

# Zeitschrift für Wortbildung / Journal of Word Formation

eine Open-Access-Zeitschrift /  
an Open Access Journal

1 / 2020

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**

**2020 • Jahrgang/Volume 4 • Heft/Issue 1**

## **Zeitschrift für Wortbildung Journal of Word Formation**

**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](mailto: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](mailto:ZJWF@germanistik.uni-siegen.de)

Homepage: [www.peterlang.com/zjwf](http://www.peterlang.com/zjwf)

**“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](mailto: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](mailto:ZJWF@germanistik.uni-siegen.de)

Homepage: [www.peterlang.com/zjwf](http://www.peterlang.com/zjwf)

ISSN 2367-3877

Peter Lang GmbH

Internationaler Verlag der Wissenschaften

Fehlerstraße 8, 12161 Berlin

<http://www.peterlang.com>

**PETER LANG**



This is an open access publication. This work is licensed under a Creative Commons Attribution CC-BY 4.0 license. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

# Inhalt / Contents

## Beiträge / Papers

*Susumu Kuroda*

Funktionsabspaltung des Suffixes *-ig* beim Verbbildungsmuster  
mit *-igen* .....7

*Barbara Schlücker*

Between word-formation and syntax: Cross-linguistic perspectives on  
an ungoing debate .....26

## Varia / Miscellaneous

*Petra Steiner*

Tagungsbericht / Conference Report: “DeriMo 2019. Second  
Workshop on Resources and Tools for Derivational Morphology”  
(September 19–20, 2019; Prague, Czech Republic) .....76

Tagungsankündigung / Conference Announcement: “Word-Formation  
Theories IV / Typology and Universals in Word-Formation V”  
(June 23–26, 2021; Kosice, Slovakia) .....81

## Beiträge / Papers

Susumu Kuroda

# Funktionsabspaltung des Suffixes *-ig* beim Verbbildungsmuster mit *-igen*\*

**Abstract:** This article deals with the development of *-igen* verbs in German since the Old High German period, demonstrating that this can be regarded as a process in which the adjective formation morpheme *-ig* gradually develops into a component of a word formation pattern that derives transitive verbs from nouns. An *-igen*-verb can be descended not only from an *-ig*-adjective (*würdig – würdigen*) but also from a noun without an intermediary *-ig*-adjective (*Pein – \*peinig – peinigen*). In this article, it is claimed that a word formation pattern with *-ig* develops over time. The emergence of this word formation pattern can be described as a “reanalysis” of the verb structure accompanied by a “resegmentation” of the original word structure and a semantic “remotivation” of the established unit. It is also pointed out that this development is particularly evident in the Middle High German period.

**Keywords:** suffix verb, verb derivation, Old High German, Middle High German, reanalysis

## 1. Einleitung

Im Deutschen sind viele Verben auf *-igen* zu beobachten, die auf ein Substantiv oder ein Adjektiv bezogen werden können:

- (1) *Pein – peinigen*  
*Schaden – schädigen*  
*rein – reinigen*  
*satt – sättigen*  
*niedrig – erniedrigen*  
*einig – vereinigen*

Diesen *-igen*-Verben<sup>1</sup> wird in der deutschen Wortbildungsforschung aufgrund ihrer zwar nicht großen, aber doch nicht unbedeutenden Anzahl sowie

---

\* This contribution was partly supported by JSPS KAKENHI Grant Number 18K00550. Herrn Prof. Dr. Hans-Werner Eroms und Herrn Prof. Dr. Rüdiger Harnisch sowie zwei anonymen Gutachtern bin ich für zahlreiche wertvolle Hinweise zu Dank verpflichtet.

1 *-en* (bzw. *-n*) bezieht sich auf das Flexiv (des Infinitivs) und ist insofern nicht als Bestandteil des Verbs zu verstehen. *-en* wird hier jedoch stellvertretend für alle Flexionsformen als Markierung der verbalen Eigenschaft gebraucht, um vor allem den kategorialen Unterschied zwischen dem Adjektivbildungsmorphem *-ig* und dem Verbalisierungsmorphem *-igen* hervorzuheben.

ihres auffälligen Aufbaus (s.u. 2.3) besondere Aufmerksamkeit geschenkt. Der vorliegende Beitrag beleuchtet die hinter den *-igen*-Verben stehenden Wortbildungsmuster und ordnet in einen sprachhistorischen Zusammenhang ein.

Der diachronische Forschungsansatz ist insofern interessant, als in letzter Zeit in der Wortbildungsforschung das Bewusstsein zunehmend in den Vordergrund rückt, dass sich im Laufe der Zeit der Aufbau eines Wortes und die Funktion seiner Bestandteile verändern können. Somit stellen Wörter nicht einfach eine Kette von Morphemen dar, die ihre ursprüngliche Eigenschaft beibehalten. So sind in der Forschung die Zusammenführung von Morphemen zu einem Morphem, die Spaltung eines Morphems zu mehreren eigenständigen Morphemen oder die Ausstattung eines Morphems mit einer neuen Funktion festgestellt worden (vgl. Lass 1990, Haspelmath 1994, Munske 2002, Harnisch 2004, Harnisch 2010, Nübling 2010, Simon 2010, Wischer 2010, Kempf/Hartmann 2018; s.a. 2.3). Dem letztgenannten Veränderungstyp wird die Verwendung des Adjektivbildungsmorphems *-ig* in der Verbbildung zugeordnet (vgl. Wilmanns 1899: 455ff.). Im Folgenden wird versucht, die Distribution und Funktion von *-igen*-Verben im Mittel- und Althochdeutschen zu erfassen. Der weitere Entwicklungsverlauf bis zum Gegenwartsdeutschen wird ebenfalls angesprochen.

Das Material, auf das sich dieser Beitrag stützt, ist zum Teil der Fachliteratur entnommen. Für die Erfassung des bislang wenig dokumentierten ahd. Sprachstands wurde eine Korpusuntersuchung vorgenommen (anhand des *Referenzkorpus Altdeutsch*<sup>2</sup>, 2018 März). Für das Mhd. wurden Daten aus dem *Referenzkorpus Mittelhochdeutsch*<sup>3</sup> erhoben, aber das Ergebnis stimmte weitestgehend mit den Angaben bei Leipold (2006) überein, bis auf wenige Stellen, die im von Leipold untersuchten Material nicht enthalten sind. Aus diesem Grund unterscheiden sich im Folgenden die Angaben zum Mhd. großteils nicht von denen bei Leipold. Anders als bei Leipold wurden jedoch keine sprachgeographischen und feineren diachronischen

- 
- 2 Donhauser, Karin/Gippert, Jost/Lühr, Rosemarie (2017). *Referenzkorpus Altdeutsch*. Nutzung über ANNIS (ANNotation of Information Structure). Version 1.0. <http://www.deutschdiachrondigital.de>. Ausgewertet wurden die im Korpus enthaltenen „klassischen“ ahd. Texte mit größerem Umfang: Der ahd. *Isidor*, Otrfrids *Evangelienbuch*, der ahd. *Tatian*, Notkers Übersetzungen (*Martianus Capella*, *Categoriae*, *De consolatione philosophiae*, *De interpretatione*, Psalmen) sowie kleinere Schriften (*Ars rhetorica*, *De musica*, *De partibus logicae*, *De syllogismis*).
  - 3 Klein, Thomas/Wegera, Klaus-Peter/Dipper, Stefanie/Wich-Reif, Claudia (2016). *Referenzkorpus Mittelhochdeutsch* (1050–1350). Version 1.0. <https://www.linguistics.ruhr-uni-bochum.de/rem/>



Unterschiede erfasst. Zu einigen Punkten, die bei Leipold nicht ausreichend erklärt sind, wurden ergänzende Zusatzerhebungen durchgeführt. Die Angaben zum Gegenwartsdeutschen basieren auf Muthmann (2001). Bei der Strukturanalyse der ahd. und gegenwartsdt. Lexeme wurde auch Splett (1993, 2009) herangezogen.

## 2. Stellung der *-igen*-Verben im Gegenwartsdeutschen

### 2.1 Zwei Typen von *-igen*-Verben

Die *-igen*-Verben stellen keine homogene Gruppe dar. Unter ihnen findet sich eine Reihe von Verben<sup>4</sup>, zu denen sich synchron ein formal und semantisch verwandtes Adjektiv auf *-ig* findet. Es stehen wie in (2) neben *würdig* oder *zeitig* jeweils *würdigen* oder *zeitigen*, sodass hier eine Verbbildung aus einem *-ig*-Adjektiv parallel zu den Fällen wie *schwarz* – *schwärzen* festgestellt werden kann. Auch präfigierte Formen wie *bewilligen* oder *ermäßigen* sind möglich.

- (2) *würdig* – *würdigen*  
*zeitig* – *zeitigen*  
*willig* – *bewilligen*  
*mäßig* – *ermäßigen*  
*kräftig* – *bekräftigen*

Anders verhält es sich jedoch mit den Fällen, wo einem *-igen*-Verb synchron kein entsprechendes *-ig*-Adjektiv als Derivationsbasis zugrunde gelegt werden kann. Wellmann (1973: 105ff.) grenzt solche *-igen*-Verben von denen wie in (2) ab.

- (3) \**schädig* – *schädigen*  
\**endig* – *endigen*  
\**sättig* – *sättigen*  
\**schädig* – *beschädigen*  
\**sänftig* – *besänftigen*  
\**lobig* – *belobigen*  
\**eidig* – *vereidigen*

---

4 Die 99 *-igen*-Verben im Gegenwartsdeutschen bei Muthmann (2001) können auf 59 semantisch und formal entsprechende Adjektive mit *-ig* bezogen werden. (Als Einzellexeme verwertbare Adjektive wurden getrennt gezählt, z.B. wird *veruneinigen* dem Adjektiv *uneinig* gegenübergestellt, nicht *einig*). 55 der bei Muthmann (2001) verzeichneten *-igen*-Verben kann kein entsprechendes *-ig*-Adjektiv gegenübergestellt werden.

Der Unterschied zwischen dem zuerst genannten („ersten“) wie in (2) und diesem („zweiten“) Typ der *-igen*-Verben wie in (3) ist abgesehen von der synchronen Absenz des Basisadjektivs nicht deutlich. Ein großer Teil der *-igen*-Verben ist sowohl beim ersten als auch beim zweiten Typ präfigiert (insbesondere mit *be-*; vgl. Wellmann 1973: 107). Die Verben sind ferner bei beiden Typen meist transitiv und gehören nach Wellmann (1973: 98, 105) zur Klasse der „objektprädikativen“ Verben wie *schwärzen*, *kürzen* oder *leeren* an, die eine Eigenschaftsveränderung des Objekts ausdrücken<sup>5</sup>. Die Adjektive auf *-ig*, die den *-igen*-Verben des ersten Typs zugrunde gelegt werden können, beziehen sich ihrerseits wie in (4) in vielen Fällen auf ein synchron existierendes, nicht auf bestimmte Wortarten eingeschränktes, Basislexem. Ein bedeutender Teil der *-igen*-Verben des zweiten Typs kann auf ein formal und semantisch verwandtes Lexem bezogen werden, wobei auch hier wie in (5) die Wortart dieses Lexems nicht eingeschränkt ist. Beim ersten Typ gibt es Fälle wie *billigen* und *fertigen*, wo das zugrunde liegende *-ig*-Adjektiv nicht weiter segmentiert werden kann<sup>6</sup>. Auch beim zweiten Typ gibt es Fälle, wo synchron eine opake Struktur vorliegt (wie *behelligen*, *beschwichtigen*).

- (4) *Würde* – *würdig*  
*Zeit* – *zeitig*  
*Wille* – *willig*  
*ein* – *einig*  
*heil/Heil* – *heilig*  
*ergeben* – *ergiebig*

- (5) [Basis: Substantiv]  
*Schaden* – (\**schädig*) – *schädigen*  
*Ende* – (\**endig*) – *endigen*  
 [Basis: Adjektiv]  
*satt* – (\**sättig*) – *sättigen*  
*schön* – (\**schönig*) – *beschönigen*  
 [Basis: Verb]  
*walten* – (\**wältig*) – *bewältigen*

5 Unter den *-igen*-Verben bei Muthmann (2001) finden sich einige wenige Intransitiva wie *endigen*, *nächtigen*, *sündigen*, *übernächtigen*, *willigen*. Wellmann (1973: z.B. 104) gibt einige *-igen*-Verben an, deren Basisadjektiv ein „Subjektsprädikativ“ ist. Dabei handelt es sich allerdings um Reflexiva (wie *sich beruhigen*, *sich befeiligen*, *sich entledigen*).

6 *-ig* bei lexikalisierten Adjektiven wie *billig*, *fertig*, *schleunig*, bei denen es nicht heraussegmentierbar ist, fungiert genau genommen nicht als Suffix, wird hier aber als ein Suffix behandelt, und die Adjektive, die es enthalten, werden den suffigierten zugezählt.

Die *-igen*-Verben des zweiten Typs können als Produkte eines eigenständigen Wortbildungsmusters<sup>7</sup> angesehen werden, das einem Lexem ein *-igen*-Verb unmittelbar gegenüberstellt. Das Suffix *-ig* ist hier also funktional nicht heraussegmentierbar (s.u. 2.3). Somit stellt sich nun die Frage nach dem Status von *-ig* beim zweiten Typ der *-igen*-Verben, das sich von dem beim ersten Typ unterscheidet (vgl. schon Grimm 1826: 306-307, Jacobi 1843: 176ff., Wilmanns 1899: 111, Paul 1920: 122).

## 2.2 Besondere Stellung von *-igen*-Verben

Die Adjektivierungsfunktion kommt im Gegenwartsdeutschen nicht nur *-ig*, sondern auch einer Reihe von Suffixen wie *-bar*, *-(er)isch*, *-(er)lich*, *-haft*, *-sam*, *-mäßig*<sup>8</sup> und *-los* zu (Dudengrammatik 2016: 765ff.). Die Verben, die als Derivation aus den mit diesen Suffixen gebildeten Adjektiven anzusehen sind, sind allerdings zahlenmäßig weit weniger stark vertreten als die *-igen*-Verben.

Unter den bei Muthmann (2001) verzeichneten Verben, die auf ein formal und semantisch verwandtes suffigiertes Adjektiv bezogen werden können, bilden die *-lichen*-Verben (wie *sachlich* – *versachlichen*, *anschaulich* – *veranschaulichen*, *staatlich* – *ver-*lent*staatlichen*, *heimlich* – *verheimlichen*, *möglich* – *ermöglichen*) eine relativ große Gruppe (41 Verben, vor allem mit *ver-* präfigiert). Anders als *-ig* bei den *-igen*-Verben behält *-lich* aber seine adjektivbildende Funktion bei.

## 2.3 „Reanalyse“ bei *-igen*-Verben

Dass sich die morphologische Struktur eines Wortes historisch verändern kann, ist wie eingangs erwähnt eines der aktuell intensiv diskutierten Themen in der Wortbildungsforschung. Dies ist möglicherweise eine Resonanz auf die Grammatikalisierungsforschung seit den 1990er Jahren. Vielfach werden nämlich Entwicklungswege von Wortstrukturen analog zur Grammatikalisierung syntaktischer Einheiten erklärt, die eine Bedeutungsabstrahierung und/oder eine Fusion mit anderen Einheiten erfahren, z.B. die Entstehung der Suffixe *-heit*, *-schaft* und *-tum* oder die Herausbildung der Präpositionen *zugunsten*, *infolge*, *anstatt* aus Präpositionalphrasen (Munske 2002). Kempf/Hartmann (2018) diskutieren aus der konstruktionsgrammatischen Perspektive diskontinuierliche Wortbildungsschemata (z.B. [PREF-X-ung]<sub>N</sub> bei *Bespaßung* < *Spaß* oder [un-V-lich]<sub>A</sub> bei *ungaublich*), die sich aus mehreren Wortbildungseinheiten entwickeln.

7 Zur Erfassung von Relationen zwischen funktional/formal nahestehenden Wortbildungsmustern vgl. Riehemann (1998).

8 *-mäßig* kann als eine selbständige Form angesehen werden und nicht als ein Subtyp von *-ig*.

Andererseits ist seit Harnisch (2004) das Bewusstsein gewachsen (was ebenfalls als Reaktion auf die Grammatikalisierungsforschung aufzufassen ist), dass der Sprache das Grundbedürfnis zuzuerkennen ist, ein Gleichgewicht zwischen Form und Inhalt eines sprachlichen Zeichens herzustellen (Harnisch 2010: 5–8), und ihr eine „degrammatikalisierende“ Tendenz zur Erhöhung der strukturellen Komplexität eines Wortes nicht abgesprochen werden darf. So entsteht z.B. beim mhd. Substantiv *kint* im Laufe der Entwicklung zum Gegenwartsdeutschen eine (markierte) eigenständige Pluralform *Kinder* („Substanziierung von Bedeutung“, Harnisch 2010: 5–8). Eine entgegengesetzte Entwicklung („Semantisierung formaler Substanz“, Harnisch 2010: 5–8) kann z.B. in der Remotivierung von *-er* als Pluralmarker beim Übergang von vor-ahd. *rinder* (Sg./Pl.) zu ahd. *rind* (Sg.) – *rinder* (Pl.) gefunden werden. Auch eine eigentlich untrennbare Teilsequenz eines Wortes kann als ein Morphem mit einer eigenständigen Funktion umgedeutet werden (z.B. die als „Sekretion“ bekannte Entwicklung des Auslauts *-n* des ae. Possessivpronomens *min* / *hin* zum Marker der nicht-attributiven Form, Jespersen 1922: 384–385, vgl. auch Haspelmath 1994). Die hier genannten Vorgänge können mit Harnisch (2010: 3) als „Prozesse sprachlicher Verstärkung“ bezeichnet werden, die dem Verschleifen der Form und der Bedeutung des sprachlichen Zeichens, also „sprachlicher Abschwächung“, entgegengesetzt sind.

Die Bildung der *-igen*-Verben des ersten Typs wurde oben als eine Verbalisierung eines *-ig*-Adjektivs wie *würdig* – *würdigen* erklärt. In den neueren korpusbasierten historischen Arbeiten (Prell/Schebben-Schmidt 1996: 24, Leipold 2006: 25, Klein/Solms/Wegera 2009: 494) wird diese Wortbildungsoperation in Anlehnung an die Auffassung von Erben (2006: 81) und Wellmann (1973: 20) über die analoge Operation im Gegenwartsdeutschen als Verbderivation mittels eines latent existierenden Suffixes („Nullsuffix“) angesehen. Die vorliegende Arbeit schließt sich diesem Standpunkt an. Die Annahme eines „Nullsuffixes“ bei der Verbderivation ist nicht unproblematisch, da dieses je nach dem Morphem, mit dem es in Analogie gestellt wird, unterschiedliche Funktionen aufweisen muss, sodass eine Annahme von mehreren Nullsuffixen notwendig ist (Eschenlohr 1999: 59, Bergenholtz/Mugdan 2000: 445–446). Statt einer Konversionsanalyse (d.h. einer Veränderung des syntaktischen Status des Lexems ohne Zuhilfenahme eines Derivationsmorphems) wird hier jedoch versucht, Verbbildung einheitlich als suffixal motiviert zu beschreiben.

So kann die Bildung des ersten Typs der *-igen*-Verben wie in (6) als Anfügung eines Verbalisierungselements an die Basis erklärt werden.

- (6)  $Sünde_s - [[sünd_s] -ig_A] - [[[sünd_s] -ig_A] -en]_V$   
 $stet_A - [[stet_A] -ig_A] - ver-[[[stet_A] -ig_A] -en]_V$

Beim zweiten Typ der *-igen*-Verben fungiert *-ig* nicht adjektivbildend und ist damit beim zweiten Typ der *-igen*-Verben nicht funktional segmentierbar. Da das Vorkommen von *-ig* beim zweiten Typ der *-igen*-Verben lexikalisch festgelegt ist<sup>9</sup>, ist es als ein Element anzusehen, das ein Wortbildungsmuster mitkonstituiert. (Diese Interpretation findet man bei Wilmanns 1899: 111 und Henzen 1965: 114; Wellmann 1973: 20 stellt *-ig* funktional neben *-l*, *-el*, *-(is)ier* bei z.B. *kränkeln*, *klügeln*, *legalisieren*, *legitimieren*.). Die Entstehung dieses Wortbildungsmusters lässt sich als „Reanalyse“ der Verbstruktur beschreiben (Harnisch 2010: 19, vgl. auch Hüning 2019: 479ff.), die mit einer „Resegmentierung“ der Wortstruktur (Umfunktionierung von *-ig* und Beseitigung der Morphemgrenze zwischen *-ig* und dem Verbalisierungselement) sowie einer semantischen „Remotivierung“ der neu entstandenen Einheit als Wortbildungsmuster einhergeht.

Für die *-igen*-Verben des zweiten Typs ergibt sich folglich die Derivationsanalyse wie in (7). Bei einer Annahme eines etablierten *-igen*-Musters wird allerdings die Mehrfachinterpretation<sup>10</sup> der Struktur von *-igen*-Verben nicht auszuschließen sein, da dann die Interpretation des ersten Typs der *-igen*-Verben als der zweite Typ möglich ist. In diesem Fall würden die Verben in (6) die Derivationsanalyse wie in (6') erhalten.

- (7)  $Ende_s - [[end_s]-igen]_V (*[[end_s]-ig]_A-en]_V)$   
 $rein_A - [[rein_A]-igen]_V (*[[rein_A]-ig]_A-en]_V)$   
 (6')  $Sünde_s - [[sünd_s]-igen]_V$   
 $stet_A - ver-[[stet_A]-igen]_V$ <sup>11</sup>

Aus der Perspektive der „sprachlichen Verstärkung“ stellt sich nun die Frage, ob die Entstehung des Wortbildungsmusters mit *-igen* als „Exaptation“ (Lass 1990, Harnisch 2010: 7, Nübling 2010, Wischer 2010, Simon 2010) bezeichnet werden kann. Diese stellt einen Prozess dar, bei dem ein Wortelement ohne Funktion („junk“) mit einer neuen Funktion beladen wird, was zu einer „Reanalyse“ der Wortstruktur führt (z.B. die Verschmelzung des Perfekts und des Aorists nach dem Zerfall der *ie*. Aspektopposition bei den germ. starken Verben, Lass 1990: 83–87). Die „Exaptation“ ist also ein Teilvorgang der „Reanalyse“, stellt aber die Umfunktionierung der

9 Bei einigen *-igen*-Verben kann eine Alternanz mit der entsprechenden *-en*-Form festgestellt werden, wie *befehlen* – *befehligen* (zum Mhd. vgl. Klein/Solms/Wegera 2009: 520). Ihre Zahl ist jedoch klein und eine semantische Regelmäßigkeit lässt sich schwer erkennen (vgl. Weisgerber 1958), sodass *-ig* nicht als ein freies Morphem anzusehen ist.

10 Zur mehrfachen Strukturinterpretationsmöglichkeit eines Bildungstyps, bei dem eine Strukturreeanalyse angenommen werden kann, vgl. auch Hüning (2019: 481).

11 Hierbei handelt es sich um eine Präfigierung eines *-igen*-Verbs.

Elemente in den Fokus, deren Funktion verloren oder marginalisiert ist (Lass 1990: 80, Wischer 2010: 35). So gilt die Entstehung des *-igen*-Musters nicht als „Exaptation“, solange man diesen Begriff anders als Simon (2010: 53) nicht auf die Fälle erweitert, in denen das Morphem eine neue zusätzliche Funktion erhält und gleichzeitig die alte Funktion weiterbesteht, da *-ig* durchaus als adjektivbildendes Morphem zu erkennen ist und deshalb kein „junk“ darstellt.

Unklar bleibt noch, wie die Funktionsabspaltung von *-ig* vonstattenging. Eine sprachhistorische Beleuchtung der Distribution von *-ig* würde Klarheit über seinen Entwicklungsgang und dessen Stellenwert als Wortbildungselement verschaffen.

### 3. *-igen*-Verben im Alt- und Mittelhochdeutschen

#### 3.1 Etablierung von *-ig* als Verbalisierungselement im Mittelhochdeutschen?

Durch eine Untersuchung des mhd. Korpus wurden 124 Verben<sup>12</sup> ermittelt, die als mhd. *-igen*-Verben angesehen werden können. In der bisherigen Forschung besteht ein breiter Konsens darüber, dass die Funktionsabspaltung von *-ig* (mhd. *-ec*) bereits zu mittelhochdeutscher Zeit vollzogen war und *-ig* den Status als eigenständiges Verbalisierungselement erlangt hatte (Wilmanns 1899: 111–112, Waldherr 1906: 76, van Zuiden 1934: 188, Henzen 1965: 225, Klein/Solms/Wegera 2009: 519ff.)<sup>13</sup>. Die beiden Typen, die oben in Bezug auf die *-igen*-Verben im Gegenwartsdeutschen festgestellt wurden, können auch im untersuchten mhd. Korpus beobachtet werden. Hier sind 103 *-igen*-Verben zu finden, für die es ein formal und semantisch verwandtes *-ig*-Adjektiv gibt. Diese Verben können dem ersten Typ der *-igen*-Verben zugerechnet werden. Dabei kann das Basisadjektiv mit *-ig* wie in (8) selbst auf ein Element verschiedener Wortarten zurückgehen. Es gibt

---

12 Wenn das Präfix *ge-* als Aspektmarker gelten kann (zum Überblick über die Funktion von *ge-* vgl. Klein/Solms/Wegera 2009: 480ff.), darf ein *ge-*Verb (sofern es die entsprechende Form ohne *ge-* gibt) nicht als ein eigenständiges Lexem angesehen werden. Da die semantische Funktion von *ge-* jedoch in vielen Fällen nicht eindeutig zu bestimmen ist, ist hier diese Frage dahingestellt, und die Lexemfeststellung an Lexer (1872–1878) angelehnt. Auch die Schreibung stützt sich grundsätzlich auf Lexer (1872–1878), wobei auf die Angabe der hier irrelevanten Formenvariante wie *reinegen* – *reinigen* verzichtet wird. Die Bedeutungsangabe mhd. Lexeme wird knapp gehalten. Das Gleiche gilt für die Angaben zu ahd. Lexemen.

13 Aus diesem Grund wird in diesem Artikel der gegenwartsdeutsche Stand nicht diskutiert. (Zum Stand im Gegenwartsdeutschen vgl. Wellmann 1973).

aber auch Fälle wie *lëdec* ‚ledig, frei‘ – *erlëdigen* ‚frei machen‘, wo *-ig* nicht heraussegmentierbar ist.

(8) [Basis: Substantiv]

*kraft* ‚Kraft‘ – *kreftic* ‚kräftig‘ – *ge-kreftigen* ‚stärken‘

[Basis: Adjektiv]

*reine* ‚rein‘ – *reinec* ‚rein‘ – *reinegen* ‚reinigen‘

[Basis: Verb]

*walten* ‚Gewalt haben, herrschen‘ – *waltec* ‚mächtig‘ – *entweltigen* ‚der Gewalt berauben‘)

Die restlichen *-igen*-Verben ohne entsprechendes *-ig*-Adjektiv sind in die Gruppe der *-igen*-Verben des zweiten Typs einzugliedern. Auch diesen Verben stehen formal und semantisch entsprechende Lexeme verschiedener Wortarten gegenüber. Die Zahl dieser Verben beläuft sich aber lediglich auf 21. Bei einer Erweiterung des Bezugsmaterials für die Basissuche auf Lexer (1872–1878) wird der Anteil dieser Verben noch kleiner, nur 14 Verben lassen sich als mhd. *-igen*-Verben des zweiten Typs einstufen (Leipold 2006: 222–223, Klein/Solms/Wegera 2009: 515)<sup>14</sup>.

(9) [Basis: Substantiv]

*kriuze* ‚Kreuz‘ – *kriuzigen* ‚kreuzigen‘

*zant* ‚Zahn‘ – *zer-zanegen* ‚mit den Zähnen zerreißen‘

[Basis: Adjektiv]

*milte* ‚freundlich, barmherzig‘ – *miltigen* ‚mildern‘

[Basis: Verb]

*zwîden* ‚willfahren, erhören‘ – *zwîdigen* ‚erhören‘<sup>15</sup>

Als ahd. *-igen*-Verben des ersten Typs anzusehen sind die Verben, deren Grundform auf *-ig* (ahd. *-ag/ig*) endet und denen eine formale und semantische

14 Da das Wortbildungsmuster für die *-igen*-Verben des zweiten Typs in der Gegenwartssprache nicht produktiv ist (Wellmann 1973: 107), stellt sich die Frage, ob diese Verben als Überrest aus dem Wortschatz einer Vorstufe anzusehen sind. Allerdings bestätigen die Wörterbuchangaben (Splett 1993, Schützeichel 2012) für die mhd. *-igen*-Verben des zweiten Typs die Existenz entsprechender *-ig*-Adjektive im Ahd. nicht, auch wenn dort deren jeweils mögliche Basislexeme bezeugt sind. Wenn wir uns den 55 *-igen*-Verben des zweiten Typs in der Gegenwartssprache zuwenden, stellen wir jedoch fest, dass im Mhd. die möglichen Adjektivbasen für etliche *-igen*-Verben des zweiten Typs bezeugt sind (nach Lexer 1872–1878 für 20 gegenwartsdeutsche *-igen*-Verben des zweiten Typs). Für eine eindeutige Aussage über die Produktivität des *-igen*-Musters ist wohl eine eingehende Beobachtung der frnhd. Verhältnisse notwendig.

15 Zu *zwîdigen* vgl. Leipold (2006: 130) und Klein/Solms/Wegera (2009: 519).

Verwandtschaft mit einem *-ig*-Adjektiv zuzuerkennen ist. Im untersuchten Korpus, das überlieferungsbedingt eingeschränkt ist, fanden sich insgesamt 25 *-igen*-Verben. Für 13 unter diesen ahd. *-igen*-Verben ist im untersuchten Korpus ein selbständiges, semantisch und formal verwandtes *-ig*-Adjektiv festzustellen, dessen Basis unterschiedliche Wortarten bildet.

- (10) [Basis: Substantiv]  
*sculd / sculda* ‚Schuld‘ – *sculdîg* ‚schuldîg‘ – *sculdîgôn* ‚beschuldigen‘  
 [Basis: Adjektiv]  
*sêr* ‚traurig‘<sup>16</sup> – *sêrag* ‚traurig‘ – *sêragôn* ‚traurig machen‘  
 [Basis: Verb]  
*beran* – ‚(hervor)bringen, (Frucht) tragen‘ – *birîg* ‚fruchtbar‘ – *birîgôn* ‚fruchtbar machen‘).

Es gibt zwar auch hier Fälle wie \**sâl* – *sâlig* ‚selig‘ – *sâligôn* ‚seligpreisen; beglücken‘<sup>17</sup>, wo die Basis für das *-ig*-Adjektiv nicht belegt ist, aber bei diesen Verben weist das *-ig*-Adjektiv mit einem funktional heraussegmentierbaren *-ig* eine transparente Struktur auf. Diese Verben können als der erste Typ der *-igen*-Verben gelten.

Für 12 ahd. *-igen*-Verben wie in (11) fehlt im untersuchten Korpus ein entsprechendes Adjektiv auf *-ig*. Diese Verben können als der zweite Typ der *-igen*-Verben im Ahd. in Betracht kommen. Es gibt aber auch Verben wie *libfestîgôn* ‚lebendig machen‘ oder *bimunîgôn* ‚beschwören‘, die nicht auf ein eigenständiges Lexem zu beziehen<sup>18</sup> und als Lexikalisierungsfälle anzusehen sind.

- (11) [Basis: Substantiv]  
*krûzi* ‚Kreuz‘ – *krûzigôn* ‚kreuzigen‘, *gimunt* ‚Gedenken‘ – *gimuntîgôn* ‚gedenken‘  
 [Basis: Adjektiv]  
*rehthaft* ‚gerechtfertigt‘ – *rehthaftîgôn* ‚rechtfertigen‘  
 [Basis: Verb]  
*tîlon* ‚auslöschen‘ – *tîligôn* ‚(ver)tilgen, vernichten‘

Durch eine Basissuche mit Blick auf einen breiteren ahd. Wortschatz nach Splett (1993) können drei weitere Verben als zum ersten Typ der ahd. *-igen*-Verben gehörig eingestuft werden, sodass *-ig* als ein nicht-adjektivbildendes

16 *sêrag* kann auch auf das Substantiv *sêr* ‚Schmerz‘ bezogen werden.

17 Hierbei handelt es sich allerdings um Verben, die auf eine gleiche Basis zurückgehen: *sâligôn* ‚seligpreisen; beglücken, erfreuen‘ und *gisâligôn* ‚beglücken‘.

18 *bimunîgôn* ‚beschwören‘ ist auch semantisch schwer auf die verbale Basis *manôn* ‚(er)mahnen, erinnern, auffordern‘ zu beziehen (vgl. Splett 1993: 597).



Element bei sieben *-igen*-Verben anzunehmen<sup>19</sup> ist. In der Literatur wird zwar *-ig* im Ahd. generell nicht als Verberivationsmittel angesehen (Wilmanns 1899: 111, van Zuiden 1934: 188, Henzen 1965: 225), insofern aber im ahd. Korpus der Anteil der *-igen*-Verben des zweiten Typs größer ist als im mhd. Korpus, kann man das Eintreten der Funktionsabspaltung von *-ig* im Ahd. nicht völlig ausschließen. Dies bedeutet allerdings auch nicht, dass diese Entwicklung schon fortgeschritten ist, da das Fehlen des entsprechenden Basislexems aus der eingeschränkten Überlieferung des Ahd. herrühren kann.

Das Vorliegen von *-ig* als eigenständiges Verbalisierungselement ist auch wegen der oben bei (6) und (6') erwähnten unklaren Grenze<sup>20</sup> zwischen dem ersten und dem zweiten Typ der *-igen*-Verben schwer zu ermitteln. Angesichts des gegenwartsdeutschen Stands, wo nach einer Überprüfung der Angaben von Muthmann (2001)<sup>21</sup> 99 *-igen*-Verben des ersten Typs 55 *-igen*-Verben des zweiten Typs gegenüberstehen, kann der zweite Typ der *-igen*-Verben im Mhd. noch als Randerscheinung angesehen werden.

### 3.2 Distribution von *-igen-* und *-lichen-*Verben

Die mhd. *-igen*-Verben weisen aber immerhin ein von ihren ahd. Entsprechungen abweichendes Verhalten auf. In 2.2 wurde darauf hingewiesen, dass im Gegenwartsdeutschen die *-lichen*-Verben die zweitgrößte Gruppe der Verben aus einem suffigierten Adjektiv nach den *-igen*-Verben ausmachen.

Die *-lichen*-Verben wurden im Gegensatz zu *-igen*-Verben in der bisherigen Forschung nicht besonders beachtet. Dies geht möglicherweise auf den Umstand zurück, dass für *-lichen*-Verben grundsätzlich eine entsprechende Adjektivbasis mit *-lich* (ahd. *-lih*, mhd. *-lich*) vorhanden ist, die wie *-ig* zur Derivation eines Adjektivs aus Lexemen verschiedener Wortarten dient (wie ahd. *jung* ‚jung, neu‘ – *junglih* ‚jugendlich‘, *wehsal* ‚Wechsel‘ – *wehsallih* ‚veränderlich‘, *zeman* ‚sich geziemen‘ – *zimilih* ‚geziemend, schicklich‘, *ûf* ‚hinauf, aufwärts‘ – *ûflih* ‚oben befindlich‘; mhd. *wêrt* ‚Welt, menschengeschlecht‘ – *wêrtlich* ‚zur Welt, zum Leben gehörend, weltlich‘, *êwic* ‚ewig‘ – *êwiclich*, *êweclich* ‚ewig‘, *zëmen* ‚ziemen, passen‘ – *zimelich*,

19 Das gleiche Ergebnis liegt auch nach einer Heranziehung von Schützeichel (2012) vor.

20 Z.B. kann das Verb *reinigen* ‚reinigen‘ nicht nur auf *reinec* ‚rein‘, sondern auch auf *reine* ‚rein/Reinheit‘ oder *reinen* ‚reinigen‘ bezogen werden, vgl. Leipold (2006: 29, 222–223) und Klein/Solms/Wegera (2009: 515).

21 Bei Wellmann (1973) fehlt die Angabe zur Frequenz des ersten Typs der *-igen*-Verben.

*zimlich* ‚schicklich‘), sodass von der Bildungsanalyse wie (12) ausgegangen werden kann.

(12) [Ahd.]

*anderlîh* ‚andersartig‘ – *anderlîhhôn* ‚verändern‘

*burlîh* ‚hoch, hervorragend‘ – *burlîhhôn* ‚aufrichten, preisen‘

*guollîh* ‚herrlich, berühmt‘ – *guollîhhôn* ‚rühmen, preisen‘

[Mhd.]

*guotlich* ‚gut, gütig‘ – *guotlichen* ‚sich rühmen, übermütig sein‘

*heimelich* ‚einheimisch‘ – *heimelichen* ‚heimisch machen‘

*micHELLich* ‚groß‘ – *micHELLichen* ‚magnificare‘

Neben den *-igen*-Verben des ersten Typs und den *-lichen*-Verben gibt es auch weitere Verben, die auf ein mit einem Suffix gebildetes Adjektiv zu beziehen sind. Sie sind unterschiedlich häufig festzustellen:

Tab. 1: *Zahl der ahd. auf ein suffigiertes Adjektiv beziehbaren Verben im ausgewerteten Material*

Suffix des Basisadj.	<i>-ag/ig</i>	<i>-lîh</i>	<i>-falt</i>	<i>-haft</i>	<i>-isc</i>	<i>-lôs</i>	<i>-sam</i>
<i>type</i> -Frequenz	16	7	4	2	1	1	2
<i>token</i> -Frequenz	50	23	7	4	1	1	5

Tab. 2: *Zahl der mhd. auf ein suffigiertes Adjektiv beziehbaren Verben im ausgewerteten Material*

Suffix des Basisadj.	<i>-ig</i>	<i>-lich</i>	<i>-haft</i>	<i>-valt</i>
<i>type</i> -Frequenz	124	13	2	4
<i>token</i> -Frequenz	1061	58	2	23

In den untersuchten Korpora machen wie im Gegenwartsdeutschen die *-igen*-Verben die am häufigsten anzutreffende Gruppe aus. Den zweiten Rang nehmen die *-lichen*-Verben ein. Sowohl im Ahd. als auch im Mhd. ist die Gebrauchsfrequenz der Verben, die aus anderen suffigierten Adjektiven deriviert sind, so eingeschränkt (besonders im Ahd.), dass sie in der folgenden Diskussion ausgeklammert werden.

Die suffigierten Adjektive sind in den untersuchten Korpora weit häufiger anzutreffen als die aus den suffigierten Adjektiven derivierten Verben.

Tab. 3: *Zahl der ahd. suffigierten Adjektive im ausgewerteten Material*

Suffix des Adj.	<i>-ag/ig</i>	<i>-lîh</i>	<i>-falt</i>	<i>-haft</i>	<i>-isc</i>	<i>-lôs</i>	<i>-sam</i>
<i>type</i> -Frequenz	264	170	8	29	33	23	22
<i>token</i> -Frequenz	1295	527	115	66	182	41	66

Tab. 4: Zahl der mhd. suffigierten Adjektive im ausgewerteten Material

Suffix des Adj.	<i>-ig</i>	<i>-lich</i>	<i>-valt</i>	<i>-haft</i>
<i>type</i> -Frequenz	599	679	18	86
<i>token</i> -Frequenz	13577	7650	278	556

Auffällig ist, dass im Mhd. die *type*-Frequenz der *-lich*-Adjektive höher ist als die der *-ig*-Adjektive (vgl. auch Klein/Wegera/Solms 2009: 273). Die *token*-Frequenz von *-ig*-Adjektiven im mhd. Korpus ist zwar beträchtlich höher als die der *-lichen*-Adjektive, was allerdings auf den Gebrauch der Adjektive *heilec* ‚heilig‘ und *êwic* ‚ewig‘ zurückzuführen ist, die wegen ihrer äußerst hohen Vorkommenshäufigkeit als Ausnahmefälle gelten können. Andere *-ig*-Adjektive sind nicht dermaßen häufig<sup>22</sup>, sodass man den *-lich*-Adjektiven einen höheren Stellenwert im mhd. Wortschatz zuerkennen kann als den *-ig*-Adjektiven. Das oben beobachtete weit häufigere Vorkommen von *-igen*- als das von *-lichen*-Bildungen im Bereich der Verbderivation stellt also einen deutlichen Kontrast dar.

Hier unterscheiden sich das Ahd. und das Mhd. deutlich voneinander, da im Ahd. *-ig*-Adjektiv und *-igen*-Verben sowohl in der *type*- als auch in der *token*-Frequenz an erster Stelle stehen. Die distributionelle Independenz von *-igen*-Verben und *-ig*-Adjektiven im Mhd. deutet darauf hin, dass die Bindung von *-ig* an die ursprüngliche adjektivbildende Funktion im Laufe der Entwicklung vom Ahd. zum Mhd. merklich schwächer geworden ist.

### 3.3 Morphologische Struktur des Basisadjektivs und Gliederung der Verbbildungssystems

Es ist vorerst unklar, was zur in 3.2 beobachteten besonderen Entwicklung von *-ig* führte. Das Ausbleiben des Ausbaus der Verbbildung aus anderen suffigierten Adjektiven ist aber wenigstens auf einen Wandel im Verbalsystem zurückzuführen.

Der Verbbildung mit *-ig* ist wie in 2.1 erwähnt eine Affinität mit der transitiven Valenzstruktur zuzuerkennen. Die *-igen*-Verben, die sich im untersuchten ahd. und mhd. Korpus finden, sind durchweg transitiv: *heilag*

22 Die *token*-Frequenz der fünf häufigsten Adjektive: *-ig*-Adjektive: 1. *heilec* ‚heilig‘ (4429 *tokens*), 2. *êwic* ‚ewig‘ (1352 *tokens*), 3. *sælec* ‚selig‘ (693 *tokens*), 4. *schuldec* ‚schuldig‘ (461 *tokens*), 5. *almehtec* ‚allmächtig‘ (365 *tokens*); *-lich*-Adjektive: 1. *geistlich* ‚geistlich, fromm‘ (614 *tokens*), 2. *gotelich* ‚von gott ausgehend, sich auf gott beziehend, göttlich‘ (459 *tokens*), 3. *wërltlich* ‚zur Welt, zum Leben gehörend‘ (365 *tokens*), 4. *wunderlich* ‚wunderbar, seltsam‘ (205 *tokens*), 5. *menschlich* ‚menschlich‘ (204 *tokens*).

,heilig, geweiht‘ – *heilagôn* ‚heiligen, heilig machen‘, *giheilagôn* ‚heiligen, heilig machen‘, *bluotag* ‚blutig‘ – *bluotagôn* ‚blutig färben‘.

Im Ahd. korreliert die semantisch-syntaktische Eigenschaft eines schwachen Verbs mit seiner Konjugationsklasse (*-jan*, *-ôn* oder *-ên*). Dies ist bei den deadjektivischen Bildungen<sup>23</sup> wie in (13) besonders deutlich. Hier treten die transitiven *-jan*-Verben durch ihre faktitive Bedeutung (Wilmanns 1899: 49, 54ff., Krahe/Meid 1967: 247, Riecke 1996: 421ff.<sup>24</sup>) und die intransitiven *-ên*-Verben durch ihre inchoative bzw. durative Bedeutung (Wilmanns 1899: 49, 70, Krahe/Meid 1967: 250) hervor, während die *-ôn*-Verben sowohl transitiv als auch intransitiv sein können und semantisch unterbestimmt zu sein scheinen (Wilmanns 1899: 49, 66ff., Schaefer 1984: 357–358)<sup>25</sup>. Im untersuchten ahd. Korpus findet sich aber trotz der oben erwähnten Nähe zur Transitivität interessanterweise kein *-jan*-Verb, das aus einem *-ig*- Adjektiv abgeleitet ist.

(13) [*-jan*-Verben]

*alt* ‚alt‘ – *elten* ‚verlängern‘

*irri* ‚ruhelos, verirrt‘ – *irren* ‚verwirren, in Verwirrung bringen‘

[*-ôn*-Verben]

*niuwi* ‚neu, jung‘ – *niuwôn* ‚erneuern, wiederherstellen‘

*irri* ‚ruhelos, verirrt‘ – *irrôn* ‚(sich) irren, verwirrt sein‘

[*-ên*-Verben]

*arm* ‚gering, arm, schwach‘ – *armên* ‚arm werden, verarmen‘

*irri* ‚ruhelos, verirrt‘ – *irrên* ‚umherirren, in die Irre gehen‘

Nach der ahd. Zeit ist ein Übergang von dieser ternären Gliederung der deadjektivischen Verben zu einer bipolaren wie in (14) festzustellen, die einerseits aus meist vom Umlaut des Basisvokals<sup>26</sup> begleiteten transitiven Verben und andererseits aus umlautlosen intransitiven Verben besteht (Wilmanns 1899: 78–79, Henzen 1965: 219, Wellmann 1973: 24, Erben 2006: 82). Zu den beiden Verbgruppen passen die syntaktisch-semantischen Charakterisierungen für *-jan*-Verben (mit Umlaut) bzw. *-ên*-Verben (ohne Umlaut). Die Entwicklung zur morphologisch unterlegten dichotomischen

23 Das Bedeutungsverhältnis der Verbklassen im desubstantivischen Bereich ist wesentlich komplizierter. Vgl. Leipold (2006), Klein/Solms/Wegera (2009: 494ff.).

24 Nach Riecke (1996: 422) finden sich auch durativ-rationale *-jan*-Verben („Essiva“, vgl. auch Krahe/Meid 1967: 248), die aber nicht deutlich hervortreten.

25 Es lassen sich vor allem faktitive deadjektivische *-ôn*-Verben finden (z.B. *niuwôn* ‚erneuern, wiederherstellen‘, *ebanôn* ‚ebnen, glätten‘, *reinôn* ‚reinigen, säubern‘), sodass keine klare Abgrenzung von *-jan*-Verben möglich ist (Wilmanns 1899: 66–67, Henzen 1965: 215–216).

26 Nach Wellmann (1973: 24) ist der Basisumlaut bei 80% der transitiven deadjektivischen transitiven Verben im Gegenwartsdeutschen festzustellen. Vgl. auch Henzen (1965: 219).

Gliederung der deadjektivischen Verben lässt sich deshalb als Profilierung der *-jan-* und *-ên-*Verben zu Ungunsten von *-ôn-*Verben auffassen<sup>27</sup>.

(14) [transitiv]

*tot – töten*

*los – lösen*

*scharf – schärfen*

*schwach – schwächen*

*schwanger – schwängern*

*glatt – glätten*

*ganz – ergänzen*

[intransitiv]

*alt – veralten*

*faul – (ver)faulen*

*blass – (er)blassen*

*dunkel – dunkeln*

*gesund – gesunden*

*nah – nahen*

Interessanterweise gehören 93,9 %<sup>28</sup> der aus einem suffigierten Adjektiv gebildeten Verben (inkl. *-igen-* des ersten Typs und *-lichen-*Verben) im ahd. Korpus (wie in (15)) der Klasse der *-ôn-*Verben an. Dies steht im Gegensatz zu den aus nicht-suffigierten Adjektiven abgeleiteten Verben, die überwiegend (zu 58,6 %) der Gruppe der transitiven *-jan-*Verben angehören:

(15) *drifalt* ‚dreifaltig, -fach‘ – *drifaltôn* ‚verdreifachen‘

*frônisc* ‚herrlich, (vor)trefflich‘ – *gifrôniscôn* ‚schmücken‘

*burlîh* ‚hoch, hervorragend‘ – *burlîhhôn* ‚aufrichten, preisen‘

*ginuhtsam* ‚hinlänglich, geeignet‘ – *ginuhtsamôn* ‚genügen, ausreichen‘

---

27 Nach Wilmanns (1899: 75ff.) wurden die *-ôn-*Verben aufgrund ihrer syntaktisch-semanticen Unterspezifikation die Hauptkategorie der deadjektivischen Verben: „... aber die Thatsache, dass die Form der zweiten Conjugation überwiegt und das beliebteste Mittel wird, andere Wortarten zu verbalisieren, unterliegt keinem Zweifel; ... Schon ehe der Verfall der Endungen die Unterschiede zwischen den schwachen Conjugationen aufhob, hatte also eine von ihnen das entscheidende Übergewicht gewonnen und zwar die, deren Endungen am volltönendsten, deren Bedeutung von Anfang an am unbestimmtesten war. [d.i. 2. Klasse, der Verfasser des Artikels]“ Wilmanns (1899: 75) begründet dies damit, dass der Vielfalt der Basen eine gewisse Bedeutung zuzuerkennen ist: „... man kann es als Regel ansehen, dass ein Suffix um so lebenskräftiger ist, je mannigfacher und umfangreicher die Stämme sind, mit denen es sich verbindet.“

28 Unter insgesamt 33 Verben sind ein *-jan-*Verb (*managfalten* ‚vermehrten, ausbreiten‘) und ein *-ên-*Verb (*horawîgên* ‚verunreinigen, beschmutzen‘) festzustellen.

Diese Distribution, die in der Forschung lange bekannt ist (vgl. bereits Jacobi 1843: 169ff., Wilmanns 1899: 66, Henzen 1965: 223), zeigt, dass die Verbalisierung eines suffigierten Adjektivs grundsätzlich in den Funktionsbereich der *-ôn*-Verben fällt<sup>29</sup>. Der Anteil, den die Verben aus suffigierten Adjektiven (außer *-ig*-Adjektiven) im Gesamtbestand der deadjektivischen Verben einnehmen, bleibt wie in 2.2 gezeigt bis heute recht begrenzt. Es liegt nahe, dass die Abschwächung der *-ôn*-Verben zum Eindämmen der Bildung dieses Verbtyps führt.

Die *-ig*-Adjektive haben jedoch im Mhd. ihren Anwendungsbereich als Basis für Verbbildung massiv ausgebaut. Dies kann als ein weiteres Zeichen dafür angesehen werden, dass sich *-ig* von seiner Adjektivierungsfunktion entfernt hat. Im nächsten Abschnitt wird abschließend diese Entwicklung in einen größeren sprachhistorischen Zusammenhang gestellt.

#### 4. Verbalisierungselement *-ig* als Produkt der Reanalyse

Aus den Beobachtungen im vorigen Abschnitt kann man schließen, dass sich ein neuer Funktionsbereich von *-ig* im Laufe der Entwicklung vom Ahd. zum Mhd. allmählich entfaltet hat. Diese Entwicklung setzt sich im Fnhd. weiter fort, wo der zweite Typ der *-igen*-Verben einen Status erreicht hat, bei dem man mit Prell/Schebben-Schmidt (1996: 391) *-ig* als eine „ausdrucksseitige Signalisierung eines Objektsbezuges“ ansehen kann. Hier kann man von einer etablierten Funktionsabspaltung von *-ig* sprechen (vgl. auch Waldherr 1906: 77–78). Der vorausgehende Verlust der Unterscheidung der *-jan*-, *-ôn*- und *-ên*-Verben (Wilmanns 1899: 78, Henzen 1965: 219) verwischte den Unterschied zwischen den Konjugationsklassen, womit gleichzeitig auch die Basisbeschränkungen dreier Verbklassen nicht mehr aufrechtzuerhalten waren. Vor dem Hintergrund der Diskussion über die „sprachliche Verstärkung“, die gegen den Verschleiß sprachlicher Ausdrücke wirkt, ist die Annahme nicht abwegig, dass dieser Merkmalabbau eine Kompensation der dadurch verloren gegangenen Verbalkategorien nötig machte. Dass *-ig* durch sein häufiges Auftreten bei deadjektivischen transitiven Verben mit der Transitivität assoziiert und als Ersatz für die *-jan*-Form etabliert wurde, ist denkbar.

Die Verbreitung der *-igen*-Verben des zweiten Typs ist aber begrenzt. Im Bereich der deadjektivischen Verben im Gegenwartsteutschen bildet der zweite Typ der *-igen*-Verben eine funktional konturierte, aber nur kleine Gruppe. Ein markantes Mittel für die gegenwartsdeutsche Verbbildung ist die Präfigierung, vor allem mit *ver*-. Auch zahlreiche *-igen*-Verben des zweiten

---

29 Dass die Länge und die morphologische Struktur eines Lexems mit seiner Verträglichkeit mit einem Wortbildungsmuster korrelieren, zeigen Krott/Schreuder/Baayen (1999) und Rainer (2000: 881–882, 2005: 345–347).

Typs sind nur in Verbindung mit einer Präfigierung festzustellen (Wilmanns 1899: 112, Waldherr 1906: 77, Henzen 1965: 234ff., Wellmann 1973)<sup>30</sup>.

Die Entwicklung von *-ig* als Verbbildungsmorphem lässt sich als eine schrittweise Abspaltung einer neuen Gebrauchsweise von der ursprünglichen, nämlich der Adjektivbildung, nachzeichnen. Die adjektivbildende Funktion von *-ig* ist keineswegs verlorengegangen, aber im Laufe der Zeit treten Fälle allmählich hervor, wo *-ig* einer Charakterisierung als Adjektivbildungsmorphem nicht gerecht wird. Die Verwendung eines sprachlichen Zeichens in einem sich vom ursprünglichen unterscheidenden kategorialen Gebrauchsfeld ist generell ohne Loslösung von der ursprünglichen Bedeutung schwer vorstellbar. Diese Problematik soll aber im Kontext der „sprachlichen Verstärkung“ in einem noch größeren Rahmen diskutiert werden.

## Literatur

- Bergenholtz, Henning & Joachim Mugdan. 2000. Nullelemente in der Morphologie. In Geert Booij, Geert, Christian Lehmann & Joachim Mugdan (Hrsgg.), *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*, 435–450. 1. Halbband. Berlin & New York: de Gruyter.
- Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Duden. 2016. *Die Grammatik*. 9., vollständig überarbeitet und aktualisierte Aufl. Hg. von Wöllstein, Angelika. Berlin: Dudenverlag.
- Eschenlohr, Stefanie. 1999. *Vom Nomen zum Verb: Konversion, Präfigierung und Rückbildung im Deutschen*. Hildesheim, Zürich & New York: Olms.
- Grimm, Jacob. 1826. *Deutsche Grammatik*. Zweiter Theil. Göttingen: Dieterich.
- Harnisch, Rüdiger. 2004. Verstärkungsprozesse. Zu einer Theorie der ‚Sekretion‘ und des ‚Re-konstruktionellen Ikonismus‘. *Zeitschrift für Germanistische Linguistik* 32. 210–232.
- Harnisch, Rüdiger. 2010. Zu einer Typologie sprachlicher Verstärkungsprozesse. In Rüdiger Harnisch (Hrsg.), *Prozesse sprachlicher Verstärkung: Typen formaler Resegmentierung und semantischer Remotivierung*, 3–23. Berlin & New York: de Gruyter.
- Haspelmath, Martin. 1994. The Growth of Affixes in Morphological Reanalysis. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1994*, 1–29. Dordrecht, Boston & London: Kluwer.
- Henzen, Walter. 1965. *Deutsche Wortbildung*. 3., durchges. u. erg. Aufl. Tübingen: Niemeyer.

---

30 Die Präfigierung unterscheidet sich insofern vom Wortbildungsmuster mit *-igen*, dass sich ihr Aufkommen seit dem Ahd. nicht auf den transitiven Bereich beschränkt (Kuroda 2014). Einzelne Präfixe haben unterschiedliche satzstruktursteuernde Funktionen. Interessant ist in diesem Zusammenhang auch die hohe Affinität der auf der Komparativform basierenden Verbbildungen wie *schmal* – *schmälern* oder *lang* – *verlängern* zur transitiven Struktur.

- Hüning, Matthias. 2019. Morphological Theory and Diachronic Change. In Jenny Audring & Francesca Masini (eds.), *The Oxford Handbook of Morphological Theory*, 476–492. Oxford: Oxford University Press.
- Jacobi, Theodor. 1843. Die Bedeutung der schwachen Conjugationen. In Theodor Jacobi, *Beiträge zur Deutschen Grammatik*, 129–196. Berlin: Verlag von T. Trautwein.
- Jespersen, Otto. 1922. *Language: Its Nature, Development and Origin*. London: George Allen & Unwin.
- Kempf, Luise & Stefan Hartmann. 2018. Schema Unification and Morphological Productivity: A Diachronic Perspective. In Geert Booij (ed.): *The Construction of Words*, 441–474. Cham: Springer International.
- Klein, Thomas, Hans-Joachim Solms & Klaus-Peter Wegera Hrsgg. 2009. *Mittelhochdeutsche Grammatik*. Teil III. Wortbildung. Tübingen: Niemeyer.
- Krahe, Hans & Wolfgang Meid. 1967. *Germanische Sprachwissenschaft: III Wortbildungslehre*. Berlin: de Gruyter.
- Krott, Andrea, Robert Schreuder & Rolf Harald Baayen. 1999. Complex words in complex words. *Linguistics* 37. 905–926.
- Kuroda, Susumu. 2014. *Die syntaktische Funktion der Präfigierung im Althochdeutschen*. Heidelberg: Winter.
- Lass, Roger. 1990. How to do things with junk: Exaptation in language evolution. *Journal of Linguistics* 26. 79–102.
- Leipold, Aletta. 2006. *Verbableitung im Mittelhochdeutschen*. Tübingen: Niemeyer.
- Lexer, Matthias. 1872–1878. *Mittelhochdeutsches Handwörterbuch*. Leipzig: Hirzel.
- Munske, Horst Haider. 2002. Wortbildungswandel. In Mechthild Habermann, Peter O. Müller & Horst Haider Munske (Hrsgg.), *Historische Wortbildung des Deutschen*, 23–40. Tübingen: Niemeyer.
- Muthmann, Gustav. 2001. *Rückläufiges deutsches Wörterbuch*. Handbuch der Wortausgänge im Deutschen, mit Beachtung der Wort- und Lautstruktur. 3., überarbeitete und erweiterte Auflage. Tübingen: Niemeyer.
- Nübling, Damaris. 2010. Von *Schreiner* zu *Schreinert* oder: Auf dem Wege zu einem onymischen Suffix? *Der -ert*-Ausgang als Ergebnis eines onymischen Verstärkungsprozesses. In Rüdiger Harnisch (Hrsg.), *Prozesse sprachlicher Verstärkung: Typen formaler Resegmentierung und semantischer Remotivierung*, 129–155. Berlin & New York: de Gruyter.
- Paul, Hermann. 1920. *Deutsche Grammatik*. Band V. Teil V: Wortbildungslehre. Halle: Niemeyer.
- Prell, Heinz-Peter & Marietheres Schebben-Schmidt. 1996. *Die Verbableitung im Frühneuhochdeutschen*. Berlin & New York: de Gruyter.
- Rainer, Franz. 2000. Produktivitätsbeschränkungen. In Geert Booij, Christian, Lehmann & Joachim Mugdan (Hrsg.), *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*, 877–885. 1. Halbband. Berlin & New York: de Gruyter.
- Rainer, Franz. 2005. Constraints on Productivity. In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of Word-Formation*, 335–352. Dordrecht: Springer.



- Rainer, Franz. 2015. Mechanisms and motives of change in word-formation. In Peter O. Müller, Ingeborg Ohnheise, Susan Olsen & Franz Rainer (eds.), *Word-Formation: An International Handbook of the Languages of Europe, 1761–1781*. Berlin & Boston: de Gruyter.
- Riecke, Jörg. 1996. *Die schwachen jan-Verben des Althochdeutschen: Ein Gliederungsversuch*. Göttingen: Vandenhoeck & Ruprecht.
- Riehemann, Susanne. 1998. Type-based Derivational Morphology. *Journal of Comparative Germanic Linguistics* 2. 49–77.
- Schaefer, Christiane. 1984. Zur semantischen Klassifizierung germanischer denominaler *ōn*-Verben. *Sprachwissenschaft* 9. 356–383.
- Schützeichel, Rudolf. 2012. *Althochdeutsches Wörterbuch*. 7., durchges. u. verb. Aufl. Berlin & Boston: de Gruyter.
- Simon, Horst J. 2010. ‚Exaptation‘ in der Sprachwandeltheorie. Eine Begriffspräzisierung. In Rüdiger Harnisch (Hrsg.), *Prozesse sprachlicher Verstärkung: Typen formaler Resegmentierung und semantischer Remotivierung*, 41–57. Berlin & New York: de Gruyter.
- Splett, Jochen. 1993. *Althochdeutsches Wörterbuch*. Berlin & New York: de Gruyter.
- Splett, Jochen. 2009. *Deutsches Wortfamilienwörterbuch*. Berlin & New York: de Gruyter.
- Waldherr, Friedrich. 1906. *Die durch Ableitungssuffixe gebildeten Verba der schwachen Konjugation im 16. Jahrhundert*. Darmstadt: Ottos Hof-Buchdruckerei.
- Weisgerber, Leo. 1958. Die heutige Leistung der Verben auf *-igen*. *Verschiebungen in der sprachlichen Einschätzung von Menschen und Sachen*. Köln & Opladen: Westdeutscher Verlag.
- Wellmann, Hans. 1973. Verbbildung durch Suffixe. In Ingeburg Kühnhold & Hans Wellmann (Hrsgg.), *Deutsche Wortbildung*, 17–140. 1. Hauptteil. Düsseldorf: Pädagogischer Verlag Schwann.
- Wilmanns, Wilhelm. 1899. *Deutsche Grammatik*. 2. Abteilung; Wortbildung. 2. Aufl. Straßburg: Trübner.
- Wischer, Ilse. 2010. Sekretion und Exaptation als Mechanismen in der Wortbildung und Grammatik. In Rüdiger Harnisch (Hrsg.), *Prozesse sprachlicher Verstärkung: Typen formaler Resegmentierung und semantischer Remotivierung*, 29–40. Berlin & New York: de Gruyter.
- van Zuiden, Jozef. 1934. *Die Verba auf -igen im Deutschen*. Maastricht: Boosten & Stols.

Susumu Kuroda  
 Musashi University  
 Faculty of Humanities  
 Toyotama-kami 1-26-1, Nerima-ku  
 176-8534 Tokyo  
 kuroda.susumu@cc.musashi.ac.jp

Barbara Schlücker

## Between word-formation and syntax: Cross-linguistic perspectives on an ongoing debate<sup>1</sup>

**Abstract:** The relation between word-formation and syntax and whether they form distinct domains of grammar or not has been discussed controversially in different theoretical frameworks. The answer to this question is closely connected to the languages under discussion, among other things, because languages seem to differ considerably in this regard. The discussion in this paper focuses on nominal compounds and phrases. On the basis of a great variety of data from a total of 14 European languages, it is argued that the relation between compounds and phrases, and, more generally, between word formation and syntax, should be characterized not in terms of a categorical but instead in terms of a gradient distinction.

**Keywords:** compound, compound-phrase distinction, Construction Grammar, lexicon, lexical phrase

The relation between word-formation and syntax and whether they form distinct domains of grammar or not has been discussed controversially in different theoretical frameworks for a long time. The answer to this question is closely connected to the languages under discussion, among other things, because languages seem to differ considerably in this regard. Apparently, in some languages word-formation entities can be clearly distinguished from syntactic ones on formal grounds. In other languages this is much more difficult. The discussion in this paper focuses on nominal compounds and phrases. German is considered a prime example of a clear compound-phrase distinction and therefore often regarded as evidence for the idea of a categorical distinction between word-formation and syntax. Meanwhile in English the compound-phrase distinction is known to be highly problematic. English, for this reason, has often been regarded as counterevidence to a categorical distinction.

The paper brings together a great variety of data from a total of 14 European languages, many of which taken from language-specific studies, and evaluates them with regard to the central question. This is followed by a discussion of divergent theoretical positions. It is argued that, firstly, an unclear situation as in English is by no means the exception since in

---

1 I am grateful to three anonymous reviewers for their detailed and helpful comments. The usual disclaimers apply.

various languages a clear compound-phrase distinction cannot be drawn. Secondly, even in languages such as German the situation appears differently if one also takes into consideration lexical entities with an internal phrasal structure (analysed here as ‘syntactic compounds’). It is shown that these phrases, in addition to phrasal properties, also have morphological properties and that they can be regarded as formations “in between” compounds and phrases. Thirdly, it will be concluded that the relation between compounds and phrases, and, more generally, between word-formation and syntax, should be characterized not in terms of a categorical but rather in terms of a gradient distinction. Morphological and syntactic lexical patterns can be conceived of as a series of related, but distinct constructions. They form part of a continuum with fully morphological and fully syntactic constructions at the endpoints and various mixed patterns combining morphological and syntactic properties in between. This is in line with non-modular approaches to linguistic theory, in particular constructionist frameworks.

## 1. Introduction: The compound-phrase distinction in German and English

This paper discusses the relation between word-formation and syntax on the basis of complex lexical entities of various kinds, thereby focussing on nominal units. More specifically, it compares compounds, which are usually considered morphological objects, and phrases, thus syntactic entities, cf. (1). In addition, it also takes into consideration various kinds of phrasal lexical entities, analysed here as syntactic compounds (see Section 3 and 4), cf. (2):

### GERMAN

- |     |                       |                  |             |                    |
|-----|-----------------------|------------------|-------------|--------------------|
| (1) | a. <i>Rotwein</i>     | (lit. red wine)  | ‘red wine’  | COMPOUND           |
|     | b. <i>roter Stuhl</i> | (lit. red chair) | ‘red chair’ | PHRASE             |
| (2) | <i>rote Bete</i>      | (lit. red beet)  | ‘beetroot’  | SYNTACTIC COMPOUND |

Generally, compounds are defined by morphological (or more specifically word-formation) rules and are for this reason clearly opposed to phrases which are built according to syntactic principles. German is a typical representative of this opposition. However, cross-linguistic evidence shows that this distinction cannot always clearly be drawn. As shown in the following, in various languages entities that are considered compounds do not only have morphological but also syntactic properties, such as inflection. On the other hand, phrasal entities as in (2) also have word-formation properties.

Also, both compounds and such phrasal entities occur systematically and arise from productive patterns. These observations call for the conclusion that there is no categorial separation between word-formation and syntax. The relation between compounds and phrases, and, more generally, between word-formation and syntax, can be described more adequately in terms of gradience than by opposition.

Consider the following two examples from English:

- (3) a. *red wine*  
b. *apple tree*

Forms like those in (3) are usually considered compounds. Compounds are generally defined as complex lexical entities that consist of two (or more) stems that are combined in a morphological process. They usually function as conventional names and express a unitary meaning. Also, the semantic relation between the constituents is implicit, contrary to phrases which use inflectional markers, determiners or prepositions to explicitly express this relationship.

Although this definition of compound is probably very similar to many definitions in the literature, it is far from clear or uncontroversial (and in fact every part of it has been subject to discussion). This becomes particularly obvious if English is compared to German. Although English and German are closely related, and compounding, in particular nominal compounding, is considered a productive word-formation pattern in both languages, German clearly differs from English in that the distinction between compounds and phrases is almost always uncontroversial whereas in English this distinction is highly problematic.

To start with, compounds in German generally have forestress, cf. (4), in contrast to phrases which have end-stress, cf. (5). In English, compound stress is much more variable, and although many compounds have forestress, as in German, there are also many forms that have rightward, phrasal stress, cf. (6).

## GERMAN

	FORESTRESS	END-STRESS
(4)	' <i>Apfelkuchen</i> 'apple cake' ' <i>Rotwein</i> 'red wine'	- -
(5)	-	<i>Kuchen aus 'Äpfeln</i> 'cake from apples' <i>roter 'Wein</i> 'red wine'

## ENGLISH

	FORESTRESS	END-STRESS
(6)	' <i>apple tree</i>	<i>apple</i> 'pie
	' <i>classroom</i>	<i>silk</i> 'tie
	' <i>blackboard</i>	<i>summer</i> 'night

Second, German compounds can be distinguished from phrases by the absence of inflectional marking of the modifier, cf. (7), where the adjectival modifier *rot* 'red' is inflected for gender, case, and number in the phrase but appears in the stem form in the compound. In English, on the other hand, there is generally no prenominal adjectival inflection, neither in the compound nor in the phrase. Thus, the inflection criterion cannot be applied to English and it is impossible to decide on the basis of inflection whether forms such as *red wine* or *cold war* are to be regarded as compounds or phrases.

		COMPOUND	PHRASE
(7)	GERMAN	<i>Rot</i> Ø <i>wein</i>	<i>rot</i> <u>er</u> <i>Wein</i> 'red wine'

Finally, German compounds are uniformly marked by solid spelling as well as, to a lesser extent, hyphenation. In contrast, English compounds may be written either as one or two words or are hyphenated, cf. (8)a), and it is also known that the spelling of some words is variable, cf. (8)b).

- (8) a. *classroom, summer night, milkman, milk bottle*  
 b. *word formation, wordformation, word-formation* (e.g., Bauer 2017: 6)

All these facts are well-known and various explanations have been suggested in order to explain the behaviour of English compounds, and in particular the variability of compound stress: semantic, phonological, and syntactic factors as well as analogy etc. (cf. Plag 2006; Plag, Kunter & Lappe 2007; Bell & Plag 2013; Giegerich 2015: chap. 3; among many others).

What is more important in the present context, however, are the assumptions about the status of these complex forms that follow from these observations.<sup>2</sup> On the one hand, there are uniform analyses which assume that the formations in question are either all morphological, and thus compounds (e.g., Olsen 2000; Bell 2011; for N+N compounds), or all syntactic, and thus phrases (e.g., Spencer 2003). Such analyses need explanations for the fact that morphological objects can have syntactic properties (i.e., phrasal stress) and vice versa.

2 In addition to the stress criterion other criteria have also been discussed with respect to the compound-phrase distinction in English, such as coordination or outbound anaphora with *one*. I cannot discuss these tests here. Importantly, the forms that evolve as either compound or phrase do not always coincide with the outcomes of the stress criterion.

Meanwhile, other analyses take the stress criterion (or other criteria, for that matter, cf. fn. 2) seriously and conclude that in fact some are morphological objects (i.e. the forms with forestress) and others (with end-stress) are syntactic (e.g., Payne & Huddleston 2002; Giegerich 2004). A consequence of this, however, is that some forms which are apparently very similar, have a distinct grammatical status. Some of the pairs that are usually cited in this connection are given in (9) (from Lees 1960: 120):

	COMPOUND?	PHRASE?
(9)	a. 'apple cake	apple 'pie
	b. 'Madison Street	Madison 'Avenue

Finally, there are also uniform analyses that leave open the exact nature of the forms, e.g. Bauer (1998), Arndt-Lappe (2011). It also has been advocated that the inconclusive data are an indication of the fact that the compound-phrase distinction does not exist. Giegerich (2015), in turn, explains these data by means of an overlap between syntax and word-formation (cf. Section 4.3).

Thus, with regard to the central question about the nature of the relation between word-formation and syntax German can be regarded as evidence in favour of a clear, categorial distinction, whereas English rather provides evidence against a categorial and in favour of a gradient relationship. In order to further investigate the question, in what follows evidence of two sorts will be discussed, based on a variety of European languages: Firstly, it will be shown that a categorial compound-phrase distinction as in German is by no means the standard case since in various languages there are compounds with syntactic properties, such as phrasal stress. Thus, various languages have compounds with mixed (morphological and syntactic) properties which is problematic for the idea of a categorial distinction between word-formation and syntax. Secondly, many languages also have complex word-level entities with an internal phrasal structure, analysed here as syntactic compounds. They are mainly phrasal but also have morphological properties. For this reason, they can be regarded as formations “in between” compounds and phrases, just as compounds with mixed properties.

The discussion is based on data from a total of 14 European languages, belonging to different language families, and exhibiting, for this reason, some typological differences. Although it is not a representative selection of an areal or typological language group they nevertheless attest the wide cross-linguistic distribution of the object of investigation.<sup>3</sup> Many of the data presented here are taken from other studies, both language-specific and

3 Similar observations have also be made for other, unrelated languages, e.g., Kageyama (2001) for Japanese (see also Section 4.3).

cross-linguistic ones, with various theoretical backgrounds. Using a great number and variety of problematic formations in different, partly non-related languages, the central aim of the paper is to show how common and widespread it is to have data which are problematic for a strict demarcation between morphology and syntax. In fact, none of the languages reviewed here provides evidence in favour of a strict demarcation. It can be concluded, then, that languages with mixed formations are the rule and not the exception which, in turn, must have implications for grammatical theory. The second aim of the paper, therefore, is the theoretical discussion of the data which also includes the comparison of models from different theoretical backgrounds dealing with similar data.

The structure of the paper is as follows: Section 2 discusses compounds in various languages that have syntactic properties in addition to morphological ones. Section 3 deals with syntactic compounds, again from a cross-linguistic perspective. It also briefly presents existing analyses of syntactic compounds. Section 4 deals with the theoretical modelling of the findings of the previous sections. It introduces the idea of modelling the relationship between word-formation and syntax as a continuum of related constructions as suggested in constructionist frameworks, focusing on Parallel Architecture (Jackendoff 2002; 2010, among others) and Construction Morphology (Booij 2010). Moreover, this section also discusses the data at hand with respect to approaches that defend a demarcation view on morphology and syntax (viz. Kageyama 2001, Ackema & Neeleman 2004) and compares the idea of a continuum to that of an overlap as proposed for the English compound-phrase distinction in Giegerich (2015). The paper ends with a brief conclusion in Section 5.

## 2. Problems with compounds

Compounds are often ascribed a naming function, i.e. they are used as conventional names. The naming function of compounds can be considered a basic function of word-formation: “Words, or better lexemes, have the essential function of organizing and categorizing human experience into symbolic units, and word formation is a process that implements this function by creating new lexemes that serve to name new concepts.” (Masini 2019c: 1). In the case of compounding, this naming function is related to the classifying function of the compound modifier, since the modifier specifies a subtype of the concept denoted by the head. In this sense, compounds are lexical units (or lexical entities), when ‘lexical’ is understood as ‘stored in the lexicon/long-time memory’. However, although compounds are typically lexical units in this sense, this is not necessarily the case, firstly, because not

every compound is necessarily stored in long-time memory (i.e. occasional formations) and secondly, because compounds occasionally do not have a naming function but instead are used as descriptions, just as phrases, see Section 4.1 (cf. Schlücker 2013, for instance).

Referring to compounds as lexical units in the following makes reference to this meaning of lexical. It is important to realize that the term ‘lexical’ is notoriously ambiguous and is often used to mean ‘word-level entity’ or ‘resulting from word-formation’, in particular when word-formation is equated with the lexicon. In the following, I try to avoid this latter use of ‘lexical’.<sup>4</sup> In other words, the grammatical properties should not be confused with the lexical status (see also Section 4.2).

## 2.1 Stress

The two criteria discussed so far in order to characterize compounds and to distinguish them from phrases are stress and inflection. In German, compounds have forestress and thus modifier stress while phrases have end-stress and thus head stress (e.g., Schlücker 2012). Similar stress patterns can be found in other Germanic languages such as Dutch (e.g., Köhnlein 2019), West-Frisian (e.g., Visser & Weening 2019), Danish (e.g., Hansen 2005), Swedish (e.g., Telemann 2005), or Icelandic (e.g., Bjarnadóttir 2017), with the exception of English, of course. From a more general perspective compound stress does not necessarily mean forestress (as in the Germanic pattern) but just a stress pattern that sets compounds apart from phrases. Thus, compounds are expected to display lexical stress, being different from phrasal stress, irrespective of how lexical and phrasal stress are realized in a particular language. Polish compounds, for instance, have lexical stress on the penultimate syllable which clearly sets them apart from phrases (e.g., Szymanek 2009: 467). Greek has several compound stress patterns which largely depend on the structure of the compound and the morphological properties of the constituents, but all of them are different from phrasal stress (e.g., Ralli 2013a: 79–83). From an even broader perspective one might add that stress is only one instance of the specific phonological properties of compounds, since in other languages compounds are characterized by other phonological properties such as tonal patterns or vowel harmony (cf. Bauer 2009a: 345).

However, English is not the only language which is problematic as regards stress. Romance languages such as French, Italian, and Spanish might be

---

4 I will use it occasionally, however, in some cases when I rely on well-established terms from the literature, such as ‘lexical stress’ in the next section.



regarded as even worse because they do not have compound stress (e.g., French) or no unitary compound stress (e.g., Spanish) (cf. Rainer & Varela 1992; Van Goethem 2009; Arnaud 2015, for instance). French compounds, for instance, have word-final stress, which is in fact phrasal stress, cf. (10), where both the compound and the PP construction have end-stress.

## FRENCH

(10) *stylo-'bille* (lit. pen ball) – *stylo à 'bille* (lit. pen of ball)    both: ‘ball pen’

An obvious conclusion from this observation is that compounds in Romance cannot be characterized on phonological grounds, i.e. that stress is irrelevant to the definition of compound in Romance. However, one might also draw the conclusion that the stress criterion is indeed relevant but – as it has the value of “syntactic” rather than “morphological” – that Romance compounds are not morphological to the same extent as compounds are in other languages.

## 2.2 Inflection

The next property to be discussed is internal inflection: If compounds are composed of stems (or lexemes) no inflectional markers should be found inside. The reason for the ban on inflection is that inflection is a syntactic feature. Since compounds are word-formation entities inflectional marking is not expected inside compounds and the internal constituents should neither show agreement with the head nor vary along with the context.

This criterion is of course relevant only for those languages which have the respective inflectional categories, which does not always hold for English, as shown above. The examples in (11) show that German and Dutch compounds do not have internal inflection, which sets these forms apart clearly from the corresponding phrases. In English, on the other hand, this distinction cannot be made.

	COMPOUND	PHRASE
(11) a. GERMAN	<i>FreiØstaat</i>	<i>frei<u>e</u>r Staat</i> ‘free state’
b. DUTCH	<i>vrijØstaat</i>	<i>vrij<u>e</u> staat</i> ‘free state’
c. ENGLISH	<i>freeØ state</i>	

Internal inflection should not be confused with linking elements. It is well known that compounds in various languages (e.g., German, Dutch, Danish, Swedish, Russian, and Polish, among others) may have linking elements, cf. (12). In Greek, linking elements are even obligatory and can be regarded as a defining property of compounding (e.g., Ralli 2008).

(12) a. GERMAN	<i>Klasse</i>	-n-	<i>zimmer</i>	‘class room’
	class	LE	room	
	<i>Arbeit</i>	-s-	<i>zimmer</i>	‘study’
	work	LE	room	
b. DUTCH	<i>stad</i>	-s-	<i>raad</i>	‘city council’
	city	LE	council	

However, linking elements are not inflectional markers, although some of them are formally identical. Contrary to inflectional markers, linking elements are meaningless, thus they do not express the grammatical meaning that comes with inflectional elements, and they are invariably bound to the respective formations. In many languages linking elements are historically related to inflectional affixes but their formal distribution is different in present-day language (so, for instance, the -s- in *Arbeitszimmer* in (12)a) does not show in the inflectional paradigm of *Arbeit*). Thus, compounds with linking elements do not constitute counterexamples to the ban on internal inflection.

What is more relevant, therefore, are English and French compounds with an apparent internal plural marking, especially since English and French are generally described as not having linking elements.<sup>5</sup> Some of the pertinent English examples are given in (13). In contrast to linking elements, the -s- is meaningful and expresses plural meaning.

(13) *schools concert, suggestions box, programmes list, jobs programme*

However, these plural s-markers are not variable as is usual with genuine inflectional markers, and are thus invariably bound to the lexemes. In some cases, forms with and without plural marking co-exist but the distribution is unclear. In general, the plural forms seem to be much rarer than the singular forms (cf. Bauer 2017: 144ff). In any case, the presence or absence of the marker does not clearly depend on the meaning. Since the unmarked modifier can also receive a plural interpretation, there is no real meaning difference between a compound with a plural marking on the modifier and one without (e.g., *job|jobs file, programme|programmes list*) (although a singular interpretation seems less likely if there is a s-marker). In sum, these compounds can be considered as rather weak evidence against the ban on internal inflection of compounds.<sup>6</sup>

5 With the exception of neoclassical compounding.

6 If the inflection criterion is taken seriously, it follows that the forms in (13) are phrasal, thus NP+N structures. Bauer (2017: 148) correctly points out that this would mean that even synonymous forms such as *job programme* and *jobs programme* would have to be assigned different internal structures. This is not only

Plural marking in French is different. French compounds are usually left-headed. Plural is always marked on the head and in many cases also on the modifier, as can be seen in (14). This means that French compounds systematically have word-internal inflection. In contrast to the English examples above, this is regular inflection.

## FRENCH

- (14) a. *poisson-scie* (lit. fish.SING saw.SING) ‘sawfish’  
       *poissons-scies* (lit. fish.PL saw.PL)  
       b. *auteur-compositeur* (lit. writer.SING composer.SING)  
           ‘songwriter-composer’  
       *auteurs-compositeurs* (lit. writer.PL composer.PL)

However, in most cases the internal plural marking is not phonologically realized and cannot be heard. Although forms like *poissons-scies* and *auteurs-compositeurs* are without doubt plural forms (which becomes obvious from agreement with the verb and the choice of plural determiners) the question arises of whether the internal *s*-marker is a genuine inflectional marker (which is phonologically empty) or whether the head does not actually carry an inflectional marker and the *-s-* is just added because the orthographic norm requires it. That said, there are also forms where the internal plural is phonologically non-empty, especially (but not only) in the case of irregular plural formation. So all cases in (15) have audible internal plural marking (double-underlined) on the left constituent (which is the head in (15)a) and the modifier (both N and A) in (15)b-d)).

- (15) a. *capital-décès* (lit. capital.SING death.SING) ‘death benefit’  
       *capitaux-décès* (lit. capital.PL death.SING)  
       b. *cheval-vapeur* (lit. horse.SING steam.SING) ‘horsepower’  
       *chevaux-vapeur* (lit. horse.PL steam.SING)  
       c. *bonhomme* (lit. good.SING man.SING) ‘fellow, fool’  
       *bonshommes* (lit. good.PL man.PL)  
       d. *social-démocrate* lit. social.SING democrat.SING) ‘social democrat’  
       *sociaux-démocrates* (lit. social.PL democrat.PL)

This can be regarded as evidence for the idea that in general, French compounds (both right-headed and left-headed ones) have regular internal plural marking and that French, therefore, provides counterevidence to the ban on internal inflection in compounds.

---

counterintuitive in terms of semantics but also because the form (with/without *s*) seems to be variable. Finally, the modifier constituent may not be modified in both cases. However, if the plural formation involves NP+N structures, then premodification should be permitted.

Another interesting case are V+N compounds in Danish. Here, the verbal modifier comes in its infinitival form (cf. Bauer 2009b: 409), contrary to V+N compounding in closely related languages such as Swedish, Icelandic, Dutch, or German, where the verbal modifier appears as stem form,<sup>7</sup> cf. (16), which compares Danish and Swedish V+N compounds.

	DANISH (from Bauer 2009b)	SWEDISH
(16)	<i>rullestol</i> (lit. roll.INF chair) ‘wheelchair’	<i>rullstol</i> (lit. roll.STEM chair)
	<i>løbetid</i> (lit. run.INF time) ‘duration, running time’	<i>löptid</i> (lit. run.STEM time)
	<i>bærepose</i> (lit. carry.INF bag) ‘carrier bag’	<i>bärkasse</i> (lit. carry.STEM bag)

Infinitival marking is without doubt inflectional marking. However, infinitive forms by definition do not show agreement and do not express the kind of grammatical meaning associated with finite forms. Also, the markers in (16) are invariable. Variability is however what one would expect from genuine internal inflection (cf. Finkbeiner & Schlücker 2019). So, Danish V+N compounds provide evidence against the ban on internal inflection although a rather weak one.

Icelandic and Finnish compounds, on the other hand, have both plural marking and case marking inside nominal compounds. In Icelandic nominal compounds, the modifier constituent may appear either as stem or as inflected form (cf. Bjarnadóttir 2002; 2017). Nominal modifiers inflect for genitive, both singular and plural, whereas dative modifiers are rare. Adjectival modifiers can also exhibit genitive marking, cf. (17):

ICELANDIC (from Bjarnadóttir 2017)

(17)	<u>N+N compounds</u> with genitive marking (singular/plural) on the modifier		
	a. <i>peru</i>	<i>tré</i>	‘pear tree’
	pear.N.FEM.GEN.SG	tree.N.NEUT	
	b. <i>bóka</i>	<i>búð</i>	‘book shop’
	book.N.FEM.GEN.PL	store.N.FEM	
	<u>A+N compounds</u> with genitive marking on the modifier		
	c. <i>sjúkra</i>	<i>hús</i>	‘hospital’
	sick.ADJ.GEN.PL.INDEF	house.N.NEUT	

The markers in (17) are considered inflectional markers and not linking elements because they appear in exactly the same inflectional form that would appear in the corresponding syntactic formation. Importantly, the inflectional system of Icelandic is highly complex and extensive. Nevertheless, in all compounds and with no exceptions, these markers always have the correct inflectional form. This is different from linking elements which may

7 Some German and Dutch V+N compounds have linking elements, e.g. GERMAN *Liegestuhl* (lit. lie-LE-chair, ‘deck chair’), DUTCH *drinkebroer* (lit. drink-LE-brother, ‘heavy drinker’).

be homonymous to inflectional markers in some cases or are diachronically derived from inflectional markers in some languages. They do not generally correspond to the inflectional paradigm of the modifier, however.

The inflectional nature of the internal markers is particularly obvious in Icelandic A+N compounds, cf. (18). As can be seen from (19), they show regular agreement with the head and thus formal variation which does not occur with linking elements.

ICELANDIC (from Bjarnadóttir 2017)

(18)	<i>litli</i>	<i>finger</i>		‘pinkie, little finger’
	little.ADJ.MASC.DEF	finger.N.MASC.INDEF		
(19)	NOM.SING	<i>litlifingur</i>	NOM.PL	<i>litlifingur</i>
	ACC.SING	<i>litlafingur</i>	ACC.PL	<i>litlafingur</i>
	DAT.SING	<i>litlafingri</i>	DAT.PL	<i>litlafingrum</i>
	GEN.SING	<i>litlafingurs</i>	GEN.PL	<i>litlafingurs</i>

Given that the markers in question are indeed inflectional markers, the question arises why these forms are considered compounds and not phrases. The main reason for this is that they are compounds according to another important criterion, namely stress: they exhibit lexical stress (on the first constituent) which is different from phrasal stress in Icelandic. In addition, the word order of N+N compounds with a genitive modifier is the reverse of the word order of corresponding genitive phrases. Thus these compounds are different from regular phrasal patterns.

The last case to be discussed is Finnish. Just as in Icelandic, Finnish nominal and adjectival compounds may have inflected modifiers, both nominal and adjectival, in addition to compounds with modifier stems.<sup>8</sup> Genitive marking (both singular and plural) is most frequent but other cases, such as the locative, also occur (cf. Niemi 2009; Pitkänen-Heikkilä 2016).

FINNISH (from Niemi 2009)

(20)	a. <i>käden</i>	<i>sija</i>	(lit. hand’s place) ‘handle’
	hand.GEN	place	
	b. <i>käsi.en</i>	<i>taputus</i>	(lit. hand’s clapping) ‘hand clapping’
	hand.GEN.PL	clapping	
	c. <i>silmällä</i>	<i>pito</i>	(lit. with an eye holding) ‘supervision’
	eye.LOC	holding	
	d. <i>veren</i>	<i>punainen</i>	(lit. blood’s red) ‘blood-red’
	blood.GEN	red	

8 In Niemi’s (2009) corpus study, 14% of the nominal, 20% of the adjectival, and 22% of the verbal compound modifiers have regular internal inflection.

As in Icelandic, the main reason for classifying these forms as compounds is their consistent lexical stress which clearly distinguishes them from corresponding phrases. An additional argument is that there are possessive suffixes and clitics that can be added to Finnish nouns but are not allowed inside compounds (cf. Niemi 2009: 241f).

To sum up, compounds in various languages have internal inflection. However, not all of them are equally problematic. Invariable and infinite forms, such as English plurals or Danish verbal modifiers, may be considered less severe. French, Finnish and Icelandic compounds pose a more serious problem because they certainly have internal inflectional marking and, for this reason, clearly syntactic properties.

Again one might ask what this means for the general definition of compound. If the inflection criterion is taken seriously, it has to be concluded that these forms are not compounds. This leaves us, however, with the question of what they are, as they obviously are not phrases, either, as is clear from their lexical stress and other morphological features, cf. next section. One might also conclude that the inflection criterion is irrelevant to the definition of compound and that they are thus regular compounds. The third option, as has already been mentioned in Section 2.1, is to assume that the inflection criterion is indeed relevant for the definition of compound but, as it has the value of “phrasal” rather than “morphological” in the cases at hand, these compounds are not morphological to the same extent as other compounds.

### 2.3 Syntactic impenetrability, inseparability, and unalterability

In the literature of the past decades, a large number of criteria for the definition of compound has been discussed. At the same time, it has been repeatedly demonstrated that many of these criteria are inappropriate for a general, cross-linguistic definition of compound, either because they are difficult to define and to prove (such as the claim that compounds denote a “unitary concept”) or because they do not occur cross-linguistically (such as linking elements or recursivity). Lieber & Štekauer (2009: 8), in their introduction to the *Oxford Handbook of Compounding*, review these shortcomings. They conclude that there are basically only three properties that can be maintained. Two of them have been discussed in the preceding sections: stress (and other phonological means) and inflection. The third criterion is (what Lieber and Štekauer call) syntactic impenetrability, inseparability, and unalterability. Just as with the inflection criterion, this means that compounds are characterized by their inability to undergo syntactic processes.

The reason for this is that compounds are words. As such they fall into the domain of the Lexical Integrity Hypothesis. Early versions of the Lexical

Integrity Hypothesis go back to the 1980s (Lapointe 1980) and it has undergone several revisions since then (cf. Lieber & Scalise 2007) (see also fn. 23). The Lexical Integrity Hypothesis, in a (strong) version formulated by Anderson (1992: 84), states that syntax neither manipulates nor has access to the internal form of words (where form is understood as ‘structure’). Thus, compounds are distinct from phrases in that they cannot be separated or interrupted, unlike phrases, cf. (21) and (22). Inseparability is of course not a criterion of compounds alone, but a core criterion of wordhood (cf. Haspelmath 2011: 44). Its application to compounds proves that they are words.

## ENGLISH

	COMPOUND	PHRASE
(21)	a. <i>a <u>round</u> whiteboard</i>	<i>a <u>round</u> white board</i>
	b. <i>*a white <u>round</u> board</i>	<i>a white <u>round</u> board</i>

## GERMAN

	COMPOUND	PHRASE
(22)	a. <i>eine <u>schöne</u> Altstadt</i>	<i>eine <u>schöne</u> alte Stadt</i> (lit. a <u>beautiful</u> old town)
	b. <i>*eine <u>Altschön</u>stadt</i>	<i>eine alte <u>schöne</u> Stadt</i> (lit. an old <u>beautiful</u> town)

Also, the modifier constituent in a compound cannot be modified itself, again unlike phrases, cf. (23) (from Giegerich 2006) and (24).

	COMPOUND	PHRASE
(23)	<i>*a <u>brilliantly</u> whiteboard</i>	<i>a <u>brilliantly</u> white board</i>

	COMPOUND	PHRASE
(24)	<i>*eine <u>sehr</u> Altstadt</i>	<i>eine <u>sehr</u> alte Stadt</i> (lit. a very old town)

A specific property which is usually regarded as an argument in favour of the Lexical Integrity Hypothesis is that complex words, and thus compounds, are anaphoric islands. This means that constituents of complex words cannot be co-referent with other words in a phrase and thus they cannot anaphorically be referred to (cf. Postal 1969; Sproat 1988). Again, this is a property that distinguishes compounds from phrases. This can be seen from (25) (from Sproat 1988): While it is possible to refer back anaphorically to the embedded constituent *truck* in the phrase (cf. (25)a)), this is impossible if *truck* is the internal constituent of a compound (cf. (25)b)).

(25)	a. <i><u>Drivers of trucks</u><sub>i</sub> fill them<sub>i</sub> up with diesel</i>
	b. <i>*<u>Truck</u><sub>i</sub> <u>drivers</u> fill them<sub>i</sub> up with diesel</i>

However, it is also well known that this restriction does not generally hold and that there are numerous grammatical and acceptable cases of anaphoric reference to compound constituents. Two of them are given in (26) (from Ward, Sproat & McKoon 1991).

- (26) a. *There's a balance sheet<sub>i</sub> concern – we've never had to read it<sub>i</sub> before.*  
 b. *Bush<sub>i</sub> supporters would stay home, figuring he<sub>i</sub>'d already won.*

Ward et al. (1991) argue that this phenomenon is largely a pragmatic one and rests only partly on grammatical factors (cf. also Montermini 2006). Also, not all kinds of complex words behave in this way; it seems that compounds in general are more penetrable than derived words. Importantly, the main rationale behind the principle of anaphoric islandhood is that the internal constituents are lexemes and that lexemes as such are non-referring, contrary to noun phrases (cf. Sproat 1988). If morphological rules operate on lexemes, then this explains why referring expressions – and anaphoric reference to these expressions – are not expected in compounds. This, on the other hand, also explains why many of the grammatical examples given in the literature involve proper name constituents, as in (26)b), since proper names are inherently referential.<sup>9</sup>

Interestingly, West Frisian has a compound pattern with modifiers that are not names. Still, these modifiers are generally referential (and can be anaphorically referred to), cf. (27).

WEST FRISIAN (from Hoekstra 2002)

- (27) *kokensfljer* (lit. kitchen.GEN floor) 'floor of the kitchen'  
*loddefiem* (lit. shovel.GEN handle) 'handle of the shovel'

Thus, the modifier refers to a specific entity which is contextually salient, e.g. a specific kitchen. For this reason, these compounds can only be used definitely (cf. Hoekstra 2002). In addition, the *s/e*-markers in these examples are not linking elements but rather inflectional markers (although they are not fully identical to the regular prenominal genitive markers, either, cf. Hoekstra (2002: 230)), since they are meaningful and express a possessive (part-of) relationship. For this reason, these formations have been termed genitive compounds. Finally, they have a special stress pattern (with primary stress on the second constituent) which is lexical rather than phrasal stress but at the same time also different from the usual stress pattern of compounds in West Frisian (which is forestress, just as in the closely related

---

9 This is not to say that proper name modifiers in compounds are generally referential, since they obviously are not and anaphoric reference to proper name modifiers is often impossible for this reason, cf. Schlücker (2017; 2018).



languages of Dutch and German). Thus, West Frisian genitive compounds show of number of properties that are rather unexpected in compounds because of their syntactic nature.

To conclude, it is obvious (and has often been noted) that the principle of anaphoric islandhood in its original form is too strict. Some compounds do in fact allow reference to their internal constituents under certain circumstances. It has also been suggested that anaphoric islandhood is not categorial but rather a gradient phenomenon (cf. Montermini 2006: 129, for instance). In line with what has been suggested above regarding stress and inflection, one might conclude that anaphoric islandhood, just like inseparability and unalterability, is indeed relevant for the definition of compounds and that core morphological compounds are anaphoric islands, inseparable and unalterable. Although in our data, there are no examples of compounds that are separable or allow internal modification (i.e. are alterable),<sup>10</sup> there obviously are formations that are not anaphoric islands. They might be considered compounds, but they are not morphological to the same extent as other compounds. Thus, compounds that allow anaphoric reference to their internal constituents are more similar to phrases because this presupposes the referential interpretation of these constituents. Still, they are not phrases; they are situated in between.

## 2.4 Summary

The previous sections have evidenced that even in the small sample of languages examined here there are compounds that violate the basic criteria of compounding, viz. stress, inflection and syntactic impenetrability, i.e. anaphoric islandhood. Yet, these formations are usually considered compounds, and thus word-formation entities. Also, they are instantiations of systematic and productive patterns. They are by no means exceptional, marginal or infrequent. And every type of violation is found in several languages which are not even necessarily related.

Thus, there are both genuine morphological, prototypical compounds and those that intersect with phrases. Languages may only have fully morphological compounds, such as German, or only those that systematically lack a morphological property, such as French (regarding stress), or both, such as

---

10 Therefore, inseparability and unalterability might be regarded as essential, indispensable properties of compoundhood or, more generally, wordhood. Yet, these properties are not sufficient to distinguish compounds from phrases. As has also been observed by Haspelmath (2011: 44f) there are many inseparable combinations that are not considered words (or compounds), such as *both my* (as in *both my parents*) or *very good* (as in *very good food*).

West-Frisian (which, in addition to the genitive compounds discussed above, also has fully morphological compounds). In addition, this also shows that there is no uniform, cross-linguistically valid morphological definition of compound.

If not all formations that are considered compounds are morphological to the same extent and if compounds with syntactic properties occur regularly and cross-linguistically it follows that a categorial distinction between compounds and phrases as in German is by no means the standard case and that from a cross-linguistic point of view, the idea of a categorial distinction cannot plausibly be maintained. However, this raises the question of whether the idea of a categorial distinction can be maintained for individual languages, such as German, Dutch or Greek, that do have prototypical morphological compounds (and only those). In the following, it will be argued that these languages are in fact not fundamentally different if one takes into account so-called syntactic compounds: syntactic formations that are usually not considered compounds but that show – given the pertinent criteria – a remarkable degree of similarity with compounds. Syntactic compounds have an internal phrasal structure but in addition, they also have word-formation properties. Also, they function as conventional names, like morphological compounds. Thus, just as the compounds discussed in this section, they are patterns with mixed properties and can therefore be regarded as evidence in favour of a non-categorial relation between word-formation and syntax, even in languages that do not have “deviant” compounds, such as German.

### 3. Syntactic compounds

Syntactic compounds are phrasal formations situated “in between” compounds and phrases because they share basic properties with phrases but also differ from them with respect to other basic properties. Equally, they share basic properties with morphological compounds but also differ from morphological compounds with respect to other basic properties. Importantly, they are not idiosyncratic, isolated cases of lexicalization but rather occur systematically and can be captured by abstract, productive patterns. All languages reviewed here have syntactic compound patterns, though not necessarily the same ones.<sup>11</sup>

---

11 The terminology is quite diverse in this field. Formations such as those discussed in the following are often also referred to as multi-word expression, lexical phrase, phraseme, fixed expression, idiom and the like. However, these terms are also often used for isolated, idiosyncratic cases of lexicalization such as *pain in the neck* or *kick the bucket*. For terminological remarks see also Section 3.4.

Syntactic compounds are similar to morphological compounds since they are complex entities, made up of at least two lexemes. They have a naming function, and thus are conventional names, and just like compounds some of them have a compositional meaning while others are non-compositional. Also, they are the product of productive, abstract patterns, again like compounds. Finally, they qualify as word-level entities since they share with compounds the property of syntactic impenetrability, inseparability, and unalterability. In addition, some patterns also exhibit other properties that have been considered as defining properties of compounding before, namely lexical stress or absence of inflection. At the same time, they have an internal phrasal structure and typical phrasal properties such as phrasal stress and word-internal inflection. Also, they are written separately.

In the following, three such patterns will be discussed: N+PP patterns, N+N.GEN patterns and A+N / N+A patterns. These patterns are cross-linguistically frequent and are the patterns that have been most often discussed in the literature, although in some languages, including German and Dutch, they have received relatively little attention.

### 3.1 N+PP constructions

It is well-known that the status of compounds in Romance is highly controversial. The number of formations that are undisputedly regarded as compounds (cf. (28)) is rather low compared to N+PP constructions as in (29) (e.g., Nicoladis 2002), which are sometimes also termed compounds, since they function as conventional names and often, though not always, have a non-compositional meaning.

- (28) a. FRENCH     *stylo-bille* (lit. pen ball) ‘ball pen’  
       b. ITALIAN     *pesce-spada* (lit. fish sword) ‘sword fish’  
       c. SPANISH    *célula madre* (lit. cell mother) ‘stem cell’
- (29) a. FRENCH     *moulin à vent* (lit. mill of/for wind) ‘wind mill’  
       *fauteuil de table* (lit. chair of table) ‘(armed) chair (around a table)’  
       *homme de la rue* (lit. man of the street) ‘average person’  
       *tasse à café* (lit. cup of/for coffee) ‘coffee cup’  
       *camion de pompiers* (lit. truck of fire brigades) ‘fire truck’  
       b. ITALIAN     *casa di riposo* (lit. house of rest) ‘nursing home’  
       *cucina a gas* (lit. kitchen at gas) ‘gas stove’  
       *carta di credito* (lit. card of credit) ‘credit card’  
       c. SPANISH    *casa de campo* (lit. house of country) ‘country house’  
       *dolor de cabeza* (lit. ache of head) ‘headache’  
       *café con leche* (lit. coffee with milk) ‘white coffee’

On the one hand, formations like those in (29) are undoubtedly phrasal which is obvious from the presence of explicit syntactic markers, i.e. relational markers (prepositions) and determiners. Thus, compared with compounds, the semantic relationship between the constituents is more explicitly expressed. On the other hand, however, the formations do not allow the modification of the constituents and are inseparable, unlike phrases, cf. (30). Thus, according to the Lexical Integrity Hypothesis, they are words rather than phrases.

#### FRENCH

(30) #*homme de la grande rue* (lit. man of the big street)

(30) shows that separation does not necessarily render the phrase ungrammatical (and it is therefore marked by # instead of \*). However, the conventional meaning of the lexical pattern is no longer available and it can only be interpreted literally, i.e. as a regular phrase.

Also, they are anaphoric islands, like compounds and unlike phrases. As can be seen from example (31)a), it is impossible to refer back anaphorically to the internal constituent *table* ('table') though this is possible in a corresponding regular phrase (31)b). This obviously has to do with the fact that the constituent *table* in the (a)-sentence, being an N, is non-referential (it does not refer to a specific table), contrary to the NP *la table* in the (b)-sentence (cf. also Gunkel & Zifonun 2011; Gunkel et al. 2017: 104ff).

#### FRENCH

- (31) a. \**le fauteuil de table, que, j'ai vue hier*<sup>12</sup>  
 (lit. the chair of table<sub>i</sub> that<sub>i</sub> I saw yesterday)  
 b. *le fauteuil de la table, que, j'ai vue hier*  
 (lit. the chair of the table<sub>i</sub> that<sub>i</sub> I saw yesterday)

Finally, although these formations have syntactic markers some of them are syntactically deficient because the determiner which would be present in a fully regular syntactic phrase is missing, cf. (32) (from Gunkel & Zifonun 2011):

#### FRENCH

- (32) a. *chemise de Ø nuit* (lit. shirt of night) 'nightdress'  
 (vs. *les bruits de la nuit* 'the noises of the night')  
 b. *château d'Øeau* (lit. palace of water) 'water tower'  
 (vs. *qualité de l'eau* 'quality of the water')  
 c. *voiture de Ø sport* (lit. car of sport) 'sports car'  
*moulin à Ø vent* (lit. mill of wind) 'wind mill'

---

12 I owe this example to Kristel Van Goethem.

The missing determiner clearly distinguishes these formations from regular phrases. Due to the lack of the determiner, they are also shorter and more compact than regular phrases which makes them more similar to morphological compounds with regard to their length.

Syntactic N+PP compound patterns with similar properties can also be found in numerous other languages. English, for instance, has the productive pattern [N of N]-pattern, cf. (33).<sup>13</sup> On the surface, these phrases resemble regular nominal phrases with an embedded PP. However, as in the case of French, they differ from regular noun phrases: the word order is fixed and may not be varied, the internal constituents may not be modified and are inseparable (that is, they do not allow the insertion of additional material), they are anaphoric islands and the determiners are missing.

#### ENGLISH

(33) *prisoner of war, bird of prey, bill of complaint, weapons of mass destruction*

The same holds for phrasal patterns in Slavic languages such as Russian and Polish. Both languages have productive N+PP patterns for the formation of lexical units that are the equivalents of English compounds, cf. (34) and (35) (cf. Masini & Benigni 2012; Ohnheiser 2015; Cetnarowska 2018; Cetnarowska 2019 among others).

POLISH (from Szymanek 2009; Cetnarowska 2019)

- (34) a. *pasta do zębów* (lit. pasta for teeth) ‘toothpaste’  
 b. *chustka do nosa* (lit. kerchief for nose) ‘handkerchief’  
 c. *dziurka od klucza* (lit. hole from key) ‘keyhole’

RUSSIAN (from Masini & Benigni 2012; Ohnheiser 2015)

- (35) a. *bor’ba za suščestvovanie* (lit. struggle for existence) ‘struggle for existence’  
 b. *vid na žitel’stvo* (lit. permit for stay) ‘permit of stay’  
 c. *priznanie v ljubvi* (lit. declaration in love) ‘declaration of love’

In German and Dutch, syntactic N+PP compounds are less frequent than they are in Romance, English, Polish or Russian. Some examples are given in (36).

- (36) a. GERMAN (from Gunkel et al. 2017)  
*Haus am See* (lit. house at the lake) ‘lakeside house’  
*Anzug von der Stange* (lit. suit off the bar) ‘off-the-peg suit’  
 b. DUTCH *restaurant met tuin* (lit. restaurant with garden) ‘garden café’

13 According to Klinge (2006), this pattern is a Romance loan to English (at round 1000 AD), e.g. *prisonnier de guerre* > *prisoner of war*.

Just as in the French and English examples discussed above, these constructions are inseparable and unalterable, such that they do not allow the insertion of an additional modifying adjective, for instance, cf. (37), at least not in the intended meaning.

#### GERMAN

- (37) a. #*Haus am kleinen See* (lit. house at the small lake)  
 b. #*Anzug von der oberen Stange* (lit. suit of the upper bar)

Also, (again in the intended reading) the noun in the embedded PP is non-referential and does not allow anaphoric reference, cf. (38)a), contrary to the corresponding regular phrase, cf. (38)b).<sup>14</sup>

GERMAN (from Gunkel et al. 2017: 105)

- (38) a. #*Sie haben ein Haus am See<sub>i</sub>, der<sub>i</sub> übrigens sehr tief ist.*  
 ‘They have a lake<sub>i</sub>side house which<sub>i</sub> is very deep by the way.’  
 b. *Sie haben ein Haus an einem See<sub>i</sub>, der<sub>i</sub> übrigens sehr tief ist.*  
 ‘They have a house at a lake<sub>i</sub> which<sub>i</sub> is very deep, by the way.’

And finally, as in French (cf. (32)) some of these formations are syntactically deficient in that the otherwise obligatory determiner is missing, cf. (39).

#### GERMAN

- (39) a. *Zimmer mit Aussicht* (lit. room with Ø view) ‘room with a view’  
 b. *Dame von Welt* (lit. lady of Ø world) ‘sophisticated woman’

To summarize, various languages have N+PP patterns for creating lexical units. They share with compounds the properties of syntactic impenetrability, inseparability, and unalterability. The embedded modifier is non-referential. Thus, they are anaphoric islands and, according to the Lexical Integrity Hypothesis, qualify as words. On the other hand, they are built up according to phrasal rather than morphological principles, including relational syntactic markers and determiners, and they also have phrasal stress.

### 3.2 N+N.GEN constructions

Another phrasal pattern are complex nominals with embedded genitives. Syntactic compounds of this kind exist in Russian, Polish, Greek, Polish, German and English, among others, cf. (40)–(44).

14 As is often the case with anaphoric reference to constituents of complex words/compounds, speaker judgments are rather heterogeneous. So where according to some speakers (including myself) anaphoric reference is impossible in (38)a), this is possible according to others, including one of the reviewers. However, in any case (38)b) is judged unproblematic and better than (38)a).

Naturally, such constructions only exist in languages that have genitive case marking.<sup>15</sup> The feature of case assignment clearly indicates the syntactic nature of these formations. Also, they generally have phrasal and not lexical stress. In addition, for Greek it can be observed that they lack linking elements which are compulsory in Greek morphological compounds, which is an additional indication of their syntactic nature.

#### GERMAN

- (40) a. *Schlaf der Gerechten* (lit. sleep the.GEN.PL just.GEN.PL) ‘sleep of the just’  
 b. *Geschenk des Himmels* (lit. gift the.GEN.SG heaven.GEN.SG) ‘godsend’  
 c. *Macht der Gewohnheit* (lit. force the.GEN.SG habit.GEN.SG) ‘force of habit’

#### ENGLISH

- (41) a. *lawyer’s fee*  
 b. *mother’s milk*  
 c. *women’s magazine*

RUSSIAN<sup>16</sup> (from Masini & Benigni 2012; Ohnheiser 2015; 2019)

- (42) a. *zub mudrosti* (lit. tooth wisdom.GEN) ‘wisdom tooth’  
 b. *nomer telefona* (lit. number telephone.GEN) ‘telephone number’  
 c. *urok čtenija* (lit. lesson reading.GEN) ‘reading instruction’

POLISH (from Cetnarowska 2019)

- (43) a. *dom studenta* (lit. house.NOM student.GEN.SG) ‘dormitory, student hall of residence’  
 b. *mąż stanu* (lit. man.NOM state.GEN.SG) ‘statesman’

GREEK (from Ralli 2013b; Bağrıaçık & Ralli 2015)

- (44) a. *zóni asfalías* (lit. belt safety.GEN) ‘safety belt’  
 b. *praktorío idíseon* (lit. agency news.GEN) ‘news agency’

Just as in the case of the N+ PP patterns these genitive phrases differ from corresponding regular genitive phrases in that, firstly, they are inseparable, cf. (45), at least in the intended meaning of the syntactic compound. Thus, these phrases are not ungrammatical, but can only be interpreted as the corresponding regular phrase, e.g. fresh (cow) milk that belongs to a specific mother.

15 Obviously, English does not have genuine genitive marking anymore but rather a possessor construction.

16 Russian also has syntactic compounds with embedded datives and instrumentals, e.g. *obmen studentami* (lit. interchange student<sub>INSTR.PL</sub>) ‘student exchange program’, cf. Masini & Benigni (2012).

- (45) a. #*mother's fresh milk*  
 b. #*lawyer's expensive fee*

The Greek examples in (46) also illustrate the property of inseparability. (46)a) is similar to the English ones in (45). (46)b) relates to a special feature of Greek, namely that in regular phrases the definite article can be doubled. Due to the property of inseparability, this is however ungrammatical in the syntactic compound (cf. Bağrıaçık & Ralli 2015; Ralli 2016).

#### GREEK

- (46) a. \**i zóni meyális asfalías* (lit. the belt big.GEN safety.GEN)  
 b. \**i zóni tis asfalías* (lit. the belt the.GEN safety.GEN)

Secondly, the genitive dependent, the possessor, is not interpreted referentially, as would normally be the case in a regular possessor phrase, cf. (47)a), but rather denotes an abstract concept, just as modifiers in compounds do, cf. (47)b). In other words, it does not refer, it classifies.

- (47) a. (*the*) *lawyer's son*, (*the*) *mother's bag*  
 b. (*the*) *lawyer's fee*, (*the*) *mother's milk*

Thus, unlike in regular phrases the genitive dependent is an N and not an NP, just as the modifier in morphological compounds which, as we have seen above, cannot usually be referred to anaphorically for this reason. The same contrast can be found with syntactic genitive compounds and regular genitive phrases, cf. (48) (cf. Rosenbach 2006):

- (48) a. \**a women's magazine, who<sub>i</sub> are interested in that kind of stuff*  
 b. *a magazine for women<sub>i</sub>, who<sub>i</sub> are interested in that kind of stuff*

Thirdly, the fact that the genitive dependent is an N and not an NP can also be seen from the fact that in English, where the possessor/genitive dependent – in contrast to the other languages shown above – has prenominal position, the determiner relates to the head of the phrase and not to the genitive dependent. The example *women's magazine* in (49)a) makes this particularly obvious because the possessor/genitive dependent is in the plural, so that the singular determiner *a* can only be meant to refer to *magazine*, contrary to (49)b) (cf. Rosenbach 2006).

- (49) a. *a women's magazine*  
 b. *a woman's book*

Fourthly, in these genitive constructions the order of the constituents is fixed, thus they resist reordering, unlike with the corresponding regular genitive phrases (in most languages).



Finally, the English syntactic genitive compounds also (partially) form an exception to the above-mentioned observation that syntactic genitive compounds have phrasal stress. Rosenbach (2006: 88) notes that stress is inconclusive not only with regard to morphological compounds in English, but also these genitive patterns: Although some of them have rightward, phrasal stress, they usually have leftward (thus: lexical) stress, cf. (50). Thus, just as the stress criterion does not help to distinguish between morphological compounds and (regular) phrases in English, it is also useless for distinguishing between syntactic genitive compounds from morphological compounds and regular phrases.

ENGLISH (from Spencer 2003)

- (50) a. *'men's room, 'Broca's aphasia 'Parkinson's disease*  
 b. *Foucault's 'pendulum, Hodgkin's lym'phoma, women's maga'zine*

If we compare syntactic genitive compounds to morphological genitive compounds (i.e. compounds with genitive case marking as in Finnish and Icelandic, cf. Section 2.2) it turns out that these patterns are almost identical. Both in syntactic and morphological genitive compounds the modifier/genitive dependent is non-referential, and both constructions are inseparable and unalterable. Also, all of them have a naming function and (potentially) a non-compositional meaning. The only difference then is stress (lexical stress with the morphological compounds and phrasal stress with syntactic ones). In the case of the English genitive pattern with forestress as in (50)a) not even this difference holds, so they are structurally fully identical to morphological genitive compounds. The German syntactic genitive compound pattern is slightly different, as it also contains a syntactic marker, namely the determiner (cf. (40)), which forms an additional difference. In contrast, West Frisian genitive compounds do not fit this picture. They have a referential modifier that can be anaphorically referred to, fully in parallel with regular genitive phrases. Also, they do not function as names though they are inseparable and have lexical stress.

All in all, this shows that syntactic genitive compounds are not fundamentally different from morphological genitive compounds, but differ only gradually. Compared to syntactic genitive compounds, morphological genitive compounds are thus slightly more like prototypical morphological compounds and unlike phrases. Meanwhile, morphological genitive compounds are also different from morphological compounds proper, due to their internal inflectional marking. The West Frisian genitive compounds might be even regarded as more phrasal than syntactic genitive compounds. Again, these data suggest a series of related patterns that differ only gradually, discounting a categorial distinction between compounds and phrases.

### 3.3 A+N / N+A construction

The third cross-linguistic example of syntactic compounds are A+N / N+A patterns. They are found in Polish, Russian, Greek, French, Italian, Spanish, Swedish, Dutch, and German, among others (cf. Cetnarowska 2015; 2018; Masini & Benigni 2012; Bağrıaçık & Ralli 2015; Van Goethem 2009; Fernández-Domínguez 2019; Giegerich 2005; Koptjevskaja-Tamm 2009; Booij 2009; Schlücker & Hüning 2009; Schlücker 2014, among others). As can be seen from the examples in (51)–(58), some languages have only one of the two word orders while others, like Polish, Russian, Italian, or Spanish, have both.<sup>17</sup>

#### GERMAN

- (51) a. *saurer Regen* ‘acid rain’ [A+N]  
 b. *rote Grütze* (lit. red groats) ‘red fruit jelly’ [A+N]  
 c. *gelbes Trikot* ‘yellow jersey’ (cycle racing) [A+N]  
 d. *häusliche Gewalt* ‘domestic violence’ [A.REL+N]

#### DUTCH

- a. *zure regen* ‘acid rain’ [A+N]  
 b. *oud papier* (lit. old paper) ‘scrap paper’ [A+N]  
 c. *koninklijk huis* (lit. royal house) ‘royal family’ [A.REL+N]

#### SWEDISH (from Koptjevskaja-Tamm 2009)

- (52) a. *röda hund* (lit. red dog) ‘measels’ [A+N]  
 b. *hög hatt* (lit. high hat) ‘top hat’ [A+N]  
 c. *friska luften* (lit. fresh air) ‘out of doors’ [A+N]  
 d. *röda korset* ‘Red Cross’ [A+N]

#### POLISH (from Cetnarowska 2018)

- (53) a. *kwaśny deszcz* ‘acid rain’ [A.REL+N]  
 b. *dział finansowy* (lit. department financial) ‘financial department’ [N+A.REL]  
 c. *niedźwiedź brunatny* (lit. bear brown) ‘brown bear’ [N+A]  
 d. *panna młoda* (lit. girl young) ‘bride’ [N+A]

17 Although German and English are generally considered to have A+N sequences only, in fact some N+A combination can be also found, e.g. English *attorney general*, *heir apparent*; German *Forelle blau* (lit. trout blue, ‘blue trout’), *Sonne pur* (lit. sun pure). The English examples are French loans or loan formations. The German examples form a special pattern different from the A+N/N+A patterns discussed here. For further discussion, see Bauer (2019), Schlücker (2019).

## RUSSIAN (from Masini &amp; Benigni 2012; Ohnheiser 2015)

- (54) a. *sotovyj telefon* (lit. official telephone) ‘mobile phone’ [A.REL+N]  
 b. *železnaja doroga* (lit. iron road) ‘railway’ [A.REL+N]  
 c. *dizel’nyj motor* (lit. diesel.ADJ engine) ‘diesel engine’ [A.REL+N]

## GREEK (from Ralli 2013b; Bağrıaçık &amp; Ralli 2015)

- (55) a. *đimósios ipáilos* (lit. public servant) ‘civil servant’ [A+N]  
 b. *psixrós pólemos* ‘cold war’ [A+N]

## FRENCH (from Van Goethem 2009; Arnaud 2015)

- (56) a. *train rapide* (lit. train fast) ‘fast train’ [N+A]  
 b. *guerre froide* (lit. war cold) ‘cold war’ [N+A]  
 c. *haute tension* (lit. high tension) ‘high voltage’ [A+N]  
 d. *forêt tropicale* (lit. forest tropical) ‘rain forest’ [N+A.REL]

## SPANISH (from Rainer &amp; Varela 1992; Fernández-Domínguez 2019)

- (57) a. *media luna* ‘half moon, crescent’ [A+N]  
 b. *luna nueva* (lit. moon new) ‘new moon’ [N+A]  
 c. *escalera mecánica* (lit. staircase mechanical) ‘escalator’ [N+A.REL]

## ITALIAN (from Masini 2019b)

- (58) a. *mezza luna* ‘half moon’ [A+N]  
 b. *terzo mondo* ‘Third World’ [A+N]  
 c. *carta telefonica* (lit. card telephonic) ‘phone card’ [N+A.REL]

These formations are syntactic in that they exhibit regular inflection of the adjective (i.e. agreement with the head) and have phrasal stress. In addition, to reflect these phrasal properties, they are written in two words.

Many of the examples above contain relational adjectives, many of which have English N+N equivalents. This becomes immediately understandable from the internal structure and semantics of these adjectives. Relational adjectives are derived from nouns. For this reason, they do not denote a property on their own but rather express a relation to this noun. So, A+N / N+A constructions with relational adjectives (both morphological and phrasal ones) are in fact very much alike N+N compounds and just as N+N compounds they typically are conventional names and thus lexical entities. Maybe even more importantly, relational adjectives lack prototypical adjectival properties as they generally cannot be used predicatively, cannot be modified and are ungradable. These properties are obviously among those that have been discussed above as word-level properties, so this suggests that relational adjectives are inherently inclined to form

syntactic compounds.<sup>18</sup> With respect to English, relational A+N sequences have often been assigned a different status than other A+N sequences in the literature (cf. Sadler & Arnold 1994; Giegerich 2005; 2006; 2015, for instance).

Just as with the N+PP and genitive patterns, the A+N / N+A patterns are like morphological compounds in that they have a naming function and possibly a non-compositional meaning. Also, both formations with relational and with non-relational, qualifying adjectives exhibit some of the other word-formation properties discussed above. So, to start with, they are inseparable, cf. (59).

- (59) a. GERMAN     #*der saure anhaltende Regen* (lit. the acid permanent rain)  
       b. RUSSIAN    \**sotovyj služebnyj telefon* (lit. cellular official telephone)  
       c. GREEK      \**o đimósios, kalós, ipátilos* (lit. the public good servant)

As we have seen above in (46), in Greek the property of inseparability also shows with respect to the ban of the double definite determiner (cf. (60)a)), which is otherwise a regular option in Greek A+N phrases (cf. (60)b)):

- (60) a. \**o psixrós o pólemos* (lit. the Cold the War)  
       b. *o meyálos o pólemos* (lit. the big the war)

Secondly, and closely related to the first criterion, the order of the constituents is fixed such that the predicative position of the adjective is excluded for these phrases, unlike with regular phrases. (61) illustrates an important difference between patterns with relational adjectives and those with qualitative ones. As mentioned above, relational adjectives can never occur in predicative position. For this reason, (61)a) is ungrammatical. Qualitative adjectives, on the other hand, can generally be used in either predicative or in attributive positions. However, if they occur in a syntactic A+N compound, they are immobile and can only be used attributively. For this reason, if used predicatively, as in (61)b), they cannot receive the classifying interpretation of the syntactic compound (e.g., denoting a particular kind of jersey). However, such a phrase can receive a regular, literal interpretation, so it is not ungrammatical (indicated here by #). (This difference is also present in (59).)

---

18 The formation of morphological compounds with relational adjectives, on the other hand, is more restricted. In German, for instance, almost only non-native, Graeco-Latin relational adjectives occur in A+N compounds (e.g. *Polarluft* ‘polar air’, *Instrumentalmusik* ‘instrumental music’). In Dutch they are not available in morphological compounds at all (with a few exceptions which are probably German loans, e.g. *sociaaldemocraat* ‘social democrat’), cf. Schlücker (2014).

## GERMAN

- (61) a. RELATIONAL ADJECTIVE: \**Die Gewalt ist häuslich* (lit. The violence is domestic)  
 b. QUALIFYING ADJECTIVE: #*Das Trikot ist gelb* (lit. The jersey is yellow)

Thirdly, A+N / N+A phrases do not allow internal modification, thus the adjectival modifier may not be modified, again like in morphological compounds, cf. (62).

- (62) a. GERMAN #*sehr gelbes Trikot* (lit. very yellow jersey)  
 b. POLISH \**dział najbardziej finansowy* (lit. department most financial)  
 c. GREEK \**metria psixrós pólemos* (lit. moderately Cold War)

A final piece of evidence that has not been used here comes from coordination. Coordination is instructive since it presupposes syntactic and semantic similarity of the conjuncts. The German examples in (63) show that it is possible to coordinate two morphological compounds (a), two syntactic compounds (b), a morphological and a syntactic compound (c), but not a syntactic compound and a regular A+N phrase (d) (cf. Schlücker 2014). Again, this shows that syntactic A+N compounds are similar to morphological A+N compounds and rather unlike regular A+N-phrases.

## GERMAN

- (63) a. Coordination of two morphological compounds  
*Rot- und Gelbgold* ‘red and yellow gold’  
*Rotgold und Gelbgold* ‘red gold and yellow gold’  
 b. Coordination of two syntactic A+N compounds  
*rote und gelbe Bete* ‘red and yellow beet’  
*das gelbe und das grüne Trikot* ‘the yellow and the green jersey’  
 (cycle racing)  
 c. Coordination of a morphological and a syntactic compound  
*Kartoffel- und grüner Salat* ‘potato and green salad’  
*das gelbe und das Bergtrikot* (lit. the yellow and the mountain jersey)  
 ‘the yellow and the polka dot jersey’  
*eine wilde oder eine Scheinehe* ‘a wild or a fictitious marriage’  
 d. No coordination of a syntactic compound and a regular phrase  
 \**die häusliche und die schlimme Gewalt* ‘the domestic and the bad violence’  
 #*das gelbe und das verschwitzte Trikot* ‘the yellow and the sweaty jersey’  
 #*saurer und permanenter Regen* ‘acid and permanent rain’

Finally, there are two particularly interesting special cases in Dutch and Swedish. To start with, it can be observed that in Dutch some syntactic A+N compounds lack the inflectional marking of the adjective. Thus, they do not



Hüning (2010: 207) points to an important functional aspect of this observation: Due to the lack of the inflectional marking the form of the adjective is invariable in the syntactic compound, just like a morphological adjectival compound modifier. Consequently, the entire formation has an invariable, stable form, like morphological compounds and unlike regular phrases. Due to this stable form the syntactic compounds are better suited as linguistic signs in De Saussure's sense, i.e. as conventional mappings of a form and a meaning, than formations with a variable form. Thus, according to Hüning (2010), the lack of the inflectional marking can be regarded as functionally advantageous since it makes the phrasal formation more easily identifiable as a naming unit. This also means that Dutch has two different syntactic A+N compound patterns (with / without adjectival inflection) and those that lack adjectival inflectional marking are more similar to morphological compounds than those that have adjectival inflection.

Syntactic A+N compounds in Swedish provide another example of a special marking by means of deviant formal properties. In general, regular Swedish complex NPs are characterized by double determination. This means that definiteness is not only realized by a determiner but also by a noun suffix, cf. (67). As observed by Koptjevskaja-Tamm (2009: 133–134), in Swedish syntactic A+N compounds the adjective may either be inflected (marked for definiteness) or not and may be combined either with a definite, suffixed noun or an indefinite noun, cf. (68).

#### SWEDISH

- (67) a. *det röda korset* (lit. the red.DEF cross.DET.DEF) 'the red cross'  
 b. *den gula hatten* (lit. the yellow.DEF hat.DET.DEF) 'the yellow hat'
- (68) a. *röda korset* (lit. red.DEF cross.DET.DEF) '(the) Red Cross'  
 b. *friska luften* (lit. fresh.DEF air.DET.DEF) 'out of doors'  
 c. *röda hund* (lit. red.DEF dog.INDEF) 'measels'  
 d. *hög hatt* (lit. high.INDEF hat.INDEF) 'top hat'

What seems to be avoided in syntactic A+N compounds, however, is the regular pattern of double determination, since they do not use the definite determiner in addition to the noun suffix (cf. also Finkbeiner & Schlücker 2019). Also, Swedish syntactic A+N compounds have a special stress pattern which distinguishes them both from compounds and regular phrases. This means that in addition to the other word-formation properties of syntactic A+N compounds, such as inseparability or the ban on the modification of the adjective, compounds in Swedish have two other properties that mark them as different from regular A+N phrases, namely a special prosodic pattern and, similar to Greek, cf. (60), the absence of double determination. Importantly, just as in the case of Dutch, this last property is not only a feature peculiar

to syntactic compounds but it also means that syntactic compounds are formally closer to morphological compounds since the latter, as with simplex words, do not have double determination. Thus, they take the suffixed determiner only and not the preposed definite determiner, cf. (69).<sup>20</sup>

- (69) a. SIMPLEX NOUN: *huset* (lit. house.DET.DEF) ‘the house’  
 b. MORPHOLOGICAL COMPOUND: *höghuset* (lit. high house.DET.DEF) ‘the sky-scraper’  
*rödvinet* (lit. red wine.DET.DEF) ‘the red wine’

Thus, the ban on the preposed determiner in syntactic A+N compounds makes them formally more similar to morphological compounds.

Finally, the defective N+PP patterns that lack the otherwise obligatory determiner (e.g. French and German, cf. (32), (39)) can be viewed in the same way. Due to the lack of the determiner the phrases are shorter and more compact, which makes them more similar to morphological compounds. Also, since in many languages the form of the determiner varies depending on the gender of the noun, the pattern has a more stable, invariable form if it does not have a determiner, thus ensuring a stable form-meaning relation.

### 3.4 Theoretical approaches to syntactic compounds

The syntactic compound patterns discussed in this section (N+PP, N+N.GEN and A+N/N+A) share with morphological compounds the properties of inseparability, unalterability and impenetrability. Thus, according to the Lexical Integrity Hypothesis they qualify as words or word-level items. At the same time, they have a phrasal structure. The analysis as word-level items with an internal syntactic structure is reflected by terms such as ‘syntactic compound’, as in the present paper (see also Booij 2009; 2010; 2019 for Dutch and Greek; Schlücker 2014 for German), ‘loose multi-word compound’ (Bağrıaçık & Ralli 2015; Koliopoulou 2008 for Greek)<sup>21</sup> and ‘phrasal noun’ (cf. Masini 2009; Masini 2019b for Italian; Masini & Benigni 2012 for Russian; Cetnarowska 2018; 2019 for Polish; Booij & Masini 2015). Whereas various kinds of phrasal entities can be lexicalized and thus are lexical items (e.g., all kinds of ‘classical’ verbal idioms) this does not imply

20 The simultaneous presence of the determiner and the noun suffix with words (simplex and complex ones) is only found with demonstrative use, e.g. *den här stolen* ‘this chair’, *det huset* ‘this house’.

21 In several works (e.g., Ralli 2013b), Ralli also uses the term ‘phrasal compound’ which is somewhat misleading, however, as this term is also widely used in the literature for compounds of the structure XP+N.



that they are also word-level items. This is, however, exactly what the terms ‘syntactic compounds’ and ‘phrasal noun’ express: they are word-level items with an internal syntactic structure.<sup>22</sup>

Also, it has been observed above that the adjective in syntactic A+N compounds and the genitival non-head in syntactic genitive compounds may not be modified, contrary to regular A+N and genitive phrases. This has led to analyses of these constituents as bare adjectives and nouns (i.e., non-projecting A<sup>0</sup> and N<sup>0</sup>) (cf., e.g., for A+N: Booij 2009; Booij 2010 [Dutch and Greek]; Masini & Benigni 2012 [Russian]; Schlücker 2014 [German]; for A+N and genitive constructions: Cetnarowska 2018; 2019 [Polish]).<sup>23</sup> Interestingly, word-level analyses with an internal syntactic structure ([A<sup>0</sup> N<sup>0</sup>]<sub>N</sub><sup>0</sup>; [N<sup>0</sup> N<sup>0</sup>]<sub>N</sub><sup>0</sup>) have also been proposed for some English A+N and N+N compounds (cf. Liberman & Sproat 1992; Sadler & Arnold 1994). The referential nature of the compound modifier in West Frisian genitive compounds, on the other hand, is captured in Hoekstra (2002) by analysing these modifiers as DPs.

Analyses of word-level items with an internal syntactic structure as well as the general view on the relation between word-formation and syntax taken here are obviously very much in line with constructionist approaches to grammar, since one of the basic constructionist tenets is that grammar is non-modular and that there is no morphology-syntax divide because all linguistic structure on all levels is represented by a single representational format, namely constructions (cf. Section 4.2). So it does not come as a surprise that in recent years various constructionist analyses have been offered for the patterns discussed here for several languages (e.g., Booij 2009; 2010; Van Goethem & Amiot 2019; Masini 2009; 2019b; Cetnarowska 2018; 2019; Masini & Benigni 2012; Schlücker 2014).<sup>24</sup>

---

22 The term ‘syntactic compound’ is wider than ‘phrasal noun’ in that it is not restricted to nominal entities. Booij (2010: chap. 4), for instance, discusses quasi-incorporation as an instance of syntactic compounding in the verbal domain.

23 Importantly, although these modifiers do not allow modification, in contrast to modifiers in regular phrases, they may have inflectional marking, other than genuine compound modifiers. Accordingly, Booij (2010: 177) argues that the constraint of Lexical Integrity should be divided into two subconstraints. The first one prohibits syntactic reordering [and modification, BS] while the second one proscribes agreement and case assignment. Whereas syntactic compounds are subject only to the first subconstraint, for morphological compounds the stronger version of the Lexicalist Integrity Hypothesis applies, i.e. both subconstraints.

24 This also holds for another compound type that is problematic for a strict demarcation view on grammar, viz. phrasal compounds, e.g. *mother-in-law joke*, *not guilty plea*, *underwater construction*.

## 4. Discussion

The preceding sections have shown that

- various languages have productive patterns for the formation of morphological objects called compounds that deviate from what one would expect from morphological objects in terms of stress, inflection, or lexical integrity.
- various languages have productive syntactic patterns for the formation of lexical units, many of which exist cross-linguistically. They are characterized by special morphosyntactic and semantic properties that distinguish them from regular phrases.
- since some compound patterns also have syntactic properties and the phrasal patterns discussed above also have morphological properties, they closely resemble one another and it is impossible to draw a clear line between them. Consider, for instance, the Icelandic A+N sequences in (17)–(19) and the Dutch A+N sequences in (64) and (66): although in the Icelandic sequences the adjective is regularly inflected and in the Dutch ones the adjective invariably lacks inflectional marking, the former are considered compounds and the latter phrases. This is not only a terminological matter of different languages or different linguistic traditions. There are also other properties of the patterns that explain the respective categorisations. In the case just mentioned, it is stress.

### 4.1 A continuum of constructions

The conclusion to be drawn from these observations is that instead of a categorial distinction between compounds and phrases, these data can be more adequately described as a series of related constructions that differ from each other in single phonological, morphosyntactic and/or semantic properties. These constructions form a continuum with fully morphological constructions (genuine compounds) and fully syntactic constructions (regular phrases) at the endpoints and various mixed patterns in between.

Table 1 is a simplified, schematic representation of this continuum. It contains the five most important word-formation properties according to the discussion in the preceding sections. Since inseparability and unalterability, i.e. the ban on internal modification, always co-occur they have been bundled here. The eight columns each represent one particular combination of these properties. The order of these columns from left to right is intended to reflect the transition from fully morphological (1) to fully syntactic (7). It is important to note that these columns are **not** specific constructions or patterns as discussed in the previous sections. They serve as placeholders for illustrative purposes only, since in reality each column reflects the various constructions that have the respective properties.

Tab. 1: Word-Formation properties and continuum of constructions.

	1	2a	2b	3	4	5	6	7
Inseparable & unalterability	X	X	X	X	X	X	X	
Non-referential modifier	X	X	X	X	X			
Absence of relational markers	X	X	X	X		X	X	
No internal inflection	X		X			X		
Lexical stress	X	X				X	X	

The list of properties in Table 1 is obviously not exhaustive. They can be considered crucial properties with respect to the present discussion, but the corresponding specific patterns clearly have more, often language-specific properties, such as double determination in Swedish and Greek or the missing determiners in some N+PP constructions, for instance. In addition, there are various other language-specific properties that have been not discussed here, such as linking elements, recursivity etc. Note also that the property of non-referentiality differs from the others in that it is not a formal property proper, but rather a semantic one. Finally, not all properties are really binary. This holds in particular for non-referentiality. In addition, stress is sometimes not binary, as in the case of West Frisian genitive compounds and Swedish syntactic A+N compounds, which have stress patterns different from both phrasal stress and compound stress proper.

However, all these details could be spelled out in detail and the representation extended accordingly, so they do not form a principled problem for the underlying idea.

- (1) has all and only word-formation properties. It corresponds to prototypical morphological compounds in various languages (e.g., German, Dutch, Polish, Greek, Swedish, Icelandic, ...).
- (2a) and (2b) are of equal rank, so (2a) is not considered more morphological than (2b). The corresponding patterns have most of the word-formation properties but also one syntactic property, namely phrasal stress or internal inflection, respectively. Among the patterns that correspond to (2a) are N+N and A+N compounds with inflected modifiers in Icelandic and Finnish (e.g. Icelandic *perutré* ‘pear tree’), for instance, but also syntactic genitive compounds with forestress in English (e.g., *men’s room*). (2b) matches with a subtype of English compounds (e.g., *apple pie*) and a subtype of Romance morphological compounds (e.g., French *stylo-bille* ‘ball pen’) but also with a subtype of Dutch syntactic A+N compounds (e.g., *geheim agent* ‘secret agent’).
- (3) combines the morphological properties of inseparability/unalterability, non-referential modifier and absence of relational markers

with the syntactic properties of phrasal stress and internal inflection. Patterns with this particular combination are some Romance morphological compounds (e.g., French *poisson-scie* ‘sawfish’), syntactic genitive compounds in various languages (e.g., Polish *dom studenta* ‘dormitory’) and syntactic A+N compounds, again in various languages (e.g., German *gelbes Trikot* ‘yellow jersey’).

- (4) is similar to (3), but in addition the patterns corresponding to (4) also have relational syntactic markers, such as the syntactic N+PP compounds in various languages (e.g., French *moulin à vent* ‘windmill’) and syntactic genitive compounds in German (e.g., *Macht der Gewohnheit* ‘force of habit’).
- (5) corresponds to a subtype of proper name compounds (in various languages).
- (6) corresponds to genitive compounds in West Frisian. (5) and (6) are positioned near the syntactic endpoint of the continuum which might be surprising since they seem to have less syntactic properties than (4). However, what is important here are the referential modifiers. They do not denote concepts, as is usual with compound modifiers, but refer to specific referents. If the modifier is interpreted referentially it does not have a classifying function. Instead, it “anchors” or “identifies” the referent of the head noun. For this reason, such compounds do not have a naming function. For instance, the modifier *Berlin* in (70) does not specify a particular class of patients but rather helps to identify one particular patient by establishing a local relation (from Schlücker 2013).

#### GERMAN

- (70) *Im Vergleich zu den beiden Patienten aus Boston ist der Berlin-Patient sehr viel intensiver untersucht worden (...)*  
 ‘Compared to the two patients from Boston the Berlin patient has been examined much more extensively (...).’

This is what would usually be expected from a syntactic construction (e.g. *the patient from / in Berlin; the coat of my son, my son’s coat*) and what has often be referred to as the “descriptive” function of syntax in contrast to the naming function of word-formation (cf. Bauer 1988).

- (7), finally, are regular noun phrase patterns, such as *the coat of my son*.

The crucial note here is that there is no point where a line between compounds and phrases (word-formation and syntax) could be drawn. Although there are fully morphological and fully syntactic patterns at the endpoints, there are various mixed patterns in between. Patterns with mixed properties that have more morphological properties are called compounds and others with more phrasal properties phrases, but there is no clear, principled difference

between them (note also that (2a), (2b) and (3) comprise patterns considered “compounds” as well as “phrases”).

Thus, from a cross-linguistic perspective there is no categorial distinction between compounds and phrases, as has also been argued by Haspelmath (2011) with regard to the distinction between words and phrases in general. In addition, this also means that there is no uniform cross-linguistic definition of a compound; not only because compounds in different languages have different morphological features, e.g. regarding linking elements (cf. Ralli 2013b), but also because no uniform, cross-linguistic distinction between compounds and phrases can be established.

However, what does this mean for individual languages and in particular languages that apparently do have a clear, categorial compound-phrases distinction, such as German? English has patterns corresponding to all columns, with the exception of (6). This is little surprise given the problems of the compound-phrases distinction in English outlined in Section 1. French, in turn, has patterns in (2b), (3), (4), (5) and (7), but not in (1), due to the lack of morphological compounds with lexical stress. Finally, in German, there are patterns in (1), (3), (4), (5), and (7). So, even in a language with an (apparent) clear categorial compound-phrases distinction such as German there are not only compounds and phrases proper (i.e., (1) and (7)) but also mixed patterns in between. Thus, if syntactic compounds are taken into account the distinction between compounds and phrases, or word-formation and syntax, must generally be regarded as gradient rather than categorial.

Obviously, syntactic compounds are not equally frequent in all languages. In general, the productivity of these patterns seems to be related to the productivity of morphological compounding in ways of complementation or competition (cf. Booij 2019; Cetnarowska 2019; Masini 2019a; Schlücker 2019, for instance), and the existence of a particular phrasal form may (but need not) lead to the blocking of the formation of a corresponding morphological compound and vice versa. In German, nominal syntactic compounds are less frequent than in other languages, which can be related to the high productivity of nominal morphological compounding (cf. Barz 2007, for instance). Accordingly, nominal syntactic compounds have been given little attention in the literature when compared to other languages. In other languages, like French, it is the other way round; similarly, the relative wealth of nominal syntactic compounding in Russian and Polish has been explained in connection with the restricted productivity of nominal morphological compounding in these languages. Thus, although from the perspective of an individual language the significance of syntactic compounding for the system of that language might not be immediately apparent, as in German, the cross-linguistic perspective makes its role very clear which in turn also has consequences for the theoretical assumptions about the individual languages.

## 4.2 The relationship between word-formation and syntax

The findings from the previous sections can be summarized as follows: (1) In addition to genuine compounds and phrases there are various patterns that have both word-formation and syntactic properties. (2) There is no such thing as a clear compound-phrase distinction, neither generally nor in a single language (of those studied here). These findings suggest, as indicated by the schematic arrangement of the relevant properties in Table 1, that word-formation and syntax are not strictly demarcated domains of grammar.

What does this mean for our understanding of the relationship between morphology (more specifically word-formation), syntax and the lexicon? Two questions are crucial in this connection: (a) What is (in) the lexicon? (b) Are morphology/word-formation and syntax distinct or are they rather one single structure building domain?

Following Jackendoff (2002; 2010, among others) and his Parallel Architecture Model, the lexicon is understood here as long-term storage rather than the locus of generation of complex words. Thus, the lexicon is not identical to morphology or word-formation. In fact, “[...] lexical/non-lexical and word/syntax are not related concepts, but orthogonally distinct scales” (Bauer 2017: 173). The lexicon contains simplex words, thus linguistic signs in the core meaning of a conventional pairing of form and meaning. In addition, it also contains morphological and syntactic building principles (“rules”) since they are learned and stored just like words. These “rules” are basically represented in the same way as simplex words, namely as tripartite structures consisting of a phonological and a morphological or syntactic structure (the two dimensions of the form) and a semantic-conceptual structure. These representations are referred to as constructions (or schemas in the terminology of Booij (2010)). The main difference between these kinds of constructions is the level of abstractness: on the one hand, there are fully specified constructions, e.g., simplex and complex words (*apple*, *apple tree*). The constructions that represent the morphological and syntactic building principles, on the other hand, are abstract (e.g.,  $[X_x + X_N]_N$ ,  $[NP VP]_S$ ). Some of these abstract formal structures, in particular syntactic ones, are not inherently associated with a specific meaning and/or phonological realization. Thus, these lexical items lack a phonological and/or semantic representation and they are, in this sense, “defective” or “doubly defective” constructions (Jackendoff 2002: 180). In fact, such defective constructions correspond to traditional word-formation or phrase structure rules.

In between these two extremes are constructions with varying degrees of abstractness, such as pairings of an abstract syntactic structure with a specific semantic representation (such as the resultative construction, e.g., *Fred watered the plants flat*, a prime example of an argument structure

construction which figures prominently in Construction Grammar, cf. Goldberg (1995)) but also various kinds of ‘constructional idioms’, i.e. constructions that have both lexically specified positions and open slots, such as [N by N] (e.g., *day by day*) (cf. Jackendoff 2008). Booij (2010), in the framework of Construction Morphology, applies the concept of constructional idioms to morphological entities. The derivational pattern with the suffix *-er*, for instance, is represented as a formal structure  $[X_V+er]_N$ . This structure is mapped to the semantic representation ‘one who Vs’ as well as to a partially filled phonological structure.

Thus, constructions exist on various levels of abstraction. This also includes partially as well as fully lexically specified constructions (e.g.,  $[X_V+er]_N$  vs.  $[dream+er]_N$ ). So the lexicon is conceived of as a network of constructions of various levels of abstraction which form hierarchies and are linked by inheritance relations.

Whereas all abstract and semi-abstract constructions are lexical items, the fully specified forms that arise from these patterns may or may not be stored. Complex forms with a fully compositional meaning that can be computed online (such as regular inflection) tend not to be stored, complex forms with a non-compositional meaning are stored, just as frequent formations. Since the morphological and syntactic compounds discussed here usually are conventional names, they are stored, thus they are mostly lexical items.

Thus, both morphological and syntactic structure building principles form part of the lexicon since the abstract morphological and syntactic constructions are lexical items themselves. In this sense there is no principled difference between syntax and the lexicon. This does, however, not mean that morphology and syntax form one single structure building mechanism. Although they share some properties, they also differ with regard to others. This has been discussed in detail with respect to compounds and phrases in the preceding sections. Thus, morphology and syntax are two different domains of structure building with related but yet different properties (e.g., Jackendoff 2002: 128ff; see also Masini & Audring 2019: 387ff). However, since morphology and syntax are on a par as structure building domains and since all these structure building principles are constructions, and thus lexical items, nothing precludes the existence of constructions that combine properties from both domains, thus mixed patterns such as those discussed in the previous sections. If constructions with mixed properties are arranged according to their similarity to the core principles of each domain, this yields a continuum of constructions, as described for the compound-phrase continuum in Section 4.1.

Assuming that “everything is in the lexicon” as described above does not mean that there are no restrictions on structure building. The morphosyntactic, phonological and semantic properties are specified in such

a way that the particular structure building process is clearly defined and delimited, thus restrictions are directly built in the constructional representation. Consider, as an example, Booij's (2009; 2010) analysis of syntactic and morphological A+N compounds in (71)–(73). To the left of the double arrow is the morphosyntactic information, to the right is the semantic information. The indices  $i$ ,  $j$  and  $k$  map pieces of morphosyntactic information on the respective semantic information.<sup>25</sup>

Schema for syntactic A+N compounds (from Booij 2009: 232)

$$(71) [A_i N_j^0]_{N^0 k} \leftrightarrow [\text{NAME for SEM}_j \text{ with some relation R to SEM}_{i k}]$$

The schema for a morphological A+N compound is quite similar. The only difference is that the internal constituents are not words and the entire structure is dominated by one  $X^0$  node only, cf. (72).

Schema for morphological A+N compounds

$$(72) [A_i N_j]_{N^0 k} \leftrightarrow [\text{NAME for SEM}_j \text{ with some relation R to SEM}_{i k}]$$

Finally, stress is specified in the phonological structure, as illustrated in (73). On the left side of these schemas, there is a (greatly simplified) phonological structure. (73)a-b) differ only with regard to stress, thus (a) is a schema for compounds with forestress and (b) for end-stressed compounds. In the same way, fore-stress and endstress can be represented in other constructions, regardless of their internal morphosyntactic structure.<sup>26</sup>

$$(73) \text{ a. } [{}^i X_i Y_j] \leftrightarrow [A_i N_j]_{N^0 k} \leftrightarrow [\text{NAME for SEM}_j \text{ with some relation R to SEM}_{i k}]$$

$$\text{ b. } [X_i {}^j Y_j] \leftrightarrow [A_i N_j]_{N^0 k} \leftrightarrow [\text{NAME for SEM}_j \text{ with some relation R to SEM}_{i k}]$$

It is easy to see that entities with mixed properties, such as the Icelandic A+N sequences with lexical stress and regularly inflecting adjectival modifiers, or Dutch A+N sequences with phrasal stress and non-inflecting adjectival modifiers, could be represented accordingly. There is no principled difference between these constructions, nor compared to the representation of fully morphological or fully syntactic phrases.

25 For other A+N combinations, for instance those containing relational adjectives, the semantic representation is different. Schlücker (2014; 2016) proposes detailed semantic analyses for several groups of A+N combinations in German and Dutch. There are also alternative proposals as to how to represent the naming function instead of using the predicate NAME (cf. Masini 2009; Schlücker 2014, for instance).

26 Similar schemas have been proposed for some of the other patterns discussed above, e.g. syntactic genitive compounds in Polish (Cetnarowska 2019), syntactic N+PP compounds in Italian (Masini 2009), or French N+N sequences (Van Goethem & Amiot 2019).



### 4.3 A gradient word formation-syntax structure or a word formation-syntax divide?

In addition to constructionist analyses of syntactic compounds and other data that are problematic for the idea of a strict divide between morphology/word-formation and syntax, various proposals in defence of such a divide have been made, three of which will be briefly discussed in the following.

Kageyama (2001) discusses what he calls “phrase-like words” in Japanese (as well as, rather briefly, in English). He finds that Japanese has a group of formations that have both morphological and phrasal properties, similar to the data discussed here. These formations are analysed as ‘Word+’ (“Word Plus”) formations. The Word+ category is “intended as a new morphological category that is larger than ordinary words but still belongs to the morphological as opposed to the syntactic domain” (Kageyama 2001: 253). Thus, the Word+ category is considered a “Janus-faced category at the intersection of morphological and syntactic structures” (Kageyama 2001: 274), and this intersective position of Word+ is also schematically illustrated (p. 273). Yet, Kageyama assigns the category a clearly morphological status, since the respective formations are word-level entities. However, although it is acknowledged that Word+ formations are formations with mixed properties, the proposal does not explain why morphology and syntax are still considered distinct domains. In particular, it does not explain how syntax and morphology interact – which they obviously do, given the properties of the Word+ category – and how syntactic principles can operate on the internal structure of the Word+ formations. So although the analysis provides a detailed description and inventory of the phenomenon, it does not provide a full explanation of the underlying processes and relations.

In Ackema & Neeleman’s (2004) framework, too, morphology and syntax are considered completely separate domains of grammar. However, the model allows for two kinds of interaction between these modules. The first one is insertion: Not only can morphological structure be integrated in syntax, but also vice versa. In this way, the model can account for phrasal compounds (cf. fn. 24), for instance – another problematic case for the idea of a strict demarcation between morphology or word-formation and syntax (cf. Ackema & Neeleman 2004: 122–129). In this case, a clearly syntactic object, a phrase, is inserted in either of the positions of a clearly morphological object, a compound. The second kind of interaction is competition between morphology and syntax. Generally, both in syntax and morphology, two elements can be combined (merged) to form a complex structure. According to Ackema & Neeleman, syntactic merging blocks morphological merging and vice versa, although, all things being equal, syntax takes precedence. Complex lexical items can be underspecified in their locus of realization,

that is, whether they are realized in syntax or morphological (Ackema & Neeleman 2004: 50ff). So, for instance, particle verbs are non-uniformly realized as either morphological or syntactic, depending on their actual use. However, although in both kinds of interaction the modules can “see” each other, they are nevertheless distinct and the principles of structure building do not mix. For this reason, the model does not allow for mixed properties within one structure (this is true also in the case of phrasal compounds, since a phrase is inserted as a block into a morphological structure and the morphological structure is blind to the internal structure of the phrase). Thus, the model cannot account for the mixed properties of the patterns discussed here, and for the gradient nature of morphological and syntactic properties of these patterns.

Giegerich’s (2015) proposal, finally, deals with the compound-phrase distinction in English. In view of the various observations regarding the inconclusive status of English compounds (cf. Section 1), Giegerich notes that a strict divide of the modules cannot be maintained. However, he does not do away with the idea of distinct modules. Instead, he concludes that these modules overlap and that there are forms that are simultaneously morphological (or lexical, in Giegerich’s terminology) and phrasal. This raises the question of whether such a proposal is different from the idea of a continuum between word-formation and syntax, as sketched in the previous sections.<sup>27</sup> A closer look reveals that they are in fact fundamentally different. The difference is less obvious as long as one only considers the form level: the overlap area in Giegerich’s model, and thus the intersection of compounds and phrases, comprises the forms with at least one morphological and one syntactic property. This is identical to the set of forms that are instantiated by the patterns in the middle part of the continuum in Table 1. However, the set of forms generated in the intersection area is quite heterogeneous as it contains all forms that are neither fully morphological or fully syntactic but have any combinations of at least one morphological or one syntactic property. The difference is even more significant when it comes to the formation process when word-formation and syntax are seen as processors for generating these forms. According to Giegerich (2015: 122), these forms “(...) are generated in an ‘area’ of the modularized derivation which is simultaneously part of the lexicon and the syntax, an area therefore where the two modules overlap with each other (...).” This means, however, that there is

---

27 Of course, Giegerich is only concerned with the English compound-phrase distinction, and not with the whole range of mixed patterns as discussed here, nor with other languages. Nevertheless, the question of a comparison of the models arises even if syntactic compounds are not taken in account.

one area where different, conflicting processes (e.g., the allocation of lexical or phrasal stress) are at work at the same time. It is, however, left completely open which of the processes would “win” under which circumstances.<sup>28</sup> On the continuum view, on the other hand, each construction is fully specified for its phonological, morphosyntactic and semantic properties. Each construction differs from the other constructions by at least one phonological, morphosyntactic or semantic property. At the same time, they can be conceived of as a set of closely related constructions that collectively form the transition from fully morphological to fully syntactic constructions. Such a view does justice to the English compound-phrase distinction but also covers the entire range of morphological and syntactic compounds.

## 5. Conclusion

The paper has presented cross-linguistic evidence from various languages that (1) there are compound patterns that – in addition to pertinent word-formation (compound) properties – may also have syntactic properties, and (2) there are phrasal patterns that instantiate lexical entities, on a par with morphological compounds, analysed here as syntactic compounds. Both the deviant compounds and phrasal patterns occur cross-linguistically and systematically. These observations are taken as evidence for the idea that there is no categorial distinction between compounds and phrases, or, more generally, between word-formation and syntax. This even holds for languages which, at first sight, have a categorial compound-phrase distinction, such as German. It has been argued that the observation can be more adequately accounted for if the relation between word-formation and syntax is regarded as a continuum of related, but distinct constructions, with fully morphological and syntactic constructions at the endpoints and various mixed constructions in between. Such a view on the grammar system is in line with constructionist approaches to the language system.

## References

- Ackema, Peter & Ad Neeleman. 2004. *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.
- Anderson, Stephen R. 1992. *A-Morphous morphology (Cambridge Studies in Linguistics 62)*. Cambridge: Cambridge University Press.
- Arnaud, Pierre J. L. 2015. 38. Noun-noun compounds in French. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-Formation*, vol.

---

28 For similar criticism and further details on Giegerich’s (2015) analysis see also Bell (2018).

- Vol. I, 673–687. Berlin, München, Boston: De Gruyter. doi:10.1515/9783110246254-040.
- Arndt-Lappe, Sabine. 2011. Towards an exemplar-based model of stress in English noun-noun compounds. *Journal of Linguistics* 47. 549–585. doi:10.1017/S0022226711000028.
- Bağrıaçık, Metin & Angela Ralli. 2015. Phrasal vs. morphological compounds: Insights from Modern Greek and Turkish. *STUF - Language Typology and Universals* 68(3). 323–357. doi:10.1515/stuf-2015-0016.
- Barz, Irmhild. 2007. Wortbildung und Phraseologie. In Harald Burger, Dimitrij Dobrovol'skij, Peter Kühn & Neal R. Norrick (eds.), *Phraseologie / Phraseology*, vol. 1, 27–36. Berlin & New York: de Gruyter.
- Bauer, Laurie. 1988. *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press.
- Bauer, Laurie. 1998. When is a sequence of two nouns a compound in English? *English Language and Linguistics* 2(1). 65–86.
- Bauer, Laurie. 2009a. 17. Typology of compounds. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Compounding (Oxford Handbooks in Linguistics)*, 343–356. Oxford: Oxford University Press.
- Bauer, Laurie. 2009b. 21. IE, Germanic: Danish. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Compounding*, 400–416. Oxford: Oxford University Press.
- Bauer, Laurie. 2017. *Compounds and Compounding (Cambridge Studies in Linguistics 155)*. Cambridge: Cambridge University Press. doi:10.1017/9781108235679.
- Bauer, Laurie. 2019. Compounds and multi-word expressions in English. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 45–68. Berlin & Boston: de Gruyter. doi:10.1515/9783110632446-002.
- Bell, Melanie J. 2011. At the boundary of morphology and syntax: Noun noun constructions in English. In Alexandra Galani, Glynn Hicks & George Tsoulos (eds.), *Morphology and its interfaces*, 137–168. Amsterdam: Benjamins.
- Bell, Melanie J. 2018. Lexical deconstruction – review of Giegerich, H., *Lexical structures: Compounding and the modules of grammar*, Edinburgh Studies in Theoretical Linguistics, vol. 1: Edinburgh: Edinburgh University Press, 2015. *Morphology* 28(2). 219–227. doi:10.1007/s11525-018-9324-4.
- Bell, Melanie J. & Ingo Plag. 2013. Informativity and analogy in English compound stress. *Word Structure* 6(2). 129–155. doi:10.3366/word.2013.0042.
- Bjarnadóttir, Kristín. 2002. A Short Description of Icelandic Compounds. Institute of Lexicography, ms. <https://notendur.hi.is/~kristinb/comp-short.pdf>.
- Bjarnadóttir, Kristín. 2017. Phrasal Compounds In Modern Icelandic With Reference To Icelandic Word Formation In General. In Carola Trips & Jaklin Kornfilt (eds.), *Further investigations into the nature of phrasal compounding*, 13–48. Berlin: Language Science Press. <https://zenodo.org/record/885105>.
- Booij, Geert. 2002. Constructional Idioms, Morphology, and the Dutch Lexicon. *Journal of Germanic Linguistics* 14(4). 301–329. doi:10.1017/S1470542702000168.

- Booij, Geert. 2009. Phrasal names: A constructionist analysis. *Word Structure* 2(2). 219–240. doi:10.3366/E1750124509000427.
- Booij, Geert. 2010. *Construction morphology*. Oxford & New York: Oxford University Press.
- Booij, Geert. 2019. Compounds and multi-word expressions in Dutch. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 95–126. Berlin & Boston: de Gruyter. doi:10.1515/9783110632446-004.
- Booij, Geert & Francesca Masini. 2015. The Role of Second Order Schemas in the Construction of Complex Words. In Laurie Bauer, Lívia Körtvélyessy & Pavol Štekauer (eds.), *Semantics of Complex Words*, vol. 3, 47–66. Cham: Springer International Publishing. doi:10.1007/978-3-319-14102-2\_4.
- Cetnarowska, Bożena. 2015. The lexical/phrasal status of Polish Noun+Adjective or Noun+Noun combinations and the relevance of coordination as a diagnostic test. *SKASE Journal of Theoretical Linguistics* 12(3). 142–170.
- Cetnarowska, Bożena. 2018. Phrasal Names in Polish: A+N, N+A and N+N Units. In Geert Booij (ed.), *The Construction of Words. Advances in Construction Morphology* (Studies in Morphology Series 4), 287–313. Cham: Springer International Publishing. doi:10.1007/978-3-319-74394-3\_11.
- Cetnarowska, Bożena. 2019. Compounds and multi-word expressions in Polish. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 279–306. Berlin, Boston: De Gruyter. doi:10.1515/9783110632446-010.
- Fernández-Domínguez, Jesús. 2019. Compounds and multi-word expressions in Spanish. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 189–220. Berlin, Boston: De Gruyter. doi:10.1515/9783110632446-007.
- Finkbeiner, Rita & Barbara Schlücker. 2019. Compounds and multi-word expressions in the languages of Europe. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 1–44. Berlin & Boston: de Gruyter. doi:10.1515/9783110632446-001.
- Giegerich, Heinz J. 2004. Compound or phrase? English noun-plus-noun constructions and the stress criterion. *English Language and Linguistics* 8(1). 1–24.
- Giegerich, Heinz J. 2005. Associative adjectives in English and the lexicon-syntax interface. *Journal of Linguistics* 41. 571–591.
- Giegerich, Heinz J. 2006. Attribution in English and the distinction between phrases and compounds. In Petr Rösl (ed.), *Englisch in Zeit und Raum – English in Time and Space: Forschungsbericht für Klaus Faiss*, 10–27. Trier: Wissenschaftlicher Verlag Trier.
- Giegerich, Heinz J. 2015. *Lexical structures: compounding and the modules of grammar* (Edinburgh Studies in Theoretical Linguistics 1). Edinburgh: Edinburgh University Press.
- Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Gunkel, Lutz, Adriano Murelli, Susan Schlotthauer, Bernd Wiese & Gisela Zifonun. 2017. *Grammatik des Deutschen im europäischen Vergleich: das Nominal*

- (*Schriften Des Instituts Für Deutsche Sprache* 14). Vol. Teilband 1: Funktionale Domänen, Wort und Wortklassen. Boston: de Gruyter.
- Gunkel, Lutz & Gisela Zifonun. 2011. Klassifikatorische Modifikation im Deutschen und Französischen. In Lavric, Eva, Pöckl, Wolfgang & Schallhart, Florian (eds.), *Comparatio delectat: Akten der VI. Internationalen Arbeitstagung zum romanisch-deutschen und innerromanischen Sprachvergleich: Innsbruck, 3.–5. September 2008*, 549–562. Frankfurt am Main: Lang.
- Hansen, Erik. 2005. The standard languages and their systems in the 20th century V: Danish. In Oskar Bandle, Kurt Braunmüller, Ernst Hakon Jahr, Allan Karker, Hans-Peter Naumann, Ulf Teleman, Lennart Elmevik & Gun Widmark (eds.), *The Nordic Languages* (HSK 22), vol. 2, 1626–1634. Berlin & Boston: de Gruyter. doi:10.1515/9783110197068-061.
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80. doi:10.1515/flin.2011.002.
- Hoekstra, Jarich. 2002. Genitive Compounds in Frisian as Lexical Phrases. *Journal of Comparative Germanic Linguistics* 6(2). 227–259. doi:10.1023/A:1023677825279.
- Hüning, Matthias. 2010. Adjective + Noun constructions between syntax and word formation in Dutch and German. In Alexander Onysko & Sascha Michel (eds.), *Cognitive Perspectives on Word Formation* (Trends in Linguistics. Studies and Monographs 221), 195–215. Berlin & New York: de Gruyter.
- Jackendoff, Ray. 2002. *Foundations of language*. Oxford: Oxford University Press.
- Jackendoff, Ray. 2008. Construction after construction and its theoretical challenges. *Language* 84(1). 8–28.
- Jackendoff, Ray. 2010. *Meaning and the Lexicon: The Parallel Architecture 1975–2010*. Oxford & New York: Oxford University Press.
- Kageyama, Taro. 2001. Word Plus: The Intersection of Words and Phrases. In Jeroen Weijer & Tetsuo Nishihara (eds.), *Issues in Japanese Phonology and Morphology*, 245–276. Berlin & Boston: de Gruyter. doi:10.1515/9783110885989.
- Klinge, Alex. 2006. The Origin of Weapons of Mass Destructions. In Henning Nølle, Irène Baron, Hanne Korzen, Iørn Korzen & Hendrik H. Müller (eds.), *Grammatica. Festschrift in honour of Michael Herslund*, 233–248. Berlin: Peter Lang.
- Köhnlein, Björn. 2019. Stress in nominal compounds. <http://www.taalportaal.org/taalportaal/topic/pid/topic-14020545850975812>. (accessed 27 December 2019)
- Koliopoulou, Maria. 2008. The loose multi-word compounds of Modern Greek under the prism of construction morphology. In Nikolaos Lavidas, Elissavet Nouchoutidou, & Marietta Sionti (eds.), *New perspectives in Greek linguistics*, 213–224. Cambridge: Cambridge Scholars Publishing.
- Koptjevskaja-Tamm, Maria. 2009. Proper-name compounds in Swedish between syntax and lexicon. *Rivista di Linguistica / Italian Journal of Linguistics* 21. 119–148.

- Lapointe, Steven. 1980. *A Theory of Grammatical Agreement*. Amherst: University of Massachusetts PhD dissertation.
- Lees, Robert B. 1960. *Grammar of English nominalizations*. Bloomington: Indiana University Press.
- Liberman, Mark & Richard Sproat. 1992. The stress and structure of modified noun phrases in English. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical Matters*, 131–182. Stanford CA: Center for Study of Language and Information.
- Lieber, Rochelle & Sergio Scalise. 2007. The Lexical Integrity Hypothesis in a New Theoretical Universe. In Geert E. Booij, Luca Ducceschi, Bernard Fradin, Emiliano Guevara, Angela Ralli & Sergio Scalise (eds.), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5) 2005*, 1–24. University of Bologna. [http://www.lilec.it/mmm/wp/?page\\_id=71](http://www.lilec.it/mmm/wp/?page_id=71).
- Lieber, Rochelle & Pavol Štekauer. 2009. Introduction: Status and definition of compounding. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Compounding* (Oxford Handbooks in Linguistics), 3–18. Oxford: Oxford University Press.
- Masini, Francesca. 2009. Phrasal lexemes, compounds and phrases: A constructionist perspective. *Word Structure* 2(2). 254–271. doi:10.3366/E1750124509000440.
- Masini, Francesca. 2019a. Competition Between Morphological Words and Multi-word Expressions. In Franz Rainer, Francesco Gardani, Wolfgang U. Dressler & Hans Christian Luschützky (eds.), *Competition in Inflection and Word-Formation*, 281–305. Cham: Springer International Publishing. doi:10.1007/978-3-030-02550-2\_11.
- Masini, Francesca. 2019b. Compounds and multi-word expressions in Italian. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 153–188. Berlin & Boston: de Gruyter. doi:10.1515/9783110632446-006.
- Masini, Francesca. 2019c. Multi-Word Expressions and Morphology. *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. doi:10.1093/acrefore/9780199384655.013.611.
- Masini, Francesca & Jenny Audring. 2019. Construction Morphology. In Jenny Audring & Francesca Masini (eds.), *The Oxford Handbook of Morphological Theory*, 364–389. Oxford University Press. doi:10.1093/oxfordhb/9780199668984.013.25.
- Masini, Francesca & Valentina Benigni. 2012. Phrasal lexemes and shortening strategies in Russian: the case for constructions. *Morphology* 22(3). 417–451. doi:10.1007/s11525-011-9200-y.
- Montermini, Fabio. 2006. A new look on word-internal anaphora on the basis of Italian data. *Lingue e Linguaggio* 5(1). 127–148.
- Nicoladis, Elena. 2002. When Is a Preposition a Linking Element? Bilingual Children's Acquisition of French Compound Nouns. *Folia Linguistica* 36(1–2). doi:10.1515/flin.2002.36.1-2.45.
- Niemi, Jussi. 2009. Compounds in Finnish. *Lingue e Linguaggio* 8(2). 237–256.

- Ohnheiser, Ingeborg. 2015. Compounds and multi-word expressions in Slavic. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-Formation. An International Handbook on the Languages of Europe.*, vol. 1, 757–779. Berlin & Boston: de Gruyter.
- Ohnheiser, Ingeborg. 2019. Compounds and multi-word expressions in Russian. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 251–278. Berlin & Boston: de Gruyter. doi:10.1515/9783110632446-009.
- Olsen, Susan. 2000. Compounding and Stress: A closer look at the boundary between morphology and syntax. *Linguistische Berichte* 181. 55–69.
- Payne, John & Rodney Huddleston. 2002. Nouns and noun phrases. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge Grammar of the English Language*, 323–451. Cambridge: Cambridge University Press.
- Pitkänen-Heikkilä, Kaarina. 2016. 176. Finnish. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-Formation (HSK 40)*, vol. V, 3209–3228. Berlin & Boston: de Gruyter. doi:10.1515/9783110424942-008.
- Plag, Ingo. 2006. The variability of compound stress in English: structural, semantic, and analogical factors. *English Language and Linguistics* 10(1). 143–172.
- Plag, Ingo, Gero Kunter & Sabine Lappe. 2007. Testing hypotheses about compound stress assignment in English: a corpus-based investigation. *Corpus Linguistics and Linguistic Theory* 3(2). 199–232.
- Postal, Paul M. 1969. Anaphoric Islands. *Papers from the Fifth Regional Meeting*, 205–239. Chicago: Chicago Linguistic Society.
- Rainer, Franz & Soledad Varela. 1992. Compounding in Spanish. *Rivista di Linguistica* 4(1). 117–142.
- Ralli, Angela. 2008. Compound markers and parametric variation. *STUF – Language Typology and Universals* 61(1). 19–38. doi:10.1524/stuf.2008.0004.
- Ralli, Angela. 2013a. *Compounding in modern Greek* (Studies in Morphology 2). Dordrecht: Springer.
- Ralli, Angela. 2013b. Compounding and its locus of realization: Evidence from Greek and Turkish. *Word Structure* 6(2). 181–200. doi:10.3366/word.2013.0044.
- Ralli, Angela. 2016. 172. Greek. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-Formation (HSK 40)*, vol. V, 3138–3156. Berlin & Boston: de Gruyter. doi:10.1515/9783110424942-004.
- Rosenbach, Anette. 2006. Descriptive genitives in English: a case study on constructional gradience. *English Language and Linguistics* 10. 77–118.
- Sadler, Louisa & Douglas J. Arnold. 1994. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* 30. 187–226.
- Schlücker, Barbara. 2012. Die deutsche Kompositionsfreudigkeit. Übersicht und Einführung. In Livio Gaeta & Barbara Schlücker (eds.), *Das Deutsche als kompositionsfreudige Sprache. Strukturelle Eigenschaften und systembezogene Aspekte*, 1–25. Berlin & New York: de Gruyter.



- Schlücker, Barbara. 2013. Non-classifying compounds in German. *Folia Linguistica* 47(2). 449–480. doi:10.1515/flin.2013.017.
- Schlücker, Barbara. 2014. *Grammatik im Lexikon. Adjektiv+Nomen-Verbindungen im Deutschen und Niederländischen* (Linguistische Arbeiten 553). Berlin & Boston: de Gruyter.
- Schlücker, Barbara. 2016. Adjective-noun compounding in Parallel Architecture. In Pius ten Hacken (ed.), *The semantics of compounding*, 178–191. Cambridge: Cambridge University Press.
- Schlücker, Barbara. 2017. Eigennamenkomposita im Deutschen. In Johannes Helmbrecht, Damaris Nübling & Barbara Schlücker (eds.), *Namengrammatik* (Linguistische Berichte Sonderheft 23), 59–93. Buske.
- Schlücker, Barbara. 2018. Genitives and proper name compounds in German. In Tanja Ackermann, Horst J. Simon & Christian Zimmer (eds.), *Germanic Genitives* (Studies in Language Companion Series 193), 275–299. Amsterdam & Philadelphia: John Benjamins Publishing Company. doi:10.1075/slcs.193.10sch.
- Schlücker, Barbara. 2019. Compounds and multi-word expressions in German. In Barbara Schlücker (ed.), *Complex lexical units. Compounds and multi-word expressions*, 69–94. Berlin & Boston: de Gruyter. doi:10.1515/9783110632446-003.
- Schlücker, Barbara & Matthias Hüning. 2009. Compounds and phrases. A functional comparison between German A+N-compounds and corresponding phrases. *Italian Journal of Linguistics / Rivista di Linguistica* 21(1). 209–234.
- Spencer, Andrew. 2003. Does English have productive compounding? In Geert Booij, J. DeCesaris, Angela Ralli & Sergio Scalise (eds.), *Topics in morphology: selected papers from the Third Mediterranean morphology meeting (Barcelona, September 20–22 2001)*, 329–341. Barcelona: IULA.
- Sproat, Richard. 1988. On anaphoric islandhood. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology: approaches in modern linguistics*, 291–301. San Diego: Academic Press.
- Szymanek, Bogdan. 2009. IE, Slavonic: Polish. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Compounding* (Oxford Handbooks in Linguistics), 464–477. Oxford: Oxford University Press.
- Teleman, Ulf. 2005. The standard languages and their systems in the 20th century IV: Swedish. In Oskar Bandle, Kurt Braunmüller, Ernst Hakon Jahr, Allan Karker, Hans-Peter Naumann, Ulf Teleman, Lennart Elmevik & Gun Widmark (eds.), *The Nordic Languages* (HSK 22), vol. 2, 1603–1626. Berlin & Boston: de Gruyter. doi:10.1515/9783110197068-060.
- Tummers, José. 2005. *Het naakt(e) adjectief. Kwantitatief-empirisch onderzoek naar de adjectivische buigingsalternantie bij neutra*. Leuven: Katholieke Universiteit Leuven Proefschrift.
- Van Goethem, Kristel. 2009. Choosing between A+N compounds and lexicalized A+N phrases: The position of French in comparison to Germanic languages. *Word Structure* 2(2). 241–253. doi:10.3366/E1750124509000439.

- Van Goethem, Kristel & Dany Amiot. 2019. Compounds and multi-word expressions in French. In Barbara Schlücker (ed.), *Complex Lexical Units*, 127–152. Berlin & Boston: de Gruyter. doi:10.1515/9783110632446-005.
- Visser, Willem & Joke Weening. 2019. Stress in nominal compounds. <http://www.taalportaal.org/taalportaal/topic/pid/topic-14020545850975822>. (accessed 27 December 2019)
- Ward, Gregory, Richard Sproat & Gail McKoon. 1991. A pragmatic analysis of so-called anaphoric islands. *Language* 67(3). 439–474.

Barbara Schlücker  
Universität Leipzig  
Institut für Germanistik  
Beethovenstr. 15  
04107 Leipzig  
barbara.schluecker@uni-leipzig.de

## Varia / Miscellaneous

# Tagungsbericht / Conference Report

## “DeriMo 2019. Second Workshop on Resources and Tools for Derivational Morphology”

### 19–20 September 2019; Prague, Czech Republic (Petra Steiner)

DeriMo is an international meeting dealing with derivational morphology from the perspective of data analysis. Its second edition DeriMo 2019 was held at the Faculty of Mathematics and Physics, Charles University in Prague. The local organizers are researchers of the Institute of Formal and Applied Linguistics (ÚFAL = Ústav Formální a Aplikované Linguistiky) at the Computer Science School of the Faculty of Mathematics and Physics. Chairs of the program committee were Magda Ševčíková (ÚFAL), Zdeněk Žabokrtský (ÚFAL), Eleonora Litta Modignani Picozzi (CIRCSE, Università Cattolica del Sacro Cuore, Milan), and Marco Passarotti (CIRCSE, Università Cattolica del Sacro Cuore, Milan).

ÚFAL with its strong and long tradition from Prague School and functionalism is dedicated to approaches of Natural Language Processing based on theoretical frameworks such as Dependency Grammar or Quantitative Linguistics. This institutional background and the work on morphological tools and dictionaries were the pivots of the workshop. One of the latest developments at ÚFAL is DeriNet, a network of word-formation for the Czech language. It is one of 22 databases with derivational information listed by Kyánek (2018). 17 of these and many other derivational resources were dealt with in the presentations. The contributions presented recent developments in word-formation resources and tools, research by using morphological data, and theoretical accounts and descriptions of word-formation processes.

This workshop gathered researchers from typology, computational linguistics, classical philology, and grammatical description frameworks. The languages covered included Bulgarian, Croatian, Czech, English, Estonian, Farsi, French, German, Japanese, Latin, Polish, Portuguese, and Spanish, besides many others in cross-linguistic studies. These two days' talks dealt with a wealth of morphological data, tools, and methods not just for applications in natural language processing but also for research in morphology, typology, semantics, quantitative linguistics, language teaching, or speech therapy. Researchers with a need for high-quality morphological data should have a look at this workshop's proceedings to get an overview of

the state of the art. Most of the presented sources are publicly available. The workshops comprised 14 contributions of 24 authors. The papers, slides and one handout are available at the conference website: <http://ufal.mff.cuni.cz/derimo2019/>.

The first of two invited talks, *Cross-linguistic research into derivational networks* by Livia Körtvélyessy (P.J.Šafárik University, Košice, Slovakia), introduced to derivational networks that are multi-dimensional description frameworks for derivational processes. Körtvélyessy's project aims at comparing the derivational complexity of 40 European languages. Chains of consecutive derivations yield different levels of forms and semantic categories. The study aims at characterizing their derivational and paradigmatic properties by defining measures: The paradigmatic capacity of a word is characterized by the number of its potential derivations, the number of consecutive morphs is defined as the order. Consecutive semantic categories show the semantic shifts of the derivational processes. Another characteristic measure is the Maximum Derivational Network (MDN), which is the sum total of maximums of the derivatives of all semantic categories. The MDN is a measure for the derivational potential. The saturation value is the ratio of the number of actual derivatives and the MDN; it describes the realization of the potential. These values were calculated for each order and the sums of all orders. Only six of the 40 languages reach the maximum order of 5 for all parts of speech, among them Czech with its complex derivational morphology. German comes only to an order of 3, with a high saturation value of 45.45 compared to the Czech value of 27.91. Furthermore, the investigation shows that the choice of semantic categories depends on the order and some language-specific blocking effects. This ongoing research gives new perspectives on cross-linguistic patterns of word-formation.

Daniele Sanacore, Nabil Hathout, and Fiammetta Namer (University of Toulouse/University of Lorraine, France) presented *Semantic descriptions of French derivational relations in a families-and-paradigms framework*. They link data of Démonette, a derivational database for French (Hathout & Namer 2014), with the semantic database FrameNet by building frame-like structures for derivational paradigms. By this, they show that at least some FrameNet frames can be adapted for describing morphosemantic relations of derivational families.

Lucie Pultrová's (Charles University, Prague) contribution *Correlation between the gradability of Latin adjectives and the ability to form qualitative abstract nouns* provides strong evidence for the correlation between the gradability and the ability of adjectives to form abstract nouns by investigating the suffixes *-tas*, *-itia*, *-tudo* and *-ia* in the database Bibliotheca Teubneriana Latina III.

In *The Treatment of Word Formation in the LiLa Knowledge Base of Linguistic Resources for Latin*, Eleonora Litta, Marco Passarotti, and Francesco Mambrini (Università Cattolica del Sacro, Milan, Italy) described the union of linguistic resources and tools for Latin within a new knowledge base. Here, all entries in lexical resources and corpus tokens that refer to the same lemma are linked within the Word Formation Latin Lexicon.

In her talk *Combining Data-Intense and Compute-Intense Methods for Fine-Grained Morphological Analyses*, Petra Steiner (Friedrich Schiller University of Jena, Germany) reported on a hybrid approach for analyzing German derivational and compositional morphology. This combines deep-level morphological structures of a database derived from GermaNet and CELEX, the results of a conventional word segmenter, and frequencies of constituents within contexts drawn from a large Wikipedia corpus.

The talk by Klára Osolobě (Masaryk University, Brno) on *The Tagged Corpus (SYN2010) as a Help and a Pitfall in the Word-formation Research* describes the problems the SYN2010 corpus caused for the Dictionary of Affixes used in Czech (Slovník afixů užívaných v češtině; SAUČ). Especially graphical variants and overgeneration of morphological analyses lead to wrong entries.

In *Attempting to separate inflection and derivation using vector space representations*, Rudolf Rosa, and Zdeněk Žabokrtský (Charles University) explored the possibility of automatically classifying inflected and derived word forms. Their work builds on the working assumption of a clear-cut boundary between these categories and the contextual notion of meaning. The experiments used word embedding similarities, edit distance measures, and the combinations of both. The main hypotheses were that inflected forms are more similar and semantically closer than derived forms. The test data consists of 69,743 word forms of derivational families with 4,514 lemmas taken from DeriNet. With F1 scores around 40%, close inflections and derivations could be distinguished from a large set of different forms. A further analysis of particularly difficult cases shows that cos distance clearly separates inflected from derived forms.

The talk *Redesign of the Croatian derivational lexicon (CroDeriV)* by Matea Filko, Krešimir Šojat, and Vanja Štefanec (University of Zagreb, Croatia) describes their diligent lexicographic work in developing and enriching CroDeriV, which they expanded by 6,000 nouns and 1,000 adjectives. The morphological structures were manually analyzed by a two-level morphology, and the allomorphs of stems and affixes annotated. Other information was automatically determined, such as the base or root of the word, and the type of the word-formation process.

The last talk of the first day was from Jonáš Vidra, Zdeněk Žabokrtský, Magda Ševčíková, and Lukáš Kyjánek (Charles University). In *DeriNet 2.0*:

*Towards an All-in-One Word-Formation Resource*, the new annotation scheme of the database was introduced: a new file format permits storing new features and detailed description: especially compounds can now be described as entities with two parents. Furthermore, information about morphological categories such as aspect, gender and animacy, five semantic labels (diminutive, possessive, female, iterative, and aspect), and pseudo morphemes and bound root morphemes can be added to the knowledge base.

The second day began with the second invited talk by Fiammetta Namer and Nabil Hathout (Nancy/Toulouse), *ParaDis and Démonette, From Theory to Resources for Derivational Paradigms*. Following the prototype version *Démonettev1*, *Démonettev2* implements ParaDis (“Paradigms vs Discrepancies”), a paradigmatic model for representing morphologically complex lexical units by merging principles from lexeme-based and paradigm-based approaches. This permits aligning lexemes to sets of derivational families by multiple connections even if the relations of their formal and semantic levels are not isomorphic, as for the affix replacement for *banque* (‘bankN’), *bancaire* (‘bankAdj’), and *interbancaire* (‘interbankAdj’). The entries are sets of relations between lexemes with annotations.

In *Building a Morphological Network for Persian on Top of a Morpheme-Segmented Lexicon*, Hamid Haghdoost, Ebrahim Ansari, Zdeněk Žabokrtský and Mahshid Nikravesh (Charles University, Prague) presented two novelties: The first is a new large segmented morphological lexicon of Persian with 45,000 entries. As Persian is a morphologically rich language with many derivational affixes, this required concerted work of manual annotation and correction of the preprocessed word lists. Secondly, the word segmentation lexicon was used for building a morphological network analogous to DeriNet – DeriNet.FA. The authors experimented with weaving new and unprocessed words into the network by combining the morphological tool Morfessor with a recursive algorithm.

Lukáš Kyjánek, Zdeněk Žabokrtský, Magda Ševčíková and Jonáš Vidra (Charles University Prague) described the development of a multilingual derivational database in their talk *Universal Derivations Kickoff: A Collection of Harmonized Derivational Resources for Eleven Languages*. The collection under the title Universal Derivations (UDer 0.5) comprises a harmonization of the data from *Démonette*, DeriNet, DeriNet.ES, DeriNet.FA, DERivBase, English WordNet, EstWordNet, FinnWordNet, NomLexPT, The Polish Word-Formation Network, and Word Formation Latin. The different schemas are unified as rooted trees under DeriNet 2.0 format, other structures are represented in secondary sets. UDer combines networks of different quality and depth. For instance, DERivBase 2.0 has an average number of nodes of 1.2 per tree compared to 4.7 for DeriNet. Most derivational trees of this German part consist of singleton nodes.

Christian Curtis's (University of Washington, USA) work on *A Parametric Approach to Implemented Analyses: Valence-changing Morphology in the LinGO Grammar Matrix* is rooted in the Head-Driven Phrase Structure Grammar (HPSG). The object of investigation were verbal derivations that alter the argument structure of verbs by increasing or decreasing their valency. Curtis augmented the HPSG rule system by attaching valency-changing operations to lexical rule types. Examples and evaluation were among others on Mayan and Bantu languages, Javanese, Lakota, Japanese, and Hungarian.

The final talk of the workshop *Grammaticalization in Derivational Morphology: Verification of the Process by Innovative Derivatives* was presented by Junya Morita (Kinjo Gakuin University, Nagoya, Japan). It dealt with a contrastive investigation of meaning shifts of Japanese and English affixes such as *-er* and *-ee*. Objects were hapax legomena and other nonce words from BNC and BCCWJ (Balanced Corpus of Contemporary Written Japanese). While English derivational suffixes are ambiguous concerning the semantic roles of agent and instrument, their Japanese counterpart are not.

Kyjánek, Lukáš. 2018. Morphological Resources of Derivational Word-Formation Relations. Techn. Report. Prague: Charles University. ÚFAL. TR-2018-61.

Dr. Petra C. Steiner  
Friedrich-Schiller-Universität Jena  
Faculty of Arts  
Institute of German Linguistics  
Fürstengraben 30  
07743 Jena / Germany  
petra.steiner@uni-jena.de



# Tagungsankündigung / Conference Announcement

## “Word-Formation Theories IV / Typology and Universals in Word-Formation V” (June 23–26, 2021; Kosice, Slovakia)

Contact Person: Pavol Stekauer  
Meeting Email: [pavel.stekauer@upjs.sk](mailto:pavel.stekauer@upjs.sk)

Linguistic Field(s): Linguistic Theories; Morphology; Typology

### Meeting Description:

Another edition of our 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.

### Call for Papers:

Abstract of max. 300 words should be emailed to [livia.kortvelyessyupjs.sk](mailto:livia.kortvelyessyupjs.sk)

### Deadlines:

Submission of abstracts: 15 April, 2021

Notification of acceptance: 30 April, 2021

Submission of a registration form: 31 May, 2021

### Plenary speakers (in alphabetical order):

Mark Aronoff,

Stony Brook University, New York, USA

Jan Don, Amsterdam University, The Netherlands

Martin Everaert, Utrecht University, The Netherlands

Volker Gast, Friedrich Schiller University, Jena

Gary Libben, Brock University, USA

Brian MacWhinney, Carnegie Mellon University, USA