e.g. Van Gijn & Zúñiga 2014). In this workshop, therefore, we call for papers that, rather than trying to classify elements, focus on highlighting the parameters of variation within a language or across languages.

Questions we would like to address include the following (although they are not restricted to this list):

- Is the dichotomic distinction between morphological and syntactic elements sufficient to describe the variation of morphosyntactic units of a language or language family?
- To what extent are notions such as “word”, “affix”, “clitic”, “particle” useful for comparative or descriptive purposes?
- What parameters are required to describe the variation among morphosyntactic units in a language or language family?
- How are these parameters distributed over different elements in a language or language families?
- How do properties of morphosyntactic units evolve diachronically?
- How do properties of morphosyntactic units interact with their borrowability?
- What inconsistencies are found in descriptions of morphosyntactic units across languages and how can they be reconciled?

Please send your abstract to misfitslima2018@gmail.com before **15 April 2018**. Notification will be given on 25 April 2018.

References

- Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Bickel, Balthasar. 2007. Typology in the 21st century: major current developments. *Linguistic Typology* 11. 239–251.
- Bickel, Balthasar & Fernando Zúñiga. 2017. The ‘word’ in polysynthetic languages: phonological and syntactic challenges. In Michael Fortescue, Marianne Mithun &

- Nicholas Evans (eds.), *The handbook of polysynthesis*, 158–185. Oxford: Oxford University Press.
- Dixon, R. M. W. 1977. Some phonological rules in Yidiny. *Linguistic Inquiry* 8(1), 1–34.
- Dixon, R. M. W. & Alexandra Y. Aikhenvald, (Eds.) 2002. *Word: a cross-linguistic typology*. Cambridge: Cambridge University Press.
- Gijn, Rik van & Fernando Zúñiga. 2014. Word and the americanist perspective. *Morphology* 24 (3). 135–160.
- Haspelmath, Martin 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1), 31–80.
- Spencer, Andrew & Ana Luís. 2012. The canonical clitic. In Dunstan Brown, Marina Chumakina & Greville Corbett (eds.), *Canonical morphology and syntax*, 123–150. Oxford: Oxford University Press.
- Zwicky, Arnold M. 1977. *On clitics*. Bloomington: IULC.

