Maximilian Frankowsky

N+N-Komposita mit identischen Konstituenten im Deutschen

Reihe Germanistische Linguistik

Herausgegeben von
Noah Bubenhofer und Britt-Marie Schuster

330

Maximilian Frankowsky

N+N-Komposita mit identischen Konstituenten im Deutschen

Theorie und Empirie zu reduplikativer Komposition

DE GRUYTER

Zugl.: Dissertation, Universität Leipzig, 2023

Diese Publikation wurde unterstützt durch den Open-Access-Publikationsfonds der Universität Leipzig.

ISBN 978-3-11-131496-9
e-ISBN (PDF) 978-3-11-131541-6
e-ISBN (EPUB) 978-3-11-131575-1
ISSN 0344-6778
DOI https://doi.org/10.1515/9783111315416

Library of Congress Control Number: 2024942035

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.dnb.de abrufbar.

Druck und Bindung: CPI books GmbH, Leck

www.degruyter.com

Vorwort

Gleich und gleich gesellt sich gern.
Homer

Habe keinen Freund, der Dir nicht gleich ist.
Konfuzius

Gleiche kann man mit Gleichen sehr leicht zusammenscharen.
Cicero

Die sich gleichen, werden ein Paar.
Japanisches Sprichwort

Der Mensch scheint überzeugt davon zu sein, dass Gleiches gut zusammenpasst. Die gesamte Wissenschaft beruht auf diesem Grundsatz. Die Setzung der Eins und ihre Abgrenzung zur Null gehen damit einher, dass man Gruppen bildet aus Dingen, Elementen oder Seinszuständen, die sich gleichen. Die Guten ins Töpfchen, die Schlechten ins Kröpfchen. Das Zusammengruppieren von Gleichem geht wiederum damit einher, dass es etwas gibt, das dem, was da zusammengruppiert wird, nicht gleicht. Der Satz „A=A“ – nach Heidegger das oberste Denkgesetz – ergibt nur dann Sinn, wenn auch „A≠B“ gilt. Identität und Differenz, das Gleichsein und das Ungleichsein, sind also untrennbar miteinander verbunden. Gäbe man einfach alle Linsen ins Töpfchen, wäre am Ende nichts gewonnen.

Die Sprache als die wohl wichtigste Begleiterscheinung menschlicher Kultur beruht ebenfalls auf der Dialektik von Identität und Differenz. Einerseits ist Identität von sprachlichen Einheiten eine absolute Notwendigkeit, da sonst jedes Wort neu und den Gesprächspartner:innen unbekannt wäre. Andererseits müssen sprachliche Einheiten in einem gewissen Maße unterschiedlich sein, sonst gäbe es nur ein einziges Wort und effektive Kommunikation wäre unmöglich.

Für besonders effektive Kommunikation haben häufig die Mitarbeiter:innen von Werbeagenturen ein gutes Gespür. Auch das in dieser Arbeit beschriebene Phänomen gehört zum Werkzeugkasten der Werbebranche: Die Zusammenfügung identischer Wörter (Abbildung 1).

Die Originalversion dieses Kapitels wurde revidiert: Seite V des Vorworts wurde korrigiert. Ein Erratum ist verfügbar unter: https://doi.org/10.1515/9783111315416-014

https://doi.org/10.1515/9783111315416-203

Abb. 1: Werbedisplay an der Leipziger S-Bahn-Haltestelle „Markt“.

Sommer Sommer in Abbildung 1 beinhaltet eine unmittelbare Wiederholung. Identisches wird zusammengefügt. Für die deutsche Sprache ist das ungewöhnlich, denn entgegen den oben aufgeführten Zitaten gesellt sich gleich und gleich im Deutschen nur ungern. Besser gesagt: Die Sprecher:innen des Deutschen mögen keine unmittelbaren Wiederholungen und vermeiden tunlichst alles Redundante; jedenfalls in der Sprache.

Auch Linguist:innen stellt eine Bildung wie *Sommer Sommer* vor Probleme. Wie kommen solche Ausdrücke zustande? Ist es eine syntaktische Bildung oder eine morphologische? *Sommer* und *Sommer* sind graphisch voneinander getrennt. Zwischen ihnen steht ein Leerzeichen, beziehungsweise ein Zeilenumbruch. Für gewöhnlich spricht das dagegen, sie als Bestandteile eines Wortes aufzufassen. Andererseits besetzen die Bestandteile gemeinsam eine syntaktische Position. Bilden sie also doch ein Wort? Sind Bildungen wie *Sommer Sommer* etwas besonderes oder genauso gewöhnlich wie Komposita aus ungleichen Nomina, etwa *Dürresommer*? Einerseits ist es natürlich extravagant, wenn ein Nominalstamm mit sich selbst kombiniert wird. Andererseits wäre es aber auch etwas besonderes, wenn genau dieser Nominalstamm nicht für ein Kompositum mit *Sommer* zur Verfügung stünde.

Bildungen wie *Sommer Sommer* werfen viele weitere Fragen auf: Was bedeutet es überhaupt, *mehr Sommer Sommer* zu haben und wie verstehen die Empfän-

ger:innen der Werbebotschaft in Abbildung 1, was mit *Sommer Sommer* gemeint ist? Verstehen sie es überhaupt? Verwenden sie solche Bildungen selbst oder werden solche Ausdrücke allein von Germanistikabsolvent:innen gebildet, die es in den Marketingbereich verschlagen hat?

Ich habe mich in den letzten Jahren intensiv mit den genannten Fragen beschäftigt und so manchen *Sommer Sommer* der Fertigstellung meiner Dissertation gewidmet, aus der das Buch, das Sie gerade lesen, hervorgegangen ist. Diese Dissertation habe ich Ende 2016 an der Rheinischen Friedrich-Wilhelms-Universität Bonn begonnen und im November 2021 an der Universität Leipzig eingereicht. Ich hätte mich dieser Arbeit niemals so erfolgreich widmen können, wenn ich nicht viele liebe Menschen um mich gehabt hätte, die mich dabei unterstützt haben. Bei diesen Menschen möchte ich mich deshalb herzlich bedanken. An erster Stelle ist hier Barbara Schlücker zu nennen, die – da sind wir uns alle einig – für mich die ideale Betreuerin war, weil ich nirgends mehr über Morphologie und Komposita hätte lernen können, und weil sie *fördern* hin und wieder auch mal ohne Umlaut schreibt. Auch möchte ich mich bei Rita Finkbeiner bedanken, die meine Zweitbetreuerin war und das Thema ICCs für mich schon wunderbar vorbereitet hat. Für den Druckkostenzuschuss zur Publikation dieses Buches möchte ich mich beim Open Science Office der Universität Leipzig bedanken, ohne das das Buch wohl sehr viel weniger Leser:innen erreichen würde. Für die ideelle Förderung bedanke ich mich beim Graduiertenkolleg *Interaktion Grammatischer Bausteine* (IGRA), insbesondere bei Gereon Müller und Barbara Stiebels. Ein besonderer Dank gebührt auch Roland Schäfer und Felix Bildhauer sowie Adam Kilgarriff, Pavel Rychly, Pavel Smrz und David Tugwell, die mit DECOW, respektive deTenTen Ressourcen geschaffen haben, die für meine Arbeit von zentraler Bedeutung sind. Noreen Sell und Anton Gerasimovich haben mir dabei geholfen, mich durch die entsprechenden Daten zu kämpfen, wofür ich ihnen ganz herzlich danken möchte. Elisabeth Stanciu und Albina Töws danke ich für die Hilfe bei der Drucklegung dieses Buches. An der sehr angenehmen Atmosphäre und dem regen wissenschaftlichen Austausch während des gesamten Dissertationsprojektes haben die Mitglieder der verschiedenen Arbeitsgruppen an der Uni Bonn und an der Uni Leipzig ihren Anteil. Hier sind insbesondere Adele Baltuttis, Thomas Bertram, Anna Bliß, Eva Büthe-Scheider, Sandra Döring, Julia Fuchs, Laura Hüser, Eva Kosmata, Robert Külpmann, Karen Lehmann, Marianna Lohmann, Matthias Richter, Jan Seifert, Katerina Stathi, Naomi Truan, Aleksandra Uttenweiler, Diana Walther und Claudia Wich-Reif zu nennen, die allesamt auf unterschiedliche Art und Weise positiv auf diese Arbeit eingewirkt haben. Wertvolle Kommentare verdanke ich außerdem Jenny Audring, Martin Haspelmath und Thomas Stolz. Ich danke ferner Nico Schäfer, Thomas Goik, Robin Schweiger, Mats Middelberg und

der erweiterten Hooked-Community für zahlreiche (Hör-)belege sowie für den notwendigen kontrastiven Fokus, für den ich ebenso dem Leipziger Honigdachs sowie Satoshi Nakamoto und Keith Gill danke. Meinem Bruder Sven sowie Christian Forche bin ich für vieles dankbar, etwa dafür, dass sie sich die frühe Alphaversion dieses Buches angetan haben. Ausgesprochen dankbar bin ich schließlich auch meinen Eltern Doris und Rudolf, auf die ich mich immer verlassen kann. Meiner Frau Dan sowie meinen beiden Kindern danke ich in besonderer Verbundenheit für alles.

Berlin, im Mai 2024
Maximilian Frankowsky

Inhalt

Teil III: Theorie zu ICCs im Deutschen

Abbildungs- und Tabellenverzeichnis

https://doi.org/10.1515/9783111315416-205

Teil I: **Das Phänomen ICC**

1 Einleitung

1.1 N+N-Komposita mit identischen Konstituenten

N+N-Komposita sind Komposita, bei denen beide Konstituenten Nominalstämme sind. Es kommt bisweilen vor, dass diese beiden Konstituenten identisch sind. Solche Bildungen werden identical constituent compounds, oder kurz: ICCs, genannt (Hohenhaus 2004). In den folgenden drei Beispielen für ICCs werden Erst- und Zweitglied durch denselben Stamm realisiert:

(1) *Es hat sich aber herausgestellt, dass mein Spottini eher ein "Menschenhund" und nicht ein "**Hundehund**" ;-)) ist. D.h. er möchte eigentlich gar kein Rudel haben sondern "seine(n)" Menschen ganz für sich.*

<www.beagleweb.ch/de/spotty.htm>

(2) *Und hier klingt's halt so als hätte er schon gerne 'ne **Beziehung-Beziehung** und nicht Hauptsache irgendjemanden.*

<www.superkreuzburg.de/die-ratsherren-47, 27:30>

(3) *Durch meinen Facebookaufruf habe ich noch zwei weitere Hersteller erfahren, nämlich **fotofoto** und happyfoto, doch ich kann von beiden nicht aus eigener Erfahrung erzählen.*

<www.elena-zeitler.de/fotobuch-drucken>

1.2 Problemstellung und Zielsetzung

ICCs sind grammatische Sonderfälle, denn sie haben eine für das Deutsche sehr unübliche reduplikative Struktur. Für die Grammatikbeschreibung stellen sie deshalb ein Problem dar. Viele Kompositatheorien setzen diskrete Konstituenten voraus. ICCs werden zudem in der grammatiktheoretischen Literatur kontrovers diskutiert und dabei mal als Komposita, mal als Reduplikationen angesehen. Auch ist es bisher nicht gelungen, die Bedeutungskonstitution von ICCs angemessen darzustellen. Einschlägige Modelle, die die Transparenz kanonischer Komposita angemessen modellieren, machen für ICCs höchst unwahrscheinliche Voraussagen. Darüber hinaus ist man sich in der linguistischen Forschung nicht darüber einig, welche Rolle der Kontext für ICCs spielt. Der These, dass ICCs ihre Bedeu-

https://doi.org/10.1515/9783111315416-001

tung mit formalen Mitteln anzeigen, steht die Annahme absoluter Kontextabhängigkeit dieser Bildungen gegenüber. Die empirische Grundlage, auf der die jeweiligen Argumentationen beruhen, ist – sofern überhaupt vorhanden – meist nicht sehr solide. Viele Annahmen zu ICCs wurden bislang noch überhaupt nicht empirisch überprüft. All diese Problemstellen sind Ausgangspunkte der vorliegenden Arbeit.

Zunächst soll die Arbeit die bereits angesprochene Datenlücke füllen und eine umfassende Datenbasis bieten, auf der ICCs beschrieben und eingeordnet werden können. Ziel dieser Arbeit ist weiterhin, die Formen und Funktionen von ICCs zu beschreiben und im Anschluss daran zu überprüfen, inwieweit Ansätze der Grammatikbeschreibung ICCs angemessen darstellen können. Neben dieser Frage gehe ich folgenden Forschungsfragen nach:

– Wie produktiv ist die Bildung von ICCs im Deutschen?
– Sind ICCs Komposita?
– Wie erhalten ICCs ihre Bedeutung?

1.3 Abgrenzung des Untersuchungsgegenstandes

Der Untersuchungsgegenstand dieser Arbeit sind morphologische Bildungen, deren Konstituenten von ein und demselben Nominalstamm realisiert werden. Zu syntaktischen Bildungen, in denen zwei identische Nomina direkt aufeinander folgen (*wenn Kinder Kinder kriegen)* oder Bildungen, in denen die Bestandteile nicht nominal sind (*zum letztenletzten Abschluss*), werden keine Aussagen getroffen. Sie werden lediglich bei der Beschreibung der Datenkodierung erwähnt. Auch Verbindungen, in denen die identischen Nominalstämme nicht adjazent sind, wie die N-P-N-Konstruktion (*Stück für Stück*, *Tag für Tag* [Jackendoff 2008]), sind nicht Teil dieser Untersuchung.

Ferner ist wichtig zu erwähnen, dass Identität der Konstituenten in dieser Arbeit immer die Identität von Semantik und Form meint. Ausgeschlossen von der vorliegenden Untersuchung sind also sowohl Bildungen, bei denen die Konstituenten nur eine gleiche Semantik haben (*Einzelindividuum*), als auch solche, in denen lediglich die (graphematische) Form gleich ist (*August-August* ‘ein im August geborener Mensch namens August’).

Diese Arbeit behandelt auch keine Komposita, in denen zwei Konstituenten zwar bedeutungs- und formgleich sind, aber in einem mehr als zwei Konstituenten umfassenden Kompositum auftreten (*Kindergartenkind*). Solche Komposita sind auch dann nicht Teil der Untersuchung, wenn die beiden identischen Konstituenten, wie etwa in Beispiel (4), adjazent sind:

(4) *Schon im Elternhaus wird ihnen suggeriert, daß eine Mann-Frau-Beziehung das einzig Wahre ist, eine **Mann-Mann-Beziehung** allenfalls verboten, verpönt oder krankhaft.*

<www.miraculum-aurich.de/archiv/alt/vor05/web/homepage/kunstschule/ventil/ausgaben/no07-/schwul_titel.htm>

Der offensichtlichste Unterschied zwischen Bildungen wie in Beispiel (4) und ICCs wie in (1–3) ist, dass in (4) die Identität nicht alle Konstituenten umfasst. Diese Bildung ist zudem kein N+N-Kompositum, sondern ein Phrasenkompositum. Elsen (2014: 62) nimmt bei solchen „Sonderkomposita" ein Kopulativverhältnis zwischen den beiden dem Letztglied untergeordneten Konstituenten an (*Mann-Mann-*). Doch unterscheidet sich das Kompositum in (4) von Kopulativkomposita wie *Hosenrock* oder *Dichterkomponist* dahingehend, dass nicht nur der semantische Kopf fehlt, sondern auch der syntaktische. *Mann-Mann-* kann deshalb nicht determiniert werden und daher auch nicht alleine stehen. In manchen Fällen wird die dritte, übergeordnete Konstituente zwar vorangestellt (5), oder, etwa als Wortellipse, weggelassen (6), sodass dann doch genau zwei identische Konstituenten vorliegen:

(5) *Ist die Beziehung **Mann-Mann** oder Frau-Frau weniger wert als die zwischen Frau und Mann?*

<http://bitgewitter.blogger.de/stories/1009862>

(6) *Bei Beziehungen bevorzuge ich **Mann-Mann**.*

<http://rakate.24fps.de/index.php?paged=2>

Allerdings ist in beiden Fällen die dritte Konstituente implizit gegeben. *Mann-Mann* ist in keinem der Fälle ein Wort, sondern der phrasale Teil eines Phrasenkompositums. Der Determinierer richtet sich in seiner Form nach dem implizit gegebenen Kopf des Kompositums, nicht nach *Mann* (*die Mann-Mann-Beziehung*, *die Mann-Mann*, aber **der Mann-Mann*).

1.4 Aufbau der Arbeit

In Kapitel 2 beschreibe ich zunächst den Forschungsstand zu N+N-Komposita sowie zu ICCs. Da Forschungsliteratur zu ICCs im Deutschen sehr rar ist, bleiben viele Fragen offen, die die Arbeit im Anschluss zu beantworten versucht. Die Ka-

pitel 3 bis 6 stellen den empirischen Teil der Arbeit dar. Kapitel 3 beschreibt die Durchführung der Korpusstudien. Kapitel 4 gibt einen Überblick über die Daten und nimmt eine funktional-semantische Analyse vor, die in einer Subklassifikation der ICCs mündet. Kapitel 5 referiert die Ergebnisse zu formalen Aspekten der ICC-Subtypen. Kapitel 6 fasst die Ergebnisse der Korpusstudien zusammen, beantwortet die aufgestellten Forschungsfragen und diskutiert die Ergebnisse vor dem Hintergrund der Forschungsliteratur. Kapitel 7 bis 10 ordnen die Ergebnisse in die grammatiktheoretische Diskussion ein. Kapitel 7 behandelt hier die Frage, inwieweit die Bildung von ICCs der Komposition zuzurechnen ist; Kapitel 8 diskutiert, inwieweit sie als Reduplikation aufzufassen ist. Kapitel 9 wendet einschlägige Kompositatheorien und Modelle semantischer Transparenz auf ICCs an. Zum Abschluss des grammatiktheoretischen Teils wird in Kapitel 10 der Ansatz der Relational Morphology (Jackendoff & Audring 2020a, b) hinsichtlich der Beschreibung von ICCs bewertet, für gut befunden und genutzt, um die Bedeutungskonstitution von ICCs nachzuzeichnen. Das 11. und letzte Kapitel zieht ein Fazit aus den Erkenntnissen dieser Arbeit und formuliert Fragen, die nicht beantwortet werden konnten.

2 Forschungsüberblick

N+N-Komposita mit identischen Konstituenten wurden in der Forschung bisher wenig beachtet. Dennoch hat die englisch- und deutschsprachige Forschungsliteratur viele unterschiedliche Termini für das Phänomen hervorgebracht, die jeweils unterschiedliche Aspekte von Form und Funktion der ICCs betonen. Zum einen werden für ICCs die Bezeichnungen „Selbstkomposita“ und „Eigenkomposita“ (Donalies 2011: 72, Elsen 2014: 67, Schindler 1991: 603) verwendet, vornehmlich in Handbüchern und Einführungen zur Wortbildungsmorphologie. Zum anderen gibt es die Termini „contrastive focus reduplication“ (Bross & Fraser 2020, Ghomeshi et al. 2004), „contrastive reduplication“ (Song & Lee 2015, Whitton 2006), „Double“ (Horn 1993), „lexical cloning“ (Horn 2006), „identical constituent compounds“ (Finkbeiner 2014, Hohenhaus 2004, Kentner 2017), „REAL-X-Reduplication“ (Freywald 2015, Stolz et al. 2011), „reduplicative compounding“ (Lieber 2009a: 364), „Name Doubling“ (Kauffman 2015: 3) und „Echowörter“ (Platen 2013: 53). Aufmerksamkeit bekam das Phänomen ICC vor allem durch die Studie von Ghomeshi et al. (2004), die englische ICCs wie *salad-salad* ‘prototypischer Salat’ beschreibt (The salad-salad paper).

Nur wenige Studien widmen sich dabei dezidiert N+N-Komposita mit identischen Konstituenten im Deutschen. Davon wiederum gehen nur sehr wenige das Thema empirisch an, weshalb es nur wenige gesicherte Erkenntnisse zu Form und Funktion von ICCs gibt. Ich stelle in diesem Kapitel zunächst N+N-Komposita vor, zu denen ICCs gehören (2.1). Im Anschluss stelle ich chronologisch die bisher veröffentlichten Forschungstexte vor, die sich gezielt oder zumindest in nennenswertem Maße mit ICCs im Deutschen auseinandersetzen (2.2). Danach gehe ich auf Texte ein, die ICCs nur am Rande erwähnen (2.3) sowie auf die Literatur zu ICCs im Englischen und anderen Sprachen, da vor allem die Literatur zu ICCs im Englischen sehr viel umfangreicher ist als die zu ICCs im Deutschen (2.4). Zum Abschluss des Kapitels ziehe ich ein Fazit zum Forschungsstand (2.5).

2.1 N+N-Komposita im Deutschen

2.1.1 Erstglied, Zweitglied und die Elemente dazwischen

In einem N+N-Kompositum verbinden sich zwei nominale Stämme zu einer morphologischen, prosodischen, graphematischen und semantischen Einheit:

https://doi.org/10.1515/9783111315416-002

(7) *Schnee* + *Ball* =*Schneeball*

(8) *Haus* + *Katze* =*Hauskatze*

Die Konstituenten sind im Kompositum weder trennbar (*eine schöne Hauskatze, *eine Haus schöne Katze*) noch ohne Bedeutungsveränderung vertauschbar (*Hauskatze ≠ Katzenhaus*, Olsen 2015, Schlücker 2012). Zudem ist N+N-Komposition rekursiv. Ihr Ergebnis steht also nach dem Wortbildungsprozess wieder für die Bildung eines N+N-Kompositums zur Verfügung (*Schneeballsystem*, *Hauskatzenrasse*).

N+N-Komposita lassen sich semantisch-funktional und formal in Kopf und Modifikator einteilen. Formal bedeutet diese Einteilung, dass das Kompositum seine grammatischen Eigenschaften vom grammatischen Kopf, im Deutschen die rechte Konstituente, erhält. Es gilt das Prinzip der Rechtsköpfigkeit. Einem Kompositum wie *Hauskatze* wird so das grammatische Genus von *Katze* zugewiesen (Femininum) und nicht das von *Haus* (Neutrum). Auch weitere Eigenschaften wie die Flexionsklasse des Gesamtkompositums leiten sich von der rechten Konstituente ab.

Semantisch-funktional bedeutet die Einteilung in Kopf und Modifikator, dass bei der N+N-Komposition meist Subklassen gebildet werden und das Zweitglied der semantische Kopf ist, der die dem Kompositum übergeordnete Klasse bezeichnet. Das Erstglied, der Modifikator, nennt das Subklassifikationsmerkmal. Die Formel, nach der man die Konzeptvereinigung in N+N-Komposita beschreiben kann, lautet in etwa ‘X, das mit Y zu tun hat’ (Zifonun 2010: 128). Im Beispiel *Schneeball* wäre die Bedeutungsangabe also ‘ein Ball, der mit Schnee zu tun hat’. Komposita sind dadurch Hyponyme ihrer Köpfe: Ein *Schneeball* ist eine bestimmte Art Ball, nämlich einer, der mit Schnee zu tun hat, und eine *Hauskatze* ist eine bestimmte Art Katze, nämlich eine, die etwas mit einem Haus zu tun hat.

Erst- und Zweitglied in einem N+N-Kompositum realisieren also die sehr unterschiedlichen Funktionen Kopf und Modifikator. Die Nominalstämme des Deutschen treten nicht gleichermaßen in diesen beiden Funktionen auf. Es gibt korpusbasierte Evidenz dafür, dass die meisten Nomina häufiger die Modifikatorposition als die Kopfposition realisieren (Brunner et al. 2021: 22). Zudem gibt es lexemspezifische Präferenzen: Manche Lexeme treten häufiger in der Kopfposition, manche eher in der Modifikatorposition von N+N-Komposita auf (Bauer et al. 2019: 50, Brunner et al. 2021). Diese Präferenz für eine der zwei Positionen lässt sich sowohl hinsichtlich lexikalisierter Komposita als auch hinsichtlich neu gebildeter Komposita nachweisen (Baayen 2010, Tarasova 2013) und hat auch Auswirkungen auf die Verarbeitung der Komposita. Komposita sind demnach kognitiv

leichter zu verarbeiten, wenn die involvierten Nomina die für sie übliche Rolle als Kopf oder Modifikator einnehmen, die Erstglieder also üblicherweise Erstglieder, die Zweitglieder üblicherweise Zweitglieder sind (Baayen 2003, Schreuder & Baayen 1997).

Welche Position ein Nomen bevorzugt, hängt von unterschiedlichen Faktoren ab. Zum einen wird der Type-Frequenz eines Lexems und somit dem Lexikon Einfluss zugeschrieben (Bauer 2001b, Hay & Baayen 2002). Wenn ein Nomen bereits in sehr vielen Komposita als Kopf auftritt, ist die Wahrscheinlichkeit, dass es auch in Neubildungen als Zweitglied verwendet wird, höher als bei solchen, die vor allem als Modifikator vorkommen. Diese übernehmen in Neubildungen entsprechend eher die Funktion des Modifikators (Bauer et al. 2019: 54, Maguire et al. 2010). Auch ist die Frage, ob Nomina überhaupt eine Präferenz für Erst- oder Zweitglied haben, von der Frequenz abhängig. Nomina, die besonders häufig in Komposita verwendet werden, die also eine große Kompositafamilie haben, zeigen im Allgemeinen keine Präferenz und werden in Kopf- und Modifikatorposition gleichermaßen verwendet. Nomina, die generell seltener in Komposita auftreten, tendieren hingegen zu einer der beiden Positionen (Brunner et al. 2021: 22). Zum anderen werden morphosyntaktische Faktoren wie Komplexität oder eine unklare Wortartenzuordnung als Faktoren diskutiert. Tarasova (2013) zeigt etwa für das Englische, dass Nomina, die Eigenschaften von Adjektiven aufweisen und in Wörterbüchern sowohl als Nomina als auch als Adjektive geführt werden (*chocolate*, *lemon*, *future*), eine Tendenz zur Modifikatorposition haben.

Darüber hinaus wird der Semantik der Lexeme ein Einfluss auf ihre Verwendung als Erst- oder Zweitglied attestiert. Hier ist zum einen die semantisch-thematische Kategorie der Nomina ausschlaggebend. Stoffsubstantive treten besonders häufig in Modifikatorposition auf, während Lexeme, die Artefakte bezeichnen, signifikant häufiger in Kopfposition auftreten. Dadurch ergibt sich beispielsweise das frequente Interpretationsmuster „x^{ARTIFACT} is made of $y^{\text{SUBSTANCE}}$“ (Brunner et al. 2021: 25), zu dem etwa das oben genannte *Schneeball* gehört. Brunner et al. (2021) konnten zudem korpusbasiert nachweisen, dass Polysemie ein Faktor ist, der darüber entscheidet, ob ein Nomen in N+N-Komposita eher die Kopf- oder eher die Modifikatorposition einnimmt. Polyseme Nomina treten demnach häufig in Kopfposition auf; Lexeme mit nur einer lexikalischen Bedeutung präferieren hingegen die Modifikatorposition. Dieses Ergebnis entspricht der bereits erwähnten Tatsache, dass die Bildung von Hyponymen eine der Hauptfunktionen von N+N-Komposition ist. Komposita im Allgemeinen dienen der Benennung neuer Subkonzepte (Schlücker 2012: 16). Mit diesen Subkonzepten können die Sprecher:innen Bedeutung genauer anzeigen als mit den Konzepten der jeweiligen Basen. Polyseme Lexeme sind nun aber in doppelter Hinsicht unbe-

stimmt, weil sie neben der pragmatischen Unbestimmtheit (Vagheit), die allen Ausdrücken zukommt, die semantische Unbestimmtheit (Ambiguität) aufweisen. Sie bedürfen deshalb in besonderem Maße der Spezifizierung durch eine Modifikatorkonstituente (Brunner et al. 2021: 23).

Kopf und Modifikator eines N+N-Kompositums bilden aber nicht nur semantisch, sondern auch morphologisch eine Einheit. Vergleicht man N+N-Komposita mit einer syntaktischen Verbindung zweier Nomina, zeigen sich viele formale Unterschiede:

(9) *Das ist der **Doktorvater**.* ➔ Morphologische Bildung

(10) *Sie erinnert sich des **Doktorvaters**.* ➔ Morphologische Bildung

(11) *Das ist des **Doktors Vater**.* ➔ Syntaktische Bildung

In den Bildungen in (9–10) zeigt sich die morphologische Natur zum einen dadurch, dass die Bildungen lexikalisch integer sind. Die erste Konstituente ist für syntaktische Operationen nicht zugänglich. Nur das letzte Glied flektiert nach dem jeweiligen syntaktischen Kontext. Versetzt man die im Nominativ stehende NP in (9) in den Genitiv (10), zeigt also nur das Zweitglied den Flexionsmarker *-s*. Das Erstglied zeigt keinen Flexionsmarker. Die Bildung flektiert als Ganzes; innerhalb des Kompositums findet keine kontextuelle Flexion statt. In (11) hingegen steht das erste Nomen der NP im Genitiv. Das Genitivattribut ist in diesem Fall als Archaismus dem nominalen Kopf der Nominalphrase vorangestellt und da *Doktor* Teil dieses vorangestellten Genitivattributs ist, zeigt es den entsprechenden Flexionsmarker *-s*.

In Komposita sind die Erstglieder nicht Teil einer vorangestellten Phrase. Doch auch an den Erstgliedern in N+N-Komposita finden sich bisweilen formale Elemente. Zwischen den Konstituenten werden in manchen Fällen Elemente hinzugefügt (*Arbeitstier*, **Arbeittier*), in anderen Fällen getilgt (*Wollschuhe*, **Wolleschuhe*). Solche Fugenelemente sind der „Verbindungskitt zwischen zwei Einheiten“ (Donalies 2011: 54). Die Fugenelemente des Deutschen sind (in unterschiedlicher Auftretenshäufigkeit) *-s-*, *-(e)n-*, *-(e)r-*, *-e-*, und *-ens-* teils in Kombination mit dem Umlaut (Kopf 2018: 33, Neef 2009: 392, Schlücker 2012: 9). Häufig gehen die Fugenelemente historisch auf Wortformen des Flexionsparadigmas zurück (Haase 1989: 23), sind aber synchron „semantisch leer“ (Donalies 2011: 45). Funktionslos sind sie jedoch nicht. In der Literatur werden unterschiedliche Funktionen von Fugenelementen diskutiert: Nach Fuhrhop (2000: 212) zeigen Fugenelemente die Morphologisierung der Komposita an und unterscheiden sie formal von

syntaktischen Verbindungen (*Richtung weisend* versus *richtungsweisend*). Auch wird angenommen, die Elemente dienten der Aussprachheerleichterung (Busch & Stenschke 2007: 87, Donalies 2005: 45, Michel 2009: 337, Wegener 2005: 117). Ferner wird Fugenelementen die Funktion zugeschrieben, Bedeutung unterscheiden zu können, etwa die zwischen einer determinativen und einer kopulativen Lesart von N+N-Komposita (siehe Besprechung von Beispiel 27). Nübling und Szczepaniak (2009) modellieren die Verwendung von Fugenelementen stattdessen komplexer und nehmen für die Allo-Formen unterschiedliche Distributionen an. Für die Verwendung des „häufigste[n], produktivste[n] und [...] am stärksten von der Flexion entkoppelte[n] Fugenelement[es] *-s-*“ nehmen sie etwa an, dass sein Auftreten von der wortphonologischen Qualität des Erstglieds gesteuert wird (Nübling & Szczepaniak 2009: 220). Weichen Wörter vom Muster Trochäus mit Reduktionssilbe ab, wird eher verfugt, um dem Hörer die Dekodierung zu erleichtern:

> Der HörerIn wird gerade dann, wenn das Ende des phonologischen Wortes schlecht zu erkennen ist, eine materielle Unterstützung geliefert. Indem -s- seinerseits gleichzeitig das phonologische Wortende komplexer werden lässt [...], unterstützt es aktiv die rechte Wortrandverstärkung.
>
> (Nübling & Szczepaniak 2009: 220)

Fugenelemente bilden allerdings keine homogene Klasse und ihnen wird nicht nur eine einzige Funktion zugeschrieben. Michel schreibt dazu etwa:

> Es darf mittlerweile als gesichert gelten, dass sich Fugenelemente polyfunktional verhalten, was in erster Linie auf der phonetisch-phonologischen, morphologischen und semantischen Ebene zum Ausdruck kommt.
>
> (Michel 2009: 337)

In vielen Fällen sind die Fugenelemente formal mit dem nominalen Teil von Genitivattribut-NPs identisch:

(12) *Das ist das **Téufelshuhn**.* ➔ Morphologische Bildung

(13) *Das ist des **Téufels Húhn**.* ➔ Syntaktische Bildung

In (13) ist *Teufels* Teil des Genitivattributs *des Teufels*. Die beiden unterstrichenen nominalen Einheiten der NP sind also nicht morphologisch, sondern syntaktisch verbunden. In (12) hingegen liegt ein Kompositum vor. Weil Kompositionsstammform und Genitivform von *Teufel* aber formgleich sind, wird die morphologische Natur nicht durch die An- oder Abwesenheit interner Marker deutlich.

In der Phonologie ist allerdings ein Unterschied hörbar. Die Erstgliedbetonung in (12) zeigt, dass die beiden Nomina *Teufel* und *Huhn* ein Kompositum bilden (morphologisierender Kontrastakzent, Eisenberg 2002: 353) und die entsprechende Nomenverbindung als semantische Einheit interpretiert werden muss. Die Betonung in (13) entspricht nicht diesem Akzentmuster und markiert die syntaktische Verbindung. Allerdings folgen nicht alle N+N-Komposita diesem Akzentmuster. Beispielsweise weisen augmentativ-evaluative N+N-Komposita wie *Ríesenpárty* (Grzega 2004) und Kopulativkomposita wie *Ősterreich-Úngarn* Doppelbetonung auf.

Im Schriftbild zeigt zudem die Zusammenschreibung an, dass (9–10) sowie (12) Komposita sind, während das Spatium in (11) und (13) markiert, dass die beiden Nomina zu unterschiedlichen syntaktischen Positionen gehören. Im Deutschen ist die normgerechte Schreibung von Komposita die Zusammenschreibung (Fleischer & Barz 2012: 127). Neu gebildete Komposita weisen allerdings die Tendenz auf, die Konstituenten durch den Bindestrich oder das Spatium zu trennen (Scherer 2012).

2.1.2 Modifikatoren, Modifikation und Modifikationsrelation

Die funktionale Einteilung der N+N-Komposita in Kopf und Modifikator beruht auf der Beobachtung, dass Erst- und Zweitglied von N+N-Komposita auf sehr unterschiedliche Weise dazu beitragen, einen Bezug zu Konzepten und Entitäten herzustellen. Wie bei Substantiven im Allgemeinen besteht auch bei N+N-Komposita die wesentliche Funktion in der Referenz. Der Terminus „Referenz“ bezeichnet dabei den Vorgang, die Aufmerksamkeit des Kommunikationspartners auf ein Objekt, beziehungsweise eine Entität zu lenken (Lyons 1977: 184, Searle 1969: 81). Auf welches Objekt die Aufmerksamkeit gelenkt werden kann, hängt stark von der Denotation eines Substantivs ab, also der „Kategorie, oder Menge, seiner potenziellen Referenten“ (Löbner 2011: 28). Hinzu kommen weitere Aspekte wie etwa textuelle Mittel (etwa Anaphern, Determinatoren), physisches Verweisen (Deixis) oder das geteilte Weltwissen der Sprecher:innen. Für die Beschreibung des gesamten Verweisungsprozesses werden häufig auch die Begriffe „Identifikation“ und „Spezifikation“ verwendet. Bei der „Objektsreferenz“ („prototypische Referenz“, Dressler & Mörth 2012: 225) genannten Art des Verweisens wird der Referent demnach in einem Diskurs-Universum identifiziert und dadurch spezifi-

ziert (*das ist das Teufelshuhn*).[1] Die referierenden Ausdrücke, die also auf eine klar identifizierte Entität verweisen, können anaphorisch oder kataphorisch wieder aufgenommen werden (Karttunen 1969).

N+N-Komposita referieren in der Regel in Form des gesamten Kompositums. Allerdings tragen die beiden Konstituenten eines N+N-Kompositums auf unterschiedliche Weise zu seiner Bedeutung bei und beeinflussen, worauf der Ausdruck verweist. Während die Kopfkonstituente in einem N+N-Kompositum für gewöhnlich auf ein Konzept referiert (Schlücker 2012: 15), kann der im Deutschen durch das Erstglied repräsentierte Modifikator unterschiedliche Funktionen ausüben. Er kann restriktiv sein, wenn nämlich der Referent oder das Denotat des Modifikanden restringiert wird, und er kann nicht restriktiv sein, wenn er also keinen Einfluss auf den Referenten oder das Denotat des Modifikanden hat.

Unter den restriktiven Modifikatoren gibt es klassifikatorische oder referenzidentifizierende. Wird das gesamte Konzept, beziehungsweise die Denotation des Modifikanden eingeschränkt, ist die Modifikation klassifikatorisch. Hier wird mitunter davon gesprochen, dass der Modifikator generisch verweist, beziehungsweise referiert, das heißt, dass die Menge der potentiellen Referenten des Modifikators nicht über die Denotation des Substantivs hinaus eingeschränkt ist. Dies wird „denotation restriction“ oder „type restriction“ genannt (Bauer 2006a, Downing 1977, Koptjevskaja-Tamm & Rosenbach 2005: 9). Wie in den Beispielen *Schneeball* und *Teufelshuhn* dient das Erstglied als Modifikator der Identifizierung der Subklasse und verweist generisch auf allgemeine Vertreter (Gunkel & Zifonun 2009, 2011, Kürschner 1974: 97f., Ortner & Müller-Bollhagen 1991: 28–38, Pavlov 1983: 44f.) beziehungsweise auf die Klasse der entsprechenden Gegenstände (Fleischer & Barz 2012: 130). So referiert *Schnee* in *Schneeball* nicht auf ein bestimmtes Auftreten von *Schnee*; *Teufel* in *Teufelshuhn* nicht auf einen bestimmten Teufel. Stattdessen referieren die beiden Erstglieder auf *Schnee* und *Teufel* im Allgemeinen.

Komposita mit solchen Modifikatoren werden auch Determinativkomposita genannt. Bei Determinativkomposita verweisen die modifizierenden Erstglieder generisch auf eine Klasse an Entitäten. Durch die Modifikation entsteht ein neues und eigenständiges Subkonzept. Die Subklassifikation und die kognitive Dimension dieses Prozesses lässt sich mit den Prinzipen der Kategorisation von Rosch (1978) beschreiben. Rosch unterscheidet drei Abstraktionsstufen, denen die Kate-

1 Unter „Definitheit“ wird neben dieser „Identifizierbarkeit“ noch „Verankerung“ sowie „Exklusivität“ oder „Unikalität / Gesamtheit“ verstanden (Lavric 2000: 25, Löbner 2011: 287). Zum Unterschied zwischen diesen Auffassungen und zur Diskussion einer Definitheitsdefinition siehe Lyons (1999), insbesondere Kapitel 1 und 7.

gorien der Basisebene („basic level“), der übergeordneten Ebene („superordinate level“) und der untergeordneten Ebene („subordinate level“) entsprechen. Während das Konzept BALL (und das entsprechende Nomen *Ball*) auf der Basisebene angesiedelt ist, sind die Subkonzepte von BALL (und die entsprechenden Nomina, etwa *Schneeball und Tennisball*) auf der untergeordneten Ebene zu verorten (Rosch 1978: 32). Die Subkonzeptbildung von N+N-Komposita ist in Abbildung 2 dargestellt.

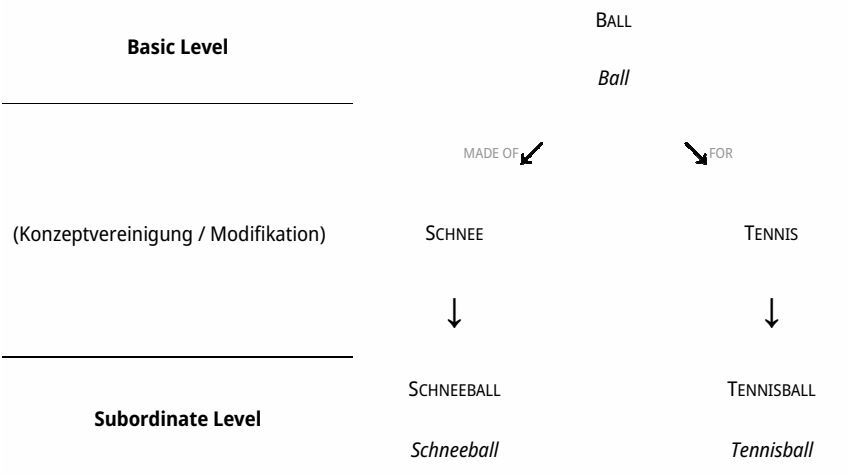

Abb. 2: Subkonzeptbildung von N+N-Komposita.

Für die morphologische Modifikation gilt diese Klassifikationsbedeutung als Defaultfall. Die Subklassifikation entsteht unabhängig davon, welcher syntaktischen Kategorie der Modifikator angehört. So sind *Schneeball* und *Softball* Hyponyme zum übergeordneten, vom Kopf bezeichneten Konzept BALL und Kohyponym zueinander. Insbesondere das Wortbildungsmuster N+N-Komposition hat von sich aus diese abstrakte klassifikatorische Bedeutung. N+N-Komposita geben somit üblicherweise eine Antwort auf die Frage: „Welche Art von X?“.

In klassifikatorischen N+N-Komposita wie *Teufelshuhn* entsteht im Zuge der Modifikation zusätzlich eine semantische Verengung (auch: semantische Spezialisierung) der Denotation des Zweitgliedes. Die Extension des modifizierten Elements wird weiter reduziert. In *Teufelshuhn* wird nämlich nicht nur ein Subkonzept gebildet, sodass das Kompositum anders als *Huhn* nicht mehr auf jedes fasanenartige, landwirtschaftlich genutzte Tier referieren kann (ein *Teufelshuhn* ist eine bestimmte Art *Huhn*, nämlich eines, das etwas mit *Teufel* zu tun hat). Es

findet zusätzlich eine semantische Spezialisierung statt, wodurch das Konzept idiosynkratisch festgelegte Bedeutungsmerkmale erhält. *Teufelshuhn* kann deshalb nicht auf alle Hühner, die etwas mit Teufeln zu tun haben (Aussehen, Farbe usw.) referieren, sondern nur auf Hühner einer ganz bestimmten Untergattung der Haubenhühner.

In klassifikatorischen N+N-Komposita modifiziert die erste Konstituente das Konzept des Zweitgliedes allerdings nicht immer gleich. *Schnee* in *Schneeball* informiert über das Material des Balls, *Fuß* in *Fußball* hingegen über das Instrument, mit dem man für gewöhnlich auf den Ball einwirkt, *Tennis* in *Tennisball* über den Sport, für den der Ball verwendet wird. Die Qualität der logischen Verbindung zwischen erster und zweiter Konstituente, also die Festlegung, in welcher Hinsicht das Erstglied das Zweitglied subklassifiziert, ist von Kompositum zu Kompositum unterschiedlich. Vor allem spontan gebildete N+N-Komposita sind zunächst immer semantisch unterspezifiziert (Bauer 1978, Downing 1977). Das von Heringer diskutierte Beispiel *Fischfrau* ermöglicht beispielsweise mehr als zehn Lesarten (etwa ‘Frau des Fischs’ oder ‘Frau, die Fisch verkauft’, Heringer 1984: 2). Erst der Äußerungskontext disambiguiert die okkasionelle Bildung (Peschel 2002: 298). Der Grad der Unterspezifizierung variiert allerdings, sodass die Interpretation neu gebildeter N+N-Komposita von Eigenschaften der Konstituenten abhängt, unter anderem etwa von der Polysemie der zugrundeliegenden Basislexeme (Schäfer & Bell 2020). Auszuklammern sind in dem Zusammenhang außerdem Rektionskomposita, bei denen das Kopfnomen deverbal ist und damit Informationen zur Relation aus der Valenz des zugrundeliegenden Verbs erbt, etwa *Zeitungsleser* oder *Geldwäscherfahndung* (Olsen 1986: 78ff.).

N+N-Komposita umfassen also mehr als die Summe der Konstituentenbedeutungen. Es kommt die Wortbildungsbedeutung hinzu, „also jene[r] Teil der (lexikalischen) Gesamtbedeutung, der durch den Wortbildungsvorgang entsteht“ (Barz 2009: 673). Zwischen den Konstituenten von N+N-Komposita besteht eine semantische Relation (auch Modifikationsrelation oder thematische Relation genannt), die sehr unterschiedliche Verhältnisse zwischen Erst- und Zweitglied ausdrücken kann. Das Schema in Abbildung 1 zeigt für *Schneeball* die Relation MADE OF, für *Tennisball* die Relation FOR. Solche Relationen sind nicht formal repräsentiert, in den meisten Fällen nicht strukturell vorhersagbar und müssen gelernt oder mithilfe des Kontexts inferiert werden (Booij 2009: 322). Die Gründe hierfür sowie die Mechanismen der Kompositasemantik sind theoretisch vielfältig beschrieben worden (siehe Kapitel 9). Wegen der nicht explizierten Modifikationsrelation gehen die Zusammenfügungen mit einer extremen Verdichtung von Informationen einher.

> N+N compounding [...] allows encoding complex concepts through an extremely compressed format, making use of as little space as language enables to.
>
> (Fernández-Domínguez 2010: 49)

Häufig wird angenommen, dass die semantischen Relationen die klassifikatorische Bedeutung von Komposita bedingen. Allerdings bilden manche Subtypen der nominalen Komposition auch ohne ein breites Set semantischer Relationen Subklassen. In A+N-Komposita etwa benennt das Adjektiv Eigenschaften und modifiziert das Kopfnomen direkt, ohne dass eine implizite semantische Relation hergeleitet werden muss (Schlücker 2014: 119ff.). Die Modifikationsstruktur von Komposita wie *Heißgetränk* oder *Vollmond* ist gleichbedeutend mit den entsprechenden A+N-Phrasen *heißes Getränk* und *voller Mond.* Nichtsdestotrotz beschreiben diese N+N-Komposita nicht bloß, vielmehr benennen sie und haben eine klassifikatorische Bedeutung (‘eine bestimmte Art Getränk’, ‘eine bestimmte Art Mond’).

In N+N-Komposita mit klassifikatorischen Modifikatoren sind die Modifikationsrelationen sehr vielfältig. Es gibt zahlreiche Versuche, diese Vielfalt zu systematisieren. Dazu werden die in Komposita formal nicht vorliegenden Relationen in Form von Präpositionen (*Ball aus Schnee*), syntaktischer Kontexte (*ein Ball besteht aus Schnee*) oder spezifischer Prädikate (*Schnee* MAKE *Ball*, *Ball* MADE OF *Schnee*) hinzugefügt. Levi beschreibt in Bezug auf englische Determinativkomposita neun mögliche Relationen, die weitgehend den traditionellen semantischen Kategorien entsprechen. So entspricht CAUSE der Kategorie ‘kausativ’ (*drug death* ‘ein Tod, der durch eine Droge herbeigeführt wurde’), ABOUT der Kategorie ‘topik’ (*tax law* ‘ein Gesetz über Steuern’) und HAVE der Kategorie ‘possessiv’ (*apple cake* ‘ein Kuchen mit Äpfeln’).

Bei solchen Systematisierungsversuchen werden mal mehr, mal weniger semantische Kategorien angenommen, die die Relation zwischen Erst- und Zweitglied charakterisieren (Eichinger 2000, Jackendoff 2010a, Levi 1978, Ortner & Müller-Bollhagen 1991, Ryder 1994). Hatcher (1960) nimmt (für das Englische) vier Relationen an; Eichinger (2000: 118) „in etwa zehn“, Jackendoff (2010) dreizehn (ebenfalls für englische Komposita), inklusive so komplexer Relationen wie der von *bike helmet* („helmet to be worn while riding a bike“). Ortner und Müller-Bollhagen (1991: 145ff.) beschreiben sogar zweiunddreißig unterschiedliche semantische Relationen in Nominalkomposita des Deutschen. Verschiedene empirische Studien weisen aber darauf hin, dass Sprecher:innen bei der Interpretation unbekannter N+N-Komposita generell nur auf einige wenige semantische Relationen zurückgreifen, vor allem MADE OF, HAS und LOCATED (Hohenhaus 2005, Krott et al. 2009, Krott et al. 2010: 385).

Keiner der bisherigen Systematisierungsversuche findet in der Kompositaforschung uneingeschränkte Zustimmung. Auch bei sehr umfangreichen Sets von Relationen, etwa dem von Ortner und Müller-Bollhagen (1991), gibt es viele Komposita, die sich keiner der Relationen zuordnen lassen. Bauer (2006a: 722) nennt in Bezug auf dieses Problem das Kompositum *Spaghettiwestern*, dessen interne Relation so spezifisch ist ('der Western aus einem Land, das durch die Menge an Spaghetti, die dort konsumiert wird, näher bestimmt ist'), dass keine der in der Literatur formulierten Relationen es abbilden kann. Das Kompositum aber einfach als Ausnahme anzusehen ist auch nicht legitim, weil es weitere Komposita gibt, die genau dieselbe Relation aufweisen, etwa *Gulaschkommunismus* ('der Kommunismus aus einem Land, das durch die Menge an Gulasch, die dort konsumiert wird, näher bestimmt ist', Bauer 2006a: 722). Bereits 1942 betont Jespersen (1942: 143ff.) die theoretisch unbegrenzte Anzahl semantischer Relationen in N+N-Komposita und die Unmöglichkeit, sie zu systematisieren (ebenso: Botha 1968). Dennoch wird häufig auf konkrete Sets an semantischen Relationen zurückgegriffen und dies auch als sinnvoll erachtet, da die semantisch-strukturelle Vielfalt der N+N-Komposita mit ihnen zumindest näherungsweise systematisiert werden kann (Schäfer 2018: 125).

Neben den klassifikatorischen Modifikatoren gibt es solche, die nicht generisch auf eine Klasse, sondern selbst auf ein Objekt in der Welt referieren. Ein solcher Modifikator wird dann als restriktiv bezeichnet (Gunkel & Zifonun 2009, 2011), da er die Referenz eines Nomens auf eine bestimmte Entität (Person, Gegenstand oder Sachverhalt) im Äußerungskontext beschränkt ('token restriction', Rosenbach 2007, Rosenbach 2019, „restricting the reference of a nominal", Seiler 1978). Der Modifikator hilft somit bei der Identifikation des Referenten:

> [I]dentifying modifiers help to fix the reference of the noun phrase, that is, they contribute to the identification of the referent of the noun phrase. [They] relate to the referential or existential status of an entity in the discourse, signalling that the addressee is presumed to be (un-)familiar with the NP referent.
>
> (Schlücker 2013: 459)

Ein solcher Modifikator wird identifizierender oder referenzspezifizierender Modifikator genannt. Er trägt zur Beantwortung der Frage „Welcher NP-Referent?" bei. In Komposita, in denen die Erstkonstituente auf diese Art referenziell interpretiert wird, finden sich als Modifikatoren häufig Eigennamen, da diese inhärent definit sind (Schlücker 2012: 16):

(14) *Nach dem **Reus-Freistoß** aus halblinker Position köpfte Piszczek unhaltbar für Ulreich rechts unten ins Tor.*

Das Kompositum in (14) ersetzt eine syntaktische Konstruktion (*der Freistoß von Reus*) durch eine morphologische und verdichtet so die Informationen im Text. Die Bildungen werden daher auch deiktische Komposita oder Textfunktionskomposita genannt. In *Reus-Freistoß* referiert das Erstglied auf eine konkrete Person im Äußerungskontext. Es ist deshalb die Referenz des Kopfnomens, die dadurch spezifiziert wird. Es wird ein ganz bestimmter Freistoß identifiziert.

Im Zuge der Modifikation in N+N-Komposita können die Kopfnomen also klassifizierend (Type restriction, *Schneeball*) oder referenzspezifizierend (identifizierend, Token restriction, *Reus-Freistoß*) modifiziert werden. Diese zwei Funktionen entsprechen zwei der drei Modifikationsrelationen, die klassischerweise für Modifikation in Nominalphrasen angenommen werden (etwa Teyssier (1968): identification (id), qualification (qual), classification (class)). Die Unterscheidung zwischen referenzspezifizierenden und klassifikatorischen Modifikatoren betrifft deren Referenzleistung. Während die Kopfkonstituente in einem N+N-Kompositum auf ein Konzept referiert, können die Erstkonstituenten spezifisch (*Reus* in *Reus-Freistoß*) oder nicht-spezifisch (*Schnee* in *Schneeball*) verweisen. Verweisen sie spezifisch, sind den Rezipient:innen konkrete Referenten textuell (Anapher) oder physisch (Deixis) präsent. Verweisen sie nicht-spezifisch, wird hingegen nicht auf konkrete Referenten Bezug genommen, sondern auf eine Gattung, eine Klasse oder ein Konzept.

Es gibt mehrere Möglichkeiten, festzustellen, welche Art der Modifikation jeweils vorliegt. Zum einen gibt die Determination des Artikels einen Hinweis darauf, ob ein Subkonzept zum Konzept des Kopfnomens vorliegt oder nicht. In Beispiel (14) ist das Kompositum durch einen definiten Artikel determiniert (*der Reus-Freistoß*). In Beispiel (15) aber ist *Reus* ein klassifizierender Modifikator, da das Kompositum eine bestimmte Art Freistoß bezeichnet, nämlich einen, der etwa durch eine bestimmte Technik oder Flugbahn gekennzeichnet ist (Sortenlesart). Ein solcher Freistoß kann auch von einem anderen Fußballspieler ausgeführt werden:

(15) *Platziert. Hart. Unberechenbar. Die Freistöße von Marco Reus sind in der ganzen Liga gefürchtet. Wir verraten Dir, wie auch Dir ein **Reus-Freistoß** gelingt.*

<https://www.bravo.de/sport/die-freistoesse-von-marco-reus-so-schiesst-marco-reus-seine-freistoesse-222149.html>

Obgleich das Kompositum formal mit dem in (14) identisch ist, ist das Erstglied nicht mehr als referenzspezifizierend zu interpretieren, sondern notwendigerweise ein klassifikatorischer Modifikator, da referenzspezifizierende Modifikatoren stets ein definites Kompositum bilden, das einen ganz konkreten Fall (Token) identifiziert. Der indefinite Artikel ist mit solchen Komposita inkompatibel, weil die „indefinite Determination noch einmal eine echte Teilmenge innerhalb [der Menge der möglichen Referenten] ausgliedert" (Lavric 2000: 52). Referenzspezifizierende Modifikatoren sollen einen konkreten Fall identifizieren, indefinit generischen Nominalphrasen aber „wohnt [...] die Fähigkeit inne, eine gesamte Klasse zu repräsentieren" (Lavric 2000: 48).[2] Die Beispiele (14) und (15) verdeutlichen außerdem, dass Eigennamen als Erstglieder in N+N-Komposita nicht notwendigerweise referieren, sondern genauso gut klassifikatorische Modifikatoren sein können.

Eine weitere Möglichkeit, zu prüfen, ob eine Modifikation der Spezifizierung eines Falls oder eines Typs dient, ist die anaphorische Bezugnahme. Die Anapher ist neben der Deixis das wichtigste Mittel der definiten Referenz (Hawkins 1978, 1984, Vater 1984).[3] Wird ein konkreter Fall spezifiziert, hat der Modifikator einen konkreten Referenten im Äußerungskontext, auf den Bezug genommen werden kann (Coulmas 1988). Im Beispiel (14) ist es möglich, auf das Erstglied *Reus* Bezug zu nehmen. In Beispiel (16) geschieht dies in Form eines Pronomens; in (17) mit einer referenzidentischen Nominalphrase:

(16) *Nach dem **Reus**$_i$-**Freistoß** aus halblinker Position, den er$_i$ schnell ausgeführt hatte, köpfte Piszczek unhaltbar für Ulreich rechts unten ins Tor.*

(17) *Nach dem **Reus**$_i$-**Freistoß** aus halblinker Position, den das Schlitzohr$_i$ schnell ausgeführt hatte, köpfte Piszczek unhaltbar für Ulreich rechts unten ins Tor.*

Im Falle eines klassifikatorischen Modifikators ist diese Bezugnahme nicht (18–19) oder nur unter bestimmten Bedingungen (etwa bei weiterer kontextueller Anreicherung) möglich (20–21):[4]

2 Generell können aber auch grammatisch indefinit bezeichnete Referenten im Diskurs definit sein (Givón 1993: 215).

3 Zur Diskussion der Definitionen und Unterscheidungskriterien zwischen Anapher und Deixis siehe Consten (2004: 4ff.).

4 Andere Faktoren, die eine anaphorische Bezugnahme auch bei klassifikatorischen Modifikatoren begünstigen, liegen etwa vor, wenn der Referent semantisch transparent (Ward et al. 1991) oder generell ein prominentes Thema im Text ist (Coulmas 1988).

(18) **Wir verraten Dir, wie auch Dir ein **Reus$_i$-Freistoß** gelingt, über den er$_i$ staunen würde.*

(19) **Wir verraten Dir, wie auch Dir ein **Reus$_i$-Freistoß** gelingt, über den das Schlitzohr$_i$ staunen würde.*

(20) *Wir verraten Dir, wie auch Dir ein **Reus$_i$-Freistoß** gelingt, über den er$_i$, also Marco Reus selbst, staunen würde.*

(21) *Wir verraten Dir, wie auch Dir ein **Reus$_i$-Freistoß** gelingt, über den der Star selbst$_i$ staunen würde.*

Restriktive Modifikatoren können also die Menge der möglichen Referenten eines Kompositums entweder durch die Einschränkung der Referenz oder der Denotation reduzieren.

Ist eine Modifikation nicht-restriktiv, wird also weder die Referenz noch die Denotation des Modifikanden restringiert, ist die Modifikation qualifizierend. Die qualifizierenden Modifikatoren entsprechen der letzten noch verbleibenden der drei Modifikationsrelationen, die klassischerweise für Modifikation in Nominalphrasen angenommen werden („qualification“, Teyssier 1968). Die qualifizierenden Modifikatoren geben eine Antwort auf die Frage: „Wie ist X?“.

Bei der adjektivischen Modifikation von Nomina geschieht diese qualitative Modifikation üblicherweise in Form von A+N-Phrasen:

(22) *Das **weiße Kaninchen** mümmelt Möhrchen.*

Die Adjektive schreiben einem Gegenstandsbegriff, im Beispiel dem Konzept Kaninchen, Eigenschaften zu, etwa die, eine bestimmte Farbe zu haben (*weißes Kaninchen*). Die Modifikation in der A+N-Phrase ist nicht-restriktiv, weil ein bereits eindeutig identifizierbares Kaninchen zusätzlich als weiß beschrieben wird. Weder Konzept (Denotat) noch Referenz des Nomens werden durch die Modifikation beschränkt. Der Satz in (22) ist in dem Fall äquivalent zu folgendem Satz:

(23) *Das **Kaninchen**, das übrigens **weiß** ist, mümmelt Möhrchen.*

Auch A+N-Phrasen können allerdings restriktiv modifizieren und klassifikatorische Modifikatoren haben. Klassifikatorisch sind A+N-Phrasen, wenn sie Benennungseinheiten sind und also das Denotat des Bezugsnomens restringieren, wie

etwa *saure Sahne* oder *gelbes Trikot* (Schlücker 2014: 147). Andersherum können auch Komposita nicht-restriktive, qualitative Modifikatoren haben:

(24) *Als neuralgischer Punkt kann der Mainzer Hauptbahnhof gelten, wenn* ***Normal-Pendler****, Bush-Zugbenutzer und Demonstranten aufeinander prallen werden.*
(Schlücker 2013: 474)

Hier ist *Normal-Pendler* das Äquivalent zur Phrase *normale Pendler*. Das Adjektiv hat eine rein qualitative, den Referenten des Bezugsnomens beschreibende Funktion. In N+N-Komposita ist die Modifikation aber meist restriktiv und der Modifikator hat einen Einfluss auf die Referenz des Kopfes.

Für das Englische wird allerdings diskutiert, ob auch adnominale Nomina qualitative Modifikatoren sein können. Durch die Abwesenheit morphologischer Marker sowohl an adjektivischen als auch an nominalen Modifikatoren sei im Englischen nur schwer zu entscheiden, ob in Bildungen wie *cat food* oder *student performance* ein adjektivischer oder ein nominaler Modifikator vorliegt (Payne & Huddleston 2002: 537). Die traditionelle Schulgrammatik nimmt an dieser Stelle einen adjektivischen Modifikator an (Koptjevskaja-Tamm & Rosenbach 2005: 15). Spencer (2003, 2005) interpretiert hier das Erstglied als relationales, denominales Adjektiv, das sich syntaktisch wie ein Adjektiv verhält, aber die morphosyntaktischen Eigenschaften eines Nomens bewahrt.

Im Deutschen kommen die Erstglieder der bereits erwähnten augmentativ-evaluativen Komposita einer qualitativen Modifikation am nächsten (Pittner & Berman 2006). Hier wird die vom Zweitglied bezeichnete Entität mithilfe der ersten Konstituente gesteigert, also quantitativ modifiziert (25), oder bewertet, also qualitativ modifiziert (26):

(25) *Ich bekam erst einmal* ***Mordsprobleme****, weil ich ungefragt an die Unterlagen meiner Mutter gegangen war.*

<http://www.wochenendvati.de/archives/2315>

(26) *Ronnie James Dio ist nicht sonderlich großgewachsen, hat aber eine* ***Hammerstimme****. Was ist sein Geheimnis?*

<http://www.rockhard.de/megazine/heftarchiv/online-megazine/tonezone/47864.html>

Auch hier identifizieren die Erstglieder *Mord* und *Hammer* kein Subkonzept der Konzepte PROBLEM und STIMME und dienen auch nicht der Spezifizierung des Re-

ferenten. Die Modifikatoren sind hier stattdessen eine nähere Beschreibung oder Charakterisierung des zweiten nominalen Elements. Anders als bei rein qualitativen A+N-Phrasen mit nicht-restriktiven Modifikatoren wird hier aber eine Bewertung des Sprechers über die konkreten Referenten dieser Konzepte ausgedrückt („expressive compounds", Meibauer 2013, Rijkhoff 2008, 2010). Die Modifikation benennt also keine Qualität, die dem Referenten unabhängig vom Sprecher zukommt.

Schließlich gibt es noch N+N-Komposita, bei denen augenscheinlich überhaupt keine Modifikation stattfindet, da der semantische Kopf nicht ohne Weiteres zugeordnet werden kann:

(27) *Es ist nicht nur zum Sitzen da, sondern auch zum Schlafen, ein Sofa mit Schlaffunktion also. Ein modernes* ***Bettsofa*** *hat nichts mehr mit den früher üblichen Notbehelfen zu tun.*

<http://www.matratzen1.de/tag/sofa-mit-schlaffunktion/>

Kopulativkomposita wie *Bettsofa* sind nach dem Mustern 'X, das gleichzeitig Y ist' beziehungsweise 'X, das gleichzeitig Y und Z ist' zu paraphrasieren („both N1 and N2", Jackendoff 2009: 123, „B which is also A", Marchand 1969: 41). Ein *Bettsofa* ist also ein 'Bett, das gleichzeitig Sofa ist' oder ein 'Möbelstück, das gleichermaßen Bett und Sofa ist'. Die Konstituenten scheinen gleichrangig zu sein. In jedem Fall besteht ein Unterschied solcher Komposita zu determinativen Komposita vom Typ *Hauskatze* (` 'Haus, das auch eine Katze ist').

Wie oben bereits angesprochen wird den Fugenelementen mitunter die Funktion zugeschrieben, die jeweilige Bedeutung unterscheiden zu können, dergestalt, dass bei determinativen Komposita eine Tendenz zur Verfugung bestehe, bei kopulativen hingegen zur Nullfuge (Dressler et al. 2001: 51, Neef & Borgwaldt 2012). Ein verfugtes Kompositum wie *Komponistenmaler* wird also tendenziell die Bedeutung 'eine Person, die einen Komponisten malt' zugeordnet, dem unverfugten Äquivalent *Komponist-Maler* hingegen die Bedeutung 'eine Person, die Komponist und Maler gleichzeitig ist'. Empirische Studien zeigen hier aber für das Deutsche, dass die Sprecher:innen auch solche Komposita wie *Bettsofa* oder *Komponist-Maler* determinativ interpretieren, also von einer wortinternen Modifikation ausgehen (Breindl & Thurmair 1992, Klos 2011: 271). Insofern gehören diese meist Kopulativkomposita genannten Bildungen also zu den Komposita mit klassifikatorischer Modifikation.

Es stellt sich nun die Frage, inwieweit ICCs die hier besprochenen formalen und funktionalen Merkmale sowie die Modifikationsarten von N+N-Komposita aufweisen. ICCs zeigen mit ihrer reduplikativen Struktur ein zusätzliches, sehr

auffälliges formales Merkmal. Dies ist allerdings nicht der einzige Unterschied zwischen ICCs und anderen N+N-Komposita.

2.2 ICCs im Deutschen

ICCs gelten als N+N-Komposita (Finkbeiner 2014, Hohenhaus 2004, Kentner 2017). Demnach sind sie eine morphologische Zusammensetzung von Nominalstämmen. Lange wurden ICCs in linguistischen Texten bloß erwähnt, um ihre Unmöglichkeit zu proklamieren, um also Beschränkungen für Kompositionsprozesse zu definieren, die eben gerade ICCs ausschließen. So heißt es etwa bei Kürschner (1974):

> Es gelten [für Nominalkomposita] folgende Bedingungen: (i.) A` B; d.h. als A und B müssen unterschiedliche Lexikoneinträge gewählt werden, um Formen wie **Gartengarten*, **Bleistiftbleistift* auszuschließen.
>
> (Kürschner 1974: 148)

Auch Erben gibt an, dass „für Komposita im Deutschen normalerweise die Regel von der Kombination u n g l e i c h e r Morpheme bzw. Lexeme“ gelte (Erben 1981: 39, Hervorhebung im Original). Möglicherweise ist die Annahme, dass Komposita mit identischen Konstituenten nicht existieren, der Grund dafür, dass sich die germanistische Linguistik bisher so wenig mit ICCs befasst hat. Ich beschreibe nun chronologisch die wenigen Forschungstexte, die sich den ICCs im Deutschen widmen.

2.2.1 Günther (1979)

Ende der 1970er Jahre erscheint die erste Studie, die sich mit ICCs beschäftigt. Die Studie von Günther erhebt empirische Daten zu ICCs, beschäftigt sich allerdings nicht dezidiert mit ihnen, sondern mit zweigliedrigen N+N-Komposita im Allgemeinen. Für den empirischen Teil der Studie interpretieren und paraphrasieren Proband:innen unbekannte N+N-Komposita ohne Informationen zum Kontext („compound meaning interpretation task“, Onysko 2014: 73). Die kombinatorische Erstellung der Stimuli bedingt, dass den Proband:innen unter anderem Bildungen präsentiert werden, in denen die Konstituenten identisch sind, etwa *Holzholz* und *Frauenfrau*. Günther weist mit seinen Ergebnissen erstmals nach, dass ICCs in den meisten Fällen interpretiert werden können. Zudem zeigt er, dass bei diesen Interpretationen nicht immer ein Determinationsverhältnis zwischen Erst- und Zweitglied besteht:

> In 2/3 der Interpretationsversuche [...] wurden diese Bildungen als Determinativkomposita aufgefaßt. In 1/3 der Antworten aber – und dies betraf gerade solche SSN [ICCs; M.F.], die prima facie nicht interpretierbar erschienen – wurde ganz offenbar nach dem Vorbild der Genitivbildungen wie Fest der Feste, Buch der Bücher, Tag der Tage etc. interpretiert, was zu Angaben führte wie *Hammerhammer* 'Superhammer', 'Überhammer'; *Frauenfrau* 'besonders typische Frau', 'Überfrau', usw.
>
> (Günther 1979: 271)

Günthers Studie liefert den frühesten Beleg dafür, dass Sprecher:innen die ICCs nicht nur als Determinativkomposita interpretieren, sondern auch als Genitivbildungen, die eine Art Steigerung oder Intensivierung ausdrücken. Auf die Studie von Günther folgt lange keine weitere Studie, die sich in relevantem Maße mit dem Thema ICC auseinandersetzt. Erst seit Mitte der Neunzigerjahre befasst sich die linguistische Forschung auch explizit mit ICCs.

2.2.2 Hohenhaus (1996, 1998, 2004, 2007, 2015)

Die Arbeiten von Hohenhaus zu ICCs im Deutschen ignorieren weitestgehend ICCs wie *Kindeskind* oder *Zinseszins* und widmen sich vornehmlich ICCs wie *Beziehung-Beziehung* in (2), die eine Prototypenlesart aufweisen. In seinen Arbeiten widerspricht Hohenhaus der von Ghomeshi et al. (2004) hervorgebrachten Einschätzung, im Deutschen gebe es keine Komposita wie *salad-salad*, das sie für das Englische beschreiben. Nach Hohenhaus sind solche ICCs im Deutschen mindestens genauso häufig wie im Englischen (Hohenhaus 1996: 59, 1998: 255f., 2004: 319, ebenso: Stolz et al. 2011: 202). Hohenhaus führt zudem den formal motivierten Terminus ICC 'identical constituent compound' für diese Bildungen ein und klassifiziert sie damit erstmals und in Abgrenzung zu Ghomeshi et al. (2004) als Komposita. Der Begriff ICC, wie Hohenhaus ihn verwendet, bezieht sich dabei ausschließlich auf ICCs mit Prototypenlesart.

Hohenhaus gibt zudem, in Anlehnung an Horn (1993: 48), eine Formel an, nach der solche ICCs im Deutschen zu interpretieren sind: „an XX is a proper/prototypical X" (Hohenhaus 2004: 301). Er konkretisiert damit das von Günther genannte Interpretationsschema („Genitivbildungen wie Fest der Feste", Günther 1979: 271) hinsichtlich einer Prototypenlesart. Der Verweis auf den Prototypen geschieht in Anlehnung an die Prototypentheorie Roschs (1975: 455). Das Konzept, das ein solches ICC ausdrückt, ist also der eindeutigste Vertreter, das beste Beispiel seiner Klasse. Für das englische Beispiel *salad-salad* von Ghomeshi et al. (2004) visualisieren Song und Lee (2015) die Prototypenstruktur wie folgt:

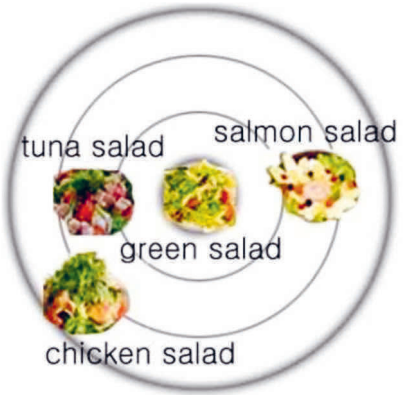

Abb. 3: Prototypenstruktur der Kategorie SALAT (Song & Lee 2015).

Grundlage dieser Einteilung ist eine Umfrage, die unter englischen Muttersprachler:innen durchgeführt wurde (Song & Lee 2015). Die Teilnehmer:innen mussten hier für verschiedene Konzepte einen Prototypen nennen. Für das Konzept SALAT wurde zumeist grüner Salat als prototypischster Vertreter genannt (Song & Lee 2015: 449). Das englische *salad-salad* verweist also auf das Konzept eines prototypischen, nämlich grünen, Salats (Abbildung 2).

In den Arbeiten von Hohenhaus dienen die ICCs unter anderem als Beispiel für „nonce formations", also für Wortbildungsprodukte, die für sehr konkrete Situationen gebildet werden und sich deshalb nicht im Lexikon festsetzen können (Hohenhaus 2007: 25ff., 2015: 275). Im Deutschen seien ICCs, anders als im Italienischen, wo das Muster vollständig grammatikalisiert sei (Wierzbicka 1991), generell nicht lexikalisierbar. Es gebe nur ein einziges etabliertes Lexem (*Film-Film* 'richtiger Film, keine Dokumentation') das aber bewusst vom deutschen Fernsehsender Sat.1 gebildet worden und daher eine „künstliche" Bildung sei. *Film-Film* sei das einzige ICC, das „item-familiar" (im Sinne von Meys 1975) ist. Alle anderen ICCs im Deutschen seien nur „type-familiar", also über das ICC-Muster und die von Hohenhaus genannte Formel interpretierbar ohne die Möglichkeit, ins Lexikon des Deutschen zu gelangen. Hohenhaus formuliert bezüglich *Film-Film* die These, dass das Wort von den Sprecher:innen des Deutschen trotz seines institutionalisierten Status als spontane Bildung der mündlichen Konversation wahrgenommen werde. Dies erkläre auch, warum die Sprecher:innen die Bildung so häufig humorvoll ironisch oder gar spöttisch verwenden würden: Sie nähmen den Versuch des Fernsehsenders, ICCs außerhalb ihrer natürlichen Domäne („outside its natural domain", Hohenhaus 2004: 319) zu verwenden, als gescheitert wahr.

Hohenhaus bespricht einen möglichen semantisch-funktionalen Unterschied zwischen ICCs mit Prototypenlesart und kanonischen N+N-Komposita. So liege es nahe, dass in ICCs keine Kopf-Modifikator-Relation und somit auch keine Modifikationsrelation vorliegt (Hohenhaus 2004: 301). Im genannten Beispiel *Salat-Salat* wäre dann also das Zweitglied nicht vom Erstglied näher bestimmt. Eine semantische Relation wie OF, FOR oder HAVE wäre nicht anwendbar. Diese Besonderheit verneint Hohenhaus aber und nimmt an, dass ICCs genau wie andere N+N-Komposita eine rechtsköpfige Kopf-Modifikator-Relation aufweisen und dadurch auch eine Form der Subkonzeptbildung darstellen. Das Erstglied ist demnach ein klassifikatorischer Modifikator, der das Denotat des zweiten Bestandteiles einschränkt (Hohenhaus 2004: 299, ebenso: Kentner 2017: 240). Hohenhaus ist allerdings der Auffassung, dass ICCs nicht mit den üblichen semantischen Relationen paraphrasiert werden können, sondern nur gemäß der von ihm genannten Formel (Hohenhaus 2004: 301).

Der Kontext sei für die Anwendung der Formel aber entscheidend. Durch ihn werde eine Lesart aus einer Reihe möglicher Lesarten des Lexems identifiziert. Nach Hohenhaus (2004: 301) treten Bildungen wie *Salat-Salat* vor allem in kontrastiven „co-text frames“ auf, in denen ein prototypischer mit einem weniger prototypischen Vertreter der Kategorie kontrastiert wird (‘not X, but XX’). Dies deckt sich mit der Funktion, die für ICCs im Englischen angenommen wird, dass nämlich die Verwendung eines ICCs eine Reparatur darstellt:

> Identical Constituent Compounding is best characterized as a repair strategy, by which a speaker recognizes that there is a potential for being misunderstood – or that they have been explicitly misunderstood – and attempts to reorient the listener toward a more precise construal in line with the speaker’s intent.
>
> (Benjamin 2018: 21)

Widlitzki (2016) führt für das Englische etwa das Beispiel an, dass das Nomen *wife* in bestimmten Kontexten mithilfe von *WIFE-wife* disambiguiert wird, um es von *WORK-wife* abzugrenzen.

Hohenhaus legt mit seinen Texten die seit Günther (1979) ersten, und über lange Zeit auch die einzigen, theoretischen Auseinandersetzungen mit ICCs im Deutschen vor. Doch sind seine Arbeiten keine empirischen und behandeln noch dazu nur ICCs vom Typ *Salat-Salat*. Die Arbeiten von Finkbeiner füllen diese beiden Forschungsdesiderate.

2.2.3 Finkbeiner (2014)

Ähnlich wie die Studie von Günther (1979) untersucht auch Finkbeiner ICCs in Isolation. Im Zuge einer compound meaning interpretation task paraphrasieren 40 Teilnehmer:innen 16 nicht-lexikalisierte N+N-Komposita mit identischen Konstituenten. Finkbeiner kann dabei die Ergebnisse von Günther (1979) replizieren: In ungefähr 60% der Fälle können die Proband:innen den unbekannten Formen eine konkrete Bedeutung zuordnen. Anders als in der Studie von Günther zeigen die Paraphrasen allerdings in 64% der auszuwertenden Fälle die von Hohenhaus angenommene Prototypenlesart.[5] Nur in weniger als 20% der Fälle interpretieren die Proband:innen die Komposita determinativ. Finkbeiners Studie liefert also Evidenz dafür, die Prototypenlesart als Defaultlesart von ICCs anzusetzen.[6] Finkbeiner zieht daraus den Schluss, dass die Sprecher:innen des Deutschen kontextfreies Wissen über die Interpretation von ICCs haben. Sie positioniert sich mit ihren Ergebnissen allerdings nicht zu der von Hohenhaus aufgeworfenen Frage, ob ICCs sich im Lexikon des Deutschen festsetzen können, und spricht vorsichtiger von „words in performance“ (Finkbeiner 2014: 206).

Abseits der empirischen Ergebnisse stellt Finkbeiner auch theoretische Überlegungen zu Form und Funktion von ICCs an. Bezüglich der funktionalen Aspekte nimmt Finkbeiner eine weitere Lesart an, die nicht auf einen guten Vertreter einer Klasse abzielt (gewöhnlich’ Blumen, keine ausgefallenen), sondern durch Kategorienidentität entsteht (echte Blumen, keine Plasteblumen). So dienten in einigen Bildungen die ICCs vor allem dazu, auszudrücken, dass der Referent auch tatsächlich Teil der Klasse ist. Während beispielsweise das Lexem *Blume* nicht nur auf Vertreter der Klasse Blumen referiert, sondern auch auf Klassen nichtpflanzlicher Objekte wie beispielsweise Kunstblumen, könne das ICC verdeutlichen, dass der Referent auch wirklich der Klasse der Blumen angehört. Die Prototypenlesart nennt Finkbeiner „Prot“, die Lesart der Kategorienidentität „Real“.

Darüber hinaus stellt Finkbeiner die These auf, dass sich die Funktionen, die die Konstituenten in ICCs mit Prototypenlesart übernehmen, von denen unterscheiden, die sie in ICCs wie *Kindeskind* ausüben (Finkbeiner 2014: 188). Trügen in *Kindeskind* beide Konstituenten zur Gesamtbedeutung des Kompositums bei (‘Kind des Kindes’), sei in ICCs wie *Salat-Salat* hingegen die Gesamtbedeutung keine Kombination der beiden Konstituentenbedeutungen. Stattdessen werde die

5 Die beiden von Finkbeiner unterschiedenen, nicht-determinativen Lesarten PROT und REAL sind hierbei zusammengerechnet.

6 Die Stimuli werden allerdings allesamt ohne Fugenelemente präsentiert und forcieren damit – wie später noch zu sehen sein wird – die Prototypenlesart.

Gesamtbedeutung nur von der lexikalischen Bedeutung der zweiten Konstituente bestimmt.

> X2 contributes its lexical meaning, and X1 contributes some kind of functional meaning.
> (Finkbeiner 2014: 188)

Die erste Konstituente diene also nur dazu, das Denotat des Kopfes auf seinen prototypischen Kern zu beschränken. Somit ließen sich solche ICCs nicht wie andere N+N-Komposita in Kopf und Modifikator einteilen, sondern in einen lexikalischen und einen funktionalen Teil. Finkbeiner charakterisiert den gesamten Prozess daher als funktional, weshalb er der totalen Reduplikation in anderen Sprachen gleiche (Finkbeiner 2014: 188). Wie schon Hohenhaus nimmt Finkbeiner aber dennoch bei ICCs mit Prot- oder Real-Lesart eine klassifikatorische Bedeutung an (Finkbeiner 2014: 188). *Beziehungbeziehung* in (2) ist eine bestimmte Art Beziehung; *Salat-Salat* eine bestimmte Art Salat.

Finkbeiner beschreibt fernerhin formale Unterschiede zwischen ICCs und anderen N+N-Komposita. Demnach träten in ICCs die Basislexeme im Erstglied in ihrer Nennform auf (*Liebe*) und nicht – wie in N+N-Komposita üblich – in ihrer Kompositionsstammform (*Liebes*). Die Verwendung des Fugenelementes unterbleibe (*Liebe-Liebe*, Finkbeiner 2014: 185ff.). Finkbeiner sieht hierin sogar ein (einseitiges) Distinktionsmerkmal: Gibt es ein Fugenelement am Erstglied, wird also statt der Nennform die Kompositionsstammform verwendet, sei die Prototypenlesart ausgeschlossen (Finkbeiner 2014: 187).

Auch zu flexionsmorphologischen Aspekten äußert sich Finkbeiner. So verhielten sich ICCs hinsichtlich der Flexion wie determinative N+N-Komposita (Finkbeiner 2014: 186). Flexionselemente träten demnach nicht an der ersten Konstituente auf; das ICC flektiere stets als Einheit (28):

(28) *Wein-Weine / *Weine-Weine / *Weine-Wein*
(Finkbeiner 2014: 186)

In Ermangelung empirischer Studien zur Flexion von ICCs kann Finkbeiner allerdings nicht entscheiden, ob tatsächlich alle ICCs als Ganzes flektieren oder die ICCs womöglich Flexionsmarker am Erstglied erlauben.

Hinsichtlich des Akzentmusters von ICCs mit Prot- oder Real-Lesart geht Finkbeiner davon aus, dass sie die übliche Kompositionsbetonung aufweisen (Finkbeiner 2014: 187), sie also auf der ersten Konstituente betont werden (*Salát-Salat*, *Beźiehung-Beziehung*). In Bezug auf den Kontext schließlich unterstreicht Finkbeiner, dass der Kontrast nicht unbedingt im Kotext etabliert werden, sondern sich lediglich im Common Ground der Konversation befinden muss (Fink-

beiner 2014: 191). Dafür sprechen auch ihre empirischen Ergebnisse, nach denen die Proband:innen auch ohne Kontext eine klare Präferenz für eine der Lesarten haben.

Finkbeiner selbst nennt ihre Studie eine „Pilotstudie“ (Finkbeiner 2014: 182). Für die wissenschaftliche Auseinandersetzung mit ICCs stellt die Studie jedoch einen Meilenstein dar. Sie liefert die bis heute einzigen empirischen Daten zu ICCs (determinativen und solchen mit Real- oder Prot-Lesart) seit der Studie von Günther aus dem Jahre 1979, und zudem Evidenz für die These, dass ICCs auch außerhalb eines konkreten Äußerungskontextes interpretiert werden können. Da die Studie auf Paraphrasen von Proband:innen fußt, sagt sie allerdings selbstredend nichts über ICCs aus, die die Sprecher:innen auch tatsächlich verwenden. Diesen Aspekt berücksichtigt die Publikation von Freywald.

2.2.4 Freywald (2015)

Freywald (2015) verwendet Korpora zur Beschreibung von ICCs im Deutschen. Freywald bezieht Beispiele für ICCs aus drei verschiedenen Quellen, nämlich aus dem Richling-Korpus (2008), dem DeWaC (Baroni et al. 2009) und von Google. Anhand von 24 ausgewählten Beispielen versucht Freywald, formale und funktionale Eigenschaften der Bildungen zu beschreiben. Im Ergebnis ordnet sie ICCs, anders als Hohenhaus und Finkbeiner, nicht der Komposition zu, sondern setzt einen eigenen, reduplikativen Wortbildungsprozess an. Der Begriff „REAL-X-Reduplication“, den sie von Stolz et al. (2011) adaptiert, bezieht sich neben dem Verweis auf die Reduplikation auch auf die Bedeutungsseite von ICCs. Freywald setzt also die Real- und nicht die Prot-Lesart als grundlegend an.

Freywald geht auch auf die Frage der Modifikation in ICCs ein und schreibt: „we are dealing with a genuine modifier-head structure here“ (Freywald 2015: 921, 925, 941). Wie Hohenhaus sieht sie also in keiner Weise einen Unterschied zwischen ICCs mit Real- oder Prot-Lesart und gewöhnlichen N+N-Komposita:

> One might wonder whether it is just that a new semantic relation has been added to the repertoire of possible ways of interpreting otherwise ‘normal’ compounds.
>
> (Freywald 2015: 923)

Hinsichtlich der formalen Merkmale von ICCs mit Real- oder Prot-Lesart nimmt Freywald eine Gegenposition zu Finkbeiner ein. In einigen ihrer Beispiele tritt das Erstglied nicht in der Nennform auf, sondern beinhaltet formale Elemente (*Blumen<u>n</u>blumen*). Freywald interpretiert solche Beispiele nicht als verfugte Komposita, sondern als Komposita mit Flexionselementen am Erstglied und spricht sich somit

auch gegen die Auffassung Finkbeiners aus, nach der ICCs nur als Ganzes flektieren (Freywald 2015: 925). Durch diese Klassifikation der Elemente am Erstglied erklärt sich auch, warum Freywald ICCs nicht als Komposita, sondern als das Ergebnis eines Reduplikationsprozesses betrachtet.

Auch in Bezug auf die Schreibung weichen Freywalds ICC-Beispiele von anderen N+N-Komposita ab (Freywald 2015: 916). Ihre Korpusbeispiele sprechen dafür, dass die beiden Konstituenten von ICCs häufig durch den Bindestrich, das Spatium oder auch die Binnenmajuskel getrennt werden, was nicht der normgerechten Schreibweise entspricht, nach der nominale Komposita zusammengeschrieben werden (Fleischer & Barz 2012: 127).

In einer Annahme zur Bedeutungskonstitution spricht sich Freywald dafür aus, dass allein der Kontext die Bedeutung eines ICCs bestimmt (Freywald 2015: 923). Sie nimmt also die Position ein, die auch der Auffassung von Hohenhaus zugrunde liegt, nach der ICCs außerhalb eines lizensierenden Kontextes nicht verwendet werden können (Hohenhaus 2007: 25ff., 2015: 275). Diese Position steht in einem Widerspruch zu den Ergebnissen Finkbeiners, die für eine kontextfreie und dennoch einheitliche Interpretation von ICCs durch die Sprecher:innen sprechen.

Die von Freywald verwendeten Beispiele stellen keine große empirische Datenbasis dar. Sie zeigen aber, dass die Frage, ob ICCs mit Real- oder Prot-Lesart hinsichtlich Verfugung, Flexionsverhalten und Schreibung von N+N-Komposita im Allgemeinen abweichen, noch nicht beantwortet ist.

2.2.5 Kentner (2017)

Die Studie von Kentner beschäftigt sich nicht ausschließlich mit ICCs, sondern geht der weiter gefassten Frage nach, inwieweit Eigennamen mit reim- und ablautreduplikativer Struktur systematisch gebildet werden. Die von Kentner untersuchten Bildungen sind also vornehmlich keine ICCs, denn die Konstituenten dieser Komposita sind nicht vollkommen identisch, sondern partiell redupliziert. In den von ihm untersuchten Ablautreduplikationen wie *Mipsmops* alterniert der Vokal, in Reimreduplikationen wie *Hansipansi* der Konsonant. Hier könnte man einwenden, dass auch ICCs wie *Kindeskind* der Fugenelemente wegen nicht vollständig identisch sind. Doch ist dies nicht das Definitionskriterium für ICCs. *Kindeskind* ist ein N+N-Kompositum, weil es eine Wortbildung aus zwei nominalen Stämmen darstellt, und es ist ein ICC, weil diese beiden Stämme identisch sind. Die von Kentner beschriebenen Bildungen sind allerdings keine Komposita aus zwei nominalen Stämmen, da nur eine Konstituente einen Nominalstamm dar-

stellt (*Hans*, *Mops*). Der jeweils andere Bestandteil der Bildung ist hingegen kein Nominalstamm (*-pansi, Mips-*) und kann eher als Affix beschrieben werden, das seine Segmente teilweise aus dem Nominalstamm erhält, an den er affigiert wird.

Kentner geht aber auch auf ICCs ein und stimmt Hohenhaus und Finkbeiner darin zu, die Bildungen als Komposita anzusehen. Er bespricht darüber hinaus das Phänomen, dass Sprecher:innen mithilfe von ICCs, die ein Anthroponym als Basis haben (*Sabinesabine*), Pseudonyme erstellen (Kentner 2017: 244). Die Basis werde allein zur formalen Erweiterung im Zuge der Namensbildung verdoppelt (ebenso: Dürscheid 2005: 48). Kentners Studie liefert mithilfe von Akzeptabilitätstests allerdings Evidenz dafür, dass bei der Bildung von Eigennamen, die selbst auf Eigennamen basieren, die Reimreduplikation (*Sabinepabine*) vorherrscht und die totale Reduplikation, also die Verwendung von ICCs (*Sabinesabine*) eher dispräferiert ist (Kentner 2017: 252f.).

2.2.6 Bross und Fraser (2020)

In der bis dato jüngsten wissenschaftlichen Veröffentlichung zu ICCs argumentieren Bross und Fraser auf der Grundlage von Introspektion und informeller Befragung von Einzelpersonen dafür, dass ICCs vom Typ *Kaffee-Kaffee* nicht adjektivisch modifiziert werden können, abgesehen von „Real Intensifications“ (Bross & Fraser 2020: 6). Als Real Intensifications sehen die Autor:innen Adjektive an, die die Prototypikalität des Referenten des Substantivs semantisch unterstützen (*richtiger Mann-Mann*). Diese Besonderheit, adjektivische Modifikation zu blockieren, die in keiner der anderen Publikationen zu ICCs im Deutschen erwähnt wird, lasse sich den Autor:innen zufolge damit erklären, dass das Erstglied bereits einen „adjectival flavor by itself“ habe (Bross & Fraser 2020: 6). Bei contrastive focus reduplications (= ICCs mit Real- oder Prot-Lesart) sei darum die adjektivische Modifikation blockiert, ein Umstand, den die Autor:innen mithilfe generativer Beschreibungsansätze zu erklären versuchen.

Ferner nehmen Bross und Fraser an, dass in diesen Bildungen Erst- und Zweitglied stets gleich flektieren müssen (2020: 3). ICCs seien also lexikalisch nicht integer, da die kontextuelle Flexion in die Bildungen hineinreiche und die entsprechenden Flexionsmarker damit auch am Erstglied obligatorisch seien (29):

(29) *Männer-Männer / *Mann-Männer / *Männer-Mann*
(Bross & Fraser 2020: 3)

Auch diese Annahme der Autor:innen widerspricht den in früheren Forschungstexten vertretenden Auffassungen. Denn nach Finkbeiner (2014: 186) flektieren ICCs nicht am Erstglied, sondern nur als Einheit. Freywald (2015: 925) geht bei formalen Elementen am Erstglied zwar auch von Flexionsmarkern aus, hält diese aber nicht für obligatorisch.

Für ICCs verwenden Bross und Fraser zudem, in Anlehnung an Ghomeshi et al. (2004), den Begriff „contrastive focus reduplication" und sprechen sich energisch gegen die Annahme aus, dass es sich bei Bildungen wie *Kaffee-Kaffee* um Komposita handelt (Bross & Fraser 2020: 5). Allen von Bross und Fraser getroffenen Annahmen liegen allerdings keine empirischen Daten zugrunde.

2.3 Publikationen, die ICCs am Rande erwähnen

Zunächst gibt es die bereits erwähnten Texte, die ICCs nur verwenden, um zu illustrieren, dass N+N-Komposita aus unterschiedlichen Stämmen gebildet sein müssen. Dem widersprechend führen manche Einführungen in die (Wortbildungs-)Morphologie sowie manche Handbücher zur Wortbildung ICCs als Wortbildungsprozess auf. Hier wird allerdings meist nur auf einige der bis dato erschienenen Publikationen verwiesen und mithilfe einiger Beispiele eine kurze Einschätzung zu Form, Funktion sowie eine Einordnung von ICCs gegeben. Donalies (2011), Elsen (2014) sowie Fleischer und Barz (2012) verwenden etwa die Bezeichnungen „Selbstkomposita" und „Eigenkomposita" und ordnen sie den Determinativkomposita zu. Ihre Konstituenten seien rein zufällig identisch (Donalies 2011: 72, Elsen 2014: 67, ebenso: Fleischer & Barz 2012: 96).

ICCs wurden außerdem von präskriptiver, beziehungsweise sprachpflegerischer Seite aus berücksichtigt. Kurz nach der Studie Günthers, die zeigt, dass die Sprecher:innen ICCs sinnvoll interpretieren können, werden ICC-Belege aus überregionalen Zeitungen in linguistischen Veröffentlichungen besprochen. Erben (1981) führt in seinem Text zu Neologismen die Bildung *Mannmann* mit der Bedeutung 'männlicher Mann' aus der Süddeutschen Zeitung vom 25.08.1978 auf und nennt sie „ungewöhnlich" (Erben 1981: 39). Der Wortbildungstyp sei zudem „nur in wenigen Bildungen rechtssprachlichen Charakters vertreten" (Erben 1981: 39). Erbens Ausführungen zu dem Phänomen gehen aber über Stilkritik nicht hinaus.[7]

7 Der Autor wirkt in seinem Beitrag durchaus enttäuscht von seinen Kollegen, wenn er schreibt, dass „in neuerer Fachsprache auch Bildungen wie *Wissenschaftswissenschaft* (Lieb) gewagt werden". Dies sei „höchstens deshalb verwunderlich, weil dies von linguistischer Seite geschieht" (Erben 1981: 40).

Des weiteren erwähnen einige Studien zu ICCs in anderen Sprachen auch die Möglichkeit der ICC-Bildung im Deutschen. So geht ein Text von Schindler der Frage nach, ob der sprachtypologisch so häufige Prozess der Reduplikation auch für das Deutsche angesetzt werden kann. Schindler behandelt wie Kentner vor allem Reim- und Ablautreduplikationen, geht aber auch auf ICCs ein. Er beschreibt in einem Absatz die Bildungen *Holzholz*, *Helfershelfer* und *Kindeskind*, für die er den Terminus „determinative Selbstkomposita" verwendet. Er stellt außerdem fest, dass diese Bildungen „z.T. mit Fugenelementen" auftreten und ordnet sie der Komposition zu (Schindler 1991: 603).

Ebenso beschreibt Mau (2002) sprachliche Wiederholungen im Allgemeinen und geht dabei auch auf die Aussage von Hohenhaus ein, dass ICCs ebenso im Deutschen existieren. Er beschreibt ICCs zudem im Deutschen in Anlehnung an Wierzbickas (1991: 267) funktional-pragmatische Beschreibung englischer ICCs. Im Gegenteil zu englischen ICCs wiesen ICCs im Deutschen laut Mau keine emotionale Komponente auf (Mau 2002: 219).

Die bereits mehrfach erwähnte Publikation von Ghomeshi et al. (2004) untersucht ICCs im Englischen, etwa *salad-salad* 'green salad as opposed to salads in general'. Die Autor:innen prägen den Begriff „contrastive focus reduplication". Mit der Aussage, das Deutsche gehöre nicht zu den ICC-Sprachen, nehmen sie auch zum Deutschen Bezug (Ghomeshi et al. 2004: 312).

Aus typologischer Perspektive behandelt Stolz Reduplikation als Wortbildungsprozess und bespricht im Zuge dessen auch die Bildung von ICCs im Deutschen. Das der Mediensprache entstammende *Filmfilm* ordnet er ähnlich wie Hohenhaus als „unikale, mediensprachliche Idiosynkrasie" ein (Stolz 2006: 116). Stolz äußert darüber hinaus die Ansicht, dass „W[ort]I[teration] von Substantiven im Deutschen nicht musterhaft" ist (Stolz 2006: 116). In Bezug auf die Bildungen vom Typ *Salat-Salat* äußert er die Ansicht, dass zwischen Erst- und Zweitglied eine gewöhnliche Modifikator-Kopf-Relation vorliege (Stolz et al. 2011: 203).

Einige Texte, die nicht dezidiert ICCs behandeln, führen zudem ICCs mit Eigennamenbasis auf, die ansonsten nur Kentner (2017) im Detail bespricht. So gibt es Publikationen aus der Wirtschaftslinguistik,[8] die ICCs zu den Wortbildungsmustern der Markennamenbildung zählen. Zu erwähnen ist hier zum einen ein Text

8 Die sich zu Beginn des 20. Jahrhunderts (teilweise im Umfeld der Prager Schule) entwickelnde Disziplin „Wirtschaftslinguistik" untersucht(e) die Wirtschaftssprache als Kommunikationsmittel. Die Wirtschaftslinguistik konnte sich allerdings nicht als eigene Disziplin halten. Heute werden die linguistischen Themen der Wirtschaftslinguistik (Wirtschaftssprache) in der Fachsprachenforschung behandelt, die betriebswirtschaftlichen Themen (Unternehmenskommunikation) in der Kommunikationswissenschaft (Steinhauer 2000: 203).

von Platen (2013), der eine Zusammenstellung von Techniken der Markennamenbildung ist und N+N-Komposita mit identischen Konstituenten als „Echowörter" (Platen 2013: 53) mit Monoreferenz erwähnt. Die Funktion solcher ICCs sieht er vor allem in der Expressivität, die bei solchen Echowörtern entsteht (Platen 2013: 53). Platen subsumiert diese ICCs dabei unter die Konzeptformen und Eigennamenbildungen und sieht solche Bildungen, etwa *Dior-Dior* (referierend auf ein Parfum des Pariser Luxusgüterherstellers *Dior*), als eine Variante der emphatischen Wiederholung und affektiven Reduplikation an, die „bei der Produktnamensgebung gezielt eingesetzt" werde (Platen 2013: 54).

Zudem gibt es ein sprachtypologisches Übersichtspapier von Kauffman (2015), das am Rande auch ICCs behandelt, nämlich solche, die Raumnamen als Basis haben und wiederum Toponyme (Ortsnamen) bilden. Für solche Bildungen nennt Kauffman für das Deutsche etwa *Baden-Baden* (Kauffman 2015: 4):

> Place names around the world also employ reduplication in such examples as *Baden-Baden* (Germany), *Puka Puka* (Cook Islands), and *Tawi Tawi* (Philippines).
>
> (Kauffman 2015: 4)

Zudem gibt es, etwa bei Pseudonymen im Internet, ICCs, deren Basen Anthroponyme sind, etwa das bereits erwähnte *Sabinesabine*. Kauffman nennt den Prozess dahinter „Name Doubling" und erkennt bei solchen Bildungen eine soziale Funktion:

> Name Doubling (Reduplication) [is] primarily used for close relationships, imparting an endearing quality that implies a quality of likability. This form is common in English and Chinese. Take some English nicknames for example: *Jon-Jon*, *Lou-Lou* […]. They add an affable quality to the way an individual addresses a friend or family member.
>
> (Kauffman 2015: 3)

Alle genannten Texte zu ICCs mit Eigennamenbasis sind allerdings weder empirische Studien, noch ausführliche theoretische Auseinandersetzungen mit ICCs. Sie diskutieren etwa nicht die Frage, ob die Bildungen „ohne Umweg über die Semantik" (Nübling et al. 2015: 27), referieren und also wie Eigennamen „ohne inhärente Bedeutung" (Leys 1989: 150) sind. Alle Annahmen zu Form und Funktion dieser ICCs stützen sich zudem auf bloße Introspektion. Über einzelne Beobachtungen und Erwähnungen hinaus fanden solche ICCs keine Beachtung seitens der germanistischen Sprachwissenschaft.

2.4 Publikationen zu ICCs in anderen Sprachen

Auch in anderen Sprachen werden ICCs untersucht, und zwar vornehmlich solche, die auf Prototypen referieren. Diese ICCs sind etwa für das Italienische (Medici 1959), Spanische (Roca & Suñer 1998), Französische (Mau 2002) und Niederländische (van Santen & Booij 2017) beschrieben. Die meisten Publikationen zu ICCs gibt es zum Englischen. In der jüngsten Vergangenheit ist hier eine Vielzahl von Texten erschienen, die wiederum weitestgehend ICCs mit Prototypenlesart behandeln. Zu erwähnen sind hier – neben den bereits erwähnten Veröffentlichungen von Horn (1993), Ghomeshi et al. (2004) und Hohenhaus (2004) – Dray (1987) , Whitton (2006), Lee (2007), Stolz et al. (2011), Rossi (2011), Song und Lee (2015), Huang (2015), sowie die korpusbasierte Studie von Widlitzki (2016). Mit Ausnahme von letzterer sind diese Texte allerdings rein theoretische Auseinandersetzungen, die sich ohne empirische Datengrundlage mit dem Thema „Contrastive Focus Reduplication“ befassen. Auf diese Publikationen wird hier nicht im Detail eingegangen. Die Texte sind allesamt Beschreibungen von ICCs mit Prototypenlesart im Rahmen konkreter pragmatischer, kognitiver oder generativer Theorien. Die Annahmen und Einordnungen können aber nur schwer auf ICCs im Deutschen übertragen werden. Eine solche theoretische Einordnung von ICCs im Deutschen wird in dieser Arbeit erst auf der Grundlage der empirisch erhobenen Daten erfolgen. Die wesentlichen Punkte dieser Texte, die die allgemeineren Aspekte von ICCs betreffen, sind aber in den Texten zum Deutschen enthalten und wurden bereits referiert.

Die Korpusstudie von Widlitzki verdient hingegen eine nähere Betrachtung, da ihre Datenbeschreibung Annahmen zu ICCs im Deutschen motiviert. Widlitzki legt ihrer Studie drei große Korpora zugrunde, nämlich das Blog Authorship Corpus (Schler et al. 2006), das Corpus of Contemporary American English (Davies 2010) sowie das Corpus of American Soap Operas (Davies 2012). Insgesamt erhebt Widlitzki 157 nominale ICCs. Sie untersucht auf dieser Grundlage vor allem soziodemographische Aspekte und zeigt, dass vor allem jüngere Menschen Gebrauch von Contrastive Focus Reduplication machen und Frauen eher dazu neigen als Männer (Abbildung 4).

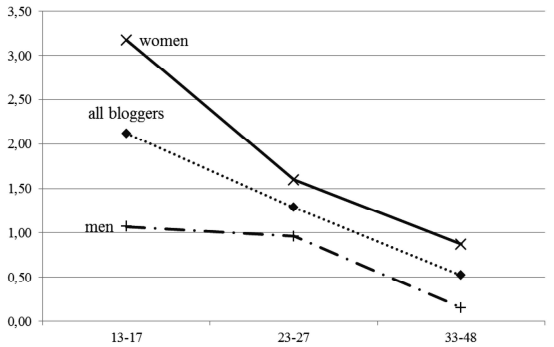

Abb. 4: Verwendung von Contrastive Focus Reduplication als Funktion der Faktoren Alter & Geschlecht (Blog Corpus, Schler et al. 2006) aus Widlitzki (2016: 136).

Ein weiteres Ergebnis aus der Studie ist, dass die Bedeutung der Bildungen im Kontext weiter spezifiziert wird. So verwenden die Sprecher:innen in den meisten Fällen, nämlich in 69% der Fälle, kontrastierende Elemente sowie in 39% der Fälle Synonyme oder Paraphrasen, um den Adressat:innen die Interpretation zu erleichtern. Zwar behandelt Widlitzkis Studie allein englische Bildungen und ihre Ergebnisse betreffen zudem Contrastive Focus Reduplication im Allgemeinen, also auch adjektivische, verbale und adverbiale Bildungen. Für die Untersuchung von ICCs im Deutschen stellt die Studie dennoch eine Vergleichsfolie dar. Es stellt sich die Frage, ob sich die Disambiguierungsstrategien, die Widlitzki für Contrastive Focus Reduplication im Englischen klassifiziert und quantifiziert, in ähnlicher Weise bei ICCs im Deutschen beobachten lassen.

2.5 Fazit zum Forschungsstand

Die wissenschaftliche Auseinandersetzung mit ICCs im Deutschen ist überschaubar. Es fehlen bisher vor allem empirische Arbeiten. Kentner (2017) erhebt in seiner Akzeptabilitätsstudie Daten zur Ablaut- und Reimreduplikation, aber nicht zum Phänomen ICC. Freywald (2015) verwendet zwar unter anderem Korpora, bezieht aus diesen Daten aber lediglich die Beispiele, die ihre theoretischen Überlegungen begleiten. Die Arbeit von Bross und Fraser (2020) ist ebenfalls eine rein theoretische. Die nicht explizit ICCs behandelnde Untersuchung Günthers aus den Siebzigerjahren und die Pilotstudie von Finkbeiner (2014) sind daher die bisher einzigen empirischen Studien zu ICCs im Deutschen. ICCs können daher nur schwer zu kanonischen N+N-Komposita in Bezug gesetzt werden. Es bleiben bis

dato viele grundlegende Fragen unbeantwortet: Sind ICCs im Deutschen ein Randphänomen? Inwieweit teilen ICCs die formalen und funktionalen Eigenschaften von N+N-Komposita? Sind ICCs überhaupt Komposita? Wie erhalten sie ihre Bedeutung, allein über den Kontext? Wie lässt sich die Modifikation innerhalb der ICCs charakterisieren? Findet überhaupt Modifikation statt?

Der nun folgende empirische Teil der vorliegenden Arbeit soll die Datengrundlage zu ICCs im Deutschen verbessern. Bevor aber Überlegungen zu den genannten Fragestellungen angestellt und die dahinterstehenden Prozesse grammatiktheoretisch eingeordnet werden, muss eine ausreichende Datenbasis erhoben worden sein. Im folgenden Kapitel wird beschrieben, welche Fragen mithilfe der Datenbasis beantwortet werden sollen, um welche Daten es sich handelt und mit welcher Methodik sie erhoben wurde.

Teil II: **Empirie zu ICCs im Deutschen**

3 Beschreibung der Korpusstudien

Dieses Kapitel stellt den Einstieg in den empirischen Teil der Arbeit dar. Zwei korpusbasierte Untersuchungen liefern Daten zu ICCs und anderen Dopplungen, die in geschriebenen Texten vorkommen. Mit diesen Belegen kann man Aussagen zu formalen und funktionalen Aspekten der ICCs treffen und eine grammatiktheoretische Einordnung der ICCs vornehmen. Die Korpusdaten werden sowohl qualitativ als auch quantitativ ausgewertet. Zu Beginn des Kapitels fasse ich zunächst zusammen, welche Forschungsfragen sich im Laufe des Forschungsüberblicks entwickelt haben (3.1). Daraufhin stelle ich die beiden verwendeten Korpora (3.2) und die Methodik der Datenerhebung (3.3) vor.

3.1 Forschungsfragen

Die Korpusstudien sollen Antworten auf folgende die Form und die Bedeutung von ICCs betreffende Fragestellungen liefern:

In Bezug auf die Semantik und Funktion von ICCs:

1. Wie lassen sich zu ICCs Subgruppen bilden?
2. Welche Funktionen üben ICCs aus?

In Bezug auf die Produktivität der ICCs:

3. Welche Art von ICC wird besonders häufig verwendet?
4. Wie produktiv sind die jeweiligen ICC-Arten; also wie geneigt sind Sprecher:innen, neue Bildungen auf der Basis bestehender ICCs zu bilden?

Produktivität bedeutet hier nicht Frequenz, sondern ist eine Funktion aus den Merkmalen Transparenz, Regularität (Abwesenheit von Ausnahmen) und Generalität (Welche Basen erlaubt das Muster?, Baayen 1992, 2009). Zwar ist Produktivität ein inhärent diachrones Merkmal und betrifft das Potential eines Wortbildungsmusters. Die vorliegende Arbeit untersucht allerdings ICCs im Gegenwartsdeutschen und in den entsprechenden Kapiteln, in denen dieses Merkmal von ICCs untersucht wird, wird Produktivität deshalb auf der Grundlage der Stamm-Beleg-Verhältnisse sowie der Anzahl der Hapax legomena zu beurteilen versucht. Hier folge ich der Ansicht von Bauer und Kolleg:innen (2019):

https://doi.org/10.1515/9783111315416-003

> Productivity is shown by, and potentially also encouraged by, a number of low-frequency items in a single schema, more often than by single exemplars of high frequency.
> (Bauer et al. 2019: 76)

In Bezug auf die lexikalische Struktur und mögliche Selektionsbeschränkungen des ICC-Schemas gibt es folgende Fragestellungen:

5. Können Lexeme identifiziert werden, die besonders häufig als ICC auftreten?
6. Welche Eigenschaften haben die Basen, die besonders häufig als ICC auftreten?

Hinsichtlich dieser Fragen sollen drei Aspekte untersucht werden. Erstens wird der semantische Aspekt der Polysemie / Homonymie erfasst, denn es liegt nahe, dass dieser einen Einfluss darauf hat, ob und wie häufig ein Nomen Basis für ICC-Bildung ist. Brunner und Kolleg:innen (2021) konnten beispielsweise zeigen, dass die Mehrdeutigkeit von Nomina bedingt, ob sie in Komposita präferiert als erste oder zweite Konstituente auftreten. Es ist darüber hinaus möglich, dass ICCs im Deutschen – wie es für englische ICCs angenommen wird (Benjamin 2018, Widlitzki 2016) – vor allem dann auftreten, wenn das entsprechende Basislexem disambiguiert werden muss, was vor allem bei mehrdeutigen Nomina nötig ist. Zweitens wird untersucht, wie sehr die formalen Merkmale der Basen die Chancen von Nomina erhöhen, in ICCs vorzukommen. Hierzu sollen mehrere Merkmale der Komplexität erfasst werden. Da ICCs das sprachliche Material der Basis verdoppeln, ist es vorstellbar, dass die Basen der ICCs nicht allzu komplex sein dürfen. Kurze, monosyllabische und monomorphematische Nomina eignen sich womöglich besser für die Verwendung in ICCs als komplexe Basen mit viel sprachlichem Material. Drittens ist der Faktor Frequenz zu untersuchen. In der linguistischen Forschung wurde nachgewiesen, dass die Frequenz sprachlicher Einheiten einen großen Einfluss auf sprachliche Strukturen haben kann (Haspelmath 2006, 2008).

Im Forschungsüberblick wurde darüber hinaus dargestellt, dass ICCs oft zugeschrieben wird, hinsichtlich der flexionsmorphologischen Eigenschaften, der Verfugung und der Schreibung von anderen N+N-Komposita abzuweichen, und dass diese formalen Merkmale auch dabei helfen, die Dekodierung der Bildungen durch die Adressat:innen zu unterstützen (Bross & Fraser 2020, Finkbeiner 2014, Freywald 2015). In Bezug auf die morphosyntaktische Struktur sowie die genannten formalen Eigenschaften der ICCs selbst gibt es also folgende Forschungsfragen:

7. Weichen ICCs hinsichtlich der Verwendung von Fugenelementen von kanonischen Komposita ab?
8. Verhalten sich ICCs hinsichtlich der Verwendung von Fugenelementen einheitlich oder sind Fugenelemente ein formales Mittel, die jeweilige Semantik der ICCs anzuzeigen?
9. Weichen ICCs hinsichtlich der internen Flexion und lexikalischen Integrität von kanonischen Komposita ab?
10. Verhalten sich ICCs hinsichtlich der Flexion einheitlich oder ist das Flexionsverhalten ein formales Mittel, die jeweilige Semantik der ICCs anzuzeigen?
11. Weichen ICCs hinsichtlich der Schreibung von kanonischen Komposita ab?
12. Verhalten sich die ICCs hinsichtlich der Schreibung einheitlich oder ist die Schreibung ein formales Mittel, die jeweilige Semantik der ICCs anzuzeigen?

Schließlich stellen sich folgende Forschungsfragen in Bezug auf die Kontextabhängigkeit von ICCs und die Disambiguierungsstrategien der Sprecher:innen, die ebenfalls in der einschlägigen Literatur zu englischen und deutschen ICCs diskutiert werden (Hohenhaus 2004, 2015, Widlitzki 2016):

13. Muss die Verwendung von ICCs immer durch den Kontext lizensiert werden und wenn ja, in welchem Umfang?
14. Mit welchen kontextuellen Mitteln unterstützen die Sprecher:innen die Adressat:innen bei der Interpretation der ICCs?
15. Welche weiteren sprachlichen Mittel gewährleisten, dass die Adressat:innen ein ICC korrekt interpretieren?

Zur Beantwortung dieser Forschungsfragen werden zwei Korpusstudien durchgeführt. Diese haben eine unterschiedliche Abfragemethodik und ihnen liegen zwei unterschiedliche Korpora zugrunde. Durch die unterschiedliche Methodik tragen die zwei Korpusstudien in unterschiedlichem Maße zu der Beantwortung der Forschungsfragen bei und ergänzen sich dabei gegenseitig. Die Abfrage des ersten Korpus erfolgt lexembasiert und erfasst daher nur die Dopplungen einer begrenzten Anzahl Basislexeme. Diese Abfrage liefert deshalb weder Daten zu phonologischen Dopplungen noch erfasst sie Dopplungen von Nomina, die nicht auf der vordefinierten Liste stehen. Durch das klar definierte und auf wenige Lexeme limitierte Itemset ist es aber möglich, die Spatienschreibung zu berücksichtigen. Außerdem können auf diese Weise Daten zur Frequenz und zur Bedeutung der untersuchten Basislexeme erhoben werden. Diese erste Korpusstudie dient somit vor allem der Beantwortung der Forschungsfragen 4 und 6. Auch für die Behandlung der Forschungsfragen 13 bis 15, die die kontextuelle Verankerung der ICCs

betreffen, wird diese erste Korpusabfrage zugrundegelegt, da eine solche qualitative Untersuchung sehr zeitaufwendig und darum nur auf der Grundlage der kleineren Datenbasis zeitökonomisch möglich ist.

Die zweite Korpusstudie basiert nicht auf einem vordefinierten Itemset, sondern erfasst mithilfe eines regulären Ausdrucks alle im Korpus auftretenden Wiederholungen von Zeichenketten. Die große Datenmenge, die dieses Vorgehen hervorbringt, lässt es nicht zu, zu den Basislexemen externe Informationen zur Frequenz und Bedeutung hinzuzufügen. Auch kann der Kontext für die einzelnen Belege nicht qualitativ ausgewertet werden. Stattdessen liefert aber vor allem diese Studie Daten zur Beantwortung der Forschungsfragen 1 bis 3, die die Verwendung des ICC-Schemas und die Produktivität im Allgemeinen betreffen. Durch die offene Abfragemethodik ermöglicht die zweite Korpusstudie zudem eine geeignetere Datenbasis, um die Fragen 5, sowie die Fragen 7 bis 12 zu formalen Aspekten von ICCs zu beantworten.

Für die Beantwortung einiger Aspekte werden die Datensätze aus den beiden Korpusabfragen kombiniert. Das betrifft vornehmlich die Forschungsfragen 1 und 2, die die Funktionen von ICCs behandeln, die Forschungsfrage 11 und 12, die die Schreibung von ICCs betreffen sowie Forschungsfrage 15 zu weiteren Möglichkeiten, wie ICCs richtig verstanden werden können.

3.2 Korpora

3.2.1 DECOW16

Das erste Korpus, das untersucht wird, ist das DECOW16 (Schäfer 2015, Schäfer & Bildhauer 2012) in der Fassung vom 10.08.2017. Dieses Korpus umfasst knapp 20 Milliarden Token, die aus über 17.000 Texten stammen.[9] Das Korpus enthält linguistisches Material aus dem deutschsprachigen World Wide Web, da alle Texte aus deutschsprachigen Internetseiten akquiriert wurden. Die Texte sind aus urheberrechtlichen Gründen nicht als Ganzes zugänglich. Stattdessen werden die Sätze unzusammenhängend präsentiert, sodass das Korpus letztlich einen „bag of sentences“ (http://corporafromtheweb.org/category/corpora/german/) darstellt.

9 Die Internetpräsenz des Korpus macht hier widersprüchliche Angaben zur Token- und Textanzahl. An einer Stelle wird eine Tokenanzahl von 20.495.087.352 angegeben (http://corporafromtheweb.org/decow16/) an anderer Stelle hingegen ist von 19.835.843.151 Token die Rede (https://www.webcorpora.org/bonito/run.cgi/corp_info?corpname=decow16a&struct_attr_stats=1&subcorpora =1). Auch die Angaben zur Textanzahl weichen leicht voneinander ab.

Für die hier vorgestellte Korpusuntersuchung ist das aber kein großes Problem, da der von DECOW16 bereitgestellte Kotext ausreicht, um die Belege angemessen zu kodieren. Für die weitere qualitative Untersuchung können die Belege über die URL bis zu ihrem Ursprung zurückverfolgt werden, sofern die Ursprungsseite noch existiert oder mithilfe des Internetarchivs (https://archive.org) rekonstruiert werden kann. In den Fällen, in denen das nicht der Fall ist, kann über die Einstellungen der Abfrageoberfläche NoSketchEngine der KWIC-Kotext erweitert werden, sodass mehrere aufeinanderfolgende Sätze einsehbar sind. So können die Belege in jedem Fall angemessen kodiert werden.

DECOW16 eignet sich zum einen für das hier vorgestellte Forschungsvorhaben, da das Korpus sehr groß ist. ICC-Daten (beziehungsweise CR-Daten) aus Korpora zu extrahieren wird allgemein als sehr schwierig beschrieben (für das Französische: Blauth-Henke 2008, für das Englische: Hohenhaus 2004, Widlitzki 2016: 125). Die Größe des Korpus verspricht trotz dieser schwierigen Auffindbarkeit von ICCs viele Belege. Zudem schreibt Widlitzki: „CR is notoriously difficult to find in reference corpora“ (Widlitzki 2016: 124). DECOW16 beinhaltet allerdings vornehmlich Texte aus computervermittelter Kommunikation und Texte, die interaktionsorientiert und größtenteils unredigiert sind. Viele Texte stammen nicht von professionellen Schreiber:innen, sondern von Laien, die privat kleinere Webseiten oder Blogs erstellen. Solche Texte oszillieren konzeptuell zwischen Mündlichkeit und Schriftlichkeit (Beck 2006: 93, Beißwenger & Storrer 2008: 300, Storrer 2018: 220). Das ist wichtig, da ICCs an solche dialogorientierte Situationen gebunden sind (“bound to highly specific conversational situations”, Finkbeiner 2014: 205). Durch die begrenzte Reichweite und private Natur der Texte besteht zudem für einen Teil der Verfasser:innen kein wirtschaftlicher Zwang zur normgerechten Sprache. All diese Umstände fördern den Anteil konzeptionell mündlicher Sprache im Internet (Koch & Oesterreicher 1985, 1990, Schlobinski & Siever 2000: 57), sodass dieses Korpus trotz der medialen Schriftlichkeit auch ungesteuerte Kommunikation enthält, die innovative sprachliche Techniken weniger hemmt als das beispielsweise bei Korpora überregional gedruckter Zeitungen der Fall wäre (Albert 2013: 163ff.).

3.2.2 deTenTen13

Das zweite untersuchte Korpus ist das deTenTen13 in der Fassung von 2013 (Jakubíček et al. 2013). Dieses Korpus umfasst wie das DECOW16 zirka 20 Milliarden

Token, die aus über 50.000 Dokumenten stammen.[10] Das Korpus bietet ebenfalls Material aus dem deutschsprachigen World Wide Web und eignet sich aus denselben Gründen wie das DECOW16 für die Untersuchung des Phänomens ICC. Die Sketch Engine, über die das Korpus zugänglich ist, bietet aber eine genauere CQL-Abfrage (Corpus Query Language) als die NoSketch-Engine, über die das DECOW16 abgefragt wird. Auf diese Weise lassen sich genaue Bedingungen für den Output definieren.

3.3 Methodik

3.3.1 DECOW16

3.3.1.1 Datenerhebung

Die Korpusuntersuchung im DECOW16 wurde lexembasiert durchgeführt. Eine strukturbasierte Abfrage des Korpus, bei der nach morphologisch gebildeten Verbindungen mit adjazent identischen nominalen Einheiten gesucht würde, liefert zwar große Datenmengen, erfordert aber einen großen Filteraufwand. Die Annotation des DECOW16 ist nicht genau genug, um einen Output genau dieser Strukturen zu generieren. Es käme also zu Suchergebnissen, die alle möglichen Formen von Dopplungen beinhalteten. Da bei der Untersuchung dieses Korpus auch die Spatienschreibung berücksichtigt wird, wäre die Anzahl der Falsch-Positiv-Fehler enorm. Die Daten wären daher nur mit größtem Zeitaufwand auszufiltern, zu klassifizieren und zu kodieren. Zudem erlaubt die Weboberfläche des DECOW16 nicht die benötigten regulären Ausdrücke und Attribute, die Stammallomorphien erfassen können. Dadurch würden einseitig diejenigen Basen, die abweichende Kompositionsstammformen haben (*Mann*, *Männ-*), nicht erfasst. Aus diesen Gründen wurde im DECOW16 lediglich eine vordefinierte Liste nominaler Basen auf reduplikative Muster hin abgefragt.

Diese Liste basiert auf der World Loanword Database (WOLD, Haspelmath & Tadmor 2009). Die WOLD beinhaltet 1814 Basiskonzepte aus 24 semantischen Feldern. Tabelle 1 gibt einen Eindruck davon, wie diese semantischen Felder aufgebaut sind.

10 Auch hier macht die Internetpräsenz des Korpus widersprüchliche Angaben zur Token- und Textanzahl. Auf der Seite der Sketch Engine wird die Tokenanzahl einmal mit 16,5 Milliarden angegeben, auf der Infotafel innerhalb der Sketch Engine-Maske hingegen auf genau 19.808.173.163 beziffert (https://www.sketchengine.eu/deTenTen13-german-corpus).

Tab. 1: Semantische Felder in der WOLD (ausgewählt: „Animals“)

	Semantische Felder	Anzahl der Bedeutungen			Basiskonzepte
1	The physical world	86			
2	Kinship	102			
3	Animals	155	→	3.11	the animal
4	The body	171		3.12	male(2)
5	Food and drink	103		3.13	female(2)
6	Clothing and grooming	68		3.15	the livestock
7	The house	53		3.16	the pasture
8	Agriculture and vegetation	128		3.18	the herdsman
9	Basic actions and technology	112		3.19	the stable or stall
10	Motion	89		3.20	the cattle
11	Possession	53		3.21	the bull
12	Spatial relations	77		3.22	the ox
13	Quantity	47		3.23	the cow
14	Time	62		3.24	the calf
15	Sense perception	55		3.25	the sheep
16	Emotions and values	63		3.26	the ram
17	Cognition	58		3.28	the ewe
18	Speech and language	53		3.29	the lamb
19	Social and political relations	76		3.32	the boar
20	Warfare and hunting	48		3.34	the sow
21	Law	28		3.35	the pig
22	Religion and belief	40		3.36	the goat
23	Modern world	58		...	...
24	Miscellaneous function words	29			

Die in der WOLD auf Englisch repräsentierten Konzepte wurden von mir ins Deutsche übersetzt und, sofern sie im Deutschen nominal ausgedrückt werden, in die Abfrageliste übernommen. Dies trifft auf 1034 dieser Konzepte zu. Die entsprechenden Nomina bilden die Liste der Items, die im Rahmen der vorliegenden Studie untersucht werden.

Die Auswahl der Nomina mithilfe der WOLD bietet einige Vorteile: Zunächst gewährleistet die Liste eine heterogene Zusammenstellung von Konzepten, sodass die Ergebnisse der Korpusuntersuchung nicht nur auf einen beschränkten Teil

des Lexikons zutreffen. Die Konzepte umfassen etwa ebenso Konkreta wie Abstrakta und ebenso Belebtes wie Unbelebtes. Des Weiteren ist die Auswahl der Konzepte in der WOLD nicht anfällig für Frequenzunterschiede innerhalb einer Einzelsprache. Dadurch werden im Deutschen hochfrequente Nomina wie etwa *Auto* erfasst, aber auch niedrigfrequente wie beispielsweise *Laken*. Außerdem werden durch dieses konzeptbasierte Vorgehen Nomina des Deutschen unabhängig von ihrer morphologischen Komplexität ausgewählt. Komplexe Nomina wie *Regierung* oder *Schraubendreher* sind ebenso enthalten wie die Simplizia *Polizei* und *Haus*.

Zu allen 1034 Lexemen wurden zudem grammatische sowie die Frequenz und Bedeutung betreffende Informationen erfasst. Hierzu wurde das Wortschatz-Korpus der Universität Leipzig konsultiert (Goldhahn 2012, Quasthoff & Richter 2005). Zu jedem Item wurde auf dieser Grundlage das Genus sowie die Morphem-, Graphem-, und Silbenanzahl bestimmt. Diese Merkmale sollen eine automatisch auszuwertende Größe für die Komplexität der ICCs bilden und helfen bei der Beantwortung von Forschungsfrage 6 zu den formalen Eigenschaften, die die Verwendung eines Nomens als ICCs fördern oder hemmen. Für sich genommen sagt jedes dieser Merkmale nur begrenzt etwas über die Komplexität der untersuchten Nomina aus. Zusammengenommen können sie aber bei der Auswertung der Daten einen Eindruck davon geben, ob möglicherweise die Bildung von ICCs bevorzugt mit besonders kurzen Stämmen geschieht. Zudem wurde zu den belegten ICCs die Zeichenanzahl bestimmt, um auch den Einfluss dieses Faktors bei der Auswertung zu berücksichtigen.

Außerdem wurde dem Wortschatz-Korpus – ebenfalls in Bezug auf Forschungsfrage 6 – die Anzahl der Dornseiff-Bedeutungsgruppen (Dornseiff 1934, Dornseiff et al. 2010) entnommen. Die Dornseiff-Bedeutungsgruppen bieten für die deutsche Sprache eine Einteilung der Wörter in semantische Bedeutungsgruppen und ermöglichen eine Eingrenzung des Verwendungsbereichs der Lexeme. Die Dornseiff-Bedeutungsgruppen unterscheiden sich von den semantischen Feldern des WOLD dahingehend, dass sie alle Verwendungsbereiche eines Lexems umfassen und ein Lexem somit meist mehr als nur einer Bedeutungsgruppe zuordnen. Ein Lexem, das in vielen Bedeutungsgruppen auftritt und also in vielen unterschiedlichen Lebensbereichen verwendet wird, hat demnach eine weniger spezifische Semantik als ein Lexem, das nur in einer oder wenigen Bedeutungsgruppen vertreten ist. Auf diese Weise kann ein etwaiger Zusammenhang zwischen der Anzahl an Bedeutungsgruppen eines Lexems und seiner Auftretenshäufigkeit in den ICC-Konstruktionen ermittelt werden.

Auch die Frequenz der Nomina wurde erhoben. Hierfür wurden zwei unterschiedliche Quellen konsultiert: das digitale Wörterbuch der deutschen Sprache

(DWDS) (Klein & Geyken 2010) und wiederum der Leipziger Wortschatz (Goldhahn 2012, Quasthoff & Richter 2005). Beide Quellen teilen die Lexeme nach ihrer relativen Frequenz in Frequenzklassen ein. Das DWDS unterscheidet hierbei 7 Klassen, wobei Klasse 1 die niedrigste, Klasse 7 die höchste Frequenzklasse ist. Der Leipziger Wortschatz ordnet die Lexeme in 24 Frequenzklassen ein, wobei hier Klasse 1 die höchste und die Klasse 24 die niedrigste Frequenzklasse darstellt. In der Auswertung kann durch diese Frequenzdaten untersucht werden, ob es Frequenzeffekte gibt, dergestalt, dass die Frequenz eines Lexems einen Einfluss auf seine Auftretenshäufigkeit in der ICC-Konstruktion hat.

Die Abfrage des Korpus erfolgte über die NoSketchEngine (www.webcorpora.org, Legacy interface, Version 2.36.7-SkE-2.160.3-CA1.93.18). Zu den 1034 Items wurden mithilfe der zwei Abfragestrings [x*x*] und [x*_x*] Doppelformen verschiedener Schreibweisen erfasst. Der Platzhalteoperator * steht hierbei für 0 bis n Zeichen beliebigen Typs.[11] Die zwei Abfragen [hund*hund*] und [hund*_hund*] lieferten demnach beispielsweise alle adjazent mehrfachen Auftreten des Lexems *Hund*, also alle Fälle, in denen auf *Hund* mindestens einmal unmittelbar wieder *Hund* folgt, seien die Konstituenten durch Fugenelemente (*Hundshund*), Bindestriche (*Hund-Hund*), Flexive (*Hundeshundes*), Spatium (*Hund Hund*) oder andere Zeichen (*Hundihund*) getrennt. Erfasst wurden auf diese Weise auch alle Wortformen des Lexems sowie alle Kombinationen von Groß- und Kleinschreibungen.[12] Zusätzlich zu Doppelverbindungen wurden so auch beliebig vielfache Verbindungen identischer Nomina erfasst, die hinterher aber aus den Daten entfernt wurden. Für Nomina mit Stammallomorphie wurden die Abfragen zusätzlich mit den von der Zitierform abweichenden Stämmen durchgeführt, beim Lexem *Wolle* also etwa zusätzlich mithilfe der Strings [woll*woll*] und [woll*_woll*].

Die Ergebnisse konnten nicht mithilfe von NoSketch auf morphologische Verbindungen beschränkt werden. Das einzige Kriterium für eine solche maschinelle Filterung wäre die Schreibung gewesen, dergestalt, dass alle durch ein Leerzeichen getrennten Belege als syntaktisch interpretiert und aussortiert worden wären. Da das orthographische Kriterium für die Bestimmung von Komposita nur wenig zuverlässig ist und die Spatienschreibung auch für morphologische Verbindungen verwendet wird (Fleischer & Barz 2012: 192, Scherer 2012), habe ich

11 Grundformoperatoren wie etwa „&“ wurden nicht benötigt, da die NoSketchEngine grundsätzlich alle Flexions- und Wortbildungsformen liefert.

12 Die NoSketch-Engine erlaubte während des Zeitraums der Abfrage wegen technischer Probleme seitens der Anbieter ausschließlich die Eingabe von Minuskeln, lieferte aber Ergebnisse unabhängig von der Groß- und Kleinschreibung.

mich für die zeitaufwendigere Lösung entschieden, auch die syntaktischen Verbindungen manuell zu kodieren. Nur so lässt sich die Anzahl der Falsch-Negativ-Fehler reduzieren und der Recall erhöhen (Perkuhn et al. 2012: 40).

Nicht kodiert, sondern gänzlich manuell herausgefiltert wurden lediglich diejenigen Falsch-Positiv-Fehler, in denen die identischen Konstituenten durch andere Elemente als morphologische Marker getrennt sind (etwa durch andere Konstituenten: *Hundesuchhund*[13]) sowie Wörter, die Teil einer SEO-Strategie,[14] etwa einer Schlagwortwolke (Tagcloud[15]) sind. Bei der Abfrage des Items *Auto* erhält man beispielsweise unter anderem folgenden Beleg:

(30) *...Kraftfahrzeug . auto autohaendler* ***auto autos auto*** *billig auto billig kaufen auto auto gebraucht online auto gebrauchtwagen....*

<www.yabeba-mobile.de/auto26gateway.htm>

Weil die Aneinanderreihung der Wörter im Rahmen der SEO-Strategie aber vollständig automatisch geschieht, sollten die daraus entstehenden Wortfolgen nicht als durch Sprachbenutzer:innen produzierte Sprache angesehen werden. Treffer dieser Art wurden nicht in den Datensatz aufgenommen.

Außerdem herausgefiltert wurden Belege, in denen die Basen mehr als einmal wiederholt werden:

(31) *wann's in dem Zahn wieder losgehen wird ;(* ***Angst Angst Angst***

<www.yabeba-mobile.de/auto26gateway.htm>

Der Abfragestring beschränkt die Dopplungen nicht auf eine Wiederholung und liefert daher vor allem Repetitionen, in denen ein Nomen wie in (31) dreimal hintereinander vorkommt.
Zusätzlich zu den erhobenen reduplikativen Belegen werden bei Bedarf, nämlich wenn es für Vergleiche aufschlussreich ist, Nomina im Allgemeinen auf be-

13 Der Platzhalteroperator „*“ liefert zwar solche Falsch-Positiv-Fehler, garantiert aber, dass auch Kombinationen adjazent identischer Lexeme, die durch abweichend geschriebene Fugenelemente/Flexive oder durch Sonderzeichen getrennt sind, in die Ergebnisse eingehen.
14 Unter SEO (engl.: Search Engine Optimization) versteht man eine Reihe technischer und inhaltlicher Maßnahmen, das Ranking einer Webseite in Online-Suchmaschinen zu verbessern. SEO ist ein wichtiger Bereich des Suchmaschinenmarketings und daher weit verbreitet.
15 Tagclouds sind eine Methode zur SEO-Optimierung, bei der eine Liste aus Schlagwörtern erstellt wird, die gleichzeitig auf der Seite eine Menge an seitenweiten internen Links darstellen. Diese Links treten im linearen Textcode der Seite als bloße Aneinanderreihung von Wörtern auf.

stimmte Eigenschaften hin abgefragt. Dies betrifft die Häufigkeit, zu der Nomina im Korpus in doppelte Anführungszeichen gesetzt werden, die Verwendung von direkt vorausgehenden attributiven Adjektiven sowie die Kookurrenz mit den Negationswörtern *kein* und *nicht*. Um diese Informationen zu erheben, wird eine entsprechende Beschränkung auf nominale KWIKs angewendet und mithilfe der NoSketchEngine ausgewertet.

3.3.1.2 Kodierung

Im Anschluss an die Filterung wurden die ICC-Belege manuell und automatisch kodiert. Die automatische Kodierung wurde mithilfe von iWork Numbers (Version 6.1) vorgenommen und umfasste das Vorhandensein von Markern an Erst- und Zweitglied (Ja/Nein), deren Form (*-er*, *-n*, ...) sowie eventuelle Determinierer der ICCs. Bei Letzteren wurden lediglich die den ICCs unmittelbar vorangehenden Determinierer erfasst. Schließlich wurde noch die Schreibweise (Zusammenschreibung / Bindestrich / Binnenmajuskel / Spatium / Sonderzeichen je mit Verteilung der Groß- und Kleinbuchstaben) automatisch kodiert.

Für die manuelle Kodierung wurde für jeden Beleg ein KWIC-Kotext von 140 Zeichen sowie die URL der dazugehörigen Webseite (falls vorhanden) festgehalten. Auf dieser Grundlage wurde die Numerusinformation kodiert (SG/PL), das Genus (f./m./n.), die Art der morphologischen Marker (Flexionsmarker, Fugenelement), die syntaktische Funktion im Satz (Subjekt, Objekt, Prädikativum), das im Satz verwendete Prädikat, der Kopf einer etwaigen Präpositionalphrase, in die die NP eingebettet ist, sowie eventuelle Modifikatoren (Attribute und Partikeln). Außerdem wurde die Funktion, beziehungsweise die Bedeutung der Wiederholung manuell kodiert. Die Kriterien, die bei der Kodierung zugrundegelegt wurden, entsprechen jenen, die im Rahmen der Methodenbeschreibung zur deTenTen13-Korpusstudie in Kapitel 3.3.2.2 genauer beschrieben werden.

Ich beschreibe nun die automatisch erfolgte Kodierung der Erstgliedmarker etwas ausführlicher, da in der Literatur zu ICCs kontrovers darüber diskutiert wird, ob ICCs Fugenelemente tragen können oder nicht. Als Fugenelement wird, auch wegen der gewählten automatischen Kodierung, alles gewertet, „was über die Form des Nom. Sg. eines substantivischen Determinans hinausgeht" (Eisenberg 2020: 247). Für die Unterscheidung zwischen flexionsähnlichen Markern und Fugenelementen am Erstglied werden dennoch verschiedene Stufen oder Grade einer präzisen Fugenelementbestimmung angenommen. Manche Elemente am Erstglied, die über den Nom.Sg.-Stamm hinausgehen, sind flexionsaffixanaloge Einheiten und bilden gemeinsam mit dem Stamm eine Kompositionsstammform, die identisch mit einer Paradigmenposition im Flexionsparadigma des entsprechenden Substantivs ist.

Ist ein Element am Erstglied mit dem im Beleg vorliegenden Flexionsmarker am Zweitglied identisch, kann das Element am Erstglied sowohl als Fugenelement als auch als Marker, der flexionsähnliche Bedeutung trägt, angesehen werden. Beispielsweise können das *-n-* in Belegen wie *Katzen-Katzen* (32) oder das *-e-* in Belegen wie *Hundehunde* (33) ebensogut Fugenelemente wie Flexionsmarker sein. Es kann in diesen Fällen nicht ausgeschlossen werden, dass die Elemente eine Art Pluralmarker darstellen, die die Entität, auf die das Erstglied verweist, hinsichtlich Nicht-Einzahl markieren.

(32) *Ich mag am liebsten so richtige Katzen-Katzen.*

(33) *Die Eigenschaft der Hunde-Hunde ist, dass sie auf andere Hunde bezogen sind.*

Ein solches Element wird in dieser Korpusstudie – der Formgleichheit mit dem Flexionsmarker am Zweitglied wegen – "flexionsaffixanaloges Element" genannt, und zwar ungeachtet der Frage, ob diese Marker auch tatsächlich als Flexionsmarker fungieren.

Auch Fugenelemente, die sich von dem Flexionsmarker am Zweitglied des jeweiligen Belegs unterscheiden, können theoretisch flexionsähnliche Bedeutung tragen. In Belegen wie *Katzen-Katze* oder *Hundehundes* haben *-n-* und *-es-* also nur eine gewisse Wahrscheinlichkeit, Fugenelemente zu sein. Obwohl sie sich vom Marker an N2 unterscheiden, könnten sie dennoch auch als Flexionsmarker angesehen werden, da die entstehende Form des Erstgliedes Teil des Pluralparadigmas der entsprechenden Basislexeme ist (*die Katzen*, *die Hunde*). Solch ein Element wird "paradigmatisches Fugenelement" genannt (Fuhrhop 1998: 195, Nübling 2009: 198).

(34) *Eine Katzen-Katze interessiert sich für andere Katzen.*

(35) *Die Eigenschaft eines Hunde-Hundes ist, auf andere Hunde bezogen zu sein.*

Eindeutig frei von flexionsähnlicher Bedeutung sind hingegen Elemente, die nicht formgleich mit einem der Affixe im Flexionsparadigma des entsprechenden Basislexems sind. In Belegen wie *Katzes-Katzen* oder *Arbeits-Arbeit* sind die Elemente keine potentiellen Flexionsaffixe, die entstehenden Formen **Katzes* und **Arbeits* nicht Teil des Paradigmas der Lexeme *Katze* und *Arbeit*. Elemente dieses Typs werden "unparadigmatische Fugenelemente" genannt und entsprechend anders kodiert als die paradigmatischen Elemente.

3.3.2 deTenTen13

3.3.2.1 Datenerhebung

Das deTenTen13-Korpus wurde anders als das DECOW16 nicht über eine vorformulierte Itemliste abgefragt. Stattdessen wurde eine CQL-Abfrage durchgeführt, die im deTenTen13 auch komplexe reguläre Ausdrücke erlaubt (Jakubíček et al. 2010). Dadurch kann sehr präzise nach reduplikativen Mustern gesucht werden. Im Gegensatz zur lexembasierten Abfrage im DECOW16 sind hier allerdings Belege der Spatienschreibung ausgeschlossen.

Das deTenTen13 wurde über die Sketch Engine abgefragt (Kilgarriff et al. 2004 https://www.sketchengine.eu, new interface version). Die CQL-Abfrage der Engine erlaubt umfassende reguläre Ausdrücke und Attribute. Als CQL-Abfragestring habe ich den String [ascii="([[:alpha:]]{3,})e?[nr]?s?-?\1e?r?n?s?"] verwendet. Dieser String liefert als Ergebnis alle Wörter mit zwei adjazent identischen Konstituenten im Korpus. Der String besteht aus mehreren logischen Operatoren, die zusammengenommen dazu führen, dass das Suchergebnis alle aus ASCII-Zeichen bestehenden Zeichenketten beinhaltet ([[:alpha:]]), die aus mindestens 3 Zeichen bestehen ({3,}), und die sich in derselben Reihenfolge genau einmal wiederholen (\1). Zwischen der Zeichenkette und ihrer Wiederholung dürfen optional verschiedene Formen von Fugenelementen und Flexionsmarkern auftreten. Die Operatorkombination (e?[nr]?s?) erlaubt als Fugenelemente und Flexionsmarker die Grapheme <e>, <n>, <r> und <s> sowie in derselben Reihenfolge beliebige Kombinationen dieser Grapheme (abgesehen von der Kombination <nr> und <rn>), also <ens>, <ers>, <es>, <er>, <ns> und <en>.[16] Auf diese Weise werden alle für das Deutsche beschriebenen Fugenelemente erfasst (Nübling & Szczepaniak 2008: 3). Ebenso darf die Zeichenkette an dieser Stelle optional einen Bindestrich aufweisen (-?), und zwar auch in Kombination mit den Fugenelementen. Nach der Wiederholung der Zeichenkette dürfen außerdem optional als Flexionmarker die Grapheme <e>, <r>, <n> und <s> auftreten (e?r?n?s?), sowie Kombinationen die-

16 Durch den Ausschluss von <ern> könnte hier der Eindruck entstehen, dass potentiell intern flektierende ICCs, deren Basen zu den starken nicht-Feminina gehören, durch diese Operatorenkombination ausgeschlossen wären (etwa die Form *Kindern-Kindern*). Dem ist aber nicht so. Man sollte erwarten, dass hier stets auch das Zweitglied im Dativ Plural steht und nicht bloß das Erstglied (*ich gebe die Bonbons nur richtigen Kindernkindern*, aber: **ich gebe die Bonbons nur richtigen Kindernkind*, **ich gebe die Bonbons nur einem richtigen Kindernkind, *das Verhalten eines richtigen Kindernkindes*). Die Form *Kindernkindern* wird durch den Abfragestring in jedem Fall erfasst, da die Engine in dem Fall nicht *Kind*, sondern *Kindern* als zu wiederholenden Stamm ansieht. *Kindernkindern* ist für die Engine eine Wiederholung der Zeichenfolge *Kindern* und ist daher im Ergebnis enthalten.

ser, also etwa <er>, <en>, <es>, <ers>, <ens>, <erns>, <ns> oder <rs>. Durch die Toleranz von <e> als Flexionsmarker wird gleichzeitig sichergestellt, dass durch Schwa-Tilgung entstandene Kompositionsstammformen wie etwa *Woll-* des Lexems *Wolle (Wollschuhe*) erfasst werden, da *Woll-* in der potentiellen Form *Wollwolle* als Stamm interpretiert wird, an dessen identischem Zweitglied das *-e* angehängt wird.

Als Standardattribut wurde „ascii“ gewählt.[17] Zum einen neutralisiert dieses Attribut Groß- und Kleinschreibung und konvertiert <ß> zu <ss>. Zum anderen stript es alle Sonderzeichen des Deutschen, neutralisiert also den Unterschied zwischen den umlautfähigen Vokalen des Deutschen und den zugehörigen Umlauten. Dadurch lassen sich auch Dopplungen erfassen, in denen ein Teil der Dopplung aufgrund einer Kompositionsstammform oder einer Flexionsform umgelautet ist, etwa *Männermann* oder *Buchbücher*, da sie als *mannermann* bzw. *buchbucher* repräsentiert sind. Diese Funktion der Sketch Engine hilft, die Anzahl falscher Negative zu reduzieren.

Die gewählte CQL-Abfrage bietet viele Vorteile. Sie liefert die Ergebnisse nicht auf der Grundlage vordefinierter Lemmata, sondern operiert blind auf ASCII-Zeichen und dadurch sehr flexibel. Sie toleriert von der Norm abweichende Schreibvarianten potentieller ICCs (*Kafee-Kafee*), sämtliche Schreibweisen mit Ausnahme der Spatienschreibung (*Mann-Mann, MannMann, Mannmann, MANN-MANN, mannmann* etc.) und die im Deutschen am häufigsten vorkommenden Fugenelemente und Flexionsmarker, und zwar auch in Kombination mit Bindestrichen (*Mannes-Männer*). Ein weiterer Vorteil der CQL-Abfrage ist, dass das Suchergebnis von der Wortart, die der Tagger einer Zeichenkette zuordnet, unabhängig ist. Dadurch wird das Ergebnis von fehlerhaften POS-Tags nicht beeinflusst. Der wichtigste Vorteil der gewählten Methode ist aber, dass das Ergebnis von jeglicher vordefinierter Wortliste unabhängig ist. Regional oder im Substandard verwendete Ausdrücke (*bälgerbalg*) sind im Ergebnis ebenso enthalten wie standarddeutsche (*Kind-Kind*), niedrigfrequente (*Epilog-Epilog*) ebenso wie die des Kernwortschatzes (*Erdeerde*), Lehn- und Fremdwörter (*Achievement-Achievements*) ebenso wie Wörter des Erbwortschatzes (*Mädchenmädchen*). Die CQL-Abfrage liefert also ein realistisches Bild davon, welche Wörter in den Texten des deutschsprachigen Internets eine reduplikative Struktur aufweisen.

Die Vorteile der Abfrage liegen also darin, dass sie nur wenige falsch-negative Fehler verursacht. Das gewählte Abfragedesign hat jedoch auch Nachteile, die vor

17 Ich möchte an dieser Stelle einen besonderen Dank an den Support der Sketch Engine und die Softwareentwickler des Sketch Engine-Teams aussprechen, die eigens für die Bedürfnisse dieser Studie das Attribut „ascii“ hinzugefügt haben.

allem zu falsch-positiven Fehlern führen. So bedingt die Tatsache, dass die Engine durch das ASCII-Attribut nur sensitiv für Zeichen, nicht aber für Wörter ist, dass jegliche Wiederholung von mehr als zwei Zeichen zu einem Match mit dem formulierten String führt. Das Ergebnis umfasst daher auch Treffer, die nicht im engeren Sinne zu von Menschen produziertem Text zu zählen sind (*tcptcptcptcptcptcp*, *xxxxxxxxxxx*) sowie Wörter und Wortformen, bei denen eine reduplikative Struktur bloß auf Graphemebene besteht (*Kerker, Berber, Testes*). Die beabsichtigte Toleranz gegenüber Fugenelementen und Flexionsmarkern führt überdies zu Treffern, in denen die Grapheme <e>, <n>, <r> und <s> und ihre Kombinationen nicht als Fugenelemente oder Flexionsmarker fungieren, sondern Teil des Stammes sind (*Phosphor*, *Moormoos, Wetter-Wetten*). Auch erzeugt die Gleichsetzung von Umlauten und den ihnen zugrundeliegenden Vokalen ein erhöhtes Auftreten von Komposita, deren Konstituenten nicht auf dasselbe Lexem zurückgehen (*Burgen-Bürger*, *Bären-Bar, Wässerwasser*). Schließlich führt auch das fehlende POS-Tagging zu falsch-positiven Fehlern (*altes-altes*). Die Effekte dieser verschiedenen falsch-positiv-Fehler fördernden Einstellungen akkumulieren sich zu weiteren Fehlern, bei denen sich teilweise erst nach längerem Nachdenken erschließt, warum sie für die Sketch Engine eine Übereinstimmung mit dem Abfragestring bedeuten (*nennen*, *endenden*, *umpumpen*, *Schülerschule*).

Falsch-negative Fehler betreffen vor allem ICCs mit Spatienschreibung, die ebensowenig vom Abfragestring erfasst werden wie Basen, die aus weniger als drei Zeichen bestehen (*Ei*). Auch ICCs, die durch andere Zeichen als die definierten Fugenelemente oder den Bindestrich getrennt sind, etwa durch Sonderzeichen wie <#> oder <*>, sind nicht im Suchergebnis enthalten. Dieser Umstand lässt sich aber ohne Weiteres nicht ändern und wurde für diese Korpusstudie in Kauf genommen. Die ohnehin schon immense Datenmenge wäre unter zusätzlicher Berücksichtigung dieser Fälle nur mit größtem Zeitaufwand zu verarbeiten. Die Spatienschreibung hätte etwa syntaktische Verbindungen (*…was gut ist ist dass…*) und Iterationen (*sehr sehr schön*) erfasst. Der Verlust möglicher ICCs wie *Mann Mann* wird teilweise durch die auf DECOW16 basierende Korpusstudie aufgefangen, die die Spatienschreibung berücksichtigt. Potentielle ICCs wie *Eiei* oder *Bö-Bö* werden hier indes vollends ignoriert, was angesichts der wenigen digraphen Nomina des Deutschen aber nicht allzu sehr ins Gewicht fällt.

3.3.2.2 Kodierung

Die Filterung und Kodierung der Daten wurde teilweise automatisch, teilweise manuell vorgenommen. Die manuelle Filterung war notwendig, weil eine rein automatische Fehleridentifikation basierend auf dem POS-Tagging (RFTagger, Schmid & Laws 2008) und der Word Sketch Grammar der Sketch Engine zu feh-

leranfällig gewesen wäre. POS-Fehler wurden deshalb manuell herausgefiltert. Das Herausfiltern von Doubletten und reinen Zeichenrepetitionen sowie die Stammidentifikation und die Kodierung formaler Merkmale (grammatische Marker, Schreibung, Wortlänge) wurde hingegen automatisch mithilfe von iWork Numbers (Version 6.1) vorgenommen. Manuell kodiert wurden ferner die semantische / funktionale Relation zwischen den Konstituenten. Da die automatische Kodierung der Schreibung nicht die Binnengroßschreibung erfassen konnte, wurde diese, nur für ICCs mit Prototypenlesart, zusätzlich manuell kodiert.

Zur Auswertung der deTenTen13-Abfrage wurde mithilfe der Sketch Engine eine Frequenzliste aller durch den Abfragestring erfassten Zeichenkombinationen heruntergeladen. Es handelt sich bei diesen Formen nicht um Wortformen im engeren Sinne. Sketch Engine fasst lediglich Zeichenkombinationen, die exakt identisch sind, zu einem gemeinsamen Type zusammen. Formen wie *ende-ende* und *endeende* bilden also zwei unterschiedliche Types, ebenso wie *endeende* und *endeendes*. Eine automatische, mit Numbers durchgeführte Stammidentifikation ordnete diese Formen Stämmen, beziehungsweise Lexemen zu. Da die automatische Erkennung nicht auf einer Stammliste basiert, ordnet sie auch Bestandteile, die keine Wortstämme sind, einem Stamm zu. Für die phonologischen und syntaktischen Wiederholungen sind dies keine Stämme, sondern die Basis des Wiederholungsprozesses. So ist etwa bei den phonologischen Wiederholungen die Zeichenfolge *Hiti* nicht der Stamm von *Hiti-Hiti*, bei syntaktischen *alter* nicht der Stamm von *alter*, *alter (Mann)*. Für diese Studie sind vor allem die morphologischen Verbindungen interessant. Hier ordnet die Stammidentifikation Formen sinnvoll Lexemen zu. Beispielsweise werden die Wortformen *Kindeskind*, *Kindkind* und *Kindeskindern* zuverlässig dem Stamm *Kind* zugeordnet.

Die Kodierung der Daten in Hinblick auf die Semantik und die Funktion der Wiederholung ist nur manuell möglich, da meist nur eine Interpretation der Formen mithilfe des Kotextes zur korrekten Klassifikation führt. In einigen wenigen Fällen ist die Zahl der Treffer allerdings zu hoch, um eine manuelle Kodierung vornehmen zu können. Dies betrifft vor allem das Verb *nennen* und das ICC *Baden-Baden*. Bei diesen hochfrequenten Lexemen wäre es sehr zeitaufwendig, alle Fälle manuell zu kodieren. Doch betrifft dieses Problem nur 94 Types (0,3%). Für diese Types wurde die manuelle Kodierung der Bedeutung aus zeitökonomischen Gründen nur näherungsweise vorgenommen. Mithilfe der Sketch Engine wurden Zufallsstichproben mit einer Größe von 1000 Treffern erstellt, zu denen die Kodierung der Funktion / Bedeutung dann manuell vorgenommen wurde. Das Ergebnis wurde im Folgenden auf die Gesamttrefferzahl der jeweiligen Types hochgerechnet. Mit einer Ausnahme (*Mannesmann* enthielt einige ICC-Treffer mit Prototypenlesart) war das allerdings gar nicht nötig, da alle 1000 Treffer der Zufallsstichpro-

be dieselbe Bedeutungskodierung verlangten. So betraf die Form *nennen* in allen Fällen der Zufallsstichprobe das Verb *nennen* und wurde folglich als nichtnominaler Beleg aussortiert; *Kindeskinder* immer die flektierte Form des Determinativkompositums und *Barabara* immer die Fehlschreibung des Vornamens *Barbara*. Natürlich lässt sich nicht ausschließen, dass durch dieses Vorgehen einige Belege eine falsche Kodierung erhalten. Es ist aber unwahrscheinlich, dass die Zufallsstichproben so ungünstig erstellt worden sind, dass diese nicht erfassten Belege von Relevanz für das jeweilige Lexem sind.

Es ist notwendig, näher auf die Kriterien, die der semantisch-funktionalen Kodierung zugrundeliegen, und die damit zusammenhängenden Probleme einzugehen. Belege, die für den Untersuchungsgegenstand irrelevant sind, wurden als Fehler kodiert. Das betrifft Fälle, die nicht als repetitive oder reduplikative Formen im engeren Sinne gelten können, da sie bloße Wiederholungen eines Zeichens sind (*xxxxxxxxx*). Außerdem wurden solche Fälle als Fehler kodiert, die nur auf Graphemebene reduplikativ sind (*Kerker*), nur aufgrund der im Abfragestring enthaltenen Toleranzen erfasst wurden (*Phosphor*) oder aus zwei gänzlich unterschiedlichen Stämmen zusammengesetzt sind (*Reis-Reise*).

Über solche Fehler-Fehler hinaus wurden aber auch sinnvolle und von Sprecher:innen beabsichtigte Wiederholungen als Fehler kodiert, sofern sie für den Untersuchungsgegenstand irrelevant sind. Dies betrifft phonologische, morphologische und syntaktische Verbindungen. Die zwei zentralen Unterscheidungskriterien für die Einteilung sind:

1. Stellen die Bestandteile der Bildung sowie ihr Ergebnis Wortstämme dar?
2. Sind die Belege nominal?

Als phonologische Verbindung (PHON-Fehler) wurden Fälle kodiert, bei denen weder der Reduplikant noch die Basis ein Wortstamm ist, sondern allein die gesamte Wortform:

(36) *„Och, ist der süß“, rief ich und begann dem **Wauwau** zu streicheln.*

<http://www.sweetamoris.de/forum/t31229,1-schwarz-rot-cas-ff.htm>

(37) *Es gibt sie also doch noch; Filme, die ohne viel **Tamtam** Spannung erzeugen?*

<http://blog.sodomedia.de/category/filmblog/filmkritik-und-review/genre/thriller/page/2/>

(38) *wo sie die Idee hatten, vom Fahrrad aufs **Tuk-Tuk**, also die typische südostasiatische Autorikscha, umzusteigen.*

<http://www.kulturkreis-badlauterberg.de/index.php/presseartikel>

(39) ***Renren**, der Facebook-Konkurrent aus China, hat heute sein Debüt an der New Yorker Börse gegeben.*

<http://www.geld-trader.de/category/wirtschaft/page/5/>

Häufig sind die Basen dieser Lexeme Lautmalereien, die entweder originär im Deutschen (36) oder in einer Gebersprache (37–38) Geräusche imitieren. Da die nachgeahmten Geräusche selbst rhythmische Wiederholungen sind, imitiert auch die Repetition der Lautmalerei die Wiederholung dieser Geräusche. (37) ist etwa aus dem Französischen entlehnt und ahmt ursprünglich das Geräusch einer mit einem filzüberzogenen Klöppel angeschlagenen Bronzescheibe nach.[18] Die Wiederholung sprachlichen Materials geschieht also ursprünglich auf der Grundlage der Phonologie. Erst in einem weiteren Schritt wurden diese bereits repetierten Interjektionen in Nomina konvertiert. Einen anderen Fall markiert (39). Hier ist keine Lautmalerei beteiligt; stattdessen führt ein grammatischer Prozess in der Gebersprache Mandarin zur Wiederholung von *ren* ‘Mensch’. Für Sprecher:innen des Deutschen sind Formen wie in (39) aber mit solchen in (36–38) gleichzusetzen: Eine Segmentierung der Form in zwei Bestandteile ist allein auf der Grundlage des Lautkörpers (des Schriftbildes) möglich, nicht auf der Grundlage lexikalischen oder grammatischen Wissens. Für die Kodierung als phonologische Verbindung ist dabei unerheblich, ob die beiden Bestandteile des Wortes in der jeweiligen Gebersprache Wortstatus besitzen. Die Bildungen sind außerdem nicht das Ergebnis eines produktiven Wortbildungsprozesses. Die Bedeutung ergibt sich nicht aus einer morphologischen Konstruktion. Die Basen in (36–39) können darum auch nicht einfach ersetzt werden (**die Miaomiao*, **der Knackknack*). Für die meisten Sprecher:innen des Deutschen sind diese Wörter nicht analysierbar und werden daher wie Simplizia behandelt.

Als das Ergebnis eines syntaktischen Prozesses (SYN-Fehler) definiere ich hingegen die Wiederholung sprachlichen Materials, bei der sowohl der Reduplikant als auch die Basis im Deutschen Wortstatus besitzen, das Ergebnis des Prozesses aber keine Wortform darstellt, weil die Form über die Wortebene hinausgeht:

18 „Tamtam“, in: Pfeiffer (1995) , https://www.dwds.de/wb/Tamtam>, abgerufen am 10.10.2021.

(40) *Ich sagte ihr das ich sie trotzdem melden werde, denn das er tot **ist-ist** deine Schuld.*

<http://www.dashaustierforum.de/kaninchen-allgemeines/koenne-kaninchen-astma-bekommen-t7277.html>

(41) *Es ist ein monotoner Takt: **Bam-Bam***

<http://www.wlz-fz.de/Welt/Kultur/Uebersicht/(offset)/2760/>

(42) ***menschmensch**. Trannypartys zu veranstalten wird heutzutage immer schwieriger.*

<http://zoe-delay.de/tag/mia-de-mar>

(43) *Wenn man schon mit die spinnen die Briten anfängt, dann sollte man auch etwas dazu schreiben und nicht über die USA und den IWF und **bla-bla**.*

<http://bankingclub.de/news/Die-spinnen-die-Briten/?p=4>

Bei Wiederholungen wie in (40) verläuft zwischen den Bestandteilen eine syntaktische Grenze, sodass das Ergebnis der Dopplung größer als ein Wort ist. In diese Kategorie der syntaktischen Verbindungen fallen außerdem Interjektionen, die iteriert werden (41–42). In (42) wird *mensch* als Interjektion verwendet und zum Ausdruck der Emphase repetiert. Auf den ersten Blick sind solche Belege den phonologischen Verbindungen wie etwa in *Wauwau* in (36) sehr ähnlich. Der syntaktische Charakter der Verbindung wird aber zum einen dadurch deutlich, dass sie nicht flektiert werden kann. Die Interjektion *mensch* wird in *menschmensch* also nicht zu einem Nomen zurücktransponiert und bleibt, etwa im Gegensatz zu *Hundehund* in (1) oder *WauWau* in (36), Interjektion. Außerdem ist die Wiederholung bei Belegen wie *menschmensch* nicht notwenig auf eine Wiederholung beschränkt. Bei einer mehrfachen Iteration (*menschmenschmensch*) trüge die Wiederholung dieselbe Bedeutung wie bei der einfachen (*menschmensch*). Bei *Wauwau* kann der Bestandteil *wau* hingegen nicht mehr als zweimal hintereinander auftreten, ohne dass der Bildung eine andere Bedeutung zukäme.

Darüber hinaus gibt es nicht wenige Belege, die zwischen Syntax und Morphologie einzuordnen sind (MORPH-SYN-Fehler):

(44) *So läßt sich die Maschine schnell von **Luft-Luft** auf Luft-Boden-Lenkwaffen umrüsten.*

<http://www.bredow-web.de/ILA_2006/Military/MiG-29/mig-29.html>

(45) *Und wir sind täglich Beobachter der **Hund-Hund**-Beziehung, die uns regelmäßig in Verzückung versetzt*

<http://aspa-ev.de/31-0-Erziehungstipps.html>

(46) *Oder für einen exotischen Abend zu zweit. Denn ohne Couscous kein **Kusskuss***

<http://www.portal-mg.de/index.php/Arvelle.de/Bücher+Zeitungen+und+Magazine/Bücher+Sachbuch+Ratgeber+Essen+und+Trinken+Allgemeine+K/start20>

(47) *Dudh Chiya, hierzulande besser bekannt als Chai-Tee – was übersetzt soviel wie **Tee-Tee** bedeutet.*

<http://kaffee-blog.maskal.de/tee/neu-im-kaffee-online-shop-nepal-tee>

Stefanowitsch (2007) beschreibt Fälle wie in (44) und (45) als asyndetische Dopplungen. Meist werden sie den kopulativen Komposita zugeordnet (Bisetto & Scalise 2005, Lieber 2009b). Es wird ausgedrückt, dass der Referent der Basis zweimal vorliegt, häufig im Rahmen eines übergeordneten Kompositums mit Nomina, die ein Verhältnis zwischen zwei Entitäten ausdrücken, etwa *Beziehung* oder *Interaktion*. Der Kopf des Kompositums kann wie in (44) weggelassen werden. In einigen Fällen, wie beispielsweise in (46), liegt hingegen kein Kompositum vor. *kusskuss* ist nicht Teil eines Kompositums, in das es eingebettet wäre. Man könnte hier von einem reduplikativ ausgedrückten, nominalen Dual sprechen[19] oder von echten, also wirklich kompositionalen Dvandva (siehe Kapitel 9.2.1). Die Möglichkeit, dass mit ICCs auch Dual oder andere periphere reduplikative Lesarten, die nicht unter die Kategorie Prot oder Real fallen, ausgedrückt werden könnten, sieht schon Finkbeiner (2014: 194). Ob man angesichts des Nischendaseins dieses Phänomens aber tatsächlich diese Annahme trifft, also etwa eine nominale Kategorie „Dual“ annehmen sollte, kann an dieser Stelle nicht entschieden werden.

19 Im Gegensatz zur häufig kontraikonischen Markierung des Duals (etwa im Sorbischen: *nanaj* ‘zwei Väter’ – *nanojo* ‘Väter’ (Stolz 1988: 481)) wäre ein nominaler Dual, der durch Reduplikation ausgedrückt wird, geradezu ein Paradebeispiel für Ikonizität.

Beispiel (47) exemplifiziert, wie formal identische Nomina in einem ganz bestimmten Kontext adjazent auftreten, wenn nämlich Sprecher:innen darüber diskutieren, dass in Bildungen wie *Chai-Tee* Lehnwörter mit ihren deutschen Übersetzungen komponiert werden. Häufig wird dabei die Bedeutungswiederholung kenntlich gemacht, indem der entlehnte Teil des Kompositums durch ein schon länger im deutschen Wortschatz befindliches Wort ersetzt wird. Das Ergebnis dieser Sprachkritik könnte man wie *kusskuss* als asydetische Dopplung verstehen ('Tee und Tee' = das Erstglied bedeutet Tee und das Zweitglied bedeutet auch Tee) oder als morphologische Bildung ('Tee der Sorte Tee'). Da dies nicht immer eindeutig entschieden werden kann, ordne ich auch solche Übersetzungsbildungen zwischen Syntax und Morphologie ein.

Als zielsprachliche, morphologische Verbindung definiere ich die Wiederholung sprachlichen Materials, bei der (wie bei den syntaktischen Verbindungen) sowohl die Basis als auch der Reduplikant im Deutschen Wortstatus besitzen. Das Ergebnis der Zusammensetzung ist hierbei genau auf der Wortebene anzusiedeln. Im Gegensatz zu den syntaktischen Verbindungen resultiert das Ergebnis des Prozesses also in genau einem Wort, einem Stamm mit lexikalischer Integrität. In vielen Fällen erkennt man Letztere daran, dass die Bildung auch als Einheit flektiert. Die morphologischen Verbindungen sind aber nicht immer nominale ICCs und entsprechen damit oft nicht dem Untersuchungsgegenstand. Bildungen, bei denen das Ergebnis nicht nominal, sondern adjektivisch ist (*neu-neu* 'ganz neu'), werden als adjektivisch (POS-FehlerA) kodiert; liegt eine verbale Basis vor (etwa *schlafenschlafen* 'wirklich schlafen') als verbal (POS-FehlerV). Dazu gibt es Fälle, bei denen ein unflektierter Verbstamm (Inflexiv) wiederholt wird, um Durativität auszudrücken (Freywald 2015), etwa *freufreu* (POS-FehlerV (SR)).

Die restlichen morphologischen Verbindungen sind ICCs. Eine genaue Beschreibung und Subklassifikation dieser Bildungen wird im folgenden Kapitel 4 im Rahmen der Besprechung der semantisch-funktionalen Aspekte vorgenommen. Tabelle 2 fasst alle in diesem Kapitel beschriebenen Kriterien für die Kodierung zusammen.

Tab. 2: Kriterien zur Kodierung der Wiederholungsfunktion

Label	Kriterien für die Kodierung	Beispiele
Fehler	Nicht von Menschen produzierter Text Nur im Schriftbild, nicht im Lautbild reduplikativ Nur durch Toleranzen im Abfragestring erfasst	*xxxxxxxxx* *Kerker* *Phosphor* *Reis-Reise*
PHON-Fehler	Basis < Wort Ergebnis = Wort	*Barbar* *Tamtam*
SYN-Fehler	Basis = Wort Ergebnis > Wort	*hurrahurra* *ein alter alter Mann* *er sagte blabla,* *piep-piep machte es*
MORPH-SYN-Fehler	Basis = Wort Ergebnis < Wort Repetition syntaktisch gebildet, aber in Kompositum eingebettet (Asyndetische Dopplung)	*winwin* *die Beziehung Mann-Mann* *der Abstand Achse-Achse* *Horst-Horst-Sprechchöre*
POS-FehlerA	Basis = Wort Ergebnis = Wort POS = A / Adv.	*neu-neu* 'ganz neu' *zuhausezuhause* 'im Elternhaus'
POS-FehlerV	Basis = Wort Ergebnis = Wort POS = V	*schlafen-schlafen* 'wirklich schlafen'
POS-FehlerV (SR)	Basis = Wort Ergebnis = Wort POS = V Durativität / Iterativität wird ausgedrückt	**freufreu** **schmatzischmatzi** *kraule-kraule*
ICC	Basis = Wort Ergebnis = Wort POS = N	*Zinseszins* 'Zins auf den Zins' *Hundehund* 'Hund, der auf andere Hunde bezogen ist'

4 Ergebnisse zu semantisch-funktionalen Aspekten von ICCs

4.1 Überblick über die Daten

Die Abfrage im DECOW16 lieferte basierend auf den 1034 Items über 40.000 Belege von adjazent identischen Nomina. Nach der kriteriengestützten Datenfilterung verbleiben 26.166 Belege in den Daten. Den weitaus größten Anteil an den DECOW16-Daten haben Belege syntaktischer Verbindungen. Sie machen mehr als Dreiviertel des Gesamtergebnisses aus. Dazu gehören adjazente Nomina, zwischen denen eine syntaktische Grenze verläuft, beispielsweise wenn in Sätzen mit Verbletztposition formal identische Verbkomplemente unmittelbar nebeneinander stehen (…*wenn Kinder Kinder bekommen*). Außerdem werden in Anlehnung an Gil (2005) Repetitionen und Iterationen den syntaktischen Verbindungen zugerechnet. Weitere 9% lassen sich nicht eindeutig der Syntax oder Morphologie zuordnen. Die restlichen 15% der Daten sind ICCs.

Die Abfrage im deTenTen13 lieferte insgesamt 1.830.356 Datenpunkte zu 29.186 Types. Den größten Anteil der Rohdaten machen die Fehler aus. Aus diesem Grund ist die oben genannte Anzahl von knapp 1,8 Millionen Belegen zunächst irreführend, da 1.003.061 Belege beispielsweise allein auf die Verbform *nennen* zurückgehen. Auch andere Falsch-Positiv-Fehler wie die Form *endenden* des deverbalen Adjektivs (10116 Belege) oder die Bezeichnung *entente* ‘Bündnis zwischen dem Vereinigten Königreich, Frankreich und Russland im ersten Weltkrieg’ (5800 Belege) sind für viele Datenpunkte verantwortlich. Insgesamt machen falschpositive Fehler 1.204.853 Belege und somit knapp zwei Drittel der Rohdaten aus. Diese Belege wurden allesamt als Fehler kodiert und werden bei der weiteren Auswertung nicht berücksichtigt. Knapp 10% der deTenTen13-Rohdaten sind ICCs und werden im Folgenden weiter untersucht.

Tabelle 3 gibt die Verteilung der nicht als Fehler aussortierten Belege in den beiden Korpora sowie veranschaulichende Belege wieder.[20]

20 Weitere 8,3% der deTenTen13-Daten gehen auf POS-Fehler zurück (N=51.708).

https://doi.org/10.1515/9783111315416-004

Tab. 3: Gesamtverteilung der Belege (DECOW16A und deTenTen13)

Art der Verbindung	Grund der Dopplung	Beispiele	Anteil am Gesamtergebnis (DECOW16)	Anteil am Gesamtergebnis (deTenTen13)
syntaktisch	Adjazente Verbkomplemente	*Tja…wenn Kinder Kinder bekommen!*	76,0 % (N=19.894)	16 % (N=100.073)
	Repetition	*wir haben Hunger, Hunger*		
	Iterierte Interjektion	*menschmensch, ihr seid ja alle fleißig*		
syntaktisch / morphologisch	Asyndetische Dopplung, Übersetzung	*die Beziehung Mann-Mann*	8,9 % (N=2335)	3,7 % (N=23.173)
		Chai-Tee heißt also eigentlich Tee-Tee		
morphologisch	ICCs	*er ist ein Menschen-Mensch, einer, der offen auf Leute zugeht*	15 % (N=3.937)	28,9 % (N=180.700)
phonologisch	Stammlose Wortbildung	*vom Fahrrad aufs Tuk-Tuk umsteigen*	0 % (N=0)	43,1 % (N=269.849)

In der Gegenüberstellung der Auszählung beider Korpora zeigt sich die unterschiedliche Methodik, mit der die Daten erhoben wurden. Während das vordefinierte Itemset bei den DECOW16-Daten dazu führt, dass keine phonologischen Dopplungen enthalten sind, reduziert der Ausschluss der Spatienschreibung die syntaktischen Verbindungen in den deTenTen13-Daten erheblich.

Die Daten sollen nun weiter analysiert und die ICC-Belege subklassifiziert werden. ICCs sind formal sehr vielfältig. Manche Belege bestehen aus bloßen Stämmen, manche zeigen morphologische Marker an Erst- und / oder Zweitglied. Sie werden mal wie kanonische N+N-Komposita zusammengeschrieben, mal sind die Konstituenten graphisch voneinander getrennt. Teils bestehen die Belege aus Simplizia, teils aus morphologisch komplexen Stämmen. Auch auf funktionaler Ebene bilden ICCs keine homogene Klasse. Manche ICCs sind endozentrisch, manche exozentrisch. Mal zeigen sie ein Determinationsverhältnis und eine semantische Relation zwischen den Konstituenten, mal nicht. Teils bilden sie eine semantische Subklasse zu ihrem Basislexem, teils ist die ICC-Bedeutung völlig anders als die der Basis. In der bisherigen Forschung zu ICCs wurden funktional sehr unterschiedliche Bildungen beschrieben und diese finden sich auch in den erhobenen Korpusdaten wieder. Es kommen aber auch ICCs vor, die Funktionen ausüben, die in der Literatur bisher überhaupt nicht behandelt werden.

ICCs sind also formal und funktional sehr vielfältig. Die Beschreibung und Kodierung der Daten muss diese Vielfalt erfassen und systematisch Gruppen bilden, um Erkenntnisse zum Phänomen ICC zu ermöglichen. Eine Herausforderung ist hier, dass augenscheinlich keine klare Passung formaler und funktionaler Merkmale besteht. Es ist beispielsweise nicht möglich, die von Finkbeiner (2014) und Freywald (2015) als Distinktionsmerkmal beschriebene Kompositionsfuge als formales Unterscheidungskriterium zwischen determinativen ICCs und solchen mit Prot- oder Real-Lesart heranzuziehen, weil nicht alle verfugten ICCs determinativ und nicht alle determinativen ICCs verfugt sind. Zudem hat der Forschungsüberblick gezeigt, dass es in der Literatur zu ICCs weder hinsichtlich formaler Merkmale noch hinsichtlich funktionaler Merkmale einen Konsens gibt. Einerseits werden ICCs als Nicht-Wörter angesehen (Hohenhaus 2015); andererseits gibt es Evidenz dafür, dass ICCs in den Köpfen der Sprecher:innen Subkonzepte evozieren, die kontextfrei funktionieren (Finkbeiner 2014). Zu den formalen Merkmalen von ICCs gibt es bisher überhaupt keine belastbare empirische Evidenz. In Bezug auf die Verfugung und das Flexionsverhalten vertreten die Autor:innen der bisherigen theoretischen Arbeiten sogar entgegengesetzte Auffassungen. Für die weitere Beschreibung und Kodierung der als ICCs identifizierten Daten ist auch diese Forschungslücke eine Herausforderung.

Es gibt mehrere Möglichkeiten, diesen Herausforderungen zu begegnen. Man kann eindeutige formale Merkmale von ICCs definieren und untersuchen, ob sie mit unterschiedlichen Funktionen der Bildungen korrelieren. Man kann die ICCs aber auch anhand semantischer, funktionaler und referenzieller Eigenschaften klassifizieren und daraufhin ermitteln, ob die Sprecher:innen diese Eigenschaften mithilfe formaler Merkmale markieren. Letzteres Vorgehen wird in dieser Arbeit gewählt. Nichtsdestotrotz wird an den Stellen, an denen interessante Unterschiede deutlich werden, auch die andere Perspektive eingenommen und etwa danach gefragt, welchen Anteil determinative Bildungen an den verfugten ICCs haben. Es führt in den meisten Fällen aber zu einer anschaulicheren Datenbeschreibung, die Belege danach zu klassifizieren, welche Funktion die ICC-Bildung ausübt.

4.2 Kriterien der Subklassifikation

Auf formaler Ebene bestehen ICCs aus zwei identischen Einheiten. Die Sichtung der Korpusdaten zeigt nun, dass sich ICCs darin unterscheiden, ob den zwei formalen Einheiten, die denselben Stamm realisieren, auch zwei konzeptuelle Einheiten gegenüberstehen, die dasselbe Konzept realisieren. Die Subklassifikation teilt die ICCs somit nach ihrem Transparenz- und Kompositionalitätsgrad in Gruppen ein. Eine ausführliche Beschreibung der Begriffe Transparenz und Kom-

positionalität nehme ich in Kapitel 9.2 vor. Um die Subklassifikation der ICCs nachvollziehen zu können, genügen aber zunächst die in diesem Kapitel besprochenen Merkmale der ICCs. Maßgeblich für die Subklassifikation ist das Verhältnis von Erst- und Zweitglied zum Basislexem und somit, da die Konstituenten identisch sind, zugleich das Verhältnis von Erst- und Zweitglied zueinander. Die Subklassifikation beruht also auf der Frage:

➡ Realisieren beide Konstituenten gleichermaßen das Konzept des Basislexems?

Die Antwort auf diese Frage, und somit auch die Entsprechung zwischen Form- und Bedeutungsseite, hängt zum einen mit der Klassifikationsfunktion zusammen, die die Defaultfunktion nominaler Komposition ist (Bauer 2006a, Schlücker 2013: 469), zum anderen mit der semantischen Köpfigkeit. Die Subklassifikation ist der Normalfall der N+N-Komposition und auch ICCs haben mitunter einen klassifikatorischen Modifikator, der den begrifflichen Kern des ICCs beschränkt, sodass ein Unterbegriff erzeugt wird (Type Restriction). Diese Subkonzeptbildung geschieht in der Regel dadurch, dass zwei nominale Konzepte zueinander in Beziehung gesetzt, beziehungsweise miteinander verbunden werden (Olsen 2015: 365, 2019: 103). Die Bildungen folgen dann der Formel ‘X, das mit Y zu tun hat’ (Zifonun 2010: 128). In Form der beiden unterschiedlichen Variablen ist in der Formel bereits die Annahme enthalten, dass die beiden Konzepte nicht auf dasselbe verweisen und also zwei unterschiedliche konzeptuelle Einheiten vorliegen. Kanonische N+N-Komposita entsprechen dieser Annahme. In *Schneeball* liegt zum einen das Konzept SCHNEE, zum anderen das Konzept BALL vor. In ICCs werden Erst- und Zweitglied aber von ein- und demselben Stamm realisiert, sodass nicht zwei vollkommen distinkte Konzepte zueinander in Beziehung gesetzt werden. Stattdessen kann ein- und derselbe Stamm jeweils zwei nominale Teilkonzepte realisieren, also Konzepte, die gemeinsam das Konzept des Kompositums ausmachen. In ICCs kann also ein- und dasselbe Konzept, das der Basis, zweimal vorliegen. Bei solchen ICCs fungiert das Erstglied als klassifikatorischer Modifikator. Durch das Vorliegen zweier nominaler Teilkonzepte ist die Bedeutung des ICCs genauso aufgebaut wie die Form: Es liegen zwei (nahezu) identische Einheiten vor. Beide Konstituenten realisieren gleichermaßen das Konzept des Basislexems.

In jedem ICC, das sich nach der Filterung noch in den Daten befindet, werden Erst- und Zweitglied vom gleichen Stamm realisiert. Die deskriptive Bedeutung der Nomina, die Erst- und Zweitglied zugrundeliegen, ist also gleich; sie haben außerhalb des ICCs identische Denotate. Als Teil des ICCs aber können sie durchaus unterschiedliche Inhalte haben. In *Hundehund* sind die Teilkonzepte mit dem

Konzept des Basislexems identisch, verweist das Konzept HUND einmal auf eine konkrete Entität (Hund, der Artgenossen bevorzugt), und einmal generisch auf die Klasse der Hunde. Es liegt ein- und dasselbe Konzept zweimal vor. In manchen ICCs verweist aber nur eine der Konstituenten auf ein Konzept, das dem des Basislexems entspricht. Bei *Beziehung-Beziehung* aus Beispiel (2) evoziert das Zweitglied das Konzept des Basislexems; das Erstglied aber trägt keine lexikalische Bedeutung zum Kompositum bei (Finkbeiner 2014: 188). In *fotofoto* in (3) wiederum verweist das Zweitglied nicht auf das Konzept der Basis FOTO, sondern auf einen Onlineshop und somit nicht auf etwas, das Teil des Denotats des Zweitgliedes ist. Es liegt kein semantischer Kopf vor. Die Bedeutung solcher ICCs ist anders aufgebaut als die Form: Es liegt nur einmal das nominale Konzept der Basis vor.

Vier Beispiele sollen die genannten Unterschiede zwischen ICCs, nach denen die ICC-Subklassifikation vorgenommen wurde, verdeutlichen.[21] Die sprachlichen Elemente im Kotext, mit denen das Kopfkonzept des ICCs identifiziert werden kann, sind jeweils doppelt unterstrichen, die Elemente, die das Konzept des Erstgliedes anzeigen, sind einfach unterstrichen. Im ersten Beispiel gibt der Kontext Aufschluss darüber, dass sowohl Erstglied als auch Zweitglied das Konzept der Basis evozieren. Das ICC hat also einen das nominale Konzept der Basis realisierenden Modifikator und einen semantischen Kopf:

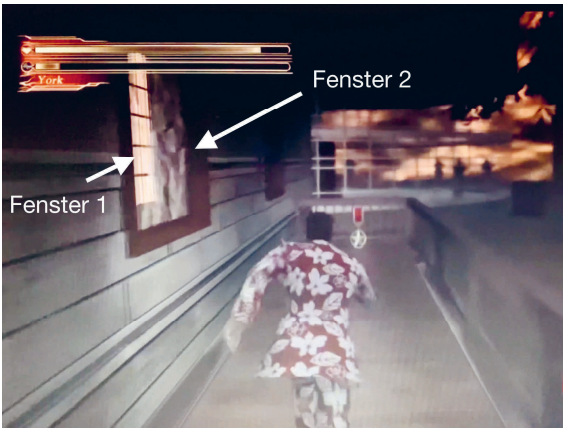

Abb. 5: Das ICC *Fenster-Fenster* im Kontext.

21 Abbildung 5 und 6 sowie die Beispiele (48), (53) und (54) dienen der Veranschaulichung und sind nicht Teil der Korpusdaten. Alle weiteren Beispiele sind hingegen den erhobenen Daten entnommen.

(48) *Ich möchte nicht, dass meine Fenster in irgendeiner Art und Weise dreckig werden also machen Sie bitte ein Glas vors Fenster. Also ein* ***Fenster-Fenster****?*

<Dem Spil zum Trots! - Deadly Premonition #37 - Time to Drei, 9:23, https://www.youtube.com/watch?v=r8jba5nUMgw&list=PLgtM00s3pkmVBaOA-p7SIN7N-E4wlZ29jW&index=37>

Sowohl das Erstglied als auch das Zweitglied realisieren das nominale Konzept der Basis. Das ICC wird vom Sprecher als 'Glas vor einem Fenster' paraphrasiert. Der Modifikator in *Fenster-Fenster* verweist generisch auf die Klasse der Fenster. Im Äußerungskontext wird das Erstgliedkonzept konkretisiert. *Fenster* (*dass meine Fenster*) nimmt auf das Objekt Bezug, vor dem *Glas* angebracht wird (in Abbildung 5 mit Fenster 1 beschriftet). Als generisch verweisender Modifikator ist diese konkrete Realisierung im Äußerungskontext zwar nicht notwendig und lässt sich damit erklären, dass nominale Konzepte typischerweise in Ausdrücken mit referentieller Funktion verwendet werden (Hopper & Thompson 1984, 1985). Im Beispiel (48) verdeutlicht diese Konkretisierung aber, dass das Modifikatorkonzept tatsächlich ein nominales Konzept ist, denn es ist möglich, dass das Konzept im Äußerungskontext (also in gewisser Weise ontologisch) vorliegt.

Auch das Zweitglied evoziert das Konzept der Basis. *Glas* nimmt kataphorisch Bezug auf den Kopf des ICCs. Da *Glas* und *Fenster* koreferent sein können und es in diesem Beispiel auch sind, zeigt diese Bezugnahme, dass das ICC einen semantischen Kopf in Zweitgliedposition hat, der auf ein Fenster referiert (Fenster 2 in Abbildung 5). Zudem zeigt *Fenster-Fenster* die LOC-Relation, die im Kotext durch eine Präposition realisiert wird (*ein Glas* ***vors*** *Fenster*). Ein *Fenster-Fenster* ist eine Art Fenster, nämlich eines, das sich örtlich an einem Fenster befindet, so wie ein *Dachfenster* ein Fenster ist, das sich an einem Dach befindet, oder ein *Autofenster* eines, das sich an einem Auto befindet. Zwei nominale Konzepte sind über die LOC-Relation verbunden. Um ein Subkonzept zum Konzept des Basislexems zu bilden, liegen zwei nominale Konzepte vor und werden zueinander ins Verhältnis gesetzt.

Das zweite Beispiel veranschaulicht, dass in ICCs mitunter nicht beide Konstituenten gleichermaßen das Konzept des Basislexems realisieren:

(49) *Die Autoren wie auch Mutter Beimer wollten einen kräftigen Mann haben, so einen '****MannMann****'. Er sollte einen Bart haben, musste Humor haben und souverän sein.*

<http://www.alzd.de/category/duft-parfum/gruenes/page/5>

Bei *MannMann* zeigt sich an der kataphorischen Bezugnahme (*einen kräftigen Mann*), auf welche Entität sich *MannMann* bezieht und dass es sich beim Referenten des ICCs um eine bestimmte Art Mann handelt. Auch hier wird also ein Subkonzept gebildet und auch hier liegt im Zweitglied das Konzept der Basis vor. Im Unterschied zu *Fenster-Fenster* geschieht die Subkonzeptbildung aber nicht durch das In-Beziehung-Setzen zweier nominaler Konzepte. Stattdessen entspricht dem Erstglied nicht das nominale Konzept der Basis; in *MannMann* wird somit nicht zweimal das nominale Konzept MANN evoziert. Eine weitere Instanz eines Mannes, die dem Modifikator entsprechen würde, ist nicht realisierbar und liegt in diesem Beispiel auch im Kotext nicht vor. Es ist neben dem Mann, den das ICC bezeichnet, kein weiterer Mann gegeben. In der Formel 'X, das mit Y zu tun hat' kann *Mann* nur für X eingesetzt werden.

Im dritten Beispiel liegt überhaupt kein Konzept vor:

(50) *ich bin auf keinen fall ausländerfeindlich, aber es kann eifnach nicht sein, dass man, wie* ***sabinesabine*** *schon sagte, nach Deutschland kommt und denkt man kann hier leben ohne sich anzupassen.*

<http://www.www.webnews.de/17184/al-qaeda-video-stammte-erfurt>

Bei *sabinesabine* in (50) ist das Zweitglied kein appellativer Nominalstamm, sondern ein Eigenname und beinhaltet daher kein Konzept, zu dem das ICC ein Subkonzept bilden könnte. *Sabinesabine* bezeichnet keine bestimmte Art Sabine, sondern einfach eine Person namens *Sabine*.

Das vierte Beispiel zeigt, dass in ICCs auch zwei nominale Konzepte vorliegen können, ohne dass diese Teilkonzepte äquivalent zum Konzept des Basislexems sind:

Künstler zertrümmern Auto im Pavillon Hannover

Anders als bei vielen anderen Kulturveranstaltungen ist hier der Eindruck, dass an diesem Abend Schrott produziert wird, durchaus erwünscht. In „AutoAuto!" wird ein Auto zertrümmert. Zwei Herren bearbeiten es mit Fäusten, Schleifmaschinen, Metallstäben, Vorschlaghämmern.

Von Ronald Meyer-Arlt

Abb. 6: Das ICC *AutoAuto* im Kontext (https://www.haz.de/Nachrichten/Kultur/Uebersicht/Kuenstler-zertruemmern-Auto-im-Pavillon-Hannover).

Auf das Basislexem *Auto*, also das potentielle Hyperonym des ICCs, wird im Kotext des Belegs Bezug genommen (*es*). Doch wird im Kotext deutlich, dass das ICC auf etwas referiert, das nicht Auto, sondern Kulturveranstaltung ist. Das Zweitglied ist somit nicht semantischer Kopf des Kompositums, es realisiert nicht das Konzept der Basis. Das Erstglied allerdings realisiert das Konzept der Basis, denn es verweist auf ein Auto, das in die Show, auf die das ICC referiert, involviert ist. In der Formel 'X, das mit Y zu tun hat' kann *Auto* für Y eingesetzt werden. *AutoAuto* referiert auf etwas, das mit einem Auto zu tun hat. Das ICC kann also als 'Show mit einem Auto' beschrieben werden. Den zwei identischen Bestandteilen auf der Formebene stehen somit zwar zwei nominale Konzepte gegenüber (es gibt eine Show und es gibt ein Auto), doch ist nur eines davon mit dem Konzept des Basislexems deckungsgleich. Das Erstglied verweist generisch auf Autos, die Teil der Show sind. Die Show selbst hat aber keine formale Entsprechung und das ICC *AutoAuto* kann somit als exozentrisches Kompositum beschrieben werden.

Die bisher genannten ICCs *Fenster-Fenster*, *MannMann*, *sabinesabine* und *Auto-Auto* unterscheiden sich also hinsichtlich der Entsprechung zwischen formaler und konzeptueller Ebene. Bei der nun folgenden Subklassifikation der ICCs wird zusätzlich berücksichtigt, welche der zwei Konstituenten das Konzept der Basis realisiert. Die Subklassifikation der ICCs fußt somit auf den folgenden zwei Kriterien:

I. Das Erstglied evoziert das nominale Konzept des Basislexems.
II. Das Zweitlied evoziert das nominale Konzept des Basislexems.

Die Subklassifikation hinsichtlich der genannten Kriterien ermöglicht es, die unterschiedlichen Funktionen der Konstituenten von ICCs sowie der ICCs selbst zu unterscheiden. Im Folgenden werden die ICC-Subtypen und ihre Funktionen näher beschrieben.

4.3 ICC-Subtypen

Sind die Kriterien I. und II. erfüllt, handelt es sich bei den Bildungen um ICCs, in denen die zwei Konstituenten auf zwei nominale Konzepte verweisen und diese beiden Konzepte dem Konzept des Basislexems entsprechen:

(51) *Es geht aber noch viel schlimmer; ich "kenne" einen **Lehrer-Lehrer**, also Jemanden, der Grundschullehrer ausbildet.*

<http://www.rechtsanwalt-news.de/kanzleialltag/schwierige-oder-einfache-mandanten-nach-berufsgruppen/comment-page-1>

(52) *Auch Redakteure von Publikumszeitschriften sind an Verkaufszahlen gebunden. Darüber hinaus folgen sie ihren "**Erwartungserwartungen**", ihren Vorstellungen dessen, was das Publikum lesen will.*

<http://www.tucottbus.de/theoriederarchitektur/wolke/deu/Themen/061+062/Petrow/petrow.htm>

Diese ICCs sind Determinativkomposita, denn sie bilden Subkonzepte, indem zwei Begriffe zueinander in Beziehung gesetzt werden. Das Erstglied ist dabei das Determinans, das Zweitglied das Determinatum. Der erste ICC-Subtyp wird darum Det-ICC genannt.

In (51) zeigt der Kotext, dass das Erstglied ein klassifikatorischer Modifikator ist, der generisch auf die Klasse der Lehrer verweist (*Grundschullehrer*). Das Erstglied evoziert also ein nominales Konzept, dass zudem dem Konzept des Basislexems entspricht. Gleichzeitig hat das ICC einen semantischen Kopf, da *Lehrer-Lehrer* Hyponym zu *Lehrer* ist. Das Zweitglied referiert auf eine Entität (im Kontext durch *Jemanden* repräsentiert), auf den auch das Basislexem referieren kann. Auch hier ist das Konzept der Konstituente also mit dem der Basis identisch. Analog zu *Lehrer-Lehrer* ist *Erwartungserwartungen* in (52) gebildet. Die Bildung hat einen semantischen Kopf in Zweitgliedposition. Im Kotext wird das durch *Vorstel-*

lungen deutlich. *Erwartungserwartungen* hat außerdem einen klassifikatorischen Modifikator, der ebenfalls auf dem Konzept der Basis beruht. In diesem Fall wird dies durch die Phrase *was das Publikum lesen will* deutlich. Generell ist diese kontextuelle Entsprechung zum Erstglied allerdings nicht obligatorisch, da der Modifikator nur generisch verweist.

Beide ICCs zeigen außerdem semantische Relationen, die aus dem Zueinander-ins-Verhältnis-Setzen zweier nominaler Konzepte resultiert. Bei *Lehrer-Lehrer* macht der Kontext durch die Semantik des Verbs *ausbilden* deutlich, wie die beiden Konzepte ins Verhältnis gesetzt werden, es sich bei einem *Lehrer-Lehrer* nämlich um einen Lehrer handelt, der auf andere Lehrer einwirkt. Das Verhältnis entspricht der semantischen Relation FOR. Zwischen den Konstituenten von *Erwartungserwartungen* besteht eine ABOUT-Relation. In diesem Fall kann man annehmen, dass das ICC seine Modifikationsrelation aus der verbalen Basis *erwarten* erhält, aus dem die Basis *Erwartung* konvertiert ist. Det-ICCs können alle möglichen Modifikationsrelationen haben. Bei *Kindeskind* liegt eine OF-Relation vor, bei *Städtestadt* (‘Stadt aus Städten’ = Ruhrgebiet) eine MAKE-Relation (Levi 1978), beziehungsweise die Relationen MADE oder COMP (Jackendoff 2016: 28f.). Auf der Basis der Modifikationsrelationen könnte man Det-ICCs also wiederum subklassifizieren, was in dieser Arbeit aber aus Zeit- und Platzgründen unterbleibt. Abbildung 7 visualisiert die semantischen, modifikatorischen und referenziellen Eigenschaften der Det-ICCs sowie ihrer Konstituenten in Anlehnung an das semiotische Dreieck (Ogden & Richards 1923):

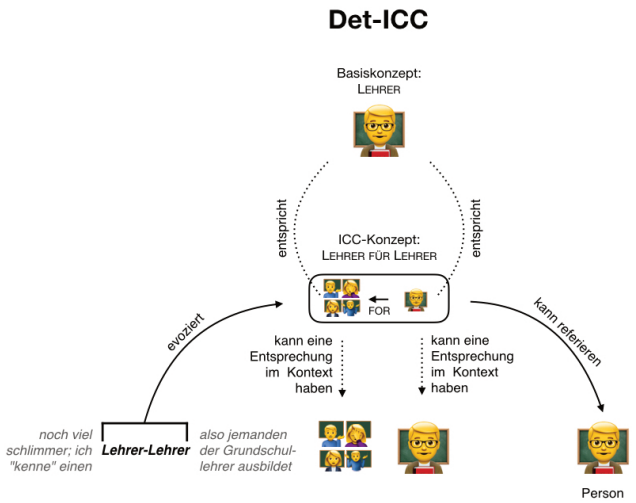

Abb. 7: Semantik und Referenz des Det-ICCs *Lehrer-Lehrer* und seiner Konstituenten.

Wird in einem ICC bloß Kriterium II. erfüllt, haben die Bildungen zwar einen semantischen Kopf, der dasselbe nominale Konzept wie die Basis evoziert, aber keinen Modifikator, der dem Konzept der Basis entspricht:

(53) *Daft Punk, Random Access Memories (2013). Dieses Album ist in seiner Gänze vielleicht eben das letzte als **Album-Album** angelegte Großwerk der zeitgenössischen Tanzmusik geworden.*

<https://www.zeit.de/kultur/musik/2021-02/daft-punk-trennung-dance-duo-frankreich-electro-pop-alter>

(54) *Und am nächsten Tag sind wir dann zur Oma meiner Freundin gefahren, die nochmal 200, 300 km weit weg war und da war dann das was Du auch sagst, nämlich Essen und kurze Pause und Essen und kurze Pause und Essen und kurze Pause. Und ich habe meine Familie da schon so drauf vorbereitet und meinte: das ist so ne richtige **Oma-Oma**, die auch immer noch nachfragt und immer nochmal nochwas aufn Teller.*

<Die Ratsherren #27, 15:03, http://www.superkreuzburg.de/podcasts/dieratsherren>

Album-Album bezeichnet ein *Großwerk*, ein richtiges, prototypisches Album; eine *Oma-Oma* eine richtige, prototypische Oma. Ebenso benennt das zuvor besprochene *MannMann* (49) eine Art von Mann, nämlich einen, der richtig, prototypisch ist. Solche ICCs haben also die von Finkbeiner beschriebene Prot- beziehungsweise Real-Lesart, weshalb ich diese ICCs Prot-ICCs nenne. Die Bildungen können mit Hohenhaus Formel „an XX is a proper/prototypical X" paraphrasiert werden.

Wie bei Det-ICCs liegt auch hier eine Subklassifikation vor, da es sich bei einem *Album-Album* um eine bestimmte Art Album handelt und bei *Oma-Oma* um eine bestimmte Art Oma. Dem ICC kommt die Funktion der begrifflichen Subklassifikation zu. Wie schon bei den Det-ICCs wird auch bei *Album-Album* das Konzept, das das Gesamtkompositum evoziert, im Kontext des Belegs zusätzlich erklärt. Das ICC referiert auf eine Entität in der Welt, nämlich auf einen Tonträger, beziehungsweise auf die Musik, die sich darauf befindet. In den gegebenen Beispielen verweisen die Zweitglieder zudem auf jeweils im Kontext etablierte Konzepte und Entitäten (*Album*, *Oma*) und sind die Köpfe der jeweiligen Komposita. Das Erstglied allerdings führt lediglich zu einer Konstruktion, die dann dem semantischen Kern des Kompositums prototypische Eigenschaften zuweist. In einem Prot-ICC evoziert das Erstglied somit, anders als bei einem Det-ICCs wie *Lehrer-Lehrer*, nicht das nominale Konzept des Basislexems. *Album-Album* ist kein 'Album für ein Album' im Sinne einer FOR-Relation oder ein 'Album aus (anderen)

Alben' im Sinne einer MAKE-Relation, weil es keine durch das Erstglied repräsentierte konzeptuelle Einheit gibt, auf die die konzeptuelle Einheit des Zweitgliedes bezogen sein könnte. Auch bei *Oma-Oma* gibt es kein zweites Konzept einer Oma, zu dem das Konzept der Kopfkonstituente in Beziehung gesetzt würde. Die Erstglieder von *Album-Album* und *Oma-Oma* können deshalb auch keine Entsprechung im Kotext haben, etwa in Form von mit dem Basislexem koreferenten Nominalphrasen, und auch keine ontologische Entsprechung im Äußerungskontext.

Es lassen sich zudem bei Prot-ICCs keine semantischen Relationen anwenden, weil nicht beide Konstituenten auf nominale Konzepte Bezug nehmen. Bei *Album-Album* und *Oma-Oma* sind darum semantische Relationen wie FOR oder OF nicht anwendbar (zur Einordnung der Prot-ICC-Erstglieder und alternativen Analysen hierzu, siehe Kapitel 4.2.3). Die semantischen und referenziellen Eigenschaften dieser ICCs sowie ihrer Konstituenten sind in Abbildung 8 dargestellt.

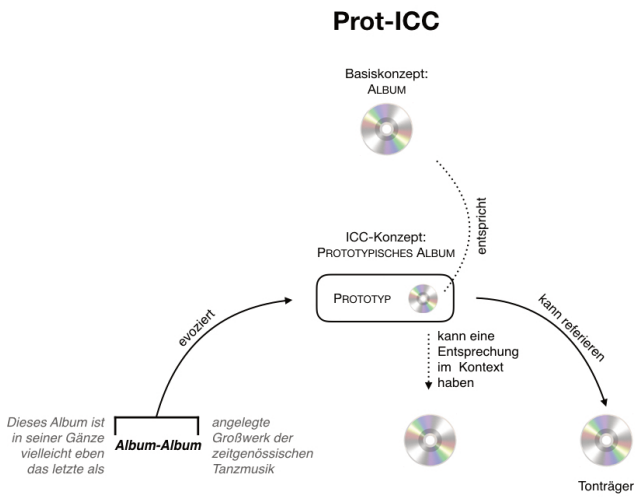

Abb. 8: Semantik und Referenz des Prot-ICCs *Album-Album* und seiner Konstituenten.

Erfüllen ICCs nur Kriterium I., nicht aber Kriterium II., haben die Bildungen einen klassifikatorischen Modifikator aber keinen semantischen Kopf. Nur das Erstglied, nicht aber das Zweitglied entspricht dem Konzept des Basislexems:

(55) *Poster XXL bietet auch Fotobücher an die von Seiten der Qualität nichts zu beanstanden haben. Leider hatten sie bei mir das letzte Mal massive Lieferschwierigkeiten, was bei einem persönlich nicht so tragisch ist aber für einen Kunden würde ich dort nicht bestellen. Durch meinen Facebookaufruf*

habe ich noch zwei weitere Hersteller erfahren, nämlich ***fotofoto*** *und happyfoto, doch ich kann von beiden nicht aus eigener Erfahrung erzählen.*

<http://www.elena-zeitler.de/fotobuch-drucken>

(56) *Nö-Startup mit größtem Online-Garten-Sortiment im Land – Das Startup* ***GartenGarten*** *mit Sitz in Gars am Kamp (Niederösterreich) baut bereits seit drei Jahren ein Online Gartencenter fast ausschließlich mit heimischen Produkten auf.*

<https://www.derbrutkasten.com/gartengarten-no-startup-mit-grostem-online-garten-sortiment-im-land>

Anders als bei den beiden appellativen ICC-Subtypen Det-ICC und Prot-ICC wird die Referenz in den Beispielen (55–56) sowie bei *AutoAuto* in Abbildung 6 nicht primär über Konzepte hergestellt, weil den Bildungen ein semantischer Kopf fehlt. Die NP *zwei weitere Hersteller* in (55) macht deutlich, dass das ICC auf einen Hersteller referiert und nicht auf eine Entität, die zum Denotat des Basislexems *Foto* gehört. Dasselbe stellen die NPs *Startup* und *Gartencenter* im Beispiel *GartenGarten* heraus. Es handelt sich beim Bezeichneten nicht um einen Garten, sondern um einen Konzern. Auch hier fungiert das Zweitglied des ICCs also nicht als semantischer Kopf des Kompositums und entspricht somit nicht dem Konzept des Basislexems.[22]

Die ICCs in den Beispielen (55–56) sowie in Abbildung 6 sind nur teilweise durch ihre Konstituenten motiviert und evozieren keine vollständigen Konzepte. Deshalb wird die Referenz auf Entitäten auf der außersprachlichen Ebene bei diesen ICCs nicht über Konzepte hergestellt. Die ICCs referieren stattdessen direkt auf genau eine Entität im außersprachlichen Kontext. Das Konzept der Basen ist für die Referenz irrelevant. *AutoAuto* referiert als Eigenname weitestgehend ohne Umweg über ein Konzept direkt auf ein Unterhaltungsprodukt. Es gehört als Warenname zu den Ergonymen (Objektnamen). Wegen der Mono- und Direktreferenz solcher ICCs und aufgrund der Tatsache, dass der Modifikator dennoch zur Kompositumsbedeutung beiträgt, sind diese Bildungen als teilmotivierte Eigennamen zu beschreiben. Ich nenne diese ICCs deshalb Name-ICCs. Der Terminus

22 Man könnte auch annehmen, dass es das Zweit- und nicht das Erstglied eines Name-ICCs ist, das die deskriptive Bedeutung aufweist. In Anbetracht der Tatsache, dass nominale Ausdrücke im Deutschen für gewöhnlich rechtsköpfig sind, Name-ICCs aber nicht, *fotofoto* also nicht auf ein Foto, *GartenGarten* nicht auf einen Garten verweist, halte ich die Annahme, dass in Name-ICCs kein semantischer Kopf vorliegt, für nachvollziehbarer.

Name-ICC berücksichtigt, dass die Bildung solcher ICCs funktional eine Form der Onymisierung darstellt. Appellativa (*Auto*, *Foto*, *Garten*) werden onymisiert und identifizieren und individualisieren in der Folge einen Referenten.

Gewöhnlich werden im Zuge der Onymisierung auch die semantischen Merkmale der Ausdrücke getilgt (Nübling et al. 2015: 50). In allen drei Beispielen wird aber deutlich, dass die ICCs durchaus über Semantik verfügen. Die jeweiligen Basen *Auto*, *Foto* und *Garten* stellen zwar nicht die semantischen Köpfe der ICCs dar, tragen aber zur Gesamtbedeutung der Bildungen bei. Angesichts der Einteilung in Modifikator und Kopf in kanonischen N+N-Komposita kann man annehmen, dass diese Semantik vom Erstglied ausgeht. Im Kontext des Belegs *AutoAuto* wird deutlich, dass das, worauf das ICC referiert, mit einem Auto zu tun hat (*wird ein Auto zertrümmert*). Der Referent wird also konkretisiert. Bei *fotofoto* wird durch die NP *Fotobücher* im Kotext deutlich, dass das Erstglied des ICCs dem Konzept des Basislexems entspricht. Der Hersteller wird hinsichtlich der Produkte, die er herstellt (Fotoprodukte), konkretisiert. In gleicher Weise zeigt das Erstglied in *GartenGarten* an, dass der Referent irgendetwas mit Garten zu tun hat. Die Konstituenten in Modifikatorposition evozieren also durchaus das nominale Konzept der jeweiligen Basen. Diese Teilkonzepte geben den Adressat:innen aber nur einen vagen Eindruck davon, womit der Referent zu tun hat, taugen wegen des fehlenden Kopfes jedoch nicht zu einer Referenz. Die Sprecher:innen müssen die Referenz der Name-ICCs bereits kennen, um sie verwenden zu können. Die Semantik der Erstglieder motiviert diese Eigennamen bloß nachträglich und kann als Gedächtnisstütze fungieren. Es kommt nicht von ungefähr, dass alle aufgeführten Beispiele Bezeichnungen sind, die im Vermarktungskontext gebildet wurden (*fotofoto*, *GartenGarten* = Konzernnamen, *AutoAuto* = Unterhaltungsprodukt).

Wegen des fehlenden semantischen Kopfes kann man Bildungen wie *AutoAuto* als exozentrische Komposita beschreiben. Zudem ist *AutoAuto* monoreferent. Der Ausdruck referiert auf ein definites, spezifisches Objekt, nicht auf eine Menge. *AutoAuto* kann nicht auf andere Shows, in denen Autos zertrümmert werden, referieren, sondern allein auf die Unterhaltungsshow von Christian von Richthofen und Stefan Gwildis. Die semantischen und referenziellen Eigenschaften der Name-ICCs sowie ihrer Konstituenten sind in Abbildung 9 dargestellt.

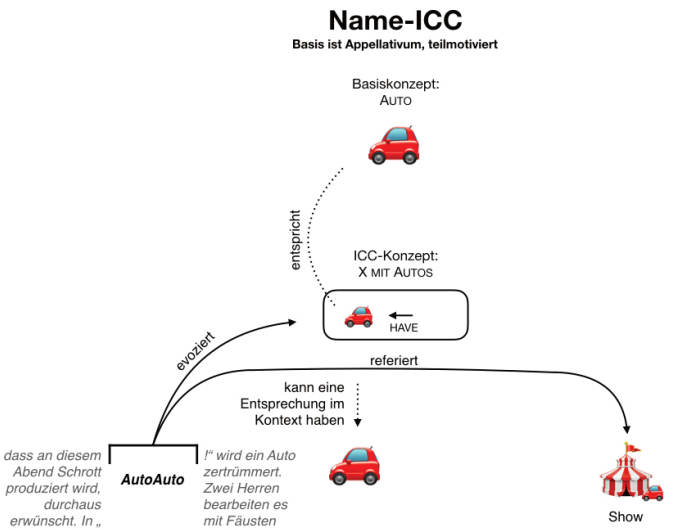

Abb. 9: Semantik und Referenz des Name-ICCs *AutoAuto* und seiner Konstituenten.

Der vierte ICC-Subtyp umfasst ICCs, bei denen weder Kriterium I noch Kriterium II erfüllt sind. Die Bildungen haben weder einen semantischen Kopf noch einen Modifikator, der dem Basislexem entspräche:

(57) *die Jungs sind sowohl einzeln als auch als Komplettpaket unterwegs, rücken Locations ins rechte Licht, strotzen vor manpower und/oder spielen nächtelang von deep bis techno. Special Guest: **PunktPunkt** aus Hamburg. Die Herren Mo und Jo collagieren seit 2009 aus den wirklich guten Dingen der technoiden Musik-Landschaft eine Stimmung, ein Gefühl.*

<http://www.detail-kiel.de/detail/2013/06>

Auch *PunktPunkt* evoziert wie die bisher beschriebenen Name-ICCs kein Konzept, über das die Referenz auf die entsprechende Entität gelänge. *PunktPunkt* hat keine Bedeutung, sondern nur eine Referenz, nämlich die auf ein DJ-Duo aus Hamburg. In ICCs dieses Typs trägt keine der Konstituenten zur Gesamtbedeutung bei. Das ICC ist ohne Semantik und die Bedeutung der Basis völlig irrelevant für den Referenten, weshalb auch solche ICCs Name-ICC sind. Im Unterschied zu den bisher besprochenen Name-ICCs ist die Beziehung zwischen der Semantik des Basislexems und der des ICCs aber gänzlich arbiträr. *PunktPunkt* ist keine bestimmte Art Punkt und hat auch nichts mit einem Punkt zu tun.

Auch die Name-ICCs können also weiter subklassifiziert werden. Name-ICCs mit Appellativa als Basis können unterteilt werden in solche, bei denen das Appellativum als Modifikator eine Bedeutungskomponente trägt und das Konzept des Basislexems realisiert, und solche, in denen das Konzept der Basis überhaupt nicht zur Geltung kommt. Die Bedeutung von *Punkt* hat nichts mit dem DJ-Duo, auf das das Kompositum referiert, zu tun und gibt auch keine Information dazu, womit das DJ-Duo zu tun hat. Die Semantik von *Punkt* ist also für die Referenz des Name-ICCs vollkommen irrelevant, spielt auch bei der Motivation der Referenz keine Rolle und der Ausdruck ist somit völlig arbiträr mit dem Musikact aus Hamburg verknüpft. Anders ausgedrückt: Der Name *PunktPunkt* weist ebenso viel Semantik auf wie der Familienname *Müller*. In beiden Fällen wurde die ursprüngliche Semantik ('winziger Fleck, kleines schriftliches Zeichen'; 'Handwerker, der Getreide mahlt') im Zuge der Onymisierung gelöscht. Da beide Konstituenten kein nominales Konzept evozieren, gibt es innerhalb des ICCs auch keine semantische Relation und es findet auch keine Subklassifikation statt. Die semantischen und referenziellen Eigenschaften dieser Name-ICCs sowie ihrer Konstituenten sind in Abbildung 10 dargestellt.

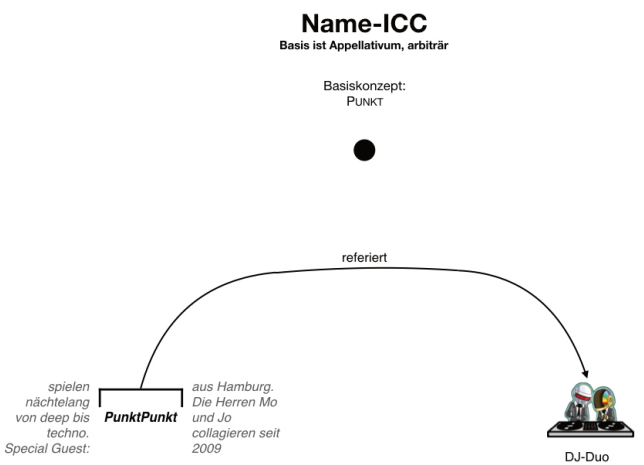

Abb. 10: Semantik und Referenz des Name-ICCs *PunktPunkt* und seiner Konstituenten.

Im fünften ICC-Subtyp ist die Basis selbst ein Eigenname. Das ICC *Kurt-Kurt* in Abbildung 11 ist monoreferent wie die anderen Name-ICCs und referiert auf ein bestimmtes Berliner Kulturzentrum. Die Basis ist ein Anthroponym, das für gewöhnlich auf eine männliche Person referiert. Das spielt aber für die Referenz auf

das Kulturzentrum keine Rolle. Basis und Name-ICC sind nicht referenzidentisch. Kein Bestandteil im ICC verweist auf ein Konzept oder eine Entität im Kontext.

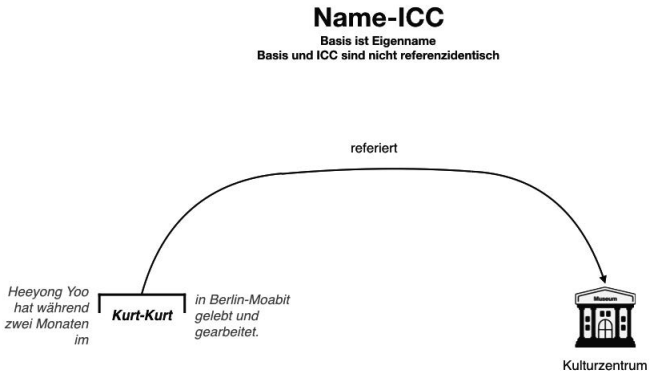

Abb. 11: Semantik und Referenz des Name-ICCs *Kurt-Kurt* und seiner Konstituenten.

Der sechste und letzte ICC-Subtyp funktioniert genau wie *Kurt-Kurt*, nur dass die Eigennamenbasis mit dem ICC referenzidentisch ist (Abbildung 12).

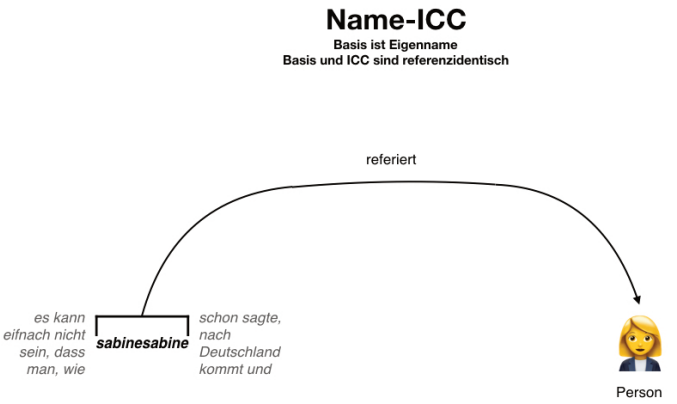

Abb. 12: Referenz des Name-ICCs *sabinesabine* und seiner Konstituenten.

Welche der zwei Konstituente nun eine Entsprechung auf außersprachlicher Ebene hat, lässt sich nicht entscheiden. Keine der Konstituenten evoziert ein Konzept. In diesem Name-ICC-Subtyp wird der Eigenname allein zur formalen Erweiterung und Variation verdoppelt (Kentner 2017: 252f.).

Die vier letztgenannten ICC-Typen stellen Subtypen der Name-ICC-Konstruktion dar und werden zu einer Gruppe zusammengefasst. *AutoAuto, fotofoto*, *GartenGarten*, *PunktPunkt*, *Kurt-Kurt* und *sabinesabine* evozieren allesamt als Gesamtkomposita kein Konzept zur Herstellung von Referenz. Stattdessen werden sie zur Mono- und Direktreferenz auf eine Entität in der Welt genutzt (Heusinger 2012: 425, Nübling et al. 2015: 17) und dienen der Identifizierung und Individualisierung von Referenten. Diese ICCs fasse ich deshalb unter dem Terminus Name-ICC zusammen. Somit gibt es drei grundlegende ICC-Typen: Det-ICCs, Prot-ICCs und Name-ICCs.

4.4 Einordnung der Prot-ICC-Erstglieder

Im vorigen Unterkapitel vertrete ich die Position, dass die Erstglieder in ICCs wie *MannMann*, *Album-Album* oder *Oma-Oma* nicht die Konzepte der jeweiligen Basen evozieren, sondern rein funktional sind. Die vorgenommene Unterscheidung zwischen Det-ICCs und Prot-ICCs hängt davon ab, inwieweit man in der Eigenschaftszuschreibung der Prot-ICCs einen kategorialen Unterschied zu den zahlreichen Modifikationsrelationen in determinativen ICCs sieht. Hierzu gibt es unterschiedliche Analysen, denn die Entscheidung, wie die Prot-ICC-Erstglieder einzuordnen sind, ist nicht trivial.

Finkbeiner (2014: 194) unterscheidet die ICC-Lesarten Prot- und Real ausdrücklich von determinativen Lesarten. Auch Hohenhaus stellt einer determinativen Lesart eine Lesart gemäß seiner Formel „an XX is a proper/prototypical X“ gegenüber. Der Unterschied zu den determinativen ICCs ist offensichtlich. Zweifelsfrei liegen bei Det-ICCs wie *Lehrer-Lehrer* zwei Instanzen des Konzeptes LEHRER vor. Auf konzeptueller Ebene gibt es einen Lehrer, der eine (andere) Menge von Lehrern unterrichtet. Aus diesem Grund kann im Kontext eines solchen Belegs fakultativ auf zwei unterschiedliche dem Konzept LEHRER entsprechende Entitäten verwiesen werden. Demgegenüber, und auch das ist eine unproblematische Annahme, liegen bei Prot-ICC wie *Album-Album* nicht zwei Alben vor, sondern bloß eines. Wenn das Konzept ALBUM in einem ICC *Album-Album* doch zweimal vorliegt, ist das ICC automatisch determinativ zu lesen und auch die Modifikationsrelationen determinativer Lesarten sind dann möglich. Ein Beispiel hierfür ist *Album-Album* in Abbildung 13, das die Mini-CD in der Abbildung bezeichnet und sich in etwa als ‘Album im Album’ im Sinne einer LOC-Relation paraphrasieren ließe.

Abb. 13: Eine Mini-CD als Album im Album *Wir wollen nur deine Seele* der Berliner Punkband *Die Ärzte.*

In determinativen ICCs können also im Äußerungskontext zwei Entitäten, die dem Basiskonzept (=Konzept des Basislexems) entsprechen, vorliegen, in ICCs mit Prototypenlesart hingegen nicht. Für diese Beobachtung gibt es zwei mögliche Erklärungen, die in der Literatur zu ICCs zum Teil schon erwähnt werden, und die sich wie folgt zusammenfassen lassen:

1. Die Erstglieder in Prot-ICCs sind rein funktional und deshalb ohne lexikalische Bedeutung.
2. Die Erstglieder in Prot-ICCs tragen zwar lexikalische Bedeutung, sind aber keine Nomina.

Die erste Analyse nimmt Finkbeiner (2014: 188) vor. Sie analysiert das Erstglied in Prot-ICCs als ein Element, das keine lexikalische Bedeutung, sondern nur eine Funktion hat, nämlich die, die Prototypenbedeutung anzuzeigen. Als rein funktionales Element wäre das Erstglied eines Prot-ICCs also ein abstraktes grammatisches Morphem, dessen grammatische Bedeutung die Markierung der Prototypenlesart ist und dessen Form erst über den Prozess der Reduplikation festgelegt wird. Übernimmt man Finkbeiners Analyse der Prot-ICC-Erstglieder, lässt sich leicht erklären, warum im Äußerungskontext von Prot-ICCs keine zwei Entitäten, die dem Basiskonzept entsprechen, vorliegen können. Das Erstglied dient allein dazu, das Muster Prot-ICC zu realisieren und evoziert ergo kein Teilkonzept, das im Äußerungskontext eine ontologische Entsprechung haben könnte.

Die zweite mögliche Ursache für die Besonderheit der Prot-ICCs erwähnen Bross und Fraser (2020: 6). Sie sprechen von einem „adjectival flavor“ der Erstglieder solcher ICCs, gehen hier also von einer eher lexikalischen Bedeutung aus (ebenso zum Englischen: Ghomeshi et al. 2004: 343, Stolz et al. 2011: 202ff.). Adaptiert man Bross und Frasers Analyse, kann man den Unterschied zwischen Erstkonstituenten in Det-ICCs und denen in Prot-ICCs mit Rückgriff auf die Unterscheidung zwischen spezieller und charakterisierender Generizität näher beschreiben („particular“ versus „characterizing“, Krifka 2003: 180, Krifka et al.

1995: 2ff., Löbner 2011: 280). Bei der speziellen Generizität verweisen Modifikatoren generisch auf Vertreter eines nominalen Konzeptes („reference to an entity that is related to specimens"), bei der charakterisierenden Generizität werden hingegen Eigenschaften, beziehungsweise Generalisierungen über diese Entitäten ausgedrückt („express generalizations about sets of entities or situations", Krifka 2003: 180). Krifkas Termini beziehen sich auf Sätze und darin befindliche Nominalphrasen, wobei er die charakterisierende Generizität als ein Phänomen ansieht, das stets den gesamten Satz betrifft. Pallottino und Ihsane (2015) beziehen die Begriffe aber auf Nomina und visualisieren diese verschiedenen Ebenen der Generizität wie folgt (Abbildung 14):

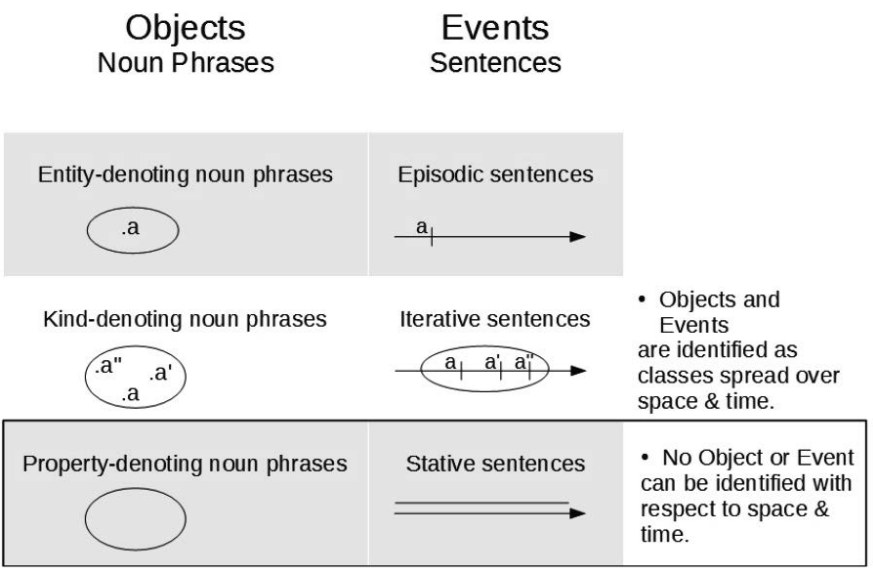

Abb. 14: Verschiedene Ebenen der Generizität (Pallottino & Ihsane 2015: 3).

Diese Unterscheidung lässt sich auch auf die Erstglieder von N+N-Komposita anwenden. Während das Erstglied *Reus-* im referenzspezifizierenden Kompositum *Reus-Freistoß* in (14) auf eine konkrete Entität, nämlich auf die Person Marco Reus, referiert („entity-denoting") und *Fenster* im Det-ICC *Fenster-Fenster* auf die Klasse der Fenster verweist („kind-denoting"), verweist *Album* im Prot-ICC *Album-Album* auf Eigenschaften („property-denoting"). Die Det-ICC-Erstglieder evozieren also nominale Konzepte und als „kind-denoting noun phrases" können ihnen Lebewesen, Gegenstände und Sachverhalte mit räumlicher und zeitlicher Aus-

dehnung zugeordnet werden (Pallottino & Ihsane 2015: 3). Erstglieder von Prot-ICCs sind „property-denoting noun phrases“, darum liegt ihnen kein nominales Konzept zugrunde, sondern ein adjektivisches. Ein *Album-Album* im Sinne von ‘richtiges, prototypisches Album’ ist ein Album mit den üblichen Eigenschaften eines Albums, ein *Album-Album* im Sinne von ‘Album im Album’ ist hingegen ein Album, das ein weiteres Album beinhaltet, zu dem ein nominales Konzept vorliegt.

Die Entscheidung, welche der zwei Analysen man anwendet, geht mit unterschiedlichen Ansichten zur Abgrenzung zwischen Prot-ICCs, Det-ICCs und N+N-Komposita im Allgemeinen einher. Nimmt man mit Finkbeiner (2014) rein funktionale Erstglieder in Prot-ICCs an, unterscheiden diese Erstglieder die Bildungen kategorial von anderen N+N-Komposita, in denen die Erstglieder stets lexikalische Bedeutung tragen. Nimmt man hingegen an, dass die Erstkonstituenten in Prot-ICCs adjektivisch (und damit lexikalischen Gehalts) sind, besteht zwischen Prot-ICCs und Det-ICCs sowie zwischen Prot-ICCs und N+N-Komposita im Allgemeinen ein eher gradueller Unterschied. Das Merkmal, eine Eigenschaftszuschreibung in Bezug auf den Referenten des Kopfkonzeptes vorzunehmen, ist nämlich kein Alleinstellungsmerkmal von Prot-ICC-Erstgliedern. Auch in N+N-Komposita mit unterschiedlichen Konstituenten existiert diese Art der Modifikation. Ähnlich wie das Erstglied von *Oma-Oma* führt beispielsweise das Erstglied in *Omabrille* dazu, dass der Brille bestimmte Eigenschaften zugeschrieben werden (Material aus Horn, dicke Gläser, Brillenbänder an den Bügeln, …). In beiden Fällen, sowohl bei *Oma-Oma* als auch bei *Omabrille*, werden Eigenschaften des Referenten des Zweitgliedkonzeptes spezifiziert.

Auch wirken Modifikationsrelationen bei Prot-ICCs anwendbar und rücken sie in die Nähe kanonischer N+N-Komposita. Die von Jackendoff (2009: 191) genannte Funktion KIND etwa wirkt sehr ähnlich zum Verhältnis zwischen Erst- und Zweitglied in Prot-ICCs. Als Beispiel für die Funktion führt Jackendoff *ferryboat* an, das also ein ‘Boot vom Typ Fähre’ ist. Ebenso könnte man argumentieren, dass eine *Oma-Oma* eine ‘Oma vom Typ Oma’ ist.

Zudem tritt in seltenen Fällen auch bei Erstgliedern von N+N-Komposita mit unterschiedlichen Konstituenten charakterisierende Generizität auf. Manche Kopulativkomposita können etwa ähnlich wie *Oma-Oma* paraphrasiert werden. Ein *Bettsofa* ist beispielsweise ein ‘Sofa mit den Eigenschaften eines Bettes’.

Es sprechen aber einige Argumente gegen die Analyse, dass die Erstkonstituenten in Prot-ICCs mit denen in kanonischen N+N-Komposita gleichzusetzen sind. Erstens sind Kopulativkomposita wie *Bettsofa* eher nach den Mustern ‘X, das gleichzeitig Y ist’ oder ‘X, das gleichzeitig Y und Z ist’ zu paraphrasieren („both N1 and N2“, Jackendoff 2009: 123, „B which is also A“, Marchand 1969: 41). Ein *Bettsofa*

ist also eher ein ‘Bett, das gleichzeitig Sofa ist’ oder ein ‘Möbelstück, das gleichermaßen Bett und Sofa ist’. Eine *Oma-Oma* ist aber keine ‘Oma, die gleichzeitig Oma ist’ beziehungsweise keine ‘Oma, die gleichermaßen Oma und Oma ist’. Zweitens erlauben die meisten Kopulativkomposita nicht die Paraphrase nach dem Muster ‘Y mit den Eigenschaften von X’. Ein *Dichter-Komponist* ist etwa kein Komponist mit den Eigenschaften eines Dichters, sondern eine Person, die den Beruf des Dichters und den des Komponisten gleichermaßen ausübt. Auch andere übliche Beispiele für Kopulativkomposita sind eher über das Muster von Marchand (1969) als über das von *Oma-Oma* zu paraphrasieren (*Muttertier*, *Singer-Songwriter*, *Strumpfhose*). Drittens ist selbst die eingangs erwähnte Eigenschaftsparaphrase bei *Bettsofa* nicht äquivalent zu der von *Oma-Oma*: Ein *Bettsofa* ist kein ‘Sofa mit den typischen Eigenschaften eines Bettes’. Die Zuschreibung prototypischer Eigenschaften entsteht nicht durch das Erstglied, sondern ist idiosynkratischer Bestandteil der Konstruktion Prot-ICC, die durch das Element in Erstgliedposition entsteht.

Auch funktioniert die Eigenschaftszuschreibung in determinativen N+N-Komposita anders als in Prot-ICCs. Eine *Oma-Oma* ist eine ‘Oma mit typischen Eigenschaften einer Oma’, eine *Omabrille* aber keine ‘Brille mit typischen Eigenschaften einer Oma’. Andersherum ist eine *Omabrille* eine ‘Brille, die typisch ist für eine Oma’, eine *Oma-Oma* ist aber keine ‘Oma, die typisch ist für eine Oma’. Der Unterschied liegt hier in der Anzahl nominaler Konzepte. Im Gegensatz zum Erstglied des N+N-Kompositums *Omabrille* liegt zum Prot-ICC-Erstglied kein nominales Konzept vor. Eine *Omabrille* ist eine Entität (eine Brille), die einer anderen Entität (einer Oma) gehört, beziehungsweise eine Brille, von der angenommen wird, dass sie von Menschen aus der Gruppe der Omas getragen wird. Sprecher:innen inferieren Eigenschaften des Objektes (der Brille) aus der Erfahrung heraus, dass bestimmte Eigenschaften dieses Objektes genau dann zu beobachten sind, wenn die Träger:innen dieser Objekte ältere Frauen sind. Eine *Oma-Oma* ist aber keine Oma, die typisch für eine andere Oma ist, beziehungsweise eine Oma, von der angenommen wird, dass ihre Charakteristika mit anderen Menschen aus der Klasse der Omas zu tun hat. Die Eigenschaft, generisch im Sinne von „property-denoting“ zu sein, kommt also nur dem Erstglied eines Prot-ICCs zu, nicht aber dem Erstglied von N+N-Komposita wie *Omabrille.* Charakterisierende Generizität der ersten Konstituente ist also, vor allem in der Gegenüberstellung mit den anderen ICC-Typen, ein Alleinstellungsmerkmal der Prot-ICCs. Die Erstkonstituenten von Det-ICCs und teildeskriptiven Name-ICCs evozieren stets das nominale Konzept der Basis.

In dieser Arbeit vertrete ich gewissermaßen eine Kombination der beiden Analysen von Finkbeiner (2014) und Bross und Fraser (2020) und nehme an, dass das Erstglied in einem Prot-ICC ein funktionales Element ist, das zu einer Eigen-

schaftszuschreibung innerhalb des ICCs führt. Die Frage, ob die Erstglieder von Prot-ICCs nun als adjektivische oder als rein funktionale Elemente anzusehen sind, ist damit letztlich unerheblich für die Unterscheidung Det-ICC/Prot-ICC. Auch wenn man die Erstglieder in Prot-ICCs als lexikalische Einheiten ansieht, entsprechen sie nicht dem Konzept des Basislexems. Die klare Unterscheidung zwischen Prot-ICCs und Det-ICCs, beziehungsweise N+N-Komposita im Allgemeinen bleibt in jedem Fall gewahrt. Die Unterscheidung zwischen Det- und Prot-ICCs anhand Kriterium I ist also gerechtfertigt.

4.5 Vor- und Nachteile einer konzeptbasierten ICC-Subklassifikation

Wie im Forschungsüberblick dargestellt, zielen bisherige Versuche der Differenzierung einzelner ICC-Typen auf die Unterscheidung zwischen determinativen ICCs und Prototypen-ICCs ab (Finkbeiner 2014). Das Kriterium für die Unterscheidung ist hier stets die semantische Relation zwischen den Konstituenten. Abgesehen davon, dass Name-ICCs bei diesen Subklassifikationsansätzen bisher überhaupt nicht berücksichtigt werden, bergen die bisherigen Ansätze die Gefahr, wegen der unscharfen Grenze zwischen den Modifikationsrelationen beliebig zu sein. Folgendes Beispiel verdeutlicht dies:

(58) *Ist das der „**Chefchef**", sprich der Geschäftsführer selbst?*

<http://www.foraus.de/forum/showthread.php?3786-3-Welche-betrieblichen-Bedingungen-können-die-Ausbildung-am-Arbeitsplatz-erschweren&s=b3761cc97415a6b8f222c12815e 59314>

Will man dieses ICC anhand der semantischen Relation zwischen den Konstituenten klassifizieren, ist schwer zu entscheiden, welche Relation vorliegt. *Chefchef* kann sowohl im Sinne eines Prot-ICCs paraphrasiert werden ('der richtige Chef') als auch im Sinne eines Determinativkompositums mit einer OF- oder FOR-Relation ('der Chef der Chefs, der Chef des Chefs, der Chef für Chefs'). Man könnte zudem, wie Freywald (2015: 923) und Kentner (2017), die Prototypenlesart einfach als eine weitere semantische Relation und somit Prot-ICCs als Subtyp der Determinativkomposita verstehen. Fest steht, dass in beiden ICC-Typen Subkonzepte gebildet werden und das Erstglied daran maßgeblich beteiligt ist.

Und doch gibt es Unterschiede zwischen Prot-ICCs und Det-ICCs. Wie beschrieben evozieren letztere zwei nominale Konzepte, während bei Prot-ICCs nur ein nominales Konzept vorliegt. Aus diesem Grund können nur bei Det-ICCs die

Konzepte beider Konstituenten wieder aufgenommen werden und als ontologische Einheiten im Kontext konkret vorliegen. Bei Prot-ICCs ist das nicht der Fall, weshalb nur das Kopfkonzept im Kontext wieder aufgenommen und als ontologische Einheiten realisiert werden kann. Auch sind die Modifikationsrelationen bei Det-ICCs vielfältig. Prot-ICCs hingegen sind auf die Zuschreibung prototypischer Eigenschaften begrenzt. Ein Subklassifikationsansatz, der Prot-ICCs den Det-ICCs beiordnet, ignoriert diese Unterschiede.

Der von mir gewählte Ansatz, ICCs in Subklassen zu gruppieren, bietet demgegenüber entscheidende Vorteile. Erstens behebt er das Problem unscharfer Grenzen zwischen semantischen Relationen, weil er gar nicht erst versucht, die Relationen zu klassifizieren und als Unterscheidungskriterium zu verwenden. Zweitens berücksichtigt mein Ansatz den zentralen Unterschied zwischen Det- und Prot-ICCs, indem er die Frage, ob die Erst- und Zweitglieder Konzepte evozieren, als Unterscheidungskriterien verwendet. Drittens ist es mit den zwei Unterscheidungskriterien meines Ansatzes möglich, Name-ICCs angemessen zu beschreiben und wiederum zu subklassifizieren. Nur bei Name-ICCs evoziert das Zweitglied kein dem Basislexem entsprechendes Konzept. Viertens können die Ausprägung der Kriterien im konkreten Kotext der ICCs klar festgestellt werden. Beim ICC *Chefchef* im obigen Beispiel wird im Kontext beispielsweise auf keine weitere Entität eines Chefs verwiesen als auf die, auf die das ICC verweist, und das ICC setzt nicht zwei nominale Konzepte in Bezug, sondern schreibt der Entität, also dem Chef, Eigenschaften zu. Dass dies die prototypischen Eigenschaften eines Chefs sind, ist eine Bedeutungskomponente, die durch die Konstruktion ICC entsteht. Ein ICC mit demselben Stamm als Basis kann aber auch Det-ICC sein:

(59) *Mit einem eingeschriebenen Brief informierten sie die Chefs ihres Chefs über die Zustände in ihrem Betrieb. Am selben Tag, an dem die Beschwerde die* ***Chefchefs*** *erreichte, wurde sie ins Büro ihres Vorgesetzten zitiert.*

<http://bruellen.blogspot.de/2005_09_01_archive.html>

Auch hier kann mithilfe des Kontextes klar bestimmt werden, ob Erst- und Zweitglied auf zwei verschiedene Konzepte verweisen. Da mit *Chefs* eine Entität eingeführt wird, die sich von der, auf die das Konzept des ICCs verweist, unterscheidet, ist der Beleg eindeutig ein Det-ICC. Das Erstglied hat hier keinen adjektivischen Charakter, sondern verweist auf das Konzept einer Person, die ein Chef ist.

Wie wichtig der Kontext für die Kodierung der ICCs ist, zeigen die Beispiele (60–62). Die ICCs in diesen Beispielen sind formal identisch mit den bereits besprochenen Beispielen *Hundehund*, *Oma-Oma* und *AutoAuto.* In einem veränderten Kontext gehören sie zu anderen ICC-Typen:

(60) *Ich meine so nen richtigen **Hundehund**. Einen Schäferhund zum Beispiel.*

(61) *Auch meine Oma hatte eine Oma. Das war dann also quasi die **Oma-Oma**.*

(62) *Das **AutoAuto** – Künstlerin baut richtiges Auto aus über 2000 Spielzeugautos.*

Ungeachtet ihrer formalen Merkmale gehören die Bildungen in einem veränderten Kontext anderen ICC-Typen an. Das Det-ICC *Hundehund* aus (1) ist in (60) ein Prot-ICC, das Prot-ICC *Oma-Oma* aus (54) ist in (61) Det-ICC, ebenso wie das Name-ICC *AutoAuto* aus (62). Nach dem gewählten Ansatz stützt sich die Kodierung der Belege weder auf formale Merkmale der ICCs noch auf syntaktische Merkmale der ICC-NP oder des Satzes, in das das ICC eingebettet ist. Aus dem Kontext wird für die Subklassifikation einzig die semantische Köpfigkeit herangezogen sowie die Information, ob das Erstglied das Konzept des Basislexems realisiert.

Doch auch der hier gewählte Ansatz der ICC-Subklassifikation ist nicht ohne Schwierigkeiten. Die Abgrenzung der Prot-ICCs basiert darauf, dass das Erstglied nicht das Konzept des Basislexems realisiert. Ein Problem für meinen Subklassifikationsansatz sind deshalb alle Belege, bei denen nicht einfach zu entscheiden ist, ob im Erstglied das Konzept des Basislexems vorliegt. Dies ist etwa bei Nomina der Fall, die (auch) Stoffsubstantive sind:

(63) *Baguettes, Pasteten, Würste, Käse, Porzellangeschirr und echte **Glasgläser** - und das Ganze angerichtet auf dem eigens mitgebrachten Picknicktisch mit vier integrierten Sitzen.*

<http://blog.szon.de/szon/index.php?/archives/P59.html>

(64) *Suche: Echte **Holz-Hölzer**, nicht nur für Ästheten: Persimmon Golfschläger von Bagger Vance. Nicht nur die Optik überzeugt, auch die Spielleistung stellt herkömmliche Metallhölzer in den Schatten. Dabei liefern diese Persimmons ein wunderbares Feedback und belohnen durch eine maximale Ballkontrolle. Passend dazu werden diese massiven Holzköpfe mit traditionellen Stahschäften oder den topaktuellen Hickory-Schäften geliefert.*

<https://www.golfershouse.de/pages/innovationen/persimmon-golfschlaeger.php>

(65) *ich kann nur sagen, was ich für mich "Milchkaffee" nenne: Filterkaffee mit besonders viel Milch. Also,* ***Milchmilch****, nicht Kaffeemilch.*

<http://cre.fm/cre119>

Hier ist ein semantischer Unterschied zwischen Erst- und Zweitglied erkennbar. Das Erstglied erhält die Materiallesart des Stoffsubstantivs, das Zweitglied ist hingegen metonymisch verschoben (Material –> Gegenstand) und verweist als Individualsubstantiv auf ein abgegrenztes Objekt (Trinkgefäß). Das Problem mit solchen Bildungen ist, dass nicht ganz einfach zu entscheiden ist, ob Stoffsubstantive dasselbe Konzept, nämlich das der Basis, evozieren. Deshalb können ICCs mit Stoffsubstantiven als Basis sowohl als Det- als auch als Prot-ICCs aufgefasst werden. *Glasgläser* sind 'Gläser aus Glas'. HAVE- oder MAKE-Relationen sind problemlos anwendbar. Allerdings ähnelt diese Bildung in ihrer Verwendung den Prot-ICCs dergestalt, dass sie den Bedeutungsgehalt der Basis auf seinen prototypischen Kern beschränken. Die Dopplung verstärkt das Konzept selbst und setzt den Referenten in einen Kontrast zu weniger prototypischen Vertretern. Unterstützt wird der Eindruck einer Prototypenlesart durch den Zusatz, dass es sich um *echte* Glasgläser handelt. Analog dazu sind *Holz-Hölzer*, also Golfschläger aus Holz, Gegenstände, die aus einem Material bestehen, dessen Substantiv durch semantische Verschiebung neben dem Stoffsubstantiv auch Individualsubstantiv ist. Auch hier wird *echt* adjektivisch attribuiert. Auch *Milchmilch* kann man sowohl als determinativ als 'Milch aus Milch' als auch als Prot-ICCs 'richtige Kuhmilch' paraphrasieren. Im Unterschied zu *Glasgläser* und *Milchmilch* stellen *Holz-Hölzer* aber nicht den Normalfall in ihrer Gegenstandsklasse dar. Ein (Trink-)Glas besteht für gewöhnlich aus Glas, eine Milch aus Milch. Ein Golfschläger besteht aber in der Regel aus Metall. Im Kotext des Belegs werden die *Metallhölzer* denn auch als die *herkömmlichen* bezeichnet.

In der Regel führt die Bildung eines ICCs auf Basis eines Stoffsubstantivs zu einer Prototypenbedeutung, weil die Referenten der Individualsubstantive, zu denen sie transponiert werden können, üblicherweise aus dem Material bestehen, auf das die Stoffsubstantive verweisen (*ein Holz-Holz* 'Latte aus echtem Holz', *eine Milchmilch* 'eine Milch aus Milch, Kuhmilch', *ein Eisen-Eisen* 'Golfschläger aus Metall' oder 'Eisenstange aus richtigem Eisen wie Moniereisen, Bewehrungsstahl', *ein Gips-Gips* 'Verband aus Calciumsulfat-Dihydrat'). Ist das nicht der Fall, besteht also wie bei den *Holz-Hölzern* 'Golfschläger' oder einem *Quecksilber-Quecksilber* 'Thermometer mit echtem Quecksilber' der Gegenstand, auf den das Zweitglied referiert, üblicherweise gerade nicht aus dem Material, auf das das Erstglied verweist, ist neben der Materiallesart aber auch die Real-Lesart möglich: 'richtiges Holz, wie es sich eigentlich gehört, nämlich aus echtem Holz', 'richtiges Thermo-

meter, so wie es ursprünglich mal erfunden wurde'. Die Klassifikation solcher Bildungen verlangt also immer eine Entscheidung, deren Kriterien weit über die zuvor genannten Kriterien I. und II. hinausgehen.

Hinsichtlich der Kodierung solcher Belege wurde wie folgt vorgegangen: Aus dem Material zu bestehen, das das Stoffsubstantiv bezeichnet, ist nur eine Art der Eigenschaftszuschreibung für das jeweils gegebene Objekt. Es liegt in jedem Fall nur einmal das nominale Konzept der Basis vor (Glas, Golfschläger, Getränk). Belege dieser Art werden somit als Prot-ICC kodiert, weil sie semantisch rechtsköpfig sind und das Erstglied dem Referenten des Zweitgliedes prototypische Eigenschaften zuweist. Diese Setzung führt zwar dazu, dass auch N+N-Komposita distinkter Konstituenten wie *Glasfenster* oder *Holzeimer* als Komposita zu werten wären, bei denen das Erstglied als Stoffsubstantiv charakterisierend generisch ist und darum keine Entsprechung im Kontext haben kann. Gegenüber anderen ICC-Arten mit Stoffsubstantiven als Basis sind Prot-ICCs allerdings eindeutig abzugrenzen. Die entsprechenden Belege werden nur dann als Det-ICCs klassifiziert, wenn das Erstglied eindeutig ein nominales Konzept evoziert, das sich als Gegenstand und nicht nur als Material im Äußerungskontext manifestiert, also etwa *Glasglas* als Verweis auf ein Trinkglas, das sich in einem anderen Trinkglas befindet.

Problematisch sind auch alle anderen Fälle, in denen die Bestandteile wie in den Beispielen *Glasgläser*, *Holz-Hölzer* oder *Milchmilch* eine semantische Verschiebung aufweisen. Auch bei Bildungen, deren Basis kein Stoffsubstantiv ist, ist die Identität der Konstituenten mitunter diskussionswürdig:

(66) *Früher hiessen die Mückenmittel "Autan", heute heissen sie "**Bremsenbremse**", "Antibrumm" oder "Mückweg".*

<http://www.renate-prior.de/html/juli2003.htm>

Das Beispiel *Bremsenbremse* in (66) zeigt einen Fall, in dem die Konstituenten nicht nur semantisch verschoben, sondern gänzlich divergent sind. Es liegt Homonymie vor. Das Erstglied *Bremse* im Sinne von 'blutsaugende Stechfliege' leitet sich von ahd. *breman*, mhd. *Bremen* 'brummen, brüllen' ab, das Zweitglied in der Bedeutung 'Hemmvorrichtung' hingegen stammt von mnd. *Premese* 'Maulholz, Zügel'. Es stellt sich also die Frage, ab welchem semantischen Unterschied zwischen Erst- und Zweitglied man Fälle wie *Bremsenbremse* nicht mehr den ICCs, sondern nur noch allgemeiner den N+N-Komposita zuordnet.

Es lässt sich schwer abgrenzen, ab wann die Semantik eines ICC-Bestandteils so weit verschoben ist, dass man von zwei unterschiedlichen Lexemen und nicht nur von unterschiedlichen Lesarten sprechen kann. Die Grenze zwischen ICCs und bloßen N+N-Komposita wird hier zugunsten der ICCs gezogen. In beiden

Fällen, also sowohl wenn die Konstituenten eines N+N-Kompositums wegen Polysemie semantisch nicht ganz deckungsgleich sind, als auch wenn sie wegen Homonymie semantisch ungleich sind (*Bremsenbremse*), gelten die entsprechenden Bildungen in dieser Arbeit als ICCs. Erst wenn zu der semantischen Verschiebung eine formale hinzukommt, etwa bei Komposita wie *Bären-Bar*, *Reis-Reise* oder *Wagen-Waage*, werden die Bildungen nicht als ICCs angesehen.

Bremsenbremse zeigt ein weiteres Problem auf: Es kann determinativ interpretiert werden im Sinne von ‘eine Bremse für die Bremsen = ein Schutz gegen Mücken’. Gleichzeitig ist das ICC ein Markenname. Zwar ist ein Markenname nicht im strengen Sinne monoreferent. Zum einen kann er, wie etwa *Tempo*, deonymisiert und als Gattungsbegriff für alle Produkte der Klasse verwendet werden. Zum anderen referiert ein Markenname nicht auf genau eine Entität in der Welt, sondern auf eine Menge an Entitäten, die lediglich nahezu vollständig identisch sind, nämlich auf alle Insektenschutz-Sprühflaschen des Herstellers. Als Markenname kann *Bremsenbremse* dennoch als Eigenname angesehen werden. Die semantische Verschiebung von *Bremse* im Sinne von ‘Bremsvorrichtung’ hin zu ‘Mittel gegen etwas (Insekten)’ bedingt, dass der semantische Kopf des Kompositums wie bei den Name-ICCs außerhalb des Kompositums liegt. Das macht es sehr schwierig, zu entscheiden, ob das ICC semantisch rechtsköpfig ist, ob das ICC also ein Det- oder Name-ICC ist. In solchen Fällen gilt es zu entscheiden, ob der Referent des ICCs, in diesem Fall das Mückenspray, noch dem Denotat des Basislexems *Bremse* zuzurechnen ist, und inwieweit Metaphern und Metonymien berücksichtigt werden.

Dieser Aspekt der Grenzziehung zwischen Det- und Name-ICCs wird in dieser Arbeit wie folgt behandelt: ICCs, in denen das Kopfnomen metaphorisch auf die vom ICC bezeichnete Entität verweist, weil zwischen ICC-Referent und Referent des Basislexems Similarität hinsichtlich eines Konzeptes besteht, sind keine Name-ICCs. *Bremsenbremse* ist also zunächst ein Det-ICC, weil das Kopfnomen *Bremse* metaphorisch auf die vom ICC bezeichnete Entität *Mückenspray* verweisen kann. In diesem Fall besteht Similarität hinsichtlich des Konzeptes des Aufhaltens und Schützens. ICCs wie *AutoAuto* oder *PunktPunkt* sind hingegen immer Name-ICCs, da es auch unter Zuhilfenahme von Metaphern nicht möglich ist, auf die entsprechenden Entitäten (Show, DJ-Duo) zu verweisen.

Eine weitere Gruppe von ICCs, bei denen die Unterscheidungskriterien nicht ohne Weiteres zu einer eindeutigen Einordnung führen, sind solche, bei denen eine semantische Relation des Beinhaltens eine Rolle spielt:

(67) *Wie groß war die Erleichterung, als ich im vergangenen Jahr mehrere Umzugskartons an Büchern an Reseller verkauft, den Rest verschenkt, und den **Rest-Rest** (also die, die nun wirklich niemand wollte) weggeworfen hatte.*

<http://blogweise.junfermann.de/2012/06/06/e-books-zu-teuer>

(68) *SPOILER: Das Ende hat mich überrascht, sogar positiv. Und zwar, weil es so "negativ" ist. Naja, nicht das **Ende-Ende**, aber kurz davor, als Vincent Marie festgekettet in dem italienischen Krankenhaus zurücklässt.*

<http://www.mokita.de/blog/2010/07/20/film-vincent-will-meer>

Beide ICCs lassen sich unter Anwendung einer OF-Relation determinativ lesen: ‘der Rest vom Rest’, ‘das Ende vom Ende’. Die Bedeutung von *Rest* gibt das DWDS wie folgt an:[23]

- etw., was beim Verbrauch, Verzehr von etw. übriggeblieben ist
- etw., was von etw. weitgehend Verschwundenem, Geschwundenem noch vorhanden ist
- etw., was von etw. Vergangenem, Zerstörtem, Verfallenem, Abgestorbenem noch vorhanden ist; Überrest
- letztes (nur noch zu reduziertem Preis verkäufliches) Stück von einer Meterware

Rest verweist also darauf, dass es einen deutlichen Mengenunterschied gibt zwischen dem, was von einem Gegenstand, einer Menge oder Masse vorlag, und dem, was noch vorliegt. Demnach gibt es zwei unterschiedliche Konzepte, die die Kompositakonstituenten hervorrufen. Im Beispiel (67) verweist das Erstglied auf die Menge der Gegenstände, die nach dem Verkauf *mehrerer Umzugskartons an Büchern* noch da ist, nämlich ein *Rest*, das Zweitglied verweist auf die Menge der Gegenstände, die noch übrig ist, nachdem auch dieser *Rest verschenkt* worden ist. Dass hier Modifikator und Kopf zweimal das Konzept der Basis hervorrufen, spricht für die Annahme eines Det-ICCs. *Rest-Rest* verweist aber gleichzeitig darauf, dass der Prozess des Verschwindens von etwas und des Übrigbleibens eines Teiles davon erneut stattfindet. Der Unterschied zwischen dem, was einst da war und dem, was davon noch vorhanden ist, wird größer und somit das Konzept REST verstärkt. Im genannten Beispiel zeigt sich das zusätzlich dadurch, dass *Rest-Rest*

23 „Rest“ das Digitale Wörterbuch der deutschen Sprache, <https://www.dwds.de/wb/Rest>, abgerufen am 21.03.2021.

in Kontrast zu *Rest* gesetzt wird. Der Beleg weist also auch eine Prototypenbedeutung auf.

Rest-Rest verhält sich ähnlich wie das oben besprochene *Chefchef* in (59), nur dass der Referent des ICCs zum Zeitpunkt der Äußerung nicht mehr Teil der übergeordneten Gruppe ist. Der *Chefchef* in (59) ist der Chef einer Menge an Chefs, *Rest-Rest* ist der Rest einer Menge, die aber nicht mehr da ist. Die Entscheidung, ob zwei unterschiedliche Konzepte (Gruppe und Bestandteil) gegeben sind, wird in solchen Fällen mithilfe des Kontextes für jedes Beispiel gesondert bestimmt. Ähnlich wie bei *Chefchef* wird ein ICC den Det-ICCs zugeordnet, wenn die übergeordnete Gruppe im Kotext gegeben ist; andernfalls den Prot-ICCs. Im gegebenen Beispiel ist *Rest-Rest* demnach Det-ICC und äquivalent zu *Chefchef* in (59). In anderen Fällen, wenn nämlich wie bei *Chefchef* in (58) unabhängig von einem weiteren „Rest“ ‘richtiger Rest’ gemeint ist, das Erstglied also kein nominales Konzept hervorruft, wird das ICC als Prot-ICC klassifiziert.

Auch die Abgrenzung zwischen Det- und Name-ICCs, sowie zwischen Prot- und Name-ICCs bedarf in manchen Fällen der näheren Erläuterung. Name-ICCs sind ICCs mit Monoreferenz. Sie verweisen nicht über Konzepte, sondern direkt auf eine Entität wie etwa einen Menschen oder ein Produkt. Die Relation zwischen den Konstituenten ist leer, weil dem Zweitglied kein Konzept entspricht. In machen Fällen überlappen sich in einem ICC allerdings die Name-ICCs und Prot-ICCs zugeschriebenen Funktionen (69):

(69) *War der „Sat.1 **FilmFilm**“ nicht mal der Platz für die ganz großen Blockbuster?*

<https://taz.de/Kolumne-Fernsehen/!5060726>

FilmFilm ist hier einerseits Prot-ICC, denn es bezeichnet einen ‘richtigen Film’, einen Kinofilm/Blockbuster. Andererseits ist *FilmFilm* der Name eines Unterhaltungsproduktes. Die Überschneidung der Funktionen von Prot-ICC und Name-ICC kommt nicht von ungefähr. Marketingabteilungen von Unternehmen machen sich Prot-ICCs zunutze und versuchen so, ihre Produkte als beste Vertreter ihrer Klasse zu vermarkten. Als Name für ein Produkt verwendet, handelt es sich in (69) also um ein onymisiertes Appellativum. Die Dopplung in beiden ICC-Typen dient der Expressivität und Intensivierung. Damit entsprechen beide ICC-Typen der in der Sprachtypologie vorherrschenden Annahme, dass Reduplikationen immer die Kernbedeutung der Intensivierung zukommt (Botha 1988: 100, Cardona 1988: 254, Stolz et al. 2011: 10). Das für die Subklassifikation gewählte Kriterium bildet den Unterschied zwischen Prot-ICCs und Name-ICCs ohne Rückgriff auf Eintragungen in Markenregistern ab. Je nach Verwendung von *FilmFilm* ist die Bildung Prot-

oder Name-ICC. Im gegebenen Beispiel (69) ist *FilmFilm* ein Prot-ICC, weil hier klar ein semantischer Kopf vorliegt, das Erstglied aber nicht das Konzept FILM evoziert und es im Kontext also keinen zweiten Film geben kann. In Fällen, in denen kein semantischer Kopf vorliegt, *FilmFilm* etwa eine Filmzeitschrift bezeichnet und damit äquivalent zu *AutoAuto* oder *GartenGarten* ist, wird es als Name-ICC klassifiziert.

4.6 Zusammenfassung

Die Korpusstudien haben gemeinsam eine Datenbasis von über 180.000 ICC-Belegen hervorgebracht. In diesem Kapitel wurden die semantisch-funktionalen Aspekte dieser ICCs beschrieben, Kriterien für die Subklassifikation aufgestellt und die Belege hinsichtlich dieser Kriterien subklassifiziert, nämlich I. dem Vorliegen des Basiskonzeptes im Erstglied sowie II. dem Vorliegen des Basiskonzeptes im Zweitlied (= semantischer Kopf). Nach diesen Kriterien lassen sich drei ICC-Typen unterscheiden: Sind I. und II. erfüllt, handelt es sich bei den Bildungen um Det-ICCs. Diese bilden ein Subkonzept zum Basislexem und setzen zwei Konzepte ins Verhältnis. Ist nur II. erfüllt, handelt es sich um Prot-ICCs. Auch sie bilden ein Subkonzept zum Basislexem, allerdings mithilfe eines rein funktionalen Erstgliedes. Ist nur I. erfüllt, liegen Name-ICCs vor, die teildeskriptiv sind, also eine Information dazu enthalten, welches Konzept im Kontext des Referenten eine Rolle spielt. Sind weder I. noch II. erfüllt, sind die Bildungen arbiträre Name-ICCs. Name-ICCs bilden generell keine Subkonzepte, sondern referieren direkt auf Entitäten und werden deshalb zu einer Klasse zusammengefasst. Problematisch an dieser Subklassifikation ist vor allem die Einordnung der Prot-ICC-Erstglieder, da bei Lexemen wie *Glas* oder *Rest* das Vorliegen eines nominalen Konzeptes nicht immer eindeutig bestimmt werden kann. Der Ansatz bietet aber viele Vorteile und kann die erhobenen Belege in den meisten Fällen klar einem ICC-Typ zuordnen. Im folgenden Kapitel werden die Daten nun hinsichtlich der formalen Eigenschaften dieser drei ICC-Subtypen analysiert. Anders als bei der Besprechung der Ergebnisse zu Semantik und Funktion von ICCs werden die formalen Aspekte der zwei Korpusstudien getrennt voneinander besprochen.

5 Ergebnisse zu formalen Aspekten von ICCs

5.1 DECOW16

5.1.1 Überblick über die Daten und die verwendeten statistischen Tests

Die als ICCs klassifizierten Daten verteilen sich nach der Subklassifikation wie folgt auf die drei ICC-Typen:

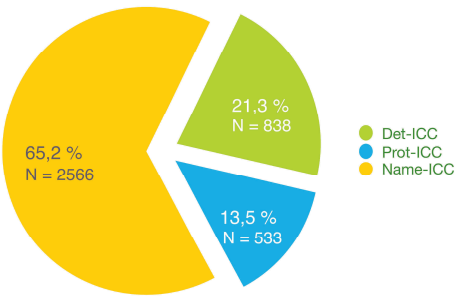

Abb. 15: Verteilung der ICC-Typen (DECOW16).

Den größten Anteil an den 3937 ICC-Belegen haben die Name-ICCs (N = 2566). Det-ICCs (N = 838) und Prot-ICCs (N = 533) machen zusammen nur etwas mehr als ein Drittel der ICC-Belege aus.

Die ICC-Belege verteilen sich auf 400 Stämme. 39% von 1034 der untersuchten Basislexeme treten also in ICCs auf. Im Korpus ließen sich zu 96 der 1034 untersuchten Stämme Det-ICC-Belege finden (9%), zu 175 Stämmen Prot-ICC-Belege (17%) und zu 277 Stämmen Name-ICC-Belege (27%). 432 Stämme (42%) traten nicht in ICCs, aber in anderen (meist syntaktischen) Verbindungen auf. Zu 202 Stämmen (20%) konnten überhaupt keine Belege ermittelt werden. Hierzu gehörten etwa die Stämme *Sojasauce* und *Wirbelsäule*.

Die Analyse der Daten aus DECOW16 untersucht zum einen die Eigenschaften der Basislexeme, die den ICC-Typen zugrundeliegen, zum anderen die Eigenschaften der ICC-Belege selbst. Für die Vergleiche der Basislexeme werden immer zwei Gruppenvergleiche durchgeführt: Erstens werden die Basen der drei ICC-Typen (Det-ICC, Prot-ICC, Name-ICC) miteinander verglichen, zweitens werden die Stämme, die in ICCs vorkommen (ICC-Basen), mit denen verglichen, die nicht in ICCs vorkommen (Nicht-ICC-Basen).

https://doi.org/10.1515/9783111315416-005

Die statistischen Tests, die die Vergleiche jeweils begleiten, richten sich nach dem Skalenniveau der jeweiligen Variablen. Zu Variablen, die in den Daten intervallskaliert sind, werden über die deskriptive Statistik hinaus Tests zur statistischen Signifikanz der Unterschiede sowie zur Effektstärke durchgeführt. Dies betrifft die die Basislexeme betreffenden Variablen FREQUENZKLASSE.DWDS (FK.DWDS), FREQUENZKLASSE.LEIPZIG (FK.LEIPZIG), ZEICHENANZAHL, SILBENANZAHL, MORPHEMANZAHL, ANZAHL.DORNSEIFF sowie die die Belege betreffende Variable BELEGLÄNGE. Bei den entsprechenden Vergleichen wird eine einfaktorielle ANOVA gerechnet, sofern drei Gruppen miteinander verglichen werden (Vergleich der ICC-Typen Det-ICC, Prot-ICC, Name-ICC). Bei den Vergleichen zwischen zwei Gruppen (ICC-Basen und Nicht-ICC-Basen) wird ein t-Test gerechnet. In den Fällen, in denen die Gruppen gemäß dem Shapiro-Wilk-Test nicht normalverteilt sind und die Voraussetzung der Varianzhomogenität verletzt ist, wird auf Welchs ANOVA, beziehungsweise Welchs t-Test zurückgegriffen. Zu Variablen, die in den Daten nominalskaliert sind, werden allein die Häufigkeiten berichtet. Dies betrifft die Variablen SCHREIBUNG, FUGENELEMENT, FLEXIONSMARKER, ARTIKEL, PRÄPOSITION, PRÄDIKATIV sowie ADJEKTIVISCHE MODIFIKATION. Hinsichtlich der Vergleiche zwischen den Merkmalen der Basislexeme wird zudem eine binomiale logistische Regression gerechnet, um den Einfluss der genannten Faktoren auf die ICC-Bildung einschätzen zu können und um zu gewährleisten, dass die Faktoren voneinander unabhängig sind. Für die Durchführung aller statistischen Tests sowie der binomialen logistischen Regression wurde IBM SPSS Statistics (Version 27.0.0.0) verwendet.

5.1.2 Analyse zu den Eigenschaften der Basislexeme

5.1.2.1 Frequenz

Für alle Gruppenvergleiche werden die Daten zu den Frequenzklassen im DWDS (Klein & Geyken 2010) und im Leipziger Wortschatz (Goldhahn 2012, Quasthoff & Richter 2005) zugrundegelegt. Tabelle 4 zeigt die deskriptive Statistik, die Teststatistik sowie die Gruppenvergleiche.

Tab. 4: Vergleich der mittleren Frequenzklasse der Basislexeme (DWDS und Leipziger Wortschatz)

Deskriptive Statistik						
Variable	**ICC-Typ**	**N**	***M***	***SD***	***SEM***	
FK.DWDS	Det-ICC-Basen	96	4,67	0,675	0,069	
	Prot-ICC-Basen	175	4,56	0,716	0,054	
	Name-ICC-Basen	27	4,25	0,855	0,051	
	ICC-Basen	400	4,32	0,818	0,041	
	Nicht-ICC-Basen	634	3,35	1,149	0,046	
FK.LEIPZIG	Det-ICCs	96	9,66	2,061	0,210	
	Prot-ICCs	175	9,95	2,252	0,170	
	Name-ICCs	277	11,17	2,722	0,164	
	ICC-Basen	400	10,89	2,625	0,131	
	Nicht-ICC-Basen	634	14,32	3,717	0,148	
ANOVA (Unterschied Det-/Prot-/Name-ICC-Basen)						
Variable	**Statistik**	***df1***	***df2***	***p***	η^2	
FK.DWDS	14,589	2	269,848	0,000	0,049	
FK.LEIPZIG	20,614	2	273,965	0,000	0,068	
t-Test (Unterschied ICC-Basen/Nicht-ICC-Basen)						
Variable	**T**	***df***	***p***	***Cohen's d***		
FK.DWDS	-15,715	1017,209	0,000	1,033		
FK.LEIPZIG	17,360	1019,061	0,000	3,338		
Games-Howell post-hoc Test (Gruppenvergleiche)						
Variable	**Vergleich**	**Differenz**	***SEM***	***p***	**95% Konfidenzintervall**	
					Untergrenze	**Obergrenze**
FK.DWDS	Prot-ICC–Det-ICC	-0,107	0,088	0,444	-0,31	0,10
	Name-ICC–Det-ICC	-0,418	0,086	0,000	-0,62	-0,21
	Name-ICC–Prot-ICC	-0,311	0,075	0,000	-0,49	-0,14
	ICC–Nicht_ICC	-0,963	0,061	0,000	-1,083	-0,842
FK.LEIPZIG	Prot-ICC–Det-ICC	0,292	0,271	0,527	-0,35	0,93
	Name-ICC–Det-ICC	1,510	0,266	0,000	0,88	2,14
	Name-ICC–Prot-ICC	1,217	0,236	0,000	0,66	1,77
	ICC–Nicht_ICC	3,430	0,198	0,000	3,042	3,817

Die Frequenz der Basislexeme, gemessen an der Frequenzklasse im DWDS und im Leipziger Wortschatz, unterscheidet sich zwischen den drei ICC-Typen. Die Det-ICC-Basen sind im Schnitt frequenter als die Prot-ICC-Basen, die wiederum frequenter als die Name-ICC-Basen sind. So zeigt Welchs einfaktorielle ANOVA basierend auf den Daten des DWDS einen signifikanten Unterschied zwischen der durchschnittlichen Frequenzklasse der ICC-Basen.[24] Die Effektstärke nach Cohen (1988) liegt bei $f = .22$ und entspricht einem kleinen Effekt. Der Games-Howell post-hoc Test zeigt hier allerdings nur für den Vergleich zwischen den Basen von Name-ICCs und Prot-ICCs und zwischen den Basen von Name-ICCs und Det-ICCs einen signifikanten Unterschied in der Frequenzklasse. Der Unterschied zwischen den Basen von Det-ICCs und Prot-ICCs ist hingegen nicht signifikant. Welchs einfaktorielle ANOVA zeigt zudem, dass es in den Daten vom Leipziger Wortschatz einen statistisch signifikanten Unterschied zwischen der durchschnittlichen Frequenzklasse von Det-, Prot- und Name-ICC-Basen gibt. Die Effektstärke nach Cohen (1988) liegt bei $f = .27$ und entspricht einem mittleren Effekt. Der Games-Howell post-hoc Test zeigt einen signifikanten Unterschied in der Frequenzklasse zwischen den Basen von Name-ICCs und Prot-ICCs sowie zwischen denen von Name-ICCs und Det-ICCs, aber wiederum nicht zwischen den Basen von Det-ICCs und Prot-ICCs. Die durchgeführten Tests und Vergleiche bedeuten, dass es zwischen den ICC-Typen kleine Unterschiede in der mittleren Frequenzklasse der Basen gibt. Name-ICC-Basen gehören, basierend auf beiden Frequenzdatenbanken, im Schnitt einer niedrigeren Frequenzklasse an als Det- und Prot-ICC-Basen.[25]

Ein weitaus größerer Unterschied in der durchschnittlichen Frequenzklasse liegt vor, wenn man die ICC-Basen mit den Nicht-ICC-Basen vergleicht. Welchs t-Test zeigt hier einen signifikanten Unterschied zwischen den beiden Gruppen, wobei die ICC-Basen im DWDS im Schnitt 0,96 Frequenzklassen höher eingeordnet sind als die Nicht-ICC-Basen, im Leipziger Wortschatz liegen sie 3,43 Frequenzklassen höher als die Basen ohne ICCs.[26] Die Effektstärke nach Cohen (1988) liegt bei $f = 1.67$ und entspricht einem großen Effekt. Die Basen, die in ICCs vorkommen, sind also im Schnitt deutlich frequenter als die Basen, zu denen keine ICCs belegt sind. Die Boxplots in Abbildung 16 visualisieren alle besprochenen Frequenzunterschiede.

---

24 Die Gruppen sind nicht normalverteilt (Shapiro-Wilk Test $p < .001$) und die Voraussetzung der Varianzhomogenität ist nicht gegeben (Levenetest $p < .001$), für beide Vergleiche wurde deshalb Welchs ANOVA gerechnet.

25 Es sei noch einmal daran erinnert, dass im Wortschatz Leipzig eine hohe Frequenzklasse für eine niedrigere Durchschnittsfrequenz steht. Dessen ungeachtet spreche ich bei den Vergleichen der besseren Verständlichkeit wegen stets von einer „niedrigeren Frequenzklasse".

26 Welchs t-Test wird gerechnet, weil die Gruppen nicht normalverteilt sind (Shapiro-Wilk Test $p < .001$) und die Voraussetzung der Varianzhomogenität verletzt ist (Levenetest $p < .001$).

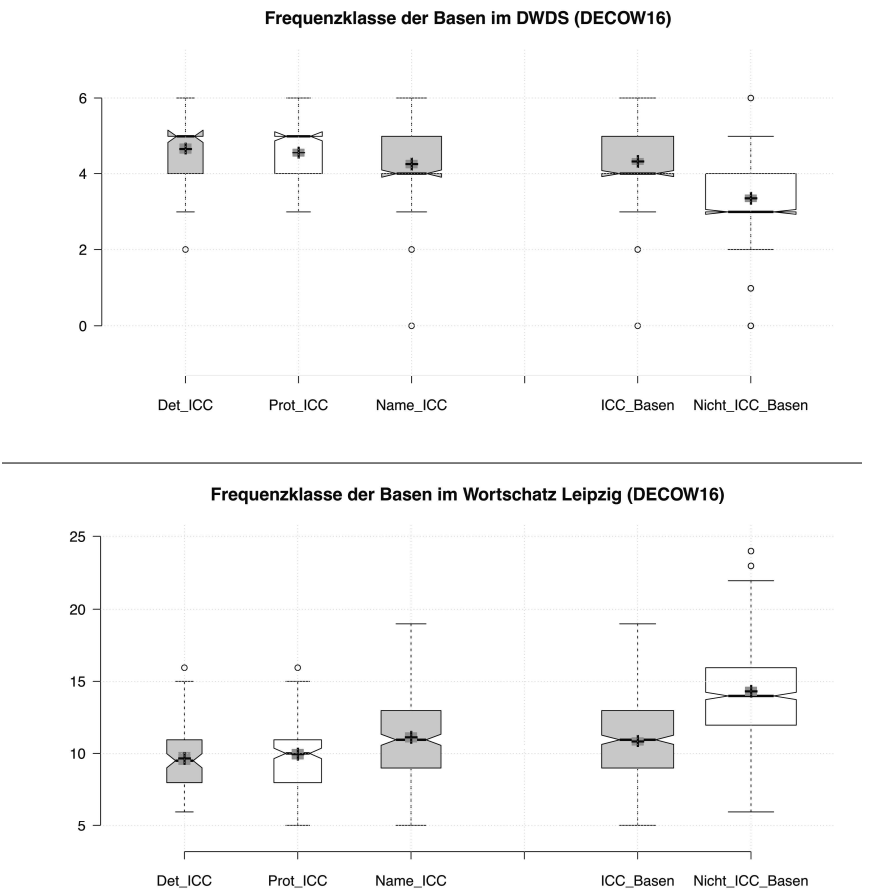

Abb. 16: Unterschiede zwischen den ICC-Typen in Bezug auf die durchschnittliche Frequenzklasse der Basislexeme (DWDS, Leipziger Wortschatz).[27]

5.1.2.2 Komplexität

Tabelle 5 zeigt die deskriptive Statistik, die Teststatistik sowie die Gruppenvergleiche für die Vergleiche zur Komplexität.

27 Für diese und die folgenden Abbildungen gilt: Die Boxen entsprechen dem Bereich, in dem die mittleren 50% der Daten liegen (oberes und unteres Quartil), die Breite der Boxen ist proportional zur Quadratwurzel der Beleganzahl. Die Antennen erstrecken sich über das 1,5-fache des Interquartilabstandes, Punkte stehen für Ausreißer, Mittellinien für die Mediane, Kreuze für die Mittelwerte. Graue Balken um die Kreuze und Kerben in den Boxen an der Stelle des Medians zeigen jeweils ein 95%-Konfidenzintervall des Mittelwerts/des Medians an (Chambers et al. 1983).

Tab. 5: Vergleich zur Komplexität gemessen an der Anzahl der Zeichen, Silben und Morpheme der Basislexeme

Deskriptive Statistik					
Variable	**ICC-Typ**	**N**	***M***	***SD***	***SEM***
ZEICHENANZAHL	Det-ICC-Basen	96	5,38	1,416	0,145
	Prot-ICC-Basen	175	5,31	1,519	0,115
	Name-ICC-Basen	277	5,23	1,614	0,097
	ICC-Basen	400	5,45	1,697	0,085
	Nicht-ICC-Basen	634	7,45	2,898	0,115
SILBENANZAHL	Det-ICC-Basen	96	1,58	0,574	0,059
	Prot-ICC-Basen	175	1,59	0,670	0,051
	Name-ICC-Basen	277	1,60	0,621	0,037
	ICC-Basen	400	1,65	0,644	0,032
	Nicht-ICC-Basen	634	2,36	0,899	0,036
MORPHEMANZAHL	Det-ICCs	96	1,14	0,344	0,035
	Prot-ICCs	175	1,11	0,453	0,034
	Name-ICCs	277	1,14	0,365	0,022
	ICC-Basen	400	1,15	0,420	0,021
	Nicht-ICC-Basen	634	1,56	0,716	0,028

ANOVA (Unterschied Det-/Prot-/Name-ICC-Basen)					
Variable	**Statistik**	***df1***	***df2***	***p***	η^2
ZEICHENANZAHL	0,368	2	260,653	0,692	–
SILBENANZAHL	0,041	2	255,092	0,960	–
MORPHEMANZAHL	0,166	2	251,584	0,847	–

t-Test (Unterschied ICC-Basen/Nicht-ICC-Basen)				
Variable	**T**	***df***	***p***	***Cohen's d***
ZEICHENANZAHL	14,040	1026,817	0,000	2,503
SILBENANZAHL	14,751	1015,690	0,000	0,810
MORPHEMANZAHL	11,627	1027,030	0,000	0,618

Games-Howell post-hoc Test (Gruppenvergleiche)						
Variable	**Vergleich**	**Differenz**	***SEM***	***p***	**95% Konfidenzintervall**	
					Untergrenze	**Obergrenze**
ZEICHENANZAHL	Prot-ICC–Det-ICC	-0,066	0,185	0,931	-0,50	0,37
	Name-ICC–Det-ICC	-0,144	0,174	0,687	-0,56	0,27
	Name-ICC–Prot-ICC	-0,078	0,150	0,864	-0,43	0,28
	ICC–Nicht_ICC	2,008	0,143	0,000	1,727	2,288
SILBENANZAHL	Prot-ICC–Det-ICC	0,011	0,077	0,989	-0,17	0,19
	Name-ICC–Det-ICC	0,020	0,069	0,957	-0,14	0,18
	Name-ICC–Prot-ICC	0,009	0,063	0,990	-0,14	0,16
	ICC–Nicht_ICC	0,709	0,048	0,000	0,615	0,803
MORPHEMANZAHL	Prot-ICC–Det-ICC	-0,021	0,049	0,903	-0,14	0,09
	Name-ICC–Det-ICC	0,002	0,041	0,999	-0,10	0,10
	Name-ICC–Prot-ICC	0,023	0,041	0,840	-0,07	0,12
	ICC–Nicht_ICC	0,411	0,035	0,000	0,342	0,480

Die Basislexeme der drei ICC-Typen unterscheiden sich nicht hinsichtlich ihrer Komplexität. Bei der durchschnittlichen Zeichen-, Silben- und Morphemanzahl der Basen zeigen sich zwar leichte Unterschiede. So sind Name-ICC-Basen im Hinblick auf die Zeichenanzahl im Schnitt etwas kürzer als Prot-ICC-Basen, die wiederum kürzer sind als Det-ICC-Basen. Eine einfaktorielle ANOVA zeigt aber, dass diese Unterschiede nicht signifikant sind.[28] Auch die Unterschiede hinsichtlich der durchschnittlichen Silben- und Morphemanzahl sind nicht signifikant.

Vergleicht man jedoch ICC-Basen und Nicht-ICC-Basen, gibt es Unterschiede hinsichtlich der Komplexität. Basen, zu denen keine ICCs gefunden wurden, weisen dabei eine höhere Komplexität auf als die ICC-Basen. Nicht-ICC-Basen bestehen im Schnitt aus zwei Zeichen mehr als ICC-Basen und haben 0,7 Silben und 0,4 Morpheme mehr. Auffällig ist auch, dass ICC-Basen in der Regel monomorphematisch sind. 87% der ICC-Basen sind monomorphematisch (Nicht-ICC-Basen: 56%); nur knapp 1% der ICC-Basen beinhaltet mehr als 2 Morpheme (Nicht-ICC-Basen: 11%). Alle Unterschiede, sowohl der Unterschied in der Zeichenanzahl, als auch der in der Silben- und Mor-

28 Bei den Vergleichen zur Komplexität ist die Voraussetzung der Varianzhomogenität in jedem Fall gegeben (Levenetest: $p < .468$ (Zeichenanzahl), $p < .358$ (Silben), $p < .640$ (Morpheme). Deshalb wurde eine ANOVA gerechnet.

phemanzahl sind signifikant. Die Effekte zur Zeichen- und Silbenanzahl können als stark gelten (f = 1.25, respektive f = .41); der zur Morphemanzahl ist ein mittlerer Effekt (f = .31). Abbildung 17 veranschaulicht diese Komplexitätsunterschiede der Basen.

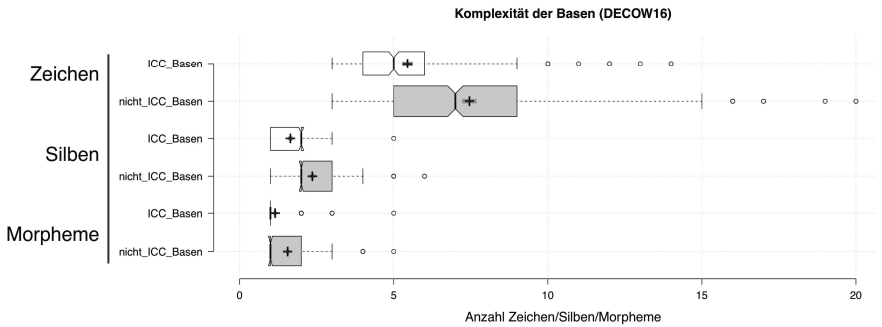

Abb. 17: Komplexitätsunterschiede zwischen ICC-Basen und Nicht-ICC-Basen.

5.1.2.3 Bedeutungsgruppen

Tabelle 6 zeigt die deskriptive Statistik, die Teststatistik sowie die Gruppenvergleiche für die Vergleiche zu der Anzahl der Dornseiff-Bedeutungsgruppen.

Tab. 6: Vergleich der Anzahl der Dornseiff-Bedeutungsgruppen

Deskriptive Statistik					
Variable	**ICC-Typ**	**N**	***M***	***SD***	***SEM***
ANZAHL-DORNSEIFF	Det-ICCs	96	4,31	2,754	0,281
	Prot-ICCs	175	3,33	2,171	0,164
	Name-ICCs	277	3,31	2,307	0,139
	ICC-Basen	400	3,34	2,319	0,116
	Nicht-ICC-Basen	634	2,04	2,016	0,080

ANOVA (Unterschied Det-/Prot-/Name-ICC-Basen)					
Variable	**Statistik**	***df1***	***df2***	***p***	η^2
ANZAHL-DORNSEIFF	5,481	2	236,862	<0,01	0,025

t-Test (Unterschied ICC-Basen/Nicht-ICC-Basen)				
Variable	**T**	***df***	***p***	***Cohen's d***
ANZAHL-DORNSEIFF	-9,207	761,269	0,000	2,139

Games-Howell post-hoc Test (Gruppenvergleiche)						
Variable	**Vergleich**	**Differenz**	***SEM***	***p***	**95% Konfidenzintervall**	
					Untergr.	**Obergr.**
ANZAHL-DORNSEIFF	Prot-ICC–Det-ICC	-0,987	0,326	0,008	-1,76	-0,22
	Name-ICC–Det-ICC	-1,002	0,313	0,005	-1,74	-0,26
	Name-ICC–Prot-ICC	-0,015	0,215	0,997	-0,52	0,49
	ICC–Nicht_ICC	-1,297	0,141	0,000	-1,574	-1,021

ICC-Basen unterscheiden sich hinsichtlich der Anzahl der Dornseiff-Bedeutungsgruppen, an denen sie teilhaben. Det-ICC-Basen haben an ungefähr einer Dornseiff-Bedeutungsgruppe mehr teil als Prot- und Name-ICC-Basen. Welchs einfaktorielle ANOVA zeigt, dass innerhalb der ICC-Basen ein signifikanter Unterschied besteht.[29] Die Effektstärke von f = .162 entspricht einem kleinen Effekt. Games-Howell post-hoc Tests zeigten einen signifikanten Unterschied in der Anzahl an Dornseiff-Bedeutungsgruppen zwischen Name-ICC-Basen und Det-ICC-Basen sowie zwischen Prot-ICC-Basen und Det-ICC-Basen. Der Unterschied zwischen Prot- und Name-ICC-Basen ist nicht signifikant.

ICC-Basen und Nicht-ICC-Basen unterscheiden sich hinsichtlich der Anzahl der Dornseiff-Bedeutungsgruppen deutlicher. ICC-Basen haben im Schnitt an 3,34 Bedeutungsgruppen teil, Nicht-ICC-Basen hingegen nur an 2,04. Dieser Unterschied ist signifikant und der Effekt stark (f = 1.07). Abbildung 18 zeigt die Unterschiede in Bezug auf die Dornseiff-Bedeutungsgruppen.

29 Auch bei den Vergleichen zu den Bedeutungsgruppen wurde Welchs einfaktorielle ANOVA gerechnet, weil die Voraussetzung der Varianzhomogenität verletzt wurde (Levenetest $p < .01$).

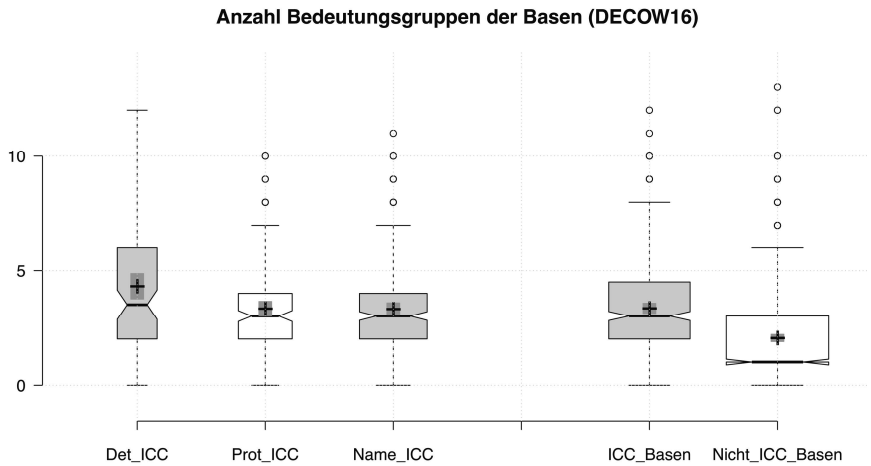

Abb. 18: Unterschiede in der Anzahl der Dornseiff-Bedeutungsgruppen von ICC-Basen und Nicht-ICC-Basen.

5.1.2.4 Binomiale logistische Regression zu den Basisvergleichen

ICC-Basen sind im Mittel frequenter, weniger komplex und an mehr Dornseiff-Bedeutungsgruppen beteiligt als Nicht-ICC-Basen. Nun könnten diese Unterschiede aber alle auf einen einzigen Faktor zurückzuführen sein. Eine solche Analyse wird in der linguistischen Forschung etwa hinsichtlich des Faktors Frequenz diskutiert. Der Einfluss, den Faktoren auf sprachliche Strukturen haben, könnte demnach bloß ein Epiphänomen des Faktors Frequenz sein (Haspelmath 2006, 2008). Hinsichtlich der ICC-Basen könnte man in ähnlicher Weise argumentieren, dass hochfrequente Nomina generell kürzer und weniger komplex sind als niedrigfrequente und wegen der häufigen Verwendung auch an mehr Bedeutungsgruppen teilhaben. Um zu gewährleisten, dass diese Unterschiede zwischen ICC-Basen und Nicht-ICC-Basen nicht letztlich alle auf ein und dasselbe Phänomen zurückgehen, um also auszuschließen, dass die beschriebenen Variablen alle dasselbe messen, wird zusätzlich zu den Gruppenvergleichen eine binomiale logistische Regression gerechnet. Mit der Regression werden die Faktoren, die für die Zugehörigkeit eines Lexems zu den Gruppen ICC-Basen und Nicht-ICC-Basen eine Rolle spielen, identifiziert und hinsichtlich ihrer Stärke spezifiziert.

In dem hier beschriebenen Modell werden fünf Variablen (ZEICHENANZAHL, SILBENANZAHL, ANZAHL.DORNSEIFF, FK.DWDS, FK.LEIPZIG) untersucht. Hinsichlich der Multikollinearität, also der Korrelation der Faktoren, stellt sich dabei heraus, dass einzig zwischen den Variablen FK.DWDS und FK.LEIPZIG eine hohe Korrelation

besteht (r = .84), was besagt, dass die beiden Variablen dasselbe messen. Bei der Frequenz der Basislexeme ist dies ja aber auch erwartbar und belegt bloß die Validität der beiden Frequenzdatenbanken DWDS und Leipziger Wortschatz. Die Korrelationen zwischen den übrigen Prädiktoren sind gering (r < .70). In der statistischen Literatur gehen diesbezüglich selbst die konservativsten Autoren erst bei Werten ab .80 von Multikollinearität aus (etwa: Abu-Bader & Pryce 2006, Dattalo 2013). Dass die Korrelationen der Variablen allesamt den Wert .70 nicht überschreiten, deutet also darauf hin, dass Multikollinearität die Analyse nicht konfundiert. Von den beiden Frequenzvariablen wurde wegen der Multikollinearität FK.DWDS aus dem Modell genommen.

Das binomial logistische Regressionsmodell ist statistisch signifikant, $\chi^2(5) = 350.07$, $p < .001$ mit einer Varianzaufklärung von Nagelkerkes $R^2 = .390$ (Nagelkerke 1991). Gemäß den Empfehlungen von Backhaus et al. (2015) entspricht dies einer akzeptablen bis guten Varianzaufklärung. Die Anpassungsgüte wurde mit dem Hosmer-Lemeshow-Test überprüft, der eine hohe Anpassungsgüte zeigt, $\chi^2(8) = 16.14$, $p > .05$ (Hosmer et al. 2013). Der Gesamtprozentsatz korrekter Klassifikation ist 73.6%, mit einer Sensitivität von 62.0% und einer Spezifität von 80.9%. Von den vier Variablen, die in das Modell aufgenommen wurden, sind drei signifikant, nämlich ZEICHENANZAHL ($p < .01$), SILBENANZAHL ($p < .001$), sowie FK.LEIPZIG ($p < .001$). ANZAHL.DORNSEIFF ($p = .336$) hat hingegen keinen signifikanten Einfluss auf die prädiktive Leistung des Modells. ZEICHENANZAHL hemmt das Vorkommen einer Basis in der ICC-Konstruktion mit einem Odds von 0.835 (95%-KI[0.746, 0.935]), genauso wie SILBENANZAHL, mit einem Odds von 0.562 (95%-KI[0.412, 0.766]) und FK.LEIPZIG mit einem Odds von 0.740 (95%-KI[0.697, 0.785]). Alle Modellkoeffizienten und Odds können Tabelle 7 entnommen werden.

Tab. 7: Binomiale logistische Regression zu den Vergleichen der Basislexeme

Binomiale logistische Regression								
Regressor	**B**	**SE**	***Wald***	***df***	***p***	***Odds***	**95% Konfidenzintervall**	
							Untergrenze	**Obergrenze**
ZEICHENANZAHL	-0,180	0,058	9,760	1	0,002	0,835	0,746	0,935
SILBENANZAHL	-0,576	0,158	13,252	1	0,000	0,562	0,412	0,766
ANZAHL.DORNSEIFF	0,036	0,036	0,957	1	0,328	1,036	0,965	1,113
FK.LEIPZIG	-0,301	0,030	97,613	1	0,000	0,740	0,697	0,785

Die Variablen, die die ICC-Bildung eines Basislexems bedingen, üben den Einfluss also jeweils unabhängig von den anderen Faktoren aus. Es gibt formale und die Frequenz betreffende Besonderheiten der Lexeme, die ICCs bilden. ICC-Basen sind im Schnitt deutlich frequenter als Nicht-ICC-Basen und bestehen aus weniger Zeichen und Silben als Nicht-ICC-Basen. Auch unter den ICC-Typen konnten Unterschiede nachgewiesen werden. Unter den ICC-bildenden Basen sind Det- und Prot-ICC-Basen im Schnitt frequenter als Name-ICC-Basen. Die drei ICC-Typen unterscheiden sich allerdings nicht hinsichtlich der Komplexität ihrer Basen oder der Anzahl der Dornseiff-Bedeutungsgruppen. Im Folgenden geht es nun um die Unterschiede und Merkmale der ICC-Belege.

5.1.3 Analyse zu den Eigenschaften der ICC-Belege

5.1.3.1 Länge

Tabelle 8 zeigt die deskriptive Statistik, die Teststatistik sowie die Gruppenvergleiche für die Vergleiche zur Beleglänge.

Tab. 8: Vergleiche zur Beleglänge (DECOW16)

Deskriptive Statistik						
Variable	**ICC-Typ**	**N**	***M***	***SD***	***SEM***	
BELEGLÄNGE	Det-ICCs	838	13,63	2,868	0,099	
	Prot-ICCs	533	11,44	2,785	0,121	
	Name-ICCs	2566	10,08	3,291	0,065	
ANOVA (Unterschied Det-/Prot-/Name-ICC-Basen)						
Variable	**Statistik**	***df1***	***df2***	***p***	η^2	
BELEGLÄNGE	452,344	2	1307,540	0,000	0,173	
Games-Howell post-hoc Test (Gruppenvergleiche)						
Variable	**Vergleich**	**Differenz**	***SEM***	***p***	**95% Konfidenzintervall**	
					Untergr.	**Obergr.**
BELEGLÄNGE	Prot-ICC–Det-ICC	-2,185	0,156	0,000	-2,55	-1,82
	Name-ICC–Det-ICC	-3,554	0,118	0,000	-3,83	-3,28
	Name-ICC–Prot-ICC	-1,369	0,137	0,000	-1,69	-1,05

Die ICC-Belege der drei ICC-Typen unterscheiden sich deutlich in der Zeichenanzahl. Die Wortlänge, gemessen an der Zeichenanzahl der Belege,[30] unterscheidet sich für die verschiedenen Bedingungen Det-ICC, Prot-ICC und Name-ICC (Abbildung 19). Welchs einfaktorielle ANOVA zeigt, dass es einen statistisch signifikanten Unterschied zwischen der Beleglänge von Det-, Prot- und Name-ICCs gibt.[31] Die Effektstärke nach Cohen (1988) liegt bei $f = .46$ und entspricht einem starken Effekt. Der Games-Howell post-hoc Test zeigte einen signifikanten Unterschied in der Wortlänge zwischen allen Gruppen. Name-ICC-Belege sind im Schnitt knapp 1,4 Zeichen kürzer als Prot-ICC-Belege, die wiederum 2,2 Zeichen kürzer als Det-ICCs sind. Somit sind Det-ICC-Belege knapp 3,6 Zeichen länger als Name-ICC-Belege. Abbildung 19 visualisiert die Längenunterschiede.

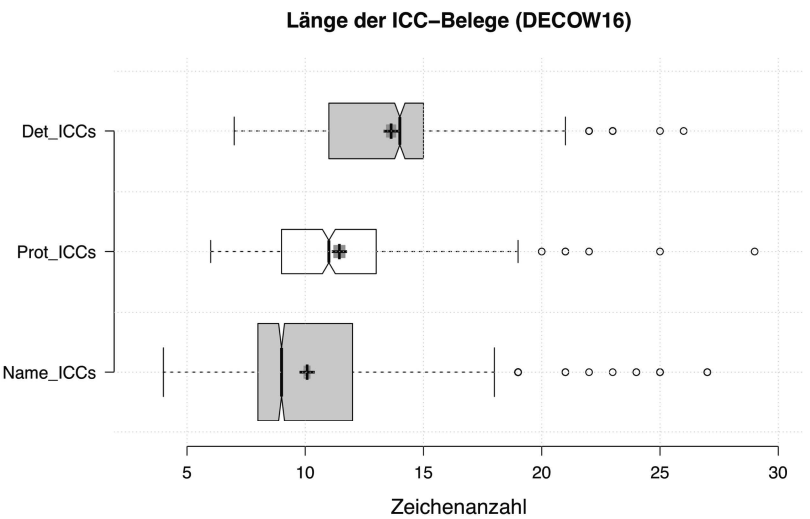

Abb. 19: Unterschiede zwischen den ICC-Typen in Bezug auf die durchschnittliche Zeichenanzahl der Belege (DECOW16).

30 Das Spatium wurde bei der Erfassung der Zeichenanzahl als Zeichen gewertet. Für die noch folgenden Vergleiche ist das relevant, da der Längenunterschied zwischen Name-ICCs, die den größten Anteil an der Spatienschreibung haben, und den anderen ICC-Typen ansonsten noch größer ausgefallen wäre.

31 In allen Vergleichen zur Wortlänge wurde Welchs einfaktorielle ANOVA gerechnet, da die Gruppen gemäß dem Shapiro-Wilk Test nicht normalverteilt ($p < .001$) waren und die Voraussetzung der Varianzhomogenität verletzt wurde (Levenetest $p < .001$).

5.1.3.2 Schreibung

ICCs weichen in Bezug auf die Schreibung von der normgerechten Zusammenschreibung ab. Die Konstituenten sind häufig durch den Bindestrich, die Binnengroßschreibung sowie das Spatium grafisch getrennt. Hierbei unterscheiden sich die drei ICC-Typen deutlich. Während bei den Det-ICCs die Zusammenschreibung überwiegt, weisen Prot- und Name-ICCs hier im Vergleich einen größeren Anteil der Bindestrichschreibung (Prot-ICCs) und Spatienschreibung (Name-ICCs) auf. Zudem wird bei Prot- und Name-ICCs die Binnengroßschreibung zur Trennung der Konstituenten eingesetzt. All das führt dazu, dass Prot-ICCs nur in etwas mehr als einem Drittel, Name-ICCs sogar in deutlich weniger als einem Drittel der Fälle zusammengeschrieben werden (Abbildung 20).

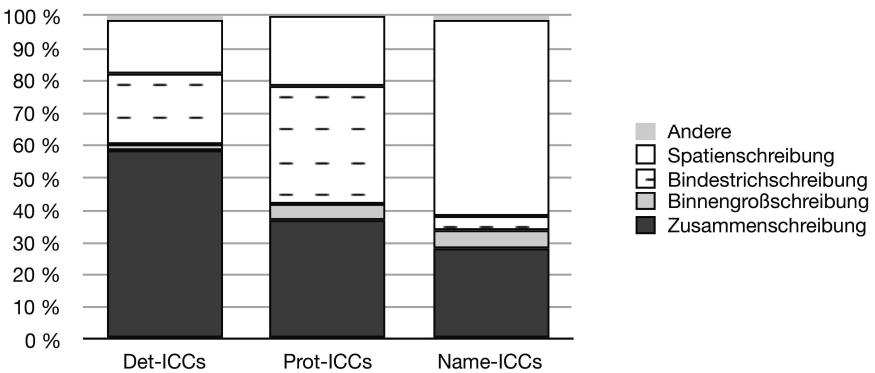

Abb. 20: Schreibung von Det-, Prot- und Name-ICCs (DECOW16).

Eine weitere Besonderheit von ICCs hinsichtlich der Schreibung ist die häufige Verwendung von doppelten Anführungszeichen. Im DECOW16 werden Nomina im Schnitt nur in 0,7% der Fälle in doppelte Anführungszeichen gesetzt. ICCs umgeben hingegen in knapp 10% der Fälle doppelten Anführungszeichen. Hier zeigt sich außerdem ein Unterschied zwischen den ICC-Typen. Det-ICCs erscheinen zu 6,8%, Prot-ICCs zu 8,3% und Name-ICC sogar zu 11% in doppelten Anführungszeichen.

5.1.3.3 Fugenelemente

Auch in Bezug auf die Fugenelemente verhalten sich die ICC-Typen unterschiedlich. ICCs im Allgemeinen tragen ähnlich häufig Fugenelemente am Erstglied wie kanonische N+N-Komposita, nämlich in 29% der Fälle. Bei kanonischen N+N-

Komposita nimmt Wellmann (1975) 26,5% Verfugung an, Nübling und Szczepaniak (2009) gehen von 35% aus. Interessanterweise unterscheiden sich die drei ICC-Typen hinsichtlich der Verfugung erheblich voneinander. Abbildung 21 visualisiert die Verteilung der (wie in 3.3.1.2 beschrieben klassifizierten) Marker an der ersten (N1) und zweiten (N2) Konstituente für die drei ICC-Typen.

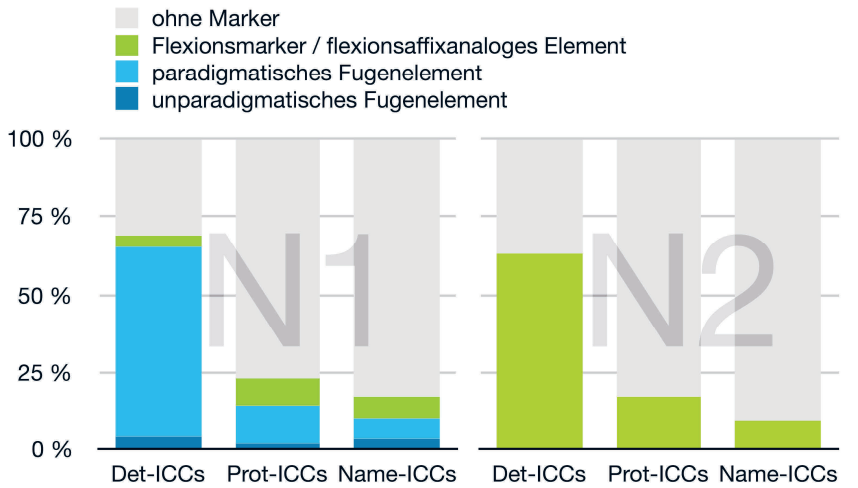

Abb. 21: Verteilung der Marker an Erst- und Zweitglied der drei ICC-Typen (DECOW16).

Det-ICCs weisen in nahezu Dreiviertel der Fälle ein Element am Erstglied auf. Prot-ICCs hingegen weisen nur in 23% der Fälle ein Element an N1 auf, Name-ICCs nur in 17%. Bei Prot-ICCs sind das in 15%, bei Name-ICCs nur in 10% der Fälle paradigmatische Fugenelemente oder unparadigmatische Fugenelemente.

Vor allem überrascht hier, dass Prot-ICCs überhaupt nennenswert Elemente am Erstglied tragen, wird ihnen diese Eigenschaft doch meist abgesprochen. Nach dem oben beschrieben Ansatz, Marker am Erstglied in paradigmatische und unparadigmatische Fugenelemente einzuteilen, treten bei Prot-ICCs in 2% der Fälle am Erstglied unparadigmatische Fugenelemente auf (*Arbeitsarbeit*). In 13% der Fälle treten Elemente auf, die als paradigmatische Fugenelemente gelten können (*Katzenkatze*). In 9% der Fälle zeigen die Erstglieder der Prot-ICCs Elemente, die sowohl Fugenelemente als auch Flexionsmarker sein könnten und als flexionsaffixanaloge Elemente angesehen werden können (*Quellen-Quellen*).

5.1.3.4 Flexionsmarker

Auch Flexionsmarker am Zweitglied sind mit 17%, respektive 9% bei den Prot- und Name-ICCs viel seltener als bei Det-ICCs, die zu 62% einen Flexionsmarker am Zweitglied tragen (Abbildung 21). Dies ist Evidenz dafür, dass Prot-ICCs und Name-ICCs generell seltener in die obliquen Kasus oder in den Plural gesetzt werden. Im Plural stehen sie etwa nur zu 18% (Prot-ICCs), beziehungsweise 7% (Name-ICCs); Det-ICCs hingegen in 62% der Fälle.

5.1.3.5 Syntax und Kontext

Die ICC-Typen verhalten sich auch syntaktisch unterschiedlich. Unterschiede zeigen sich hier bei der adjektivischen Modifikation, dem Artikelgebrauch und der Verwendung von Präpositionen.

Zunächst zum Artikelgebrauch. Abbildung 22 zeigt die Anteile, zu denen die drei ICC-Typen im Korpus mit vorangehenden definiten und indefiniten Artikeln auftreten. Im Diagramm ebenfalls aufgeführt ist die Verteilung, zu der Nomina im Korpus im Allgemeinen mit vorangehenden definiten und indefiniten Artikeln stehen.

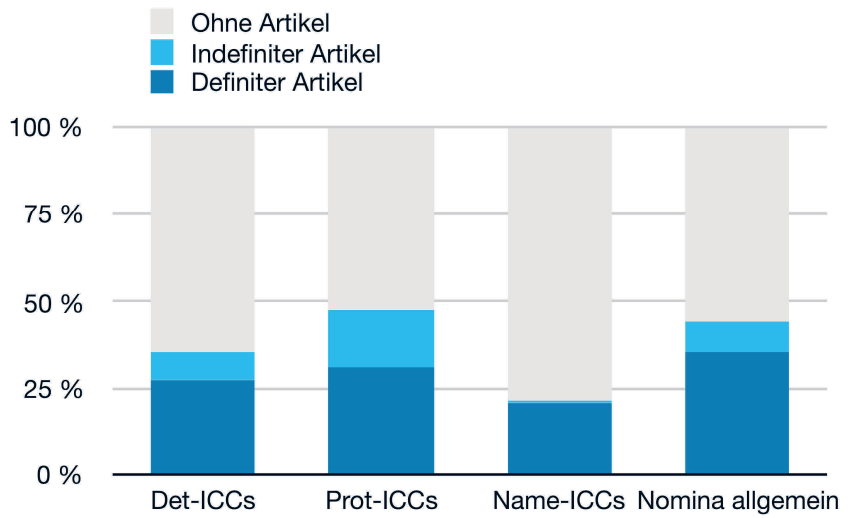

Abb. 22: Definite und indefinite Artikel an ICCs und Nomina im Allgemeinen (DECOW16).

ICCs werden im Schnitt seltener mit Artikeln verwendet als Nomina im Allgemeinen. Allerdings unterscheiden sich auch hier die ICC-Typen untereinander deut-

lich. Prot-ICCs übersteigen den Wert, den Nomina im Korpus im Allgemeinen erzielen. Name-ICCs unterschreiten den Wert zum prozentualen Artikelgebrauch von Nomina im Allgemeinen besonders stark. Nur in einem Fünftel der Fälle treten Name-ICCs mit einem Artikel auf. Das Verhältnis von indefinitem zu definitem Artikelgebrauch ist bei ICCs etwas höher als bei Nomina im Allgemeinen, nämlich 1:3 (1:4 bei Nomina im Allgemeinen). Aber auch hier gibt es wieder große Unterschiede zwischen den ICC-Typen: Prot-ICCs werden ungewöhnlich häufig mit indefinitem Artikel verwendet, Name-ICCs hingegen so gut wie überhaupt nicht.

In manchen Fällen, in denen Name-ICCs einen Artikel bei sich haben, weicht das durch den Artikel ausgedrückte Genus, und damit das Genus des Name-ICCs, vom Genus des Basislexems ab (70–72):

(70) *Darauf folgte das "**Hundehund**". Ein geniales Lied.*

<www.deutsche-mugge.de/index.php/live-berichte/2010/2459-stumpen-liest-in-berlin.html>

(71) *Zu den größeren Clubs zählen das **Tiger Tiger**, das Liquid und das Digital.*

<https://bw.fh-muenchen.de/internationales/partnerhochschulen/erfahrungsberichte_2/ws09 10>

(72) *Der Einschlag mit der linken Hüfte am Asphalt war die Hölle und ließ mich nur sehr verzögert wieder aufstehen um nach dem verdrehten Lenker meiner **GasGas** zu greifen und weiter zu fahren.*

<www.erzbergrodeo.at/leiwand/szene?artikel_kategorie=&artikel_id-=121154938523 4>

Das morpholexikalische Genus der Basis konfligiert hier mit dem Klassengenus, also der Klasse, auf das das Name-ICC verweist. Nach dem Letztgliedprinzip der Genuszuweisung müssten *Hundehund* und *Tiger Tiger* eigentlich das morpholexikalische Genus der letzten Konstituente des Kompositums erhalten, also Maskulina sein (Köpcke & Zubin 2009: 139). Das Genus wird hier aber offensichtlich, wie bei Eigennamen im Allgemeinen, referenziell vergeben, also in Abhängigkeit von der Objektklasse (Referenzdomäne), auf die die Bildung verweist (Fahlbusch & Nübling 2014, Köpcke & Zubin 2009: 144). Bei *Hundehund* (Klasse der Lieder) und *Tiger Tiger* (Klasse der Nachtclubs / Restaurants) ist das das Neutrum. Auch bei *GasGas* zeigt der Determinierer nicht das morpholexikalische Genus Neutrum, sondern das Femininum an (Klasse der Motorräder). Im Fall von Det- und Prot-ICCs entspricht das Genus dagegen immer dem der Basis.

Meist wird der Konflikt zwischen dem morpholexikalischen Genus der Basis und dem Klassengenus bei Name-ICCs aber entschärft, indem die Objektklasse explizit genannt und dem ICC vorangestellt (73) oder als Kompositumsletztglied angehängt (74) wird, sodass der Artikel dann mit diesem hinsichtlich Genus kongruiert:

(73) *Das Theaterstück "**Hase Hase**" ist eine skurrile Begegnung mit brisanten Themen.*

<https://www.waldorfschule-oberursel.de/index.php?tg=articles%26idx=More-%26topics=147%26article=329%26htf=Kom%C3%B6die_mit_viel_Tiefgang.html>

(74) *Califorlia war hungrig und durstig und auch musste sie bei ihrem **GasGas**-Motorrad die Kette nachspannen, ihren Luftfilter reinigen und Benzin nachfüllen.*

<https://wwwjessys-aprilia.ch/index.php?option=com_content&view=article&id-=139%3A13&catid=37%3Aadvent2012&Itemid=84>

Die Beobachtung, dass das Genus von Name-ICCs von dem der Basis abweicht, lässt sich daher nur schwer quantifizieren. Name-ICCs treten nur selten mit einem Artikel auf, an dem man das Genus des ICCs erkennen könnte. Unter diesen Fällen sind zudem nur wenige, in denen das Genus des Name-ICCs von der Basis abweicht. Es ist also nicht möglich, hierin evidenzbasiert ein Merkmal von Name-ICCs zu sehen.

Auch im Hinblick auf die Verwendung von Präpositionen unterscheiden sich die ICC-Typen.[32] Abbildung 23 gibt einen Überblick über die 15 Präpositionen, die mit ICCs am häufigsten eine Präpositionalphrase bilden.

32 Die syntaktische Kategorie von *als* und *wie* ist umstritten (für einen Überblick siehe Bliß 2017). Die beiden Wörter gelten meist nicht als Präpositionen im engeren Sinne, sondern als Adjunktoren. Zählt man sie hingegen zu den Präpositionen, muss man annehmen, dass Präpositionen den Nominativ regieren können. Der Einfachheit halber werden *als* und *wie* aber im Zuge dieser Ergebnisbesprechung zu den Präpositionen gezählt. Streng genommen ist hier aber von Präpositional- und Adjunktphrasen einleitenden Elementen die Rede.

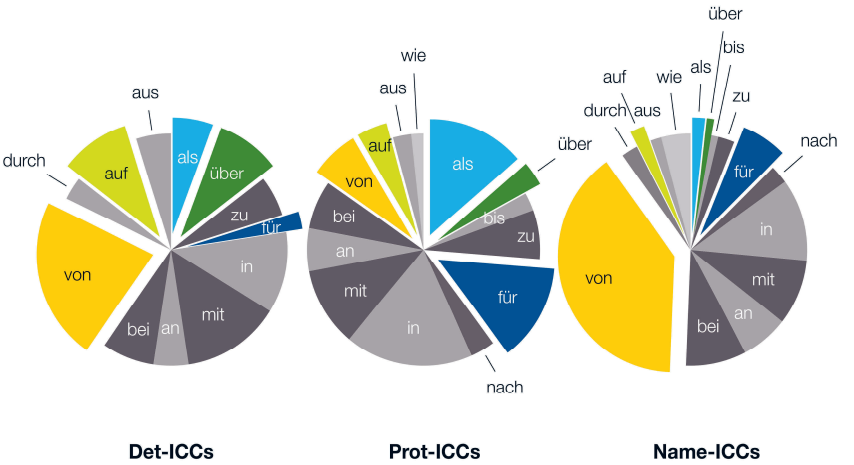

Abb. 23: Verbindung von ICCs mit den 15 häufigsten Präpositionen.

In den Phrasen mancher Präpositionen kommen alle ICC-Typen ähnlich häufig vor (grau gefärbte Flächen in Abbildung 23); andere Präpositionen verbinden sich mit den ICC-Typen unterschiedlich häufig (farbliche Flächen in Abbildung 23). So wird *von* vor allem in Name-ICCs verwendet, für die Verwendung mit *als* und *für* sind hingegen Prot-ICCs prädestiniert. Die Präpositionen *auf* und *über* finden prozentual am häufigsten mit Det-ICCs Verwendung. Eine Besonderheit von Name-ICCs ist, dass vor allem *von* und *in* auftreten und abseits davon besonders viele Präpositionen keine / nur eine marginale Rolle spielen (*durch*, *auf*, *aus*, *als*, *über*, *bis*, *zu* und *nach*).

Eine syntaktisch-funktionale Besonderheit der Prot-ICCs ist, dass sie doppelt so häufig wie Det- und Name-ICCs in der Funktion des Prädikativs verwendet werden. Für diesen Vergleich wurde ein regulärer Ausdruck verwendet, der jegliche Konstruktion mit den Kopula *sein*, *bleiben* und *werden* erfasst, in denen die ICC-Belege in prädikativer Funktion vorkommen (fakultativ mit Determinierer und Adjektiven, Abbildung 24).[33] Während Det-ICCs nur zu 6% Prädikative sind und Name-ICCs nur zu 5%, sind es Prot-ICC-Belege in 12% der Fälle. Dieses Ergebnis passt zu der Beobachtung, dass Prot-ICCs ungewöhnlich häufig mit indefinitem Artikel verwendet werden.

33 [lemma="sein|bleiben|werden"][]{0,2}[tag="ART.*"]?[tag="ADJ.*"]{0,2}[ascii="ICC-Typ"].

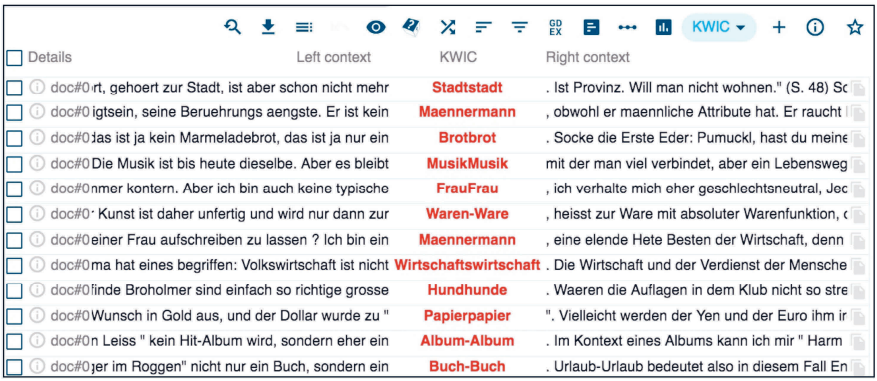

Abb. 24: Prot-ICCs in prädikativer Funktion.

Einen Unterschied zwischen den ICC-Typen markiert ferner die Kookkurrenz der Negationswörter *kein* und *nicht*. Bei Nomina im Allgemeinen finden sie sich in 0,7% der Fälle in einem Textfenster von 3 Wörtern links und rechts vom KWIK. Det- und Name-ICCs liegen hier mit 2%, respektive 3% etwas höher. Bei Prot-ICCs finden sich diese Negationswörter dagegen in 10% der Fälle.

Zudem gibt es Unterschiede zwischen den ICC-Typen bei der adjektivischen Modifikation. Tabelle 9 gibt an, wie häufig die ICC-Typen von Adjektiven modifiziert werden, sowie wichtige adjektivische Lemmata dieser Modifikation.

Tab. 9: Adjektivische Modifikation der ICCs (DECOW16)

Substantivtyp	Fälle mit attributivem Adjektiv	Häufigste Lemmata
Nomina	575.157.018 (21,5%)	–
ICCs	269 (6,8%)	–
Det-ICCs	68 (8,1%)	*sogenannt* (9%)
Prot-ICCs	107 (20,1%)	*richtig* (13%), *rein* (5%), *wirklich* (5%)
Name-ICCs	94 (3,7%)	*lieb* (13%)

Auch hier stechen wieder die Prot-ICCs heraus. Ihnen geht in mehr als 20% aller Fälle ein attributives Adjektiv voran. Das legt nahe, dass Prot-ICCs häufiger als Det- und Name-ICCs mithilfe von Attributen näher beschrieben werden und entspricht der in der Literatur besprochenen Kontextabhängigkeit der Prot-ICCs. Der Kotext der Prot-

ICCs wird daher zusätzlich zu den bisher beschriebenen quantitativen Methoden qualitativ hinsichtlich der Verankerung von Prot-ICCs im Kontext untersucht.

Der Blick auf den erweiterten Kotext der Prot-ICCs lässt den Schluss zu, dass die Sprecher:innen ihre ICC-Verwendung als Gelegenheitsbildung begreifen und darum kontextuell vor- und nachbereiten. In 76% der Fälle findet sich im Kontext eine der vier in Tabelle 10 aufgeführten Strategien, den Adressat:innen die Interpretation zu erleichtern (kontextuelle Anreicherung). In vielen Fällen werden diese Strategien kombiniert. Tabelle 10 zeigt Beispiele und Häufigkeiten zu den vier Strategien.

Tab. 10: Strategien der kontextuellen Anreicherung von Prot-ICCs

Strategie	Beispiel	N	%
Kontrastierende NP	*Calcium-Sand oder auch einfach nur "Sand Sand"*	221	41 %
Erklärung	*Kinder-Kinder, also die Altersstufe zwischen Kleinkind und Teenager*	170	32 %
Fokuspartikeln, Negationswörter	*keine Frau-Frau* *nur ein Buch-Buch*	86	16 %
Anreichernde attributive Adjektive	*ein selbstverliebtes Mädchen-Mädchen* *eine typische Frau-Frau*	107	20 %

In den meisten Fällen findet die kontextuelle Anreicherung über NPs statt, die einen Kontrast zum Prot-ICC markieren. Sie referieren auf eine Entität, die eben gerade nicht das ist, worauf das Prot-ICC referiert. Meist, nämlich zu 59%, geht die NP dem ICC voran (75); mitunter, nämlich zu 34%, folgt die NP dem ICC (76). In manchen Fällen (7%) wird sowohl vor der Verwendung des Prot-ICCs als auch danach eine kontrastierende NP verwendet (77).

(75) *ich hab einen Freund. Also... nicht jetzt einen **Freund Freund**, aber er ist doch ... ganz nah.*

<http://www.doctorsdiaryfanforum.de/t1808f215-Story-von-Fran-5.html>

(76) *Es war kein **Ende-Ende** sondern ein Neubeginn-Ende.*

<https://www.elbenwald.de/spot/news/magische-aktion-zum-harry-potter-zum-kinostart>

(77) *Dafür bin ich ein zu großer Trampel und Rowdie (okay, war Madita auch, aber sie hatte den großen Vorteil, auszusehen wie ein **Mädchenmädchen** und nicht gerade wie die Läuse-Mia.*

<http://www.thepandafck.de/page/19>

In 32% aller Prot-ICC-Belege erklären die Sprecher:innen in Form von Hauptsätzen, eingebetteten Sätzen oder komplexen NPs mit attribuierten Adjektiven, was das Prot-ICC bezeichnet:

(78) *Die Location war der absolute Knaller, nämlich auf der Cap San Diego, diesem wunderbaren "**Schiff Schiff**" Als "Schiff Schiff" bezeichne ich Wassergefährte, die einfach so typisch nach Schiff aussehen, wie es eben nur geht. ein Schiff, wie ein Kind es malt. Ein Schiff wie es schiffiger nciht aussehen kann.*

<http://www.doctorsdiaryfanforum.de/t1808f215-Story-von-Fran-5.html>

In der Regel (zu 65%) folgt diese Erklärung, die oft mit *also* oder *ich meine* eingeleitet ist, dem Prot-ICC, seltener (in 28%) geht sie ihm voran, in sehr seltenen Fällen (7%) wird sowohl vor als auch nach dem ICC eine Erklärung des Gemeinten gegeben.

Fokuspartikeln und Negationswörter finden sich in 16% der Fälle, anreichernde Adjektive, die in attributiver Funktion zum Prot-ICC stehen, in 20%. Dies ist ein weiteres interessantes Ergebnis, da die Verwendung von attributiven Adjektiven bei Prot-ICCs umstritten ist. Bross und Fraser nehmen an, dass adjektivische Attribution bei Prot-ICCs nicht möglich ist, mit Ausnahme von Real Intensifications wie *richtig* oder *echt* (Bross & Fraser 2020, siehe Kapitel 2.2.6). In der Hälfte der Fälle, in denen Prot-ICCs adjektivische Attribute haben, sind diese Adjektive auch tatsächlich Real Intensifications. Folgende Adjektive werden hier als Real Intensifications angesehen: (*wasch-*)*echt*, *eigentlich*, (*stink-*)*normal*, *ordentlich*, *rein*, *richtig*, *total*, *typisch*, *üblich*, *ultimativ* und *wirklich*. Diese machen aber nur die Hälfte der Fälle aus. Entgegen Bross und Frasers Annahme werden Prot-ICCs genauso oft attributiv durch gewöhnliche Adjektive modifiziert wie durch Real Intensifications.

5.1.4 Zusammenfassung der Ergebnisse (DECOW16)

Die aus dem DECOW16 erhobenen Daten liefern Evidenz zu formalen, semantischen und funktionalen Aspekten von ICCs. Zum einen zeigen sich Unterschiede zwischen den Basen, die den ICCs zugrundeliegen, zum anderen zwischen den ICC-Belegen selbst.

Die Eigenschaft eines Nominalstammes, ICCs zu bilden, korreliert mit konkreten formalen Merkmalen dieses Stammes. Der Faktor Frequenz hat hier offenbar einen Einfluss. Die in ICCs verwendeten Stämme sind knapp 3,5 Häufigkeitsklassen im Leipziger Wortschatz und knapp eine Frequenzklasse im DWDS höher angesiedelt als die Stämme, zu denen es keine ICCs gibt. Auch die Komplexität der Stämme, gemessen an Graphem-, Silben- und Morphemanzahl, hat offenbar einen Einfluss darauf, ob die Sprecher:innen mit dem entsprechenden Stamm ein ICC bilden. Vor allem kurze und morphologisch einfache Stämme werden in der Konstruktion ICC verwendet, was sich vornehmlich darin zeigt, dass ICC-Basen im Schnitt mehr als zwei Zeichen kürzer sind als Nicht-ICC-Basen. Der semantische Faktor, der mithilfe der Dornseiff-Bedeutungsgruppen erfasst wurde, scheint hingegen keine Rolle bei der ICC-Bildung zu spielen. ICC-Basen haben zwar an etwas mehr Bedeutungsgruppen teil als Nicht-ICC-Basen und dieser Unterschied ist auch signifikant. Doch zeigt die binomiale logistische Regression, dass dieser Faktor kein Prädiktor dafür ist, ob ein Nominalstamm ein ICC bildet oder nicht. Die Unterschiede zwischen den jeweiligen ICC-Typen sind außerdem nicht sehr groß. Die Basen, die die drei ICC-Typen verwenden, sind formal ungefähr gleich komplex. Es gibt nur einen tatsächlich gemessenen Unterschied zwischen den ICC-Basen: Name-ICCs werden auch mit etwas niedrigfrequenteren Basen gebildet als Det- und Prot-ICCs.

Auch bei den Belegen zeigen sich Unterschiede. Die drei ICC-Typen unterscheiden sich hinsichtlich ihrer Länge, Schreibung, Morphologie und Syntax. Det-ICCs sind im Schnitt fast 4 Zeichen länger als Name-ICCs. Det-ICCs werden zudem doppelt so häufig zusammengeschrieben wie Name-ICCs, deutlich seltener in Anführungszeichen gesetzt und weisen mehr als sechsmal so häufig Fugenelemente und Flexionsmarker auf. Interessanterweise nehmen Prot-ICCs hier in allen Vergleichskriterien eine Zwischenposition ein.

Hinsichtlich der Syntax stechen vor allem Name-ICCs heraus. Während Det- und Prot-ICCs annähernd so häufig wie Nomina im Allgemeinen mit Artikeln verwendet werden, treten Name-ICCs nur halb so häufig mit Artikel auf. Indefinite Artikel fehlen bei ihnen fast gänzlich. Auch bei der Einbettung in Präpositionalphrasen stechen die Name-ICCs heraus. Jeder ICC-Typ hat mindestens eine Präposition, mit der er besonders häufig verwendet wird. Bei den Name-ICCs ist das

von, bei den Prot-ICCs *als* und bei den Det-ICCs *auf* und *über*. Name-ICCs haben aber, abgesehen von *in*, eine besondere Abneigung gegen die anderen Präpositionen, sodass 10 der 15 häufigsten Präpositionen eine sehr geringe oder sogar überhaupt keine Rolle spielen. Auch attributive Adjektive treten bei Name-ICCs nahezu überhaupt nicht auf. Im Kontrast dazu werden Prot-ICCs unter den ICC-Typen am häufigsten durch Adjektive modifiziert, fast ebenso häufig wie Nomina im Allgemeinen.

Diese Besonderheit der Prot-ICCs fügt sich in das Gesamtbild von Prot-ICCs als kontextabhängigen Bildungen. Die qualitative Datenanalyse konnte zeigen, dass Sprecher:innen Prot-ICCs auf vielfältige Weise im Kontext anreichern, um eine korrekte Interpretation durch die Gesprächspartner:innen zu gewährleisten. Diesbezüglich konnten vier sprachliche Mittel identifiziert werden, die die Sprecher:innen nutzen, um die Bedeutung der ICCs hin zum Prototypen zu vereindeutigen: Kontrastierende Nominalphrasen, Erklärungen in Form von (eingebetteten) Sätzen, Fokus- und Negationspartikeln sowie anreichernde attributive Adjektive.

Diese Korpusstudie liefert viele Erkenntnisse zu ICCs. Allerdings konnten bisher keine Aussagen über die Produktivität der ICC-Typen gemacht werden. Eine Beschreibung etwa der Hapax legomena oder der generellen Stamm-Beleg-Verhältnisse ist ob der Datengrundlage, die aus einem vordefinierten Set an Items besteht, nicht möglich. Auch zur Gesamtheit der Stämme, mit denen ICCs gebildet werden, kann wegen des begrenzten Itemsets keine Aussage getroffen werden. Es ist außerdem möglich, dass manche der bisher beschriebenen Merkmale der ICCs bloß auf eben jenes limitierte Itemset zurückzuführen sind. Die im Folgenden beschriebene Studie kann einige dieser noch offenen Fragen beantworten. Zudem kann diese zweite Studie die bisher gewonnenen Erkenntnisse mit weiterer Evidenz unterstützen.

5.2 deTenTen13

5.2.1 Überblick über die Daten und die verwendeten statistischen Tests

Die deTenTen13-Daten umfassen 180.700 ICC-Belege, die sich auf 4426 Wortformen sowie auf 2.895 Lexeme / Stämme verteilen. Abbildung 25 zeigt die Häufigkeitsverteilung der ICC-Belege auf die ICC-Typen.

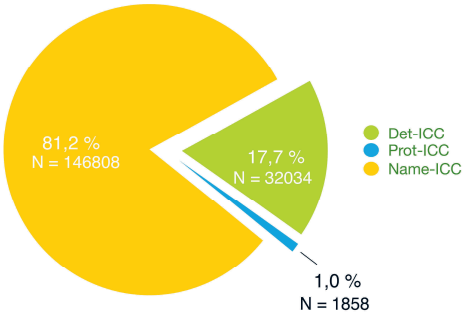

Abb. 25: Verteilung der morphologischen Verbindungen (deTenTen13).

Mit der Analyse der Daten aus deTenTen13 soll die Produktivität der drei ICC-Typen in Form des Stamm-Beleg-Verhältnisses und der Hapax legomena miteinander verglichen werden. Zudem werden die Belege, die Wortformen sowie die Stämme der drei ICC-Typen hinsichtlich ihrer Länge verglichen, um die Ergebnisse zu den Längenunterschieden zu überprüfen. Auch sollen die Ergebnisse der ersten Studie zur Verwendung von Fugenelementen und Flexionsmarkern sowie zur Schreibung überprüft werden.

Die statistischen Tests, die die Vergleiche jeweils begleiten, richten sich wiederum nach dem Skalenniveau der jeweiligen Variablen. Zu Variablen, die in den Daten intervallskaliert sind, werden über die deskriptive Statistik hinaus Tests zur statistischen Signifikanz der Unterschiede sowie zur Effektstärke durchgeführt. Dies betrifft die Variablen BELEGLÄNGE, WORTFORMLÄNGE sowie BASENLÄNGE. Hier wird Welchs ANOVA gerechnet, sofern die Gruppen gemäß dem Shapiro-Wilk Test nicht normalverteilt sind und die Voraussetzung der Varianzhomogenität verletzt wird. Zu Variablen, die in den Daten nominalskaliert sind, werden allein die Häufigkeiten berichtet. Dies betrifft die Variablen FUGENELEMENT, FLEXIONSMARKER und SCHREIBUNG. Auch beim Vergleich zur Produktivität der ICC-Typen sowie der Analyse der Stammverteilung werden die Stämme und Belege lediglich ausgezählt. Für die Durchführung aller statistischen Tests wurde IBM SPSS Statistics (Version 27.0.0.0) verwendet.

5.2.2 Analyse zu den Eigenschaften der ICC-Typen

5.2.2.1 Stamm-Verteilung und Hapax legomena

Durch die offene Abfragemethodik ist es in dieser Korpusstudie möglich, die ICC-Typen hinsichtlich des Stamm-Beleg-Verhältnisses zu vergleichen. Hier zeigt sich, dass sich das Stamm-Beleg-Verhältnis der Prot-ICCs stark von dem der beiden anderen ICC-Typen unterscheidet. Bei den Prot-ICCs bildet ein Stamm im Schnitt 4 Belege, bei den Det-ICCs hingegen 29, bei den Name-ICCs 88. Zum einen sind die Stämme bei den Prot-ICCs generell für weniger Belege verantwortlich als die der Det- und Name-ICCs, zum anderen findet sich bei den Prot-ICCs ein höherer Anteil nur einmal auftretender Belege / Stämme, die nur einen Beleg liefern. Abbildung 26 vergleicht die Stamm-Verteilung der zehn Stämme, die innerhalb der drei ICC-Arten die meisten Belege hervorbringen.

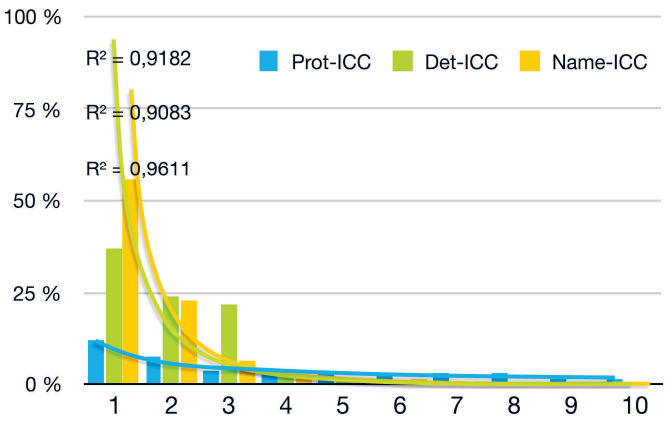

Abb. 26: Vergleich der 10 produktivsten Stämme der drei ICC-Typen hinsichtlich des Anteils an den Gesamtbelegen des jeweiligen ICC-Typs.

Alle ICC-Arten haben eine exponentiell verlaufende Stamm-Beleg-Verteilungskurve. Für alle ICC-Typen sind einige wenige Lexeme für sehr viele Belege verantwortlich; der Großteil der Stämme produziert hingegen nur wenige Auftreten von ICCs. Allerdings ist bei den Prot-ICCs die Produktivität der Stämme nicht ganz so disparat. Die Leistungstrendlinie in Abbildung 26 zeigt, dass die zehn Stämme, die bei den Det- und Name-ICCs die meisten Belege hervorbringen, einen viel größeren Anteil der jeweiligen Gesamtbelege ausmachen als bei den Prot-ICCs. Diese zehn Stämme sind bei den Det- und Name-ICCs die Basis für

knapp 90% der entsprechenden Gesamtbelege. Bei den Prot-ICCs sind die zehn Stämme, die die meisten Belege hervorbringen, hingegen nur für 43% aller Belege verantwortlich. Stämme wie etwa *Chef*, *Film* oder *Künstler* dominieren die Prot-ICC-Daten also nicht so sehr wie *Zins*, *Helfer* und *Kind* die Det-, beziehungsweise *Baden*, *Tom* und *Mann* die Name-ICC-Daten. Tabelle 11 zeigt die Beleganzahl der genannten Stämme für die drei ICC-Typen.

Tab. 11: Belegzahlen der zehn Stämme, die die meisten Belege hervorbringen

Top 10	**Det-ICC**			**Prot-ICC**			**Name-ICC**		
	Stamm	**N**	**% aller Fälle**	**Stamm**	**N**	**% aller Fälle**	**Stamm**	**N**	**% aller Fälle**
1	*zins*	11.941	37,3%	*chef*	221	11,9%	*baden*	81.849	55,8%
2	*helfer*	7.743	24,2%	*film*	146	7,8%	*tom*	33.617	22,9%
3	*kind*	6.932	21,6 %	*kuenstler*	71	3,8%	*mann*	9.095	6,2%
4	*freund*	678	2,1%	*maedchen*	68	3,7%	*zauberer*	2675	1,8%
5	*kompetenz*	441	1,4%	*literatur*	59	3,2%	*clown*	1898	1,3%
6	*erwartung*	184	0,6%	*glas*	58	3,1%	*heim*	1840	1,3%
7	*schranke*	158	0,5%	*frau*	56	3,0%	*gas*	497	0,3%
8	*erbe*	115	0,4%	*mann*	55	3,0%	*money*	467	0,3%
9	*stadt*	79	0,2%	*theater*	31	1,7%	*book*	459	0,3%
10	*kandidat*	77	0,2%	*mitte*	30	1,6%	*beri*	447	0,3%
Insg.		**28.348**	**88,5%**		**795**	**42,7%**		**81.849**	**90,5%**

Diese Beobachtung wird ergänzt durch den Vergleich hinsichtlich der nur wenig produktiven Stämme. Der Prot-ICC-Typ weist einen höheren Anteil an Hapax legomena auf, also Belege und Stämme, die im Korpus nur einmal (für den jeweiligen ICC-Typ) belegt sind. Tabelle 12 gibt einen Überblick über die Produktivität innerhalb der drei ICC-Arten.

Tab. 12: Unterschiede im Stamm-Beleg-Verhältnis der drei ICC-Typen (deTenTen13)

ICC-Typ	Token	Stämme	Ø Belege / Stamm	Hapax-Belege		Hapax-Stämme	
Det-ICC	32.034	1.115	28,73	1.151	3,6%	598	53,6%
Prot-ICC	1.858	458	4,06	466	25,1%	262	57,2%
Name-ICC	146.808	1.669	487,96	947	0,6%	637	38,2%
Gesamt	180.700	2.895	62,42	–	–	–	–

In Tabelle 12 ist die Gesamtanzahl der ICC-bildenden Stämme (2895) niedriger als die Summe der Stämme aus den drei ICC-Typen. Das hängt damit zusammen, dass manche Stämme in mehr als nur einem ICC-Typ vorkommen. Allerdings betrifft dies nur wenige Stämme. Der größte Teil der Stämme ist Basis nur eines ICC-Typs. Abbildung 27 gibt einen Überblick über die Verteilung aller 2895 Stämme auf die drei ICC-Typen.

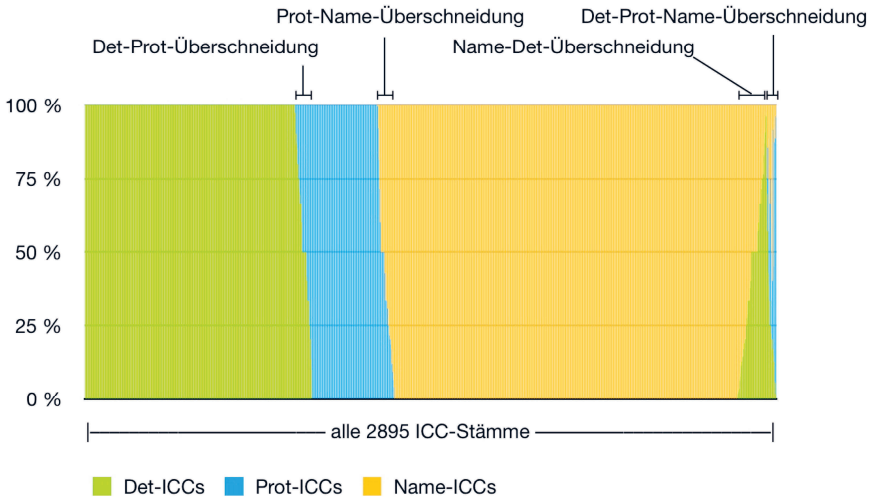

Abb. 27: Prozentuales Auftreten der ICC-Stämme in den drei ICC-Typen.

Zirka 90% der Stämme werden exklusiv in nur einem der ICC-Typen verwendet. 30% der Stämme bilden nur Det-ICCs, 10% nur Prot-ICCs und 50% nur Name-ICCs. Hier zeigt sich zum einen die geringere Produktivität der Stämme in der Prot-ICC-Konstruktion. Machten Prot-ICCs beim Vergleich der Belege noch weniger als 1%

der Daten aus, sind es beim Vergleich der Stämme fast 10%. Name-ICCs hingegen machen im Vergleich der Anzahl der Stämme nur etwas mehr als die Hälfte aller Stämme aus; beim Vergleich der Belege waren es noch über 80%.

Zum anderen kann man beobachten, dass nur 10% der Daten nicht exklusiv einen ICC-Typ bilden, sondern in mindestens zwei ICC-Typen auftreten. So bilden 3% der Stämme Det- und Prot-ICCs (in Abbildung 27 der Übergang grün –> blau), 2% Prot- und Name-ICCs (blau –> gelb), 4% Name- und Det-ICCs (gelb –> grün) und 1% alle drei ICC-Typen. Die allermeisten Stämme haben also eine klare Präferenz für einen der drei ICC-Typen. Von den 10% der Stämme, die mehr als nur einen ICC-Typ bilden, tun dies nur wenige mit einer ausgeglichenen Verteilung. Der Stamm *Mensch* tritt im Korpus etwa sowohl als Prot-ICC (79) als auch als Det-ICC (80) 16 mal auf:

(79) *Als Marketing-Mensch muss ich sagen: Wie unprofessionell kann PR-Arbeit sein? Als **Mensch-Mensch** muss ich fragen: Wie dreist und unwürdig kann ein Unternehmens seine Beschäftigten behandeln.*

<https://www.mind-steps.de/2013/02/21/amazon-und-social-responsibility-nichts-geht-mehr>

(80) *ich bin ein **Menschenmensch**. Ich brauche echte Menschen um mich rum.*

<https://daslauteunddasleise.blogger.de/topics/Festgehalten>

Nur sehr wenige Stämme treten in allen drei ICC-Konstruktionen auf, etwa *Buch*:

(81) Det-ICC mit *Buch* als Basis
***Bücherbücher** sind, wie der Name bereits andeutet, Bücher in denen Bibliotheken, Buchmenschen, Buchherstellung, Bücher, Schreiben und Lesen das zentrale Thema sind oder eine große Rolle spielen.*

<http://www.buecher-wiki.ch/index.php/BuecherWiki/Bibliophilie>

(82) Prot-ICC mit *Buch* als Basis
*So handlich das Format, so fest der Einband und gediegen das Papier – ein Buch, das schon haptisch seine Begeisterung für's Gedruckte mitteilt (wie viele Bücher von Zweitausendeins, und noch mehr die früher im damals noch eigenständigen Haffmans-Verlag erschienenen). Solche Bücher brauchen kein Kindle zu scheuen, solche Bücher sagen: Ich bin noch da, wenn die letzten Pixel längst im digitalen Nirvana verglüht sind. So ein **Buch-Buch** verrät natürlich auch gleich: Ich bin von einem Autor-Autor.*

<https://www.britcoms.de/page/10>

(83) Name-ICC mit *Buch* als Basis
*Archiv des Themenkreises '**Buchbuch**'. Heute endlich Zeit gehabt, den Artikel "¿Qué se hizo de Luis Harss?" des Romanciers Tomás Eloy Martínez zu lesen.*

<http://www.umblaetterer.de/category/buchbuch/page/29>

Die Visualisierung der Daten in Abbildung 27 zeigt also, dass sich die Stämme generell sehr komplementär auf die drei ICC-Typen verteilen und also selten mehr als nur einen ICC-Typ bilden. Man könnte nun versucht sein, in dieser Beobachtung einen bloßen Artefakt-Effekt zu sehen. Wie berichtet, tritt ein großer Teil der Stämme überhaupt nur einmal auf (Hapax-Stämme). Diese Stämme haben also scheinbar gar nicht die Möglichkeit, in den anderen ICC-Typen vorzukommen, was wiederum bedeuten würde, dass das Bild der komplementären Stammverteilung nur durch eine verzerrte Datendarstellung entsteht.

Selbstverständlich ist dies aber nicht so. Auch die Stämme, die sich überhaupt nur einmal als ICC in den Daten befinden, haben natürlich trotzdem prinzipiell die Möglichkeit, in einem anderen ICC-Typ aufzutreten. Mit einer Frequenz von 1 in einem bestimmten ICC-Typ vorzukommen ist keine arithmetisch bedingte Notwendigkeit. Aber selbst wenn man nur Stämme in die Darstellung mit einbezieht, die eine gewisse Mindestfrequenz in der Konstruktion ICC aufweisen, zeigt sich ein ähnliches Bild zu dem in Abbildung 27. Abbildung 28 zeigt das prozentuale Auftreten in den ICC-Typen nur derjenigen Stämme, die mehr als 3 Belege liefern.

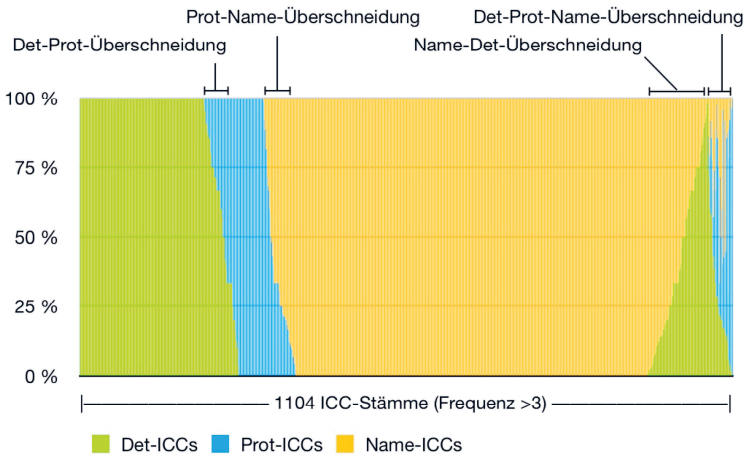

Abb. 28: Prozentuales Auftreten der ICC-Stämme in den drei ICC-Typen (nur Stämme mit mehr als 3 Auftreten).

Von den Stämmen, die mehr als dreimal als ICC vorkommen, werden 77% exklusiv in nur einem der ICC-Typen verwendet. 19% der Stämme bilden nur Det-ICCs, 4% nur Prot-ICCs und 54% nur Name-ICCs. Nur 23% der Stämme treten in mindestens zwei ICC-Typen auf. So bilden 5% der Daten Det- und Prot-ICCs (in Abbildung 29 der Übergang grün –> blau), 5% Prot- und Name-ICCs (blau –> gelb), 9% Name- und Det-ICCs (gelb –> grün) und 4% alle drei ICC-Typen.

5.2.2.2 Länge

Auch in den deTenTen13-Daten unterscheiden sich die drei ICC-Typen hinsichtlich ihrer Länge (Tabelle 13).

Tab. 13: Deskriptive Statistik, Teststatistik sowie Gruppenvergleiche zur Länge

Deskriptive Statistik					
Variable	**ICC-Typ**	**N**	***M***	***SD***	***SEM***
Beleglänge	Det-ICCs	32034	12,16	2,355	0,013
	Prot-ICCs	1858	11,13	3,624	0,084
	Name-ICCs	146808	9,78	2,524	0,007
Wortformlänge	Det-ICCs	1805	14,3856	4,74000	0,11157
	Prot-ICCs	688	12,4956	3,81829	0,14557
	Name-ICCs	2111	10,6002	3,39197	0,07383
Basenlänge	Det-ICCs	1115	6,67	2,450	0,073
	Prot-ICCs	458	5,94	1,985	0,093
	Name-ICCs	1669	5,06	1,730	0,042

ANOVA					
Variable	**Statistik**	***df1***	***df2***	***p***	η^2
Beleglänge	13142,779	2	4817,399	0,000	0,117
Wortformlänge	411,205	2	1862,190	0,000	0,157
Basenlänge	191,599	2	1178,953	0,000	0,155

Games-Howell post-hoc Test (Gruppenvergleiche)						
Variable	**Vergleich**	**Differenz**	***SEM***	***p***	**95% Konfidenzintervall**	
					Untergrenze	**Obergrenze**
BELEGLÄNGE	Prot-ICC–Det-ICC	-1,028	0,085	0,000	-1,23	-0,83
	Name-ICC–Det-ICC	-2,381	0,015	0,000	-2,42	-2,35
	Name-ICC–Prot-ICC	-1,353	0,084	0,000	-1,55	-1,15
WORTFORMLÄNGE	Prot-ICC–Det-ICC	-1,88996	0,18341	0,000	-2,3202	-1,4597
	Name-ICC–Det-ICC	-3,78541	0,13378	0,000	-4,0991	-3,4717
	Name-ICC–Prot-ICC	-1,89545	0,16322	0,000	-2,2785	-1,5124
BASENLÄNGE	Prot-ICC–Det-ICC	-0,726	0,118	0,000	-1,00	-0,45
	Name-ICC–Det-ICC	-1,608	0,085	0,000	-1,81	-1,41
	Name-ICC–Prot-ICC	-0,882	0,102	0,000	-1,12	-0,64

Die Unterschiede treten ungeachtet dessen auf, ob die einzelnen Belege, die Wortformen oder die Basen verglichen werden. Die Wortlänge, gemessen an der Zeichenanzahl der Belege, unterscheidet sich für die verschiedenen Bedingungen Det-ICC, Prot-ICC und Name-ICC (Abbildung 29). Name-ICC-Belege sind im Schnitt 1,4 Grapheme kürzer als Prot-ICC-Belege, die wiederum durchschnittlich ein Graphem kürzer als Det-ICCs sind. Welchs einfaktorielle ANOVA zeigt, dass es einen statistisch signifikanten Unterschied zwischen der Beleglänge von Det-, Prot- und Name-ICCs gibt.[34] Die Effektstärke nach Cohen (1988) liegt bei $f = .364$ und entspricht einem mittleren Effekt. Der Games-Howell post-hoc Test zeigte einen signifikanten Unterschied in der Beleglänge zwischen allen Gruppen.

34 In allen Vergleichen zur Wortlänge wurde Welchs einfaktorielle ANOVA gerechnet, da die Gruppen gemäß dem Shapiro-Wilk Test nicht normalverteilt ($p < .001$) waren und die Voraussetzung der Varianzhomogenität verletzt wurde (Levenetest $p < .001$).

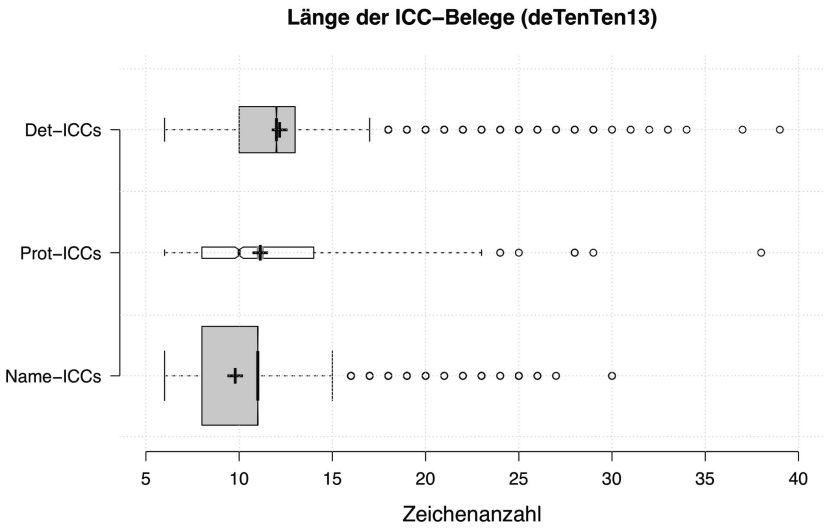

Abb. 29: Unterschiede zwischen den ICC-Typen in Bezug auf die durchschnittliche Zeichenanzahl (Belege).[35]

Die Unterschiede und Effekte sind noch stärker, wenn man nicht die Länge der Belege, sondern die Länge der in den ICC-Typen vertretenen Wortformen vergleicht und somit den Einfluss hochproduktiver Stämme beschränkt (Abbildung 30). Welchs einfaktorielle ANOVA zeigt, dass sich die Wortformlänge statistisch signifikant für die verschiedenen Bedingungen der ICC-Typen unterscheidet. Die Effektstärke nach Cohen (1988) liegt bei f = .432 und entspricht einem starken Effekt. Der Games-Howell post-hoc Test zeigte einen signifikanten Unterschied in der Wortformlänge zwischen allen Gruppen. Die Länge der Wortformen ist also, gemessen an der Zeichenanzahl der Wortformen, wie schon beim Vergleich der Belege, am geringsten bei Name-ICCs. Name-ICCs sind hier 1,9 Grapheme kürzer als Prot-ICCs, die wiederum 1,9 Grapheme kürzer sind als Det-ICCs. Somit sind Name-ICC-Wortformen im Mittel fast vier Grapheme kürzer als Det-ICCs.

35 Für diese und die folgenden Abbildungen gilt: Die Boxen entsprechen dem Bereich, in dem die mittleren 50 % der Daten liegen (oberes und unteres Quartil), die Breite der Boxen ist proportional zur Quadratwurzel der Beleganzahl. Die Antennen erstrecken sich über das 1,5-fache des Interquartilabstandes, Punkte stehen für Ausreißer, Mittellinien für die Mediane, Kreuze für die Mittelwerte. Graue Balken um die Kreuze und Kerben in den Boxen an der Stelle des Medians zeigen jeweils ein 95%-Konfidenzintervall des Mittelwerts/des Medians an (Chambers et al. 1983).

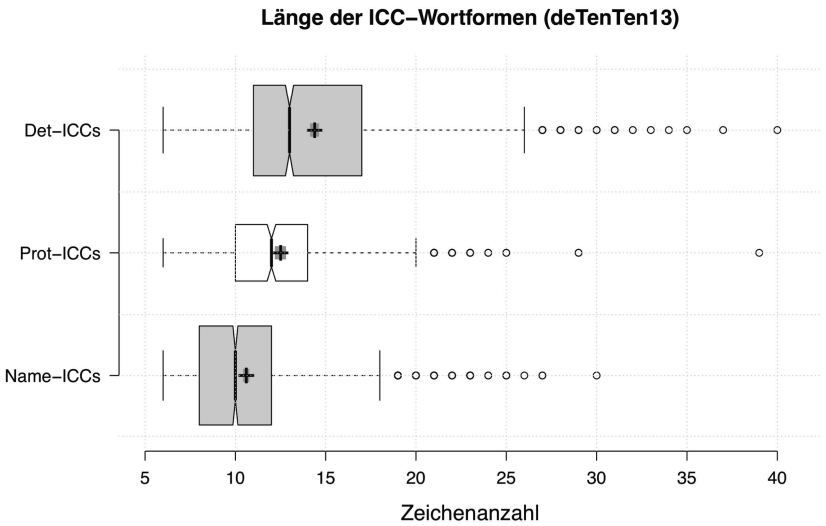

Abb. 30: Unterschiede zwischen den drei ICC-Typen in Bezug auf die durchschnittliche Zeichenanzahl (Wortformen).

Auch wenn man zum Längenvergleich zwischen den ICC-Typen die Basen heranzieht, gibt es signifikante Unterschiede (Abbildung 31). Welchs einfaktorielle ANOVA zeigt hier, dass sich die Länge der Basen statistisch signifikant für die verschiedenen Bedingungen der ICC-Typen unterscheidet. Die Effektstärke nach Cohen (1988) liegt bei f = .36 und entspricht einem mittleren Effekt. Der Games-Howell post-hoc Test zeigt einen signifikanten Unterschied in der Basenlänge zwischen allen Gruppen. Auch bei den ICC-Basen ist also die Zeichenanzahl der Basen am geringsten bei Name-ICCs. Im Vergleich mit den Prot-ICC-Basen sind Name-ICC-Basen 0,9 Grapheme kürzer, im Vergleich mit den Det-ICC-Basen 1,6 Grapheme.

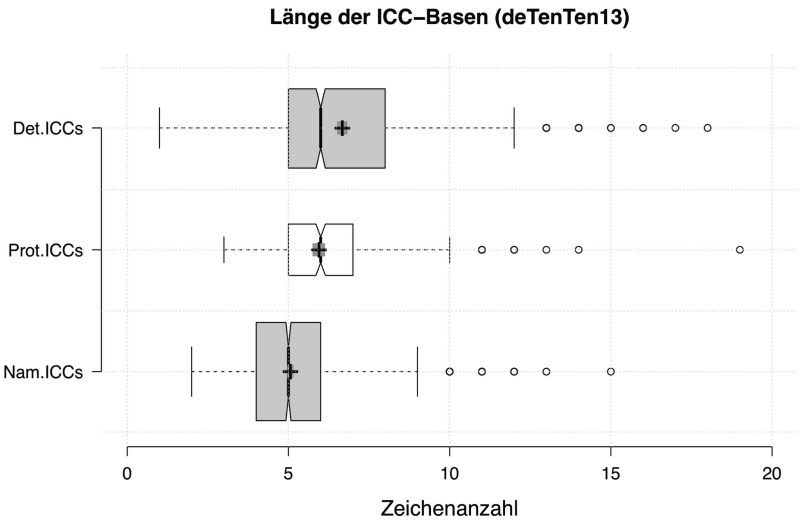

Abb. 31: Unterschiede zwischen den drei ICC-Typen in Bezug auf die durchschnittliche Zeichenanzahl (Basen).

Wäre nur die durchschnittliche Zeichenanzahl der ICC-Belege, nicht aber die der Wortformen und Basen unterschiedlich, wäre nicht auszuschließen, dass diese Unterschiede allein auf eine hohe Zeichenanzahl hochfrequenter ICCs wie etwa *Helfershelfer* oder *Kompetenzenkompetenz* zurückzuführen sind, die das Bild verzerren, weil sie die Mittelwerte der Det-ICCs nach oben ziehen. Doch auch wenn man den Einfluss solcher hochfrequenten ICCs beschränkt, indem alle Wortformen (und damit auch *Helfershelfer*) nur einmal gezählt werden, bleiben die Effekte bestehen und werden sogar noch stärker. Zwischen den Längen der ICC-Wortformen liegen jeweils fast 2 Zeichen. Det-ICCs sind 3,8 Grapheme länger als Name-ICCs. Das Ergebnis, dass ebenfalls die Wortformen sich in der Länge unterscheiden, spricht dagegen, dass ein Unterschied in der Zeichenanzahl allein mit einer hohen Zeichenanzahl der besonders hochfrequenten ICCs zusammenhängt.

Man könnte außerdem annehmen, dass der Unterschied in der Zeichenanzahl zum Teil mit unterschiedlicher Schreibung zusammenhängt, vorausgesetzt, dass Det-ICCs häufiger mit dem Bindestrich geschrieben werden und deswegen mehr Zeichen aufweisen. Doch dass sich nicht nur die Belege und Wortformen der drei ICC-Typen in der durchschnittlichen Länge unterscheiden, sondern auch die Basen, die den ICCs zugrundeliegen, hier also keine Bindestriche zur Trennung der

Konstituenten vorliegen können, spricht gegen diese Vermutung. Wie im folgenden Abschnitt noch zu sehen sein wird, hat der Einsatz des Bindestriches stattdessen sogar den gegenteiligen Effekt, dass nämlich Det-ICCs im Schnitt durch die Abwesenheit des Bindestriches mehr Grapheme aufweisen als Prot- und Name-ICCs.

5.2.2.3 Schreibung

Die drei ICC-Typen unterscheiden sich hinsichtlich ihrer Schreibung. Abbildung 32 zeigt die Verteilung der zwei automatisch erfassten Schreibvarianten (Bindestrichschreibung und Zusammenschreibung) für die drei ICC-Typen.

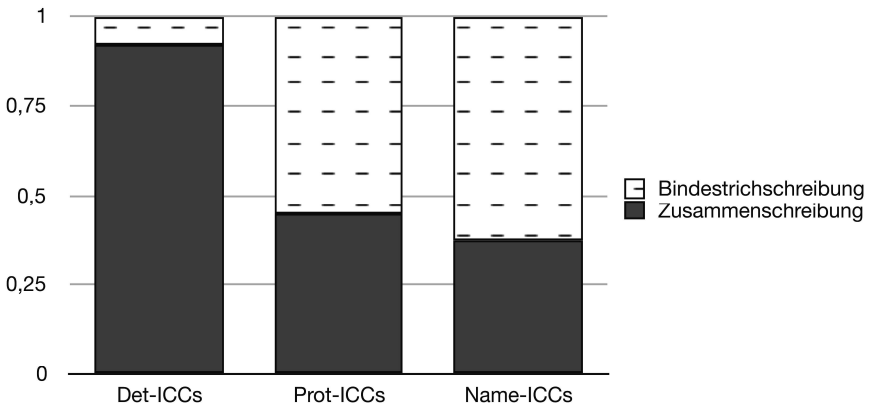

Abb. 32: Schreibung der drei ICC-Typen (deTenTen13).

Det-ICCs zeigen in 92% der Fälle die übliche Kompositaschreibung, werden also fast immer zusammengeschrieben. Bei Prot-ICCs (45% Zusammenschreibung) und Name-ICCs (28% Zusammenschreibung) werden die Konstituenten hingegen in den meisten Fällen mit einem Bindestrich getrennt.

5.2.2.4 Fugenelemente

Die Längenunterschiede der Belege lassen sich teilweise darauf zurückführen, dass die drei ICC-Typen in unterschiedlichem Maße Fugenelemente und flexionsaffixanaloge Elemente aufweisen. Abbildung 33 zeigt die Verteilung der Marker, die an Erst- (N1) und Zweitglied (N2) der drei ICC-Typen auftreten.

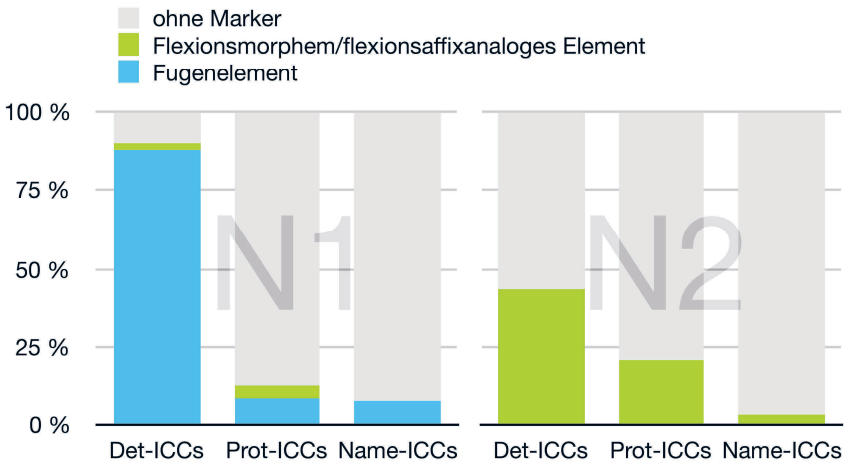

Abb. 33: Verteilung der Marker an Erst- und Zweitglied der drei ICC-Typen (deTenTen13).[36]

In knapp 88% der Fälle haben Det-ICCs ein Fugenelement, das das Erstglied erweitert, aber nicht mit dem Marker am Zweitglied identisch ist. Bei Prot-ICCs sowie bei Name-ICCs findet sich ein solches Element in weniger als 8% der Fälle. Man kann zunächst nicht ausschließen, dass die großen Unterschiede in der Häufigkeit der Verfugung damit zu tun haben, dass bei den Det-ICCs, wie in Kapitel 5.2.2.1 beschrieben, wenige Stämme sehr viele Belege generieren, dergestalt, dass beispielsweise die vielen auf *Helfer* und *Kind* basierenden Det-ICCs für den hohen Anteil Fugenelemente verantwortlich sind (*Helfershelfer*, *Kindeskind*). Doch selbst wenn man den Einfluss solch hochfrequenter Stämme begrenzt und jede Wortform nur einmal zählt, bleiben die Verfugungsunterschiede zwischen den ICC-Typen bestehen. Schaut man allein auf die gebildeten Wortformen (also unabhängig davon, wie viele Belege sie generieren), wird der Unterschied in der Verfugung zwischen Det-, Prot und Name-ICCs allerdings kleiner. Det-ICC-Wortformen sind zu 40% verfugt, Prot-ICC-Wortformen zu 17% und Name-ICC-Wortformen zu 7%. Diese Häufigkeitsverteilung spricht aber ebenso dafür, dass sich die ICC-Typen hinsichtlich der Verfugung sehr unterschiedlich verhalten.

Auch hinsichtlich der Form des Fugenelements unterscheiden sich die drei ICC-Typen. Bei Det-ICCs wird in fast 70% der verfugten Belege *-es-* verwendet, in 30% *-s-*. Somit ist der allergrößte Teil der verfugten Det-ICC-Belege mit dem Fu-

36 Die Unterscheidung in paradigmatische und unparadigmatische Fugenelemente entfällt hier leider, weil der Abgleich der Marker mit den Affixen im Flexionsparadigma der Lexeme wegen der großen Anzahl der Stämme zu aufwendig gewesen wäre.

genelement *-(e)s-* gebildet. Es herrscht also dasselbe Fugenelement vor, das auch für Nominalkomposita im Allgemeinen als das häufigste und produktivste gilt (Nübling & Szczepaniak 2009: 220). Ähnlich ist das bei Name-ICCs, wobei hier *-e-* mit 2,1% noch eine nennenswerte Rolle spielt. Ganz anders stellt sich die Verfugung bei den Prot-ICCs dar. Hier ist *-(e)n-* das am häufigsten verwendete Fugenelement, das fast für die Hälfte der verfugten Belege verantwortlich ist. Auch *-er-* und *-e-* sind bei den Prot-ICCs relevanter als bei den anderen ICC-Arten. Allerdings mag das damit zusammenhängen, dass Prot-ICCs generell seltener sind und die Formen noch dazu nur selten verfugt werden.

Blickt man auf die ICC-Stämme, also ungeachtet dessen, wie viele Belege sie generieren, spielen auch bei Det-ICCs andere Fugenelemente als *-(e)s-* eine Rolle. Die Elemente *-e-* und *-er-* verfugen 9% respektive 6% der verfugten Det-ICC-Stämme, wiederum 28% übernimmt *-(e)n-*. Doch auch bei den Stämmen dominiert bei den Det-ICCs das Fugenelement *-(e)s-*. Bei mehr als der Hälfte der verfugten Det-ICC-Stämme sind die Stämme durch dieses Fugenelement getrennt. Bei den verfugten Prot-ICC-Stämmen ist die Verteilung der Fugenelemente recht ähnlich wie bei den verfugten Prot-ICC-Belegen. Bei den Name-ICC-Stämmen sind alle Fugenelemente vertreten, wobei *-e-* und *-(e)s-* hier überwiegen. Sowohl für Prot-ICCs als auch für Name-ICCs gilt aber, dass generell wenige Stämme verfugt sind. Abbildung 34 zeigt die prozentuale Verteilung der Fugenelemente für die drei ICC-Typen.

Auch im Flexionsverhalten unterscheiden sich die ICC-Typen (Abbildung 33). In über 70% der Fälle flektieren Prot-ICCs überhaupt nicht overt, enthalten also keine Elemente über den Stamm hinaus. Bei den Name-ICCs treten sogar in 88% der Fälle keine Marker am Stamm auf. Die Det-ICCs sind hingegen nur sehr selten, nämlich in 7% der Fälle, auf den Stamm reduziert. Auch hier nehmen Prot-ICCs also wieder eine Zwischenposition ein zwischen Name-ICCs auf der einen Seite und Det-ICCs auf der anderen.

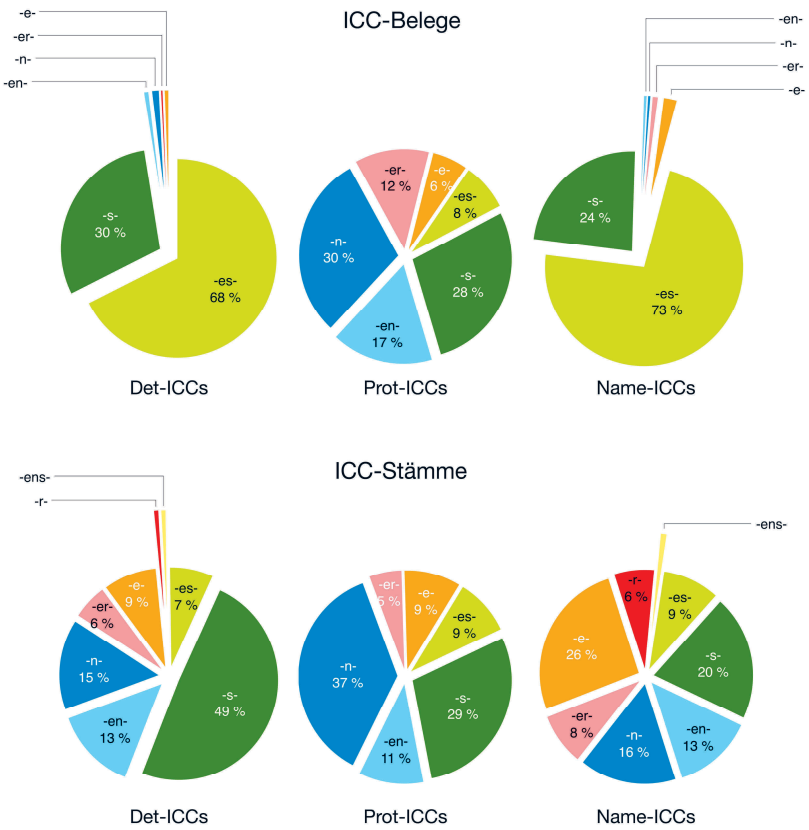

Abb. 34: Prozentuale Verteilung der Fugenelemente für die drei ICC-Typen.

In Bezug auf die Position der Flexionsmarker verläuft hingegen eine Grenze zwischen Det- und Name-ICCs auf der einen und Prot-ICCs auf der anderen Seite. Sofern overte Flexionsmarker vorliegen, flektieren Det-ICCs und Name-ICCs zu 95% respektive 91% als Einheit. Es zeigt also allein das Zweitglied einen Flexionsmarker. Prot-ICCs hingegen flektieren, wenn sie denn overt flektieren, nur in 77% der Fälle als Einheit, in den restlichen Prot-ICC-Belegen findet sich am Erstglied ein Element, das mit dem Flexiv am Zweitglied formidentisch ist.

Auch in Bezug auf totale Reduplikation, also die vollkommene Formgleichheit von Erst- und Zweitglied (siehe Kapitel 8), verhalten sich die ICC-Typen unterschiedlich. Bei Name-ICCs ist die Reduplikation in 89%, bei Prot-ICCs in rund 75%, bei Det-ICCs hingegen nur in 9% der Fälle eine totale Reduplikation.

5.2.3 Zusammenfassung der Ergebnisse (deTenTen13)

Die aus dem deTeTen13-Korpus erhobenen Daten liefern Evidenz zur Produktivität der ICC-Typen und bestätigen zudem viele der in der ersten Korpusanalyse gewonnenen Ergebnisse: Name-ICCs sind der im Mittel kürzeste ICC-Typ, Prot-ICC-Belege sind im Schnitt wiederum kürzer als Det-ICC-Belege. Name-ICCs sind so gut wie nie verfugt, Prot-ICCs nur selten, Det-ICCs hingegen fast immer. Name-ICCs sind nahezu immer frei von Markern, Prot-ICCs tragen in einem Fünftel der Fälle ein Flexiv am Zweitglied, Det-ICCs in fast der Hälfte der Fälle. Bei Prot- und Name-ICCs ist deshalb die Reduplikation in den meisten Fällen total, bei Det-ICCs partiell. Bei Name- und Prot-ICCs trennt meist ein Bindestrich die Konstituenten graphisch voneinander, Det-ICCs weisen größtenteils Zusammenschreibung auf.

Die mithilfe des deTeTen13-Korpus gewonnenen Daten zeigen aber auch weitere Unterschiede zwischen den ICC-Typen, auf die hin die DECOW16-Daten nicht analysiert werden konnten. So unterscheiden sich die drei ICC-Typen hinsichtlich ihrer Produktivität. Prot-ICC sind im Korpus am wenigsten vertreten. Sie machen nur knapp 1% der ICC-Belege aus. Bei Det- und Name-ICCs sind zudem wenige Stämme für extrem viele Belege verantwortlich, bei Prot-ICCs besteht kein so großer Unterschied in der Produktivität der einzelnen Stämme. Generell bilden die Prot-ICC-Stämme weniger Belege als die Det- und Name-ICC-Stämme. Der größte Teil der Prot-ICC-Stämme ist durch jeweils nur einen Beleg vertreten. Da die Abfrage in dieser Korpusstudie nicht auf einem vordefinierten Set an Stämmen beruht, sondern ICCs unabhängig von ihrer Basis erfasst, können zudem Aussagen über das Auftreten der Stämme in den drei ICC-Typen gemacht werden. Es zeigt sich, dass die Stämme nicht gleichermaßen Det-, Prot- und Name-ICCs bilden. Es gibt vielmehr eine nahezu komplementäre Verteilung der Stämme auf die ICC-Typen. Stämme, die beispielsweise für die Bildung eines Det-ICCs verwendet werden, treten mit hoher Wahrscheinlichkeit nicht in Prot- oder Name-ICCs auf. 90% der Stämme sind auf einen der ICC-Typen beschränkt. Nur 1,5% der Stämme treten in allen ICC-Typen auf.

Im folgenden Kapitel werden die Ergebnisse zu Semantik, Funktion und Form von ICCs zum bisherigen Forschungsstand und zu den zuvor formulierten Forschungsfragen in Bezug gesetzt.

6 Diskussion der Ergebnisse

6.1 Forschungsfragen zu Semantik und Funktion von ICCs

Die erste Forschungsfrage zu Semantik und Funktion von ICCs lautet:

1. Wie lassen sich zu ICCs Subgruppen bilden?

Die ICCs lassen sich semantisch-funktional in Subgruppen einteilen. Das Vorliegen des Basiskonzeptes in Erst- und Zweitkonstituente kann für die Einteilung zugrundegelegt werden. Dadurch zeigen sich drei Hauptfunktionen von ICCs, die durch die drei ICC-Typen Det-ICC, Prot-ICC und Name-ICC ausgeübt werden. Die Kriterien führen zu einer klaren Einteilung in drei ICC-Typen. Nur in Det-ICCs gibt es eine 1:1-Entsprechung von formaler und konzeptueller Dopplung, dergestalt, dass Erst- und Zweitglied durch ein eigenes nominales Konzept repräsentiert sind, das dem des Basislexems entspricht. Nur in Prot-ICCs ist es ferner so, dass das Erstglied rein funktional ist und kein nominales Konzept repräsentiert, das Zweitglied hingegen als semantischer Kopf der Bildung das nominale Konzept des Basislexems realisiert. In Name-ICCs schließlich ist es genau andersherum, dass also nur im Erstglied das nominale Konzept der Basis vorliegt, im Zweitglied aber nicht, weshalb das Zweitglied hier auch nicht der semantische Kopf der Bildung ist. Dass das Erstglied das nominale Konzept der Basis zeigt, ist bei Name-ICCs allerdings optional. Bei vielen Name-ICCs evoziert keine der Konstituenten das Konzept der Basis.

Die zweite Forschungsfrage zu Semantik und Funktion von ICCs lautet:

2. Welche Funktionen üben ICCs aus?

Die drei ICC-Typen Det-ICC, Prot-ICC und Name-ICC realisieren jeweils eine der drei Funktionen, die ICCs ausüben können. Erstens können ICCs ein Subkonzept zum Basiskonzept bilden und dabei zwei nominale Konzepte zueinander in Beziehung setzen, etwa räumlich, zeitlich oder logisch. Det-ICCs erfüllen damit exakt die Funktion der Determinativkomposita (Donalies 2011: 72, Ortner & Müller-Bollhagen 1991: 124), weshalb sie diesen auch eindeutig zugeordnet werden können. Das Konzept der Kopfkonstituente wird in Det-ICCs mithilfe des Modifikatorkonzeptes modifiziert und gleichzeitig subklassifiziert. Im ICC *Fenster-Fenster* ‘Fenster vor einem Fenster’ wird beispielsweise ein Subkonzept zu *Fenster* gebildet, indem das Konzept FENSTER räumlich mit einem weiteren Auftreten des Konzeptes FENSTER in Beziehung gesetzt wird. Zwischen den Kompositakonstituenten

https://doi.org/10.1515/9783111315416-006

dieser ICCs besteht durch dieses In-Beziehung-Setzen nominaler Konzepte, wie in anderen determinativen N+N-Komposita, eine Modifikationsrelation, etwa OF, FOR oder LOC. Dass die beiden Konstituenten solcher ICCs identisch sind, kann, wie in der Literatur zu ICCs häufig angenommen (Donalies 2011: 72, Elsen 2014: 67, Fleischer & Barz 2012: 96), als zufällig gelten. Im Vergleich mit N+N-Komposita, in denen die Konstituenten nicht identisch sind, kommt Det-ICCs nämlich keine zusätzliche Funktion zu. *Fenster-Fenster* ist also wie das N+N-Kompositum *Autofenster* eine Subklassifikation mithilfe eines Modifikatorkonzeptes. Diese Entsprechung von N+N-Komposita und Det-ICCs hinsichtlich der Subkonzeptbildung ist in Abbildung 35 dargestellt.

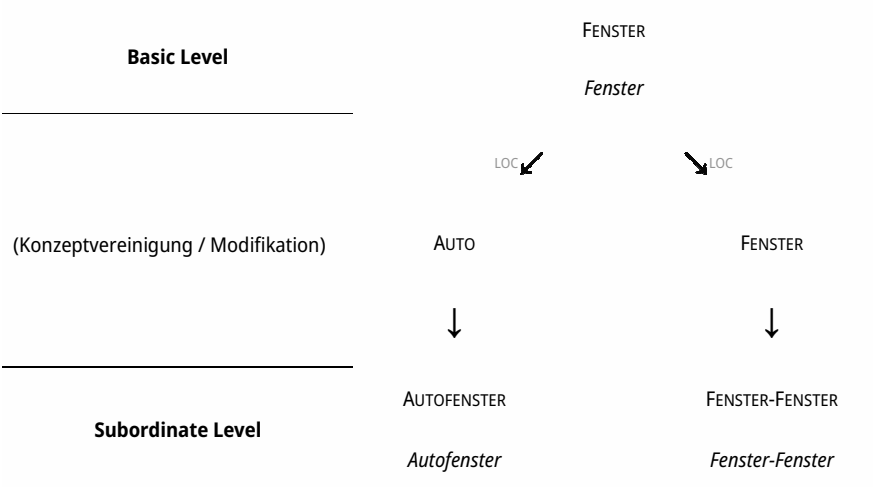

Abb. 35: Subkonzeptbildung von N+N-Komposita und Det-ICCs.

Des Weiteren können ICCs das Konzept der Basis auf seinen prototypischen Kern beschränken. In *Oma-Oma* etwa wird das Konzept OMA auf eine 'prototypische Oma' beschränkt, ein *Album-Album* ist analog dazu ein 'richtiges, prototypisches Album', ein *MannMann ein* 'richtiger, prototypischer Mann'. Diese ICCs werden dieser Prototypenbedeutung wegen Prot-ICCs genannt. Sie bilden ebenfalls Subkonzepte zum Konzept des Zweitgliedes. Prot-ICCs gehören damit wie Det-ICCs und N+N-Komposita im Allgemeinen zu den klassifikatorischen Komposita:

> [Prot-]ICCs share with canonical compounds the semantic effect of subcategorizing a concept.
>
> (Finkbeiner 2014: 188)

Das Verhältnis zwischen den Kompositakonstituenten in Prot-ICCs ist allerdings, anders als in determinativen N+N-Komposita und in Det-ICCs, kein In-Beziehung-Setzen zweier Konzepte. Dem Erstglied entspricht nicht das Basiskonzept. Stattdessen kann man es, je nach Analyse, entweder als Realisierung eines adjektivischen Konzeptes ansehen, sodass es bloß auf Eigenschaften des zugrundeliegenden Konzeptes verweist (Bross & Fraser 2020: 6) oder als ausschließlich funktionales Element ohne jegliche lexikalische Bedeutung (Finkbeiner 2014: 188). Im Zuge der Prot-ICC-Bildung führt es, ungeachtet der gewählten Analyse, zu einer Konstruktion, durch die dem Konzept der Basis, beziehungsweise dem entsprechenden Referenten, prototypische Eigenschaften zugeschrieben werden. Diese ICCs können mit Hohenhaus (2004: 301) Formel „an XX is a proper/prototypical X" paraphrasiert werden. Die Formel ist für die Beschreibung dieser Bildungen unter anderem deshalb angemessen, weil sie auf Adjektive zurückgreift („proper/prototypical"). Es sind die üblichen, (proto)typischen Eigenschaften des Basislexems, die dem Kopfkonzept des Prot-ICCs zugeschrieben werden. Anders ausgedrückt: Die Verwendung eines Prot-ICCs zeigt an, dass keine vom Basiskonzept abweichenden Eigenschaften vorliegen und der Referent nur die üblichen Eigenschaften des zugrundeliegenden Konzeptes aufweist. Omas, Alben und Männer, die nicht bloß die das Konzept generell kennzeichnenden Merkmale, sondern darüber hinaus Besonderheiten aufweisen, werden aus dem Referenzbereich ausgeschlossen.

In der bisherigen Forschung zu Prot-ICCs ist die Annahme verbreitet, der intendierte Referent eines Prot-ICCs werde in ein Kontrastverhältnis zu nicht-intendierten, ausgeschlossenen Referenten gesetzt. Diese Annahme zeigt sich unter anderem in der Verwendung von Termini, die den Kontrast hervorheben: „contrastive focus reduplication" (Bross & Fraser 2020, Ghomeshi et al. 2004), „contrastive reduplication" (Song & Lee 2015, Whitton 2006). Genau genommen besteht allerdings kein Kontrast zwischen den beiden einander gegenübergestellten Entitäten, beziehungsweise Konzepten. Die Gegenüberstellung von Prot-ICC-Referenten mit den Referenten des entsprechenden Basislexems kann mit Rückgriff auf die kognitive Relation „Same-Except" angemessener dargestellt werden. Diese Annahme bedarf eines kurzen Exkurses.

Die Same-Except-Relation geht auf William James zurück, der diese Relation wie folgt beschreibt:

> The perception of likeness is practically very much bound up with that of difference. That is to say, the only differences we note as differences, and estimate quantitatively, and arrange along a scale, are those comparatively limited differences which we find between members of a common genus. […] To be found different, things must as a rule have some commen-

> surability, some aspect in common, which suggests the possibility of their being treated in the same way.
>
> (James 1890: 528)

Culicover und Jackendoff (2012) veranschaulichen die verschiedenen Arten, auf die zwei Entitäten ähnlich sein können, mithilfe der in der Psychologie entwickelten *Wug*-Zeichnungen (Berko-Gleason 1958). Ein *Wug* ist eine schemenhafte, vogelähnliche Figur, die auch in der linguistischen Forschung vielfältig eingesetzt wird.

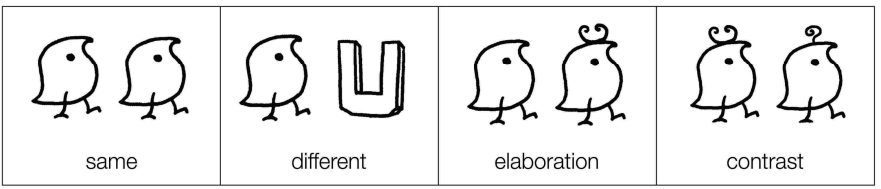

Abb. 36: Verschiedene Ähnlichkeitsrelationen zwischen *Wugs* (Culicover & Jackendoff 2012: 305f.).

Nach Culicover und Jackendoff (2012) können zwei Entitäten nicht nur gleich (same) oder völlig unterschiedlich (different) sein. Neben der völlig unterschiedlichen Erscheinung können zwei Entitäten auch im Rahmen der „Same-Except"-Relation unterschiedlich sein, also auf eine Weise, bei der die Entitäten nur einen begrenzten Unterschied, etwa in einer einzigen Eigenschaft, aufweisen. Teilen die Entitäten alle Eigenschaften, hat aber eine der Entitäten eine zusätzliche, besondere Eigenschaft, bezeichnet man das Verhältnis zwischen ihnen demnach als Elaboration (elaboration). Kontrast schließlich liegt dann vor, wenn beide Entitäten eine über das Grundkonzept hinausgehende besondere Eigenschaft haben, diese zusätzliche Eigenschaft aber unterschiedlich ausgestaltet ist (contrast).

Der Elaboration entspricht also in etwa die semantische Relation der Hyponymie: Der linke Wug im Feld *elaboration* in Abbildung 36 wäre demnach mit *Wug* zu bezeichnen, *Wug* damit Hyperonym zu einem den rechts daneben stehenden Wug bezeichnenden Hyponym, etwa *Geweih-Wug*. Dem Kontrast hingegen entspricht die semantische Relation der Ko-Hyponymie: Der linke Wug im Feld *contrast* in Abbildung 36 wäre demnach als *Geweih-Wug* Kohyponym zu einem den rechtsstehenden Wug bezeichnenden Ausdruck, etwa *Locken-Wug*. Culicover und Jackendoffs Terminologie folgend sind die Konzepte von N+N-Komposita Elaborationen der jeweiligen Kopfkonzepte. Ein *Schneeball* ist beispielsweise ein Ball, bei dem eine über das Konzept BALL hinausgehende Eigenschaften spezifiziert wird. Kontrast besteht hingegen zwischen N+N-Komposita

mit identischem Kopf, etwa *Schneeball* und *Lederball*. Die Konzepte beider Komposita gehen über das Konzept BALL hinaus, insofern sie zum Merkmal 'Material' unterschiedliche Ausprägungen aufweisen.

Hier zeigt sich ein bedeutender Unterschied zwischen Prot-ICCs und N+N-Komposita im Allgemeinen sowie zwischen Prot-ICCs und Det-ICCs im Besonderen. Prot-ICCs stehen, anders als etwa *Geweih-Wug* und *Locken-Wug*, nicht in einem Kontrastverhältnis zu N+N-Komposita mit demselben Kopf. Der *Wug-Wug* im Feld *elaboration* in Abbildung 36 ist der linke Wug, also der normale Wug, wie ihn Linguist:innen für gewöhnlich kennen. Der rechte Wug im selben Feld ist hingegen eine Elaboration dieses *Wug-Wugs*.

Das bedeutet freilich nicht, dass das Konzept, das ein Prot-ICC evoziert, identisch mit dem Basiskonzept ist. Hyperonyme können auf Entitäten referieren, auf die auch die jeweiligen Hyponyme referieren können. *Ball* kann beispielsweise wie sein Hyponym *Schneeball* auf einen Ball aus Schnee referieren, *Wug* auf dieselbe Zeichnung wie *Locken-Wug*. *Wug-Wug* kann aber nicht auf die zu *Locken-Wug* gehörige Zeichnung referieren. Der Grund dafür ist, dass auch die Bildung eines Prot-ICCs eine Information zum Basiskonzept hinzufügt, nämlich die Information, dass der Referent keine Eigenschaften aufweist, die über die üblichen Eigenschaften des Basiskonzeptes hinausgehen.

Damit verringert die Bildung eines Prot-ICCs die Vagheit des jeweiligen Grundkonzeptes. Mit dem Ausdruck *Wug* kann auf den nicht-modifizierten (nicht-elaborierten) Wug verwiesen werden. Mit dem Ausdruck *Wug-Wug* ist hingegen garantiert der nicht-modifizierte Wug gemeint. In *Wug* wird weder formal noch konzeptuell modifiziert. Die Funktion des Prot-ICCs *Wug-Wug* ist nun, dieses Nicht-Vorliegen einer konzeptuellen Modifikation mit formalen Mitteln als gesetzt und relevant hervorzuheben. Analog dazu bildet *Album-Album* ein Subkonzept zu ALBUM, genau wie andere N+N-Komposita und Nominalkomposita mit *Album* als Kopfkonstituente, etwa *Debutalbum*, *Vorgängeralbum*, *B-Seiten-Album* oder *Live-Album.* Im Gegensatz zu letzteren findet die Subklassifikation in *Album-Album* aber ohne das Hinzufügen neuer Teilkonzepte statt. Wie in Kapitel 4 dargestellt, ist die erste Konstituente rein funktional und gewährleistet die klassifikatorische Funktion, die Komposita als Defaultfunktion zukommt. Es liegt aber derselbe Stamm wie im Kopf zugrunde, was bei Prot-ICCs anzeigt, dass das Basiskonzept nicht mithilfe eines weiteren Konzeptes modifiziert wird, sondern lediglich, in gewisser Weise neutral, subklassifiziert wird. Prot-ICCs sind damit rein subkonzeptbildend. Es liegt keine semantisch-thematische, sondern eine rein funktionale Subklassifikation vor. Die Subkonzeptbildung von Prot-ICCs im Vergleich mit kanonischen N+N-Komposita ist in Abbildung 37 dargestellt.

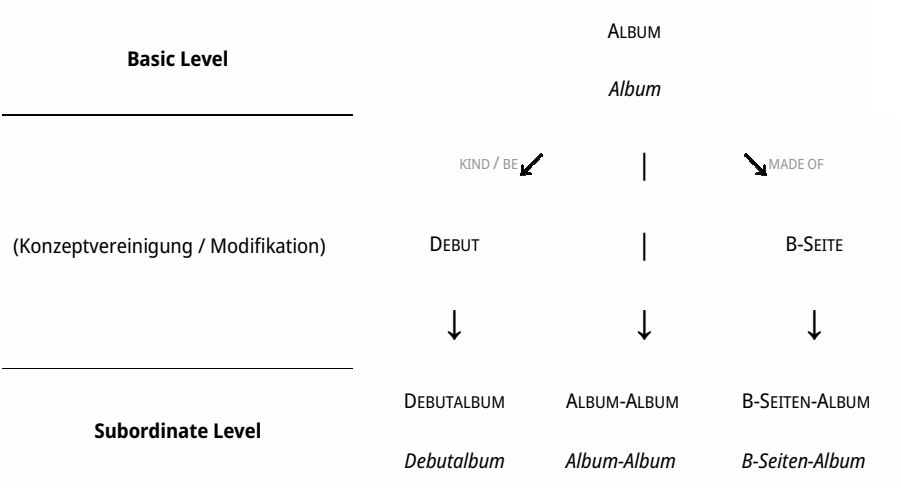

Abb. 37: Subkonzeptbildung von N+N-Komposita und Prot-ICCs.

Statt eines Kontrastes zwischen dem Prot-ICC-Referenten und dem Referenten des Basisnomens, beziehungsweise zwischen Prot-ICC-Konzept und Basiskonzept, zeigt die Bildung eines Prot-ICCs also an, dass eine Subklassifikation des Basiskonzeptes stattgefunden hat, ohne das Basiskonzept hinsichtlich anderer Konzepte zu modifizieren. Formal liegt eine N+N-Komposition vor. Semantisch-funktional allerdings bleibt die bei N+N-Komposita übliche konzeptuelle Verschiebung aus. Stattdessen wird das Basiskonzept ohne Zuhilfenahme eines weiteren Konzeptes subklassifiziert. Mit der Bildung eines Prot-ICCs wird also ein *Basic-Level-Konzept* in ein *Subordinate-Level-Konzept* konvertiert und dabei die Vagheit des Grundkonzeptes in Bezug auf mindestens eine Bedeutungskomponente beseitigt.

Prot-ICCs stehen damit in einem völlig anderen Verhältnis zu ihren Basiskonzepten als Det-ICCs und andere Determinativkomposita. Dieser Unterschied lässt sich genauer veranschaulichen, wenn man die Arbeiten zur Modifikation in Komposita von Spalding und Kolleg:innen (2019) berücksichtigt. In ihrer Studie weisen sie sogenannte Modifikationseffekte für N+N-Komposita nach („modification effects“, Spalding et al. 2019). Sprecher:innen, denen eine zuvor unbekannte semantische Information zu einem Nomen gegeben wurde, übertragen diese Information nicht vollständig auf Komposita, in denen das gegebene Nomen Kopfnomen ist.

> [P]roperties generally true of the unmodified head noun become less true of the modified head (modification effect), while properties generally false of the unmodified head become less false of the modified head (inverse modification effect).
>
> (Spalding et al. 2019: 2)

Im Experiment halten Proband:innen etwa die Information, dass Vögel Sesambeine besitzen („birds have sesamoid bones"), bei Amseln (in der Studie von Spalding et al. (2019) als englisches Kompositum *blackbird* präsentiert) für weniger wahrscheinlich als beim Überbegriff 'Vogel'.[37] Mit der Subklassifikation geht also die Gewissheit, dass bestimmte Eigenschaften vorliegen, teilweise verloren. Der Grund für solche Modifikationseffekte ist, dass Menschen im Allgemeinen erwarten, dass unterschiedliche Namen für Dinge auch mit tatsächlich zugrundeliegenden semantischen Unterschieden einhergehen („different names for things reflect real underlying semantic differences", Spalding et al. 2019: 2).

Angesichts der semantisch-funktionalen Besonderheit von Prot-ICCs, neutral zu subklassifizieren, verspricht ihre Modifikation nun vollkommen entgegengesetzt zu funktionieren. Hinsichtlich der Prot-ICCs muss die Erklärung, die Spalding et al. (2019) für die Modifikationseffekte formulieren, ergänzt werden. In den Stimuli, die in den Studien von Spalding und Kolleg:innen verwendet werden, wird den Grundkonzepten (im Beispiel dem Konzept VOGEL) in Form der Modifikatoren (im Beispiel das Adjektiv *black*) ein weiteres Konzept hinzugefügt. Diese Studien zeigen, dass die Sprecher:innen berücksichtigen, dass das Basiskonzept dadurch verschoben wird und halten deshalb für weniger wahrscheinlich, dass die mit dem Basiskonzept verbundenen Merkmale auch beim Referenten des entsprechenden Kompositums vorliegen. Bei Prot-ICCs hingegen werden dem Konzept der Basis keine weiteren Konzepte hinzugefügt. Die Eigenschaften der Referenten der Basisnomina kommen den Referenten von Prot-ICCs im Vergleich mit N+N-Komposita im Allgemeinen deshalb nicht zweifelhafter, sondern unzweifelhafter zu. Einem *Winter* kommt meist die Eigenschaft zu, von Schneefall begleitet zu sein, in einem *Winter-Winter* aber fällt in jedem Fall Schnee. *Omas* haben fast immer graue Haare, *Oma-Omas* aber unbedingt. Die Merkmale des Basiskonzeptes werden bei der Prot-ICC-Bildung also nicht so sehr verstärkt als vielmehr verifiziert. In einem *Winter-Winter* fällt nicht außergewöhnlich viel Schnee, eine *Oma-Oma* hat keine außergewöhnlich grauen Haare. Die Merkmale, die die Spre-

37 Eine Bedingung der Studie etwa führt die Oberkategorie ('Vogel') mit der Formel „almost all X have/do… Y" ein. Die Proband:innen schrieben daraufhin den unmodifizierten Basisnomina durchschnittlich in 91,1% der Fälle die jeweiligen semantischen Eigenschaften zu, den entsprechenden durch modifizierte Nomina (in Form von Komposita) repräsentierten Kategorienvertretern hingegen nur in 66,7% der Fälle (Spalding et al. 2019: 4).

cher:innen mit dem Basiskonzept verbinden (Schnee, graue Haare), werden stattdessen lediglich bestätigt. Man kann also vermuten, dass die Modifikation in Prot-ICCs andere Effekte zeigt und bei den Sprecher:innen nicht die Modifikationseffekte hervorruft, die Spalding und Kolleg:innen für N+N-Komposita nachgewiesen haben, sondern, im Gegenteil, negative Modifikationseffekte auftreten.

Da die Testitems in der Studie von Spalding et al. (2019) aber sämtlich nichtidentische Konstituenten haben, ist die Existenz negativer Modifikationseffekte bis dato freilich nur eine Hypothese. Die Ergebnisse von Spalding et al. (2019) ermöglichen aber eine neue Perspektive auf die in dieser Arbeit beschriebenen theoretischen Unterschiede zwischen Prot-ICCs und determinativen N+N-Komposita: Die Markierung einer besonderen Modifikationsart ist das wichtigste funktionale Alleinstellungsmerkmal der Prot-ICCs. Prot-ICCs zeigen an, dass das Basiskonzept nicht modifiziert ist. Mögliche Termini hierfür wären etwa „Nullmodifikation", „Nicht-Modifikation" oder „neutrale Subklassifizierung". Bislang ist diese Art der (Nicht-)Modifikation aber noch nicht gut erforscht. Zu nennen wären hier etwa die Arbeiten von Horn (1993, 2002, 2005, 2018).

In gewisser Weise besteht hier eine Ähnlichkeit zwischen Prot-ICCs und anderen klassifikatorischen Komposita, in denen die Erstkonstituente keine Bedeutung trägt. Auch bei N+N-Komposita wie *Betazelle* oder *Röntgenstrahlen* zeigt der Modifikator lediglich an, dass ein Subkonzept gebildet wird (Schlücker 2014: 118). Doch unterscheiden sich Prot-ICCs von diesen Komposita in manchen Aspekten: In N+N-Komposita wie *Betazelle* oder *Röntgenstrahlen* fungiert „N1 lediglich als Namensgeber für das vom Kompositum bezeichnete Subkonzept" (Schlücker 2014: 119) und trägt nicht zur Erschließung der Kompositionsbedeutung bei. Zwischen den Konstituenten besteht die Relation commemorative ('Y ist nach X benannt'). Die erste Konstituente eines Prot-ICCs gibt der Subklasse hingegen nicht bloß einen Namen, sondern reduziert zusätzlich die Vagheit des Basiskonzeptes.

Dazu bedarf es aber mehr als nur irgendeines Erstgliedes, nämlich eines, das identisch mit dem jeweiligen Zweitglied ist. Nur dadurch kann die Prototypenlesart angezeigt werden. Insofern sind Prot-ICCs im Gegensatz zu klassifikatorischen Komposita wie *Betazelle* oder *Röntgenstrahlen* auch als Beispiele für kreativen Sprachgebrauch anzusehen.[38] Sie stellen eine Form der bewussten Manipulation sprachlicher Form im Alltagsdiskurs dar. Prot-ICCs zeigen ähnlich wie Wortspiele oder Metaphern das selbstreferenzielle Potenzial und damit Jakobsons (1960) „poetische Funktion" der Sprache. „[They] focus on the message for its own sake" (Jakobson 1960: 356). Prot-ICCs lenken die Aufmerksamkeit auf sich, steigern die

38 In der Forschung zu kreativer Sprache finden insbesondere Aspekte der Wiederholung Beachtung (beispielsweise bei Carter 1999, 2004, 2007 sowie Tannen 1989).

Erfahrung und das Vergnügen der Gesprächsteilnehmer:innen und rufen evaluative und ästhetische Urteile hervor. Aus diesem Grund sind Prot-ICCs selbst also mitnichten neutral. Zudem ist Prototypikalität sehr schwer zu definieren. Es gibt also keine eindeutig richtige Antwort auf die Frage, was denn nun der 'eigentliche, richtige' *Winter*, *Mann*, *Film* ist, von dem aus man neutral subklassifiziert. Prot-ICCs werden deshalb sehr häufig auch wertend verwendet und mit der Festlegung, was ein Prot-ICC bedeuten soll (*Buchbuch* 'klassische Literatur, kein Groschenroman' versus 'Buch für die gemütliche Lektüre, keine klassische Literatur') sagen Sprecher:innen immer auch etwas über sich selbst aus. Der Begriff der neutralen Subklassifizierung muss also im Sinne des von mir ebenfalls vorgeschlagenen Terminus „Nicht-Modifikation" verstanden werden, dergestalt, dass im Zuge einer Prot-ICC-Bildung ein Basislexem ohne Zuhilfenahme eines weiteren nominalen Konzeptes subklassifiziert wird.

Prot-ICCs haben noch mehr Gemeinsamkeiten mit anderen Kompositatypen und Konstruktionen. Teilweise überlappt sich die Funktion der Prot-ICC-Bildung beispielsweise mit der indefiniten N-*von*-N-Konstruktion. Die Bildungen, beispielsweise *ein Bilderbuch von einem Mann* oder *ein Engel von einer Frau* (Foolen 2004), entsprechen in der Denotatseinschränkung und Expressivität bisweilen den Prot-ICC-Bildungen *Mannmann* und *Fraufrau.* Allerdings referiert die indefinite N-*von*-N-Konstruktion nicht immer auf einen Prototypen, etwa bei den Phrasen *dieses Nest von einer Hauptstadt* oder *so ein Apparat von Karton* (Foolen 2004: 79). Prot-ICCs hingegen schreiben dem Referenten immer die üblichen Eigenschaften des Basiskonzeptes zu.

Zudem erfüllt die Konstruktion Prot-ICC mit der Zuschreibung prototypischer, üblicher Eigenschaften gewissermaßen spiegelbildlich die entgegengesetzte Funktion der für das Englische beschriebenen Stereotypennegation (Andreou 2017, Horn 2002, 2005, Zimmer & Carson 2018). Bei der Stereotypennegation zeigen Affixe wie das nominale *un-* im Derivat *unpolitician* an, dass das Bezeichnete zwar Teil der Kategorie des Basislexems ist, ihm aber wichtige Attribute fehlen.[39] Ein *unpolitician* ist demnach ein Politiker ohne Führungskraft, Wahlkampftaktiken oder Korruptionanfälligkeit. Auch im Deutschen gibt es das nominale Präfix *Un-*, das eine meist negative, pejorative Wertumkehr des Basislexemes hervorruft (*Unkosten* 'schlechte Kosten', *Unmensch* 'schlechter Mensch'). Darüber hinaus lassen sich aber auch solche Wortbildungen finden, in denen wie im Englischen Stereotypennegation vorliegt, in denen also der Referent des Basislexems als untypischer Vertreter gekennzeichnet wird.

39 Andreou (2017) nimmt an, dass diese Attribute nicht vollständig fehlen, sondern nur ein Wertwechsel stattgefunden hat.

(84) *Warum viele Ägypter in al-Baradei ihren Retter sehen, ist unschwer zu begreifen: Er kommt von außen und ist nicht durch irgendwelche Kompromisse und Absprachen an die politischen Parteien gebunden. […] Der Journalist Issandr al-Amrani bezeichnet ihn deshalb als* ***„Unpolitiker“****.*

<Le Monde diplomatique vom 09.07.2010, https://www.monde-diplomatique.de/pm/2010/07/09.mondeText.artikel,a0010.idx,0>

Bei der Bildung von Prot-ICCs wie *Oma-Oma* geschieht nun genau das Gegenteil. Es wird die Abwesenheit von Eigenschaften angezeigt, die das Konzept in Richtung anderer Konzepte verschieben würden. Auf diese Weise verengen Prot-ICCs den Bedeutungsfokus der Basis, reduzieren das Denotat auf das Prototypische und beschränken so die Referenzmöglichkeiten auf prototypische Vertreter.

Es gibt aber auch Konstruktionen, die nicht nur in bestimmten Fällen mit Prot-ICCs bedeutungsgleich sind oder nur eine ähnliche Funktion erfüllen, sondern nahezu vollständig deckungsgleich mit der Semantik von Prot-ICCs sind. So sind Phrasen aus einem Nomen und einem Suffixderivat auf *-mäßig*, in denen das Kopfnomen der NP identisch mit der Basis des Derivats ist, von der Bedeutung her identisch mit den entsprechenden Prot-ICCs. Die Phrase *omamäßige Oma* ist etwa das phrasale Äquivalent zum Prot-ICC *Oma-Oma.* Auch A+N-Phrasen mit anderen denominalen Adjektiven (etwa solchen auf *-haft* oder *-typisch*) können eine solche Funktion ausüben:

(85) a. *Sie ist so eine* ***omamäßige Oma****.*

b. *Das ist so ein* ***albumhaftes Album****.*

c. *Er ist so ein* ***manntypischer Mann****.*

Die Suffixe *-typisch*, *-haft und -mäßig* fungieren in (85) jeweils als syntaktisches Verbindungsstück (Wellmann 1998: 554) und setzen zwei Nomina „weitgehend ohne semantische Anreicherung“ (Fleischer & Barz 2012: 308) in Beziehung.

Man kann Produktivität von Wortbildungsprozessen aus einer onomasiologischen Perspektive betrachten (Štekauer 2000) und hinsichtlich der ICC-Bildung fragen: Warum wählt ein:e Sprecher:in die morphologische, warum die phrasale Realisierung? Welche Funktion erfüllt ein Prot-ICC, wenn es doch eine scheinbar bedeutungsgleiche Konstruktion bereits gibt? In beiden Konstruktionen wird das Konzept des nominalen Kopfes auf den prototypischen Kern reduziert. In beiden Fällen muss auf Basis des Weltwissens spezifiziert werden, wie genau die Bedeutungskomponente ‘typisch, richtig’ für das entsprechende Nomen ausgestaltet ist.

Um die Frage zu beantworten, welchen Mehrwert Prot-ICCs gegenüber Phrasen wie denen in (85) bieten, muss man sich den generellen Unterschied zwischen Phrasen und Komposita anschauen. Die Funktion einer A+N-Phrase wie *omamäßige Oma* kann als eine NP beschrieben werden, in der das Denotat des Kopfnomens restriktiv qualitativ modifiziert wird. Eine phrasale Bildung hat die qualitative Modifikation als Defaultfall, bildet also kein Subkonzept. Es besteht nun aber mitunter aus diversen Gründen für die Sprecher:innen die Notwendigkeit, eine morphologische Bildung zu verwenden. Zum einen ist eine morphologische Bildung vonnöten, wenn zur Verdichtung, etwa im Nominalstil, (weitere) phrasale (Adjektiv-)Attribute vermieden werden sollen („form compression", Schlücker & Hüning 2009: 224). Zum anderen benötigen Sprecher:innen eine morphologische Bildung dann, wenn sie die defaultmäßig den Komposita zukommende Klassifikationsfunktion nutzen wollen. Für eine morphologische Bildung stehen Adjektive auf *-mäßig*, *-typisch* oder *-haft* aber nicht zur Verfügung, da A+N-Komposita im Deutschen generell der Bildungsbeschränkung unterliegen, dass morphologisch komplexe Erstglieder unzulässig sind (Schlücker 2012: 9).[40]

(86) a. *Sie ist so eine* ****Omamäßigoma***.

b. *Das ist so ein* ****Albumhaftalbum***.

c. *Er ist so ein* ****Manntypischmann***.

A+N-Komposita wie diese sind also wegen der genannten Bildungsrestriktionen nicht möglich. Prot-ICCs ermöglichen nun, die semantisch-funktionalen Eigenschaften der A+N-Phrase mit den syntaktischen und semantisch-funktionalen Eigenschaften der A+N-Komposita zu verbinden. Prot-ICCs kommt als Komposita defaultmäßig die Funktion der Subkonzeptbildung zu. Dennoch können sie, wie die genannten A+N-Phrasen, dem Kopfnomen Eigenschaften zuschreiben. Zudem ist die Position des pränominalen attributiven Adjektivs frei und kann für eine weitere Beschreibung der Referenten genutzt werden. Die Daten aus den Korpusstudien konnten zeigen, dass die Sprecher:innen von diesem Vorteil der Prot-ICC-Konstruktion Gebrauch machen, denn in vielen Belegen werden Prot-ICCs zusätzlich adjektivisch modifiziert.

40 Zu dieser Restriktion gibt es einige Ausnahmen. Nicht-native Relationsadjektive etwa können trotz ihrer morphologischen Komplexität A+N-Komposita bilden (*Regionalliga*, *Nuklearkatastrophe*). Zudem gibt es in der Gegenwartssprache die Tendenz, dass die Bildungsbeschränkung auch abseits von Relationsadjektiven aufgehoben wird (*Endloslockdown*, *Schnörkellossieg*) und A+N-Komposita mit komplexen Erstgliedern möglich werden (Schlücker 2012: 9, 2014: 29ff.).

(87) a. *Sie ist so eine alte **Oma-Oma**.*

b. *Das ist so ein schönes **Albumalbum**.*

c. *Er ist so ein machomäßiger **Mannmann**.*

Möglich wird diese qualitative, morphologische Modifikation durch die Prototypenlesart, die durch das Kompositionserstglied entsteht. Prot-ICCs sind zwar klassifikatorische Komposita, die funktionale, nicht auf nominale Konzepte verweisende Erstglieder haben. Doch haben die Erstglieder eine Ähnlichkeit zu qualitativen Modifikatoren, insofern sie Attribute des Referenten des Zweitgliedes spezifizieren.

Diese Annahme zu Prot-ICCs wirkt zunächst widersprüchlich. Klassifikatorische Komposita sind Komposita mit klassifikatorischen Modifikatoren, also Modifikatoren, die Subkonzepte zum Kopfkonzept bilden. Qualitative Modifikatoren hingegen schreiben dem Modifikanden Eigenschaften zu. In Prot-ICCs spezifiziert die Modifikation Eigenschaften des Konzeptes oder des Referenten des Zweitgliedes, ist allerdings restriktiv, beeinflusst also das Denotat der Basis, wodurch die Bildungen klassifikatorisch sind. Anders ausgedrückt: Prot-ICCs verwenden – wie etwa auch die augmentativ-evaluativen Komposita (Schlücker 2013: 458) – die primäre Klassifikationsfunktion der N+N-Komposita parasitär. Durch die Klassifikationsfunktion, die N+N-Komposita defaultmäßig zukommt, können die Prot-ICCs Subkonzepte bilden, obwohl (beziehungsweise während) das Erstglied dem Konzept oder dem Referenten der Kopfkonstituente Eigenschaften zuweist. A+N-Phrasen mit denominalen Adjektiven vermögen das nicht, weil sie defaultmäßig nicht klassifizieren, sondern qualitativ modifizieren.

Diese Interpretation der Erstglieder in Prot-ICCs ist nicht im Einklang mit der klassischen, kategorialen Einteilung der Modifikationsrelationen, die Teyssier (1968) für die Modifikation in Nominalphrasen annimmt. Teyssier betont gerade den Unterschied zwischen identification (ID), qualification (QUAL) und classification (CLASS). Dass eine klassifikatorische Funktion der Erstglieder in Komposita andere Modifikatorfunktionen nicht notwenigerweise ausschließt, ist allerdings bekannt. Koptjevskaja-Tamm (2013: 273) bespricht beispielsweise Komposita, die referenzspezifizierende (identifizierende) Modifikatoren haben, aber gleichzeitig klassifikatorisch sind, etwa *Hitler-Bärtchen.* Obwohl das Erstglied auf eine konkrete Person referiert, ist das Kompositum dennoch klassifikatorisch. Ein weiteres Beispiel hierzu wäre etwa *China-Hacker* (Schlücker 2013: 462). Die klassifikatorische Funktion eines Modifikators kann in solchen Fällen nicht immer eindeutig von der identifizierenden Funktion unterschieden werden. Insofern sind Prot-

ICCs nur ein weiterer Hinweis darauf, dass die Klassifikationsfunktion zwar die grundlegende, nicht aber die einzige Funktion der Modifikatoren in nominalen Komposita ist.

Prot-ICCs haben aber neben der Funktion, Prototypeneigenschaften mithilfe morphologischer Bildungen zuzuschreiben, ein weiteres Alleinstellungsmerkmal gegenüber Phrasen wie in (85). Die Subklassenbildung geschieht auf sehr auffällige Art und Weise. Mit Prot-ICCs weisen die Sprecher:innen darauf hin, dass kein unbestimmter Bereich des Denotats eines Lexems gemeint ist, sondern ein ganz konkreter, dergestalt, dass das Konzept bloß subklassifiziert, aber nicht durch ein weiteres Konzept verschoben wird. Dabei leisten Prot-ICCs etwas, was Bauer (2000) für alle spielerischen und innovativen Wortbildungsprodukte annimmt:

> [P]layful formations [...] and some literary creations [...] may go beyond the bounds of normal rules specifically to gain effect.
>
> (Bauer 2000: 838)

Wenn die Sprecher:innen des Deutschen diesen auffälligen, extravaganten Ausdruck der phrasalen Variante vorziehen, mag das auch eben dieser Auffälligkeit geschuldet sein (Frankowsky 2022). In einer Phrase wie *omamäßige Oma* sind die beiden identischen Stämme nicht adjazent. Das Prot-ICC *Oma-Oma* hingegen rückt die Bildung durch die im Deutschen sehr auffällige reduplikative Struktur in den Mittelpunkt. Sprecher:innen wählen außergewöhnliche Ausdrücke genau dann, wenn auch eine außergewöhnliche, nicht-typische Situation beschrieben wird (Levinson 2000: 136). Prot-ICCs sind außergewöhnliche Ausdrücke, die auf die neutrale Subklassifizierung aufmerksam machen, also anzeigen, dass das Denotat eines Lexems wider Erwarten nicht in Hinblick auf ein weiteres Konzept zu modifizieren ist.

Die hier vorgenommene Analyse der Prot-ICCs macht auch die Entscheidung obsolet, ob ein konkretes Prot-ICC eine Prototypenbedeutung (Hohenhaus 2004), eine auf Kategorienidentität basierende Bedeutung (Freywald 2015) oder eine auf Kontrast basierende Bedeutung (Ghomeshi et al. 2004) hat. Kategorienidentität ist neben anderen Eigenschaften wie der Form oder der Größe nur eine – wenn auch sehr wichtige – Eigenschaft, die über die Nähe und Distanz zum prototypischen Kern entscheidet. Eine Blume, die tatsächlich ein Vertreter der Kategorie BLUME ist, besitzt durch diese Kategorienidentität Eigenschaften, durch die sie näher an den abstrakten Prototypen heranrückt, etwa die Eigenschaft der Belebtheit. Man kann also die Kategorienidentität nur als einen von vielen Aspekten der Prototypikalität ansehen. Auch die Annahme, dass bei der Verwendung von Prot-ICCs stets ein Kontrast (im Sinne einer Gegenüberstellung) zwischen zwei Entitäten bestehe, ist bereits im Begriff des Prototypen enthalten. Ein guter Vertreter einer

Klasse ist genau deshalb ein guter Vertreter der Klasse, weil andere Fälle schlechtere Vertreter der Klasse sind. Ohne die Existenz schlechterer Vertreter – also Subkonzepten, die sich vom Grundkonzept entfernt haben – können Prot-ICCs also nicht auf einen Prototypen verweisen. Die Daten meiner Korpusstudien konnten zeigen, dass die schlechteren Vertreter im Kotext häufig genannt werden. Letztlich gehen aber alle funktionalen Beschreibungen der Prot-ICCs (Prototyp, Kategorienidentität, Kontrast) davon aus, dass die Dopplung in *Oma-Oma* auf eine Oma verweist, die, im Gegensatz zu anderen möglichen Referenten, bestimmte Eigenschaften aufweist. Diese Eigenschaften werden durch das reduplikative Muster eines Prot-ICCs betont.

Prot-ICCs erfüllen damit eine wichtige Funktion. Sie klassifizieren Entitäten mithilfe des Weltwissens der Sprecher:innen. In Kommunikationssituationen ist es eine wichtige Funktion, Referenten Prototypikalität zu- oder abzusprechen und somit anzuzeigen, dass ein Referent voll und ganz dem entspricht, was man im Allgemeinen mit dem Konzept der Basis verbindet. So zeichnen sich soziale Gruppen unter anderem dadurch aus, dass ihre Mitglieder eine gemeinsame soziale Identität besitzen und im Zuge der Identitätsbildung Gruppenmitglieder als prototypisch oder weniger prototypisch gekennzeichnet werden. Die Sozialwissenschaft nennt solche Zuschreibungen „Gruppenprototypen“ (Hogg et al. 1993). Der Prototyp ist bei jedem Gruppenmitglied kognitiv repräsentiert und das Ergebnis gruppeninterner Synchronisation.[41] Mit Prot-ICCs steht den Sprecher:innen ein Wortbildungsmuster zur Verfügung, mit dem sie Prototypikalität in Form von klassifikatorischen, morphologischen Bildungen zuweisen können.

Die dritte und letzte Funktion von ICCs, zu der die Korpusstudien Belege liefern konnten, ist die der Mono- und Direktreferenz. Zwei identische Nominalstämme bilden dazu eine mophologische Verbindung, bei der kein Bestandteil der Bildung eine Semantik hat, über die die Referenz zur Entität zustande kommt. Diese ICCs sind Eigennamen und werden Name-ICCs genannt. Die Wortkörper dieser ICCs evozieren allesamt als Gesamtkomposita kein Konzept zur Herstellung von Referenz. Stattdessen werden sie zur Mono- und Direktreferenz auf eine Entität in der Welt genutzt (Heusinger 2012: 425, Nübling et al. 2015: 17). Meist referieren Name-ICCs „ohne Umweg über die Semantik“ (Nübling et al. 2015: 27), sind also „ohne inhärente Bedeutung“ (Leys 1989: 150) und dienen der Identifizierung und Individualisierung von Referenten. In diesem Name-ICC-Subtyp wird der Eigenname, wie von Kentner (2017) beschrieben, allein zur formalen Erweiterung und Variation verdoppelt. Dies ist ein Unterschied zur Verdopplung von Anthro-

41 Giessner et al. (2014) zeigen das etwa am Beispiel von Personalmanagement in der Wirtschaft, Zepp et al. (2013) am Beispiel von Fußballmannschaften.

ponymen etwa im Englischen, wo häufig die Kurzformen von Personennamen verdoppelt werden (Aziz & Nolikasari 2020: 46), was pragmatische Funktionen ausübt, etwa den Ausdruck von Nähe (*Lou-Lou*).

Im Deutschen haben solche ICCs also vor allem die Funktion, das sprachliche Material eines Eigennamens zu erweitern. Diese Funktion ist in vielerlei Hinsicht wichtig. Bei der Erstellung von Benutzernamen für Online-Plattformen besteht beispielsweise der Druck, ein Pseudonym zu kreieren, das noch nicht von anderen Nutzer:innen verwendet wird. Da auf großen Plattformen aber sehr viele Benutzer:innen denselben Vornamen haben, bedarf es oft einer formalen Erweiterung (Dürscheid 2005: 48, Kentner 2017: 244). Die hier beschriebenen Daten zeigen, dass aus dieser Not heraus unter anderem ICCs entstehen. Angesichts der anhaltenden Entwicklung, dass sich Nutzer:innen von Online-Plattformen auf immer weniger Plattformen konzentrieren, wird es schwieriger, einen nicht bereits vergebenen Nutzernamen zu erstellen, was dazu führt, dass bei der Erstellung von Nutzernamen auch vermehrt auf Name-ICCs mit Eigennamenbasis zurückgegriffen wird. Auf YouTube kann man beispielsweise jeden nur erdenklichen Vornamen zweimal hintereinander ins Suchfeld eingeben und findet einen entsprechenden Kanal oder Benutzer dieser Bildung, etwa *Sabine Sabine* (Abbildung 38).

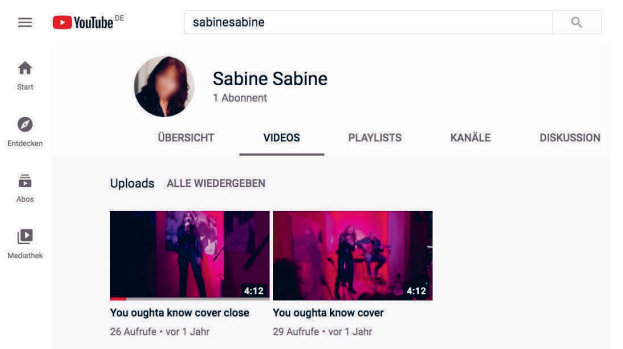

Abb. 38: Das Name-ICC *Sabine Sabine* als Eigenname für einen YouTube-Kanal.

Optional kann die Basis eines solchen ICCs den Referenten in Bezug auf ein Merkmal spezifizieren. In *AutoAuto* wird der Name einer Show beispielsweise dahingehend motiviert, dass Autos in der Show eine zentrale Rolle einnehmen. In anderen Name-ICCs ist das Verhältnis zwischen der Basis und dem Referenten hingegen arbiträr. *PunktPunkt* beispielsweise referiert auf ein DJ-Duo, für das das Konzept PUNKT irrelevant ist. Name-ICCs kommt somit die Funktion zu, (optional teildeskriptive) Eigennamen auf der Basis von Appellativa zu bilden.

Die Frage nach der Funktion von ICCs kann also mit Rückgriff auf die drei ICC-Typen beantwortet werden. Erstens können ICCs die Funktion ausüben, die Determinativkomposita im Allgemeinen erfüllen, nämlich die Modifikation eines Konzeptes mithilfe eines weiteren Konzeptes. Zweitens können ICCs die Funktion ausüben, eine Nicht-Modifikation, beziehungsweise eine neutrale Subklassifikation des Basiskonzeptes anzuzeigen, wodurch die Vagheit des betreffenden Konzeptes reduziert und das Vorliegen der für das Konzept üblichen Merkmale bestätigt wird. Drittens können ICCs die Funktion ausüben, Appellative in Eigennamen zu überführen, die optional teilmotiviert sind oder das sprachliche Material von Eigennamen erweitern.

6.2 Forschungsfragen zur Produktivität von ICCs

Zwei Fragen betreffen die Produktivität der ICC-Subtypen. Die erste lautet:

3. Welche Art von ICC wird besonders häufig verwendet?

Die zwei Korpusstudien liefern Evidenz dafür, dass Komposita mit identischen Konstituenten abseits der wenigen lexikalisierten Fälle wie *Kindeskind* oder *Baden-Baden* existieren, ICCs also neu gebildet werden. Besonders häufig werden Name-ICCs gebildet. Sie machen in beiden Korpusstudien jeweils den Großteil der Daten aus. Die zweite Forschungsfrage, die die Produktivität von ICCs betrifft, kann mit Rückgriff auf die drei ICC-Typen beantwortet werden und lautet:

4. Wie produktiv sind die jeweiligen ICC-Arten; also wie geneigt sind Sprecher:innen, neue Bildungen auf der Basis bestehender ICCs zu bilden?

Generell sind die meisten ICCs nicht lexikalisiert und ein großer Teil der erhobenen Daten besteht zudem aus Hapax legomena. Das spricht dafür, dass die Schemata sehr produktiv zur Wortbildung verwendet werden. Die meisten nur einmal belegten Stämme sind Prot-ICCs. Vor allem bei den Hapax-Belegen besteht ein großer Unterschied zwischen Prot-ICCs und den anderen ICC-Typen. Während mehr als ein Viertel aller Prot-ICCs Hapax legomena sind, sind es bei Det-ICCs nur 4%, bei Name-ICCs weniger als 1%. Dazu passt das Ergebnis, dass die Stämme im Prot-ICC-Schema relativ ausgewogen auftreten. Es gibt zwar einige Stämme, die besonders häufig als Prot-ICC verwendet werden, etwa *Film* oder *Chef*, das Gros der Prot-ICC-Belege verteilt sich aber auf viele unterschiedliche Stämme. Die Konstruktion Prot-ICC bringt also in Relation zu den Prot-ICC-Gesamtbelegen mehr unterschiedliche Typen hervor als die Konstruktionen Det- und Name-ICC zu ihren je-

weiligen Gesamtbelegen. Wenn ein Wortbildungsprozess viele Typen hervorbringt, die jeweils aber nur wenige Belege liefern, spricht das für eine hohe Produktivität des Prozesses (Baayen 1989, 1992, 2009). Die vielen Hapax legomena bei den Prot-ICCs sind also ein Hinweis auf eine hohe allgemeine Anwendbarkeit des Schemas.

6.3 Forschungsfragen zu lexikalischer Struktur und Selektionsbeschränkungen

Forschungsfrage 5 betrifft die lexikalische Struktur des ICC-Schemas:

5. Können Lexeme identifiziert werden, die besonders häufig als ICC auftreten?

Zunächst einmal ist festzuhalten, dass nahezu jedes Nomen die Basis für ein ICCs sein kann. Wie in Tabelle 12 aufgeführt, gibt es allein in den deTenTen-Daten über 1500 Nomina, die Name-ICCs bilden, über 1000 Nomina, die Det-ICCs bilden und immerhin fast 500 Nomina, die Prot-ICCs bilden. Die Suche nach semantischen Beschränkungen ist deshalb nicht einfach. Auf den ersten Blick stechen zwar vor allem Menschenbezeichnungen ins Auge, und zwar sowohl bei den Det-ICCs (*Helfershelfer*, *Kindeskind*, *Freundesfreund*, *Kandidatenkandidat*) als auch bei den Prot-ICCs (*Chefchef*, *Künstlerkünstler, Mädchenmädchen*, *Mannmann*). Allerdings finden sich Nomina vieler anderer semantischer Klassen unter den Basislexemen der beiden ICC-Typen: Abstrakta (Det-ICC: *Zinseszins*, *Kompetenzenkompetenz, Hungerhunger*; Prot-ICC: *Mitte-Mitte* 'genau die Mitte', *Zeit-Zeit* 'Qualitätszeit', *Ende-Ende* 'Bereich nach dem Abspann eines Films'), zählbare, nicht-belebte Konkreta (Det-ICC: *Tütentüte* 'Tüte für Tüten', *Ballball* 'Ball, den man auf Bälle wirft', *Backup-Backup* 'eine Festplatte mit einem Backup von einem Backup'; Prot-ICC: *Busbus* 'kein Shuttlebus', *PizzaPizza* 'keine Dönerpizza', *Haushaus* 'kein Doppelhaus/Ökohaus'), Kollektiva/Stoffsubstantive (Det-ICC: *Mengenmenge* 'Menge verschiedener Mengen', *Kohlekohle* 'Kohle, die zum Abbau von Kohle benötigt wird', *Gesellschaftsgesellschaft* 'Gesellschaft aus Gesellschaften'; Prot-ICC: *Farbe-Farbe* 'Primärfarbe', *Müll-Müll* 'kein Trash', *Papier-Papier* 'Dollarnote ohne Wert/Schreibpapier, kein Klopapier').

Bei den Name-ICCs ist die Frage nach semantischen Selektionsbeschränkungen natürlich ein wenig anders zu stellen, da Namen keine Semantik haben und also das Ergebnis der Name-ICC-Bildung stets direkt auf eine Entität in der Welt verweist. Die direkt referierten Entitäten sind hier vor allem Personen, Produkte, Institutionen und Unternehmen. Die Basen selbst weisen aber ebenso wie die Basen von Det- und Prot-ICCs eine große semantische Vielfalt auf und lassen sich den Menschenbezeichnungen (*Mannesmann*, *Clown-Clown*, *Kindkind*) und ande-

ren belebten Nomina (*Walwal*, *Hasehase*, *Hummelhummel*), den nicht-belebten zählbaren Konkreta (*Bongobongo*, *Ballball*, *Eimereimer*), den Kollektiva/Stoffsubstantiven (*Gasgas*, *Pastapasta*, *Nussnuss*) sowie den Abstrakta (*Urlauburlaub*, *Geschichtegeschichte*, *Fitnessfitness*) zuordnen. Es ist also bei keinem der drei ICC-Typen ohne weiteres möglich, zu bestimmen, welche Nomina nicht als Basis auftreten können.

Auf die Frage, welche Basen in den ICCs besonders häufig auftreten, liefern die Daten aus dem deTenTen13-Korpus aber eine Antwort. Bei den Det-ICCs treten, wie nicht anders zu erwarten war, die Basen der lexikalisierten ICCs, also *Zins*, *Helfer* und *Kind* sowie die Basen von in bestimmten Sprecher:innengruppen lexikalisierten ICCs, nämlich *Kompetenz* und *Erwartung*, besonders häufig auf. Die genannten Stämme machen zusammen 88% der Det-ICC-Belege im deTenTen13 aus. Bei den Prot-ICCs sind die häufigsten Basislexeme *Chef*, *Film*, *Künstler*, *Mädchen* und *Literatur*. Sie bilden zusammen 30% der Prot-ICC-Tokens im deTenTen13. Im Gegensatz zu den frequentesten Det-ICCs sind diese ICCs nicht lexikalisiert.[42] Die häufigsten Name-ICC-Basen sind *Baden*, *Tom*, *Mann*, *Zauberer* und *Clown*. Hier können nur die ICCs *Baden-Baden* und *TomTom* als lexikalisiert gelten.

Forschungsfrage 6 betrifft mögliche Selektionsbeschränkungen der ICCs:

6. Welche Eigenschaften haben die Basen, die besonders häufig als ICC auftreten?

Semantische Selektionsbeschränkungen für ICCs zu definieren ist, wie oben erläutert, schwierig. Etwas einfacher ist es, die Frequenz und die Form betreffende Präferenzen der ICC-Konstruktionen zu beschreiben. Ein Ergebnis der Korpusstudien ist, dass ICC-Bildung vorwiegend auf der Basis hochfrequenter Lexeme geschieht. Die erste Korpusanalyse konnte mit Rückgriff auf zwei Frequenzdatenbanken Evidenz dafür liefern, dass zwischen den Basen, die ICCs bilden, und denen, die keine ICCs bilden, ein Unterschied von mehreren Frequenzklassen liegt. Interessant ist dieses Ergebnis vor allem vor dem Hintergrund der im Forschungsüberblick beschriebenen Präferenz von Nominalstämmen für die Kopf- beziehungsweise Modifikatorposition. Brunner et al. (2021) konnten nachweisen, dass hochfrequente Nomina in einem mehr oder weniger ausgeglichenen Verhältnis in Erst- und Zweitgliedposition vorkommen. Für die Bildung eines ICCs ist

42 Für die Einschätzung, ob ein ICC lexikalisiert ist, wurde in dieser Arbeit auf die Onlineversion des Duden zurückgegriffen (https://www.duden.de/). Die vier frequentesten Det-ICCs *Zinseszins*, *Helfershelfer*, *Kindeskind* und *Kompetenzkompetenz* sind hier verzeichnet; ebenso die Name-ICCs *Baden-Baden* und *TomTom*. Demgegenüber gibt es zu keinem der Prot-ICCs einen Eintrag.

das förderlich. Den Sprecher:innen ist ein hochfrequentes Nomen nicht nur generell vertrauter als ein niedrigfrequentes; es ist, dem Ergebnis von Brunner et al. (2021) nach, auch gleichermaßen als Erstglied wie als Zweitglied bekannt. Es eignet sich dadurch besser für ein Kompositum, in dem es gleichzeitig Erst- und Zweitglied realisiert, als ein Nomen, das eine klare Präferenz für eine der beiden Positionen hat und den Sprecher:innen deshalb in der jeweils anderen Position unbekannt ist. Die Frequenzeffekte, die Brunner et al. (2021) nachweisen konnten, liefern also eine Erklärung dafür, warum hochfrequente Nomina in den auf DECOW16 basierenden Korpusdaten derart dominieren.

Abseits der Frequenzeffekte konnte die auf DECOW16 basierende Korpusstudie zudem Evidenz dafür liefern, dass ICC-bildende Stämme im Mittel weniger komplex sind als solche, die keine ICCs bilden, und weniger Zeichen, Silben und Morpheme aufweisen. Dies ist angesichts der Tatsache, dass das sprachliche Material in ICCs verdoppelt wird, nicht weiter verwunderlich. Es steht dem sprachlichen Ökonomieprinzip entgegen, allzu lange und komplexe Stämme für die Bildung von ICCs zusätzlich noch zu verdoppeln. Es ist aber wohl nicht angemessen, entsprechende morphosyntaktische oder phonologische Beschränkungen für ICCs zu formulieren, da auch Det-ICCs wie *Zusammenfassungs-Zusammenfassungen*, Prot-ICCs wie *Vergewaltigung-Vergewaltigung* oder Name-ICCs wie *Beobachterbeobachter* in den abgefragten Korpora vorkommen. Dennoch sprechen die erhobenen Daten dafür, dass weniger komplexe Lexeme eine größere Wahrscheinlichkeit haben, in ICCs verwendet zu werden.

6.4 Forschungsfragen zu formalen Merkmalen der ICC-Typen

Die morphologischen (Distinktions-)Merkmale betreffend gibt es folgende Forschungsfragen:

7. Weichen ICCs hinsichtlich der Verwendung von Fugenelementen von kanonischen Komposita ab?
8. Verhalten sich ICCs hinsichtlich der Verwendung von Fugenelementen einheitlich oder sind Fugenelemente ein formales Mittel, die jeweilige Semantik der ICCs anzuzeigen?
9. Weichen ICCs hinsichtlich der internen Flexion und lexikalischen Integrität von kanonischen Komposita ab?
10. Verhalten sich ICCs hinsichtlich der Flexion einheitlich oder ist das Flexionsverhalten ein formales Mittel, die jeweilige Semantik der ICCs anzuzeigen?

Eine Erkenntnis aus den beiden Korpusstudien ist, dass die drei unterschiedlichen Funktionen, die ICCs ausüben, mit jeweils unterschiedlichen morphologischen Merkmalen der Bildungen einhergehen. Die Ergebnisse der beiden Korpusstudien zu diesen formalen Unterschieden entsprechen sich weitgehend. Die auf dem begrenzten Stammset der DECOW16-Studie basierenden Ergebnisse hinsichtlich Morphologie und Schreibung konnten durch die auf deTenTen13 basierende Studie bestätigt werden.

In vielen Aspekten sind Det-ICCs und Name-ICCs Antipoden, etwa in Bezug auf die Länge. Det-ICCs sind im Schnitt knapp 4 Zeichen länger als Name-ICCs. Teilweise, aber nicht ausschließlich, lässt sich dieser Längenunterschied mit dem Anteil der durch Flexionsmarker und Fugenelemente erweiterten Stämme und Basisstämme erklären. Det-ICCs zeigen Fugenelemente und Flexionsmarker und damit eine Ähnlichkeit mit kanonischen N+N-Komposita. Verglichen mit diesen treten die Marker bei Det-ICCs sogar sehr viel häufiger auf. Gut die Hälfte der Det-ICCs ist overt flektiert. Name-ICCs zeigen hingegen nahezu ausnahmslos Nullflexion – eine Eigenschaft, die für Namen im Allgemeinen charakteristisch ist (Nowak & Nübling 2017, Nübling 2012). Det-ICCs tragen zudem in 90% der Fälle Fugenelemente. Auch das ist sehr viel mehr als bei N+N-Komposita im Allgemeinen, welche nur zu knapp einem Drittel verfugt sind (zu 35% bei Nübling & Szczepaniak 2009, zu 26,5% bei Wellmann 1975). Im Gegensatz dazu zeigen Name-ICCs vornehmlich Nullfugen. Det-ICC-Stämme haben also eine deutliche Tendenz, durch Flexive und Fugenelemente erweitert zu werden. Name-ICC-Stämme hingegen werden nicht erweitert, weder durch Fugenelemente am Erstglied noch durch Flexionsmarker am Zweitglied. Sie bestehen somit in fast allen Fällen aus bloßen, identischen Nominalstämmen. Für all diese die Länge und die formalen Marker an Erst- und Zweitglied betreffenden Eigenschaften der ICC-Typen liefern beide Korpusstudien Evidenz.

Diese Unterschiede zwischen Det- und Name-ICCs können als Zeichen für die Substantivklassen dieser beiden ICC-Typen gewertet werden: Det-ICCs sind Appellativa. Name-ICCs sind Namen. Dass sich die Konstituenten von Det-ICCs uneingeschränkt und häufig mit Fugenelementen und Flexionsmarkern verbinden, führt zu einer Verfremdung der Stämme. Die bisherige Forschung zu ICCs und anderen Kompositionsphänomenen kann einen Hinweis darauf geben, was der Grund für diese häufige Verfremdung ist: die identische phonologische Struktur der drei ICC-Schemata. Angesichts der Koexistenz von Det-ICCs und Prot-ICCs wurde zum einen dem Flexionsverhalten (Bross & Fraser 2020: 3, Freywald 2015: 925) und zum anderen den Fugenelementen (Finkbeiner 2014: 189) die Funktion zugesprochen, die Prototypenlesart zu hemmen und die determinative Lesart zu fördern. In ähnlicher Weise wurde diese diskriminierende Funktion von Fugenelementen im Deutschen schon für die Disambiguierung von Kopulativkomposita beschrieben

(Becker 1992: 12, Neef & Borgwaldt 2012: 32). Neef und Borgwaldt (2012: 32) gehen beispielsweise davon aus, dass ein verfugtes Kompositum wie *Komponistenmaler* ausschließlich die Bedeutung ‘eine Person, die einen Komponisten malt’ tragen kann, während das unverfugte Äquivalent *Komponist-Maler* auf ‘eine Person, die Komponist und Maler gleichzeitig ist’ verweist. Auch bei Det-ICCs besteht die Notwendigkeit, die determinative Lesart auf diese Weise zu forcieren. Der Gebrauch von Fugenelementen und Flexionsmarkern hilft, die Prototypen- und Eigennamenlesarten von ICCs auszuschließen.

Name-ICCs verhalten sich hinsichtlich der Morphologie genau entgegengesetzt. Ihr flexionsmorphologisches Verhalten kann man als Minimalflexion oder Nullflexion bezeichnen. Sie sind Eigennamen und gehören damit zu einer Substantivklasse, die sich im Deutschen deutlich in Richtung Deflexion bewegt (Nübling 2012). Die Vermeidung von Stammerweiterungen lässt sich mit dem Prinzip zur Schonung des Wortkörpers erklären. Demnach werden stammaffizierende Elemente wie Umlaute oder Flexionssuffixe von Sprecher:innen vermieden, um den Stamm nicht zu verfremden. Das ist zum einen wichtig, um den Adressat:innen die Worterkennung nicht zu erschweren, etwa wenn die Wörter für die Hörer:innen einen niedrigen Bekanntheitsgrad haben. Zum anderen aber spielt die Schonung des Wortkörpers eine wichtige Rolle in der Grammatik der Eigennamen (Ackermann 2016, Nübling 2005, Nübling & Schmuck 2010, Schmuck 2017). Marker aller Art werden demnach bei Eigennamen vermieden, um die Stabilität des Namenskörpers zu gewährleisten. Unter anderem deshalb treten etwa in Eigennamenkomposita keine Fugenelemente auf (Schlücker 2017). Zudem gehört, etwa bei Personen- und Ortsnamen, „der Name […] zur Identität der Person, des Ortes“ (Lüssy 1974: 189). Name-ICCs zeigen sich in diesen morphologischen Aspekten als die Eigennamen, die sie sind. Sie werden nicht overt flektiert, um den zur Identität einer Person, einer Firma oder eines Produkts gehörigen Namen nicht durch Verfremdung zu beschädigen.

Neben der Schonung des Wortkörpers verweist das grammatische Genus von Name-ICCs auf eine weitere morphologische Besonderheit, die Name-ICCs mit Eigennamen im Allgemeinen teilen. Det-ICCs (ebenso Prot-ICCs) folgen ausnahmslos dem Letztgliedprinzip der Genuszuweisung, nach dem das Genus eines Kompositums durch seine letzte Konstituente, den grammatischen Kopf, bestimmt wird (Köpcke & Zubin 2009: 139). Bei Name-ICCs hingegen besteht in manchen Fällen ein Konflikt zwischen dem morpholexikalischen Genus der ICC-Basis und dem Klassengenus der Referenzdomäne. Das Genus wird hier mitunter, wie für Eigennamen charakteristisch, referenziell zugewiesen, also in Abhängigkeit von der Objektklasse, auf die die Bildung verweist (Fahlbusch & Nübling 2014: 73ff., Köpcke & Zubin 2009: 144). *Das Tiger Tiger* etwa ist Neutrum, da die Referenzdo-

mäne (Klasse der Nachtclubs / Restaurants) das Neutrum zuweist. In den meisten Fällen wird das Genus von Name-ICCs aber gar nicht erst sichtbar, weil Targets wie attributive Adjektive oder Determinierer, die das Genus anzeigen würden, bei Name-ICCs nur sehr selten auftreten. Auch das ist ein Merkmal, das Eigennamen im Allgemeinen zukommt, und wird am Ende dieses Kapitels genauer behandelt.

Man kann also festhalten, dass sowohl der Verwendung von Fugenelementen als auch den Flexionsmarkern die Funktion zukommt, die drei ICC-Typen voneinander zu unterscheiden. Dies betrifft vornehmlich Det- und Name-ICCs, da diese zwei ICC-Typen sich hinsichtlich der formalen Marker gegensätzlich verhalten.

Die letzten beiden Forschungsfragen zu formalen Merkmalen betreffen die Schreibung der ICCs:

11. Weichen ICCs hinsichtlich der Schreibung von kanonischen Komposita ab?
12. Verhalten sich die ICCs hinsichtlich der Schreibung einheitlich oder ist die Schreibung ein formales Mittel, die jeweilige Semantik der ICCs anzuzeigen?

Zunächst einmal ist festzuhalten, dass ICCs generell viele Schreibvarianten aufweisen. Die Schreiber:innen machen regen Gebrauch von Bindestrichen, Majuskeln und Leerzeichen. Dies sei am Beispiel der mit *Mann* gebildeten ICCs illustriert: Hier gesellen sich zu der für Komposita regelhaften Schreibung *Mannmann* (ungeachtet der unterschiedlich flektierten und verfugten Formen) die Varianten *mann Mann*, *Mann-Mann*, *MANN-Mann*, *Mann Mann*, *MannMann* und die Zusammenschrebung *mannmann.*[43] Ob ein solcher Variantenreichtum nun aber die ICCs besonders betrifft oder online verwendeten Ad-hoc-Komposita im Allgemeinen zukommt, kann an dieser Stelle nicht entschieden werden.

Hinsichtlich der Unterscheidung der drei ICC-Typen kommt der Schreibung zudem eine diskriminierende Funktion zu, denn auch in Bezug auf die Schreibung stehen Det-ICCs und Name-ICCs in Kontrast zueinander. Det-ICCs werden in den Datensätzen überwiegend (DECOW16), beziehungsweise fast ausschließlich (deTenTen13) zusammengeschrieben. Nur in 7,9% der Fälle werden sie in den deTenTen13-Daten mit einem Bindestrich geschrieben. Das entspricht in etwa dem Anteil, zu dem N+N-Komposita auch in der lektorierten Schriftsprache einen Bindestrich aufweisen (7,6%, Grube 1976: 209, Kopf 2017) und ist nur geringfügig

43 Nimmt man die unterschiedlich flektierten und verfugten Formen hinzu, ergeben sich natürlich deutlich mehr Varianten. Die mit *Buch* gebildeten ICCs treten beispielsweise in 14 Formen auf, nämlich *Buch-Buch*, *Buch-Bücher*, *Buch-Buchs*, *Bücher Bücher*, *Bücher-Bücher*, *Bücher-Buches*, *Buchbuch*, *BuchBuch*, *Buchbuches*, *Bücherbuch*, *BücherBuch*, *Bücherbücher*, *Bücherbüchern*, *Bücherbuches.*

höher als in von Sprecher:innen ad hoc gebildeten N+N-Komposita (5,8%, Borgwaldt 2013).[44] Det-ICCs weichen also in der Regel nicht von der normgerechten Zusammenschreibung ab und verhalten sich wie kanonische N+N-Komposita.

Bei Name-ICCs dagegen zeigt sich das für Eigennamen generell charakteristische Merkmal, dass sie sich „prinzipiell der Orthographie entziehen“ (Kempf et al. 2020: 2) und abweichenden Gebrauch von Syngraphemen aufweisen (Nübling et al. 2015: 89). In beiden Datensätzen werden bei über 70% der Name-ICCs die Konstituenten getrennt, und zwar vornehmlich durch den Bindestrich und das Spatium. Hier kann man eine Brücke zu den Ergebnissen zur Verfugung schlagen. Denn Spatium, Bindestrich und Binnenmajuskel nehmen in Name- (und auch Prot-) ICCs wortintern eine graphische Grenzmarkierung vor, die bei den Det-ICCs die Fugenelemente übernehmen. Sowohl Fugenelemente als auch Bindestriche und Spatien gliedern Komposita (Kopf 2018: 349). Doch während Fugenelemente gliedern, ohne die graphematische Einheit des Wortes zu stören, gliedern Spatium und Bindestrich gerade durch die deutlichere Konturierung der Konstituenten. Bei den Name-ICCs trägt die Getrenntschreibung, ähnlich wie die Nullfuge und die Nullflexion, dem Schonungsbedarf der Eigennamen Rechnung (Vogel 2017). Die Bildungen werden gegliedert, ohne dass weiteres sprachliches Material in Form von Fugenelementen die Sichtbarkeit der Konstituenten erschwert. Der Wortkörper der einzelnen Konstituenten wird graphisch geschont und sie sind direkt als solche erkennbar (Fuhrhop 2008: 202ff.). Auch auf der Ebene der Schreibung zeigen Name-ICCs also überdeutlich, dass sie keine Appellativa, sondern Eigennamen sind.

Spatium, Bindestrich und Binnenmajuskel führen aber außerdem zu einer Abweichung von der normgerechten Kompositaschreibung. Dies entspricht zum einen der Eigenschaft von Eigennamen, die generell zu orthographischen Abweichungen tendieren. Als Eigennamen, von denen „graphische Devianz […] fast erwartet“ wird (Kempf et al. 2020: 2), werden Name-ICCs zusätzlich durch die Getrenntschreibung als etwas Besonderes markiert. Der Bindestrichschreibung etwa wird zugeschrieben, generell bei markierten Strukturen aufzutreten (Buchmann 2015: 236ff.). Durch die Abweichung von der normgerechten Schreibung zeigen sich Name-ICCs somit als extravagante (Haspelmath 1999), kreative Bildungen (Frankowsky & Schlücker [angenommen]), bleiben dadurch eher im Gedächtnis und erfüllen auch damit Ansprüche, denen Eigennamen genügen müssen.

Auch den Sprecher:innen ist bewusst, dass ICCs, und hier vor allem Name- und Prot-ICCs, extravagante Bildungen sind. Dass sich die ICC-Typen diesbezüglich aber graduell unterscheiden, zeigt ein weiterer Aspekt der Schreibung. Auch das

44 In Borgwaldts (2013) schriftlichem Produktionsexperiment war den Proband:innen aufgetragen, bestehende Substantive zu neuen Komposita zu kombinieren.

Setzen von Anführungszeichen geschieht bei Det- und Name-ICCs unterschiedlich häufig. Nomina im Allgemeinen werden im DECOW16-Korpus in weniger als 1% der Fälle in Anführungszeichen gesetzt. Name-ICCs weisen hier mit 11% einen ungleich höheren Anteil auf. Det-ICCs (6,8%) und Prot-ICCs (8,3%) liegen allerdings nicht allzu weit dahinter. Generell können Anführungszeichen, abseits direkter Zitate, auch pragmatische Funktionen übernehmen. So werden sie mitunter als „Abschirmeinrichtung“ (Zienkowski 2014: 515) bezeichnet, die dazu dient, den oder die Sprecher:in vor etwaigen riskanten Behauptungen oder anderen Sprechakten zu schützen, die Konversationsmaximen verletzen könnten (Zienkowski 2014: 515). Anführungszeichen sind zudem als Indikatoren metapragmatischen Bewusstseins Teil der expliziten Metasprache (Verschueren 2012: 60f.). Insofern ist es nicht verwunderlich, dass Komposita mit einer für das Deutsche sehr ungewöhnlichen reduplikativen Struktur (siehe Kapitel 8), bei der Erst- und Zweitglied von ein und demselben Stamm realisiert wird, zur Rechtfertigung vor Kommunikationspartner:innen häufig als deviant gekennzeichnet werden. Der Unterschied in der Verwendung von Anführungszeichen, der zwischen den einzelnen ICC-Typen besteht, lässt sich zum einen damit erklären, dass die zu rechfertigende reduplikative Struktur bei Det-ICCs durch die Verwendung von Fugenelementen gebrochen, beziehungsweise abgeschwächt wird. Zum anderen kann man vermuten, dass Det-ICCs als gewöhnliche Determinativkomposita (Vorhandensein diverser semantischer Relationen, Vorliegen zweier nominale Konzepte, Klassifikationsfunktion) auch funktional weniger abweichend sind als Prot- und Name-ICCs und deshalb seltener explizit metasprachlich behandelt werden müssen. Auch die Schreibung ist also ein formales Mittel, die jeweilige Semantik der ICCs anzuzeigen. Auch hier betrifft dies vor allem Det- und Name-ICCs, da vor allem die Schreibungen dieser zwei ICC-Typen stark voneinander abweichen.

Die Ergebnisse zu den formalen Merkmalen von ICCs zeigen ein Phänomen besonders deutlich: Det- und Name-ICCs stehen sich konträr gegenüber, während sich Prot-ICCs in vielen der bisher genannten Punkte in gewisser Weise zwischen Det- und Name-ICCs verorten. Det-ICCs sind besonders lang, Name-ICCs besonders kurz. Prot-ICCs stehen bei der durchschnittlichen Länge der Bildungen genau zwischen Det- und Name-ICCs. Det-ICCs sind nahezu immer verfugt, Name-ICCs fast nie. Prot-ICCs sind hingegen mal verfugt, mal nicht. Det-ICCs zeigen fast immer Flexionsmarker am Zweitglied, Name-ICCs bleiben unflektiert. Prot-ICCs tragen zumindest in knapp einem Viertel der Fälle Marker an Erst- und/oder Zweitglied. Auch in der Schreibung nehmen Prot-ICCs eine Mittelposition ein. Sie werden seltener als Det-ICCs, aber häufiger als Name-ICCs zusammengeschrieben.

6.5 Forschungsfragen zu Syntax und Kontextabhängigkeit von ICCs

Zu syntaktischen Unterschieden zwischen den ICC-Typen wurden nicht a priori Forschungsfragen formuliert; dennoch sind die Ergebnisse der Korpusstudien hierzu interessant und geben Aufschluss über die syntaktischen Merkmale von Det-, Prot- und Name-ICCs. Die wichtigsten Vergleichsaspekte sind hier das Auftreten mit definiten und indefiniten Artikeln, die Negation, die adjektivische Modifikation sowie das Auftreten in Prädikativkonstruktionen. Bei all diesen Vergleichsaspekten stechen, anders als bei den am Substantiv selbst markierten formalen Merkmalen, die Prot-ICCs deutlich heraus. Diese Aspekte betreffen die Forschungsfragen zur Kontextabhängigkeit von ICCs und zeigen Prot-ICCs in einer exponierten Stellung:

13. Muss die Verwendung von ICCs immer durch den Kontext lizensiert werden und wenn ja, in welchem Umfang?
14. Mit welchen kontextuellen Mitteln unterstützen die Sprecher:innen die Adressat:innen bei der Interpretation der ICCs?

Ein Vergleichsaspekt zur Syntax ist die Häufigkeit, zu der die ICCs mit einem (adjazent pränominalen) Artikel verwendet werden. Prot-ICCs sind der ICC-Typ, der am häufigsten mit einem solchen Artikel steht. In knapp der Hälfte aller Belege geht den Prot-ICCs ein Artikel direkt voran und damit deutlich öfter als das bei Det-ICCs der Fall ist. Auch ist bei Prot-ICCs der indefinite Artikel sehr viel häufiger als bei den anderen ICC-Typen und als bei N+N-Komposita im Allgemeinen. Det-ICCs stehen hier wiederum häufiger mit einem Artikel als Name-ICCs und treten darüber hinaus sowohl mit definiten als auch mit indefiniten Artikeln auf. Name-ICCs dagegen haben nur in etwas mehr als einem Fünftel der Fälle einen Artikel bei sich, nur sehr selten ist dieser Artikel indefinit.

Diese Beobachtung passt sich in das Bild ein, dass Name-ICCs zur Substantivklasse der Eigennamen gehören. Sie haben Monoreferenz, identifizieren also stets ein konkretes Objekt, sind inhärent definit und ergo mit dem indefiniten Artikel inkompatibel (Nübling et al. 2015: 100). Name-ICCs treten zwar, wenn auch seltener als bei Det- und Prot-ICCs, mit dem definiten Artikel auf, der hier wie bei Eigennamen generell redundant ist (Heusinger 2010: 89, ^ übling 2011: 110). Dies ist aber kein ungewöhnliches Verhalten für Eigennamen. Nur bestimmte Namenklassen im Deutschen, etwa Städte- oder Ländernamen, benötigen keinen Definitartikel (Nübling et al. 2015: 79). Zudem kann der Definitartikel bei der Verwendung mit Eigennamen verwendet werden, wenn der Referent kontrastiv

präsentiert wird (Heusinger 2010: 89) oder der Definitartikel bestimmte Funktionen übernimmt und beispielsweise Distanz oder Emotionalität (*Ich kann diesen Schultze einfach nicht leiden!*) ausgedrückt werden soll (Heusinger 2012: 463, Kolde 1995: 406, Lakoff 1974, Nübling et al. 2015: 79ff.). Der geringe Anteil definiter Artikel bei Name-ICCs spricht also nicht gegen, sondern für die Auffassung, dass sie typische Vertreter der Eigennamen sind. Selbst die Verwendung des Indefinitartikels ist bei Eigennamen nicht unmöglich, sondern lediglich wiederum mit ganz konkreten Funktionen verbunden (Kolde 1992, 1995). So kann ein indefiniter Artikel auch in Kombination mit einem Eigennamen Nicht-Einzigkeit anzeigen (*Selbst ein Mozart hätte das nicht viel besser machen können*) und einen neuen Diskursreferenten einführen, der „anders als bei dem demonstrativen Gebrauch unbekannt und nicht salient ist" (Heusinger 2010: 89). Im Allgemeinen ist der Indefinitartikel aber bei Eigennamen selten. Diese Eigenschaft teilen Name-ICCs offenbar mit anderen Arten von Eigennamen.

Die häufige Verwendung vor allem indefiniter Artikel bei Prot-ICCs passt zu der folgenden Beobachtung: Prot-ICCs sind häufiger (im DECOW16 in 12% der Fälle) Teil einer Prädikativkonstruktion als die anderen ICC-Typen. Das Ergebnis zur häufigen Verwendung des Indefinitartikels und die Tendenz der Prot-ICCs zur Prädikativkonstruktion passen insofern zueinander, als der indefinite Artikel in Prädikativkonstruktionen sehr häufig, beziehungsweise der Regelfall ist (Eisenberg 2006: 463, Engel 1996: 528, Erben 1988: 309). Diese beiden Aspekte zusammengenommen kann man also bei Prot-ICCs eine erhöhte Tendenz feststellen, dass sie in der Konstruktion *Y ist/bleibt/wird ein XX* auftreten. Weil Eigennamen generell, „da per se referentiell, nicht […] prädikativ auftreten" (Nübling et al. 2015: 83) können, sind auch die Name-ICCs der ICC-Typ, der am seltensten (zu 5%) in Prädikativkonstruktionen auftritt. Auch hierzu passt das zuvor beschriebene Ausbleiben indefiniter Artikel bei Name-ICCs.

Des Weiteren sind Prot-ICCs in Bezug auf die adjektivische Modifikation auffällig. Die ICC-Typen treten unterschiedlich häufig mit attributiven Adjektiven auf. Bei Name-ICCs zeigt sich hier einmal mehr der Eigennamencharakter dieser Bildungen. Üblicherweise können Eigennamen nicht restriktiv modifiziert werden, da sie ja bereits „auf einen Gegenstand referieren [und also] nicht spezifiziert werden muss, auf welchen Gegenstand der Ausdruck zutrifft" (Sturm 2005: 74). Dies schränkt die Verwendung attributiver Adjektive bei Eigennamen ein und resultiert bei Name-ICCs darin, dass sie nur in 3,7% der Fälle mit einem (unmittelbar vorangehenden) attributiven Adjektiv auftreten. Prot-ICCs werden hingegen recht häufig, nämlich in über 20% der Fälle, durch attributive Adjektive modifiziert. Dieses Ergebnis spricht somit deutlich gegen die Annahme, nach der adjektivische Modifikation bei Prot-ICCs ausgeschlossen sei (zum Englischen: Benjamin

2018: 33, zum Deutschen: Bross & Fraser 2020). Zur Hälfte werden Prot-ICCs durch Real Intensifications (*richtig*, *echt*) modifiziert und angereichert (vgl. Bross & Fraser 2020: 6), zur Hälfte durch Adjektive, die nicht zu den Real Intensifications zählen (*groß*, *selbstverliebt*). Die von Bross und Fraser (2020) auf der Basis eines englischsprachigen Korpus formulierte Beschränkung ist offensichtlich im Deutschen nicht gegeben.

Bei Prot-ICCs wird nicht nur die adjektivische Modifikation zur kontextuellen Anreicherung verwendet. Sie werden zudem knapp dreimal so häufig wie Nomina im Allgemeinen, Det- und Name-ICCs mithilfe von *kein* oder *nicht* in einen kontrastiven Kontext gestellt. Generell lässt sich also beobachten, dass Prot-ICCs syntaktisch am meisten von den beiden anderen ICC-Typen abweichen.

Die Beobachtung, dass Prot-ICCs eine morphologische und graphematische Zwischenposition einnehmen, syntaktisch aber herausstechen, interpretiere ich wie folgt: Prot-ICCs haben durch die beschriebene morphologische und graphematische Mittelstellung kein direkt am Wort markiertes formales Mittel, sich als Prototypenkomposita zu profilieren. Ihre formalen Eigenschaften sind keine eindeutigen Hinweise darauf, dass sie mit einer Prototypenlesart zu interpretieren sind. Sie greifen stattdessen auf syntaktische und kontextuelle Mittel zurück. Die qualitative Analyse der Prot-ICC-Belege zeigt weitere Strategien, wie Prot-ICCs kontextuell angereichert werden. Mithilfe von Attributsätzen, die häufig mit *also* eingeleitet sind, wird bei einem Drittel aller Prot-ICCs die Prototypenlesart explizit gemacht. Diese Strategien entsprechen weitgehend denen, die Widlitzki (2016) – ebenfalls korpusbasiert – für Prot-ICCs und andere CFRs im Englischen beschreibt. Das Bestreben der Sprecher:innen, Prot-ICCs kontextuell anzureichern, spricht für die These, dass Prot-ICCs unterspezifiziert und stark vom Kontext abhängig sind (Hohenhaus 1998, 2015). Prot-ICCs greifen besonders auf syntaktische und kontextuelle Mittel zurück (vermehrter Artikelgebrauch, häufiger Einsatz von attributiven Adjektiven und Negationswörtern, Verwendung von disambiguierenden (Neben-)sätzen). Doch kann kontextuelle Anreicherung nicht die alleinige Antwort auf die Frage sein, wie Prot-ICCs verstanden werden. Mehrere Studien sprechen dafür, dass Sprecher:innen auch ohne Kontext übereinstimmend eine der verschiedenen ICC-Schemata auswählen (Finkbeiner 2014, Günther 1979). Deshalb muss eine Antwort auf die letzte Forschungsfrage gefunden werden:

15. Welche weiteren sprachlichen Mittel gewährleisten, dass die Adressat:innen ein ICC korrekt interpretieren?

Eine mögliche Antwort auf diese Frage besteht in der Annahme, dass sich ICC-Bildungen gegenseitig blockieren. Die Ergebnisse der auf deTenTen13 basierenden Korpusstudie zeigen, dass Stämme nur selten in mehr als einem ICC-Schema ver-

wendet werden. Diesen Umstand kann man allerdings nicht ohne Weiteres damit erklären, dass sich ICCs direkt gegenseitig blockieren, denn die meisten ICCs sind nicht lexikalisiert und ein großer Teil der erhobenen Daten besteht aus Hapax legomena. Die klare Aufteilung der Stämme zu den drei ICC-Schemata kann also nicht durch etwaige ICC-Einträge im mentalen Lexikon erklärt werden. Deshalb kann die Frage, wie Prot-ICCs von den Sprecher/Hörer:innen als solche erkannt werden, anhand der präsentierten Ergebnisse nicht zufriedenstellend beantwortet werden. Im weiteren Verlauf der Arbeit werden hierzu aber Hypothesen aufgestellt. Alle bisher behandelten Vergleichsaspekte, die Unterschiede zwischen den ICC-Typen hervorgebracht haben, werden in Tabelle 14 noch einmal zusammengefasst.

Zu Beginn des nun folgenden dritten Teils dieser Arbeit wird zunächst auf der Grundlage der besprochenen formalen und funktionalen Ergebnisse die Frage behandelt, was für ein sprachlich-grammatischer Prozess überhaupt vorliegt, wenn ein ICC gebildet wird. Dabei wird sich herausstellen, dass dies keine triviale Frage ist. Det-ICCs wie *Zinseszins* und *Freundesfreund* wird zwar einhellig zugeschrieben, das Ergebnis eines Kompositionsprozesses zu sein, doch bei Prot-ICCs und Name-ICCs besteht hierüber keine Einigkeit. Das folgende Kapitel 7 bespricht die vorherrschende Einordnung von ICCs als Komposita.

Tab. 14: Vergleich ICC-Typen zu Frequenz, Komplexität, Morphologie, Syntax, Schreibung

Vergleichsaspekt		**Det-ICCs**	**Prot-ICCs**	**Name-ICCs**
Frequenz	Frequenz der ICCs	mittel***	niedrig***	hoch***
	Frequenz der Basen	hoch*	mittel*	niedrig*
Komplexität	Zeichenanzahl der Belege	12,2**	11,1**	9,8**
	Zeichenanzahl der Wortformen	14,4**	12,5**	10,6**
Morphologie	Verfugung	90 %**	13 %**	8 %**
	(overte) Flexion	43 %**	21 %**	3 %**
	Genus	lexikalisch*	lexikalisch*	lexikalisch/ referenziell*
Syntax	Adjektivische Modifikation	8,1%* (vor allem *sogenannt*)	20,1%* (vor allem *richtig*)	3,7%* (vor allem *lieb*)
	Artikelverwendung	35 %* (74 %* definit)	48 %* (65 %* definit)	22 %* (93 %* defnit)
	Präpositionen	*auf*, *über**	*als*, *für**	*von**
	Negationswörter	3 %*	10 %*	2 %*
	Verwendung als Prädikativum	6 %*	12 %*	5 %*
Schreibung	Graphische Trennung der Konstituenten	8 %**	55 %**	63 %**
	Anführungszeichen	6,8 %*	8,3 %*	11 %*

* = Vergleich basierend auf DECOW16-Daten

** = Vergleich basierend auf deTenTen13-Daten

*** = Vergleich basierend auf DECOW16- und deTenTen13-Daten

Teil III: **Theorie zu ICCs im Deutschen**

7 ICCs als Komposition

Wie im Forschungsüberblick dargestellt, wird ICC-Bildung mitunter nicht als Komposition, sondern als Reduplikationsprozess angesehen. Mit den Ergebnissen aus den Korpusstudien lässt sich zu dieser Kontroverse nun eine Aussage tätigen. In diesem Kapitel wird zunächst Komposition als ein sprachtypologisch weit verbreiteter morphologischer Prozess dargestellt (7.1), ehe die Definitionskriterien bestimmt werden, die für Komposition grundlegend sind (7.2). Zum Abschluss des Kapitels wird mithilfe dieser Kriterien eine Einschätzung dazu abgegeben, ob ICCs Komposita sind (7.3).

7.1 Komposition in den Sprachen der Welt

Komposition ist die Bildung eines neuen Wortes durch die Verbindung mindestens zweier bereits vorhandener Wörter bzw. Wortstämme. Übereinzelsprachlich gibt es eine ungeheure Vielfalt an Phänomenen, für die in der wissenschaftlichen Beschreibung der Begriff „Komposition“ verwendet wird. Es ist sehr schwierig, wenn nicht unmöglich, eine universale Definition für Komposition zu formulieren, die für alle Einzelsprachen gleichermaßen anwendbar ist (Dressler 2006: 24, Lieber & Štekauer 2009: 4). Dressler nennt zwei Gründe dafür:

> [U]niversal definitions [of compounding] are not only theory-dependent […] but also cross-linguistically never watertight – in many languages there are exceptions or fuzzy transitions to non-compounding.
>
> (Dressler 2006: 24)

Mit der genannten Abhängigkeit einer Kompositionsdefinition vom jeweils verwendeten linguistischen Framework benennt Dressler bloß eine Tatsache, die es beim Verfassen linguistischer Forschungstexte zu berücksichtigen gilt. Ein wirkliches theoretisches und praktisches Problem ist dahingegen die Beobachtung, dass die Stämme oder Wörter, die verbunden werden, in jeder Einzelsprache andere grammatische Eigenschaften haben, unterschiedlichen syntaktischen Kategorien angehören und die Kompositionsprozesse sich in jeder Einzelsprache mit anderen Phänomenen überschneiden. Der Begriff „Komposition“ wird für unterschiedliche Phänomene aus morphologisch sehr unterschiedlichen Sprachen verwendet:

https://doi.org/10.1515/9783111315416-007

(88) Englisch (Schlücker 2020: 28)
red *wine*
rot Wein
'Rotwein / roter Wein'

(89) Russisch (Kazakovskaya 2017: 66)
город герой
gorod *geroy*
Stadt Held
'Heldenstadt'

(90) Japanisch (Kageyama 2009: 516)
欧 米
oo- *bei*
Europa Amerika
'der (politische) Westen'

(91) Mandarin Chinesisch
电 脑
diàn *nǎo*
Elektrizität Gehirn
'Computer'

Alle Beispiele zeigen zweigliedrige Nominalkomposita. Die Beispiele reflektieren also nicht einmal das volle Spektrum dessen, was unter dem Begriff „Komposition" behandelt wird, sondern bloß die Prozesse, bei denen Nomina gebildet werden. Und obwohl der Bereich, dem die Beispiele angehören, schon sehr eingegrenzt ist, gibt es hinsichtlich vieler Merkmale bedeutende Unterschiede. In manchen Sprachen werden im Zuge des Prozesses Wörter, in manchen Sprachen hingegen Stämme oder Wurzeln verbunden. Im Chinesischen (91) sind die Bestandteile gebundene, lexikalische Wurzelmorpheme, im Englischen (88) und Russischen (89) hingegen freie lexikalische Morpheme. Auch die interne Struktur der Bildungen unterscheidet sich von Einzelsprache zu Einzelsprache. In den Beispielen zum Englischen (88), Japanischen (90) und Chinesischen (91) ist der semantische Kopf rechts, im Russischen (89) hingegen links.

Ein großer Unterschied zwischen den Sprachen, aus denen die aufgeführten Beispiele stammen, betrifft außerdem die Rolle, die der Morphologie zukommt. Das Deutsche und das Russische haben ein umfangreiches Kasus- und Genussystem, wodurch die Form der Konstituenten variabel ist. Im Genitiv Singular kom-

men im Russischen Beispiel etwa Endungen hinzu (город**а**-геро**я**). Die Komposita lassen sich auf diese Weise auch flexionsmorphologisch von Phrasen unterscheiden. Ähnlich ist es im Deutschen. Hier ist *Rotwein* eindeutig Kompositum und benennt das Konzept ROTWEIN. Die entsprechende Phrase *roter Wein* hingegen kann beschreibend auch für Weine verwendet werden, die eine rote Farbe haben, aber keine Rotweine sind, beispielsweise für Portwein.

Im Englischen ist eine solch klare Einteilung auf der Basis formaler Merkmale nicht ohne weiteres möglich. Wegen der weitgehenden Formgleichheit von Kompositum und Phrase ist hier oft nicht eindeutig zu erkennen, ob es sich bei einer Bildung um eine syntaktische Bildung (Phrase) oder eine morphologische (Kompositum) handelt. In der Literatur wird zwar das Betonungsmuster der Bildungen zurate gezogen, dergestalt, dass Komposita im Englischen an der Erstgliedbetonung zu erkennen seien, Phrasen daran, dass das zweite Glied betont wird (Bloomfield 1933, Marchand 1969). In der neueren Literatur zu dem Thema wird dieses Kriterium aber kritisch gesehen, da letztlich nicht ohne Weiteres vorhersagbar ist, in welchen Fällen englische Bildungen Erst- und wann Zweitgliedbetonung erhalten (Bauer 1998, Schmerling 1971). Für ein verlässliches Modell müssen immer auch andere Faktoren berücksichtigt werden, etwa die Syntax, Semantik und Analogie (Giegerich 2004, 2015, Plag 2006).

Auch im Chinesischen lässt sich in Ermangelung flexionsmorphologischer Prozesse die Art der Bildung nicht eindeutig bestimmen. Hier kommt der Komposition zudem eine ganz andere Funktion zu als im Deutschen, Englischen und Russischen. Die Komposition hat im Chinesischen vornehmlich die Funktion, silbenphonologische Defizite auszugleichen. Die Funktion, neue Wörter zu bilden, die in den anderen genannten Sprachen vorherrschend ist, steht im Hintergrund (Li & Thompson 1981: 14).

Die Grenze, die Komposita von Derivaten und Phrasen trennt, verläuft also in allen Sprachen anders und unterschiedlich deutlich. Selbst in Sprachen, die mithilfe formaler Merkmale eine recht eindeutige Einteilung in Komposita und Phrasen erlauben, gibt es produktive Prozesse, deren Ergebnisse sich diesbezüglich nicht eindeutig zuweisen lassen. Die Unterscheidung zwischen Komposita und Phrasen muss deshalb eher graduell als kategorial verstanden werden (Schlücker 2020). Übereinzelsprachlich bleiben zudem immer Unterschiede zwischen den Kompositionsprozessen. Deshalb muss jede Basisdefinition von Komposition für die Untersuchung einer Einzelsprache modifiziert werden. An dieser Stelle soll darum nicht der unmögliche Versuch unternommen werden, Komposition sprachübergreifend eindeutig und detailliert zu definieren. Stattdessen werden ICCs zu einer allgemeinen Basisdefinition in Bezug gesetzt, die für alle Kompositionsprozesse gilt. Im Detail betrifft die weitere Bezugnahme dann aber auch die

Merkmale, die speziell für N+N-Komposita im Deutschen gelten. Inwieweit diese Kriterien in anderen Sprachen angewandt werden (können), spielt dabei keine Rolle.

7.2 Definitionskriterien für Komposition (im Deutschen)

Viele Merkmale der (N+N-)Komposition wurden bereits in Kapitel 2.1 behandelt. Doch nur einige dieser Merkmale eignen sich auch als Definitionskriterien für Komposita im Allgemeinen. Generell sind Komposita lexikalische Einheiten, die aus der Verbindung von mindestens zwei lexikalischen Morphemen bestehen. Diese Annahmen finden sich so oder so ähnlich in den meisten Versuchen, Komposita zu definieren (Elsen 2014: 293, Payne 1997: 93). Bauer definiert Komposita wie folgt:

> A compound is the formation of a new lexeme by adjoining two or more lexemes.
> (Bauer 2003: 40)

Ein Wort wie *Grünfink* ist also ein Kompositum, weil es aus den beiden Lexemen *grün* und *Fink* besteht. Sowohl *Grünfink*, als auch die Konstituenten *grün* und *Fink* haben eine eindeutige Bedeutung, eine gewisse Gebrauchsfrequenz, sind in Lexika verzeichnet und können somit als Lexeme gelten. Für das Deutsche muss man noch ergänzen, dass die beiden Konstituenten lexikalische Stämme sind. Konfixe wie *-thek* in *Bibliothek* oder *Spielothek* haben zwar eine lexikalische Bedeutung, stellen aber keine Wortstämme dar.[45] Die folgende Definition ist hinsichtlich der Stämme genauer:

> We can now define a compound as a lexical unit made up of two or more elements, each of which can function as a lexeme independent of the other(s) in other contexts, and which shows some phonological and/or grammatical isolation from normal syntactic usage.
> (Bauer 2001a: 695)

Die Definition von Bauer (2001a) ist allerdings noch immer nicht zufriedenstellend, da Lexeme genauso gut in Form von Phrasen gebildet werden können. *Schwarze Mamba* etwa bezeichnet wie *Schwarznatter* eine Schlange. Beide Ausdrücke sind lexikalische Einheiten (lexical units) und beide zeigen zusätzlich eine

45 Ein Sonderfall der Komposition ist im Deutschen zudem die Konfixkomposition, bei der zwar lexikalische Einheiten in Form von Konfixen zusammengefügt werden, dabei aber gebundene Einheiten, etwa *pädagog-*, entstehen. Manche der Konfixe „im engeren Sinne" sind „auf das Vorkommen in Konfixkomposita spezialisiert" (Eisenberg 2006: 242).

gewisse grammatische Isolation, also Unterschiede zum rein syntaktischen Gebrauch. Zwischen den beiden Bestandteilen dürfen in beiden Fällen beispielsweise keine anderen Wörter stehen (**die Schwarzgefährlichnatter ≠ die gefährliche Schwarznatter*; *die schwarze, gefährliche Mamba ≠ die gefährliche schwarze Mamba*).

Dass es sich bei *Schwarze Mamba* um eine Phrase handelt, wird spätestens in Kontexten deutlich, in denen kontextuelle Flexion gefordert ist. Während das Erstglied der Phrasen in (92) Flexive trägt, die sich je nach syntaktischem Kontext unterscheiden, bleibt es bei *Schwarznatter* in (93) unflektiert und nur das Zweitglied zeigt Flexive kontextueller Flexion. Die lexikalische Integrität bleibt gewahrt und *Schwarznatter* ist also ein Kompositum.

(92) a. *Die **Schwarz<u>e</u> Mamba** ist die längste Giftschlange Afrikas.*

b. *Die Zahl der **Schwarz<u>en</u> Mambas** sinkt bedrohlich.*

(93) a. *Die **Schwarznatter** gehört zu den harmlosen Schlangen in Nordamerika*

b. *Der Zahl der **Schwarznatter<u>n</u>** sinkt bedrohlich*

Gaeta und Ricca schlagen deshalb eine Kombination des Merkmals [±lexikalisch] mit dem Merkmal [±morphologisch] vor (Gaeta & Ricca 2009). In beiden Fällen *Schwarznatter* und *Schwarze Mamba* liegt zwar das Merkmal [+lexikalisch] vor, da beide Ausdrücke im Lexikon verzeichnet sind. *Schwarznatter* ist allerdings morphologisch gebildet, *Schwarze Mamba* nicht. Um Komposita von lexikalischen Phrasen abzugrenzen, muss also noch das morphologische Kriterium hinzugezogen werden. Zudem sind es im Deutschen Wortstämme, die Komposita bilden.

Das Kriterium [±lexikalisch] ist hier nicht misszuverstehen als ‘aus lexikalischen Einheiten bestehend’. Es gibt Zusammensetzungen, die für gewöhnlich als Komposita klassifiziert werden, bei denen ein Element aber keine lexikalische Bedeutung trägt, etwa die in Kapitel 2 besprochenen Komposita mit Eigennamen als Erstglied (*Reus-Freistoß*). Im Ansatz von Gaeta und Ricca ist eine Zusammensetzung aber dann [+lexikalisch], wenn der Referent stabil ist, eine einheitliche Bedeutung hat und eine nicht zu vernachlässigende Frequenz vorliegt (Gaeta & Ricca 2009: 39). [+lexikalisch] heißt hier also, dass der Prozess lexikalische Einheiten hervorbringt und also der Erweiterung des Lexikons dient. Ich nehme darüber hinaus und in Anlehnung an die oben genannten Definitionen von Bauer an, dass es eine grundlegende Eigenschaft von Komposita ist, aus lexikalischen Einheiten zu bestehen. Eigennamenkomposita wie *Reus-Freistoß* entsprechen diesem

Kriterium weniger als Komposita wie *Schneeball*, die keinen Eigennamen beinhalten. Beide Komposita, *Reus-Freistoß* und *Schneeball*, sind aber morphologische Zusammensetzungen.

Der übrigen, in Kapitel 2.1 näher besprochenen Eigenschaften von N+N-Komposita eingedenk gibt es nun also folgende (Definitions-)merkmale für Komposition im Deutschen:

Die Form betreffend:

A. Die Bildung ist morphologisch: Es gibt keine wortinterne Flexion. (definitorisch)
B. Die Konstituenten sind Wortstämme. (definitorisch)
C. Die rechte Konstituente ist der grammatische Kopf. (definitorisch)
D. Die linke Konstituente ist betont.
E. Der Prozess ist rekursiv.
F. Die Bildung wird zusammengeschrieben.

Die Bedeutung betreffend:

G. Das Kompositum und die Konstituenten tragen lexikalische Bedeutung. (definitorisch)
H. Die rechte Konstituente ist der semantische Kopf.
I. Das Kompositum bezeichnet eine Subklasse zum Konzept des Zweitgliedes.
J. Zwischen Erst- und Zweitglied besteht eine semantische Relation.

Die Definitionskriterien müssen unterschiedlich gewichtet werden. Wenn ein Wortbildungsprozess (grammatisch) rechtsköpfige, morphologische Bildungen aus lexikalischen Wortstämmen bildet (A, B, C, G), sind alle Definitionsmerkmale von Komposition im Deutschen gegeben und stellen in gewisser Weise also die Minimalanforderungen an Komposita im Deutschen dar. Dass die linke der beiden Komponenten betont ist (D), die Komponenten in der Schreibung nicht getrennt werden (F) oder dem Kompositum eine Klassifikationsfunktion zukommt (I), ist im Vergleich dazu weniger grundlegend. In Kapitel 2 wurden etwa Komposita behandelt, die Doppelbetonung aufweisen (*Őesterreich-Úngarn*) oder keine Klassifikationsfunktion haben (*Reus-Freistoß*). Nach den besprochenen Definitionsversuchen definieren also nur A, B, C und G Komposita. Die übrigen Merkmale sind hingegen nicht definitorisch für Komposition im Allgemeinen, sondern eher charakteristische Merkmale des vorherrschenden Kompositionssubtyps, nämlich der determinativen N+N-Komposita. Da diese Kriterien in der Diskussion über

ICCs und deren Zugehörigkeit zur Komposition eine Rolle spielen, werden sie hier zusätzlich aufgeführt. Sie stellen aber, gemäß dem Ansatzes von Gaeta und Ricca (2009), keine Definitionskriterien für Komposition im Allgemeinen dar. Bei der Gewichtung gilt also: A, B, C und G sind definitorisch, D, E, F, H, I, J sind häufige Merkmale von N+N-Komposita. [A, B, C, G]>[D, E, F, H, I, J]. Des Weiteren gilt A=B=C=G sowie D>E>F>H>I>J.

7.3 ICC-Bildung als Komposition im Deutschen

ICCs werden meist als Komposita beschrieben. Bei Det-ICCs gibt es hierzu keine Gegenposition. Die reduplikative Struktur gilt oft als rein zufällig (Donalies 2011: 72, Fleischer & Barz 2012: 96). Wenn man die aufgestellten Definitionskriterien ansetzt, zeigt sich, dass Det-ICCs die Zusammensetzung von lexikalischen Wortstämmen sind (B, G) und als morphologische Bildungen lexikalische Integrität aufweisen (A). Die ersten Konstituenten sind zudem betont (D); grammatischer (C) und semantischer Kopf (H) sind, wie für Komposita im Deutschen üblich, rechts. Schließlich ist der Prozess rekursiv (*Kindeskindeskinder*, E) und die Schreibung der Det-ICCs ist, basierend auf den hier beschriebenen Korpusstudien, in über 90% der Fälle die Zusammenschreibung (F). Schließlich haben Det-ICCs noch eine Klassifikationsfunktion (I) und es lässt sich recht einfach eine der häufig vorkommenden semantischen Relationen zwischen Erst- und Zweitglied etablieren (J), etwa OF wie bei *Kindeskind.* Det-ICCs sind also eindeutig Komposita.

Bei Prot-ICCs liegt der Fall nicht so klar. Die Einordnung der Prot-ICCs wird darum auch kontrovers diskutiert. Bereits die uneinheitliche Terminologie spiegelt wider, dass Bildungen wie *Oma-Oma* und der Prozess dahinter unterschiedlich eingeordnet werden. Die Bildungen werden als Komposita (identical constituent compounds, Finkbeiner 2014, Kentner 2017), als das Ergebnis eines phonologischen Kopierprozesses (Stolz 2018: 203f.), oder eines eigenen morphologischen (Freywald 2015, Gil 2005 zum Englischen) oder syntaktischen (Ghomeshi et al. 2004 zum Englischen) Reduplikationsprozesses verstanden.

Für die Annahme eines Kompositionsprozesses argumentieren neben Hohenhaus (2004) vor allem Kentner (2017) und Finkbeiner (2014). Die Annahme eines Kompositionsprozesses stützt sich zum einen auf die Erstgliedbetonung der Bildungen (D). Zum anderen befinde sich wie bei kanonischen Komposita der grammatische (C) sowie der semantische Kopf (H) rechts, wodurch auch die für Komposita im Deutschen grundliegende Modifikationsrichtung gegeben sei. Der erste Teil fungiert also als Modifikator, der die Bedeutung des zweiten (identischen) Teiles einschränkt (Hohenhaus 2004: 299, Kentner 2017: 240). An den Ergebnissen der Korpusstudien zeigt sich allerdings, dass, abgesehen von manchen

Belegen einiger ICCs wie *Chefchef*, *Glasglas* und *Rest-Rest*, die üblichen semantischen Relationen wie OF, FOR oder HAVE nicht angewendet werden können (J), beziehungsweise, dass die Relation auf die Eigenschaftszuschreibung reduziert ist. Allerdings trifft dieser Umstand auch auf A+N-Komposita zu. Finkbeiner (2014: 188) und Hohenhaus (2004) gehen ferner davon aus, dass Prot-ICCs trotz fehlender semantischer Relation die Funktion der Subklassenbildung zukommt (I). Nimmt man alle Kriterien für die Kompositabestimmung zusammen, gebührt Prot-ICCs „a place in the realm of German compounding" (Kentner 2017: 241).

Ein Argument gegen die Analyse von Prot-ICCs als Komposita ist, dass die Bildungen mitunter Binnenflexion aufwiesen (Bross & Fraser 2020: 3, Freywald 2015: 925). Das sei für N+N-Komposita ungewöhnlich. Dieses Argument (gegen die Kompositaanalyse) ist aber aus mehreren Gründen nicht valide. Zum einen widersprechen sich die Annahmen der bisherigen Forschung dazu und die formalen Elemente zwischen den identischen Stämmen werden meist nicht als Flexive angesehen (vgl. Finkbeiner 2014). Zum anderen liefern die durchgeführten Korpusstudien keine Evidenz dafür, dass Prot-ICCs Binnenflexion aufweisen. Zwar zeigen etwa die deTenTen13-Daten, dass Prot-ICCs nur in 77% der Fälle als Einheit flektieren, also nur das Zweitglied einen Flexionsmarker zeigt. In den restlichen 23% Prot-ICC-Belegen, in denen sich am Erstglied ein formales Element findet, ist das Element zum Flexiv am Zweitglied formidentisch und könnte als Argument gegen die Kompositaanalyse ins Feld geführt werden. Doch zeigen Prot-ICCs nur zu 13% überhaupt ein Element am Erstglied, sodass die Fälle, in denen Prot-ICCs ein Element am Erstglied tragen, das theoretisch ein Flexiv sein könnte, nur 89 von 1858 Fällen, also 4,8% aller Prot-ICC-Belege betrifft. Wenn man allerdings argumentiert, dass es sich in 4,8% der Prot-ICC-Belege theoretisch um Fälle handeln könnte, in denen das Element am Erstglied ein wortinterner Flexionsmarker ist, müsste man Binnenflexion auch als Argument für N+N-Komposita im Allgemeinen gelten lassen. Auch in kanonischen N+N-Komposita gibt es Fälle, in denen das Fugenelement am Erstglied mit dem Flexionsmarker am Zweitglied formidentisch ist (*die Wahrung des Kindeswohles*, *an Wochenenden*). Dieses Argument gegen die Kompositaanalyse von Prot-ICCs ist also auch schon dann ein recht schwaches, wenn man die Ergebnisse aus den beiden Korpusstudien nicht berücksichtigt. Die Korpusstudien zeigen darüber hinaus aber, dass die Grundlage für dieses Argument, nämlich ein Element am Erstglied, das möglicherweise ein Flexiv sein könnte, nur sehr selten vorliegt.

Der Bildung von Prot-ICCs wird außerdem die Besonderheit zugeschrieben, nicht rekursiv zu sein (Finkbeiner 2014: 187) was gegen Kriterium E verstieße. Auf ähnliche Weise begründet Botha (1988: 82), dass die von ihm untersuchten Reduplikationen im Afrikaans nicht als Komposita anzusehen seien, weil sie keine hie-

rarchische, rekursive Struktur aufweisen. Ebenso könnte man annehmen, dass *Oma-Oma* mit der Bedeutung 'richtige, prototypische Oma' nicht erneut mit *Oma* kombiniert werden kann, *Zinseszins* als Selbstkompositum hingegen als Possessivkompositum wieder für eine Kombination mit *Zins* bereitsteht: *Zinseszinseszins* 'Zins auf Zinseszinsen'. Da die zuvor berichteten Korpusstudien nur die einmalige Wiederholung von Nominalstämmen betreffen, kann auf der Grundlage der erhobenen Daten keine Aussage über die Rekursivität der ICC-Bildung getroffen werden.

Das Argument der fehlenden Rekursivität habe ich in 7.2 aber nur schwach gewichtet. Das Kriterium ist nämlich nicht allzu hilfreich, wenn man Komposition von anderen morphologischen Prozessen unterscheiden können will, da auch bei kanonischer Komposition nicht in allen Fällen Rekursivität gegeben ist. Allein N+N-Komposita zeichnen sich durch weitreichende Rekursivität aus (*Katzenhauskratzbaum*). Es ist in der Tat fraglich, ob ICCs wie **Oma-Omaoma* oder **Oma-Oma-Oma-Oma* eine Oma bezeichnen können, die noch prototypischer als eine *Oma-Oma* ist. Oder anders ausgedrückt: Es kann bezweifelt werden, dass es Sprachbenutzer:innen notwendig erscheint, ein so konkret eingeordnetes Subkonzept noch weiter zu spezifizieren.[46] Abseits von solchen Triplikationen steht ein ICC-Stamm wie *Oma-Oma* nach seiner Bildung aber womöglich für weitere Wortbildungsprozesse zur Verfügung (?**Omaoma-Schürze*).

Ebenso fragwürdig als Argument gegen die Kompositaanalyse ist das Kriterium der Schreibung (F). Im Gegensatz zu kanonischen Komposita, die meist als ein Wort geschrieben werden (*Liebesbrief*) sei die übliche Schreibweise von ICCs nach Finkbeiner (2012: 188) die Bindestrichschreibung (*Liebe-Liebe*), die Prot-ICCs also graphisch von kanonischer Komposition entfernt. Finkbeiners Annahme, die sich auf die 14 Beispiele aus Freywald (2015) stützt, von denen 7 die Bindestrichschreibung aufweisen, erhält mit den Ergebnissen aus den Korpusstudien dieser Arbeit weitere empirische Evidenz. In Prot-ICCs werden die Konstituenten überwiegend durch den Bindestrich (oder auf andere Art) getrennt; das zeigen sowohl die DECOW16- (63% Getrenntschreibung) als auch die deTenTen13-Daten (55% Getrenntschreibung).

46 Ein Hörbeleg aus einem Podcast (hooked fm #174, www.hookedmagazin.de/blog/2018/06/04/hooked-fm-174-mario-tennis-aces-dragon-quest-xi-deadpool-2-fallout-76-pokemon-und-mehr) lässt allerdings vermuten, dass die Prot-ICC-Bildung sich tatsächlich als rekursiver Prozess zeigt, sofern es nur eine entsprechende Situation gibt, die eine solche Bildung sinnvoll erscheinen lässt. Im Hörbeleg verweist der Sprecher darauf, dass das *Remake* eines bestimmten Films auch unter den typischen Film-Remakes ein sehr prototypischer Vertreter ist, und verwendet dazu die Bildung *Remakeremake-Remakeremake*. Offensichtlich können Prot-ICCs also durchaus die Basis einer weiteren Prot-ICC-Bildung sein.

Doch selbst auf dieser größeren Datenbasis ist die Schreibung kein valides Argument gegen die Kompositaanalyse, da die Schreibung mit Bindestrich, Binnenmajuskel oder Spatium auch bei kanonischen Komposita zunimmt (Fleischer & Barz 2012: 192ff., Scherer 2012). Selbst wenn die Schreibung von ICCs von der kanonischer Komposita abwiche, wäre dies eher ein Argument für den Ad hoc-Charakter der Bildungen und keines gegen die Kompositaanalyse.

Es werden noch viele weitere Argumente gegen die Kompositaanalyse angeführt, die nicht zu den oben aufgestellten Definitionsmerkmalen gehören. Etwa das Argument, dass auch Wortarten wie Pronomina, Verben (94) oder Adverbien (95) Basis des Prozesses sein können.

(94) *Wie ist das eigentlich wenn ich 2 Jahre arbeite und dann ein Jahr in Elternzeit gehe, muss ich dann noch ein Jahr **arbeitenarbeiten** um verbeamtet auf Lebenszeit zu sein?*
(Freywald 2015: 917)

(95) *was ich jetzt mach? **jetztjetzt** dir schreiben, und sonst studieren.*
(Freywald 2015: 917)

Üblicherweise stünden diese Wortarten nicht für Komposition zur Verfügung, weshalb sich Freywald (2015: 928) gegen die Kompositaanalyse ausspricht. Finkbeiner (2014: 186) verweist aber zurecht auf den Umstand, dass dies kein valides Argument gegen die Kompositaanalyse ist, da das Deutsche produktiv Komposita mit Pronomina und Adverbien bildet, beispielsweise *Wir-Gefühl* oder *Linkskurve* (Finkbeiner 2014: 186). Auch als Zweitglieder sind nicht-genuin nominale Wortarten keine Seltenheit:

(96) *Zu sehr hat sie die Regeln der Gesellschaft verinnerlicht, um sie noch in Frage stellen zu können. Das **Gesellschafts-Wir** ist zu ihrem eigenen **Über-Ich** geworden.*

<http://www.hamburgtheater.de/056660a49013b2d01.html>

(97) *SPD **fürgegen** Asyl – Unterbezirk beschloß Spagat in der Art.16-Debatte*

<taz.am Wochenende vom 12. 9. 1992, Inland, S. 26>

Des Weiteren bezieht sich die Prototypenreduplikation im Englischen auch systematisch auf Phrasen, etwa *SLEEPING-TOGETHER-sleeping-together* (Ghomeshi et al. 2004: 321), was vordergründig gegen Kriterium B verstieße. Um dieses Argu-

ment auch für das Deutsche zu stützen, müsste aber erst einmal Prototypenreduplikation im Deutschen nachgewiesen werden, bei der der Prozess auf syntaktische Bildungen angewendet wird. Hierzu liefert die Abfrage der zwei Korpora keine umfassenden Daten, da die deTenTen13-Daten die Spatiumschreibung ausschließen und die auf DECOW16 fußenden Daten zwar die Spatiumschreibung beinhalten, aber auf eine begrenzte Liste Nomina beschränkt sind. Dieses Argument ist allerdings auch abgesehen von einer solchen Überprüfung nicht gültig, da das Deutsche hochproduktiv Phrasalkomposita bildet.

Als Argument gegen die Kompositionsanalyse gilt außerdem die Annahme, dass die Fuge zwischen den beiden Konstituenten in ICCs nicht verfugt wird. Die Verwendung eines Fugenelementes fördere hingegen die determinative Lesart:

(98) *Freund-Freund*
'richtiger Freund'
(Freywald 2015: 925)

(99) *Freundesfreund*
'Freund eines Freundes'
(Freywald 2015: 925)

(100) *Mann-Mann*
'richtiger Mann'
(Bross & Fraser: 3)

(101) *Männermann*
'an Männern interessierter Mann'
(Bross & Fraser 2020: 3)

Die Annahme, dass ICCs mit Prototypenbedeutung wie in (98) und (100) unverfugt bleiben, wurde bisher allein durch theoretische Überlegungen nahegelegt (Kentner 2017: 240). Die in dieser Arbeit beschriebenen Korpusstudien konnten für diese Annahme keine Evidenz liefern. Ganz im Gegenteil: Prot-ICCs zeigen Fugenelemente, und zwar zu einem nicht geringen Anteil. In den DECOW16-Daten sind 23% der Prot-ICC-Belege verfugt, in den deTenTen13-Daten 13%. Angesichts einer durchschnittlichen Verfugung kanonischer N+N-Komposita von 26,5% (Wellmann 1975), beziehungsweise 35% (Nübling & Szczepaniak 2009), verhalten sich Prot-ICCs hinsichtlich der Kompositionsfuge also nicht sonderlich auffällig. Auf keinen Fall kann man von einem systematischen Ausbleiben der Verfugung bei Prot-ICCs sprechen.

Doch selbst wenn man die Position vertritt, dass der kleine Unterschied in der Verfugung zwischen Prot-ICCs und N+N-Komposita im Allgemeinen relevant ist, wäre das dennoch kein gültiges Argument gegen die Kompositionsanalyse. Viele andere Kompositionsarten verfugen gar nicht oder nur in bestimmten Fällen. Phrasenkomposita etwa kommen gänzlich ohne Fugenelement aus (Hein 2015). In Eigennamenkomposita wird ebenfalls systematisch nicht verfugt (*Mannsbild* versus *Thomas-Mann-Bild*), wodurch diese einen formalen Unterschied zu Appellativa markieren (Schlücker 2017, Schlücker & Ackermann 2017: 324). Überhaupt können Fugenelemente nur dann auftreten, wenn das Erstglied ein Substantiv- oder Verbalstamm ist (Fleischer & Barz 2012: 185). Und selbst die meisten Substantivkomposita haben kein Fugenelement (Nübling & Szczepaniak 2009: 196, Wellmann 1975). Dennoch wird diese Eigenschaft den Prot-ICCs zugeschrieben und ist das am meisten angeführte Argument gegen die Kompositionsanalyse:

> [C]ompounding in German often involves linking elements which are banned in CR [=Prot-ICCs, M.F.]. […] despite all these problems with a compound analysis, some authors still defend this view.
>
> (Bross & Fraser 2020: 5).

Bross und Fraser (2020) stützen ihre Annahmen zur Verfugung in Prot-ICCs allerdings nicht auf empirische Daten und es kann angesichts der Ergebnisse der hier präsentierten Korpusstudien als widerlegt gelten, dass Prot-ICCs keine Fugenelemente aufweisen.

Bross und Fraser (2020: 7) nehmen des Weiteren an, dass Prot-ICCs nicht modifiziert werden können (102) und formulieren für den nominalen Bereich die syntaktische Beschränkung „adjectival modification is completely banned in CR".

(102) *Ich will einen (*schwarzen/kostenlosen/heißen) kaffee-Kaffee.*
(Bross & Fraser 2020: 7)

Dies gelte allerdings nicht für Real Intensifications, also Adjektive, die bloß die Prototypenbedeutung der ICCs unterstreichen. (Bross & Fraser 2020: 7).

(103) *Paul ist so ein richtiger mann-Mann.*
(Bross & Fraser 2020: 7)

Dieser Bann adjektivischer Modifikation spreche gegen eine Kompositionsanalyse, da bei Komposita normalerweise alle Modifikationen erlaubt seien. Dass Real Intensifications keiner Restriktion unterliegen, weise zudem darauf hin, dass

diese Beschränkung keine phonologische sein könne. Die Beschränkung fuße vielmehr auf der Tatsache, dass die Position, in der für gewöhnlich eine Adjektivphrase steht, bereits besetzt sei, da das Prot-ICC selbst die „Aura eines Modifiers" (Ghomeshi et al. 2004, ebenso: Stolz et al. 2011: 202ff.) habe, beziehungsweise einen „adjectival flavor" (Bross & Fraser 2020: 6). Diese Argumentation beruht auf der allgemeineren Diskussion in der Kompositaforschung zum Englischen, dass manche Erstglieder von N+N-Komposita auch als Adjektive angesehen werden können (Bell 2012, Tarasova 2013).

Die empirische Grundlage für Bross und Frasers Annahmen ist allerdings allein das von Ghomeshi und Kolleg:innen (2004) zur Verfügung gestellte Corpus of English contrastive focus reduplications,[47] das keine Daten zum Deutschen enthält. Die Annahme, das adjektivische Modifikation bei Prot-ICCs im Deutschen nicht möglich sei, rührt daher wohl allein vom Englischen her, für das in der Literatur ein Bann adjektivischer Modifikation angenommen und mithilfe des Korpus von Ghomeshi und Kolleg:innen auch empirisch untermauert wurde (Benjamin 2018: 33, Ghomeshi et al. 2004). Die nominalen der insgesamt 334 im Korpus enthaltenen Beispiele zum Englischen sind nämlich tatsächlich allesamt nicht adjektivisch modifiziert.

Benjamin (2018) führt (für das Englische) folgende Erklärung dafür an, warum Prot-ICCs nicht modifiziert werden können:

> Modification by adjectives […] is, by and large, blocked by Identical Constituent Compounding. This follows from the same principle that realizes these utterances as prototypical; Identical Constituent Compounds prompt for the imagination of some conceptual prototype with pre-established features and boundaries; if a prototypical apple is red and fist-sized, then a **green* ***APPLE****-apple* or a **giant* ***APPLE****-apple* would no longer be prototypical.
>
> (Benjamin 2018: 33, Hervorhebung im Original)

Bross und Frasers Argument inferiert also Merkmale von Prot-ICCs im Deutschen auf der Grundlage der Bildungen im Englischen. Es ist aber angesichts der morphosyntaktischen Unterschiede der zwei Sprachen schwierig, die adjektivische Modifikation im Deutschen mit der im Englischen gleichzusetzen. Zwar sind die Verbindungen aus Adjektiven und Nomina in Mehrwortausdrücken des Englischen und Deutschen jeweils Abfolgen von Modifikator und Kopf (*high mountains – hohes Gebirge*). Doch ist die Grenze zwischen morphologischen und syntaktischen Verbindungen im Englischen sehr viel schwieriger zu ziehen als im Deutschen. Anders als im Deutschen geht kein Betonungsmuster eindeutig mit dem Vorliegen von Komposita oder Phrasen einher und wegen der weitgehenden Fle-

47 http://home.cc.umanitoba.ca/~krussll/redup-corpus.html

xionslosigkeit kann auch die Flexion des Modifikators nicht als Hinweis auf die formale Einordnung dienen. Die Unterscheidung wird deshalb für das Englische mitunter auf Basis der Schreibung getroffen (Copestake et al. 2002, Erman & Warren 2000). Die Frage, ob ein Prot-ICC adjektivisch modifiziert wird oder nicht, lässt sich aber im Englischen nicht so einfach entscheiden wie im Deutschen. Im Deutschen zeigt das Betonungsmuster der Erstgliedbetonung (*hohes Gebírge* versus *Hóchgebirge*) sowie die Flexion des pränominalen Adjektivs (*hohes Gebirge* versus *Hoch_gebirge*) eindeutig an, ob das Adjektiv attributiv syntaktisch modifiziert oder morphologisch (Bauer 1998, Giegerich 2006, Schlücker 2012). Die Gleichsetzung adjektivischer Modifikation der englischen Prot-ICCs mit den deutschen, die mit Bross und Frasers Argument einhergeht, ist angesichts der Unterschiede der zwei Sprachen schwierig und die Ergebnisse zu Prot-ICCs im Englischen sagen darum nichts zur adjektivischen Modifikation deutscher Prot-ICCs aus.

Die Daten aus den hier berichteten Korpusstudien zum Deutschen sprechen denn auch dafür, dass der für das Englische angenommene Bann adjektivischer Modifikation nicht für das Deutsche gilt. Basierend auf den DECOW16-Daten werden Prot-ICC-Belege in 20% der Fälle durch attributive Adjektive modifiziert. Nur in der Hälfte der Fälle, in denen Prot-ICCs adjektivische Attribute haben, sind diese Adjektive Real Intensifications. Entgegen Bross und Frasers Annahme werden Prot-ICCs im Deutschen also durchaus attributiv modifiziert und zwar durch ‘gewöhnliche’ Adjektive wie *schön* oder *kreativ*. Dieses Argument gegen die Kompositionsanalyse basiert also offensichtlich auf einer Fehleinschätzung.

Finkbeiner (2014) nennt eine Reihe weiterer Eigenschaften, die Prot-ICCs von kanonischen Komposita unterscheiden. Bei manchen Nomina ist für die Komposition eine vom Stamm abweichende Kompositionsstammform vorgesehen, etwa bei *Liebe*. Tritt das Lexem in Komposita auf, erhält es die Form *Liebes-* (*Liebesbrief, Liebesheirat* – **Liebebrief, *Liebeheirat*, Finkbeiner 2014: 185ff.). Mit anderen Worten: Das Fugenelement *-s-* ist bei Komposita mit *Liebe* als Erstglied obligatorisch. In ICCs hingegen trete das Lexem in seiner Nennform *Liebe* auf (*Liebe-Liebe*). Wird stattdessen die Kompositionsstammform verwendet, sei die Prototypenlesart ausgeschlossen (Finkbeiner 2014: 187, ebenso: Freywald 2015: 926). Diese Einschätzung wird allerdings, wie schon die Einschätzungen zum Fugenelement generell, nicht vollumfänglich von den vorgestellten Korpusstudien bestätigt. Eine Wortform wie *Liebesliebe* kann durchaus die Bedeutung ‘richtige, prototypische Liebe’ tragen.

Ein letzter, häufig diskutierter Aspekt, der meist für die Kompositionsanalyse hervorgebracht wird, ist der Wortakzent von Prot-ICCs. Die Betonung liege auf dem Erstglied, was dem Betonungsmuster von N+N-Komposita entspräche, und nicht auf dem zweiten Glied, wie es für phrasale Konstruktionen charakteristisch

ist (Finkbeiner 2014: 187, Freywald 2015: 925, Hohenhaus 2004: 323). Bisher gibt es hierzu allerdings keine empirische Evidenz, da Prot-ICCs noch nicht an Daten der gesprochenen Sprache untersucht wurden.[48]

Hinsichtlich der formalen Kriterien spricht nichts dagegen, Bildungen wie *Oma-Oma* als Komposita aufzufassen. Die diskutierten Argumente sprechen entweder nicht gegen die Kompositaanalyse (eingeschränkte Rekursivität), konnten mithilfe der Korpusdaten widerlegt werden (keine lexikalische Integrität) oder beides (abweichende Kompositionsstammformen, Binnenflexion, fehlende Fugenelemente, Bann adjektivischer Modifikation). Viele formale und funktionale Eigenschaften von Kompositionsprodukten liegen in den besprochenen Bildungen hingegen vor. Die Position des grammatischen (C) und semantischen Kopfes (H) ist wie in kanonischen Nominalkomposita rechts, die Konstituenten sind freie Stämme (B) und die Bildungen sind lexikalisch integer (A).

Den dritten ICC-Typ, Name-ICC, kann man nicht uneingeschränkt der Komposition zurechnen. Hinsichtlich fast aller Definitionskriterien wirken Name-ICCs uneindeutig. Die Bildungen sind zwar morphologisch und die Konstituenten Wortstämme, sodass die Kriterien A und B erfüllt sind. Doch bestehen hinsichtlich der Erfüllung der meisten anderen Kriterien mindestens Zweifel.

Die semantisch-funktionalen Kriterien erfüllt die Name-ICC-Bildung allesamt nicht. Zwar tragen die Konstituenten lexikalische Bedeutung. Diese geht aber während des Bildungsprozesses größtenteils oder gänzlich verloren. Als Eigennamen haben Name-ICCs zudem oft nur eine Referenz aber keine Semantik. In Beispielen wie *AutoAuto* liegt zwar eine lexikalische Information in Form des Modifikators vor. Der semantische Kopf, der anzeigen würde, dass es sich um eine Show handelt, ist aber nicht gegeben / befindet sich außerhalb des Kompositums. Das ICC ist also exozentrisch und erfüllt nicht das Kriterium H. Das Beispiel *PunktPunkt* übt selbst diese Funktion nicht aus. Kriterium G zur lexikalischen Bedeutung eines Kompositums kann hier nicht als gegeben betrachtet werden. Auch eine Subklasse (I) entsteht durch die Bildung eines Name-ICCs nicht. Eine semantische Relation zwischen den Konstituenten liegt ebenfalls nicht vor (J). Semantisch-funktional gesehen sind Name-ICCs also keine Komposita. Mitunter weichen Name-ICCs sogar hinsichtlich des grammatischen Kopfes (C) von kanonischer Komposition ab (*das Tiger-Tiger* = Disco, *die GasGas* = Motorrad). Von den definitorischen Kriterien erfüllen Name-ICCs aber immerhin zwei vollumfänglich. Da die Bildungen morphologisch sind und

48 Im Laufe meines Dissertationsprojektes habe ich allerdings eine erste Datenbasis für eine Bewertung des Betonungsmusters von Prot-ICCs zusammengestellt. Zu Illustrationszwecken in Vorträgen habe ich, größtenteils basierend auf Podcasts, 50 Prot-ICCs als Audiodateien erfasst. Alle Hörbelege dieser Sammlung zeigen die Erstgliedbetonung.

aus Wortstämmen bestehen, sind Kriterium A und B erfüllt. Doch trägt die Name-ICC-Bildung keine (oder nur eine sehr unvollständige) lexikalische Bedeutung.

In Bezug auf die Analyse von ICCs lässt sich zusammenfassen: Det-ICCs erfüllen alle aus der Literatur abgeleiteten Definitionskriterien und Merkmale für Komposition und sind damit Komposita. Prot-ICCs weichen in einigen Punkten ab, erfüllen aber alle definitorischen Kriterien, sind also ebenfalls Komposita. Name-ICCs hingegen bestehen zwar aus Wortstämmen, erfüllen aber nicht alle definitorischen Kriterien. Zudem weisen sie die übrigen Merkmale von Komposition nicht auf. Sie können somit nur sehr eingeschränkt als Komposita angesehen werden. Tabelle 15 gibt einen Überblick darüber, welcher ICC-Typ welche Kriterien und Merkmale erfüllt.

Tab. 15: Die drei ICC-Typen als Komposition

Kriterium / Beschreibung		**definitorisch**	**Det-ICC**	**Prot-ICC**	**Name-ICC**
		Formal			
A	Die Bildung ist morphologisch.	+	+	+	+
B	Die Konstituenten sind Wortstämme.	+	+	+	+
C	Die letzte Konstituente ist der grammatische Kopf.	+	+	+	○
D	Die erste Konstituente ist betont.	–	+	+	○
E	Der Prozess ist rekursiv.	–	+	○	–
F	Die Bildung wird zusammengeschrieben.	–	+	–	–
		Semantisch-funktional			
G	Kompositum und Konstituenten tragen lexikalische Bedeutung.	+	+	+	○
H	Die letzte Konstituente ist der semantische Kopf.	–	+	+	–
I	Das Kompositum bildet eine Subklasse	–	+	+	–
J	Semantische Relation zwischen Konstituenten	–	+	–	–

In der Literatur werden ICCs allerdings nicht nur ob ihrer Zugehörigkeit zur Komposition besprochen. Auch Ähnlichkeiten mit Reduplikationsprozessen wer-

den hervorgehoben. Der Ansatz, Bildungen wie *Oma-Oma* oder *Baden-Baden* nicht als Komposition, sondern als Reduplikation einzuordnen, wird im folgenden Kapitel 8 besprochen.

8 ICCs als Reduplikation

ICCs werden mitunter als das Ergebnis eines phonologischen Kopierprozesses (Stolz 2018: 203f.) oder eines eigenen morphologischen (Freywald 2015, Gil 2005 zum Englischen) oder syntaktischen (Ghomeshi et al. 2004 zum Englischen) Reduplikationsprozesses verstanden. Die Annahme, reduplikative Wortbildungen im Deutschen seien stets das Ergebnis eines phonologischer Kopierprozesses vertreten etwa Hentschel und Weydt (2013). Demnach geschieht Reduplikation im Deutschen immer „durch Verdopplung einer Silbe, mit oder ohne Vokalwechsel“ (Hentschel & Weydt 2013: 23). Auch Nübling (2004) spricht von Reduplikation als einer „quantitative Erweiterung […] also eine Form der Kopie“ (Nübling 2004: 26). Nübling behandelt die Dopplung von Interjektionen und nennt Bildungen wie *igittigitt* „eine Erscheinung der Phonologie“, da die betroffenen Elemente „frei von morphologischen Strukturen“ sind. Hier „bewirkt die Reduplikation […] eine Intensivierung“ (Nübling 2004: 26f.). Die Annahme, dass Reduplikation im Deutschen ein phonologischer Kopierprozesses sei, betrifft aber vornehmlich Bildungen, deren Basen keine freien Stämme sind (*Pinkepinke*). ICCs als Zusammensetzungen von Wortstämmen werden meist nicht als das Ergebnis eines phonologischen Prozesses bezeichnen, wohl aber wird Reduplikation als phonologischer Prozess angesehen (Marantz 1982, Stolz 2018: 203f., Wilbur 1973). Man kann ICCs also entweder als Hinweis darauf ansehen, dass Reduplikation eben nicht notwendigerweise ein phonologischer Kopierprozess ist, oder sie, wie Kentner (2017), nicht zu den Reduplikationsprozessen zählen.

Ich werde nun die Einordnung von ICCs darstellen, die neben der Kompositionsanalyse in der Literatur dominiert. Diese Einordnung sieht in der Bildung von ICCs einen syntaktischen oder morphologischen Reduplikationsprozess. Dazu werde ich zunächst das sprachliche Phänomen der Reduplikation vorstellen (8.1), im Anschluss daran Definitionskriterien aufstellen, um die Reduplikation möglichst klar von anderen Prozessen abzugrenzen (8.2), und mithilfe dieser Kriterien bewerten, inwieweit die Dopplungsphänomene im Deutschen (8.3) und ICCs (8.4) zur Reduplikation gehören.

8.1 Reduplikation in den Sprachen der Welt

Reduplikation ist ein sprachlicher Wiederholungsprozess, der in vielen Sprachen zu beobachten ist. In der Literatur ist man sich allerdings nicht einig, um was für einen Prozess es sich überhaupt handelt. Er wird mal als syntaktisch (Bollée 1978, Ghomeshi et al. 2004), mal als morphologisch (Booij 2012, Rubino 2005) und mal

https://doi.org/10.1515/9783111315416-008

als phonologischer Kopierprozess (Marantz 1982, Stolz 2018, Wilbur 1973) aufgefasst. Auch innerhalb dieser unterschiedlichen Auffassungen ist man sich über die Natur des Prozesses nicht einig. So wird innerhalb der vorherrschenden Auffassung, Reduplikation sei ein morphologischer Prozess, mal von einem flexionsmorphologischen (Hurch et al. 2008: 2) und mal von einem derivationsmorphologischen Prozess ausgegangen (Goodwin Gómez & van der Voort 2014: 4f.). Hinzu kommt, dass die Abgrenzung zwischen Reduplikation und ähnlichen Prozessen wie etwa der Repetition nicht immer leichtfällt. Mitunter werden die beiden Prozesse sogar überhaupt nicht unterschieden. Diese Uneinigkeiten und die ungleichen Begriffsgrenzen müssen beseitigt werden, ehe man bewerten kann, ob ICCs zur Reduplikation gehören oder nicht.

> Nothing is more natural than the prevalence of reduplication, in other words, the repetition of all or part of the radical element.
>
> (Sapir 1921: 76)

Für jedes semiotische System ist die Wiederholung eine Notwendigkeit. Die Musik, der Finanzverkehr, das Alltagsleben der Menschen, alles beruht darauf, dass sich Dinge und Sachverhalte wiederholen. „Repetition is […] oiling the waters of social interaction“ (Brown 1999: 223). Auch Sprache ist durch Wiederholung gekennzeichnet. Alle Sprachen beruhen darauf, dass Zeichen identisch wiederverwendet werden und in allen Sprachen finden sich Wiederholungsphänomene. Die Gründe für eine Wiederholung können sehr unterschiedlich sein. So werden Laute beispielsweise wiederholt, wenn sie in mehreren Wörtern vorkommen, Wörter, wenn auf einen entsprechenden Referenten mehrmals verwiesen wird, Sätze, wenn eine fremde Äußerung zitiert wird. In all diesen Fällen liegt eine Systematik, ein bestimmter Zweck der Wiederholung vor. Meist findet die Wiederholung nicht unmittelbar statt, sondern es stehen andere Elemente zwischen den identischen Lauten, Wörtern oder Sätzen. Der Wiederholung sprachlicher Elemente kommt allerdings ein besonderer Gehalt zu, wenn sich ein Segment allzu bald wiederholt, also so, dass nur wenig oder gar nichts zwischen den identischen Elementen steht.

Wiederholen sich sprachliche Elemente adjazent, kann das zufällig oder systematisch passieren. Ein Beispiel für zufällige, adjazente Wiederholung sprachlichen Materials geben (104) und (105).[49]

49 In diesen und den folgenden Beispielen dieses Kapitels sind jeweils die sich wiederholenden Einheiten unterstrichen. Da in diesem und dem folgenden Kapitel Beispiele unterschiedlicher Sprachen behandelt werden, wird in der ersten Zeile zudem die jeweilige Sprache angegeben.

(104) Deutsch (Artikel 16a, Absatz 2 des Grundgesetzes der Bundesrepublik Deutschland)
Die Staaten außerhalb der Europäischen Gemeinschaften, auf die die Voraussetzungen des Satzes 1 zutreffen, werden durch Gesetz, das der Zustimmung des Bundesrates bedarf, bestimmt

(105) Deutsch (https://tatoeba.org/deu/sentences/show/2320944)
Die, die die, die die Dietriche erfunden haben, tadeln, tun unrecht.

In (104) sind Relativpronomen und Determinierer mehr oder weniger zufällig formgleich. Aufgrund syntaktischer Eigenschaften des Deutschen stehen sie unmittelbar hintereinander. Sie tun dies aber nicht notwendigerweise. Andere Elemente (etwa Adverbien wie *ferner*) können diese Adjazenz unterbrechen, ohne dass sich die Bedeutung der formgleichen Wörter ändern würde. Den Bestandteilen dieser Wiederholung kommt hier also lediglich die Bedeutung zu, die ihnen auch in anderen Kontexten innewohnt. Die unmittelbare Wiederholung sprachlichen Materials selbst trägt hier keine Bedeutung. Ebenso in (105). Hier sind Pronomina, Determinierer sowie ein Teil des Substantivs formgleich. Auch hier führen syntaktische Regeln dazu, dass sie unmittelbar aufeinander folgen, auch wenn sich der/die Verfasser:in in diesem Fall wohl der Wiederholung bewusst war und der Text eben wegen der häufigen Wiederholung der Lautfolge [di:] erstellt wurde.

Die Reduplikation ist nun ein besonderer Fall der Wiederholung. Sie ist dadurch gekennzeichnet, dass sie nicht zufällig, sondern systematisch auftritt. Es ist die Wiederholung selbst, die grammatische oder semantische Informationen trägt, wie etwa im Twi, wo Wiederholung systematisch zur Transposition verwendet wird:

(106) Twi (Moravcsik 1978: 324)

abo	*abó*
Steine	Steine
‘steinig’	

Die Verdopplung eines Nomens führt im Twi systematisch zu einem Adjektiv. Geschieht eine Wiederholung derart systematisch, funktional und bedeutungstragend, kann man das von unsystematischen Fällen mithilfe des Terminus „Reduplikation“ begrifflich abgrenzen.

Es sind sehr unterschiedliche Phänomene auf verschiedenen sprachlichen Ebenen, die unter dem Begriff „Reduplikation“ beschrieben werden. Das liegt zum

einen an der schieren formalen und funktionalen Vielfalt der Wiederholungsphänomene in den Sprachen der Welt. Zwar gibt es Versuche, in den Doppelstrukturen einen gemeinsamen, ikonischen Bedeutungskern, der allen Reduplikationsphänomenen innewohnt, auszumachen. Rozhanskiy (2015) nimmt etwa ein Set von Bedeutungen an, von denen Reduplikation immer wenigstens eine (oder eine Kombination dieser) ausübt, nämlich Ähnlichkeit, Quantitativität, Pejoration, Emphase und lexikalischer Klassenwechsel (Rozhanskiy 2015: 996ff.). Selbst er resümiert aber zum Abschluss seiner Studie:

> [W]ithin the range of reduplication meanings the choice is unmotivated and depends both on the language and on the particular form.
>
> (Rozhanskiy 2015: 1014)

Zum anderen wird die Unübersichtlichkeit der Reduplikationsphänomene durch die bereits angedeutete terminologische Vielfalt innerhalb der linguistischen Beschreibung noch verstärkt. Ein Versuch, diesen Missstand aufzuheben, ist die Terminologie der „Graz Database on Reduplication“ (Hurch 2005), die der Base-reduplicant-correspondence theory folgt. Demnach wird der Teil einer Simplexform, der kopiert wird, als Basis der Reduplikation bezeichnet. Der kopierte Teil heißt Reduplikant und das Ergebnis dieses Prozesses wird reduplizierte Wortform genannt.[50] In (106) ist *abo* also die Simplexform und gleichzeitig die Basis des Reduplikationsprozesses, *abó* der Reduplikant und *abo-abó* die reduplizierte Wortform. Dass Simplexform und Basis nicht immer zusammenfallen, sieht man im folgenden Beispiel (107):

(107) Hausa (Newman 2000: 432)

a.	*bindigà:*	*bindig-o:gi:*
	‘Waffe’	‘Waffen’
b.	*fannì:*	*fann-o:ni:*
	‘Kategorie’	‘Kategorien’

Hier ist *bindigà*: die Simplexform, *-g-* die Basis und gleichzeitig der Reduplikant des Reduplikationsprozesses. In manchen Fällen unterscheiden sich sowohl Simplexform als auch Basis, Reduplikant und reduplizierte Wortform voneinander. Deshalb ist die terminologische Unterscheidung zwischen den beteiligten Elemen-

50 Zur Kritik an dieser Terminologie siehe Inkelas und Zoll (2005: 10f.).

ten der Reduplikation wichtig. Im Folgenden beschreibe ich Wiederholungsprozesse durchgängig mit der Terminologie der Graz Database on Reduplication.

Die bereits genannten und die noch folgenden Beispiele (106–111) entstammen sprachtypologisch sehr unterschiedlichen Sprachen, denen in der Literatur Reduplikation zugeschrieben wird:

(108) Mandarin Chinesisch (Zhu et al. 1995: 191)

張	張	桌子
zhāng	*zhāng*	*zhuōzi*
CL_{flach}	CL_{flach}	Tisch

'Tische'

(109) Mandarin Chinesisch (Packard 2000: 88)

皮	肤
pí	fū
Haut	Haut

'Haut'

(110) Bretonisch (Stolz et al. 2011: 497)

Tost	*-tost*	*dezho*	*emañ*	*egile*	*bremañ*	*just*
nah	nah	zu.3PL	sein.3SG	der andere	jetzt	nur
en	*tu*	*all*	*d'*	*ar*	*vodenn.*	
in	Seite	anderer	von	DEF	Zelt	

'Der andere ist ihnen jetzt sehr nah, nur auf der anderen Seite des Zeltes.'

(111) Französisch (Stolz et al. 2011: 508)

des	*milliers*	*et*	*des*	*milliers*

'Tausende und Abertausende'

Im Hausa (107) wird in der Nominalflexion ein Suffix *-o:Ci:* angehängt, in dem der Laut *C* aus der letzten Silbe der Basis kopiert wird (Inkelas & Zoll 2005: 2). Das Suffix enthält somit wiederholtes Lautmaterial, wodurch eine nominale Basis pluralisiert wird. Im Mandarin Chinesischen wird zu demselben Zweck der Nominalklassifikator dupliziert (108). Ebenfalls im Mandarin treten sogenannte Ko-Komposita (symmetrische Komposita, redundante Komposita) auf (109). Diese Komposita haben im Chinesischen – wie in vielen anderen Sprachen auch – die

Funktion, durch Wiederholung von Bedeutung die vielen homophonen Wurzelmorpheme der Sprache zu disambiguieren (Arcodia 2007, Wälchli 2007).[51] Das Beispiel aus dem Bretonischen (110) demonstriert, wie die Dopplung von *tost* 'nah' zu einer Verstärkung der Basisbedeutung führt; im Französischen (111) geschieht durch die Dopplung der Phrase *des milliers* 'tausende' ebenfalls eine solche Augmentation.

Die Beispiele verdeutlichen, wie vielfältig die Phänomene sind, die in der Forschung Reduplikation genannt werden. Sie unterscheiden sich in dem, was wiederholt wird, sowie in ihrer Ebene, Funktion, Struktur und Qualität. Am offensichtlichsten unterscheiden sich die Beispiele darin, welche Seite des sprachlichen Zeichens wiederholt wird. In (107) wird nur die Form, aber keine Bedeutung wiederholt, in (109) nur die Bedeutung, nicht aber die Form. In den übrigen Beispielen (106, 108, 110, 111) wird sowohl die Form als auch die Bedeutung wiederholt.

Auch in Bezug auf die Form allein gibt es viele Unterschiede: Zum einen findet in (106–111) die Wiederholung sprachlichen Materials auf unterschiedlichen Ebenen statt. In (107) wird lediglich ein Laut wiederholt, sodass man von bloßer phonologischer Kopie sprechen kann. Newman nennt den Prozess deshalb „pseudo reduplication" (Newman 1989: 249). In (108) zeigt sich ein morphologischer Prozess, bei dem gebundene grammatische Morpheme wiederholt werden, ebenso in (106) und (110), wo Wörter wiederholt werden. In (111) schließlich wird eine ganze Phrase wiederholt, was aufgrund der Tatsache, dass Einheiten oberhalb der Wortebene wiederholt werden, häufig als syntaktischer Prozess bezeichnet wird („syntaktische Reduplikation", Stolz 2018: 270, Stolz et al. 2011: 148ff., Wierzbicka 1986). Zum anderen unterscheiden sich die oben genannten Beispiele darin, ob die Formseite komplett (106, 108, 110, 111), nur partiell (107) oder überhaupt nicht (109) wiederholt wird und ob dies adjazent (106, 108–110) oder nur mittelbar (107, 111) geschieht.[52]

51 Stolz et al. (2011: 62) verweisen darauf, dass in solchen Synonymiekonstruktionen, die vor allem in asiatischen Sprachen auftreten, im Gegensatz zu Reduplikationskonstruktionen nicht alle semantischen Merkmale gleich sind, sondern sich die Konstituenten etwa in Bezug auf Register oder Stratum unterscheiden.

52 In der Forschung zur Reduplikation wird meist sehr viel genauer unterschieden. Mel'čuk (1996) etwa unterscheidet sechs „dimensions" einer Reduplikationstaxonomie; ebenso (aber abweichend von Mel'čuk) beschreiben Stolz et al. (2011) sechs Kriterien. Zu unterscheiden sind hier insbesondere completeness, exactness, contiguity, continuity und relative position zu nennen. In Anbetracht der Tatsache, dass dies eine Arbeit zum Deutschen und keine typologische ist, belasse ich es aber bei dieser rudimentären Differenzierung.

Schließlich übt die Wiederholung in den Beispielen sehr unterschiedliche Funktionen aus. In (107) wird Material wiederholt, um einen Onset für die Suffixsilbe zu erhalten. Der Prozess genügt hier also silbenphonologischen Anforderungen. (106) zeigt Wiederholung als Wortbildungsprozess, in (108) wird Material wiederholt, um Plural, also eine grammatische Kategorie, auszudrücken. In (109) trägt die Komposition synonymer Morpheme dazu bei, Homophonie zu vermeiden und die vielen durch phonologischen Abbau entstandenen gleichlautenden Wurzelmorpheme des Chinesischen zu disambiguieren (Bai et al. 2008: 695, Li & Thompson 1981: 14). In (110) und (111) hat die Wiederholung eine intensivierende, also semantisch-pragmatische Funktion.

Dass all diese Prozesse trotz der offensichtlichen Unterschiede der Reduplikation zugeordnet werden, verdeutlicht, dass man klare Definitionskriterien braucht, um den Unterschieden zwischen diesen Wiederholungsphänomenen gerecht zu werden. Von den Definitionskriterien hängt auch ab, wie man das Phänomen Reduplikation versteht. Ist Reduplikation für Sprache ebenso grundlegend wie die Wiederholung im Allgemeinen und somit als Universalie einzuordnen (Hagège 1982, Moravcsik 1978, Rubino 2005) oder ist sie in den Sprachen der Welt ungleich verteilt und damit eher als areal beschränkt (Stolz et al. 2011) anzusehen? Nach letzterer Einschätzung ist etwa Westeuropa ein Gebiet, in dem Reduplikation nahezu abwesend ist (Goodwin Gómez & van der Voort 2014: 1, Rubino 2005: 115, Vacek 1988: 111). In Rubinos sprachtypologischer Studie sind 70% der europäischen Sprachen ohne Reduplikation. In keiner andere Region fehlt so vielen Sprachen dieses Merkmal.[53] Für die Annahme einer Universalie spricht sich dagegen Hagège aus. Er sieht Reduplikation als „widely used strategy“ an (Hagège 1982: 25f.). Auch sprechen die Daten von Moravcsik (1978), deren Sprachsample 90 Einzelsprachen umfasst, für die Annahme einer Universalie.[54] Festzuhalten ist in jedem Fall, dass Sprachen ohne Reduplikation wohl in der Minderheit sind. In Rubinos Sprachsample sind von 367 Einzelsprachen nur 56, also nur zirka 15%, ohne Reduplikation. Somit stellen die europäischen Sprachen eher die Ausnahme von der Regel dar, nach der die Sprachen der Welt mit großer Mehrheit Gebrauch von Reduplikation machen.

53 Stolz et al. (2011. 130f.) nennen hier noch Südamerika, wo 12 von 31 Sprachen keine Reduplikation aufweisen sowie den nördlichen Polarkreis, in dem keine Sprache Reduplikation hat. Aufgrund dieser Verteilung sprechen die Autor:innen hier von einem Nord-Süd-Gefälle der Reduplikationssprachen.

54 Moravcsik selbst geht nicht so weit, von einer Universalie zu sprechen. Sie wird mit dieser Aussage aber häufig in Verbindung gebracht (Raimy 2000).

Die Gültigkeit dieser Aussage – und auch aller weiteren Annahmen zur Zugehörigkeit von ICCs zur Reduplikation – hängen aber davon ab, ab wann man bei einem Wiederholungsphänomen von Reduplikation spricht. Wenn man mit Reduplikation jegliche Wiederholung sprachlichen Materials ungeachtet der Ebene, Funktion, Struktur und Qualität meint, ist sie zweifelsohne eine Sprachuniversalie. Und mit dieser Setzung wären auch ICCs im Deutschen Reduplikation. Folgt man aber einer engeren und anhand deutlicher Merkmale klar beschreibbaren Begriffsauffassung, kommt man womöglich zu dem Schluss, dass die ICCs im Deutschen nicht zur Reduplikation zu rechnen sind. Im folgenden Unterkapitel werden also zunächst eindeutig feststellbare Definitionskriterien für Reduplikation aufgestellt, um in der Folge bewerten zu können, ob ICCs und andere Dopplungsphänomene im Deutschen nach dieser Definition zur Reduplikation zu zählen sind.

8.2 Definitionskriterien für Reduplikation

Der Terminus Reduplikation hat ein lateinisches Etymon (‘Verdopplung, Wiederholung’, entlehnt im 19. Jh. aus spätlateinisch *reduplicātio*),[55] das keine Angabe dazu macht, was genau verdoppelt wird, die Bedeutung oder die Form. Diese Ungenauigkeit des wissenschaftlichen Terminus wirkt bis heute fort. Die konstruktionsgrammatisch orientierte Morphological Doubling Theory (Inkelas & Zoll 2005) sieht Reduplikation als morphologische Konstruktion an, in der zwei Leerstellen mit semantisch und syntaktisch äquivalenten Konstituenten besetzt werden (Inkelas & Zoll 2005: 48); formal orientierte Theorien hingegen gehen von einem Kopierprozess aus, den ein abstraktes Reduplikationsfeature in der syntaktischen Struktur auslöst, und zwar vollkommen unabhängig von Semantik und Lexikon (Kobele 2006: 247). Auch unter den Ansätzen, die auf semantischer Wiederholung basieren, ist fernerhin unklar, ab wann Bedeutung als wiederholt gelten kann. Semantische Äquivalenz kann dabei sehr weit definiert sein und nicht nur Identität, sondern auch Synonymie und Antonymie umfassen. Abraham (2005: 548) zählt etwa bereits Derivate wie *redish* oder Syntagmen wie *red speckled* zur Reduplikation, da hier letztlich Teilbedeutungen redundant sind. So eine Annahme versteht Reduplikation also vor allem als semantische Dopplung und fasst auch solche Dopplungen als Reduplikation auf, bei denen die Konstituenten

55 „Reduplikation“, Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/wb/Reduplikation>, abgerufen am 13.09.2019.

formal distinkt sind (Inkelas & Zoll 2005: 7). Auch für Moravcsik ist es nicht unbedingt notwendig, dass die Bedeutungen der beteiligten Einheiten eines Reduplikationsprozesses identisch sind:

> Constituents to be reduplicated may in principle be definable [...] either by their meaning properties only, or by their sound properties only, or in reference to both.
>
> (Moravcsik 1978: 303f.)

Die synonyme Komposition im Chinesischen (109) zählte nach dieser Auffassung ebenfalls zu den Reduplikationen. Die phonologischen Merkmale der Lautketten *pí* und *fū* in Beispiel (109) sind zwar sehr unterschiedlich, die Bedeutung ist jedoch für beide Morpheme gleich, nämlich 'Haut'. Inkelas und Zoll (2005) nehmen daher an, dass die Identität in der Reduplikation stets semantisch und nicht phonologisch ist. Die Konstituenten sind demnach zwei Töchter, die eine parallele Beziehung zur Konstruktion haben (zur Mutter). Demnach stehen die Konstituenten also nicht direkt miteinander in Verbindung, sondern nur über die Mutter, können eigene Kophonologien haben und müssen folglich auch nicht formidentisch sein (Inkelas & Zoll 2005: 76). So erklären Inkelas und Zoll auch, dass Konstituenten etwa wegen Allomorphie formal unterschiedlich und dennoch reduplikativ sein können (Inkelas & Zoll 2005: 50ff.). Die Phonologie kann dieser Ansicht nach der Semantik folgen und die phonologische Kopie als Extremfall der Assimilation formal identische Konstituenten hervorbringen, tut dies aber nicht notwendigerweise (Inkelas & Zoll 2005: 198). Im Deutschen wäre diesem Ansatz nach auch Wörter wie *Babylein*, *dunkelschwarz* oder *Sanddüne* (partielle Bedeutungs-)Reduplikationen, deren Konstituenten formal distinkt sind. In Bezug auf ICCs würde mit diesem Ansatz zudem das Problem gelöst, dass in ICCs durch etwaige Fugenelemente manchmal ein formaler Unterschied zwischen den Konstituenten besteht. Man könnte die Bildungen zu den Reduplikationen zählen, obgleich die Konstituenten nicht gänzlich formgleich sind.

Auch formale Theorien akzeptieren mitunter die Nichtidentität der Konstituenten in der Reduplikation, erklären diese allerdings ohne die Reduplikation als im Kern semantisch anzusehen. Menn und MacWhinney formulieren zum Beispiel den Repeated Morph Constraint, also einen Bann konsekutiv homophoner Morpheme (Menn & MacWhinney 1984). Nach streng auf phonologischer Identität beruhenden Ansätzen, etwa dem Wilburs (1973), wären Bildungen wie *dunkelschwarz* oder *pífū* 'Haut+Haut=Haut' hingegen gewöhnliche Komposita. Auch Wälchli (2007: 103ff.) kommt nach seiner Untersuchung dieser Ko-Komposita zu dem Schluss, dass es entscheidende Unterschiede zwischen Reduplikationsphänomenen und der Ko-Komposition gibt. Die in der Typologie als Reduplikationen

beschriebenen Phänomene richten sich beispielsweise nach konkreten phonologischen oder morphologischen Regeln, Ko-Komposita hingegen nicht.

Kritik am semantischen Ansatz üben auch Stolz et al. (2011). Wenn allein die Wiederholung von Semantik vonnöten sei, um ein Phänomen Reduplikation zu nennen, müsse man erklären, warum nicht andere semantische Relationen, etwa Antonymien, die Annahme von Reduplikation rechtfertigen (Stolz et al. 2011: 38). Inkelas und Zolls (2005) semantischer Ansatz ist denn auch diese Extremposition, nach der weder die Form noch die Bedeutung von Basis und Reduplikant identisch sein muss und stattdessen bereits eine antonymische Beziehung zwischen den Konstituenten ausreicht, um Reduplikation anzunehmen (Inkelas & Zoll 2005: 61f.). Nach Stolz et al. (2011) öffne diese weite Reduplikationsdefinition „die Büchse der Pandora" (Stolz et al. 2011: 5), da auf diese Weise beliebig viele Konstruktionen unter dem Begriff der Reduplikation behandelt werden könnten. Wälchli (2007) weist ferner darauf hin, dass, wenn Reduplikation allein auf der Semantik fußt, schwer zu erklären ist, warum dieser Prozess an der Wortgrenze Halt macht.

Stolz et al. (2011) nehmen stattdessen die Setzung vor, dass Reduplikation immer die Wiederholung des gesamten sprachlichen Zeichens umfasst (Stolz et al. 2011: 30). Dies ist denn auch die vorherrschende und auch in dieser Arbeit vertretene Auffassung. Dies vorausgesetzt wird Reduplikation in weiten Teilen der Linguistik dann auch recht ähnlich aufgefasst. Die folgenden vier Definitionen verdeutlichen dies:

(a) In Reduplication, part of the base or the complete base is copied and attached to the base.
(Haspelmath 2002: 24)

(b) Reduplication is the doubling of some part of a morphological constituent (root, stem, word) for some morphological purpose.
(Inkelas 2012: 355)

(c) 'Reduplication' is the systematically and productively employed repetition of words or parts of words for the expression of a variety of lexical and grammatical functions.
(Schwaiger 2015: 467)

(d) [Reduplication is] the systematic repetition of phonological material within a word for semantic or grammatical purposes.
(Rubino 2005: 11)

Nach (a) liegt Reduplikation dann vor, wenn eine Basis (oder Teile davon) kopiert und an die Basis angefügt werden, nach (b), wenn ein Teil eines morphologischen Bestandteils zu morphologischen Zwecken verdoppelt wird, nach (c), wenn systematisch und produktiv Wörter (oder Teile davon) zu lexikalischen oder grammatischen Zwecken wiederholt werden und nach (d), wenn phonologisches Mate-

rial systematisch zu semantischen oder grammatischen Zwecken wiederholt wird. Fünf Punkte sind hier hervorzuheben, da sie als Definitionskriterien in den Definitionen dienen: die formale Dopplung, die Systematik, die Adjazenz, der Verweis auf das Wort und der semantisch-lexikalische / grammatische Zweck.

Das Kriterium der formalen Dopplung ist relativ einfach festzustellen. Haspelmath belässt es in seiner Definition sogar dabei und nimmt zu den anderen Punkten Systematik, Wortebene und Zweck nicht Stellung. Ein gemeinsamer Aspekt der Definitionen von Inkelas und Zoll, Schwaiger und Rubino ist die Beschränkung auf die Wiederholung innerhalb eines Wortes. In der Mehrzahl der Forschungsliteratur wird der Terminus Reduplikation allein für morphologische Prozesse verwendet. Die Wiederholung von Lauten über Wörter hinweg und die Wiederholung von Sätzen etwa zur Zitation fallen also gemeinhin nicht unter Reduplikation.[56]

Auf den eingangs erwähnten Aspekt der Systematik gehen die Definitionen von Schwaiger und Rubino ein. Zwar ist Wiederholung in gewisser Weise immer systematisch und dient einem Zweck. Bei der Reduplikation ist es aber die Wiederholung selbst, die Bedeutung trägt. Sie geschieht nicht zufällig (wie etwa die Wiederholung von Lauten in verschiedenen Wörtern), sondern gezielt zur Vermittlung von Informationen. Ein weiterer gemeinsamer Punkt der Definitionen von Schwaiger und Rubino ist, dass die Wiederholung semantisch-lexikalischen oder grammatischen Zwecken dient. Durch Reduplikation verändert sich also die lexikalische oder grammatische Bedeutung eines Wortes, sodass im Sinne Lakoffs und Johnsons (1980b) mit einem Mehr an Material ikonisch auch ein Mehr an Bedeutung einhergeht, also die Bedeutung der Basis etwa in Form von Pluralität, Iterativität, Verteilung oder Intensivierung erweitert wird (ebenso: Botha 1988: 153, Freywald & Finkbeiner 2018: 9, Moravcsik 1978: 317, Rossi 2011: 157).[57] Wichtig ist dieser Punkt, um Reduplikation von Wiederholungen phonologischen Materials innerhalb eines Wortes ohne dieses Mehr an Bedeutung abzugrenzen, etwa wenn Laute oder Lautverbindungen zufällig in verschiedenen Konstituenten eines komplexen Wortes vorkommen (*Herzschmerz*).

56 Hurch et al. (2008) wie auch Maas (2005) gehen allerdings davon aus, Reduplikation sei „a formal linguistic device that can be used at all levels of linguistic structure“ (Hurch et al. 2008: 2, Maas 2005: 395).

57 Moravcsik (1978) etwa sieht stets eine „increased quantity“ gegeben. Botha (1988) sieht ebenfalls [+increased] als Kernmerkmal, Rossi (2011) nennt als Hauptfunktionen intensify/narrow the meaning of the non-reduplicated form und expand the process denoted by the non-reduplicated form.

All diese Definitionskriterien bergen allerdings Ungenauigkeiten. Bereits bei der rein formalen Beschreibung von Haspelmath stellt sich die Frage, wie viel der Basis kopiert werden muss, um die Annahme von formaler Reduplikation zu rechtfertigen. Analog hierzu bleibt die Frage offen, ob zwischen Basis und Reduplikant weiteres sprachliches Material stehen darf. Nur Haspelmaths Definition erwähnt das Kriterium der Adjazenz.[58] Ohne dieses Kriterium wären Bildungen wie etwa *Kindergartenkinder* im Deutschen zumindest auf formaler Ebene ebenfalls zu den Reduplikationen zu zählen. Das Kriterium der Adjazenz entscheidet zudem darüber, ob Phänomene wie in (107) und (111) unter Reduplikation fallen. Zur Lösung dieser Problematik könnte man eine Subklassifikation in totale und partielle Reduplikation vornehmen. Wenn, wie etwa im Hausa (107), ein einziger Laut (in diesem Fall der konsonantische Onset der unmittelbar vorangehenden Silbe) systematisch wiederholt wird, liegt partielle Reduplikation vor.

Diese Einteilung hilft beim Versuch einer möglichst genauen Definition aber nur weiter, wenn man das Kriterium der Systematik hinzuzieht. Nur weil die Wiederholung eines Lauts im Hausa nach bestimmten Mustern und also systematisch geschieht, lässt sie sich sinnvoll zu den partiellen Reduplikationen zählen. Das *-g- in bindig-o:gi*: ist eine partielle Reduplikation, da auch in Fällen wie (107b) der entsprechende Konsonant wiederholt wird. Mit dem Verweis auf die Systematik lässt sich rechtfertigen, warum in anderen Fällen, in denen sprachliches Material unsystematisch wiederholt wird, nicht von Reduplikation ausgegangen werden sollte (*Herzschmerz, …die die Dietriche…*). Damit wären nun die ersten zwei in der Forschung diskutierten und sinnvoll anwendbaren Definitionskriterien identifiziert:

Reduplikation liegt vor, wenn…

I. …sprachliches Material (total oder partiell) wiederholt wird.
II. …die Wiederholung in einer Einzelsprache systematisch zum Ausdruck von Bedeutung geschieht.

Funktionen, Ebenen und Bedeutungen von Reduplikation sind schwieriger zu beschreiben als die rein formale Feststellung und quantitative Klassifikation von systematischer Wiederholung. Welche Funktion genau muss eine Wiederholung erfüllen? Und welche Wortdefinition setzt man an, um Wiederholung auf mor-

58 Im weiteren Verlauf der jeweiligen Texte wird deutlich, dass die Autoren dieses Merkmal aber wohl stillschweigend voraussetzen.

phologischer von Wiederholung auf syntaktischer oder phonologischer Ebene zu trennen? Im Folgenden werden diese beiden Aspekte genauer betrachtet.

In Bezug auf die Funktion besteht eine Schwierigkeit vor allem darin, semantische und grammatische Zwecke von pragmatischen oder stilistischen Zwecken abzugrenzen. Diese Grenzziehung ist Teil von Gils (2005) Versuch, Reduplikation von Repetition zu unterscheiden. Gil (2005: 33) führt sechs Kriterien auf und unternimmt somit den bisher genauesten Versuch, Reduplikation zu definieren (Tabelle 16).

Tab. 16: Kriterien der Unterscheidung Repetition / Reduplikation (nach Gil 2005: 33, Übersetzung MF)

Kriterium	Repetition	Reduplikation
Einheit des Ergebnisses	oberhalb der Wortebene	auf oder unter der Wortebene
Kommunikative Verstärkung	kann vorliegen	liegt nicht vor
Interpretation	ikonisch oder fehlend	arbiträr oder ikonisch
Intonationsdomäne des Ergebnisses	innerhalb einer oder mehrerer Intonationsgruppen	innerhalb einer Intonationsgruppe
Kontiguität der Kopie	zusammenhängend oder zerlegt	zusammenhängend
Anzahl der Kopien	zwei oder mehr	üblicherweise zwei

Nach Gil (2005) liegt Reduplikation vor, wenn sprachliches Material innerhalb einer Intonationsgruppe ununterbrochen genau einmal wiederholt wird und dieser Prozess ohne pragmatischen Wert Wörter oder Wortformen produziert, deren Interpretation ikonisch oder arbiträr ist. All diese Merkmale weist das Beispiel (106) aus dem Twi auf. Bei Repetition hingegen können auch mehrere Intonationsgruppen und mehr als eine Wiederholung beteiligt sein und es entstehen Einheiten über der Wortebene, deren Interpretation ikonisch sein kann. Ein Beispiel für eine solche Repetition wäre etwa das Rephrasieren (Gülich & Kotschi 1996), bei dem Sprecher Sequenzen etwa zur Aussagebekräftigung oder Verständnissicherung wiederholen. Hier wird also zum einen angenommen, dass die Basis bei der Reduplikation genau einmal kopiert wird. Zum anderen sieht Gil als wichtiges Unterscheidungskriterium die semantische Ebene an. Bei der Reduplikation ändert die Wiederholung die Basisbedeutung, bei der Repetition hingegen wird die Basisbedeutung lediglich wiederholt (ebenso: Goodwin Gómez & van der Voort

2014: 2), wodurch es zu einer kommunikativen Verstärkung kommt („communicative reinforcement", Gil 2005: 33).[59] Als weiteres Definitionskriterium für Reduplikation gilt also, dass der Bedeutungszuwachs primär der Grammatik und Lexik, und nicht der Kommunikationssituation dient. Diese beiden Merkmale eignen sich, um Reduplikation als grammatischen Prozess von Repetition als pragmatisch-stilistischem Prozess abzugrenzen. Zwei weitere Definitionskriterien für Reduplikation lauten also:

Reduplikation liegt vor, wenn…

III. …die Basis bei der Reduplikation genau einmal kopiert wird.
IV. …der Bedeutungszuwachs primär der Grammatik und Lexik (und nicht der Kommunikationssituation) dient.

Eine zentrale Annahme, die sowohl bei Gil als auch bei den anderen Definitionsversuchen vorliegt, ist die Unterscheidung der grammatischen Ebenen. Wie später Inkelas und Zoll (2012) und Schwaiger (2015) sehen es bereits Rubino (2005) und Gil (2005) als entscheidendes Kriterium an, dass bei der Wiederholung ein Wort und keine Einheit oberhalb des Wortes entsteht. Es wird also weithin angenommen, dass Reduplikation ein morphologischer Prozess ist und die daraus entstehenden Einheiten komplexe Wörter sind, da die individuellen Konstituenten für Flexion, Derivation oder Syntax nicht mehr zugänglich sind (Botha 1988: 11f.).[60] Als letztes Definitionskriterium gilt also:

Reduplikation liegt vor, wenn…

V. …ein morphologischer Prozess komplexe Wörter bildet.

Gil selbst merkt an, dass dieses Kriterium voraussetzt, dass man Wörter und Wortgrenzen klar bestimmen kann (Gil 2005: 33). Dies ist aber, vor allem über

59 Gil geht bei einigen Kriterien davon aus, dass bestimmte Merkmale bei der Reduplikation vorliegen müssen, bei der Repetition hingegen vorliegen können. Somit fasst er Reduplikation als Subtyp der Repetition auf. Eine solche Position nehmen auch Inkelas und Zoll (2005: 1) ein. Sie sehen in einer Wiederholung sprachlichen Materials zunächst immer eine Repetition. Ist die Dopplung allerdings grammatischer Natur und geschieht sie innerhalb von Wörtern, nehmen sie den Repetitions-Subtyp Reduplikation an.

60 In generativen Ansätzen wird zudem angenommen, dass lediglich „major lexical categories" für Reduplikation zur Verfügung stehen (Aronoff 1976: 21).

Einzelsprachen hinweg, schwierig und wird mitunter als Grund dafür genannt, dass Reduplikation eben nicht zufriedenstellend von anderen Phänomenen unterschieden werden kann (Stolz & Levkovych 2018: 30).

Die Abhängigkeit von der Wortdefinition ist aber nicht das einzige Problem der Definitionsversuche von Gil und anderen Sprachtypolog:innen. So finden sich in vielen Sprachen Phänomene, die Eigenschaften von Repetition und Reduplikation gleichermaßen aufweisen. Verdoppelte Verbalphrasen im Türkischen beispielsweise sind als Phrasen der Repetition zuzuordnen. Der Prozess weist aber viele Eigenschaften der Reduplikation auf, wie Identität, Adjazenz und die Beschränkung auf einen Reduplikanten (Erbaşı 2018).[61] Die Annahme, Reduplikation sei allein auf Wörter oder Einheiten unter der Wortebene beschränkt, ist also nicht so ohne Weiteres haltbar. Auch das Kriterium, dass bei der Reduplikation genau einmal wiederholt wird, das neben Gil auch Moravcsik (1978: 315) anführt, eignet sich nicht optimal dazu, Reduplikation zu definieren. Das Kriterium verliert zum einen seine Diskriminierungskraft, sobald einfach repetiert wird, etwa bei der asyndetischen Reihung im Deutschen: *komm, komm!* (Schindler 1991: 602). Zum anderen gibt es Fälle multipler Reduplikation, in denen also mehr als einmal wiederholt wird und dennoch alle anderen Kriterien für die Annahme von Reduplikation sprechen (Stolz & Levkovych 2018: 58).

Alle bisher besprochenen Kriterien helfen für sich genommen nicht dabei, Reduplikation kategorial erfassen zu können. Freywald und Finkbeiner (2018) führen das etwa vor Augen, indem sie Dopplungsphänomene im Deutschen und Englischen beschreiben, die sich im Graubereich zwischen Reduplikation und Repetition bewegen.

(112) *Mindestlohn hin, Mindestlohn her, unser Hauswein bleibt weiterhin gewohnt günstig.*
(Freywald & Finkbeiner 2018: 11)

Diese Konstruktion in (112) unterliegt klaren semantischen und grammatischen Beschränkungen und die Wiederholung trägt innerhalb der Konstruktion klar eine eigene Bedeutung. Beides spricht dafür, das Phänomen der Reduplikation zuzuordnen. Die Wiederholung umfasst allerdings mehr als ein Wort und ihr kommt zudem eine pragmatische Funktion zu, Eigenschaften, die Gil der Repeti-

61 Erbaşı (2018) löst dieses Dilemma, indem er die Verbalphrasen im Türkischen als „mid-sized structures“ ansieht, die das Ergebnis eines Prozesses sind, der zwischen Reduplikation und Repetition liegt.

tion zuspricht. Das englische Beispiel *fixer-upper*, bei dem das Derivationssuffix *-er* doppelt vorkommt, zeigt ferner, dass sich bei der Reduplikation die Semantik der Basis nicht notwendigerweise ändert, sondern auch bloß die Bedeutung eines Affixes, in diesem Fall die Agentivität, verstärkt werden kann (Lensch 2018).

Eine klare Grenzziehung zwischen Reduplikation und ähnlichen Prozessen ist also nicht kategorial möglich. Die in der Forschung diskutierten Merkmale Vollständigkeit, Wortebene, Adjazenz, Basisbedeutungsveränderung und die Beschränkung auf einen Reduplikanten sind weder notwendige, noch hinreichende Definitionsmerkmale. Kategoriale Einteilungsversuche scheitern stets an der Vielfalt der in den Sprachen der Welt beobachtbaren Wiederholungsphänomene. Zudem wirkt auch die Annahme, wonach die Wortgrenze das Hauptunterscheidungsmerkmal von Reduplikation und angrenzenden Phänomenen ist, willkürlich, da sich Wiederholungsphänomene auf syntaktischer und morphologischer Ebene oft in vielen anderen Punkten sehr ähnlich sind.

In der jüngeren Forschung werden darum vermehrt graduelle Beschreibungsansätze gewählt (Freywald & Finkbeiner 2018, Stolz & Levkovych 2018, Stolz et al. 2011). Wiederholungsphänomene werden hierbei nicht kategorial, sondern relativ der Reduplikation zugeordnet. Wie schon bei der Klassifikation der ICCs als Komposita in Kapitel 7 werden ICCs auch zur Reduplikation relativ in Beziehung gesetzt. Die in dieser Arbeit behandelten Merkmale, die ein Phänomen dem prototypischen Kern der Reduplikation näherbringen, sind die folgenden, aus der Literatur abgeleiteten Kriterien A–F. Sie betreffen sowohl die Form als auch die Bedeutung einer Wiederholung:

Die Form betreffend:

A. Die Basis und der angefügte Reduplikant sind formgleich.
B. Der Reduplikant wird unmittelbar vor oder nach der Basis angefügt.
C. Der Reduplikant wird genau einmal an die Basis angefügt.

Die Bedeutung betreffend:

D. Die Wiederholung selbst ist Träger von Bedeutung.
E. Die Basis und der angefügte Reduplikant sind bedeutungsgleich.
F. Die entstehende Bedeutung ist grammatischen, nicht pragmatischen Charakters.

Diese Kriterien sind flexibel anzuwenden. Jedes Kriterium ist also in sich wiederum graduell. Die Form- und Bedeutungsgleichheit (A, E) liegt mal mehr vor, wie etwa bei *abo*-*abó* in (106), mal weniger, wie bei *pífū* in (109). Bei partieller (formaler oder semantischer) Reduplikation liegen sie nicht so stark vor, wie bei totaler.

Die Adjazenz (B) ist bei *abo-abó* mehr gegeben als bei *bindig-o:gi:* in (107a). Dem Kriterium, dass genau eine Wiederholung vorliegt (C), kann in Potenz und Restriktion mehr (*Mindestlohn hin, Mindestlohn her*) oder weniger (*sehr, sehr, sehr, sehr gut*) entsprochen werden. Dass die Wiederholung selbst Bedeutung trägt (D) und diese eher grammatisch denn pragmatisch ist (F), ist bei *des milliers et des milliers* in (111) schließlich weniger gegeben als bei *bindig-o:gi:* in (107a).

Auch hier können die Merkmale dazu noch gewichtet werden. Wenn einem Phänomen Merkmal A fehlt, sollte das auf seine Zugehörigkeit zur Reduplikation einen größeren Einfluss haben, als wenn etwa Merkmal C fehlt. Ich gehe daher von folgender Gewichtung aus: A>B>C; D>E>F sowie [A, B, C]=[D, E, F].

Für die zu Beginn dieses Kapitels besprochenen Beispiele bedeutet diese Merkmalsidentifikation: (104) und (105) weisen nur wenige der Merkmale auf. (104) erfüllt lediglich die formalen Merkmale, (105) auch hier bloß A und B. Beiden Beispielen aber fehlen sämtliche die Bedeutung betreffenden Merkmale. Damit sind diese beiden Fälle nicht als Reduplikationen anzusehen oder höchstens in der weiteren Peripherie zu verorten. Ebenso (111), bei dem die Merkmale B, D und F fehlen. (107) fehlt lediglich Merkmal B und zu E kann keine Aussage getroffen werden, da die Basis keine Bedeutung trägt. Dem Chinesischen Beispiel (109) fehlen die wichtigen Merkmale A und E. (106) hingegen weist alle genannten Merkmale auf, allerdings mit der Einschränkung, dass Basis und Reduplikant nicht vollständig formgleich sind. Das chinesische (108) und das bretonische Beispiel (110) weisen alle Merkmale auf und kommen damit von allen vorgestellten Sprachbeispielen dem Reduplikationsprototypen am nächsten.

Diese aus der Forschung abgeleiteten Merkmale für Reduplikation umfassen nun also sechs Kriterien, von denen keines allein Wiederholungsphänomene angemessen der Reduplikation zuordnen kann. Zusammengenommen ermöglichen sie aber eine gute Einschätzung, inwieweit ein zu untersuchendes Phänomen der Reduplikation entspricht und welche Eigenschaften der Prozess mit anderen Prozessen in anderen Sprachen teilt. Diese Merkmale werden im Folgenden auf (mögliche) Reduplikationsphänomene und ICCs im Deutschen angewendet.

8.3 Reduplikation im Deutschen

In der Geschichte der deutschen Sprache gab es eine Zeit, in der Reduplikation ein regelhafter, morphologischer Prozess war. Im Germanischen gab es bei der Bildung der Präteritalformen starker Verben eine komplementäre Verteilung von Ablaut und Reduplikation: Wo Ablaut vorhanden oder möglich war, fehlte die Reduplikation; wo das Präteritum mit Reduplikation gebildet wurde, fehlte der Ablaut (Austefjord 1979, Meid 1983: 333). Für die Bildung von Präteritalformen war

Reduplikation also in einem kerngrammatischen Bereich aktiv. Reduplikation diente schon im Indogermanischen zur Perfektbildung. Am Ende der gemeingermanischen Zeit wird die Reduplikationssilbe getilgt (Mächler 2017: 55):

(113) a.	got. *l̲aí-l̲ôt*	→	*lêtan*
	‘ließ’		‘ließ’
b.	got. *haí-hait*	→	*haitan*
	‘hieß’		‘hieß’
	(Mächler 2017: 55)		

In allen folgenden Sprachstufen des Deutschen gibt es keine Reduplikation im kerngrammatischen Bereich mehr. In der Reduplikationsforschung der modernen Sprachwissenschaft dient das Deutsche darum oft dazu, die Abwesenheit dieses Merkmals in den europäischen Sprachen zu illustrieren (Rubino 2005, 2013, Stolz 2006). Das Deutsche wird hier sogar häufig als gänzlicher „reduplication avoider“ beschrieben (Freywald 2015: 905, Lohde 2006: 43, Stolz 2006: 115, Stolz et al. 2011: 565). Stolz et al. (2011) nehmen an, dass die Sprecher:innen in Sprachen wie dem Deutschen jedweder Redundanz kritisch gegenüber eingestellt sind und sich die Vermeidung von Reduplikation teilweise daraus ableitet (Stolz et al. 2011: 5). Menn und MacWhinney (1984) nehmen sogar an, dass Sprachen generell konsekutiv homophone Morpheme vermeiden. Der bisherigen Forschung nach zu urteilen ist das Deutsche aber besonders restriktiv gegenüber Reduplikation, da auch die in anderen europäischen Sprachen auftretenden reduplikativen Prozesse im Deutschen fehlen. Zu dem bereits behandelten Bretonischen Beispiel *tost tost* ‘sehr nah’ (Stolz et al. 2011: 497) gibt es im Italienischen ein Pendant, bei dem ebenfalls die Adjektivsteigerung produktiv reduplikativ ausgedrückt wird (*bella* ‘schön’, *bella bella* ‘sehr schön’, Wierzbicka 1986). Zudem werden bereits seit Jahrhunderten handlungsbezogene Nomina durch die Reduplikation von verbalen Imperativformen gebildet (*fuggi* ‘Renn!’, *fuggi fuggi* ‘Massenflucht’, Thornton 2009). Auch im Englischen existiert eine produktive Form der Reduplikation, bei der durch die Dopplung von Spitznamen (*Jon-Jon*, *Lou-Lou*) Nähe erzeugt wird („name doubling“, Aziz & Nolikasari 2020, Kauffman 2015).

Im Deutschen gilt Reduplikation hingegen als „nur ansatzweise“ existent (Schindler 1991: 597), in der Wortbildung nicht produktiv (Borgert & Nyhan 1975, Donalies 2011: 72, Neef 2009: 632, Rubino 2005) oder „bestenfalls randständig“ (Stefanowitsch 2007: 33). Nicht nur wird Reduplikation (von linguistischer Seite: Glück 2010: 552, Schindler 1991, 597) für unproduktiv gehalten, sondern auch (von Sprecher:innenseite) negativ bewertet (Stolz et al. 2011: 5). In der Tat gibt es Hinweise darauf, dass die Sprecher:innen des Deutschen eine generelle Abneigung

gegenüber jeglicher Redundanz hegen, nämlich Beispiele, in denen formale oder semantische Redundanz sanktioniert wird. Formale Wiederholung gilt etwa der Dudenredaktion (2016: 1027) als „stilistisch unschön“ und „erweckt […] leicht den Eindruck mangelnder sprachlicher Flexibilität.“ In Sätzen wie *Er reiste für einige Zeit, um Zeit zu gewinnen* rät die Dudenredaktion daher, auf Synonyme auszuweichen, also etwa ein Auftreten von *Zeit* im genannten Satz durch *Tage* zu ersetzen (Dudenredaktion 2016: 1028). Der Stilkritik populärwissenschaftlicher Texte gilt aber auch diese Wiederholung von Bedeutung als vermeidenswert. Sick (2009) etwa rät den Sprecher:innen des Deutschen, Pleonasmen in morphologischen (*Glasvitrine*) und syntaktischen Bildungen (*für gewöhnlich etwas zu tun pflegen*) durch distinktive (*Glasschrank* / *Vitrine*; *gewöhnlich etwas tun* / *etwas zu tun pflegen*) zu ersetzen (Sick 2009: 29ff.). Stolz et al. (2011) sehen solche Ausführungen präskriptiver Grammatiken und populärwissenschaftlicher Texte als möglichen Grund für eine „anti-T[otal] R[eduplication] attitude“ der Sprecher:innen an (Stolz et al. 2011: 5).

Dennoch existieren unbestritten Wörter reduplikativer Struktur im Deutschen, wie etwa *Tamtam*, *Wirrwarr* oder *Heckmeck*. Einige dieser Wörter befinden sich schon seit Jahrhunderten im deutschen Wortschatz. *Wirrwarr* beispielsweise ist bereits 1518 in Johann Geiler von Kaysersbergs ‘Buch von den Sünden des Munds’ belegt (114):

(114) *wer da zwischen zweien personen, die frid mit einander haben […], wirrewerre macht.*
(Geiler von Kaysersberg 1518: 47)

Als erster beschrieben hat solche Bildungen August Friedrich Pott (1862). Bzdęga (1965) liefert erstmals eine umfassende Sammlung und beziffert die Menge der Wörter reduplikativer Struktur im Deutschen auf 1.880, vermutet aber eine sehr viel höhere Zahl, da diese Wörter im Substandard zu verorten und darum nicht in Lexika zu finden seien (Bzdęga 1965: 4, ebenso: Fleischer & Barz 2012: 95).[62]

Die linguistische Beschreibung dieser Reduplikationen stellt aber im Deutschen eine besondere Herausforderung dar. So meint etwa Barz (2016: 2407): „They can hardly be dealt with systematically.“ Bzdęga (1965: 3) spricht von „problemreiche[n] Erscheinung[en].“ Problemreich ist vor allem die Wahl des Prozes-

62 In einer älteren Auflage sprechen Fleischer und Barz noch davon, dass reduplikative Bildungen im Deutschen in die Literatursprache gelangen können, dort aber „vorwiegend pejorativen Charakter“ haben (Fleischer & Barz 1995: 48).

ses, dem man die Bildungen zuordnen möchte. Meist wird hier von Reduplikationen gesprochen.[63] Elsen ordnet sie hingegen den Komposita zu und nennt Bildungen wie *Heckmeck* „Reduplikativkomposita“ (Elsen 2014: 29). Schindler (1991: 597) sieht in den Wörtern lediglich Relikte. Sie seien „beträchtlichen Alters“, rezente Bildungen gebe es nicht und der Prozess dahinter sei im Deutschen unproduktiv (Schindler 1991: 597). Die existenten Bildungen fielen überdies „zahlenmäßig kaum ins Gewicht“ (Schindler 1991: 598), und seien teilweise zufällig.

Legt man die in 8.2 erarbeiteten Definitionskriterien an, wird deutlich, dass diese Bildungen keine allzu prototypischen Vertreter der Reduplikation darstellen. Die Bestandteile sind zwar teils (*Wirrwarr*, *Heckmeck*) oder gänzlich (*Tamtam*) formgleich (Kriterium A). Die meisten Kriterien, vornehmlich die zur Bedeutung, liegen hingegen nicht vor. Kriterium D und E sind nicht gegeben, weil Basis und Reduplikant keine Bedeutung besitzen und nicht systematisch zur Wortbildung verwendet werden. Auch trägt die Wiederholung selbst keine Bedeutung.

Einige Arbeiten nehmen hier allerdings doch einen sehr vagen Bedeutungskern solcher Bildungen an. Vater (2010) beispielsweise sieht hier die Funktion des additives Wortspiels, bei dem „durch Addition eines Elements […] ein komischer Effekt erzielt wird“ (Vater 2010: 13). Glück (2010: 552) beschreibt die Bildungen als „Geräuschwörter […oder…] Naturlaute“, die „Kennzeichen einer Phase des Spracherwerbs (redupliziertes Silbenplappern)“ seien. Ebenso sorgt nach Fleischer und Barz (2012) die Reduplikation „sprachspielerisch für intensivierenden, ironischen oder scherzhaften Ausdruck“ (Fleischer & Barz 2012: 95). Donalies erkennt in Wörtern reduplikativer Struktur „immer etwas Legeres“ (Donalies 2011: 72) und Erben (1981: 39) reduziert die Reduplikation auf den Bereich der Lautmalerei, der Unsinnswörter und des frühkindlichen Spracherwerbs. Die Bildungen seien „abschätzige Ausdrücke der saloppen Umgangssprache“ (Erben 1981: 39). Ihre Bedeutung kann aber nicht aus den Bestandteilen plus Wiederholungsbedeutung erschlossen werden, auch weil die Basen keine Bedeutung tragen. Sprachliches Material wird stattdessen im Zuge eines rein phonologischen Kopierprozesses (Wiese 1990: 612) oder einer „lautassoziative[n] Dopplung“ (Hent-

63 Bei Wörtern wie *Tamtam* wird oft von totaler oder einfacher Reduplikation gesprochen (Bzdęga 1965: 8, Fleischer & Barz 2012: 94); alterniert wie bei *Wirrwarr* der Vokal von Ablautreduplikation, alterniert wie bei *Heckmeck* der Konsonant von Reimreduplikation (Bzdęga 1965: 9f.). Auch in in vielen neueren Standardwerken zur Grammatik und Wortbildung des Deutschen wie etwa der Dudengrammatik oder Fleischer und Barz (2012) finden derlei Prozesse unter dem Oberbegriff „Reduplikation“ Berücksichtigung. In allen genannten Bildungen trägt aber mindestens ein Bestandteil keine Bedeutung, weder bei der einfachen Reduplikation (**tam*), noch bei der Reimreduplikation (**heck, *meck*), noch bei der Ablautreduplikation (**warr*).

schel & Vogel 2009: 470) zur Stammerweiterung wiederholt. Da die formalen Merkmale hier teilweise vorliegen, kann man von einer reduplikativen Wortstruktur sprechen. Durch das Fehlen der Bedeutungsaspekte ist der Prozess hinter Wörtern wie *Tamtam*, *Wirrwarr* oder *Heckmeck* aber nicht als Reduplikation einzuordnen.

Hier kann man einwenden, dass dieser phonologische Prozess auch auf bedeutungstragende, morphologische Basen und komplexe Morpheme angewendet wird. Wiese (1990) nimmt beispielsweise folgende Analyse der Wortform *rumpumpeln* vor (115):

(115) *rumpeln* → *rum [pum] peln*

Die Form entsteht nach Wiese aus der morphologischen Basis *rumpeln*, einer Kopie des Reims der ersten Silbe [ʊm] und dem Onset der zweiten Silbe [p]. Eine segmentale Kopie resilbifiziert hier eine ebenfalls segmental kopierte, unterspezifizierte Silbe (Wiese 1990: 608). Allerdings ist das hinzutretende Element selbst kein lexikalisches Morphem (**pum*).[64] Die Dopplung geschieht rein phonologisch (Wiese 1990: 612) und erweitert sozusagen als Reparaturprozess (Kentner 2017: 237, Kirchner 2010: 1) oder als „quantitative[…] Erweiterung" (Nübling 2004: 26) das sprachliche Material des Ausdrucks.[65] Eine klare oder gar grammatische Bedeutung trägt diese Lautwiederholung nicht.[66]

Selbst wenn die Lautwiederholung systematisch geschieht, heißt das nicht automatisch, dass ihr auch eine eigene Bedeutung zukommt. In jüngerer Zeit hat beispielsweise Kentner (2017) mit einer empirischen Studie zeigen können, dass

64 Elsen (2014: 73) zählt Bildungen wie *Singsang* allerdings zur Komposition und die Konstituenten zu den Grundmorphemen (Elsen 2014: 61).

65 Wenn der Reduplikant ebenfalls ein Lexem ist (*Wimperklimper*), kann man auch, wie Kentner (2017: 245), Komposition annehmen. In manchen Fällen (*Krimskrams*) ist die Basis nur noch diachron transparent. *Krims-* stammt laut DWDS wohl von *krimmeln* 'kribbeln' ab („Krimskrams", in: Pfeiffer (1995), https://www.dwds.de/wb/Krimskrams>, abgerufen am 08.08.2018. Ab wann man hier noch einen Kopierprozess und ab wann bereits einen Kompositionsprozess ansetzt, lässt sich nicht immer ohne Weiteres bestimmen.

66 Ähnlich einzuordnen ist das Sprachspiel „Ubbi dubbi". Hierbei wird an jeden Silbenonset die Silbe /ʌb/ angefügt und betont. Beim Englischen *hello*, ist das Ergebnis beispielsweise *hubellubo*. Diese spielerische iterative Infigierung (Yu 2008) kam im Amerika des 17. Jahrhunderts auf und wurde in jüngerer Vergangenheit vor allem durch die PBS Kindersendung „Zoom" bekannt. Über weitere popkulturelle Erzeugnisse, etwa die Fernsehserie „The Big Bang Theory", gelangte das Sprachspiel auch in den deutschsprachigen Raum.

Reim- und Ablautreduplikation im Deutschen systematisch und produktiv angewendet wird, vor allem zur Bildung von Benutzernamen im Internet (etwa *Sillepille* aus dem Vornamen *Silke*). Auf der Grundlage von Akzeptabilitätstests kommt er zu dem Schluss, dass Sprecher:innen Reim- und Ablautreduplikationen nach bestimmten Regeln bilden. So akzeptierten die Versuchspersonen vor allem Bildungen mit trochäischer Silbenstruktur, bei denen der Reduplikant suffigiert ist (*Hansipansi*). Wird hingegen nur silbisch präfigiert, lehnten die Versuchspersonen die Bildungen größtenteils ab (*Hahansi*). Auch abseits von Benutzer- und Kosenamen verfügen die Sprecher:innen über sprachliches Wissen zur Bildung von Reim- und Ablautreduplikationen. So seien Formen wie *Mipsmops*, die den Ablautregeln entsprechen, einheitlich besser bewertet worden als Formen wie *Mopsmips*, die die Ablautregeln verletzen. Die Bewertungssätze, mit denen Kentner (2017) die 5-Punkt-Likertskala im Versuchsdesign beschriftet, lässt zudem den Schluss zu, dass diese Formen nicht nur auf ihre korrekte Ablautung hin bewertet wurden, sondern in Bezug auf die tatsächliche, alltägliche Verwendung der Formen durch die Sprecher:innen.[67] Kentners Studie zeigt somit, dass Bildungen wie *Mipsmops* zwar auf bestimmte Bereiche beschränkt sind, aber sehr systematisch geschehen. Trotzdem wird von der Lautdopplung keine Bedeutung transportiert.

Reduplikative Strukturen finden sich im Deutschen auf weiteren Ebenen. Oberhalb der Wortebene gibt es beispielsweise Dopplungen, die in gesprochensprachlicher, kommunikativer Praxis etwa der Planung weiterer Redesequenzen dienen oder die Turnübernahme durch den Kommunikationspartner verhindern. In der Gesprächsanalyse wurde außerdem die Technik des Rephrasierens zur Verständnissicherung beschrieben (Gülich & Kotschi 1996, Schwitalla 2006: 180):

(116) Rephrasieren (Schwitalla 2006: 179)
die ärzte tun WIRKklich ALles;
<<f>↑die TUN ALlES?
TUN ALles.>

67 Kentner (2017) formuliert hier lediglich auf Englisch: Der Maximalwert von 5 Punkten wurde auf dem Bewertungsbogen demnach mit dem Satz „I could use this word myself" spezifiziert, der Wert 4 erhielt die Formulierung "conceivable word, heard before". Der Durchschnittswert von *Mipsmops* befindet sich zwischen diesen beiden Werten, was nahelegt, dass diese Form tatsächlich als brauchbar angesehen wird.

Dabei werden, etwa in Streitgesprächen, ganze Formeln wiederholt. Im folgenden Beispiel erkennt man, dass beide Sprecher nahezu wortwörtlich gleiche Formeln wiederholen, und zwar sowohl eigene als auch Formeln des Gegenübers:

(117) Wiederholung von Formeln (Schwitalla 1996: 329)

257	A:	*bollizei" hot-*	** die bollizei hot=s io schriftlisch*
258	B:	*sa:ge was die leut sage die leut sage viel!*	
259	C:	*ja:-*	*ja gas gsacht werd gsacht werd vie:l*
260	A:	*ufigenummef*	*dann hot sich die frau" des habbe die leit*
261	C:		*gsacht werd viel!*

Auch zu den Doppelstrukturen gezählt werden Phrasen wie die X-*und*-X-Konstruktionen (118) und – wegen des identischen Anlautes – koordinierende Paarformeln:

(118) X-*und*-X-Konstruktionen (Finkbeiner 2012: 1)
a. *Schade dass die so teuer sind*
b. *Naja, teuer und teuer – wenn die Qualität stimmt*

(119) Koordinierende Paarformeln (Müller 1997: 2)
Haus und Hof

Verstärkende Funktion hat die Wortiteration, die im Deutschen vor allem bei Adverbien, Quantifikatoren und attributiven Adjektive angewandt wird:

(120) Iteration von Adverbien (Stolz 2006: 116)
ganz ganz viele tolle Referate

(121) Iteration von Quantifikatoren (Stolz 2006: 116)
viele viele tolle Referate

(122) Iteration von Adjektiven (Stolz 2006: 116)
ganz viele tolle tolle Referate

Auch andere Wortarten können verdoppelt werden. Verben (vor allem Inflexive, Freywald 2015), Diskurspartikeln und Interjektionen werden so verstärkt („intensive/augmentative Verstärkung“, Stefanowitsch 2007: 41) oder zur Emphase wiederholt. Schindler (1991: 602) spricht hier von asyndetischer Reihung (123),

Stefanowitsch (Stefanowitsch 2007: 32) von asyndetischer Dopplung (124), Kentner von lexikalischen Sequenzen („lexical sequences“, Kentner 2017: 234):

(123) Asyndetische Reihung (Schindler 1991: 602)
komm, komm! bibber bibber!

(124) asyndetische Dopplung (Stefanowitsch 2007: 36)
Ich will rufen: lauf, lauf'. Lauf weg von hier'

(125) Lexikalische Sequenzen (Kentner 2017: 244)
los los

Aber auch diese Phänomene entsprechen nicht den Kriterien für Reduplikation. So ist „[e]in semantischer Effekt der Iteration […] kaum wahrzunehmen“ (Schindler 1991: 602), die Bedeutung der Wiederholung hat keine oder nur eine sehr periphere, pragmatische Bedeutung. Auch die anderen Beispiele erfüllen nur teilweise die formalen und semantischen Kriterien. Ohne das Argument zu bemühen, dass Reduplikation notwendig ein morphologisch operierender Prozess sei (Kouwenberg & LaCharité 2015: 59, Rubino 2013: 7), kann man annehmen, dass auch diese Phänomene keine Reduplikation im Deutschen darstellen.

8.4 ICCs als Reduplikation im Deutschen

ICCs sind nun allerdings ein Phänomen, das die Annahme von Reduplikation im Deutschen möglicherweise wieder rechtfertigt. Zunächst kann man feststellen, dass Det-ICCs wie *Kindeskind* nicht das Ergebnis von Reduplikation sind. Für Schindler (Schindler 1991: 603) „bedarf es [dazu] keiner näheren Erläuterung.“ Versucht man dennoch eine Begründung dafür zu finden, dass Det-ICCs keine Reduplikationen darstellen, kann man die in 8.2 aufgestellten Kriterien für Reduplikation auf ICCs anwenden, mit dem Ergebnis, dass das wichtige, die Bedeutung betreffende Kriterium D nicht erfüllt ist. In einem ICC wie *Freundesfreunde* oder *Zinseszins* trägt die Wiederholung selbst keine Bedeutung. Auch bei ad hoc gebildeten Det-ICCs wie *Fenster-Fenster* in (48) erhält die Bildung ihre Bedeutung aus den Konstituenten, die Motivationsbedeutung beruht auf dem Kompositumsschema des Deutschen. Die Tatsache, dass die beiden Konstituenten identisch sind, ist für die Bedeutung der Bildung unerheblich, was man daran erkennt, dass *Fenster-Fenster* ‘Fenster am Fenster, Fenster für Fenster’ kein Mehr an Bedeutung enthält im Vergleich mit N+N-Komposita, in denen die Konstituenten nicht identisch sind, etwa *Autofenster* ‘Fenster am Auto, Fenster für Autos’. Da durch die

Dopplung keine Bedeutung ausgedrückt wird, erübrigt sich auch die Frage, ob diese grammatisch oder pragmatisch ist. Kriterium F fehlt der Bildung von Det-ICCs also auch. Das wichtigste die Form betreffende Kriterium ist zwar gegeben (A). Doch sind Basis und Reduplikant durch die in dieser Arbeit nachgewiesenen Fugenelemente fast nie adjazent (B) und die Rekursivität führt dazu, dass nicht immer genau einmal wiederholt wird (C). Man kann also von einer partiell reduplikativen formalen Struktur der Komposita sprechen. Auf der Bedeutungsseite sind die Konstituenten zwar bedeutungsgleich, was dem definitorischen Kriterium G genüge tut. Doch ohne eine Funktion, die die Wiederholung selbst ausdrückt, kann man nicht von Reduplikation als einem grammatischen Prozess sprechen.

Bei Prot-ICC ist das anders. Sie entsprechen zusätzlich zu den Kriterien A, B und C auch allen Kriterien der Bedeutung. Die Konstituenten sind bedeutungsgleich und – das ist das Wichtige – die Wiederholung selbst vermittelt Bedeutung. Die Prototypenbedeutung erhält ein Prot-ICC nicht allein durch das Kompositumsschema, sondern durch das Prot-ICC-Muster, das das funktionale Erstglied anzeigt. Zwar muss die Bedeutung vom Kontext unterstützt werden. Finkbeiners (2014) Studie konnte aber zeigen, dass sich die Prototypenbedeutung bereits vom Kontext emanzipiert hat und der Reduplikationsstruktur der Bildungen selbst innewohnt. Diese Dopplung dient zudem der Wortbildung und ist damit grammatisch, was wiederum Kriterium F und den gängigen Definitionen entspricht (Rubino 2005: 11, Schwaiger 2015: 467). Es sind bei Prot-ICCs also alle Definitionskriterien für Reduplikation erfüllt. Die Prot-ICC-Bildung kann darum aus guten Gründen als Reduplikation angesehen werden.

Name-ICCs hingegen sind, wie schon bei der Zuordnung zur Komposition, nicht so einfach einzuordnen wie Prot- und Det-ICCs. Sie weisen zwar ebenso eine reduplikative Struktur auf, weshalb sie mitunter als Reduplikationen angesehen werden (Kauffman 2015: 3f.). Doch ist die Frage nicht ohne weiteres zu beantworten, ob sie als Eigennamen ohne lexikalische Semantik durch ihre Struktur Bedeutung vermitteln können. Nähme man an, dass ihre reduplikative Struktur dabei hilft, den Rezipient:innen die Substantivklasse der Name-ICCs anzuzeigen (= Klasse der Eigennamen), hätte die Dopplung durchaus eine Funktion. Die entscheidende Frage ist hier, ob die reduplikative Struktur von Name-ICCs diese Bedeutung in ähnlicher Weise ohne allzu viel Hilfe kontextueller Anreicherung anzeigen kann wie die der Prot-ICCs die Prototypenbedeutung. Da es hierzu noch keine Studien gibt, kann man nicht mit Bestimmtheit sagen, dass Name-ICCs, über ihre reduplikative Struktur hinaus, auch das Ergebnis eines Reduplikationsprozesses sind. Für den Moment müssen die die Bedeutungen betreffenden Kriterien D–F als nicht erfüllt angesehen werden.

Zusammenfassend kann man sagen, dass die Nähe zu kanonischer Reduplikation bei Prot-ICCs am höchsten ist. Hier werden alle aufgestellten Definitionskriterien erfüllt und man kann Prot-ICCs deshalb als Reduplikationen ansehen. Wie in Kapitel 7 ausgeführt, erfüllen Prot-ICCs zusätzlich alle vier definitorischen Kriterien der Komposition. Allein die weniger wichtigen Kriterien, die bloß Eigenschaften darstellen, die generell mit Komposition verbunden werden (Zusammenschreibung, Rekursivität, semantische Relation zwischen den Konstituenten) sind bei Prot-ICCs nicht gegeben. Die Bildung von Prot-ICCs kann damit als Kompositionsprozess angesehen werden, der gleichzeitig Reduplikation ist.

Det-ICCs hingegen sind eindeutig keine Reduplikationen. Sie erfüllen nur wenige Merkmale der Reduplikation. Wichtig ist vor allem, dass die Wiederholung selbst keine Bedeutung trägt. In Kapitel 7 wurde herausgearbeitet, dass Det-ICCs stattdessen alle Merkmale der Komposition aufweisen. Die Bildung von Det-ICCs kann somit als Kompositionsprozess angesehen werden, dessen einzige Besonderheit ist, dass die entstehenden Bildungen eine reduplikative Struktur aufweisen. Det-ICC sind also das Ergebnis der Komposition.

Bei Name-ICCs ist die Zuordnung zur Reduplikation schwierig. Zwar entstehen auch hier, wie bei den anderen ICC-Typen, Wortbildungen mit reduplikativer Struktur, und das sogar exakter als bei Det-ICCs, da die verdoppelten Stämme meist frei von Markern und die Reduplikation damit total ist. Abgesehen davon hängt die Zuordnung der Name-ICCs zur Reduplikation aber davon ab, ob die Dopplung des Stammes die Bildungen tatsächlich als Eigennamen markiert und ob man diese Eigenschaft als grammatische Bedeutung der Wiederholung begreift. In Kapitel 7 hatte sich herausgestellt, dass Name-ICCs keine Komposita sind. Die Bildung von Name-ICCs kann somit als morphologischer Prozess zur Bildung von Eigennamen angesehen werden, bei dem die entstehenden Bildungen eine reduplikative Struktur aufweisen. Die Ergebnisse dieses Kapitels fasst Tabelle 17 zusammen.

Tab. 17: Die drei ICC-Typen als Reduplikation

	Kriterium / Beschreibung	Det-ICC	Prot-ICC	Name-ICC
	Formal			
A	Die Basis und der angefügte Reduplikant sind formgleich.	+	+	+
B	Der Reduplikant wird unmittelbar vor oder nach der Basis angefügt.	○	+	+
C	Der Reduplikant wird genau einmal an die Basis angefügt.	○	+	+
	Semantisch-funktional			
D	Die Wiederholung selbst ist Träger von Bedeutung.	–	+	○
E	Die Basis und der angefügte Reduplikant sind bedeutungsgleich.	+	+	+
F	Die entstehende Bedeutung ist grammatischen, nicht pragmatischen Charakters.	–	+	○

8.5 Fazit: ICCs – Komposition oder Reduplikation?

Um die drei ICC-Typen auch terminologisch zu den morphologischen Prozessen der Komposition und Reduplikation in Bezug zu setzen, kann man, angesichts der in Tabelle 15 und 17 zusammengefassten Merkmalsverteilungen, Det-ICCs als reduplikative Komposita bezeichnen. Während sie alle Merkmale der Komposition aufweisen, rückt sie allein die Form- und Bedeutungsgleichheit der Konstituenten in die Nähe der Reduplikation. Prot-ICCs hingegen lassen sich am besten als Kompositionsreduplikationen bezeichnen. Dieser Terminus wird zum einen dem Umstand gerecht, dass Prot-ICCs sowohl einige Merkmale der Komposition, unter anderem die definitorischen, als auch (alle) Eigenschaften der Reduplikation aufweisen. Zum anderen soll der Terminus und somit die begriffliche Unterscheidung von Prot-ICCs und Det-ICCs anerkennen, dass die Bedeutungskonstitution bei Prot-ICC anders funktioniert als bei Det-ICCs oder Nominalkomposita im Allgemeinen (siehe Kapitel 9). Die Bildung von Name-ICCs ist hingegen weder für Komposition noch für Reduplikation ein sonderlich gutes Beispiel. Zwar haben Name-ICCs eine reduplikative Struktur, doch wird von der Konstituentendopplung selbst keine Bedeutung transportiert. Ein gutes Beispiel für Komposition sind Name-ICCs angesichts eines fehlenden semantischen (und teilweise auch grammatischen) Kopfes aber auch nicht. Sie sind morphologische Verbindungen von Wortstäm-

men und lassen sich am besten als (total) reduplikative Wort-, beziehungsweise Eigennamenbildungen bezeichnen.

Viele der in Kapitel 7 und 8 behandelten Aspekte betreffen die Frage, wie den Rezipient:innen die Bedeutung von ICCs angezeigt wird. Die Korpusstudien haben hierzu die Verwendung einiger formaler Mittel nachgewiesen. So wurden Det-ICCs als nahezu ausnahmslos verfugt beschrieben, während Name-ICCs, dem Prinzip der Wortkörperschonung entsprechend, so gut wie niemals Fugenelemente zeigen. Auch die Flexion und die Schreibung unterstützen die Rezipient:innen bei der Interpretation der intendierten Bedeutung. Doch sind diese formalen Mittel nicht ausreichend, um die Diskriminierung dreier ICC-Typen zu gewährleisten. Im folgenden Kapitel wird untersucht, wie die Bedeutungskonstitution und die Semantik von ICCs angemessen beschrieben werden kann. Dazu werden ICCs im Rahmen unterschiedlicher Theorien der Kompositasemantik beschrieben.

9 ICCs und Theorien der Kompositasemantik

Die Korpusstudien haben gezeigt, dass die drei ICC-Typen sich teilweise formal voneinander unterscheiden. Es zeigte sich aber auch, dass die Prot-ICCs keine klaren formalen Merkmale haben, durch die sie sich von Det- und Name-ICCs abgrenzen, da sie hinsichtlich der Verfugung, Schreibung und Flexion eine Zwischenposition zwischen Det- und Name-ICCs einnehmen. Um die Frage zu beantworten, wie die Bedeutungen der drei ICC-Typen dennoch von den Sprecher:innen unterschieden werden können, wird in diesem Kapitel die Konstitution der ICC-Semantik anhand unterschiedlicher Kompositatheorien sowie mithilfe von Modellen zur Kompositasemantik beschrieben.

Prinzipiell gibt es zwei Möglichkeiten, wie ICCs – und Komposita im Allgemeinen – Bedeutung erhalten. Die erste Möglichkeit ist, dass die Bildungen im mentalen Lexikon abgespeichert sind. ICCs wären in dem Fall also lexikalisiert, wie etwa das Det-ICC *Kindeskind* oder das Name-ICC *Baden-Baden*, deren Bedeutung als Ganzes aus dem mentalen Lexikon der Sprecher:innen abgerufen wird, sofern der Kontext nicht eine andere Bedeutung evoziert. Ein Abgleich mit gängigen Wörterbüchern des Deutschen, etwa dem Duden oder dem DWDS, legt allerdings nahe, dass nur sehr wenige ICCs lexikalisiert sind. In der Forschung zu ICCs wird darüber hinaus davon ausgegangen, dass ICCs generell nicht lexikalisierbar sind (Hohenhaus 2007: 25ff., 2015: 275, Horn 2018: 236, Kentner 2017: 243). Es ist deshalb zu diskutieren, wie hoch die Leistung des Lexikons bei der Dekodierung von ICCs einzuschätzen ist.

Für die ad hoc gebildeten ICCs, die also keinen Eintrag im mentalen Lexikon haben, gibt es nur eine Möglichkeit der Bedeutungskonstitution, nämlich die, erschließbar zu sein. Die Erschließbarkeit ist dabei an folgende Sachverhalte geknüpft:

- Die Bestandteile von ICCs sind den Sprecher:innen bekannt.
- Die Relation zwischen den beiden Konstituenten wird von existierenden N+N-Komposita auf ICCs übertragen.
- Die drei ICC-Typen stellen für Sprecher:innen bekannte und transparente Muster zur Wortbildung dar. In Analogie zu bereits bekannten ICCs desselben Typs kann entsprechenden Neubildungen eine Bedeutung zugeordnet werden.
- Die Interpretation von ICCs und das Inferieren der Modifikationsrelation ist wie bei allen okkasionellen N+N-Komposita stark abhängig vom Kontext (Peschel 2002: 299f.) und Weltwissen der Sprachteilnehmer:innen (Booij 2009: 203).

https://doi.org/10.1515/9783111315416-009

Viele Theorien nehmen an, dass eine Kombination dieser Sachverhalte dazu führt, dass nicht-lexikalisierte Bildungen verstanden werden. Es stellt sich die Frage, in welcher Potenz diese Faktoren bei der Interpretation von ICCs wirken und welche Unterschiede hier jeweils zwischen den ICC-Typen bestehen.

Der erste oben genannte Punkt betrifft die Bekanntheit des Basisnomens und ist die Voraussetzung dafür, dass ein ICC überhaupt verstanden werden kann. Vorausgesetzt, dass die Rezipient:innen das Basisnomen sowie die Fugenelemente und Flexionsmarker des Deutschen kennen, sind alle formalen Bestandteile der ICCs bekannt und das korrekte Erschließen der Bildung ist lediglich von den anderen drei der vier oben genannten Aspekte abhängig. Wie im zweiten Punkt beschrieben, geschieht die Verbindung der zwei Konstituenten womöglich in Anlehnung an kanonische N+N-Komposita; eine bekannte semantische Relation könnte dann zur Dekodierung des neu gebildeten Kompositums angewendet werden. Die Relation könnte aber auch aus bekannten ICCs übertragen werden. Voraussetzung hierfür ist, dass es lexikalisierte Bildungen des entsprechenden ICC-Typs oder ein abstraktes Muster gibt, sodass sich die Neubildungen motivieren lassen. Der letzte Punkt verweist auf die generelle Kontextabhängigkeit von Wörtern, die auf Ad hoc-Bildungen besonders zutrifft.

Um abzuwägen, welche Rolle jeweils dem Lexikon, den Mustern, den Relationen und dem Kontext bei der Dekodierung von ICCs zukommt, werden in diesem Kapitel unterschiedliche Ansätze zur Bedeutungskonstitution von Komposita beschrieben. Zunächst gehe ich auf die zentralen Begriffe „Lexikalisierung“ (9.1), „Transparenz“ und „Kompositionalität“ ein (9.2), da diese für alle Modelle und für die Erschließbarkeit von Komposita im Allgemeinen unerlässlich sind. Im Zuge dessen werden auch ICCs zu diesen Begriffen in Bezug gesetzt. Im Folgenden geht es dann um die verschiedenen theoretischen Ansätze, Kompositasemantik zu analysieren (9.3). Daraufhin stelle ich Modelle vor, die mithilfe unterschiedlicher Methoden versuchen, die Erschließbarkeit von Komposita zu modellieren und versuche sie auf ICCs anzuwenden (9.4). Zum Abschluss des Kapitels ziehe ich ein Zwischenfazit zur Bedeutungskonstitution von ICCs (9.5).

9.1 ICCs und Lexikalisierung

Der Terminus „Lexikalisierung“ wird in der Linguistik sehr uneinheitlich verwendet. Bisweilen wird er für synchrone wie diachrone Prozesse, formbezogene wie bedeutungsbezogene, kognitive wie kollektive gleichermaßen verwendet. Aus diesem Grund möchte ich die Begriffsdefinition an sich behandeln, bevor ich die Lexikalisierung von ICCs beschreibe.

Wird der Begriff Lexikalisierung synchron verwendet, meint er „die Aufnahme [eines Ausdrucks] in den Wortbestand der Sprache“ (Bußmann 2008: 404), beziehungsweise den „Übergang ins Lexikon, ins Inventar“ (Diewald 1997: 73) also einen „Speicherungsvorgang“ (Glück 2010: 398). Aus kognitiver Sicht ist hier der Wortbestand einer einzelnen Person (Mithun 2010: 53), häufig aber auch der Wortbestand einer Sprecher:innengruppe gemeint (Gaeta & Ricca 2009: 38). Es wird nicht immer zwischen diesen beiden sehr unterschiedlichen Prozessen unterschieden und „Lexikalisierung“ für die Aufnahme eines Ausdrucks ins Lexikon verwendet, ungeachtet dessen, ob es das mentale Lexikon einer einzelnen Person oder das einer Sprachgemeinschaft ist (Montermini 2010: 83). Ich verwende in dieser Arbeit den Terminus „Memorierung“, wenn ein Ausdruck in den Wortbestand eines Menschen aufgenommen und in seinem mentalen Lexikon als Einheit repräsentiert wird. Wird ein Ausdruck in den Wortbestand einer ganzen Sprecher:innengruppe aufgenommen, also in den allgemeinen Wortschatz, spreche ich von „Lexikalisierung“ (Fiedler 2007: 21 Schlechtweg 2018: 36, Wunderlich 1986: 231). Lexikalisierung ist also kollektive Memorierung (Fiedler 2007: 21).

Auf den ersten Blick gehen die Begriffe fließend ineinander über. Je mehr Individuen dem Kollektiv angehören, desto eher ist der Memorierungsprozess von Neubildungen gleichzeitig eine Lexikalisierung. Doch muss man bei der Differenzierung der zwei Begriffe die unterschiedlichen Perspektiven berücksichtigen. Memorierung ist ein kognitiver Prozess im Gehirn eines einzelnen Menschen. Lexikalisierung als kollektive Memorierung ist hingegen ein soziales Phänomen, das Interaktion, Diskurs und viele weitere Prozesse einbezieht, die sich grundlegend von der kognitiven Dimension, bei der lexikaler Zugriff, Verankerung in semantischen Netzwerken etc. eine Rolle spielen, unterscheidet. Deshalb betrifft Memorierung immer nur die Speicherung von Einträgen im Gehirn genau eines Menschen. Der Begriff Lexikalisierung hingegen kann nur graduell verstanden werden, da einige Wörter mehr oder weniger allen Sprecher:innen einer Sprache bekannt sind (*Deutschland*, *Banane*), manche hingegen nur bestimmten Gruppen (*Eucharistie* ‘die katholische Form des Abendmahls’, *Flansch* ‘scheibenförmiges Verbindungsteil zur Abdichtung von Maschinenteilen’). Hier kann man also beliebig viele Abstufungen des Lexikalisierungsgrades annehmen. Für eine grobe Einschätzung dazu, ob ein Wort lexikalisiert ist, kann auf Wörterbücher zurückgegriffen werden, beispielsweise, wie in der vorliegenden Arbeit, auf den Duden.

Häufig wird der Terminus Lexikalisierung anders aufgefasst und meint zusätzlich zum bloßen Übergang eines Wortes in den Wortbestand noch den „Vorgang und [das] Ergebnis der Demotivierung“ (Bußmann 2008: 404, Diewald 1997: 72) und der „Bedeutungsveränderung“ (Glück 2010: 398). Wörter durchlaufen demnach notwendigerweise während der kollektiven Memorierung den Prozess

semantischer Spezifizierung und der Idiomatisierung. Dabei wandelt sich sprachliches Material über einen gewissen Zeitraum graphematisch, phonologisch, morphologisch oder semantisch (Hohenhaus 2005: 353ff., Sauer 2004: 1628ff.). Mit dem Wandel geht ein Verlust an Motiviertheit einher, da das Wortbildungsmuster, das dem Ausdruck zugrunde liegt, schwieriger erkannt wird. Das Ergebnis ist eine Form aus teilweise oder vollständig nicht mehr analysierbaren Morphemsegmenten (Hopper & Traugott 2003: 133). Nach dieser Auffassung entsteht durch Lexikalisierung notwendigerweise Idiosynkrasie (Bauer 2000: 834). Dafür kann auch ein unproduktiv gewordenes Muster verantwortlich sein:

> [A] word which is institutionalized could in principle still be formed by synchronic rules or word-formation, while one that is lexicalized is one which could no longer be generated by the normal rules of word-formation [...]. This can come about either because a given word-formation process ceases to be productive [...], or because of subsequent linguistic change affecting a compound or derivative and causing it to become opaque. Where the change is a semantic one, some scholars prefer to use the term idiomatization.
>
> (Bauer 2000: 837)

Häufig wird darüber hinaus angenommen, dass lexikalisierte Komposita nicht kompositional dekodiert werden können (Costello & Keane 2005, Lehmann 2007, Plank 1981, Teubert 1998).

Ich wende eine andere Auffassung von Lexikalisierung an und folge dabei den Ansätzen von Gaeta und Ricca (2009) sowie Libben (2006). Demnach wird ein neues Wortbildungsprodukt nicht automatisch opak dadurch, dass es von den Mitgliedern einer Sprachgemeinschaft memoriert wird. Stattdessen ist es andersherum, dass nämlich ein Ausdruck, der nicht mehr transparent ist, in der Gruppe lexikalisiert und von den Sprachteilnehmer:innen memoriert worden sein muss, um verstanden zu werden (Gaeta & Ricca 2009: 39). Libben (2006: 6) sieht als Grund für die Lexikalisierung eines Kompositums in erster Linie die Häufigkeit des Gebrauchs und weniger den Grad an Motivation an. Geht ein Wort im Zuge des häufigen Gebrauchs in den Wortschatz über, ist der Prozess, durch den es entstanden ist, für die Dekodierung seiner Bedeutung nicht mehr wichtig. Das Wort kann sich deshalb in der Folge ändern und von der Wortbildungsbedeutung entfernen. Das tut es aber nicht notwendigerweise, es kann genauso gut seine „Motivationsbedeutung“ behalten (Barz et al. 2003: 183f.) und weiterhin aus seinen Bestandteilen und dem Wortbildungsprozess abgeleitet werden. Die Motivationsbedeutung (auch „Wortbildungsbedeutung“ oder „semantisches Muster“, Motsch 2004: 27) ergibt sich aus der lexikalischen Bedeutung der beiden unmittelbaren Konstituenten und deren Beziehung zueinander. Trotz dieser Ableitbarkeit

kann ein Wort aber von den Sprecher:innen memoriert und von einer Sprecher:innengruppe lexikalisiert werden.[68]

Wenn sich die Wortbedeutung von der Motivationsbedeutung entfernt, verwende ich für diesen Prozess den Terminus „Idiomatisierung". Durchläuft ein Wort diesen Prozess, kann es nicht mehr vollständig aus seinen Bestandteilen erschlossen werden. Es wird opak und bedarf zu seinem Verständnis eines Eintrags im mentalen Lexikon. Das von Bauer (2000) angesprochene Szenario, dass ein Wortbildungsmuster unproduktiv und ein Wort deshalb opak wird, subsumiere ich unter diesen Terminus. Je nachdem, wie weit die Idiomatisierung fortgeschritten ist, ist die zugrundeliegende Zusammenfügung noch erkennbar.

Es stellt sich nun die Frage, ob ICCs deshalb verstanden werden können, weil sie lexikalisiert sind. Als Maß für den Lexikalisierungsgrad wird auch hier der Einfachheit halber das Vorhandensein eines Artikels im Duden angesehen (https://www.duden.de/). Lexikalisiert sind demnach vor allem Det-ICCs. Doch auch ihre Anzahl ist überschaubar. Einzig *Helfershelfer*, *Kindeskind*, *Zinseszins* und *Kompetenzkompetenz* sind im Duden verzeichnet. Es mag weitere Det-ICCs geben, die nur innerhalb bestimmter Sprecher:innengruppen memoriert sind. So ist etwa *Erwartungserwartung* in der soziologischen Systemtheorie die Erwartung, die sich auf die Erwartungen des Gegenübers bezieht. Doch können diese Bildungen nicht als allgemein lexikalisiert gelten. Neben diesen Det-ICCs gibt es zwei lexikalisierte Name-ICCs, nämlich *Baden-Baden* und *TomTom*. Für alle anderen ICCs gibt es im Duden keine Artikel.

Von den Prot-ICCs ist kein einziges lexikalisiert. Ihnen wird sogar die Fähigkeit zur Lexikalisierung abgesprochen:

> [N]one of the specimens found are even anywhere near being established lexical items. […] ICCs are […] bound to highly specific situational conditions, [and it is] this 'type-of-situation-dependency' which accounts for them not becoming lexicalized.
>
> (Hohenhaus 2004: 314f.)

68 Häufig führt die Verquickung von Lexikalisierung und Idiomatisierung zu bemerkenswerten Einschätzungen zum lexikalischen Status von Wörtern. Poethe (2014: 32) ist etwa der Ansicht, der von einem Filmtitel stammende Ausdruck *Keinohrhase* sei nicht lexikalisiert, da er ein Ausdruck „nichtalltäglicher, nichttypischer, ungewöhnlicher, von Erwartungen abweichender Erscheinungen" sei. Dieser Ansicht ist im Hinblick auf die allgemeine Bekanntheit des Wortes vehement zu widersprechen. Mittlerweile sind sogar Neubildungen in Analogie zu *Keinohrhase* entstanden, etwa *Grünohrhase* als Name für eine Süßigkeit. Es hängt also offensichtlich nicht ausschließlich von der Originalität eines Ausdrucks ab, ob er innerhalb einer Sprachgemeinschaft einen hohen Bekanntheitsgrad erreicht, also kollektiv memoriert wird.

Prot-ICCs werden von Hohenhaus (2004) also mit Augenblicksbildungen wie *apple-juice seat* 'Sitzplatz mit einem Apfelsaft auf dem Tisch davor' (Downing 1977) oder *Reus-Freistoß* aus (14) gleichgesetzt, deren Verwendung auf einen ganz spezifischen Kontext beschränkt ist. Prot-ICCs seien deshalb ohne einen vorbereitenden kontrastiven Kontext nicht verwendbar (Ghomeshi et al. 2004, Hohenhaus 2004: 301) und darum auch nicht lexikalisierbar (Hohenhaus 2007: 25ff., 2015: 275, Horn 2018: 236, Kentner 2017: 234). Ob Prot-ICCs wie *Film-Film*, *Ende-Ende* oder *Mädchenmädchen* angesichts der vielen Belege im DECOW16 und deTenTen13 nicht doch lexikalisiert werden können, es womöglich schon sind, soll an dieser Stelle nicht beurteilt werden. Im Duden sind diese Wörter jedenfalls nicht verzeichnet. Man kann daher davon ausgehen, dass diese Prot-ICCs in der Regel kontextuell angereichert werden (Finkbeiner 2014: 193, Rossi 2011: 164f.), der Kontext also wie bei anderen Ad hoc-Bildungen viele Informationen zu ihrer Interpretation liefert (Carston 2002, Levinson 2000, Récanati 2004).

Die lexikalisierten Name-ICCs *Baden-Baden* und *TomTom* unterscheidet von den anderen ICC-Typen, dass mit ihnen keine lexikalische Bedeutung memoriert werden muss, sondern lediglich der Referent, auf den sie sich beziehen. Kennt man die Stadt Baden-Baden und weiß, dass sie *Baden-Baden* heißt, hat man bereits den gesamten Gehalt des Name-ICCs erfasst. Doch trotz (oder sogar wegen) dieser fehlenden Semantik sind Name-ICCs womöglich schwieriger zu memorieren als andere ICCs. Da es zur Lexikalisierung und Memorierung von Name-ICCs keine Studien gibt, muss man auf die Erkenntnisse aus der allgemeinen Eigennamenforschung zurückgreifen und sie auf Name-ICCs zu übertragen versuchen.

In der Eigennamenforschung ist unumstritten, dass der mentale Zugriff auf Eigennamen mehr Ressourcen in Anspruch nimmt als der auf Appellativa (Semenza 2006). Eigennamen sind daher im Allgemeinen „eine kognitive Belastung […]" (Nübling et al. 2015: 25). Aus Sicht der Psychologie steht fest, dass

> Namen beim Lernen und Erinnern spezifische Probleme bereiten. Personennamen werden offensichtlich nicht wie andere Wörter oder Begriffe gedächtnismäßig repräsentiert und organisiert. Verweist ein Wort wie *Bäcker* auf eine Berufsbezeichnung, kann es in ein lexikalisches und begriffliches Netzwerk eingeordnet werden, das Relationen zu anderen Konzepten und Lexikoneintragungen kenntlich macht. Verweist der Name auf eine Person, scheint eine andere Repräsentation angesprochen zu werden.
>
> (Wippich 1995: 492)

Auch neuere Studien weisen in die Richtung, dass Eigennamen auf ganz andere Weise abgespeichert und abgerufen werden als Appellativa. So werden Eigennamen sehr viel langsamer verarbeitet als Appellativa (Evrard 2002); der Prozess des lexikalen Zugriffs findet zudem in anderen Arealen des Gehirns statt (Semenza 2006).

Unter anderem deshalb wird für Eigennamen ein eigenes Modul, das Onomastikon, angenommen, sodass Eigennamen nicht per se zum Lexikon gehören, sondern einen separaten Teil des Lexikons ausmachen (Debus 1977: 169, Nübling et al. 2015: 66). Erst wenn sie deonymisiert (*zwei Rembrandts*, *der Mozart des Schachs*), also zu Appellativa werden (Nübling et al. 2015: 61f.), gehören sie wie andere Appellativa zum Lexikon. Name-ICCs werden also grundsätzlich nicht lexikalisiert, sondern sind Teil des sprecher:innenspezifischen Eigennameninventars. Das bedeutet auch, dass sich Name-ICCs wie *Baden-Baden* oder *TomTom* im Wortinventar vieler Sprecher:innen festsetzen können, und zwar im Onomastikon.

Zusammenfassend ist festzuhalten, dass ICCs nur sehr vereinzelt lexikalisiert sind. Während es wenige lexikalisierte Det- und Name-ICCs gibt, sind Prot-ICCs nie lexikalisiert und womöglich auch gar nicht lexikalisierbar. In Bezug auf Name-ICCs kann man vermuten, dass die Bildungen nicht im mentalen Lexikon, sondern im Onomastikon abgespeichert sind und dies – wie bei Eigenammen im Allgemeinen – von Sprecher:in zu Sprecher:in sehr unterschiedlich ist. Bei allen drei ICC-Typen sind es also wohl nicht die Lexikoneinträge selbst, die bei der Konstitution der Semantik wirken. Stattdessen muss die Bedeutung der ICCs auf die ein oder andere Weise erschlossen werden. Hierzu sind die Begriffe der semantischen Transparenz und Kompositionalität wichtig, denn „[n]eue sprachliche […] Zeichen benötigen irgendeine Art von Transparenz, um […] interpretierbar zu sein“ (Keller 1995: 160).

9.2 ICCs und semantische Transparenz

9.2.1 Transparenz und Kompositionalität

Es ist zunächst wichtig hervorzuheben, dass die beiden Begriffe „Transparenz“ und „Kompositionalität“ nicht gleichbedeutend mit Erschließ-, Vorhersag- und Interpretierbarkeit sind. Bei Transparenz und Kompositionalität geht es um die Struktur von Komposita und ihre Passgenauigkeit zwischen Wortbildungsbedeutung und lexikalischer Bedeutung. Das Erschließen und Antizipieren der Bedeutung von Neubildungen geht aber aus kognitiver Sicht weit über diese strukturellen Aspekte hinaus. Benczes (2004: 13) macht zudem darauf aufmerksam, dass Komposita, die metaphorisch oder metonymisch verschoben sind, nicht, wie häufig angenommen (etwa bei Dirven & Verspoor 1998), notwendigerweise auch schwieriger zu erschließen sind. Schließlich sind Metaphern und Metonymien normale Alltagskonstruktionen (Lakoff & Johnson 1980a, b). Auch spielen bei der

Erschließbarkeit zahlreiche außersprachliche Faktoren eine Rolle (Štekauer 2005: 440):

> Predictability will depend on other factors, too: from one side largely on context, from the other side on conventionality, the criteria of which are familiarity of the new meaning, frequency of its use, and its fixation in the lexicon.
>
> (Kavka 2009: 40f.)

Die hier besprochenen Themen umfassen also bei weitem nicht alle Aspekte der Worterkennung und Wortinterpretation. Es kann aber angenommen werden, dass die kompositionale Struktur und die semantische Transparenz von Komposita wenn auch nicht die einzigen, so aber doch sehr wichtige Faktoren ihrer Bedeutungskonstitution sind.

Wie schon der Begriff der Lexikalisierung wird auch der Terminus der semantischen Transparenz in der Linguistik nicht einheitlich verwendet. Manche Autor:innen setzen Transparenz mit Kompositionalität gleich (Bell 2012, Giegerich 2015: 31, Schulte im Walde & Borgwaldt 2015: 1200), andere mit Motiviertheit (Glück 2010: 722). Glück versteht beispielsweise unter einer transparenten Bildung (V Motiviertheit) den „Grad der Erschließbarkeit der Bedeutung eines Wortbildungsproduktes aus den Bedeutungen seiner Teile“ (Glück 2010: 444). Genau das aber versteht etwa Bußmann unter Kompositionalität (Bußmann 2008: 267).

Für semantische Transparenz gilt in dieser Arbeit die folgende Definition von Libben und Kolleg:innen, die sich in ähnlicher Form in vielen anderen einschlägigen Texten zu dem Thema findet (etwa Laudanna & Burani 1995, Schreuder & Baayen 1995):

> Semantic transparency sets boundary conditions on whether a multi-morphemic word can be comprehended in terms of its constituent morphemes.
>
> (Libben et al. 2003: 51)

Transparenz betrifft die Frage, ob „die Bedeutung des Wortstamms vollständig in der Bedeutung der Vollform enthalten ist“ (Heide 2010: 71), also das Verhältnis zwischen der Bedeutung eines Stammes als Simplex und seiner Bedeutung als Kompositumskonstituente. Unterscheidet sich die Bedeutung der Konstituenten in einem Kompositum von der, die ihnen in Isolation innewohnt, gibt es also irgendeine Art der Bedeutungsspezialisierung oder Generalisierung einer Kompositakonstituente im Vergleich mit dem entsprechenden Basislexem, ist das entsprechende Kompositum damit weniger transparent (Benczes 2004: 13, 2006: 5). Die Frage nach der Transparenz lautet also: „A = [A in AB]?“ sowie „B = [B in AB]?“.

Für Kompositionalität wird häufig mit Rückgriff auf Frege (Frege-Prinzip) folgende Definition angegeben:

> The meaning of an expression is a function of the meanings of its parts and of the way they are syntactically combined.
>
> (Partee 1984: 218)

Kompositionalität meint also den Grad, zu dem sich die Bedeutung eines Ausdrucks mithilfe seiner Konstituentenbedeutungen erschließen lässt, also die semantische Bezogenheit zwischen der Bedeutung eines Gesamtkompositums und den Bedeutungen der Konstituenten. Die Frage nach der Kompositionalität lautet also: „AB = A+B?“ Eine klare Unterscheidung zwischen den beiden Begriffen Transparenz und Kompositionalität trifft Zwitserlood (1994):

> [S]emantic transparency is not the same as compositionality. Although the semantic relation between transparent compounds and their constituents might be easy to establish, the meaning of the compound as a whole is often more than the meaning of its component words.
>
> (Zwitserlood 1994: 366)

Kompositionalität und Transparenz sind also zwei Eigenschaften eines Kompositums, die zwar miteinander in Zusammenhang stehen, aber nicht dasselbe meinen und darum auch unabhängig voneinander vorliegen können.

Komposita teilen sich jedoch nicht kategorial in transparente und intransparente, kompositionale und nicht-kompositionale auf. Sowohl Transparenz als auch Kompositionalität werden in der Forschung mittlerweile als graduelle Merkmale aufgefasst (Costello & Keane 2005: 204, Klos 2011: 70, Plank 1981: 28, Schäfer 2018: 1). Der Grad semantischer Transparenz hängt davon ab, wie stark eine Konstituente semantisch verschoben ist. Je ähnlicher sich die Bedeutung der Stämme als Simplex und als Kompositumskonstituenten sind, desto transparenter ist das entsprechende Lexem. Hier spielen etwa Metaphern oder Metonymien eine Rolle. *Pulverschnee* ist etwa transparenter als *Eischnee*, da die Konstituente *Schnee* in *Pulverschnee* tatsächlich auf kalten Niederschlag aus Eiskristallen verweist, in *Eischnee* allerdings metaphorisch verschoben ist und lediglich auf eine schneeähnliche Masse referiert, die aber nicht kalt, nicht Niederschlag und nicht aus Eiskristallen erzeugt ist. Benczes (2004) ist hier anderer Auffassung und sieht Komposita wie *red tape* im Sinne von ‘Bürokratie’ nicht als opak an:

> Such compounds are not unanalysable, nor semantically opaque: in fact, they are just as transparent as their endocentric counterparts. […] [T]he analyses of these expressions require various cognitive linguistic tools, such as metaphor, metonymy, blending, profile determinacy, schema theory and construal.
>
> (Benczes 2004: 17)

Ich hingegen begreife Komposita, die letztgenannte Mechanismen umfassen (Metaphern, Metonymien, …), als weniger transparent als solche, bei denen die Konstituenten in ihrer lexikalischen Bedeutung auftreten, weil dies grundlegend für die gemeinhin verwendete Transparenzdefinition ist.

Der Grad der Kompositionalität hängt vor allem davon ab, wie groß das Mehr an Bedeutung ist, das der Wortbildungsprozess erzeugt (Barz 2009: 673, Schlaefer 2002: 32), also „die zusätzliche semantische Dimension […], die nicht auf die einzelnen Konstituenten zurückzuführen ist“ (Klos 2011: 70). Gibt es kein semantisches Mehr, das der Wortbildungsprozess beisteuert, gilt für Komposita „daß ihre Bedeutung aus der Bedeutung der Kompositionsglieder und nur aus ihr berechnet werden kann“ (Kanngießer 1987: 15). Auch Klos (2011) geht davon aus, dass „sich die Semantik der Zusammensetzung tatsächlich rein kompositional erschließen“ lässt, wenn „zwischen den unmittelbaren Konstituenten eine rein additive *und*-Relation“ besteht (Klos 2011: 219, Whitney 1889: §1247). In der Wortbildung kommen etwa die Dvandva im Chinesischen (126) oder im Sanskrit (127) dieser reinen Addition am nächsten:

(126) Mandarin Chinesisch (Zhu et al. 1995: 191)
父 母
fù *mǔ*
Vater Mutter
‘Vater und Mutter = Eltern’

(127) Sanskrit (Booij 2005: 79)
candra-ditya-u
Sonne-Mond-DUAL
‘Sonne und Mond’

In den meisten Texten zum Thema Kompositionalität wird nicht begründet, warum die Relation AND ein Kompositum kompositionaler macht als beispielsweise eine OF-Relation. Es soll an dieser Stelle der Versuch unternommen werden, diesen Unterschied zu begründen und darauf aufbauend Komposita in unterschiedliche Kompositionalitätsgrade einzuteilen.

Auf formaler Seite ist Komposition die Addition von sprachlichem Material. Ist die Gesamtbedeutung einer Bildung ebenfalls bloß die Summe der Konstituentenbedeutungen, ist die Bildung von Bedeutungsaddition gekennzeichnet. Der Addition von Form steht eine Addition von Bedeutung gegenüber, anders als bei einer OF-Relation. Bei Komposita mit einer AND-Relation findet auf beiden Seiten des sprachlichen Zeichens keine Vereinigung oder Verschmelzung der Bestandtei-

le statt. Die Konturierung der formalen wie auch der semantischen Bestandteile bleibt erhalten. In den Beispielen (126–127) referieren jeweils zwei formale Bestandteile auf zwei Entitäten in der Welt. Die Relation zwischen ihnen ist AND. Form und Bedeutung des Kompositums beinhalten nichts weiter als die beiden (Bedeutungen der) Konstituenten. Zusammenfügungen wie in (126–127) weisen also den höchsten Grad an Kompositionalität auf.

In der Wortbildung des Deutschen gibt es diesen Grad an Bedeutungsaddition nicht. Generell gibt es hinsichtlich solcher Dvandva eine große areale Asymmetrie, insofern diese Bildungen sprachtypologisch gesehen zwar nicht selten sind, sie aber in den europäischen Sprachen nicht vorkommen (Bauer 2001a). Bisweilen wird sogar angenommen, zusammengesetzte Strukturen wie N+N-Komposita seien generell nicht konsistent oder vollständig vorhersagbar.

> [Noun–noun compounds] do in fact manifest conventional patterns of composition: the relation a composite structure bears to its components is neither random nor arbitrary. But at the same time, a composite structure is not constructed out of its components, nor is it consistently or fully predictable. Rather than constituting a composite structure, the component structures correspond to certain facets of it, offering some degree of motivation for expressing the composite conception in the manner chosen.
>
> (Langacker 2000: 16)

Im Deutschen entsprechen Bildungen wie denen in (126–127) mit *und* koordinierte Phrasen, also syntaktische Verbindungen. In der Wortbildung des Deutschen kommen Kopulativkomposita wie *Bettsofa* in (27) der Bedeutungsaddition am nächsten. Hier liegt ebenfalls eine Addition von Konzepten und die Relation AND vor. Im Unterschied zu (126–127) sind die Konstituenten von *Bettsofa* aber referenzidentisch. Die Bedeutung des Kompositums ist eine Vereinigung der Eigenschaften beider Konstituenten (Zinsmeister 2013: 306). Auch wenn die Bedeutung solcher Kopulativkomposita in der Literatur zuweilen mit „AB = A und B“ paraphrasiert wird (etwa bei Ortner & Ortner 1984: 53), bilden die Konstituenten doch lediglich eine Gesamtbedeutung, die wiederum auf einen einzigen Referenten referiert (‘ein Möbelstück, das gleichzeitig Bett und Sofa ist’).

Die deutschen Kopulativkomposita haben somit keine pluralische Bedeutung wie bei echter Addition, obwohl das Gesamtkompositum auf der Formseite aus zwei Konstituenten besteht. Die addierten Konzepte sind also anders als bei *fù mǔ* ‘Eltern’ in (126) bereits verschmolzen. Lediglich die Determination wurde neutralisiert (Neuß 1981: 46ff.). Paraphrasen mit *und* bilden diesen Unterschied nicht ab. Wie Marchand schon in den Sechzigerjahren (in Bezug auf das Englische) schreibt, bieten sich bei der Analyse von Kopulativkomposita stattdessen Relativsätze an:

> Additive Compounds should not be analysed as 'A + B', but as 'B which is also A'.
> (Marchand 1969: 41)

Der Wortbildungsprozess fügt also den Aspekt der Referenzidentität hinzu, der in den Beispielen (126–127) nicht vorliegt. Kopulativkomposita im Deutschen sind also mitnichten kopulativ und keine Dvandva.[69] Stattdessen weisen die Kopulativkomposita des Deutschen bloß den zweithöchsten Grad an Kompositionalität auf.

Bei Determinativkomposita wie *Weinglas* beruht das Verhältnis der Konstituentenbedeutungen zur Gesamtbedeutung nicht auf Addition. *Weinglas* referiert nicht auf einen Wein und ein Glas und auch nicht auf ein Objekt, das gleichzeitig Wein und Glas ist. Gunkel und Zifonun (2009) gehen bei solchen determinativen N+N-Komposita gar von minimaler Kompositionalität aus („minimal compositionality", Gunkel & Zifonun 2009: 216). Man muss allerdings für Determinativkomposita einen gewissen Grad an Kompositionalität annehmen, da beide Konstituentenbedeutungen in die Gesamtbedeutung einfließen. Das Zweitglied ist dabei der semantische Kopf des Kompositums und für die Referenz bedeutsamer als das Erstglied. Das Erstglied ist Modifikator des Zweitgliedes und kann unterschiedliche Aspekte des Zweitgliedes genauer beschreiben und so Subkonzepte bilden. In *Weinglas* etwa beschreibt das Erstglied näher, wofür die Entität, auf die das Zweitglied referiert, verwendet wird (FOR). Bei determinativen Bildungen lautet die Relation also nicht AND, da die zwei Bestandteile nicht auf zwei Entitäten referieren, sondern zu einer Subkonzeptbildung führen, deren Ergebnis dann wiederum referiert.

Bei dieser Subkonzeptbildung haben die Konstituenten anders als bei Kopulativkomposita wie *Bettsofa* zusätzlich unterschiedliche Funktionen, nämlich Determinans und Determinatum. Zusätzlich zur Referenzidentität fügt die Komposition hier also eine Hierarchie und eine Funktionszuschreibung innerhalb des Wortes hinzu. Ich sehe dies als ein Mehr an Komplexität an, das den Grad an

---

69 Einzige Ausnahme im Deutschen ist womöglich *Strichpunkt,* vorausgesetzt, dass das Wort nicht auf ein grafisches Zeichen verweist, das aus einem Strich und einem Punkt besteht, sondern auf zwei Zeichen, nämlich einen Strich und einen Punkt. Im Gegensatz zu *Bettsofa* sind bei einem Strichpunkt beide Bestandteile konturiert und voneinander abgegrenzt. Wüster etwa unterscheidet Wörter wie *Bettsofa* („Anpaarung") von solchen wie *Strichpunkt* („Abpaarung", Wüster 1979: 51f.) terminologisch. Breindl und Thurmair (1992) zeigen allerdings experimentell, dass die Sprecher:innen des Deutschen auch solche Komposita determinativ interpretieren. Wörter wie *Strichpunkt* sind demnach also eher mit der Präposition *mit* zu paraphrasieren ('Punkt mit Strich'). Kopulativkomposita sind nach Breindl und Thurmair (1992) deshalb generell nicht existent im Deutschen (ebenso: Klos 2011: 271).

Kompositionalität des Wortes reduziert.[70] Determinativkomposita weisen also den dritthöchsten Grad an Kompositionalität auf, den morphologische Zusammensetzungen von Nomina haben können.

Gunkel und Zifonun (2009) erkennen in der reduzierten Kompositionalität von Determinativkomposita auch eine Funktion: Auf diese Weise werde der Benennungsfunktion der Bildungen Rechnung getragen:

> [I]t can be argued that reduced compositionality [...] even has a functional point. The objective is to define a certain subkind, which can be achieved in an optimal way when it is identified with as little descriptive effort as possible. This is because the less descriptive detail the identification of the subkind includes the less it resembles a general description and the less it resembles a general description the more name-like it is.
>
> (Gunkel & Zifonun 2009: 216)

Minimal kompositional sind schließlich Wörter wie *Spaghettiwestern*. Das Mehr an Bedeutung umfasst hier sehr viele Informationen, da das Muster der Zusammensetzung komplex ist: 'Y aus einem Land, das durch die Menge an X, die dort konsumiert wird, näher bestimmt ist'. Nach meiner Klassifikation weisen solche Komposita den niedrigsten Grad an Kompositionalität aller Komposita im Deutschen auf.

Die besprochenen Beispiele zu den Kompositionalitätsgraden verdeutlichen auch, dass Transparenz und Kompositionalität nicht immer zusammenfallen. Sandra (1990) zeigt das am englischen Beispiel *blackbird*. Die Bildung ist transparent, da *black* mit seiner lexikalischen Bedeutung einen Beitrag zur Gesamtbedeutung leiste. *Blackbird* ist aber nicht kompositional, weil die Gesamtbedeutung nicht aus den Bedeutungen der Bestandteile hergeleitet werden kann, da *blackbird* nicht einfach nur einen schwarzen Vogel, sondern ganz konkret eine Amsel bezeichnet (Sandra 1990: 550). Ebenso ist es bei den Beispielen *Weinglas* und *Spaghettiwestern.* Die Bestandteile sind zwar semantisch transparent. *Wein* und *Glas* in *Weinglas* entsprechen von der Bedeutung her genau den Simplizia *Wein* und *Glas.* Auch die Konstituente *Spaghetti* in *Spaghettiwestern* realisiert genau die Bedeutungskomponente, die das Lexem *Spaghetti* üblicherweise auch in Isolation

70 Man kann hier argumentieren, dass die Betonung auf der ersten Konstituente die hierarchische Struktur innerhalb des Kompositums reflektiert. Dieser Umstand ist aber eher ein Zeichen für die Ikonizität von determinativen N+N-Komposita. Die Kompositionalität ist in determinativen N+N-Komposita im Vergleich mit kopulativen Komposita oder Dvandva dennoch niedriger, weil die Synthese und Analyse von Elementen (Komposition und Dekomposition) der Funktionszuschreibung dieser Elemente zeitlich und logisch vorangeht. Wird in Komposita über die Addition hinausgehend eine Funktionszuschreibung der Bestandteile vorgenommen, ist das also immer ein zusätzlicher Vorgang.

realisiert, nämlich ‘lange, dünne Nudeln aus Hartweizengrieß’. Auch *Western* referiert genau auf das, was im Deutschen unter einem *Western* verstanden wird, nämlich ‘Film, dessen Handlung sich im Wilden Westen abspielt’. Die Beziehung zwischen den Konstituenten und dem Kompositum ist jedoch sehr komplex. Andersherum weisen Bildungen wie *Schweinehund* oder *feuchtfröhlich* als Kopulativkomposita zwar einen hohen Grad an Kompositionalität auf, bei *feuchtfröhlich* hat aber eine der Konstituenten (*feucht*) eine andere Bedeutung als das entsprechende Basislexem in Isolation, da es sich nicht auf die Eigenschaft ‘mit ein wenig Wasser benetzt, etwas nass’ bezieht, sondern auf die Eigenschaft ‘ausgelassen durch Alkoholgenuss’.[71] Bei *Schweinehund* sind sogar beide Konstituenten verschoben. Das Wort bezieht sich weder auf einen Hund, noch auf ein Schwein, sondern auf einen ‘unflätigen, minderwertigen Mann’.[72] Beide Ausdrücke sind also weniger wegen der semantischen Relation zwischen den Konstituenten schwer erschließbar, sondern aufgrund von semantischen Verschiebungen.

Fügt der Prozess der Komposition viel hinzu und wird zusätzlich die Bedeutung der Konstituenten semantisch verschoben, sind die Komposita maximal opak und nur noch schwer zu erschließen. Um sie zu verstehen, müssen die Sprecher:innen sie memoriert haben. Generell müssen solche Bildungen lexikalisiert sein, um in der Sprachgemeinschaft zu funktionieren. Bei *Buchstabe* ist es beispielsweise nur noch schwer möglich, die Wortbildungsbedeutung (aus dem Germanischen **bōks* ‘Buchstabe’ und **stabi-* ‘Stab’ / ‘Holzstäbchen mit eingeritzen Runenzeichen’[73]) ohne eine Recherche etwa in einem etymologischen Wörterbuch zu verstehen. Minimal transparent sind Wörter, die trotz ihrer Wortbildungsvergangenheit dem Gros der Sprecher als Simplizia erscheinen und deren analytische Vergangenheit dem heutigen Lexem nicht mehr anzusehen ist, etwa bei *Messer*, das im Althochdeutschen noch recht transparent / kompositional in *maz-* ‘Speise, Essen’ und *sahs* ‘kurzes Schwert’ zerlegt werden konnte (Pfeifer et al. 1993).[74]

Transparenz ist also ein skalarer Begriff (Schäfer 2018: 1). Ein graduelles Verständnis von Transparenz trägt auch der Ursprungsbedeutung des Begriffs im Sinne von lat. *transparens* ‘durchscheinend’ Rechnung. Peschel nennt Wörter durchsichtig, „deren inneren Bau man noch erkennen kann“ (Peschel 2002: 6), bei

71 „feuchtfröhlich“ in Pfeifer et al. (1993), abgerufen über https://www.dwds.de/wb/feuchtfröhlich, am 11.09.2021.

72 „Schweinehund“ in Pfeifer et al. (1993), abgerufen über https://www.dwds.de/wb/Schweinehund, am 11.09.2021.

73 „Buchstabe“ in Pfeifer et al. (1993), abgerufen über https://www.dwds.de/wb/Buchstabe, am 09.10.2018.

74 „Messer“ in Pfeifer et al. (1993), abgerufen über https://www.dwds.de/wb/Messer, am 10.10.2021.

denen man also erkennt, was drinsteckt. Je transparenter ein Kompositum, desto eher erkennt man, welche Lexeme sich hinter den Kompositakonstituenten verbergen. Ebenso muss Kompositionalität skalar verstanden werden. Über den Kompositionalitätsgrad einer Wortbildung entscheidet, wie komplex das Muster der Zusammensetzung ist. Je simpler das Muster der Zusammensetzung, desto einfacher lässt sich die Zusammensetzung wieder in seine Bestandteile zerlegen. Maximal kompositional sind Bildungen des additiven Musters, da das Muster den Konstituenten keine semantisch-konzeptuelle Struktur, etwa in Form von semantischen Relationen, hinzufügt. (Die Relation AND entspricht bloß der Addition der formalen Bestandteile). Kopulative und determinative Muster sind im Vergleich mit dem additiven Muster komplexer.[75] Einen Hinweis auf den Kompositionalitätgrad eines Kompositums kann die Länge seiner Paraphrase geben: Kann man Determinativkomposita wie *Weinglas* mithilfe bloßer Präpositionen wie *für* umschreiben, braucht *Spaghettiwestern* eine sehr viel umfangreiche Umschreibung. *Schneebesen* hingegen kann ebenso einfach umschrieben werden wie *Weinglas* ('Besen für Schnee'). Dass *Schneebesen* mit dieser Paraphrase noch nicht vollends erschlossen ist, geht auf die semantische Verschiebung der Konstituenten zurück.

Wenn geringe Kompositionalität und geringe Transparenz zusammenkommen, müssen die Bildungen lexikalisiert sein oder, bei neu gebildeten Komposita, umfassend erklärt werden. Die beiden zweitplatzierten Jugendwörter der Jahre 2008 und 2010 sind anschauliche Beispiele für diese Kombination aus minimaler Transparenz und minimaler Kompositionalität. Bei den beiden Wörtern handelt es sich um *Bildschirmbräune* ('Blässe, die daraus resultiert, dass man viel Zeit vor dem Bildschirm verbringt') und *Arschfax* ('Unterhosenetikett, das hinten aus der Hose herausragt'). Beides sind Komposita mit einer komplexen Relation zwischen Erst- und Zweitglied, bei denen zusätzlich eine Konstituente (*Bildschirm<u>bräune</u>*) oder sogar beide (*<u>Arschfax</u>*) verschoben sind. Bei vollständig opaken, ehemals komplexen Wörtern wie *Messer* ist die Paraphrase nur mit Rückgriff auf die Sprachgeschichtsschreibung möglich ('Schwert für Speisen').

Da beide Begriffe, Transparenz und Kompositionalität, skalar verstanden werden müssen und zudem unabhängig voneinander vorliegen können, kann man kann eine graduelle Transparenz- und Kompositionalitätsmatrix annehmen. Solche Ansätze gibt es bereits. Dirven und Verspoor (1998) sehen beispielsweise Komposita, die „unequivocally analysable" sind, als maximal transparent an, partiell transparente hingegen sind „still analysable but the semantic link is less apparent to see which subcategory the meaning of the compound involves." Am

75 Häufig werden diese Komplexitätsgrade schlicht vorausgesetzt (etwa in Nakov & Hearst 2013: 18).

Ende des Kontinuums stehen schließlich „darkened compounds“, bei denen metaphorische / metonymische Prozesse involviert sind (Dirven & Verspoor 1998: 60). An den Zitaten wird aber deutlich, dass bei Dirven und Verspoor die Begriffgrenzen für Kompositionalität und Transparenz anders verlaufen, da die Relation zwischen den Konstituenten unter dem Begriff „Transparenz“ behandelt wird.

Tabelle 18 stellt die Begriffe wie in diesem Kapitel definiert dar. Bildungen, bei denen die kombinierten Konzepte nach einem komplexen Muster zusammengefügt wurden, stehen eher am nicht-kompositionalen Pol, die der einfacheren Bildungsmuster am kompositionalen Pol. In einer weiteren Dimension stehen Bildungen, bei denen die Konstituenten der morpholexikalischen Bedeutung weitgehend entsprechen, am Transparenzpol. Haben sich hingegen Erst- oder Zweitkonstituente von ihrer morpholexikalischen Bedeutung entfernt, rücken die entsprechenden Komposita weiter Richtung Opakheitspol. Sind beide Konstituenten verschoben, sind die Bildungen völlig opak. Diese Einteilung entspricht der von Libben und Kolleg:innen (2003). Ergänzend (oder alternativ) könnte man die Potenz der semantischen Verschiebung berücksichtigen und eine noch feingliedrigere Transparenzeinteilung vornehmen. Hier ist es allerdings sowohl theoretisch als auch methodisch schwierig, ‘Verschiebungsgrade’ zu definieren. Das Transparenzkontinuum der Matrix in Tabelle 18 berücksichtigt deshalb nur, ob Erst- und/oder Zweitglied verschoben sind und nimmt also drei Transparenzgrade an. Auch das Kompositionalitätskontinuum kann man feiner einteilen. Im Rahmen dieser Arbeit reicht allerdings die gewählte basale Einteilung aus.

Einige Zellen in der Tabelle können nicht besetzt werden. So erscheint es unmöglich, dass vollkommen kompositionale Zusammenfügungen des additiven Musters dennoch opak sind. Auf der anderen Seite kann eine Bildung nicht transparent sein, wenn synchron kein Wortbildungsmuster erkennbar ist.

Tab. 18: Zweidimensionale Transparenz- und Kompositionalitätsmatrix

		Kompositional----------Nicht-kompositional				
		Additiv	Kopulativ	Determinativ	Determinativ (komplex)	Synchron kein Muster
transparent	N_1&N_2= $Basis_{1/2}$	父母 ‘Eltern’	*Bettsofa*	*Weinglas*	*Spaghettiwestern*	–
⋮	N_1∨N_2≠ $Basis_{1/2}$	–	*feuchtfröhlich*	*Eischnee*	*Bildschirmbräune*	–
opak	N_1&N_2≠ $Basis_{1/2}$	–	*Schweinehund*	*Schneebesen*	*Arschfax*	*Messer*

9.2.2 Transparenz und Kompositionalität von ICCs

Wie verorten sich die ICCs nun auf der Transparenz- und Kompositionalitätsmatrix? In Kapitel 4 wurden die ICCs in drei Subtypen eingeteilt und zwar anhand der Entsprechung von formaler und konzeptueller Dopplung, beziehungsweise hinsichtlich des Verhältnisses der Konstituenten zum Basislexem. Die wesentliche Frage war hier, ob den zwei formalen Einheiten, die denselben Stamm realisieren, auch zwei konzeptuelle Einheiten gegenüberstehen, die dasselbe Konzept realisieren. Ich habe bei der Erläuterung dieser Klassifikationskriterien schon erwähnt, dass diese Klassifikation die Transparenz- Kompositionalitätsgrade der Bildungen abbildet. Mit der Transparenz- und Kompositionalitätsmatrix lassen sich ICCs nun diesbezüglich genauer beschreiben.

Wie in Beispiel (126) 父母 *fùmǔ* 'Vater+Mutter = Eltern' wären diejenigen ICCs maximal kompositional, bei denen die Addition der formalen Bestandteile exakt der Addition der Bedeutungen entspricht. Am ehesten liegt das bei Det-ICCs vor. In *Fenster-Fenster* werden zwei formal identische Nominalstämme miteinander kombiniert und auch konzeptuell liegen zwei Einheiten vor, nämlich zwei Realisierungen des Konzeptes FENSTER. Allerdings referiert das ICC nicht auf zwei Fenster, sondern nur auf eines. Zudem besteht zwischen den Konstituenten die Relation LOC. Die zwei Entitäten werden also nicht bloß addiert. Sie werden stattdessen unterschiedlichen Funktionen zugeordnet, nämlich Determinans und Determinatum. Auch das macht *Fenster-Fenster* weniger kompositional. Auf dem Kompositionalitätskontinuum sind Det-ICCs also den Determinativkomposita zuzuodnen. Als Determinativkompositum ist die Bildung allerdings transparent, da sowohl Erst- also auch Zweitglied die Bedeutung des Simplex *Fenster* tragen. Es gibt aber auch einige Fälle, in denen das nicht so ist, etwa das in 4.2.5 besprochene Det-ICC *Bremsenbremse.*

Prot-ICCs sind nicht kompositional. Zum einen steht zwei formalen Bestandteilen, wie bei den Det-ICCs, nur ein Konzept, das auf nur einen Referenten verweisen kann, gegenüber. Das Prot-ICC *Oma-Oma* referiert etwa nicht auf zwei, sondern nur auf eine Oma. Zum anderen werden den Konstituenten nicht nur – wie in Det-ICCs – unterschiedliche Funktionen zugeordnet. Zum Erstglied liegt vielmehr überhaupt kein Konzept vor. Der formalen Dopplung entspricht also auch keine konzeptuelle Dopplung. Zudem sind Prot-ICCs nicht transparent, weil das Erstglied rein funktional ist und somit nicht die Semantik realisiert, die dem Nomen *Oma* als Simplex innewohnt. Allein das Zweitglied trägt die für *Oma* grundlegende Bedeutung.

Die Name-ICCs ordne ich in Kapitel 4.2.2 den exozentrischen Komposita zu. Auch diesen ICC-Typ kann man in einer Transparenz- und Kompositionalitätsmatrix einordnen und damit der generellen Kritik an einer Dichotomie 'exozentrische – endozentrische Komposita' begegnen (Dirven & Verspoor 1998). Name-ICCs müssen

als minimal transparent und minimal kompositional eingeordnet werden. Die einer Kompositionalität zuwiderlaufenden Eigenschaften, die bereits Det- und Prot-ICCs aufweisen, liegen bei Name-ICCs ebenfalls vor. Der formalen Dopplung entspricht bloß die Referenz auf eine Entität. *AutoAuto* referiert nur auf eine Show, nicht auf zwei. Zudem sind Name-ICCs nicht transparent. Beim Name-ICC *AutoAuto* trägt das Zweitglied nicht sie Semantik, die dem Simplex *Auto* innewohnt. Bei Name-ICCs wie *PunktPunkt* fügt schließlich weder Erst- noch Zweitglied dem Kompositum eine Semantik hinzu und beide entsprechen semantisch nicht dem Simplex *Punkt*. Die verschiedenen ICC-Typen ordnen sich also wie in Tabelle 19 dargestellt in die zweidimensionale Transparenz- und Kompositionalitätsmatrix ein.

Tab. 19: Einordnung der ICC-Typen in die Transparenz- und Kompositionalitätsmatrix

		Kompositional----------Nicht-kompositional				
		Additiv	Kopulativ	Determinativ	Determinativ (komplex)	Synchron kein Muster
transparent	N_1&N_2= $Basis_{1/2}$	–	–	Det-ICCs (*Fenster-Fenster*)	–	–
⋮	$N_1 \vee N_2 \neq Basis_{1/2}$	–	–	Det-ICCs (*Bremsen-bremse*)	Prot-ICCs (*Oma-Oma*)	Name-ICCs (*AutoAuto*)
opak	N_1&$N_2 \neq Basis_{1/2}$	–	–	–	–	Name-ICCs (*Punkt-Punkt*)

Diese Einordnung ist allerdings rein deskriptiv und modelliert nicht, wie ICCs ihre Bedeutung erhalten. Hierzu bedarf es einer Theorie zur Kompositasemantik sowie eines Modells, das die Erschließbarkeit und Bedeutungskonstitution von ICCs erklären kann. Im Folgenden stelle ich zunächst die einschlägigen Kompositatheorien vor und beschreibe daraufhin Modelle, die die Transparenz und Kompositionalität von Komposita zu messen versuchen. Im Zuge dessen gebe ich auch eine Einschätzung dazu ab, inwieweit die jeweiligen Modelle mit den Daten der Korpusstudien vereinbar sind, und ob sie die Konstitution von ICC-Semantik abbilden können.

9.3 Die semantische Analyse von Komposita

Die Gesamtbedeutung eines N+N-Kompositums ergibt sich aus den Konstituentenbedeutungen sowie aus der semantischen Relation, die die Konstituenten verbindet. Theorien zur Kompositasemantik unterscheiden sich nun in der Gewichtung der Einflussfaktoren Erstkonstituente, Zweitkonstituente und semantischer Relation. Man kann die unterschiedlichen Ansätze in drei Arten unterteilen: schemabasierte, relationsbasierte und Mischformen dieser beiden.

Schemabasierte Ansätzen zur Kompositasemantik behandeln die Frage, „how people combine the head noun with modifiers to come to an interpretation of the entire phrase“ (Murphy 1990: 559). Für Komposita sehen diese Ansätze also das Zweitglied als grundlegend für die Konzeptbildung an. Die semantischen Eigenschaften des Zweitgliedes bilden die Basis für die Integration mit dem Erstglied. Der Kopf öffnet ein Schema, in dessen Slot der Modifikator eingesetzt wird. Es wird also gemessen, je nach Ansatz mit unterschiedlichen Methoden, inwieweit sich die Eigenschaften des semantischen Kopfes ändern, wenn in sein Merkmalsschema ein Modifikator eingesetzt wird („slots and fillers“, Murphy 1988: 532, Wisniewski & Murphy 2005). Schemabasierte Analysen zielen also vor allem auf den Bedeutungsunterschied zwischen dem Kopf eines Kompositums und dem dazugehörigen Basislexem ab. Manche Ansätze belassen es nicht dabei und sehen

> the need for a second stage of processing beyond slot filling, i.e., ‘concept elaboration’. Once a slot is filled, world knowledge is used to refine the resulting combination.”
>
> (Olsen 2012: 2132)

Relationsbasierte Ansätze schreiben hingegen die zentrale Rolle bei der Konzeptbildung dem Erstglied sowie dem Wissen über die semantischen Relationen zu. Der Modifikator legt die Dimension fest, in der das Kompositum ein Subkonzept bildet. Wie ein Kompositum interpretiert wird, hängt nach dem relationsbasierten Ansatz damit zusammen, mit welchen semantischen Relationen der Modifikator für gewöhnlich in Komposita auftritt. Relationsbasierte Ansätze sind damit für die Beschreibung von Neubildungen geeignet, weil sie auf der Basis des Lexikons Analogieeffekte abbilden können.

Manche Modelle weisen Eigenschaften relationsbasierter und schemabasierter Ansätze auf. Diese Mischformen berücksichtigen beide Konstituenten eines Kompositums gleichermaßen. Die Mechanismen der Kompositasemantik sind auf vielfältige Weise in theoretischen Arbeiten beschrieben worden. Vier grobe Richtungen solcher theoretischen Arbeiten sind hier zu nennen (Søgaard 2005: 319): reduktionistische, transformationale, pragmatische Theorien sowie *slot-filler*-Theorien.

9.3.1 Reduktionistische Kompositatheorien

Für reduktionistische Theorien hängt die Interpretation von Komposita allein von den Konstituenten und der Art ab, wie sie verbunden sind (Hatcher 1960). Dazu wird eine sehr begrenzte Anzahl an Relationen beschrieben, etwa „A is in B", „B is in A", „A is the goal of B". Wie Søgaard (2005: 319) am Beispiel *car thief* zeigt, sind solche Theorien aber wegen der sehr wenigen Relationen, die sie zwischen den Kompositakonstituenten erlauben, sowohl arbiträr als auch unvollständig. So sind die Paraphrasen 'Auto in einem Dieb', 'Dieb in einem Auto', 'Dieb als Ziel eines Autos' und 'Dieb als Quelle eines Autos' gleichermaßen unpassend für das Kompositum *Autodieb*. Unter anderem aus diesem Grund spielen die reduktionistischen Ansätze in den aktuellen Diskussionen zur Kompositasemantik keine Rolle mehr.

9.3.2 Transformationale Kompositatheorien

Transformationale Theorien gehen davon aus, dass den Komposita längere syntaktische Konstruktionen zugrundeliegen und versuchen ergo die Kompositabedeutung aus der Syntax abzuleiten (Rhyne 1975). Dies fußt auf der Annahme, dass es ein Lexikon gibt, in dem Wörter, Affixe und andere idiosynkratische Strukturen abgespeichert werden. Die Zusammenfügung dieser Einheiten zu wohlgeformten Wortformen und Sätzen geschieht in einem Regelapparat, der häufig Syntax genannt wird (Chomsky 1965). Die Syntax ist in transformationalen Theorien somit das einzige generative System der Sprache und generiert aus den Einträgen im Lexikon einen Output für Phonologie und Semantik. Das Lexikon ist demnach in gewisser Weise (zeitlich, räumlich, konzeptuell) präsyntaktisch (Aronoff 1976, Hale & Keyser 2002, Halle 1973) und beinhaltet die Elemente, mit denen das generative Syntaxmodul operiert. Phonologische und semantische Strukturen sind in transformationalen Ansätzen also von der syntaktischen Struktur abgeleitet. Diese Grundannahme führt häufig dazu, dass Tiefenstrukturen, also Strukturen, die an der sprachlichen Oberfläche nicht sichtbar sind, angenommen werden müssen, um sprachliche Strukturen und Phänomene zu erklären.

Transformationale Ansätze der Grammatikbeschreibung werden auch zur Beschreibung von Komposition verwendet. Die Komposition ist sogar eines der ersten Phänomene, auf die man den in den 1960er-Jahren neuen Formalismus der generativen Transformationsgrammatik anwandte (Lees 1960). Lees leitet Komposita aus einer Tiefenstuktur ab, die äquivalent zu Sätzen ist, und führt etwa das Kompositum *ignition key* auf die Phrase *key which causes ignition* zurück. Diese

Annahme liegt transformationalen Ansätzen der Kompositabeschreibung generell zugrunde (etwa Levi 1978: 76f., Rhyne 1975: 36). Lees setzt dann Transformationen an, um zu erklären, dass viele Bedeutungsaspekte der Komposita (im Beispiel *ignition key* etwa die Kausalität) nicht durch syntaktische Regeln expliziert werden. Nach Lees Ansatz löscht die Transformation große Teile des generierten Materials und bestimmt die Abfolge der verbleibenden Komponenten. Lees Ansatz folgen weitere Versuche, Komposita aus einer Tiefenstuktur abzuleiten (Kay & Zimmer 1978, Warren 1978).

Der transformationale Ansatz, die Kompositasemantik zu beschreiben, wird aus unterschiedlichen Gründen kritisiert. Zunächst wird Lees Annahme, die Transformation lösche große Teile des sprachlichen Materials, innerhalb des transformationsgrammatischen Frameworks kritisiert und der Ansatz diesbezüglich weiterentwickelt. Chomsky (1965) fügt etwa hinzu, dass nur das Material gelöscht werden kann, das auch wiederherzustellen ist. In der Folge greift Levi (1978) diesen Mechanismus auf und definiert neun Recoverably Deletable Predicates (RDPs), ein Set aus Prädikaten, die bei diesem Löschvorgang verloren gehen, aber wiederhergestellt werden können (CAUSE, HAVE, MAKE, USE, BE, IN, FOR, FROM und ABOUT). Ein Problem von Levis Ansatz ist, dass er viele Kompositatypen explizit ausschließt. Weder Komposita, die metaphorisch verschoben oder anderweitig idiomatisch zu verstehen sind, noch Kopulativkomposita, noch Eigennamenkomposita lassen sich mit ihrem Ansatz beschreiben (Levi 1978: 6, Schäfer 2018: 96f.). Downing kritisiert an Levis Ansatz zudem, dass nicht klar sei, wie viel des semantischen Gehaltes durch die Formeln verloren geht, und bemerkt, dass viele Komposita sich mit keinem der RDPs beschreiben lassen (Downing 1977: 827). Auch wird die Zuordnung von Komposita zu den einzelnen Gruppen als willkürlich kritisiert (Fanselow 1981: 151ff.).

Dennoch ist Levis Ansatz der einflussreichste Versuch aller relationsbasierten Kompositatheorien und liegt vielen späteren Versuchen, Kompositasemantik zu beschreiben, implizit oder explizit zugrunde. Ó Séaghdha begründet die nachhaltige Wirkung des Ansatzes wie folgt:

> Levis proposals are informed by linguistic theory and by empirical observations, and they intuitively seem to comprise the right kind of relations for capturing compound semantics.
> (Ó Séaghdha 2008: 27)

Levis Ansatz wird in der Folge stetig weiterentwickelt. Allen (1978) versucht etwa, eine größere Vielfalt der N+N-Komposita zu erfassen, indem er eine Hierarchie der semantischen Merkmale und der Kompositakonstituenten annimmt. Sie definiert die *Variable R Condition*, eine Regel, die die semantischen Merkmale der Konstituenten und deren vielfältigen Bedeutungen berücksichtigt. Demnach er-

möglichen semantische Merkmale, die sich ergänzen, die Bildung eines entsprechenden Kompositums, während unmögliche Kombinationen der semantischen Merkmale die Bildung von Komposita hemmt.

Der Ansatz von Fanselow (1981) fügt dem Levis hinzu, dass sich die nicht explizit gegebene Relation in Komposita häufig aus den Bedeutungen der zwei Konstituenten herleiten lasse. In *Zeitungsfrau* ist beispielsweise die Bedeutung von *Zeitung* die Quelle für die inferierte Relation *zustellen*. Solche N+N-Komposita unterscheiden sich demnach also weniger deutlich von Rektionskomposita als vielfach angenommen, da erstere ebenfalls eine Information über die anzusetzende semantische Relation beinhalten. Komposita, in denen die semantische Relation nicht aus den Konstituentenbedeutungen abgeleitet werden kann, bilden nach Fanselow eine eigene, kleinere Gruppe. In Komposita wie *Küstenstadt* geht etwa aus keiner der Konstituenten die Information hervor, dass das Konzept des Zweitgliedes *Stadt* örtlich mit dem des Erstgliedes *Küste* zu verbinden ist (Fanselow 1981: 156). Die Relation solcher Komposita wird stattdessen aus fünf Basisrelationen generiert (UND, GEMACHT AUS, ÄHNELT, IST TEIL VON und IST LOKALISIERT BEZÜGLICH). Die größere Gruppe stellten aber Komposita wie *Zeitungsfrau* dar, in denen diese Basisrelationen nicht zum Einsatz kommen. Nach Fanselows Ansatz spielen also bei den meisten Komposita die stereotypen Relationen, die mit den Kompositakonstituenten assoziiert werden, die wichtigste Rolle bei der Interpretation. Zentrale Punkte aus Fanselows Ansatz werden von Brekle und später von Jackendoff adaptiert und weiter ausgearbeitet („stereotype compounds“, Brekle 1986: 42, Jackendoff 2010a).

Mancher Kritik am transformationalen Ansatz kann aber nicht begegnet werden, ohne den Ansatz als solchen gänzlich aufzugeben. Fanselow (1981) weist etwa darauf hin, dass die Reihenfolge im Erstspracherwerb dagegen spricht, dass Komposita von Phrasen abgeleitet sind, da erstere früher erworben werden als letztere. Zudem wird die Grundannahme regelbasierter Ansätze der Grammatikbeschreibung kritisch gesehen, nach der Lexikon und Syntax distinkte Module darstellen, die strikt voneinander zu trennen sind. Die damit einhergehende scharfe Grenze zwischen vollständig idiosynkratischen und vollständig regelhaften Bildungen bereitet in vielerlei Hinsicht Schwierigkeiten. Beispielsweise sind Interaktionen zwischen Phonologie, Semantik und Syntax, die sich etwa bei Idiomen zeigen, mit einer klaren Trennung von Lexikon und Syntax nicht vereinbar. Die Beispiele in (128) und (129) verdeutlichen dies:

(128) *Lass dir die Butter nicht vom Brot nehmen!*

Der Satz in (128) basiert auf einer idiomatischen Konstruktion, bei der nicht alle Positionen lexikalisch fixiert sind. Durch die fixierte Wortwahl und eine nichtkompositionale Bedeutung, die also nicht aus den Bestandteilen des Satzes abgeleitet werden kann, nimmt man nach regelbasierten Ansätzen an, dass ein solches Idiom im Lexikon abgespeichert ist. Eine strikte Unterteilung des Sprachsystems in ein Lexikon einerseits und ein Grammatikmodul andererseits kann nun aber einen Fall wie in (129) nicht erklären.

(129) *Ich will, dass wir uns die Butter nicht vom Brot nehmen lassen!*

Solche Beispiele sprechen gegen die Annahme, dass alle Prozesse im Lexikon denen in der Syntax vorangehen, während Prozesse der Phonologie und Semantik auf die syntaktischen folgen. Trotz der idiosynkratischen Bedeutung in (128–129) wirken syntaktische Regeln innerhalb dieses vermeintlichen Lexikoneintrags, sodass das Reflexivpronomen nicht mehr nach der 2.SG, sondern nach der 1.PL flektiert, das Verb *lassen* zudem nicht mehr in der 2.SG.IMP, sondern in der 1.PL.IND steht. Auch die Abfolge der Wörter ist – wie für einen mit *dass* eingebetteten Satz üblich – in Verbletztstellung geändert. All dies legt nahe, dass den Äußerungen ein phrasales Muster im Lexikon zugrunde liegt, und ist schwierig zu erklären, wenn man ein striktes Hintereinander der Module annimmt.

Die Kritik an der kategorialen Syntax-Lexikon-Einteilung führt zu Gegenentwürfen, in denen von einem Syntax-Lexikon-Kontinuum ausgegangen wird (Tabelle 20).

Tab. 20: Das Syntax-Lexikon-Kontinuum (Croft 2001: 17)

Konstruktionstyp	**Traditioneller Name**	**Beispiele**
Komplex und weitgehend schematisch	Syntax	[SUBJ *be*-TNS VERB-*en by* OBL]
Komplex und weitgehend spezifisch	Idiom	[pull-TNS NP's *leg*]
Komplex aber gebunden	Morphologie	[NOUN-S], [VERB-TNS]
Atomar und schematisch	Syntaktische Kategorie	[DEM], [ADJ]
Atomar und spezifisch	Wort / Lexikon	[*this*], [*green*]

Ein Ansatz, der zwischen den beiden Polen Lexikon und Syntax weitere Ebenen annimmt, kann angemessener erklären, warum Sprecher:innen ohne Weiteres Idiome wie das in (128) in unterschiedlichen syntaktischen Kontexten den ent-

sprechenden Regeln gemäß anpassen können, ohne dazu jeweils einen entsprechenden Lexikoneintrag zu haben. Ein Lexikon–Syntax-Dualismus vermag das nicht.

Wenngleich die Transformationsgrammatik selbst in der aktuellen Forschung nur eine untergeordnete Rolle spielt, so ist doch die Grundannahme, dass sich Sprache in einen Regelapparat (Grammatik, Syntax) einerseits und lexikalische Einheiten (Idiosynkrasien/Wörter im Lexikon) andererseits einteilen lässt, in vielen generativen Frameworks weiterhin präsent, beispielsweise in der *mainstream generative grammar* (Chomsky 1965), im *minimalist programm* (Chomsky 1993, 1995) oder in Pinkers (1997, 1998, 1999) *Item and Process*.

Weitere Kritik an transformationalen Kompositatheorien ruft der von Levi (1978) ausgehende Versuch hervor, die nicht explizierten Bestandteile der Komposita mithilfe von RDPs zu erklären. Hinsichtlich aller Versuche, die Relationen zwischen Kompositakonstituenten mithilfe vordefinierter Sets zu erfassen, wird kritisiert, dass eine solche Liste niemals vollständig sein kann und viele Komposita zu keiner der Kategorien passen (Fandrych & Thurmair 1994, Zimmer 1971). Zudem wird die Vagheit und Austauschbarkeit von fest definierten Sets an Relationen kritisiert (Søgaard 2005: 320). Ten Hacken (2019: 42ff., 1994: 44–49) merkt beispielsweise an, dass Levis RDPs viele Ambiguitäten hervorbringen, zu vage gefasst sind und bei Weitem nicht alle Komponenten der Kompositasemantik umfassen. Komposita wie *regeringsgebouw* ‘Regierungsgebäude’ und *regeringscrisis* ‘Regierungskrise’ im Niederländischen sind zum Beispiel gleichermaßen durch HAVE charakterisiert, weisen aber eine sehr unterschiedliche Relation zwischen den Kompositakonstituenten auf. Darüber hinaus seien Eigennamen und deiktische (referenzspezifzierende) Komposita vom RDP-System ausgeschlossen (Ten Hacken 2019: 43).

Auch die Bildung und Interpretation von ICCs stellt für transformationale, regelbasierte Ansätze eine Herausforderung dar. Die ICC-Typen Det-ICC, Prot-ICC und Name-ICC lassen sich zu den morphologischen Prozessen Komposition und Reduplikation in Beziehung setzen (siehe Kapitel 7 und 8). Die Bildung von Prot-ICCs ist ein Subtyp der Reduplikation und wird in Kapitel 8 als Kompositionsreduplikation bezeichnet. In dieser reduplikativen Konstruktion geht die Bedeutung ‘prototypisch’ allerdings nicht von den Konstituenten aus. Der transformationale Ansatz erlaubt nicht, die vorhersagbaren semantischen Eigenschaften der Bildungen, die nicht von den semantischen Eigenschaften der Konstituenten abgeleitet werden können, zu spezifizieren.

Angesichts der genannten Probleme, die tranformationsgrammatische und reduktionistische Ansätze bei der Beschreibung von Kompositasemantik haben, entstehen zwei Ansätze, die bis heute vornehmlich für die Analyse und Modellie-

rung von Komposita genutzt werden. Generativ werden Komposita mittlerweile mit schemabasierten *Slot-Filler*-Theorien modelliert, die auf der Annahme basieren, dass der Modifizierer die Semantik der Kopfkonstituente erweitert. Zudem begegnet der pragmatische Ansatz den Defiziten der klassischen, reduktionistischen Ansätze, indem er die Perspektive der Sprecher:innen einbezieht, etwa die Ansätze von Zimmer (1972) oder Bauer (2006b).

9.3.3 *Slot-Filler*-Theorien zur Kompositasemantik

Slot-Filler-Theorien wurden im Framework des *Generative Lexicon* formuliert und basieren auf der Annahme, dass der Modifikator eines Kompositums die Semantik der Kopfkonstituente erweitert. Der Ansatz geht von einer starken Modifikator-Kopf-Asymmetrie aus. Die Kompositakonstituenten werden dabei als Merkmalsbündel aufgefasst und der Modifikator fügt letztlich bloß ein Merkmal zur Kopfkonstituente hinzu. Zunächst muss man den Ansatz des *Generative Lexicon* verstehen, um *Slot-Filler*-Theorien auf Komposita anzuwenden.

Der Ansatz des *Generative Lexicon* geht auf Pustejovsky (1991) zurück, der annimmt, dass zwischen mehreren Ebenen semantischer Repräsentation zu unterscheiden ist, und es einer Reihe generativer Mechanismen bedarf, um neue Bedeutungen zu formen. Ein klar definiertes Set solcher generativen Mechanismen produziert im *Generative Lexicon* semantische Ausdrücke. Den Input zu diesen Mechanismen bilden vier Strukturebenen: die Argumentstruktur, die Eventstruktur, die Qualia-Struktur und die Struktur lexikalischer Vererbung. Die wichtigste Neuerung im *Generative Lexicon* gegenüber vorhergehenden generativen Ansätzen ist die Qualiastruktur. Die Qualiastruktur eines Lexikoneintrags beschreibt Pustejovsky als „the structured representation which gives the relational force of a lexical item“ (Pustejovsky 1991: 76). Sie besteht wiederum aus der formalen Quale, die die Bedeutung eines Wortes innerhalb einer Domäne unterscheidet und die Position innerhalb einer Vererbungshierarchie bestimmt, der konstitutiven Quale, die die Relation zwischen einem Objekt und seinen Teilen festlegt, dem agentiven Quale, das mitteilt, wie die semantischen Typen zustandekommen und dem telischen Quale, das den Zweck der Referenten bestimmt.

Die Qualiastruktur ist dazu da, die lexikalischen Informationen mit metaphysischen Grundannahmen zu verbinden. N+N-Komposita verwenden nun jeweils unterschiedliche Bedeutungsaspekte ihrer Konstituenten. Dies kann dazu führen, dass für neue Komposita auch neue Bedeutungsaspekte zum Lexikoneintrag der jeweiligen Konstituenten hinzugefügt werden müssen. Die Qualiastruktur begrenzt diese Notwendigkeit. Bei Komposita wie *Schussverletzung* und *Handverlet-*

zung bestehen beispielsweise trotz des gleichen Kopfes *Verletzung* die unterschiedlichen Relationen CAUSE und LOC. Dennoch müssen keine neuen Bedeutungsaspekte zum Kopf *Verletzung* eingetragen werden, weil der Modifikator im Falle von *Schussverletzung* den Argumentslot der agentiven Quale, bei *Handverletzung* hingegen den Argumentslot der konstitutiven Quale von *Verletzung* besetzt. Komposita gelten im Rahmen des *Generative Lexicon* als „specification of one of the semantic components within the qualia of the head noun" (Johnston & Busa 1999). Beim Kompositum *Apfelsaft* beispielsweise gehe demnach vom Kopfnomen eine agentive Rolle aus, nämlich die, dass ein *Saft* stets aus etwas gemacht ist. Dies motiviert die Relation MADE OF. Da der Modifikator *Apfel* als Frucht mit dieser Relation vereinbar ist, anders als beispielsweise *Kinder* in *Kindersaft*, kann der Modifikator den semantischen Typ spezifizieren, der zu einem Argument des Kopfnomens passt.

Slot-Filler-Theorien können die Vielfalt der Relationen zwischen Kompositakonstituenten angemessener abbilden als reduktionistische und transformationale Modelle. Doch auch sie erfassen nicht alle Komposita. Das Beispiel *Apfelsaft* zeigt, wie die MADE OF-Relation entsteht, und dass sie angesichts der Semantik des Kopfnomens vorherrschend ist. Dennoch gibt es, wie Søgaard (2005: 325) zeigt, Komposita, die dieser Struktur nicht entsprechen, etwa das erwähnte *Kindersaft* oder die Komposita *Magensaft* oder *Hustensaft*. Zudem ist der Ansatz nicht mit exozentrischen Komposita vereinbar, da der Kopf in solchen Komposita keine Slots öffnet, die von Bedeutungsaspekten der Modifikatoren besetzt werden könnten.

In Bezug auf ICCs sind Det-ICCs zwar mit dem Ansatz beschreibbar. In *Fenster-Fenster* beispielsweise öffnet der Kopf *Fenster* einen Argumentslot der konstitutiven Quale von *Fenster* (örtliche Spezifizierung, Teil-Ganzes), die der Bedeutungsaspekt 'Gebäudebestandteil' des Modifikators *Fenster* füllen kann. Bei Prot-ICCs aber sind die Slots, die das Kopfnomen aufmacht, irrelevant für die Interpretation des ICCs, da das Erstglied semantisch leer ist und unabhängig von der Qualiastruktur der Basis stets der Bedeutungsaspekt 'prototypisch, üblich' hinzugefügt wird. Name-ICCs sind darüber hinaus als exozentrische Komposita nicht mit dem *Slot-Filler*-Ansatz abzubilden, weil hier die Werte von Erst- und Zweitglied keine Rolle für die Referenz spielen und die Bildungen keine Spezifikationen von semantischen Komponenten innerhalb der Qualia des Kopfnomens sind.

In ähnlicher Weise wie Levi (1978) und Allen (1978) versucht auch Jackendoff, ein Set grundsätzlicher Relationen zwischen Kompositakonstituenten anzusetzen. Neu an Jackendoffs Ansatz ist, dass er über diese Sets hinaus Mechanismen entwickelt, wie aus den grundsätzlichen Relationen spezifischere Relationen generiert werden können (Jackendoff 2009: 123, 2010b: 436ff., 2016: 27ff.). In der konzeptuel-

len Information der Kompositakonstituenten sind demnach Aktionsmodalitäten („action modalities“) festgelegt, bei Artefakten etwa „proper functions“, die die Modifikationsrelation innerhalb eines Kompositums spezifizieren. Jackendoffs Ansatz geht hier über die Slots der Slot-Filler-Ansätze hinaus, da die Basisfunktionen auch zu komplexeren Relationen kombiniert werden können und sich so unendlich viele Relationen beschreiben lassen. Der Ansatz ist allerdings nicht dem generativen Framework zuzurechnen. So setzt Jackendoff etwa keine Transformationen für die Verbindung von Form und Bedeutung an. Die Unbegrenztheit der Relationen ist für seinen Ansatz daher, anders als bei den formalen Ansätzen, kein Problem. Jackendoffs Ansatz, Kompositasemantik zu modellieren, geschieht im weiteren Rahmen seines grammatiktheoretischen Beschreibungsmodells der *Parallel Architecture*, das in Kapitel 10 detailliert dargestellt wird.

9.3.4 Pragmatische Kompositatheorien

Die bisher präsentierten Ansätze, Kompositasemantik zu beschreiben, fokussieren sich allesamt auf die Semantik der formal gegebenen Konstituenten und die implizite Relation zwischen diesen. Eine mögliche Ursache für die Probleme in diesen Ansätzen ist, dass der Kontext und die verschiedenen pragmatischen Aspekte weitgehend ausgeblendet werden. Pragmatische Frameworks behandeln eben diese Rolle des Weltwissens, über das die Sprecher:innen verfügen. Die Verbindung zwischen den Konstituenten wird in pragmatischen Ansätzen der Kompositasemantik bewusst sehr vage beschrieben, etwa „A ist verbunden mit B“ („connected-with-hypothesis“, Søgaard 2005: 321). Bauer nimmt an, dass zwischen den Konstituenten eine abstrakte Formel wie etwa „there is a connection between“ besteht (Bauer 1979: 46). Die Möglichkeit zur formalen Realisierung wird somit nahezu verneint und stattdessen vor allem pragmatisches Wissen bei der Interpretation von Komposita als grundlegend angesehen (Lieber 2004: 49, Selkirk 1982: 23). Im *Conceptual Blending*-Framework wird die Ansicht formuliert, dass die overte sprachliche Struktur so wenige Hinweise darauf gibt, wie die Input-Frames integriert werden müssen, dass die Sprecher/Hörer:innen auf Kontext- und Weltwissen angewiesen sind, um das Kompositum zu verstehen (Coulson & Fauconnier 1999). „The exact relation must be inferred“ (Levinson 2000: 147), wozu es konversationeller Implikaturen bedürfe.

Grundzüge des pragmatischen Ansatzes finden sich in vielen Kompositatheorien, etwa bei Zimmer (1972), der davon ausgeht, dass die Klassifikation von N+N-Komposita stark von der Dichotomie „Benennen – Beschreiben“ abhängt. Er beschreibt die „Appropriately Classificatory Relationship“, die berücksichtigt, ob

eine Relation von den Sprecher:innen als relevant für die Klassifikation angesehen wird. Auch spätere Ansätze geben die klassische semantische Kategorisierung auf und beziehen immer mehr den/die Sprecher:in, das Weltwissen und den Kontext mit ein (Lieber 2004, Selkirk 1982).

Eine extreme Version der *Connected-With*-Hypothese macht Generalisierungen zu Komposita nahezu unmöglich. Den entsprechenden Ansätzen wird daher vorgeworfen, Systematisches generell zu ignorieren, weil die Interpretation von Komposita allein vom Kontext- und Weltwissen abhängt. Es ist in der Tat zu bezweifeln, dass etwa Coulson und Fauconniers (1999) Beispiel *stone lion* allein über den Kontext und nicht unter Zuhilfenahme des MADE OF-Musters, das so vielen N+N-Komposita zugrunde liegt, interpretiert wird. Ich stimme hier Sølgaards Kommentar zu, der hinsichtlich der allzu extremen Versionen pragmatischer Ansätze schreibt:

> [T]he semantics of compounding is not explained [...] by passing the job on to trash can pragmatics.
>
> (Søgaard 2005: 336)

Stattdessen müssen für eine angemessene Beschreibung von Kompositasemantik sowohl die kontextuellen und pragmatischen als auch die lexiko-semantischen Aspekte berücksichtigt werden. Zwar kann die Klassifikation von Komposita nach einem begrenzten Set semantischer Relationen die endgültige Bedeutung von Komposita nicht vorhersagen. Die Sets stellen aber sinnvolle Generalisierungen dar und sind weiterhin für die Analyse von Kompositasemantik sinnvoll. Hinsichtlich der ICC-Bedeutung hat etwa die Studie von Finkbeiner (2014) gezeigt, dass auch ohne Kontext und pragmatische Anreicheurng eine relativ feste Bedeutung der Ad hoc-Bildungen entsteht. Für die korrekte Interpretation eines Kompositums spielen neben dem Kontext eben auch Stereotypen, wie sie etwa von Fanselow beschrieben werden, sowie Analogien zu anderen Komposita im Lexikon eine Rolle.

9.4 Modelle zur Bestimmung der Erschließbarkeit von Komposita

Es gibt verschiedene Ansätze, die Bedeutungskonstitution von Komposita zu modellieren. Die Ansätze zielen mal auf semantische Transparenz, mal auf Kompositionalität und mal auf die generelle Erschließbarkeit von Komposita ab. Viele dieser Ansätze setzen Kompositionsbedeutung und Konstituentenbedeutung in Bezug. Die verschiedenen Modelle nutzen unterschiedliche Methoden, um zu

bestimmen, inwieweit die Konstituente eines Kompositums von seiner morpholexikalischen Bedeutung abweicht und wie sehr das Gesamtkompositum selbst eine Reduktion von Kompositionalität bedingt. Manche Modelle versuchen zudem, die Erschließbarkeit von Komposita zu quantifizieren. Hinsichtlich der Methoden, mit denen der Grad semantischer Transparenz / Kompositionalität / Erschließbarkeit bestimmt wird, teilen sich die verschiedenen Modelle grob in zwei Arten ein: solche, die auf Korpusdaten basieren, und solche, die auf experimentell erhobenen Daten basieren. Bei Letzteren sind wiederum assoziationsbasierte, bewertungsbasierte und lexikonbasierte Ansätze zu unterscheiden. Je nach Methode machen die Modelle unterschiedliche Aussagen zur Erschließbarkeit / Transparenz / Kompositionalität von ICCs.

Zinsmeister (2013) versucht, semantische Transparenz von N+N-Komposita korpusbasiert zu modellieren. Sie erfasst zu den Komposita und ihren Konstituenten Faktoren wie Textnachbarschaft, Kontextwörter und Selektionsbeschränkungen. Aus diesen Faktoren ergibt sich eine Schiefeabweichung, also ein Vektor, der die semantische Übereinstimmung von Kompositakonstituenten und Gesamtkompositum abbildet. Diesem Verfahren liegt die Annahme zugrunde, dass Wörter, die in demselben Kontext verwendet werden, die also ähnliche Textnachbarn und Kotexte haben, auch eine ähnliche Bedeutung tragen („Distributional Semantics“, Firth 1957, Harris 1954, Marconi 1997).

Wenn die Komposita in diesen (und weiteren) Faktoren ähnliche Ausprägungen haben wie ihre Konstituenten, spricht das für die Transparenz des Kompositums. Zinsmeister (2013: 309) vergleicht beispielsweise die Komposita *Milchzahn* und *Löwenzahn*. *Milchzahn* weicht hinsichtlich der genannten Faktoren weniger stark von *Zahn* ab als *Löwenzahn*. Die Semantik von *Milchzahn* überschneidet sich also stärker mit der von *Zahn* als die von *Löwenzahn*. Zinsmeister kommt auf diese Weise außerdem zu dem Ergebnis, dass Determinativkomposita in einem höheren Maße mit ihrem Kopf als mit ihrem Modifikator semantisch übereinstimmen.

Für ICCs würde das allerdings anders aussehen, da die Schiefeabweichungen wegen der Identität der Konstituenten notwendigerweise gleich wären. Es bestünde bei ICCs demnach dieselbe semantische Gleichheit zwischen Erstglied und Gesamtkompositum wie zwischen Zweitglied und Gesamtkompositum. Diese Divergenzwerte wären außerdem sehr niedrig, da das Erstglied in einem ICC automatisch denselben Wert bei Faktoren wie Textnachbarschaft oder Kontextwörter erhalten würde wie das Zweitglied, für das Zinsmeister empirisch ja generell einen hohen Wert belegen konnte. ICCs wären nach diesem Ansatz also sehr viel transparenter als Komposita mit unterschiedlichen Konstituenten.

Für Prot-ICCs ist das nachvollziehbar, weil der Wortbildungsprozess das Denotat des Zweitgliedes nicht verschiebt, sondern lediglich auf seinen prototypischen Kern beschränkt. Einem Prot-ICC kommen also vermutlich tatsächlich zu einem hohen Grad die Textnachbarn, Kontextwörter etc. zu wie dem jeweiligen Basislexem. Es ist hier lediglich eine empirische Frage, ob sich im Kontext von *Beziehung-Beziehung* beispielsweise dieselben Lemmata finden wie beim Basislexem *Beziehung* (etwa die Lemmata *zwischenmenschlich*, *Paar* und *einlassen*). Im Vergleich zwischen *Beziehung* und *Geschäftsbeziehung* kann man hingegen größere Unterschiede im Kontext erwarten, da hier ein weiterer Bedeutungsbereich hinzukommt und Lemmata wie *Marktstellung*, *profitabel* oder *Meeting* wohl für *Geschäftsbeziehung* eine größere Rolle spielen als für die Basis *Beziehung*. Der hohe Transparenzwert, den Prot-ICCs nach dem Modell von Zinsmeister auf diese Weise (zumindest theoretisch) erzielen, würde also zunächst gerechtfertigt erscheinen.

Bei Det-ICCs allerdings wäre ein im Vergleich zu Komposita unterschiedlicher Konstituenten höherer Transparenzwert nicht nachvollziehbar. Det-ICCs sind bis auf die reduplikative Form gewöhnliche Determinativkomposita. Die Kontext-Unterschiede, beispielsweise zwischen *Zinseszins* und *Zins*, sollten also zu einem recht ähnlichen Wert führen wie bei anderen Determinativkomposita mit *Zins*, etwa *Kreditzins*. Ebenso bei Name-ICCs. Durch die eigennamentypische Entkopplung von der lexikalischen Semantik der Basislexeme wäre sogar zu erwarten, dass die Schiefeabweichungen zwischen Name-ICCs und deren Basen viel höher ist als bei Komposita mit semantischem Kopf. Eine Bildung wie etwa *Punkt-Punkt* teilt also vermutlich sehr viel weniger Kontextwörter mit *Punkt* als beispielsweise *Lichtpunkt* mit *Licht* und *Punkt* teilt.

Allein über Faktoren wie Textnachbarschaft, Kontextwörter etc. lässt sich die Transparenz von ICCs also nicht modellieren. Zinsmeisters Modell trifft zu ICCs Voraussagen, die eher unwahrscheinlich sind. Die Bedeutungskonstitution von ICCs lässt sich aber auch mit anderen Ansätzen, semantische Transparenz zu modellieren, nicht beschreiben. Das Modell von Schulte im Walde und Borgwaldt (2015) etwa modelliert semantische Transparenz mithilfe von Assoziationen der Sprachbenutzer:innen. Die Verfasserinnen ermitteln experimentell, welche Assoziationen Proband:innen zu Komposita haben, und welche Assoziationen sie zu den beiden Bestandteilen der Komposita haben. Die so erhobenen Assoziationsdaten stellen die Autorinnen Transparenzurteilen zu denselben Items aus einer anderen Studie (von der Heide & Borgwaldt 2009) gegenüber und kommen zu dem Ergebnis, dass der Grad, zu dem sich die Assoziationen überschneiden, ein Indikator für die semantische Verbindung zwischen Gesamtkompositum und den beiden Konstituenten ist. *Ahornblatt* teilt beispielsweise 87 Assoziationen mit

Ahorn und 52 mit *Blatt* (Schulte im Walde & Borgwaldt 2015: 1210). Der Gesamtwert aus diesen beiden Werten korreliert mit den zuvor in von der Heide & Borgwaldt (2009) erhobenen Transparenzwerten. Bei opaken Komposita gibt es im Vergleich dazu signifikant weniger Assoziationsüberschneidungen zwischen den Konstituenten und dem Kompositum, das sie bilden (Schulte im Walde & Borgwaldt 2015: 1211). Wendet man diesen Ansatz, Kompositatransparenz zu quantifizieren, auf ICCs an, käme ICCs wiederum ein sehr hoher Grad an Transparenz zu. Da die beiden Konstituenten identisch sind, stimmen auch die Assoziationen vollständig überein. Dass dies mit einer hohen Transparenz der Bildungen korreliert, nimmt das Modell aber nur für lexikalisierte Komposita an. Da ICCs aber in den allermeisten Fällen nicht lexikalisiert sind, scheidet auch dieses Modell für die Beschreibung der ICC-Bedeutungskonstitution aus.

Dezidiert auf Okkasionalismen richtet sich dagegen das PI-Modell (präferierte Interpretation) von Zürn (2016: 116ff.). Nach diesem Modell ist eine einheitliche Interpretation von Komposita nicht nur vom Grad der Usualisierung oder vom Kontext abhängig. Stattdessen hänge die Erschließbarkeit von Nominalkomposita vor allem davon ab, ob die Konzepte der Konstituenten konvergent sind, womit gemeint ist, dass die Konzepte einen gemeinsamen Kontext haben, in dem sie auftreten (Zürn 2016: 3). Diese Annahme ist auch Teil der von Ortner und Ortner formulierten Voraussetzung für Komposita, dass nämlich die Konstituenten „semantisch und sachlogisch kompatibel“ sein müssen (Ortner & Ortner 1984: 27). Zürn relativiert diese Aussage dahingehend, dass, wenn die Konzepte der Konstituenten zueinander divergent sind, Nominalkomposita in dem Fall Kontext und Usualisierung benötigen, um einheitlich interpretiert zu werden.

Sind die Konstituentenkonzepte divergent, besteht ein Gleichgewicht der möglichen Konzeptverbindungsstrategien (Zürn 2016: 130). Besteht zwischen den Konzepten der Konstituenten aber eine semantische Nähe, gebe es eine klar präferierte Interpretation (Zürn 2016: 130). In einer experimentellen Untersuchung zeigt sie, dass ein Okkasionalismus wie *Ampelmaut* von den Rezipient:innen einheitlicher interpretiert wird als beispielsweise *Apfelgehör*, weil für die Konstituenten *Ampel* und *Maut* der gemeinsame Kontext „Straßenverkehr“ bestehe, für *Apfel* und *Gehör* hingegen gebe es keinen gemeinsamen Kontext (Zürn 2016: 13). Der Faktor der Konzeptkonvergenz bedingt die Erschließbarkeit von Komposita zusammen mit anderen Faktoren, wie etwa den Assoziationen, die die beteiligten Nomina hervorrufen (Zürn 2016: 112). Zürn nimmt für die semantischen Beziehungsverhältnisse ein Kontinuum an:

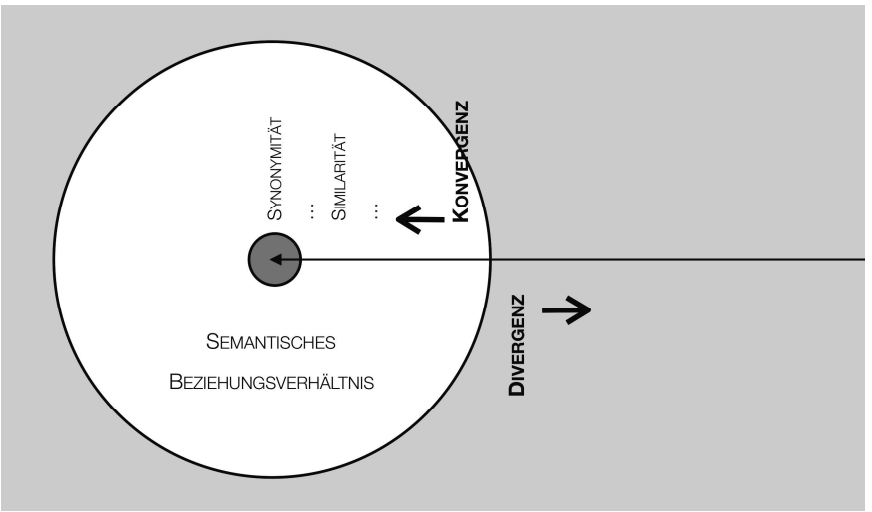

Abb. 39: Kontinuum der semantischen Beziehungsverhältnisse (Zürn 2016: 12).

Ähnlich wie beim Ansatz der distributionellen Semantik Zinsmeisters (2013) nimmt bei Zürn die Konvergenz zweier Konzepte zu, wenn sie wie *Ampel* und *Maut* über gemeinsame Auftretenskontexte und semantische Merkmale verfügen oder in einem Verhältnis semantischer Ähnlichkeit bis hin zur Synonymie stehen. Angewandt auf ICCs sagt auch das PI-Modell von Zürn eine hohe Erschließbarkeit und Transparenz für die Bildungen voraus. Die Konzepte der Kompositakonstituenten sind nicht nur ähnlich, sondern identisch. In Abbildung 39 wären ICCs also im Zentrum des Kreises zu verorten. Als Bildungen mit vollständig konvergenten Konstituentenkonzepten wären ICCs für ihre Interpretation demnach also noch weniger auf den Kontext oder eine Usualisierung angewiesen als die von Zürn bereits als hocherschließbar eingeschätzte *Ampelmaut*. Diese Vorhersage des Modells widerspricht der in der Forschung vielfach getroffenen Annahme, dass ICCs mehr als andere Komposita vom Kontext abhängig sind (Freywald 2015: 923, Hohenhaus 2007: 25ff., 2015: 275).

Das Grundproblem dieses Modells ist, dass es keine Unterscheidung zwischen Modifikatorkonzept und Kopfkonzept vornimmt. Das CARIN-Modell („Competition Among Relations in Nominals“, Gagné & Shoben 1997), und seine Weiterentwicklung RICE („Relational-Interpretation-Competitive-Evaluation“, Spalding et al. 2010), berücksichtigen diese unterschiedlichen Funktionen der Kompositakonstituenten und nehmen an, dass Kompositakonstituenten mit bestimmten Modifikationsrelationen assoziiert werden und diese Relationen auch Teil des Eintrags im

mentalen Lexikon sind. Wird ein Kompositum gebildet, konkurrieren die Relationen, die für die Kopfkonstituente gespeichert sind, mit denen, die für die Modifikatorkonstituente gespeichert sind, und beide sind in den Prozess der Interpretation involviert:

> [T]he modifier suggests relations that compete with each other for selection and the head noun then plays an important role in evaluating whether a suggested relational interpretation is a plausible meaning for the combination.
>
> (Spalding et al. 2010: 284f.)

Nach dem Modell beinhaltet die Repräsentation des Modifiziererkonzeptes also Wissen über die thematischen Relationen, mit denen der Modifizierer für gewöhnlich in konzeptuellen Kombinationen verwendet wird. Die Ansätze von Spalding und Kolleg:innen werden darum auch als „suggest-evaluate framework" bezeichnet (Brunner et al. 2021: 15, Spalding et al. 2010: 286). Ausschlaggebend für die Verbindung der Konstituentenkonzepte ist die Familie der Komposita mit demselben Modifizierer. Dieses Wissen über die Relation beeinflusst die Interpretation neuer Komposita. Wenn ein Modifizierer wie etwa *mountain* im Englischen meist mit einer LOC-Relation verwendet wird (*mountain cabin*, *mountain resort*), wenden Hörer:innen die LOC-Relation auch auf neue Komposita mit *mountain* als Erstglied an (Štekauer 2009: 435). Eine Neubildung mit *mountain* als Erstglied wird also schneller erschlossen, wenn die Bildung auch eine LOC-Relation aufweist, etwa bei *mountain stream.* Die Interpretation, die die Modifikatorkonstituente präferiert, kann aber auch abgelehnt werden, wenn die semantischen Eigenschaften der Kopfkonstituente eine solche Lesart nicht zulassen. Für das Beispiel *mountain* wäre das etwa der Fall, wenn die Kopfkonstituente *planet* oder *journal* ist und also die LOC-Relation ausgeschlossen ist (*mountain planet* b 'ein Planet am Berg', *mountain journal* b 'eine Zeitschrift am Berg'). Wird wie bei *mountain planet* und *mountain journal* eine für *mountain* unüblichere Relation angewandt, in diesen Fällen etwa HAVE und ABOUT, erschließen die Perzipient:innen die Neubildung langsamer (Gagné & Shoben 1997).

Nach dem Ansatz von Gagné (2002) entwickeln Sprecher:innen ein probabilistisches System, um anhand des Erstgliedes eines Kompositums vorherzusagen, welche semantische Relation bei neuen Komposita mit diesem Erstglied am wahrscheinlichsten ist. Nach dem CARIN-Modell spielt der Modifizierer hier eine wichtigere Rolle als der semantisch-grammatische Kopf des N+N-Kompositums. Das Modell gehört also zu den relationsbasierten Kompositatheorien.

Die semantischen Relationen in diesem Modell sind im Wesentlichen die, die Levi (1978) beschreibt. Im CARIN-Modell wird davon ausgegangen, dass diese Relationen kognitiv real sind und bei der konzeptuellen Repräsentation von Kom-

posita eine wichtige Rolle spielen. Gagné und Spalding (2009) können nachweisen, dass semantische Relationen die Verarbeitung von Komposita beeinflussen. Ein Kompositum wie *snowball*, das die semantische Relation MAKE beinhaltet ('ein Ball, der aus Schnee gemacht ist'), kann bei der Verarbeitung im Gehirn Komposita derselben Relation, etwa *snowfort,* besser voraktivieren (Priming) als Komposita, bei denen eine andere Relation zwischen den Konstituenten vorliegt. Die Stärke des Primes zeigt sich dabei konkret in kürzeren Reaktionszeiten der Versuchspersonen in Reaktionszeitexperimenten oder in einem geringen Verarbeitungsaufwand reflektierenden Kurvenverlauf im EEG. *Snowshovel* wäre demnach ein schlechterer Prime für *snowball*, da nicht die Relation MAKE, sondern die Relation FOR vorliegt. Das Wissen über die semantische Relation wird nach diesem Modell also zusammen mit der Modifikatorkonstituente im mentalen Lexikon gespeichert.

Diese Ergebnisse bedeuten nicht notwendigerweise, dass die semantischen Relationen zweier Komposita exakt identisch sein können. Die Relationen, die bei *snowball* und *snowfort* vorliegen, sind stattdessen wohl lediglich einander ähnlicher als die von *snowball* und *snowshovel.* Mit der Einschätzung, dass man semantische Relationen nicht scharf umreißen und ihre Anzahl nicht konkret bestimmen kann (etwa Eisenberg 2006: 229ff.), sind die Forschungsergebnisse von Gagné und Spalding (2009) also durchaus vereinbar.

Gagné und Shoben (1997) quantifizieren in ihrem Modell den Einfluss des Modifikators mithilfe der „strength ratio". Da die Modifikatorkonstituenten nicht in allen Komposita mit derselben semantischen Relation auftreten, stehen bei der Analyse neuer Komposita meist mehrere Kandidaten in Konkurrenz zueinander. Die Perzipient:innen errechnen also auf der Grundlage ihres mentalen Lexikons, welche semantische Relation die höchste Wahrscheinlichkeit hat, Anwendung zu finden. Ist die wahrscheinlichste Relation auch mit der Kopfkonstituente kompatibel, wird das entsprechende Kompositum hinsichtlich dieser Relation spezifiziert.

Durch die unterschiedlichen Funktionen, die die Modifikatorkonstituenten und die Kopfkonstituenten in diesem Modell übernehmen, sind ICCs nicht automatisch einfacher zu erschließen als andere Komposita. Das ist ein großer Vorteil gegenüber den zuvor vorgestellten Modellen, die gleichrangig Informationen zu Erst- und Zweitglied zur Grundlage haben. Das Modell von Gagné und Spalding (2009) ist damit weniger anfällig für die strukturelle Besonderheit von ICCs, ein und denselben Stamm zweimal zu realisieren. Bildungen wie *Beziehung-Beziehung* haben ebenso die Modifikatorkonstituente *Beziehung* wie etwa das Kompositum *Beziehungsdrama* und somit hängt auch die Erschließbarkeit von ICCs davon ab, mit welcher Relation die Modifikatorkonstituente meistens ver-

wendet wird. Allerdings ist hier vorausgesetzt, dass ICCs über eine semantische Relation verfügen. Gerade bei Prot- und arbiträren Name-ICCs ist diese Voraussetzung aber nicht gegeben.

Ein weiteres Problem ist, dass nahezu alle Kompositatheorien, seien sie transformational, generativ, pragmatisch oder den *slot-filler*-Theorien zuzurechnen, Metaphern und Metonymien bei der Beschreibung von Kompositasemantik ignorieren und entsprechende Bildungen meist als schlicht nicht analysierbar ansehen (Benczes 2006: 3). Solche Theorien sind deshalb ausschließlich auf endozentrische Komposita anzuwenden. Auch die bisher vorgestellten Ansätze, semantische Transparenz und Erschließbarkeit von Komposita zu modellieren, implementieren Metaphern und Metonymie nicht. Wie in Kapitel 4.5 sowie 9.2.2 dargestellt kommen aber beispielsweise metonymische Verschiebungen auch in ICCs vor. Die Bedeutungskonstitution von ICCs muss also Metaphern und Metonymien abbilden können.

Das Modell von Bell und Schäfer (2013, 2016) umfasst eben solche semantischen Verschiebungen und Spezifizierungen der Kompositakonstituenten und berücksichtigt außerdem, ähnlich wie das Modell von Gagné und Shoben (1997), die Verteilung der unterschiedlichen Lesarten innerhalb der Konstituentenfamilien sowie die Tatsache, dass die Bedeutung der Konstituenten eines Kompositums von der Bedeutung abweichen kann, die ihnen als freien Lexemen innewohnt. Schließlich trägt Bell und Schäfers Modell dem Umstand Rechnung, dass das Welt- und Kontextwissen der Sprachbenutzer:innen eine wichtige Rolle bei der Interpretation von Komposita spielt. Ich stelle das Modell wegen all seiner Vorzüge nun etwas detaillierter vor und wende es auf ICCs an.

Die zunächst unterspezifizierte Relation R wird angereichert und führt metaphorisch oder metonymisch[76] zu einer Verschiebung zwischen den Konstituenten A und A', beziehungsweise B und B'. (Abbildung 40).

76 Man kann durchaus hinterfragen, ob notwendigerweise diese beiden Typen der semantischen Bedeutungsverschiebung vorliegen müssen oder möglicherweise auch andere (Synekdoche, Pejorisierung) denkbar sind.

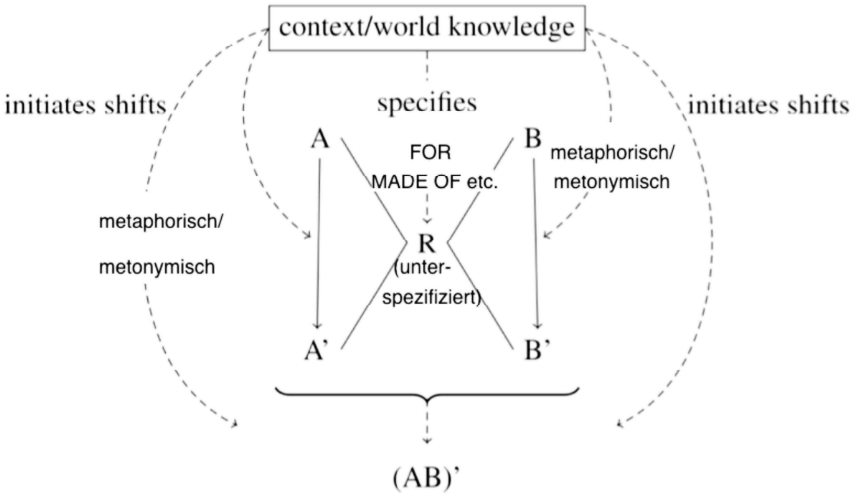

Abb. 40: Modell zur semantischen Transparenz von Nominalkomposita nach Bell & Schäfer (2013: 2, 2016: 166, leicht modifiziert).

Bell und Schäfer exemplifizieren ihr Modell anhand des englischen Kompositums *buttercup*, das eine Pflanze bezeichnet, deren Blüte die Farbe von Butter hat und von der Form her an eine Tasse erinnert. Die Bedeutung der Konstituente *butter* wird hier metaphorisch zu 'die Farbe von Butter habend' verschoben, die Bedeutung der Konstituente *cup* auf ähnliche Weise zu 'die Form einer Tasse habend'. Die Kombination wird dann wiederum metonymisch verschoben, sodass die Farbe und die Form nicht für die Blüte stehen, sondern für die gesamte Pflanze. Nachdem also zunächst die Konstituenten metaphorisch über die Konzepte FARBE und FORM verschoben worden sind (*buttercup* 'Buttertasse' V *yellowflower* 'Gelbblüte') wird in einem zweiten Schritt metonymisch die Bedeutung zur Bezeichnung der gesamten Pflanze verschoben (*yellowflower* 'Blüte' V *yellowflower* 'Pflanze'). Auf diese Weise wird die Exozentrik des Kompositums modelliert. Diese Verschiebungen werden auf der Grundlage des Kontextes und des Weltwissens der Sprecher:innen inferiert (Bell & Schäfer 2016: 166). Auch die Modifikationsrelation R wird auf dem Kontext und dem Weltwissen der Sprecher:innen basierend spezifiziert. Dadurch wird beispielsweise möglich, dass *drum sticks* eine FOR-Relation, *bread sticks* aber eine MAKE-Relation aufweist (Bell & Schäfer 2016: 166).

Angewendet auf das Det-ICC *Fenster-Fenster* wird nach diesem Modell das Lexem *Fenster*, das also sowohl in A als auch in B repräsentiert ist, semantisch zu den Konstituenten *Fenster* in A' und *Fenster* in B' verschoben. Der Äußerungskon-

text und das Weltwissen verschieben *Fenster* in A zu A' dergestalt, dass es unspezifisch jegliche Entität der Klasse der Fenster meint. *Fenster* in B wird zu B' verschoben und meint dann ein konkretes Fenster. Die zunächst unterspezifizierte Relation R wird, ebenfalls durch den Äußerungskontext und das Weltwissen, zur Relation LOC spezifiziert. Hinzuzufügen ist noch, dass im Zuge des Prozesses eine Subklassifizierung vorgenommen und den Perzipient:innen angezeigt wird. Det-ICCs funktionieren nach diesem Modell genau wie Komposita mit unterschiedlichen Konstituenten.

Angewendet auf das Prot-ICC *Oma-Oma* gibt es in diesem Modell zwei mögliche Wege der Bedeutungskonstitution. Der erste berücksichtigt, dass zwischen den Konstituenten von Prot-ICCs keine semantische Relation besteht, da keine zwei Konzepte in Beziehung gesetzt werden und das Erstglied rein funktional ist. Es wird demnach *Oma* in A zu A' verschoben, wobei die konzeptuellen Inhalte von Oma gänzlich gelöscht werden, da allein die Identität von A mit B die Prototypenbedeutung anzeigt. *Oma* in B wird zu B' verschoben, wobei das Weltwissen und der Äußerungskontext alle nicht prototypischen, nicht intendierten Eigenschaften löschen. Außerdem wird die Relation R vollkommen gelöscht, da sie in Prot-ICCs nicht vorliegt. Durch den gesamten Prozess entsteht ein Subkonzept. Wenn aber die Semantik von A und zusätzlich die Relation R gelöscht wird, beinhaltet das entstehende Subkonzept nicht mehr als die Information 'unbestimmte Entität des Konzeptes OMA'. Es ist fraglich, wie hieraus die Bedeutung entsteht, nach der eine *Oma-Oma* eine Person ist, die zur Gruppe der Omas zählt und sich durch das Vorhandensein der prototypischen Eigenschaften einer Oma von den anderen Vertretern dieser Klasse unterscheidet.

Der zweite Weg berücksichtigt, dass es eine Bedeutungsabweichung zwischen dem Lexem *Oma*, das also sowohl in A als auch in B repräsentiert ist, und den formgleichen Konstituenten *Oma* in A' und *Oma* in B' gibt. Der Äußerungskontext und das Weltwissen verschieben metonymisch (totum pro parte) *Oma* in A zu A' (Mensch eines spezifischen Verwandtschaftsverhältnisses V nur die mit diesem in Verbindung gebrachten prototypischen Eigenschaften) und *Oma* in B zu B' (Mensch eines spezifischen Verwandtschaftsverhältnisses V konkrete Person). Die zunächst unterspezifizierte Relation R wird, ebenfalls durch den Äußerungskontext und das Weltwissen, zur Relation HAVE spezifiziert. Oder anders ausgedrückt: Die Konstituente A erhält eine konzeptuelle, nicht referierende Modifikatorlesart, Konstituente B hingegen eine referenzielle, sodass eine abstrakte Bedeutung entsteht: 'X_1X_2 ist X_2, mit prototypischen Eigenschaften von X_1'. Diese Analyse beruht auf der Adjektiv-Analyse der Prot-ICC-Erstglieder von Bross und Fraser (2020), die in 4.2.3 näher erläutert wurde.

Name-ICCs scheinen als exozentrische Komposita zunächst Bell und Schäfers Beispiel *buttercup* sehr ähnlich zu sein. Bei dem Name-ICC *AutoAuto* reicht es allerdings nicht aus, die lexikalische Semantik der Konstituente B nur zu B' zu verschieben. Die Kopfkonstituente weist nicht die Semantik des zugrundeliegenden Lexems *Auto* auf. Die exozentrische Referenz kann überdies nicht sinnvoll über eine metonymische Verschiebung der zweiten Konstituente oder des Gesamtkompositums erklärt werden. Anders als bei *buttercup*, wo zunächst die Konstituenten verschoben werden (Buttertasse V Gelbblüte) und in einem zweiten Schritt metonymisch die Bedeutung zur Bezeichnung der gesamten Pflanze verschoben wird (Gelbblüte V Pflanze), liegt in Name-ICCs kein Konzept vor, das die Basis und der Referent teilen. *Auto* kann nicht über metaphorische oder metonymische Verschiebungen zu *Show* verschoben werden. Die Ausgangsbedeutung der Konstituenten muss deshalb nahezu (bei teildeskriptiven) oder vollständig (bei arbiträren) gelöscht werden. Auch die Relation R muss so spezifiziert werden, dass sie die reine Addition des Materials in A' und B' darstellt. In der Folge bleibt eine reduplikative Segmentabfolge übrig, die weitestgehend arbiträr mit einer Entität verbunden ist.

Das Modell von Bell und Schäfer bietet gegenüber den zuvor vorgestellten Modellen Vorteile. Durch die Einbindung von Kontext- und Weltwissen sowie die Annahme metaphorischer und metonymischer Verschiebung kann es die Bedeutung von ICCs sehr viel angemessener modellieren als die Modelle, die allein auf Informationen zu den Basislexemen beruhen. Doch auch in diesem Modell lassen sich allein die Det-ICCs angemessen hinsichtlich ihrer Bedeutungskonstitution beschreiben. Bei den Prot-ICCs ergeben sich bei beiden beschriebenen Wegen, die Prot-ICCs nach diesem Modell zu analysieren, Probleme. Es herrscht in der Forschung zu Prot-ICCs keine Einigkeit darüber, ob die Bildungen durch eine Relation R und die zum Kompositum gehörenden Konstituenten repräsentiert sind. Auch Beschreibungsansätze, die ICCs das Vorhandensein eines Modifikationsverhältnisses zubilligen, gehen meist davon aus, dass bei einem Prot-ICC keine der in Komposita sonst auftretenden semantischen Relationen vorliegt, sondern eine ganz eigene (Freywald 2015: 923, Hohenhaus 2004: 301). Dieser zweite Weg, die Bedeutungskonstitution von Prot-ICCs zu beschreiben, ist trotzdem von allen bisher vorgestellten Modellierungsversuchen am adäquatesten.

Allerdings müsste man das Modell insofern anpassen, als die semantische Relation in Prot-ICCs nicht vom Kontext und vom Weltwissen der Sprecher:innen spezifiziert wird. Einer solchen Anreicherung bedarf es gar nicht, denn durch die semantische Verschiebung der ersten Konstituente hin zur reinen Eigenschaftszuschreibung wäre R automatisch auf die Relation HAVE fixiert. In ähnlicher Weise ist es nicht notwendig, die Relation R im Kompositum *buttercup* mithilfe des Kon-

textes zu spezifizieren. Wenn, wie Bell und Schäfer herausstellen, *butter* zu 'having the colour of butter' verschoben wird, ist die Relation R bereits in dieser Verschiebung enthalten und also als HAVE konkretisiert. Wie bei A+N-Komposita wäre keine andere Relation möglich.

Es stellt sich aber darüber hinaus die Frage, wie das Modell den Unterschied zwischen Det-ICCs, Prot-ICCs und Name-ICCs darstellen kann. Nach Anwendung des Modells stammen alle Informationen zu den unterschiedlichen Relationen der drei ICC-Typen sowie zu den Verschiebungen der Konstituenten aus dem Weltwissen und dem Kontext. Die im Rahmen der vorgestellten Korpusstudien nachgewiesenen formalen Mittel (das Ausbleiben der Fugenelemente in Name-ICCs, die Markierung der Konstituenten bei Det-ICCs, die graphische Trennung der Konstituenten in der Schreibung) werden nicht berücksichtigt. Zudem sind Name-ICCs von der lexikalischen Semantik der Basis vollständig (*PunktPunkt*) oder weitgehend (*AutoAuto*) entkoppelt. Das Modell muss also entweder metaphorische oder metonymische Verschiebungen annehmen, die die Ausgangsbedeutung der Konstituenten sowie die Relation R vollständig löscht, sodass exozentrische Referenz möglich ist. Oder aber auch dieses Modell ist auf Komposita unterschiedlicher Konstituenten beschränkt.[77]

Ein alternativer Weg, die Bedeutung von ICCs zu modellieren, ist die Annahme eines Schemas oder Musters, das die Eigenschaften der komplexen Wörter holistisch spezifiziert (Booij 2010b). Holistisch deshalb, weil viele Bedeutungskomponenten von ICCs, bei Prot-ICCs etwa die Komponente „having the prototypical properties of" (Booij & Audring 2017), in den jeweiligen Ansätzen nicht von den Konstituenten abgeleitet werden, sondern Eigenschaften eines semantischen Schemas sind. Wie schon das CARIN-Modell und das Modell von Bell und Schäfer greift auch der musterbasierte Ansatz auf Aspekte der Wortfrequenz zurück. Der Grundgedanke ist, dass „people rely on statistical knowledge about how nouns tend to be used in combination in order to facilitate the interpretation of novel compounds" (Maguire et al. 2010: 50). Im Gegensatz zu schemabasierten Theorien lehnt der musterbasierte Ansatz die Idee ab, dass ein vollständiges konzeptuelles Schema aktiviert wird, wenn ein Nomen in Kopfposition auftritt. Stattdessen führt selektive Aktivierung von Konzeptwissen, das mit den Konstituenten verbunden ist, dazu, dass regelhafte Muster in Komposita erkannt werden. Die Nomina, die an den Komposita beteiligt sind, unterteilen Maguire et al. (2010) zunächst in semantische Kategorien, etwa PHENOMENON, EMOTION oder ARTIFACT. Basierend auf diesen Kategorien ergeben sich semantische Interpretationsmuster, die häufig

77 Der letztgenannte Punkt trifft ja auch auf Eigennamen im Allgemeinen zu. Das Modell wäre also zusätzlich nur auf Appellativa anwendbar.

vorkommen. Dadurch wird ein Kompositum wie beispielsweise *mountain bird* korrekt als 'a bird located in the mountains' interpretiert, ohne dass man weiß, was ein *mountain bird* ist (Maguire et al. 2010: 65f.). Die Kombination von Kopf und Modifikator geschieht also demnach nicht willkürlich, sondern nach Mustern, die überzufällig häufig auftreten. Ein sehr häufiges Muster ist beispielsweise das Muster „x^{ARTIFACT} is made of $y^{\text{SUBSTANCE}}$" (Maguire et al. 2010: 58), das auch vielen verwendeten Komposita zugrunde liegt, die in der vorliegenden Arbeit besprochen wurden (*Schneeball*, *Eischnee*, *Apfelsaft*, *stone lion*, *Lederball*, *Städtestadt*, *B-Seiten-Album*).

Der Grundgedanke des musterbasierten Ansatzes von Maguire et al. (2010) liegt auch anderen Ansätzen zugrunde. Das Modell von Zürn (2016) hat etwa auch zur Grundannahme, dass die Interpretation von Komposita mit dem Konzeptwissen zusammenhängt, das mit den Konstituenten verbunden ist. Maguire und Kollegen stützen diesen Mechanismus aber auf Frequenzdaten und somit auf das mentale Lexikon der Sprecher:innen. Die psycholinguistische Literatur zur Kompositaintegration spricht dafür, dass diese Verbindung zum mentalen Lexikon für ein angemessenes Modell der Bedeutungskonstitution von Komposita unerlässlich ist. Wie schon der Ansatz von Bell und Schäfer vernachlässigt aber auch der musterbasierte Ansatz von Maguire und Kollegen die formalen Mittel, die bei der Interpretation von Komposita im Deutschen eine Rolle spielen.

Man kann also festhalten, dass sich die Bedeutungskonstitution von ICCs mit den Begriffen Lexikalisierung, semantische Transparenz und Kompositionalität nicht hinreichend beschreiben lässt. Die in Kapitel 4 beschriebenen ICC-Typen lassen sich zwar auf den jeweiligen Begriffskontinua einordnen, doch sagt diese Einordnung noch nicht allzu viel darüber aus, wie sich die jeweiligen Bildungen erschließen lassen. Modelle semantischer Transparenz und Kompositionalität, die die Erschließbarkeit von Komposita zu modellieren und zu quantifizieren versuchen, helfen zwar dabei, die Besonderheiten von ICCs hervorzuheben. Die beschriebenen Ansätze sind aber meist nicht oder nur sehr eingeschränkt auf ICCs anwendbar. Die Transparenz von ICCs ist nur schwierig zu modellieren und Modelle, die die semantische Transparenz aus Informationen zu den Konstituenten errechnen, sagen stets eine unwahrscheinlich hohe Erschließbarkeit der Bildungen voraus. Modelle, die die im Lexikon gespeicherten semantischen Relationen der Konstituentenfamilien zugrunde legen, setzen hingegen voraus, dass ICCs eine solche semantische Relation besitzen – in Bezug auf Prot- und Name-ICCs eine problematische Annahme. Das Modell von Bell und Schäfer (2013, 2016) schließlich hat das Problem, dass vor allem der Kontext die Bedeutung bestimmt. Name-ICCs als exozentrische Bildungen müssen nach dem Modell zudem die Semantik ihrer Konstituenten vollkommen löschen, um ihre Referenzfunktion ausüben zu kön-

nen. Auch die Relation R muss in dem Fall gelöscht werden. Zudem werden formale Aspekte weitgehend ignoriert.

Die Annahme, dass die Eigenschaften komplexer Wörter holistisch spezifiziert werden, die Eigenschaften also nicht allein von den Konstituenten abgeleitet werden, sondern Eigenschaften eines Schemas sind, etwa der Ansatz von Maguire et al. (2010), vermeiden viele Unzulänglichkeiten der rein schemabasierten und relationsbasierten Modelle. Für die Beschreibung von ICCs sind solche Ansätze aber dennoch nicht optimal, da sie die formalen Mittel, die die Korpusstudien für ICCs nachweisen konnten, nicht berücksichtigen. Im folgenden Kapitel 10 wird nun der Versuch unternommen, alle Aspekte der vorgestellten Modelle, die sich für die Beschreibung von ICCs eignen und den Ergebnissen aus den Korpusstudien entsprechen, zu kombinieren. Ein Ansatz zur Beschreibung von ICCs darf – das ist ein Ergebnis dieses Kapitels – nicht rein mechanisch Eigenschaften der Konstituenten erfassen, die einen numerischen Wert der Erschließbarkeit ergeben. Bereits die Identität der Konstituenten ist für solche Ansätze ein Problem. Stattdessen muss ein ICC-Beschreibungsansatz die unterschiedlichen Funktionen, die den Konstituenten in Modifiziererposition und in Kopfposition zukommen, berücksichtigen. Außerdem muss der Ansatz sowohl auf Ad hoc-Bildungen als auch auf lexikalisierte ICCs anwendbar sein. Der Ansatz muss zudem eine holistische, musterbasierte Beschreibung sein, die mit den Ergebnissen der psycholinguistischen Forschung vereinbar ist und deshalb das mentale Lexikon der Sprecher:innen einbezieht. Der Ansatz muss den Äußerungskontext und das Weltwissen der Sprecher:innen berücksichtigen, aber auch die formalen Merkmale der Konstruktionen. Die Bedeutung und die Form von ICCs wird deshalb im folgenden Kapitel mithilfe eines konstruktionsgrammatischen Modells beschrieben, das all diesen Ansprüchen an einen ICC-Beschreibungsansatz gerecht wird.

10 ICCs und *Relational Morphology*

In den vergangenen Kapiteln konnten nicht alle Aspekte der Bildung und Bedeutung von ICCs erklärt werden. Es kann als gesichert gelten, dass die formalen Unterschiede zwischen den ICC-Typen bei ihrer Bedeutungskonstitution eine Rolle spielen. Doch reichen die Mittel nicht aus, drei ICC-Typen zu unterscheiden. Dass aber allein der Kontext für die richtige Interpretation eines ICCs sorgt, dagegen sprechen die bisherigen empirischen Studien zu ICCs (Finkbeiner 2014, Günther 1979). Die recht komplementäre Verteilung der Basislexeme auf die drei ICC-Typen, die die Korpusdaten nahelegen, deutet zwar darauf hin, dass ICC-Bildungen blockiert werden, wenn zu dem entsprechenden Basislexem bereits ICCs eines anderen Typs existieren. Es stellt sich aber die Frage, was diese Blockierung bedingt. Da ICCs nur in seltenen Fällen lexikalisiert sind, können es nicht die Lexikoneinträge der jeweiligen ICCs sein, die die Bildung von ICCs eines anderen Typs mit demselben Stamm hemmen. Es muss also weitere Mechanismen der Bedeutungskonstitution geben, die in Bezug auf ICCs bisher nicht beschrieben worden sind. Eine These dieser Arbeit ist, dass diese Mechanismen das Lexikon betreffen. ICCs sind nur mit Rückgriff auf das mentale Lexikon der Sprecher:innen adäquat zu beschreiben. Dieses Lexikon darf aber nicht als Sammlung von Idiosynkratischem verstanden werden, das lediglich den Input für ein generatives System liefert, sondern muss selbst als generatives System angesehen werden.

Der nun folgende Versuch, die Bedeutungskonstitution von ICCs zu fassen, geschieht im theoretischen Rahmen der *Relational Morphology* von Jackendoff und Audring (2020a, b), die auf der *Parallel Architecture* von Jackendoff (2002, 2010b) basiert. Beides sind konstruktionsgrammatische Ansätze, unterscheiden sich aber in entscheidenden Punkten von der *Konstruktionsgrammatik* (*CxG*) wie sie etwa Croft (2001) oder Goldberg (1995) vertreten, und von der *Construction Morphology* (*CxM*) von Booij (2010b). Die *Parallel Architecture* und die *Relational Morphology* können die Form und Funktion von ICCs nicht nur theoretisch erfassen. Die im empirischen Teil der Arbeit beschriebenen Daten können darüber hinaus als Evidenz für diese beiden Ansätze verstanden werden. Dieses Kapitel stellt zunächst die *Parallel Architecture* (10.1) und die *Relational Morphology* (10.2) in Grundzügen vor. Am Anschluss gehe ich darauf ein, wie Bedeutungskonstitution, semantische Transparenz und Kompositionalität in der RM modelliert werden (10.3) und beschreibe dann ICCs detailliert im Rahmen der *Relational Morphology* (10.4).

https://doi.org/10.1515/9783111315416-010

10.1 Jackendoffs *Parallel Architecture* (PA)

Jackendoffs *Parallel Architecture* (PA, Jackendoff 2002, 2010a, b) unterscheidet sich zum einen, wie andere beschränkungsbasierte Ansätze (CxG, CxM, HPSG), von generativen Ansätzen in der Frage, welchen Status grammatische Regeln einnehmen. Wie die Ansätze der CxG im Allgemeinen verortet auch die PA Regeln der Grammatik in Form von Schemata in einem erweiterten Lexikon, zusammen mit Wörtern, Mehrwortausdrücken wie Idiomen und Kollokationen und syntaktischen Konstruktionen. Die Regeln befinden sich nach der PA also nicht wie in generativen Modellen in Opposition zu den Lexikoneinträgen. Sie sind stattdessen selbst Lexikoneinträge.

Zum anderen geht die PA davon aus, dass bei der Bildung sprachlicher Strukturen kein unidirektionaler Prozess aus einem bestimmten Input einen bestimmten Output generiert, und somit eine sprachliche Form nicht die Ableitung einer anderen ist. Jackendoff nimmt zwar auch unterschiedliche, eigenständige Module des Sprachsystems an. Die Module Phonologie, Syntax und Semantik sind aber Struktursysteme, die nicht im Sinne von Ableitungen zueinander in Verbindung stehen, sondern über Verbindungsregeln. Jedes Modul hat sein eigenes Set aus generativen Regeln. Diese Systeme sind zwar unabhängig voneinander, aber über Schnittstellen miteinander verbunden. Das Schema der PA ist in Abbildung 41 dargestellt:

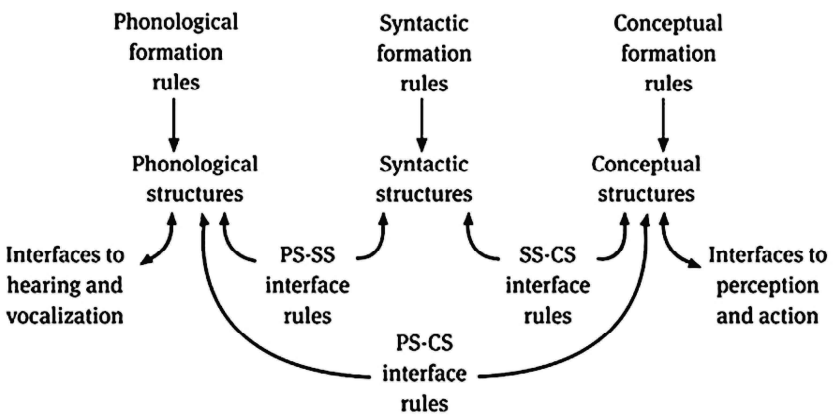

Abb. 41: *Parallel Architecture* nach Jackendoff (2002).

Alle drei Struktursysteme, also syntaktische Struktur, phonologische Struktur und semantisch-konzeptuelle Struktur, haben an lexikalischen Einheiten Anteil. Jedes

Lexem lässt sich also in diese drei Strukturebenen zerlegen. Die Schnittstellen zwischen diesen Strukturebenen werden im Formalismus über Ko-Indexierung angezeigt.[78]

Die drei Systeme haben wiederum miteinander in Verbindung stehende Subsysteme. Die phonologische Struktur besteht aus segmental-syllabischer Struktur, metrischer Struktur, Intonationsstruktur und, in Tonsprachen, Tonstruktur (Jackendoff 2010b: 2). Nur in Bezug auf die Substrukturen, die füreinander relevant sind, stehen syntaktische, phonologische und semantisch-konzeptuelle Struktur über Schnittstellen („interface links“) miteinander in Verbindung. So sind etwa Wortart und Genus für die syntaktische Struktur relevant, die phonologische, segmentale Strukturierung eines Wortes hingegen ist für die syntaktische Struktur nicht relevant.

Die wichtigste Prämisse der PA ist, dass sprachliche Strukturen, anders als in generativen Ansätzen, nicht nur durch die Syntax generiert werden, sondern durch Strukturregeln in den drei Systemen Semantik, Morphosyntax und Phonologie. Dadurch entfallen die in generativen Ansätzen notwendigen unsichtbaren syntaktischen Tiefenstrukturen. Die Syntax übernimmt in der PA mithilfe der Schnittstellenverbindungen lediglich eine Art Mittlerrolle zwischen der phonologischen und der semantisch-konzeptuellen Struktur. Doch übernimmt sie diese Aufgabe nicht notwendigerweise, da auch eine direkte Schnittstellenverbindung zwischen phonologischer und semantisch-konzeptueller Struktur besteht.

Die Syntax stellt in der PA dennoch etwas Besonderes dar, denn sie kann insofern als Kerngrammatik angesehen werden, als sie keine Verbindung zu anderen kognitiven Systemen aufweist. Das phonologische und das semantisch-konzeptuelle System sind hingegen in den Geist, die Sinne und die kognitiven Fähigkeiten des Menschen eingebettet (Abbildung 42).

78 Die Notation wird hier nicht extra behandelt, da in dieser Arbeit die Notation der *Relational Morphology* (Jackendoff & Audring 2020a, b) verwendet wird.

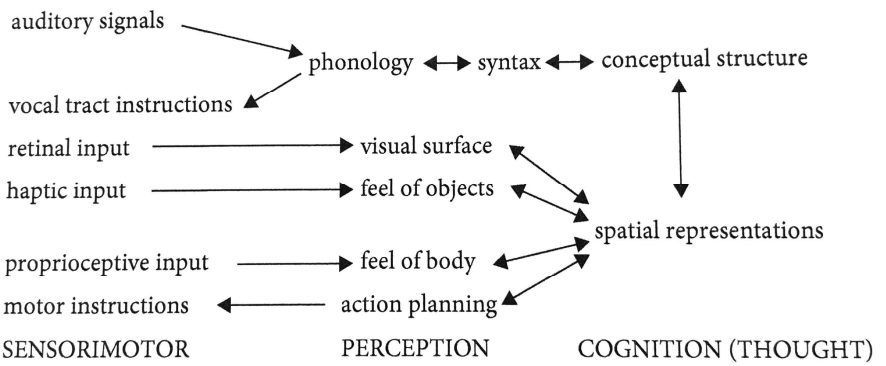

Abb. 42: Einbettung der PA in das perzeptuell-kognitive System des Menschen (Jackendoff & Audring 2020b: 8).

Das phonologische Modul ist über Schnittstellenverbindungen mit dem auditiven Teil des kognitiven Systems verbunden, die semantisch-konzeptuelle Struktur mit räumlichen Strukturen, wozu etwa Gestalt, Bewegung und Haptik gehören. Die Syntax hingegen hat eine rein grammatische, sprachinterne Funktion und interagiert nicht mit anderen kognitiven Fähigkeiten des Menschen.

Auch die Semantik ist im Rahmen der *Parallel Architecture* Teil des generativen Systems. Sie ist von Syntax und Phonologie unabhängig, aber mit ihnen über Schnittstellen verbunden. Jackendoff unterscheidet bei der Semantik nicht zwischen einer grammatischen Semantik einerseits und Welt- und Kontextwissen andererseits. Auch kennt die PA keine Teilung in semantische und pragmatische Bedeutung. Diese verschiedenen Aspekte der Bedeutung sind hingegen Teil der Subsysteme innerhalb der semantisch-konzeptuellen Struktur.

Ein Unterschied zwischen PA und anderen konstruktionsgrammatischen Ansätzen ist, dass nicht jede Konstruktion notwendigerweise bedeutungstragend ist. Transitive VPs im Englischen etwa stellen ein deklaratives Muster dar: [$_{VP}$ V – NP] ‘eine VP kann aus einem Verb und einer folgenden NP bestehen’. Dieses Muster ist aber lediglich eine Blaupause, bleibt rein syntaktisch, hat keine phonologische Struktur und geht auch nicht mit einer abstrakten Bedeutung einher. Jackendoff nennt diese Konstruktionen, die nur syntaktische, aber keine semantisch-konzeptuelle oder phonologische Struktur haben, „double defective constructions“ (Jackendoff 2002: 180). Obschon diese Konstruktionen rein syntaktisch sind, sind sie lexikalische Einheiten. Somit ist in der PA die strikte Trennung zwischen Grammatik und Lexikon aufgehoben.

Einheiten der lexikalisch spezifischeren Ebenen, etwa die Phrase *cut the grass*, erben von dieser Blaupause lediglich die Form. [$_{VP}$ V – NP] ist damit zwar

ein Lexikoneintrag, denn die Sprecher:innen müssen über das Wissen verfügen, wie Verben und Nominalphrasen im Englischen gemeinsam eine Verbalphrase bilden. Das Muster ist aber kein Saussure'sches Zeichen (Jackendoff 2002: 179ff). Jackendoff nennt solche Lexikoneinträge „pieces of stored structure" (Jackendoff 2010b: 19). Im Gegensatz zu anderen Ansätzen der Konstruktionsgrammatik, etwa dem nach Goldberg (1995), sind nach der PA nur solche abgespeicherten Strukturen Konstruktionen, die idiomatisch, nicht-kompositional sind, etwa *kick the bucket* 'den Eimer treten = verrecken, krepieren'.

Die Konstruktion ICC ordnet sich in die Hierarchie der Konstruktionsschemata ein und ist in der PA ein Subschema der N+N-Komposition [N+N], die wiederum Subschema der nominalen Komposition ist [X+N], die wiederum ein Subschema des allgemeinen Konstruktionsschemas der Komposita darstellt [X+Y]. Viele Eigenschaften der ICCs lassen sich auf diese Weise als von hierarchisch höheren Schemata ererbt auffassen: Die Rechtsköpfigkeit erben Det- und Prot-ICCs etwa vom allgemeinen Konstruktionsschema der Komposita, die Konstruktion Det-ICC erbt zudem die Möglichkeit der Verfugung vom Subschema der N+N-Komposition [N+N].

Doch nicht alle Merkmale der höheren Ebenen werden vererbt. Eigenschaften der übergeordneten Schemata werden in den jeweiligen Subschemata mit den jeweils spezifischen Eigenschaften der ICC-Typen überschrieben. So weisen die phonologische, die morphosyntaktische und die semantische Struktur von allen ICC-Konstruktionen Spezifizierungen gegenüber dem übergeordneten Schema auf, insofern sie reduplikativ sind. Außerdem gelten die (semantische) Rechtsköpfigkeit und die Möglichkeit zur Verfugung nicht bei Name-ICCs, die Konzeptvereinigung hingegen gilt nicht für Prot-ICCs.

Darüber hinaus schlägt Jackendoff für die Komposition Funktionen vor, die die Beziehung zwischen den Komponenten generieren. Die Klassifikationsbedeutung der klassifikatorischen ICC-Typen Det- und Prot-ICC ist im Wortbildungsschema [N+N] als IS A SUBTYPE OF bereits enthalten. Auch die semantische Relation zwischen Erst- und Zweitglied ergibt sich durch das Hinzuziehen von Funktionen. Dazu nennt Jackendoff 13 Basisrelationen, die im Wesentlichen die sind, die in der Literatur zu Kompositasemantik (etwa bei Levi 1978) angesetzt werden. Allerdings sieht Jackendoff die Möglichkeit vor, dass weitere, weniger übliche Relationen generiert werden und die lexikalische Bedeutung der Bildungen zusätzlich durch die sogenannte *Action Modality* spezifiziert werden, die angibt, ob das Konzept den Aktivitäten, den Funktionen, den Fähigkeiten et cetera zuzuordnen ist (Busa 1997, Jackendoff 2009). Dadurch sind bestimmte Bedeutungskomponenten, beispielsweise der Aspekt *Ziel/Zweck* nicht durch eine eigene semantische Relation gekennzeichnet, sondern als *Action Modality* Teil der lexikalischen Bedeutung der

Konstituenten. Auch die Funktion *F* kann in Jackendoffs Modell mithilfe der internen semantischen Struktur von N_1 und N_2 hergeleitet werden. Die genannten Aspekte des Modells ermöglichen, praktisch alle möglichen Komposita zu modellieren.

In jüngster Vergangenheit wurde die *Parallel Architecture* weiterentwickelt und hinsichtlich des mentalen Lexikons und der Morphologie optimiert, was in einem neuen *Relational Morphology* genannten Ansatz mündet. Durch den Fokus auf die Morphologie und ihre Interaktion mit dem mentalen Lexikon ist dieser Ansatz optimal zur Modellierung von Wortbildungsphänomenen geeignet. Die konstruktionsgrammatische Modellierung der ICCs nehme ich daher im Rahmen der *Relational Morphology* vor.

10.2 *Relational Morphology* (RM)

Relational Morphology (RM, Jackendoff & Audring 2020a; 2020b) ist ein auf der *Parallel Architecture* basierender Ansatz zur Beschreibung von sprachlichen Strukturen. Der Ansatz steht ebenfalls der Konstruktionsgrammatik (CxG) nahe und ergänzt die *Parallel Architecture* von Jackendoff (2002, 2010b) um die zentrale Frage, welche linguistischen Einheiten im Langzeitgedächtnis abgespeichert werden und in welcher Form. Dabei erweitern Jackendoff und Audring mit ihrem Ansatz CxG und PA in vielen Punkten und fokussieren die Annahme, dass den Schemata zusätzlich zur generativen Funktion noch die namensgebende relationale Funktion zukommt.

Wie schon in der PA gibt es auch in der RM keine Grenze zwischen Lexikon und Grammatik. In der RM wird zusätzlich betont, dass es auch zwischen Kern- und Peripheriephänomenen keine klare Grenze gibt. Jackendoffs und Audrings Modell bietet dabei wie die PA die Möglichkeit, rein funktionale Schemata, die keine lexikalische Bedeutung tragen (etwa die Fugenelemente im Deutschen), in die Beschreibung von Schemata zu integrieren. Die RM fasst die Verbindungen zwischen lexikalischen Einträgen zudem präziser als das der CxG eigene Konzept der *Inheritance* und erlaubt horizontale Relationen, also ein direktes Verhältnis zwischen Wörtern und Schemata ohne die Annahme eines Mutterschemas. Viele dieser Punkte sind für die Beschreibung von ICCs relevant und werden daher im Folgenden genauer beschrieben. Hierzu wird auch der Formalismus beschrieben, mit dem Jackendoff und Audring die Strukturen notieren.

Das Lexikon in RM ist, wie schon in der PA, nicht als eigene Komponente repräsentiert. Stattdessen ist das Lexikon das einzige Modul überhaupt. Es besteht aus den drei Komponenten Phonologie, Syntax und Semantik. Im Lexikon existieren Korrespondenzen zwischen den drei Komponenten, sodass ein Set aus seman-

tischer, morphosyntaktischer und phonologischer Struktur einen Lexikoneintrag darstellt. Das Nomen *Katze* in (130) besteht beispielsweise aus semantischer Struktur, also aus der Wortbedeutung (KATZE), aus phonologischer Struktur (/kat͡sə/) und aus der syntaktischen Kategorie Nomen (N). Dass diese Strukturen verbunden sind und einen Lexikoneintrag bilden, wird über tiefergestellte Indizes dargestellt. Diese Indizes ersetzen, der besseren Übersicht halber, Assoziationslinien zwischen den Komponenten und markieren die Schnittstellenverbindungen.

(130) Semantik: KATZE$_1$
Morphosyntax: N$_1$
Phonologie: /kat͡sə/$_1$

Die drei Strukturebenen agieren sowohl auf phrasaler Ebene als auch auf Wortebene und umfassen dort die entsprechenden Subsysteme. Die Morphosyntax umfasst beispielsweise Wortart, Tempus, Person, KNG, Flexionsklasse und wortinterne Konstituenten wie etwa Wurzeln, Stämme und Affixe. Sie referiert also auf X_0-Kategorien wie N, V, und A. Die phrasale Syntax bildet größere Einheiten (NP, VP, AP). Die Wortphonologie umfasst Phonotaktik, Wortakzent und Vokalharmonie. Die phrasale Phonologie beinhaltet Sandhi, Phrasenakzent und Intonationskurven. Lexikalische Semantik umfasst die Bedeutung von Wörtern, Idiomen und Schemata. Phrasale Semantik hingegen Argumentstruktur, Fokus und Skopus der Quantifikation.

Diese Strukturen werden in der RM aber nicht einfach nur beschrieben. Sie können darüber hinaus als Schema für andere Bildungen dienen und so generativ neue Strukturen hervorbringen. Die Schemata beinhalten Variablen, die diejenigen Strukturbestandteile von Lexikoneinträgen repräsentieren, die unterschiedlich sind. Werden diese Variablen mit konkretem sprachlichem Material gefüllt, entsteht ein neuer Lexikoneintrag. Dieser generative Mechanismus in der RM wird mithilfe des Begriffs *Unification* beschrieben (Shieber 1986). Die Unification ist die einzige prozedurale Regel, die RM annimmt. Unification ist eine boolische Strukturvereinigung. Im Zuge der Unification bleiben also alle Bestandteile der Ausgangsstrukturen erhalten, ohne die von ihnen geteilten Bestandteile zu verdoppeln. Die Unification der Strukturen ABCD und CDEF resultiert beispielsweise nicht, wie etwa bei der Funktion *Merge* im *minimalist programm*, in der Struktur ABCDCDEF, sondern in der Struktur ABCDEF.

Unification passiert, wie schon in der PA, indem Variablen in Schemata eingefügt werden, wodurch die Schemata eine unbegrenzte Menge an Strukturen ermöglichen. Beispiel (131) zeigt ein Pluralschema des Deutschen, das wie beim lexikalischen Eintrag *Katze* in (130) aus semantischer, morphosyntaktischer und

phonologischer Struktur besteht. Auch in diesem Schema sind die drei Strukturbestandteile über tiefergestellte Indizes verbunden. Die Notation zu diesem Schema und der Pluralform *Katzen* sieht wie folgt aus:

(131)		a. Pluralschema	b. *Katzen*
	Semantik:	$[\text{PLUR}\ (X_x)]_y$	$[\text{PLUR}\ (\text{KATZE}_1)]_3$
	Morphosyntax:	$[N_x\ PL_2]_y$	$[N_1\ PL_2]_3$
	Phonologie:	$/\ldots_x\ n_2/_y$	$/\widehat{\text{kaʦ}}\text{ə}_1\ n_2/_3$

Die Unification ersetzt die Variablen eines Schemas (131a) mit phonologischem, morphosyntaktischem oder semantischem Material. Durch diese Exemplifizierung der Variablen erhält das Schema seine generative Funktion und führt mithilfe von (130) zur Wortform in (131b). Schemata und Unification spielen also die Rolle, die traditionell die prozeduralen Regeln übernehmen. Der deklarative Formalismus der RM kann deshalb eine prozedurale Interpretation bekommen: Die generative Eigenschaft der Sprache basiert auf dem System der lexikalen Verbindungen (Jackendoff & Audring 2020b: 53). Durch diesen Mechanismus teilt RM mit der PA den Vorteil, dass Schemata, und damit die grammatischen Regeln, dasselbe Format wie Wörter haben.

Der einzige Unterschied zwischen Wörtern und Schemata besteht darin, dass bei Schemata Teile der Struktur aus Variablen bestehen. Die semantischen und morphosyntaktischen Strukturen im Pluralschema in (131a) besagen beispielsweise, dass eine Mehrzahl (PLUR) jeglicher Art von Entität (X) durch ein Nomen (N) und ein Pluralaffix (PL) ausgedrückt werden kann. Auf der Ebene der phonologischen Struktur steht die Variable … für einen beliebigen Lautkörper gefolgt vom Phonem /n/. Die Variable … ist über die Indizes $_x$ und $_y$ mit den Einheiten der morphosyntaktischen und semantischen Ebene verbunden. Das Pluralschema von *Katze* kann gemäß diesem Schema produziert werden, wenn die Variablen aus (131a) als Ausprägungen die den Indizes entsprechenden Teile im Lexikoneintrag *Katze* aus (130) erhalten. Das Ergebnis ist in (131b) dargestellt. Da es sich in (131b) um einen voll spezifizierten Lexikoneintrag handelt, wird der tiefergestellte Index $_y$ durch die numerische Konstante $_3$ ersetzt.

Die wichtigste Neuerung, die RM in die Diskussion konstruktionsgrammatischer Ansätze einbringt, ist die Annahme, dass den Schemata zusätzlich zur generativen Funktion eine relationale Funktion zukommt. Nur manche Lexikoneinträge, etwa das Pluralschema in (131a), wirken generativ, aber alle Lexikoneinträge wirken relational („Relationale Hypothese“, Jackendoff & Audring 2020a: 5). Über relationale Verbindungen stehen Wörter mit anderen Wörtern und auch mit Schemata in Verbindung. Beispielsweise ist die Form *Katzen* in (131b) einerseits

auf allen Ebenen über den Index $_1$ mit ihrer Basis *Katze* in (130) verbunden, andererseits über die Teile der morphosyntaktischen Struktur N und PL, das Phonem /n/ und den Index $_2$ mit dem Pluralschema in (131a). Wichtig ist, dass die Darstellung weder als ein Prozess mit einem Input (130) und einem Output (131b) anzusehen ist, noch das Schema (131a) als eine Abstraktion von (131b). Stattdessen sind sowohl das Wort in (130) als auch das Schema in (131a) und die Wortform in (131b) lexikalische Einträge, die im gleichen Format im Lexikon und miteinander in Verbindung stehen. Die mithilfe der Indizes notierten Verbindungen beschreiben, welche der Strukturbestandteile die lexikalischen Einträge teilen. Die Form *Katzen* in (131b) wird durch das Schema in (131a) unterstützt beziehungsweise motiviert (Jackendoff & Audring 2020a: 4).

Ein großer Vorteil von RM ist also, dass Beziehungen zwischen Wörtern und zwischen Schemata sowie zwischen Wörtern und Schemata dargestellt werden können. Die Konstruktionsgrammatik beschäftigt sich üblicherweise nicht mit diesen Horizontalverbindungen, sondern fokussiert sich auf Vererbungsrelationen zwischen Wörtern, Konstruktionen und abstrakteren Konstruktionen, also auf Vertikalverbindungen. Die RM stellt die Horizontalverbindungen zudem präzise dar, da sie nicht notwendigerweise alle Ebenen und Strukturbestandteile eines lexikalischen Eintrags umfassen, sondern nur die Bestandteile, die identisch sind. *Katze* in (130) und die später noch zu besprechende *Hauskatze* in (135b) etwa teilen nur die mit dem Index $_1$ versehenen Strukturbestandteile. Der Rest unterscheidet die beiden Nomina und hier werden keine relationalen Verbindungen markiert. Die Annahme von Horizontalverbindungen berücksichtigt auch, dass das Wissen der Sprecher:innen über ein Schema Einfluss auf die Verwendung eines anderen Schemas haben kann und ermöglicht, Wechselwirkungen zwischen Schemata darzustellen.

Die Annahme von Horizontalverbindungen und ihre Integration in das konstruktionsgrammatische Modell ist ein Versuch, den Unterschied zwischen Analogie und abstraktem Wortbildungsmuster graduell zu fassen. Dieses Ziel ist nicht neu. Booij (2010a, b: 88–93) beispielsweise beschreibt den fließenden Übergang zwischen Phänomenen rein idiosynkratischer Analogiebildungen, die auf einem ganz bestimmten Wort basieren, etwa *paniek-haas* ‘Panikhase’ basierend auf *angst-haas* ‘Angsthase’ im Niederländischen (Booij 2010b: 89), und abstrakten Mustern wie dem Schema semantisch rechtsköpfiger Komposita im Niederländischen (Booij 2010a: 3). Zwischen diesen Polen gibt es Abstufungen. Auch können sich zunächst auf Analogie basierende Wortbildungen mit zunehmender Verwendung in abstrakte Muster wandeln, beispielsweise das auf den Watergate-Skandal zurückgehende Muster in (132), das im Englischen wie im Niederländischen produktiv ist (Booij 2010b: 90, Übersetzung MF).

(132) $[[x]_{Ni}\ [gate]_{N}]_{Nj}$, ↔ [Politischer Skandal Sem_i betreffend]$_j$

Die RM nimmt nun an, dass sowohl einzelne Lexikoneinträge als auch abstrakte Schemata eine Wirkung auf das Lexikon und die Bildung von Wörtern haben. Jackendoff und Audring verwenden für die Beschreibung nur relational wirkender Lexikoneinträge allerdings nicht den Begriff der Analogie, sondern die in Kapitel 7 bereits behandelte „Same-Except-Relation" (Jackendoff & Audring 2020b: 76ff.). Die Same-Except-Relation ist eine sehr limitierte Form der Analogie. Jackendoff und Audring stellen zur Illustration die Wörter *ambition* und *ambitious* gegenüber. Die Wörter teilen sich Teile ihrer Struktur, unterscheiden sich aber in Bezug auf die Wortart und das Affix. Mithilfe der Same-Except-Relation können aber auch nicht-konkatenative Prozesse wie etwa Stammalternationen beschrieben werden. Den Plural des englischen Substantivs *goose* (*geese*) kann man beispielsweise mit der Same-Except-Relation gut modellieren, weil eine direkte Verbindung zwischen den beiden Wortformen erlaubt ist ohne die Annahme einer (abstrakten) zugrundeliegenden Form, von der *geese* abgeleitet ist. Das, was die beiden Formen unterscheidet, also eingeschränkte Übereinstimmung, wird in der Notation der RM mit Sternchen angezeigt (**, same-except-relation, Jackendoff & Audring 2020b: 76ff.).

(133)

	a. *goose* (stem)	b. *geese*
Semantik:	GOOSE$_4$	[PLUR (GOOSE$_4$)]$_5$
Morphosyntax:	N$_4$	[N$_4$ PL]$_5$
Phonologie:	/g *uw* s/$_4$	/g *i* s/$_{4,5}$

Hier unterscheidet sich die phonologische Form nur hinsichtlich der umsternten Segmente. Bei den Schemata wird mit Konstanten notiert, worin sich zwei sprachliche Elemente gleichen und mit Variablen, worin sie sich unterscheiden. Die Variablen legen auch fest, dass die Variablen einen definierten Ort haben.

Ein weiterer Vorzug von RM für die Beschreibung von Wortbildungsphänomenen hängt mit der Annahme relationaler Verbindungen zusammen. Wie in anderen konstruktionsgrammatischen Modellen, etwa der Constuction Morphology (Booij 2009: 7, 2010a), wird die Trennung von produktiven und nichtproduktiven Mustern relativiert (vgl. Pinker 1997, 1998). Abstrakte Schemata und individuelle Vorkommen koexistieren. Wörter, die regulär gebildet worden sind, können durchaus im Lexikon gelistet sein. Die Grenze zwischen produktiven und unproduktiven Mustern wird durchlässiger. In der RM drückt sich dieses graduelle Verständnis von Produktivität in der Annahme aus, dass alle lexikalischen Einheiten relational wirken, aber nur ein Subset davon auch generativ wirkt. Im

Gegensatz zu anderen konstruktionsgrammatischen Modellen modelliert die RM aber auch diesen Aspekt sehr präzise, denn die Theorie sieht keine ganze Muster umfassende Einteilung in Produktivitätsgrade vor, sondern definiert für die einzelnen Variablen eines Schemas eigene Produktivitätswerte. In der Notation wird dies mithilfe von Unterstreichungen angezeigt (134b):[79]

(134)		a. *cat*	b. Pluralschema	c. *cats*
	Semantik:	CAT_6	$[\text{PLUR}\ ([\underline{\text{INDIVIDUAL}}]_x)]_y$	$[\text{PLUR}\ ([\text{CAT}]_7)]_8$
	Morphosyntax:	N_6	$[\underline{N}_x\ \text{PL}_7]_y$	$[N_2\ \text{PL}_7]_8$
	Phonologie:	/kæt/$_6$	/$\underline{\ldots}_x$ s$_7$/$_y$	/kæt$_6$ s$_7$/$_8$

Doppelt unterstrichene Variablen sind offen und ihre Ausprägung unterliegt keinen Beschränkungen. Bei einfach unterstrichenen Variablen hingegen gibt es selektionale Restriktionen. Das Pluralschema im Englischen etwa ist beschränkt auf Individualsubstantive; das Schema in (134b) zeigt etwa, dass CAT mit der Variable INDIVIDUAL zusammengebracht werden kann, da Katzen inhärent unteilbar und klar konturiert sind.

Wie die Flexion lässt sich auch die Wortbildung über Schemata beschreiben. (135a) zeigt das Kompositionsschema (Jackendoff & Audring 2020b: 101).[80]

(135)		a. Kompositionsschema	b. *Hauskatze*
	Semantik:	$[\text{F}\ (X_x, Y_y)]_z$	$[\text{KATZE}_1, \text{IN}\ (\text{HAUS}_9)]_{10}$
	Morphosyntax:	$[_N\ N_x\ N_y]_z$	$[_N\ N_9\ N_1]_{10}$
	Phonologie:	/…$_x$ …$_y$ /$_z$	/haʊs$_9$ kat͡sə$_1$/$_{10}$

Das Schema besteht auf semantischer Ebene aus der Funktion (F) zwischen der Bedeutung einer Entität (X) und der Bedeutung einer anderen Entität (Y). Diese Variablen drücken aus, dass dies bei allen Komposita so ist und entspricht den Grundannahmen der in Kapitel 9 behandelten Modelle zur Kompositasemantik. Auf morphosyntaktischer Ebene wird ein nominales Kompositum ($_N$) aus zwei Nominalstämmen (N N) gebildet und auf phonologischer Ebene stehen Variablen

79 Auch das Pluralschema in (131) ist diesbezüglich zu konkretisieren.

80 Die semantisch-konzeptuelle Ebene von *Hauskatze* wird hier als [KATZE, IN (HAUS)] notiert, in Anlehnung an Jackendoff und Audrings (2020b: 101) Notation des niederländischen Kompositums *dorpskroeg* 'Dorfkneipe', das die Autor:innen mit [CAFÉ, IN (VILLAGE)] angeben. Alternativ könnte man solche voll spezifizierten Lexikoneinträge näher am Schema [F (X, Y)] orientieren und die semantisch-konzeptuelle Ebene von *Hauskatze* mit [IN (KATZE, HAUS)] angeben. Der Einheitlichkeit wegen wird die semantisch-konzeptuelle Ebene von Lexikoneinträgen aber im Einklang mit Jackendoff und Audrings Notation wiedergegeben.

(... ...) für beliebige Lautkörper, die über die Indizes $_x$ und $_y$ mit den Einheiten der morphosyntaktischen und semantischen Ebene verbunden sind. Fugenelemente gelten in der RM als phonologische Einheiten ohne Semantik. Sie sind morphosyntaktisch bedingt, haben aber weder eine morphosyntaktische Kategorie noch eine eigene Bedeutung. Sie erscheinen in der Notation lediglich auf phonologischer Ebene.

RM kann morphologische Prozesse sehr genau darstellen. Gleichzeitig kann man Ergebnisse der Psycho- und Neurolinguistik mit RM und die dort getroffenen Annahmen zu Horizontalverbindungen in Verbindung bringen. Die in 9.4 vorgestellten Ergebnisse von Gagné und Spalding (2009) zum Priming semantischer Relationen kann die RM mit ihrem Ansatz der geteilten Struktur gut erklären. Die Notation zu den drei Komposita sieht wie folgt aus:[81]

(136)		a. *snowfort*	b. *snowball*	c. *snowshovel*
	S:	[FORT$_{11}$; MAKE (SNOW$_{12}$)]$_{13}$	[BALL$_{14}$; MAKE (SNOW$_{12}$)]$_{15}$	[SHOVEL$_{16}$;FOR(SNOW$_{12}$)]$_{17}$
	MS:	[$_N$ N$_{12}$ N$_{11}$]$_{13}$	[$_N$ N$_{12}$ N$_{14}$]$_{15}$	[$_N$ N$_{12}$ N$_{16}$]$_{17}$
	P:	/snəʊ$_{12}$ fɔɹt$_{11}$/$_{13}$	/snəʊ$_{12}$ bɔːl$_{14}$/$_{15}$	/snoʊ$_{12}$ ʃʌvəl$_{16}$/$_{17}$

Auf phonologischer und morphosemantischer Ebene teilen *snowfort* und *snowshovel* gleich viel Struktur mit *snowball*. Auf semantisch-konzeptueller Ebene hingegen teilt *snowfort* zusätzlich die Funktion MAKE mit *snowball. snowball* und *snowfort* teilen also mehr Struktur und haben darum mehr relationale Verbindungen zueinander. So lässt sich erklären, weshalb bei *snowfort* das semantische Priming im Experiment von Gagné und Spalding (2009) stärker ist als bei *snowshovel*. Die lexikale Struktur von *snowfort* kann die Struktur von *snowball* besser voraktivieren als die von *snowshovel.*

Auch Ergebnisse zu anderen sprachlichen Ebenen aus der psycho- und neurolinguistischen Forschung können mit RM modelliert werden. So gibt es in den letzten Jahren Evidenz dafür, dass Primingeffekte stets von der Menge der geteilten Struktur abhängen, ungeachtet auf welcher Ebene Prime und Target Struktur teilen. Die Effekte unterscheiden sich jedoch, je nachdem welche Ebenen ähnliche Struktur aufweisen (Zwitserlood 2018). Auch diesen Forschungsergebnissen wird in der RM Rechnung getragen, indem Lexikoneinträge in semantische, morphosyntaktische und phonologische Struktur geschieden werden und Verbindungen zwischen Lexikoneinträgen auch hinsichtlich einzelner Elemente auf einer der Strukturen abgebildet werden können.

81 Aus Platzgründen wird in einigen der folgenden Schemadarstellungen „Semantik“ mit „S“ abgekürzt, „Morphosyntax“ mit „MS“ und „Phonologie“ mit „P“.

10.3 Erschließbarkeit und Produktivität in der RM

In der RM werden identische Strukturbestandteile über relationale Verbindungen markiert. Die RM bietet dadurch eine flexible Darstellung von lexikalischen Einträgen und ist somit auch ein Beschreibungsmodell für Erschließbarkeit, semantische Transparenz und Kompositionalität. Die RM kann dabei viele der in Kapitel 9 besprochenen Probleme bei der ICC-Analyse lösen.

Eines der drei Hauptprobleme der in Kapitel 9 behandelten Modelle ist, dass die unterschiedlichen Funktionen, die den Konstituenten in Modifiziererposition und in Kopfposition zukommen, nicht berücksichtigt werden. Deshalb ergeben sich unrealistisch hohe Transparenzwerte für ICCs in den Modellen von Zinsmeister (2013), Schulte im Walde / Borgwaldt (2015) und Zürn (2016). Das Problem des CARIN-Modells von Gagné und Shoben (1997) ist die Annahme, dass jedes N+N-Kompositum eine semantische Relation beinhaltet – eine Annahme, die bei Prot-ICCs und Name-ICCs problematisch ist. Im Modell von Bell und Schäfer (2016) schließlich geht letztlich die gesamte semantische Verschiebung innerhalb des Kompositums sowie die Qualität der Modifikationsrelation auf Kontext- und Weltwissen zurück. Alle vorgestellten Modelle sind zudem nicht sensibel gegenüber formalen Merkmalen wie etwa morphologische Marker oder die Schreibung. Die RM löst die genannten Probleme und eignet sich darum gut zur Modellierung von Form und Semantik von ICCs.

Die in Kapitel 9 vorgestellten Modelle der Kompositionalität nehmen an, dass Komposita dann kompositional sind, wenn sie viele Merkmale mit den Elementen teilen, von denen sie abgeleitet oder aus denen sie zusammengesetzt sind. Diese Grundannahme bedeutet in diesen Ansätzen auch, dass nur die Eigenschaften der Wörter spezifiziert werden, die ihnen nicht von den Elementen, aus denen sie bestehen, zukommen. Ein Wort wie *Bettsofa* ist darum kompositionaler als *Eischnee*, weil es sowohl von *Bett* als auch von *Sofa* Merkmale übernimmt. *Eischnee* wiederum ist kompositionaler als *Bildschirmbräune,* das eine sehr komplexe Beziehung zwischen den Konstituenten aufweist. Zudem ist *Eischnee* transparenter als *Schneebesen*. In Ersterem muss nur die zweite Konstituente der kompositionalen Bedeutung mit ‘Schaum’ überschrieben werden. In Letzterem sind beide Bestandteile nur über Umwege motiviert, da es sich weder um einen Besen handelt noch um ein Gerät, das für Schnee verwendet wird.

Viele Modelle verstehen Kompositionalität und Transparenz dabei als einen Aspekt der Datenkomprimierung, bei der im Lexikon nur die Bestandteile abgespeichert werden, die notwendigerweise gelistet werden müssen. Ist eine Bildung transparent und kompositional, kann sie (de-)konstruiert und aus den Bestandteilen generiert werden. Opake und nicht-kompositionale Bildungen dagegen müssen als Vollform abgespeichert werden, weil sie nicht aus den Bestandteilen zu-

sammengesetzt werden können. Diese Schlussfolgerung ist aber nicht sinnvoll, weil auch weitgehend kompositionale Formen wie etwa *Bettsofa* trotz der Redundanz der Bestandteile im Lexikon gelistet werden, sofern das Konzept hinter *Bettsofa* im Alltag der Sprecher:innen nur frequent genug ist. Wörter wie *Bettsofa*, die transparent und kompositional gebildet sind, können durchaus im Lexikon gelistet sein. Die Abspeicherung eines kompositionalen Eintrags stellt für das Gehirn keine Ineffizienz oder Energieverschwendung dar. Vielmehr fördert diese Redundanz die mentale Berechnung im Gehirn und macht sie robuster. Nur so lassen sich Frequenzeffekte bei der Kompositaverarbeitung erklären (etwa aus Sahel et al. 2008, van Jaarsveld & Rattink 1988). Das Problem der meisten Modelle semantischer Transparenz und Kompositionalität ist, dass sie diesen Aspekt, nämlich das Lexikon der Sprecher:innen, außer Acht lassen. Genau dieser Aspekt wird in der RM angemessen berücksichtigt.

In der RM wird Kompositionalität über relationale Verbindungen erklärt. *Bettsofa* (137b) hat mehr relationale Links, da es auf allen Ebenen mit *Bett* und *Sofa* (137a, c) in Verbindung steht. Im Formalismus zeigt sich das durch die Indizes $_{18}$ und $_{19}$, die für diese umfangreichen Verbindungen zwischen allen drei Strukturebenen stehen. *Schneebesen* (138b) hat hingegen weniger relationale Links, da die semantisch-konzeptuelle Ebene mit der von *Schnee* und *Besen* (138a, c) nicht verbunden ist und *Schneebesen* zudem mit *Schnee* und *Besen* nur Teile der morphosyntaktischen und phonologischen Struktur teilt. In der Notation sieht dieser Unterschied wie folgt aus:

(137)		a. *Bett*	b. *Bettsofa*	c. *Sofa*
	S:	BETT$_{18}$	[MÖBEL [SOFA$_{19}$ AND BETT$_{18}$]]$_{20}$	SOFA$_{19}$
	MS:	N$_{18}$	[$_{N}$ N$_{18}$ N$_{19}$]$_{20}$	N$_{19}$
	P:	/bɛt/$_{18}$	/bɛt$_{18}$ zoːfɑː$_{19}$/$_{20}$	/zoːfɑː/$_{19}$

(138)		a. *Schnee*	b. *Schneebesen*	c. *Besen*
	S:	SCHNEE$_{21}$	GERÄT FOR SCHAUMMASSE$_{22}$	BESEN$_{23}$
	MS:	N$_{21}$	[$_{N}$ N$_{21}$ N$_{23}$]$_{22}$	N$_{23}$
	P:	/ʃneː/$_{21}$	/ʃneː$_{21}$ beːzən$_{23}$/$_{22}$	/beːzən/$_{23}$

Messer schließlich hat gar keine relationalen Verbindungen, sondern nur Schnittstellenverbindungen, die die Strukturen zu einem Lexikoneintrag verbinden. *Messer* hat also eine Struktur analog zu *Katze* in (130).

Ferner treffen Jackendoff und Audring Annahmen dazu, wie Schemata und relationale Links erworben und produktiv angewendet werden (Jackendoff & Audring 2020b: 219ff.). Zunächst begegnen den Sprecher:innen Wörter, bei denen

sie auf (manchen der) drei Ebenen Ähnlichkeiten entdecken. Die wahrgenommenen Ähnlichkeiten führen zu relationalen Links und dazu, dass die Sprecher:innen in (mindestens zwei) Wörtern Schwestern erkennen. In der Folge leiten sie, nach einer gewissen Anzahl an Schwestern, ein Schema ab und stellen eine Hypothese dazu auf, ob dieses Schema produktiv ist. Jackendoff und Audring nennen mögliche Faktoren, die darüber entscheiden, ob die Sprecher:innen ein produktives Schema annehmen (Jackendoff & Audring 2020b: 229). Ein Schema wird demnach als produktiv angesehen, wenn

– eine hohe Anzahl kompositionaler Bildungen existiert,
– die Variablen weit gefasst sind,
– eine hohe Anzahl niedrigfrequenter Fälle existiert und
– eine niedrige Anzahl an Ausnahmen existiert.

Die in Teil II dieser Arbeit vorgestellten Korpusstudien liefern für die letzten drei genannten Faktoren relevante Daten und ermöglichen damit eine empirisch untermauerte Aussage zur Produktivität und Bedeutungskonstitution der einzelnen ICC-Schemata. Ich beschreibe nun zunächst die Bedeutung der RM für die ICCs und daraufhin die Bedeutung der ICCs (und der erhobenen Korpusdaten) für die RM.

10.4 ICCs in der *Relational Morphology*

Wenn Menschen zwei Objekte wahrnehmen, nehmen sie sowohl das wahr, was an diesen Objekten identisch ist, als auch das, was sie unterscheidet. Dieser Aspekt der Perzeption betrifft auch die Wahrnehmung sprachlichen Materials. RM richtet sich diesbezüglich explizit an den Ergebnissen der psycholinguistischen Forschung aus. Das Lexikon wird darum als reich texturiertes Netzwerk aufgefasst, in dem alle Einträge permanent auf gleiche und ungleiche Bestandteile hin abgeglichen werden. Die *Relational Morphology* bildet diesen kognitiven Prozess ab und notiert, worin sich Lexikoneinträge gleichen (im Formalismus durch Konstanten repräsentiert) und worin sie sich unterscheiden (im Formalismus durch Variablen repräsentiert). Die RM ist dadurch ein deklarativer Ansatz und eignet sich durch den simplen Mechanismus der Unification (Same-Except-Relation) optimal für die Beschreibung von ICCs. Anders als regelbasierte Ansätze müssen unterschiedliche Schemata nicht regelhaft und ausnahmslos unterschiedliche phonologische Struktur erzeugen. RM ermöglicht stattdessen Schemata, die zwar unterschiedliche semantische, aber weitgehend identische phonologische Struktur haben. Genau das ist bei ICCs der Fall.

Für die RM ist zudem die Identität der Konstituenten von ICCs kein Problem, da aus den Konstituenten keine Rückschlüsse auf die Erschließbarkeit der Bildungen gezogen werden. Auch berücksichtigt RM die unterschiedlichen Funktionen, die den Konstituenten in Modifizierer- und Kopfposition zukommen, und ist als deklarativer Ansatz, der eine generative Lesart erlaubt, auf Ad hoc-Bildungen und lexikalisierte ICCs anwendbar. Die sprachlichen Module sind ferner in den Geist, die Sinne und die kognitiven Fähigkeiten des Menschen eingebettet. RM ermöglicht so, den Äußerungskontext und das Weltwissen der Sprecher:innen zu berücksichtigen. Die zwei wichtigsten Vorteile, die die RM gegenüber anderen Modellen der Kompositasemantik begünstigen, sind aber die Berücksichtigung formaler Merkmale sowie die Verbindung zum mentalen Lexikon der Sprecher:innen.

Im Folgenden entwickele ich im Rahmen der RM Schemata, die die Funktionsweise von ICCs abbilden. Zunächst beschreibe ich die semantisch-konzeptuelle Ebene der ICC-Typen auf der Grundlage der Ergebnisse zu den semantisch-funktionalen Aspekten von ICCs aus Kapitel 4. Daraufhin implementiere ich die Ergebnisse zu den formalen Aspekten von ICCs aus Kapitel 5. Die so entwickelten Schemata und die Mechanismen der RM bieten zum Abschluss des Kapitels einen evidenzbasierten Erklärungsansatz zur Bedeutungskonstitution von ICCs.

10.4.1 Semantisch-funktionale Aspekte der ICC-Schemata

In der RM wird bisher kein eigenes Schema für ICCs beschrieben. Die Schemata zu den ICC-Typen müssen also auf der Grundlage der anderen Schemata erst entwickelt werden. Bei den Det-ICCs bietet sich das Kompositionsschema als Ausgangspunkt an, da Det-ICCs Nominalkomposita sind. Wie schon bei *Hauskatze* in (135b) werden dazu die Variablen durch die entsprechenden Strukturen der Konstituenten ersetzt. So zeigt sich, dass die spezifischen Strukturen eines Det-ICCs im Kompositionsschema in (135a) nicht genau erfasst werden und das Schema deshalb modifiziert werden muss. Beispiel (139) zeigt ein modifiziertes Kompositionsschema und eine Darstellung des Det-ICCs *Erwartungserwartung*:[82]

(139)		a. Kompositionsschema	b. *Erwartungserwartung*
	S:	$[F\ (X_x, X_x)]_y$	$[\text{ERWARTUNG}_1; \text{ABOUT}\ (\text{ERWARTUNG}_1)]_2$
	MS:	$[_N\ \underline{N}_x\ \underline{N}_x\ ]_y$	$[_N\ N_1\ N_1\ ]_2$
	P:	$/\underline{\ldots}_x\ \underline{\ldots}_x/_y$	/ɛɐ̯vaʁtʊŋ$_1$ s ɛɐ̯vaʁtʊŋ$_1$/$_2$

82 Der Anschaulichkeit wegen beginnen die Indizes an dieser Stelle wieder bei 1.

Beim ICC in (139b) zeigt Index $_1$ auf morphosyntaktischer Ebene an, dass der gebildete Nominalstamm ($_N$) aus ein und demselben Nominalstamm (N) gebildet ist, der im Wort zweimal vorhanden ist. Hierin liegt die Besonderheit, die Det-ICCs im Vergleich mit anderen determinativen N+N-Komposita aufweisen. N+N-Komposita im Allgemeinen haben zwei unterschiedliche Nominalstämme und deshalb im Formalismus Ns mit unterschiedlichen Koindizes. Für Det-ICCs ist also das allgemeine Kompositionsschema so zu modifizieren, dass die reduplikative Struktur der ICCs angemessen dargestellt wird. Das Schema in (139a) kann die Identität der Konstituenten in ICCs darstellen, da die beiden Variablen X auf semantischer Ebene wie auch die beiden tiefergestellten Indizes $_x$ genau wie die beiden Konstituenten im ICC identisch sind. Die Konstituenten stehen somit in Verbindung und die zwei Konzepte auf semantisch-konzeptueller Ebene sind – ungeachtet der Tatsache, dass sie in unterschiedlichen Positionen auftreten und womöglich auch unterschiedliche Funktionen tragen – identisch.

Die in Kapitel 2 besprochenen Beschreibungsansätze für ICCs beschäftigen sich viel mit der Frage, ob ICCs als Komposita anzusehen sind oder nicht. Die Argumente wurden in Kapitel 7 abgewogen und Det-ICCs den Komposita zugeordnet. Im Rahmen der RM aber stellt sich dieses Problem überhaupt nicht, da alle Struktureigenschaften der Wortbildungen unabhängig voneinander deklariert werden. Ein eigenes Det-ICC-Schema berücksichtigt lediglich die reduplikative Struktur auf semantischer, morphosyntaktischer und phonologischer Ebene, entspricht abseits der Identität der Variablen und relationalen Koindizes aber dem Kompositionsschema.

Auf semantischer Ebene ist eine *Erwartungserwartung* eine Erwartung, die sich auf eine (andere) Erwartung richtet. Die Funktion zwischen den identischen Konstituenten entspricht der semantischen Relation ABOUT. Auf phonologischer Ebene besteht das Det-ICC aus der Phonemverbindung /ɛɐ̯vaʁtʊŋ/ und /ɛɐ̯vaʁtʊŋ/ sowie aus dem Fugenelement /s/. Über Index $_2$ sind semantische, morphosyntaktische und phonologische Struktur verbunden.

Die Identität der Konstituenten, die über die Identität der Variablen und über die Koindexierung ausgedrückt wird, muss näher erläutert werden, denn man könnte auch dafür argumentieren, dass zwei unterschiedliche Konstituenten auf morphosyntaktischer und semantisch-konzeptueller Ebene zwei unterschiedliche Indizes erhalten. Die Konstituenten evozieren nämlich, wie in Kapitel 7 ausgeführt, zweimal ein Konzept. Im Äußerungskontext können deshalb, wie bei *Fenster-Fenster* in (48), zwei unterschiedliche ontologische Entsprechungen vorliegen. Hier muss man zwischen der semantisch-konzeptuellen Ebene eines Lexikoneintrags und dem Äußerungskontext eines Sprechers / einer Sprecherin unterscheiden. Es sind zwar zwei unterschiedliche Fenster im Äußerungskontext gemeint,

doch sind dies zwei Auftreten ein- und desselben Konzeptes FENSTER, die in Beziehung zueinander gesetzt werden. Die nominalen Konstituenten von ICCs können also mit zwei unterschiedlichen Dingen in der Welt verbunden werden, stellen aber stets zwei Auftreten eines Basislexems dar. Auch Van Santen und Booij (2017: 184) sehen in den beiden Konstituenten von ICCs (im Niederländischen) nur ein Basislexem gegeben, das zweimal formal auftritt. Die Form der Konstruktion geben sie darum mit $[[a]_{Xi}[a]_{Xi}]_X$ an.[83]

Das Konzept der Basis liegt in einem Det-ICC also zweimal vor. Dies ist bei Prot-ICCs anders. Wendet man das Kompositionsschema auf Prot-ICCs an, zeigt sich an einigen Stellen, dass das Schema für Prot-ICCs umfangreicher modifiziert werden muss als für die Darstellung von Det-ICCs. Beispiel (140) stellt das Prot-ICC *Oma-Oma* im modifizierten Kompositionsschema mit einer LIKE-Relation dar:

(140)

	Oma-Oma
Semantik:	$[\text{OMA}_3; \text{LIKE}\ (\text{OMA}_3)]_4$
Morphosyntax:	$[_N\ N_3\ N_3\ ]_4$
Phonologie:	/oːmɑː$_3$ oːmɑː$_3$/$_4$

Auf semantisch-konzeptueller Ebene passt die Relation LIKE in der Funktion der beiden identischen Konstituenten hier nicht zur Bedeutung, die Prot-ICCs zukommt. Im kontextualisierten Beispiel (54) ist *Oma-Oma* keine Oma, die wie eine (beliebige andere) Oma ist. Das Kompositionsschema muss also nicht nur im Hinblick auf die reduplikative Struktur modifiziert werden, sondern auf semantischer Ebene zusätzlich hinsichtlich der Funktion, die den Konstituenten zukommt. Eine adäquatere Darstellung sieht also wie folgt aus:

(141)

	Oma-Oma
Semantik:	$[\text{PROTOTYP}\ (\text{OMA}_3)]_5$
Morphosyntax:	$[_N\ N_3\ N_3\ ]_5$
Phonologie:	/oːmɑː$_3$ oːmɑː$_3$/$_5$

Ein Teil der semantischen Struktur ist über alle Prot-ICCs hinweg konstant, nämlich die Bedeutungskomponente PROTOTYP. Auf semantischer Ebene ist eine *Oma-Oma* eine Oma, die so ist, wie Omas üblicher- oder prototypischerweise sind. Wichtig ist, dass OMA auf dieser Ebene nur einmal vorhanden ist. Prot-ICCs unterscheiden sich von Det-ICCs darin, dass nicht zwei Auftreten eines Konzeptes zuei-

83 Die Frage, ab welcher semantischen Verschiebung von zwei Lexemen ausgegangen werden muss, wurde in Kapitel 4.2.3 diskutiert.

nander in Beziehung gesetzt werden, sondern das Basiskonzept nur einmal vorliegt und hinsichtlich prototypischer Eigenschaften spezifiziert wird. Auch die Semantik von *Oma-Oma* ist deshalb nicht die Funktion zweier Bedeutungen, sondern eine Bedeutung, die mithilfe der Funktion PROTOTYP verändert wird. Die Bedeutungsangabe PROTOTYP ist idiosynkratischer Bestandteil der Wortbildungsbedeutung.

Die Schnittstelle, die die RM – und auch die PA – zwischen semantisch-konzeptueller Struktur und räumlichen Repräsentationen annimmt, ermöglicht hier eine zusätzliche Darstellung des zentralen Unterschiedes zwischen Det- und Prot-ICCs. In Kapitel 4 habe ich argumentiert, dass das Erstglied von Prot-ICCs, anders als das von Det-ICCs, nicht ontologisch vorliegen kann. In der RM und in der PA wird angenommen, dass die Semantik von Wörtern und Schemata über die *Spatial Representations* in Verbindung mit den perzeptiven Fähigkeiten des Menschen steht (Jackendoff & Audring 2020b: 7f.). Die Vorstellung der Konzepte ist deshalb auch mit räumlichen Verhältnissen verbunden. Diese Annahme spricht dafür, dass das Basiskonzept in Det-ICCs zweimal vorliegt, in Prot-ICCs hingegen nur einmal. Bei manchen Det-ICCs, etwa bei *Fenster-Fenster*, fließt die Wahrnehmung zweier räumlich voneinander abgegrenzter, konturierter Raumausdehnungen in die Bildung des ICC-Konzeptes ein. Ein *Fenster-Fenster* ist ein Fenster, das sich räumlich vor einem anderen Fenster befindet. Ein- und dasselbe Konzept liegt auf semantisch-konzeptueller Ebene zweimal vor und kann darum auch zweimal ontologisch vorliegen. Bei *Oma-Oma* basiert das ICC-Konzept hingegen auf der Wahrnehmung einer einzigen räumlich begrenzten, konturierten Entität. Auf semantisch-konzeptueller Ebene besteht ein Prot-ICC deshalb auch aus nur einer Oma, die mithilfe des idiosynkratischen Schemabestandteils PROTOTYP modifiziert wird (PROTOTYP (OMA)). Diese Beobachtung ist ein weiterer Grund dafür, Prot-ICCs nicht wie in (140), sondern wie in (141) darzustellen. Aus semantisch-konzeptueller Ebene liegt nur ein Konzept vor, das dem Basislexem entspricht.

Dass die Konzeptbildung nicht notwendigerweise in Verbindung mit den *Spatial Representations* geschieht, zeigen aber ICCs wie *Erwartungserwartung*, das zwar Det-ICC ist und das Basiskonzept deshalb zweimal realisiert. Als Abstrakta erlauben aber weder das Basiskonzept noch das ICC-Konzept einen Verweis auf ein Objekt räumlicher Ausdehnung.

Auf der morphosyntaktischen Ebene ist die syntaktische Kategorie der Prot-ICCs das Nomen, im Schema durch ein tiefergestelltes $_{N}$ repräsentiert. Auch die Basis und der Reduplikant gehört bei allen ICCs der syntaktischen Kategorie Nomen an (N). Index $_{5}$ zeigt an, dass der gebildete Nominalstamm ($_{N}$) aus ein und demselben Nominalstamm gebildet ist, der im Wort zweimal vorhanden ist. Auf phonologischer Ebene tritt die Phonemverbindung /oːmɑː/ zweimal auf. Ein Fu-

genelement liegt auf keiner der Ebenen vor. Über Index $_6$ sind semantische, morphosyntaktische und phonologische Struktur verbunden. Von diesem ICC und anderen Prot-ICCs kann man das Schema in (142) abstrahieren:

(142) Prot-ICC-Schema

Semantik:	$[\text{PROTOTYP } (\underline{\underline{X}}_x)]_y$
Morphosyntax:	$[_N \underline{\underline{N}}_x \underline{\underline{N}}_x]_y$
Phonologie:	$/\underline{\underline{...}}_x \underline{\underline{...}}_x/_y$

Der prototypische Vertreter einer beliebigen Entität wird mithilfe eines Nomens bezeichnet, das die morphologische Verbindung zweier identischer Nomina ist. Das Basislexem ist weder formal noch semantisch restringiert. Prinzipiell können alle Nomina in dieser Struktur die Prototypikalität ihrer Referenten anzeigen. Dies wird über die doppelte Unterstreichung der Variable auf semantisch-konzeptuellen Struktur ausgedrückt ($\underline{\underline{X}}$). Auch die phonologischen Variablen können jegliche Ausprägung erhalten, solange sie nur identisch sind. Diese Identität wird über die Indizes ($_x$) angezeigt. Die in der Forschungsliteratur diskutierte Abwesenheit von Fugenelementen muss im Formalismus nicht extra notiert werden, weil Fugenelemente Phonologie ohne semantische oder morphosyntaktische Struktur sind. Auf phonologischer Ebene ist $\underline{\underline{...}}$ die Variable für die phonetische Realisierung, die in jedem Prot-ICC anders ist. Das tiefergestellte $_x$ drückt auch hier aus, dass der Lautkörper stets aus zwei identischen Teilen besteht.

Das zusätzlich auf semantischer Ebene modifizierte Kompositionsschema kann also auch Prot-ICCs in allen Facetten wiedergeben. Auch in Bezug auf Prot-ICCs können die in Kapitel 7 und 8 aufgeführten Argumente für und wider eine Kompositaanalyse in RM also berücksichtigt werden, ohne notgedrungen zu einer absoluten Zuordnung zu führen. Die Notation im Rahmen der RM in (142) enthält alle aufgeführten Argumente und zeigt schlichtweg diejenigen Strukturbestandteile, die das Prot-ICC-Schema mit dem Kompositionsschema teilt, diejenigen, die es mit Reduplikationsschemata teilt, sowie diejenigen, die spezifisch für Prot-ICCs sind. Mit dem Kompositionsschema teilt das Prot-ICC-Schema Teile der morphosyntaktischen Struktur ($[_N$ N N]); mit Reduplikationsschemata die phonologische Struktur. Beispielsweise entspricht die phonologische Struktur des Prot-ICC-Schemas exakt der, die Jackendoff und Audring für die Pluralreduplikation im Warlpiri annehmen (/…$_x$ …$_x$/) (Jackendoff & Audring 2020b: 123). Spezifisch für Prot-ICCs, und somit der Grund, ein eigenes Schema für Prot-ICCs im Lexikon anzunehmen, ist schließlich die Bedeutungskomponente (PROTOTYP) auf der semantisch-konzeptuellen Ebene.

Ähnlich wie bei den Det- und Prot-ICCs zeigt sich auch bei Name-ICCs, dass sie sich nicht allein über das Kompositionsschema sinnvoll darstellen lassen. Zwar ist auch hier die morphosyntaktische Struktur [$_{N}$ N N]. Doch ist die Semantik kein Konzept, das aus einer Funktion der Konstituenten entsteht. Stattdessen wird vom Name-ICC direkt auf eine Entität referiert, die Bedeutung der Konstituenten ist also entweder, wie teildeskriptive Name-ICCs vom Typ *AutoAuto*, nur peripher vorhanden oder sogar, wie in arbiträren Name-ICCs wie *PunktPunkt* in (143b), gänzlich irrelevant.

Name-ICCs wie *PunktPunkt* teilen allerdings viel Struktur mit dem Eigennamenschema (Jackendoff & Audring 2020b: 116):[84]

(143)		a. Eigennamenschema	b. *PunktPunkt*
	Semantik:	[X]$_{x}$	[Dɟ-Duo]$_{6}$
	Morphosyntax.:	[Eigenname]$_{x}$	[Eigenname]$_{6}$
	Phonologie:	/…/$_{x}$	/pʊŋktpʊŋkt/$_{6}$

Doch auch dieses Schema weist Unterschiede zu Name-ICCs auf. Zunächst müssen zwei wichtige Unterschiede zwischen den in der RM und den in dieser Arbeit getroffenen Annahmen zu Eigennamen erläutert werden. Zum einen nimmt das Eigennamenschema von Jackendoff und Audring eine eigene syntaktische Kategorie „Proper Noun“ an (Jackendoff & Audring 2020b: 116). In dieser Arbeit werden Eigennamen hingegen nicht als eigene Wortart angesehen, sondern als eine Klasse innerhalb der Substantive (Schlücker & Ackermann 2017). Deshalb wird ein Name-ICC in den Schemata zu Eigennamen dieser Arbeit auf morphosyntaktischer Ebene als Nomen angesehen. Zum anderen weist das Eigennamenschema von Jackendoff und Audring eine semantische Struktur auf. Es ist aber fraglich, ob es sinnvoll ist, für das Eigennamenschema eine Strukturebene namens „Semantik“ anzunehmen, da Namen für die Referenz generell keine Semantik benötigen. Hier muss man also stillschweigend voraussetzen, dass der semantische Gehalt auf der semantisch-konzeptuellen Ebene nicht über die Identifizierung einer konkreten Entität hinausgeht. Bei der Beschreibung der arbiträren Name-ICCs in Kapitel 4 habe ich den Bildungen eine semantische Struktur abgesprochen und die Referenz als ihre einzige Funktion angesehen. Hinsichtlich der Benennung der

84 Die Begrifflichkeiten wurden von mir ins Deutsche übersetzt. Zudem habe ich die Variablen der Schnittstellenverbindungen des Eigennamenschemas in x und y geändert. Jackendoff und Audring (2020b) stellen das Eigennamenschema im Kontext seiner Verbindung zum Kürzungsschema dar und verwenden für diese relationale Verbindung zwischen Schemata, wie in ihrer Notation vorgesehen, griechische Buchstaben.

Strukturebenen wird der Einheitlichkeit wegen aber bei diesen Bildungen ebenfalls von einer semantisch-konzeptuellen oder semantischen Ebene gesprochen.

Name-ICCs lassen sich nicht angemessen über das allgemeine Eigennamenschema, wie es Jackendoff und Audring beschreiben, darstellen. Die Darstellung in (143b) passt nicht zur internen morphosyntaktischen und phonologischen Struktur in Name-ICCs. Die Konstituenten sind, anders als in Eigennamen wie *Johann* oder *Sven*, für die Sprecher:innen identifizierbar. Da Name-ICCs also sowohl mit dem Eigennamenschema als auch mit dem Kompositionsschema Struktur teilen, kann man die entsprechenden Komponenten der beiden Schemata kombinieren und das Name-ICC-Schema in (144a) annehmen:

(144)	a. Name-ICCs (arbiträr)	b. *PunktPunkt*
Semantik:	$[\underline{\underline{X}}]_y$	$[\text{DJ-DUO}]_9$
Morphosyntax.:	$[_N \underline{\underline{N}}_x \underline{\underline{N}}_x]_y$	$[_N N_8 N_8]_9$
Phonologie:	$/\underline{\underline{...}}_x \underline{\underline{...}}_x/_y$	/pʊŋkt$_8$ pʊŋkt$_8$/$_9$

Das Name-ICC-Schema besteht fast ausschließlich aus Variablen, da ein Lautkörper arbiträr mit einer Entität verknüpft ist und somit nur wenig konstante Struktur im Schema vorhanden ist. Die semantische Struktur bleibt mit der im Eigennamenschema identisch, denn das ICC bildet kein Konzept, das die Funktion zweier Konstituenten ist, sondern referiert direkt auf eine Entität, im Fall von *PunktPunkt* auf ein DJ-Duo. Bei Name-ICCs gibt es also keine Bedeutungskomponente, die über alle Name-ICCs hinweg gleichbliebe. Mit jedem Name-ICC ändert sich hier die semantisch-konzeptuelle Ebene vollständig, da unterschiedliche Name-ICCs stets auch auf unterschiedliche Objekte referieren und diese Referenz die einzige Funktion der Bildungen ist. Deshalb besteht diese Ebene bei Name-ICCs gänzlich aus einer Variable. Von Eigennamen wie *Johann* oder *Sven* unterscheidet sich das Schema allein darin, dass die morphosyntaktische Struktur erkennbar ist. Da die morphosyntaktische Struktur sichtbar bleibt, erkennen Sprecher:innen wie bei den anderen ICC-Arten die Bestandteile der Wortbildung. Die zwei identischen Nominalstämme, die den Eigennamen bilden, werden darum auf dieser Ebene auch abgebildet.

Die Bestandteile der morphosyntaktischen Struktur sind weiterhin über die Schnittstellenverbindung $_9$ mit denen aus der phonologischen Struktur verbunden. Bei Name-ICCs ist die syntaktische Kategorie der Bildung $_N$. Im Vergleich mit dem Eigennamenschema bietet ein Name-ICC auf morphosyntaktischer Ebene aber mehr Struktur, da die Struktur des Nominalkompositums sichtbar bleibt. Auch hier zeigt die doppelte Unterstreichung von ($\underline{\underline{N}}$) an, dass eine offene Variable

vorliegt, also abseits der syntaktischen Kategorie keine Beschränkungen existieren und potentiell jedes Nomen im Schema verwendet werden kann.

Im Unterschied zum Det- und Prot-ICC-Schema verbindet bei Name-ICCs das tiefergestellte $_{x}$ die Ns nur miteinander, um deren Identität anzuzeigen und mit den Bestandteilen der phonologischen Struktur, weil das Name-ICC seine Phonologie aus dem Basisnomen erhält. Das tiefergestellte $_{x}$ stellt allerdings keine Verbindung zur semantischen Struktur her, da diese, wie bei Eigennamen generell, nicht besteht. In diesem Detail besteht der wichtigste Unterschied zwischen arbiträren Name-ICCs und den anderen ICC-Typen.

In Name-ICCs wie *AutoAuto* gibt es schließlich eine Verbindung zur semantischen Struktur, weil hier in Form des Modifikators deskriptive Bedeutungsanteile bestehen. *AutoAuto* referiert auf etwas, das mit einem Auto zu tun hat. In diesem Schema ist deshalb der modifizierende Bestandteil der semantisch-konzeptuellen Struktur mit dem Nominalstamm der morphosyntaktischen Struktur und der entsprechenden Variable der phonologischen Struktur verbunden. In dieser Hinsicht ähnelt das teildeskriptive Name-ICC-Schema also dem Det-ICC-Schema:

(145)		a. Name-ICCs (teildeskriptiv)	b. *AutoAuto*
	Semantik:	$[F\ (\underline{X},\ \underline{Y}_{x})]_{y}$	$[\textsc{show};\ \textsc{about}\ (\textsc{auto}_{10})]_{11}$
	Morphosyntax:	$[_{N}\ \underline{N}_{x}\ \underline{N}_{x}]_{y}$	$[_{N}\ N_{10}\ N_{10}]_{11}$
	Phonologie:	$/\underline{\ldots}_{x}\ \underline{\ldots}_{x}/_{y}$	/aʊ̯toː$_{10}$ aʊ̯toː$_{10}$/$_{11}$

Eine Verbindung zwischen der morphosyntaktischen und phonologischen Struktur einerseits und der semantischen andererseits umfasst aber nicht die Bedeutungskomponente SHOW, da das ICCs zwar anzeigt, dass der Referent mit einem Auto zu tun hat, nicht aber, dass es sich um eine Show handelt. Im Schema ist deshalb auch hier das X nicht über Indizes mit den Bestandteilen der morphosyntaktischen und phonologischen Struktur verbunden. Allein über die Schnittstellenverbindung $_{y}$, also über das lexikalische Wissen über das jeweilige Name-ICC, gelingt die Referenz auf eine Entität.

Die ICC-Schemata weisen einige Gemeinsamkeiten auf. In konstruktionsgrammatischen Ansätzen werden Gemeinsamkeiten zwischen Konstruktionen mit dem Begriff der *Inheritance* beschrieben. Speziellere Konstruktionen erben Struktur von allgemeineren. *Second-order schemas*, Paare von Schemata, teilen sich Struktur, weil sie in einem abstrakten Mutterschema enthalten sind und von ihm teilweise dieselbe Struktur erben. Die in der RM angenommenen Schwesterschemata sind flexibler, weil sie nicht ineinander enthalten sind und auch keine Mutter haben, von der sie ihre geteilte Struktur erben. Sie stehen im Lexikon gewissermaßen nebeneinander. Auf diese Weise sind horizontale Verbindungen

zwischen Schemata (relationale Verbindungen) möglich. Die drei in (146a–149a) dargestellten Schemata sind Schwestern. Mithilfe griechischer Koindizes in der Notation werden, wiederum für jede der drei Strukturebenen separat, ihre Gemeinsamkeiten markiert. Die ICC-Typen und die Verbindung zwischen ihnen lassen sich also über die folgenden Schemata darstellen:

(146)		a. Det-ICC-Schema	b. *Erwartungserwartung*
	Semantik:	$[\text{F } (\underline{\underline{X}}_{\alpha}, \underline{\underline{X}}_{\alpha})]_{w}$	$[\text{ERWARTUNG}_{1}; \text{ ABOUT } (\text{ERWARTUNG}_{1})]_{2}$
	MorpSyn:	$[_{N} \underline{\underline{N}}_{\alpha} \underline{\underline{N}}_{\alpha}]_{\beta,w}$	$[_{N} N_{1} N_{1}]_{2}$
	Phon:	$/\underline{\underline{...}}_{\alpha} \underline{\underline{...}}_{\alpha} /_{\gamma,w}$	/ɛ̪ʋaʁtʊŋ$_{1}$ s ɛ̪ʋaʁtʊŋ$_{1}$/$_{2}$

(147)		a. Prot-ICC-Schema	b. *Oma-Oma*
	Semantik:	$[\text{PROTOTYP } (\underline{\underline{X}}_{\alpha})]_{x}$	$[\text{PROTOTYP } (\text{OMA}_{3})]_{5}$
	MorpSyn:	$[_{N} \underline{\underline{N}}_{\alpha} \underline{\underline{N}}_{\alpha}]_{\beta,x}$	$[_{N} N_{3} N_{3}]_{5}$
	Phon:	$/\underline{\underline{...}}_{\alpha} \underline{\underline{...}}_{\alpha}/_{\gamma,x}$	/oːmɑːˑ$_{3}$ oːmɑːˑ$_{3}$/$_{5}$

(148)		a. Name-ICCs (arbiträr)	b. *PunktPunkt*
	Semantik:	$[\underline{\underline{X}}]_{y}$	$[\text{DJ-DUO}]_{9}$
	MorpSyn:	$[_{N} \underline{\underline{N}}_{\alpha} \underline{\underline{N}}_{\alpha}]_{\beta,y}$	$[_{N} N_{8} N_{8}]_{9}$
	Phon:	$/\underline{\underline{...}}_{\alpha} \underline{\underline{...}}_{\alpha}/_{\gamma,y}$	/pʊŋkt$_{8}$ pʊŋkt$_{8}$/$_{9}$

(149)		a. Name-ICCs (teildeskriptiv)	b. *AutoAuto*
	Semantik:	$[\text{F } (\underline{\underline{X}}, \underline{\underline{Y}}_{\alpha})]_{z}$	$[\text{SHOW}; \text{ ABOUT } (\text{AUTO}_{10})]_{11}$
	MorpSyn:	$[_{N} \underline{\underline{N}}_{\alpha} \underline{\underline{N}}_{\alpha}]_{\beta,z}$	$[_{N} N_{10} N_{10}]_{11}$
	Phon:	$/\underline{\underline{...}}_{\alpha} \underline{\underline{...}}_{\alpha}/_{\gamma,z}$	/a̪ʊ̯toːˑ$_{10}$ a̪ʊ̯toːˑ$_{10}$/$_{11}$

Die Identität der Konstituenten ist bei allen drei Schemata gegeben, was auf morphosyntaktischer und phonologischer Ebene durch α notiert wird. Außerdem haben alle Schemata auf der morphosyntaktischen Ebene eine nominale Basis $\underline{\underline{N}}$. Schließlich teilen die drei ICC-Schemata noch die gesamte morphosyntaktische (β) und phonologische Struktur (γ).

Den Gemeinsamkeiten der drei ICC-Schemata stehen nur wenige Unterschiede gegenüber. Zum einen fehlt beim arbiträren Name-ICC-Schema der Link zwischen der Phonologie/Morphosyntax und der Bedeutungsvariable in der Semantik, sodass auf dieser Ebene auch die relationale Verbindung zu den anderen Schemata (α) fehlt. Zum anderen ist im teildeskriptiven Name-ICC-Schema nur eine Konstituente mit den Variablen der morphosyntaktischen und phonologischen Ebene verbunden, da der semantische Kopf fehlt und also keine Verbindung zwischen dem Kopfkonzept und dem Basislexem besteht. Schließlich unter-

scheidet sich natürlich noch die semantische Struktur zwischen den drei Schemata, weshalb hier keine griechischen Koindizes auftauchen.

Im Folgenden werden die Ergebnisse der in Teil II vorgestellten Korpusstudien für die Repräsentation der ICC-Schemata im Sprachsystem berücksichtigt. Im Zuge dessen wird auch ein Versuch unternommen, die Frage zu beantworten, wie Schemata mit identischer morphosyntaktischer und phonologischer Struktur dennoch unterschiedliche semantisch-konzeptuelle Strukturen haben können.

10.4.2 Formale Aspekte der ICC-Schemata

Die Unterschiede zwischen den bisher formulierten ICC-Schemata betreffen nur die Semantik. Insgesamt zeigt sich in den Korpusdaten, dass die doppelten Unterstreichungen der Variablen in den Schemata angemessen sind, da zu keinem der drei ICC-Schemata erkennbare Beschränkungen der Basislexeme bestehen. Auch die Basislexeme können also nichts zur Disambiguierung der ICCs beitragen. Die in Teil II vorgestellten Korpusstudien haben allerdings auch Evidenz dafür geliefert, dass auf der Ebene der Phonologie Unterschiede zwischen den Schemata bestehen.[85] In den meisten Modellen semantischer Transparenz sowie in den meisten theoretischen Texten zu ICCs wird angenommen, dass allein der Kontext und das Weltwissen die Bildungen disambiguieren, sodass die jeweilige semantische Struktur dekodiert werden kann (Freywald 2015, Hohenhaus 2004, 2015, Kentner 2017). So schreibt Freywald (2015: 923): "Whether the […] real-X-reading is employed […], depends solely on the context." Doch erscheint das angesichts der Ergebnisse aus den Korpusstudien, die formale Unterschiede zwischen den ICC-Typen nahelegen, und angesichts der Studie von Finkbeiner (2014) unwahrscheinlich. Finkbeiners Studie weist nach, dass ICCs auch ohne Kontext mal als Det- und mal als Prot-ICCs interpretiert werden und sich die Proband:innen bei der Interpretation zudem weitgehend einig sind.

Die formalen Unterscheidungskriterien zwischen den ICCs sollten deshalb auch bei ihrer Modellierung im Rahmen der RM integriert werden. Eine Möglichkeit der formalen Disambiguierung sind die von unterschiedlicher Seite ins Spiel gebrachten Fugenelemente/Flexionsmarker am Erstglied (Bross & Fraser 2020: 3, Freywald 2015: 925). Zwar gelten Fugenelemente in der RM als Phonologie ohne

85 Da es sich bei den Korpusdaten um Daten zur geschriebenen Sprache handelt, sind diese Unterschiede selbstverständlich keine phonologischen, sondern, allgemeiner, Unterschiede in der formalen Realisierung der semantisch-konzeptuellen und morphosyntaktischen Strukturen. Auf die rein graphematischen Aspekte der ICC-Schemata gehe ich im Folgenden aber noch näher ein.

Semantik (Jackendoff & Audring 2020b: 100). Doch belegen die Korpusstudien eindeutig, dass der Unterschied in der Verwendung von Fugenelementen zwischen den Det-ICCs einerseits und Prot- und Name-ICCs andererseits deutlich ist: In mehr als 90% aller Fälle zeigen Det-ICCs Fugenelemente, Prot- und Name-ICCs hingegen in weniger als 13%, beziehungsweise in weniger als 8%. Es scheint also gerechtfertigt, von einer bedeutungsunterscheidenden Funktion der Fugenelemente zu sprechen und sie im Schema als Marker für Det-ICCs zu berücksichtigen:

(150) Det-ICC-Schema

Semantik:	$[F\ (\underline{X}_{\alpha}, \underline{X}_{\alpha})]_{w}$
Morphosyntax:	$[_{N}\ \underline{\underline{N}}_{\alpha}\ \underline{\underline{N}}_{\alpha}\ ]_{\beta,w}$
Phonologie:	$/\underline{\ldots}_{\alpha}$ *⟨ə⟩⟨ʁ⟩⟨n⟩⟨s⟩* $\underline{\ldots}_{\alpha}\ /_{\gamma,w}$

Das Auftreten der Fugenelemente relativiert dabei die Identität der phonologischen Struktur von Det-ICC-Schema einerseits und Prot- und Name-ICC-Schema andererseits. Trotz dieser formalen Abweichung der Det-ICCs von den anderen ICC-Schemata ist die phonologische Struktur der drei ICC-Schemata noch immer sehr ähnlich. Diese eingeschränkte Übereinstimmung kann mit Sternchen angezeigt werden. Die gesternten Strukturbestandteile zeigen die Abweichung der Det-ICC-Struktur von der Struktur der anderen ICC-Schemata an (Same Except).

Die Winkelklammern ⟨ ⟩ wiederum zeigen in der Notation der RM Optionales an (Jackendoff & Audring 2020b: 87).[86] Zum einen sind Fugenelemente in Det-ICCs generell optional, da nicht alle Det-ICC-Wortformen verfugt sind. Zum anderen sind die Phoneme, aus denen die Fugenelemente bestehen, jeweils optional, da in Det-ICCs die Fugenelemente *-e-*, *-(e)s-*, *-(e)r-*, *-(e)n-*, *-(e)rs-*, sowie *-(e)ns-* auftreten. Dass die Fugenelemente in Det-ICCs optional sind, bedingt, dass die phonologische Struktur des Det-ICC-Schemas nicht in jedem Fall bedeutungsunterscheidend ist. Viele Nomina, beispielsweise *Trainer* oder *Besen*, erlauben generell keine Fugenelemente. Die phonologische Struktur der drei ICC-Schemata unterscheidet sich bloß in der Möglichkeit zur Verwendung von Fugenelementen. Der Auffassung von Finkbeiner (2014: 187), ein Fugenelement am Erstglied eines ICCs schlösse die Prototypenlesart aus, ist also zuzustimmen und dahingehend zu ergänzen, dass auch Name-ICCs dadurch ausgeschlossen werden. Das Ausbleiben eines Fugenelementes zeigt also nicht das Vorliegen eines bestimmten ICC-Typs an. Vielmehr zeigt die Verwendung eines Fugenelements an, dass ein Det-ICC vorliegt.

86 Jackendoff und Audring verwenden hier die Zeichen „<“ (kleiner als) und „>“ (größer als). Da diese aber auch bei der graphematischen Transkription verwendet werden, nutze ich für Optionales die ‘eigentlichen’ Winkelklammern, also „⟨“ und „⟩“.

Die Korpusstudien belegen aber auch, dass das Auftreten von Fugenelementen in Prot- und Name-ICCs nicht ausgeschlossen ist. Stattdessen ist ein ICC mit Fugenelement bloß deutlich unwahrscheinlicher Prot- oder Name-ICC als Det-ICC. Die RM ermöglicht eine Darstellung dieser probabilistischen Unterschiede, indem neben den beschriebenen (und noch folgenden) ICC-Schemata jeweils alternative Schemata existieren. Es gibt also auch ein Prot- und ein Name-ICC-Schema, in dem die phonologische Struktur /‿…‿$_{\alpha}$ *⟨ə⟩⟨ʁ⟩⟨n⟩⟨s⟩* …$_{\alpha}$ / lautet. Die Korpusstudien liefern aber Evidenz dafür, dass diese Schemata weniger prominent sind und beim Auftreten von Fugenelementen in einer morphologischen Verbindung aus identischen Nominalstämmen eher die semantische Struktur des Det-ICC-Schemas aktiviert wird.

Ein weiterer formaler Unterschied zwischen den ICC-Typen kann allein auf der Ebene der Schreibung wahrgenommen werden. Für Alphabetschriften wird in der RM eine mittlere Repräsentationsebene angenommen, die zwischen der Phonologie und den Schriftzeichen liegt. Dies ist die orthographische Struktur (Jackendoff & Audring 2020b: 252). Die orthographische Struktur – das ist wichtig zu erwähnen – stellt hier keine graphematische Transkription der phonologischen Struktur dar. Die Zeichen der orthographischen Ebene und die der phonologischen Ebene entsprechen einander also nicht basierend auf den Phonem-Graphem-Korrespondenzregeln. Stattdessen wird die phonologische Struktur von Segmenten lexikalischer[87] Einheiten (Erstglied und Zweitglied) einer orthographischen Struktur von Segmenten lexikalischer Einheiten zugeordnet. Dies geschieht auf der Basis der Orthographie, nicht auf der Basis von Laut-Schriftzeichen-Zuordnungsregeln.

Die Schemata können um diese Strukturebene erweitert werden (151–154). Wie schon bei der Verwendung der Fugenelemente markiert aber auch die Schreibung nicht immer eindeutig einen speziellen ICC-Typen. Die Verwendung der Anführungszeichen ist bei allen ICCs zwar sehr viel häufiger als bei Nomina im Allgemeinen, sodass damit hervorgehoben werden kann, dass eine besondere Wortbildung vorliegt. Die ICC-Schemata lassen sich damit allerdings nicht voneinander unterscheiden, da alle ICC-Typen ähnlich häufig in Anführungszeichen gesetzt werden: Det-ICCs in 7% der Fälle, Prot-ICCs in 8% der Fälle und Name-ICCs in 11% der Fälle. Allerdings hilft die graphische Trennung der Konstituenten dabei, den ICC-Typ anzuzeigen. Die graphische Trennung der Konstituenten (Bindestrich, Spatium, Binnenmajuskel) tritt in den Korpusdaten sehr viel häufiger in Prot-ICCs (55%) und Name-ICCs (63%) auf als in Det-ICCs (8%). Auch wenn der Unterschied in Bezug auf die Trennung der Konstituenten weniger deutlich ist als bei der Verfugung, kann dieser Formunterschied in den Schemata berücksichtigt werden, um die Bedeutungskonstitution genauer modellieren zu können. Wie schon bei der Verfugung

87 Auf die Fugenelemente trifft das also nicht zu, da sie keine lexikalischen Einheiten sind.

funktioniert auch dieses formale Merkmal nur in eine Richtung. Die Zusammenschreibung kann nicht eindeutig anzeigen, dass ein Det-ICC vorliegt, da auch Prot- und Name-ICCs zu nicht unwesentlichen Anteilen zusammengeschrieben werden. Andersherum kann aber die Verwendung von Mitteln, die die Konstituenten graphisch trennen, anzeigen, dass kein Det-ICC vorliegt, da diese nur sehr selten solche Elemente enthalten. Die Schemata können also mithilfe der orthographischen Struktur wie folgt um ein Mittel der Bedeutungskonstitution erweitert werden:[88]

(151) a. Det-ICC-Schema

Semantik:	$[F\ (\underline{X}_{\alpha}, \underline{Y}_{\alpha})]_{w}$
Morphosyntax:	$[_{N}\ \underline{N}_{\alpha}\ \underline{N}_{\alpha}\ ]_{\beta,w}$
Phonologie:	$/\underline{\ldots}_{\alpha}$ *⟨ə_{a}⟩⟨ʁ_{b}⟩⟨n_{c}⟩⟨s_{d}⟩* $\underline{\ldots}_{\alpha}$ $/_{\gamma,w}$
Orthographie:	$<\underline{\ldots}_{\alpha}$ *⟨e_{a}⟩⟨r_{b}⟩⟨n_{c}⟩⟨s_{d}⟩* $\underline{\ldots}_{\alpha}>_{\delta,w}$

b. *Erwartungserwartung*

Semantik:	$[\text{ERWARTUNG}_{1};\ \text{ABOUT}\ (\text{ERWARTUNG}_{1})]_{2}$
Morphosyntax:	$[_{N}\ N_{1}\ N_{1}\ ]_{2}$
Phonologie:	$/\text{ɛɐ̯vaʁtʊŋ}_{1}\ \text{s}_{12}\ \text{ɛɐ̯vaʁtʊŋ}_{1}/_{2}$
Orthographie:	$<\text{Erwartung}_{1}\ \text{s}_{12}\ \text{erwartung}_{1}>_{2}$

(152) a. Prot-ICC-Schema

Semantik:	$[\text{PROTOTYP}\ (\underline{X}_{\alpha})]_{x}$
Morphosyntax:	$[_{N}\ \underline{N}_{\alpha}\ \underline{N}_{\alpha}\ ]_{\beta,x}$
Phonologie:	$/\underline{\ldots}_{\alpha}\ \underline{\ldots}_{\alpha}/_{\gamma,x}$
Orthographie:	$<\underline{\ldots}_{\alpha}$ *⟨-⟩⟨ ⟩⟨[λ, upper case]⟩* $\underline{\ldots}_{\alpha}>_{\delta,x}$

b. *Oma-Oma*

Semantik:	$[\text{PROTOTYP}\ (\text{OMA}_{3})]_{5}$
Morphosyntax:	$[_{N}\ N_{3}\ N_{3}\ ]_{5}$
Phonologie:	$/\text{o:mɑ:}_{3}\ \text{o:mɑ:}_{3}/_{5}$
Orthographie:	$<\text{Oma-Oma}_{3}>_{5}$

88 Die Repräsentation der orthographischen Struktur wird hier hinsichtlich der Großschreibung von Nomina vereinfacht dargestellt. Strenggenommen müsste im Det-ICC-Schema der Minuskel zu Beginn der zweiten Konstituente als Abweichung von der orthographischen Schreibung der Nomina gekennzeichnet und die Identität der beiden Konstituenten sowie der relationale Link zwischen den Schemata entsprechend mithilfe von ** eingeschränkt werden. Da es hier aber vor allem um die Aspekte geht, die die ICC-Schemata formal unterscheiden, werden diese Aspekte der Übersicht halber ignoriert.

(153) a. Name-ICC-Schema (arbiträr)

Semantik:	[$\underline{X}$]$_{y}$
Morphosyntax:	[$_{N}$ $\underline{N}_{\alpha}$ $\underline{N}_{\alpha}$]$_{\beta,y}$
Phonologie:	/$\underline{\ldots}_{\alpha}$ $\underline{\ldots}_{\alpha}$/$_{\gamma,y}$
Orthographie:	<$\underline{\ldots}_{\alpha}$ *⟨-⟩⟨ ⟩⟨[λ, upper case]⟩* $\underline{\ldots}_{\alpha}$>$_{\delta,y}$

b. *PunktPunkt*

Semantik:	[Dj-Duo]$_{9}$
Morphosyntax:	[$_{N}$ N$_{8}$ N$_{8}$]$_{9}$
Phonologie:	/pʊŋkt$_{8}$ pʊŋkt$_{8}$/$_{9}$
Orthographie:	<Punkt$_{8}$Punkt$_{8}$>$_{9}$

(154) a. Name-ICCs-Schema, teildeskriptiv

Semantik:	[F ($\underline{X}$, $\underline{Y}_{\alpha}$)]$_{z}$
Morphosyntax:	[$_{N}$ $\underline{N}_{\alpha}$ $\underline{N}_{\alpha}$]$_{\beta,z}$
Phonologie:	/$\underline{\ldots}_{\alpha}$ $\underline{\ldots}_{\alpha}$/$_{\gamma,z}$
Orthographie:	<$\underline{\ldots}_{\alpha}$ *⟨-⟩⟨ ⟩⟨[λ, upper case]⟩* $\underline{\ldots}_{\alpha}$>$_{\delta,z}$

AutoAuto

Semantik:	[SHOW; ABOUT (AUTO$_{10}$)]$_{11}$
Morphosyntax:	[$_{N}$ N$_{10}$ N$_{10}$]$_{11}$
Phonologie:	/aʊto:$_{10}$ aʊto:$_{10}$/$_{11}$
Orthographie:	<Auto$_{10}$Auto$_{10}$>$_{11}$

Auch auf der orthographischen Ebene zeigt in der Notation ein griechischer Buchstabe (δ) an, dass die Struktur von Prot- und Name-ICC-Schema gleich ist. Die Trennung der Konstituenten geschieht bei Det-ICCs allein über die optionalen Fugenelemente (⟨e⟩⟨r⟩⟨n⟩⟨s⟩), die den jeweiligen Elementen auf phonologischer Ebene entsprechen (⟨ə⟩⟨ʁ⟩⟨n⟩⟨s⟩) und deshalb über die Koindizes $_{a}$, $_{b}$, $_{c}$ und $_{d}$, mit ihnen verbunden sind, bei Prot-ICCs und Name-ICCs hingegen über den Bindestrich (⟨-⟩), das Spatium (⟨ ⟩) oder die Binnenmajuskel (⟨[λ, upper case]⟩), die zwischen den Konstituenten ($\underline{\ldots}_{\alpha}$) auftreten können. Bei den Binnenmajuskeln ist λ eine Variable, die alle Buchstaben umfasst, und mit der Angabe [upper case] anzeigt, dass an der Konstituentengrenze nicht die Defaultschreibung [lower case] vorliegt.

Durch die Korpusstudien können die ICC-Schemata also evidenzbasiert genauer dargestellt werden. Die Fugenelemente und die unterschiedliche Schreibung der ICC-Typen sind allerdings nur zwei Puzzleteile bei der Spezifizierung der ICCs. Sie erklären einen Teil der ICC-Bedeutungsdifferenzierung und bieten eine

Möglichkeit, Det-ICCs von Prot- und Name-ICCs zu unterscheiden. Sie sind aber nicht ausreichend, wenn es gilt, drei verschiedene Bedeutungen zu unterscheiden. Zudem kann nicht in jedem Fall auf formale Mittel der Diskriminierung zurückgegriffen werden, weil nicht jeder Stamm Fugenelemente erlaubt. Auch die Studie von Finkbeiner (2014) spricht dagegen, dass ICCs allein auf formale Mittel zurückgreifen, um die von den Sprecher:innen intendierte Interpretation zu gewährleisten. In Finkbeiner (2014) werden alle Stimuli ohne Fugenelemente oder Flexionsmarker präsentiert. Zudem werden die Konstituenten der Stimuli allesamt auf dieselbe Weise, nämlich durch den Bindestrich, graphisch voneinander getrennt. Dennoch entscheiden sich die Teilnehmer:innen recht einheitlich für eine der Lesarten. Fugenelemente und die graphische Trennung der Konstituenten sind also offensichtlich nicht die einzigen Mittel, an denen sich die Sprecher:innen orientieren, wenn sie ICC bilden und interpretieren.

Die Ergebnisse aus den Korpusstudien zeigen aber, welcher Aspekt der Bedeutungskonstitution der drei ICC-Schemata bisher unberücksichtigt blieb. In den deTenTen13-Daten zeigt sich, dass der weitaus größte Teil der Stämme Basis nur eines ICC-Typs ist. 90% der Stämme werden exklusiv in nur einem der drei ICC-Typen verwendet. Nur 1,5% der Stämme bilden alle ICC-Typen. Die allermeisten Stämme haben also eine klare Präferenz für einen der drei ICC-Typen. Die Stämme, die mehr als nur einen ICC-Typ bilden, tun dies nur selten mit einer ausgeglichenen Verteilung. In Kapitel 6 wurde keine Antwort auf die Frage gefunden, wie sich diese eindeutige Verteilung der ICC-Stämme auf die ICC-Typen erklären lässt, da sich ICC-Bildungen nicht gegenseitig blockieren können. Die meisten ICCs sind nicht lexikalisiert und ein großer Teil der erhobenen Daten besteht aus Hapax legomena. Es ist also unwahrscheinlich, dass die Belege selbst, also die tatsächliche Existenz eines anderen ICCs, die Bildung von ICCs anderer Schemata blockieren. Die klare Aufteilung der Stämme zu den drei ICC-Schemata kann nicht durch etwaige Einträge derselben morphosyntaktischen Struktur im mentalen Lexikon erklärt werden.

Die klare Zuordnung von Stämmen zu ICC-Typen lässt sich hingegen gut erklären, wenn man – wie die RM – annimmt, dass es relationale Verbindungen zwischen den Schemata gibt. Die ICCs eines Typs, denen ein:e Sprecher:in begegnet ist, verankern nach und nach ein Schema im mentalen Lexikon der Sprecher:in. ICCs mit Monoreferenz verankern das Name-ICC-Schema, das Det-ICC-Schema wird vor allem durch lexikalisierte Bildungen wie *Kindeskind* gestützt und eine Prototypenlesart von ICCs wie *Oma-Oma* führt zu einer mehr oder weniger gefestigten, eigenständigen, semantisch-konzeptuellen Struktur des Prot-ICC-Schemas. Über die relationalen Verbindungen der morphosyntaktischen, phonologischen und orthographischen Struktur sind diese drei Schemata miteinander

verbunden. Im mentalen Lexikon der Sprecher:innen stehen also die semantisch-konzeptuellen Strukturen dreier Schemata für die Dekodierung zur Verfügung.

Die Schemata geben somit auch eine Antwort auf die onomasiologisch formulierte Frage, warum ein:e Sprecher:in für das, was ausgedrückt werden soll, die morphologische ICC-Realisierung und nicht etwa eine syntaktische, phrasale Realisierung wählt. Während ein ICC auf der Basis eines ICC-Schemas ad hoc gebildet wird, sind diese Schemata im mentalen Lexikon der Sprecher:innen aktiv und die möglichen Bedeutungen der entsprechenden ICCs verfügbar. Ist ein ICC in einem konkreten Äußerungskontext ambig, weil die morphosyntaktische (β), phonologische (γ) und graphematische (δ) Struktur mit mehreren ICC-Schemata übereinstimmt, wird nach einer anderen Möglichkeit der sprachlichen Realisierung oder nach einem syntaktischen oder kontextuellen Mittel der Vereindeutigung gesucht.

Das relationale Wirken der ICC-Schemata hat also einen Einfluss darauf, ob ein ICC gebildet wird oder nicht. Wird ein ICC wegen der Existenz unterschiedlicher ICC-Schemata für zu uneindeutig befunden, wird entweder von der Bildung eines solchen ICCs abgesehen, oder es begleiten Strategien der Disambiguierung und Spezifizierung, wie sie die Korpusstudien für Prot-ICCs nachgewiesen haben, das ICC. Bei der Verwendung von Prot-ICCs machen Sprecher:innen von diesen Techniken der Erklärung und Spezifizierung außerordentlich Gebrauch. Das zeigen die Ergebnisse zur Verwendung syntaktischer Mittel kontextueller Anreicherung wie der adjektivischen Attribution oder der Kontrastierung durch Negationspartikeln. Die Prot-ICCs bedürfen dieser Mittel ganz besonders, zum einen, weil die Bildungen wortintern nicht formal als solche kenntlich gemacht werden können, zum anderen, weil das Prot-ICC-Schema nach der hier präsentierten Datengrundlage nicht so stark durch Belege gefestigt ist wie die anderen beiden ICC-Schemata.

Die Bildung *Trainer-Trainer* ist beispielsweise für ein Det-ICC prädestiniert (‘Trainer für den Trainer’), auch weil der Nominalstamm eine Relation aus der Valenz des zugrundeliegenden Verbs *trainieren* erbt. Will ein:e Sprecher:in nun aber ein Prot-ICC auf der Basis von *Trainer* bilden, muss dem Gegenüber die intendierte Lesart irgendwie angezeigt werden. Bei der Basis *Trainer* ist das aber schwierig. Der Stamm erlaubt kein Fugenelement, weshalb ein Det-ICC und ein Prot-ICC auf der Basis von *Trainer* notwendigerweise formidentisch sind. Allein die Getrenntschreibung kann einen schwachen Hinweis darauf geben, dass das Prot- und nicht das Det-ICC-Schema zur Interpretation angewendet werden muss. Die Korpusdaten zeigen denn auch, dass die Sprecher:innen ein Prot-ICC auf der Basis von *Trainer* vermeiden. *Trainer-Trainer* kommt in den Daten ausschließlich als Det-ICC vor, obwohl sich das Konzept ebenfalls gut für ein Prot-ICC eignen würde, gibt es doch in vielen Sportarten ‘richtige Trainer’, also solche, die auch

einen Trainerschein besitzen, und so genannte Teamchefs, die zwar ebenfalls die Aufgabe des Trainierens übernehmen, aber keinen Trainerschein haben.[89] Der Befund aus der auf deTenTen13 basierenden Korpusstudie, dass Stämme selten mehr als einen ICC-Typ bilden, lässt sich mithilfe von Schemata, die im mentalen Lexikon der Sprecher:innen verankert sind, gut erklären. Es ist demnach das Lexikon, das dafür sorgt, dass Sprecher:innen, die ein ICC bilden wollen, antizipieren, inwieweit ihre Ad hoc-Bildung eindeutig ist. Die Bildung und Interpretation von ICCs bedingen sich also gegenseitig, weil sowohl erstere als auch letztere auf das Lexikon und die darin befindlichen Schemata zurückgreifen.

Kontextuelle Anreicherung ermöglicht darüber hinaus, Defizite bei der formalen Markierung der Bedeutung auszugleichen und Schemata trotz semantischer Ambiguität des damit gebildeten Wortes zur Wortbildung zu verwenden. In Beispiel (78) erkennt der Sprecher etwa das Prot-ICC *Schiff Schiff* als zu uneindeutig und erklärt im Anschluss, wie die Bildung zu interpretieren ist (*auf der Cap San Diego, diesem wunderbaren "**Schiff Schiff**" Als "Schiff Schiff" bezeichne ich…*). In der Konversationsanalyse spricht man hier von einer „selbstinitiierten Reparatur“ (Kitzinger 2013). Der Sprecher kommt Verständnisschwierigkeiten des Gegenübers zuvor. ICCs sind aber nicht nur Auslöser für eine solche Reparatur; sie erfüllen diese Funktion auch selbst, gehören also zum

> set of practices whereby a co-interactant interrupts the ongoing course of action to attend to possible trouble in speaking, hearing, or understanding the talk.
>
> (Kitzinger 2013: 229)

In Beispiel (75) löst etwa ein Prot-ICC die Polysemie von *Freund* auf (*ich hab einen Freund. Also… nicht jetzt einen **Freund Freund***).

Die Unterschiede der drei ICC-Schemata auf semantisch-konzeptueller Ebene gehen also teilweise mit Unterschieden auf phonologischer und orthographischer Ebene einher, sodass die Sprecher:innen die Semantik eines ICCs formal anzeigen können. Ist dem nicht so, sind ICCs also uneindeutig, haben die Sprecher:innen zwei Möglichkeiten: Entweder wählen sie eine Ausweichkonstruktion, oder sie reichern das ICC kontextuell an. Dass eine dieser beiden Strategien notwendig ist, erkennen Sprecher:innen, weil ihr mentales Lexikon über mehrere ICC-Schemata verfügt und diese miteinander in Verbindung stehen. Diese Verbindungen zwischen den formalen und den semantisch-konzeptuellen Aspekten der drei ICC-

89 Bei der Entlassung von Jürgen Klinsmann als Trainer des FC Bayern im Jahre 2009 wurde dieser Unterschied etwa ins Feld geführt, um die Verpflichtung des ‘richtigen’ Trainers Jupp Heynckes zu begründen.

Schemata beeinflussen die Bildung von ICCs und erklären die Befunde, die die hier präsentierten Korpusstudien erbracht haben.

Die Beschreibung von ICCs profitiert also vom Beschreibungsansatz der *Relational Morphology*. Andersherum sind ICCs aber auch ein Paradebeispiel dafür, dass die *Relational Morphology* ein sinvolles Beschreibungsinstrumentarium für sprachliche Strukturen bietet. Denn viele Beschreibungsansätze scheitern gerade an den extravaganten, regelbrechenden, normabweichenden Phänomenen, zu denen ICCs zweifelsohne gehören. Hier zeigt die *Relational Morphology*, wie umfassend anwendbar das Framework ist. ICCs sind Bildungen, die weitgehend nicht lexikalisiert sind, die Regeln brechen („reduplication avoider", Verfugung, Flexion, …), die formal aus zwei identischen Bestandteilen bestehen, die aber auch selbst formal identische Doppelgänger haben können, die etwas ganz anderes bedeuten (*Buchbuch*) und die zu all dem auch noch sehr ähnlich geschrieben werden. Dass die *Relational Morphology* zu ICCs trotz allem nachvollziehbare Erklärungen liefern kann, ist beeindruckend und zeigt einmal mehr die Kraft der Konstruktionsgrammatik.

11 Fazit und Ausblick

Das Hauptanliegen dieser Arbeit war die Erforschung des grammatischen Phänomens, dass N+N-Komposita im Deutschen identische Konstituenten haben können. In der einschlägigen Literatur, die im ersten Teil der Arbeit vorgestellt wurde, gibt es bereits einige Ansätze dazu, wie das Phänomen ICC einzuordnen ist und wie die Bildungen funktionieren. Die Aussagen in den meisten Arbeiten führen ICCs auf den Zufall zurück. Demnach kombinierten die Sprecher:innen von Zeit zu Zeit zwei identische Konzepte, woraus sich N+N-Komposita ergeben, die sich in nichts von anderen N+N-Komposita unterscheiden. Darüber hinaus werden ICCs mit einer Prototypenlesart beschrieben. Hier gehen die Annahmen zu Form und Funktion sowie zur Einordnung der Bildungen in der Literatur weit auseinander. Mal werden diese ICCs als Reduplikationen und mal als Komposita beschrieben. Mal werden sie als Nichtwörter (nonce formations) angesehen, die nur für die einmalige Verwendung in einem ganz bestimmten Kontext gebildet werden und deshalb nicht ins Lexikon wandern können, mal wird ihnen Kontextunabhängigkeit bescheinigt. Mal wird ihr morphosyntaktisches Verhalten als unauffällig beschrieben, mal gelten sie als extreme Abweichler, die wortintern flektieren und keine Fugenelemente und Attribute erlauben. Zu ICCs, die Eigennamen sind, gibt es überhaupt keine einschlägige Literatur. Angesichts dieser Ausgangslage gab es für die vorliegende Arbeit also viel zu tun.

Der Zweck dieser Arbeit war in erster Linie, das Phänomen ICC mithilfe großer Datenmengen empirisch aufzuarbeiten, um viele zuvor unbeantwortete Fragen zu ICCs evidenzbasiert beantworten zu können. Dies geschah im zweiten Teil der Arbeit auf der Grundlage der mithin größten Korpora der deutschen Sprache. Sämtliche N+N-Komposita dieser Korpora, bei denen Erst- und Zweitglied identisch sind, wurden erhoben und analysiert, was zu folgenden Erkenntnissen geführt hat:

- ICCs werden vor allem auf der Basis hochfrequenter Lexeme gebildet, was sich dadurch erklären lässt, dass diese in N+N-Komposita gleichermaßen in Kopf- und Modifikatorposition vorkommen und sie diese beiden Funktionen in ICCs simultan ausüben müssen. Auch werden für ICCs kurze Lexeme bevorzugt, also Lexeme mit wenigen Zeichen und Silben.
- ICCs unterscheiden sich von anderen N+N-Komposita hinsichtlich der Schreibung und durch vermehrten Gebrauch von Anführungszeichen.
- Semantisch-funktional gesehen gibt es drei Arten von ICCs:
  - ICCs, die das Basiskonzept zweimal realisieren, also im Erstglied und im Zweitglied, und somit auf dieser Ebene als gewöhnliche, endozentrische, determinative N+N-Komposita anzusehen sind (Det-ICCs)

https://doi.org/10.1515/9783111315416-011

  - ICCs, bei denen nur das Zweitglied das Basiskonzept realisiert und das Erstglied stattdessen im Zuge einer Nullmodifikation (neutrale Subklassifizierung) zu einer Prototypenbedeutung der Bildung führt (Prot-ICCs)
  - ICCs, bei denen weder Erst- noch Zweitglied das Basiskonzept realisieren, wodurch die Bildungen Eigennamen und nicht über Konzepte mit den Referenten verbunden sind (Name-ICCs)
    (Optional kann bei letzteren das Erstglied das Basiskonzept realisieren und den Namen damit teilweise motivieren.)
- Formal unterscheiden sich diese ICC-Typen voneinander:
  - Hinsichtlich ihres morphosyntaktischen Verhaltens sind Det-ICCs und Name-ICCs Gegensätze. Det-ICCs sind fast immer verfugt und overt flektiert (formal sind also auch Det-ICCs keine gewöhnlichen N+N-Komposita). Name-ICCs sind fast nie verfugt und nur selten overt flektiert.
  - Hinsichtlich ihrer Schreibung sind ebenfalls Det-ICCs und Name-ICCs Gegensätze. Die Konstituenten werden in Det-ICCs größtenteils zusammengeschrieben, in Name-ICCs hingegen durch Spatium, Syngrapheme oder Binnenmajuskeln getrennt.
  - Det-ICCs sind besonders lang, Name-ICCs besonders kurz.
  - Prot-ICCs nehmen hinsichtlich all dieser Punkte eine Zwischenposition zwischen Det- und Name-ICCs ein. Syntaktisch und kontextuell stechen sie jedoch hervor und zeigen eine deutlich häufigere Verwendung mit Artikeln, Adjektivattributen, Negationspartikeln und Prädikativkonstruktionen.
  - Die Sprecher:innen reichern Prot-ICCs kontextuell an, indem sie die Bildungen erklären, durch Adjektivattribute vereindeutigen oder sie mithilfe kontrastierender Nominalphrasen, Fokuspartikeln und Negationswörtern in ein Kontrastverhältnis zu dem setzen, was sie nicht meinen.
- Nominalstämme, die ICCs bilden, bilden in der Regel entweder Det-, Prot- oder Name-ICCs. Nur selten kommen sie in mehreren ICC-Typen vor.

Die erhobenen Korpusdaten korrigieren also viele in der Literatur getroffene Annahmen zu ICCs. Det-ICCs sind keine gewöhnlichen Komposita, da sie weitaus häufiger verfugt werden als N+N-Komposita im Allgemeinen. Auch sind Prot-ICCs morphosyntaktisch keine extremen Abweichler und zeigen durchaus Fugenelemente und auch – sogar recht häufig – attributive Adjektive. Die bisher nahezu überhaupt nicht beschriebenen Name-ICCs wurden im Rahmen dieser Arbeit erstmals dezidiert behandelt und als Eigennamen beschrieben, die viele der Eigenschaften zeigen, die im Allgemeinen mit Namen in Verbindung gebracht wer-

den: Sie haben Direkt- und Monoreferenz, ihr Wortkörper wird geschont und sie sind morphosyntaktisch und graphematisch deviant (Nullflexion, mitunter referenzielles Genus, Getrenntschreibung).

Der dritte Teil der Arbeit widmete sich schließlich der grammatiktheoretischen Einordnung von ICCs. Zunächst wurden die ICC-Typen zu den Prozessen Komposition und Reduplikation in Bezug gesetzt. Det-ICCs zeigten sich im Zuge dessen als klare Vertreter der Komposition und wurden eingedenk der reduplikativen Struktur der Bildungen als reduplikative Komposita bezeichnet. Prot-ICCs wurden Kompositionsreduplikationen genannt, da sie Merkmale beider Prozesse gleichermaßen zeigen und sich darüber hinaus vielen Ansätzen zur Beschreibung von Kompositasemantik entziehen. Name-ICCs schließlich sind zwar morphologische Verbindungen von Wortstämmen und können durchaus als Komposita angesehen werden. Der Prozess, durch den sie zustandekommen, weicht aber in vielen Punkten sowohl von Komposition als auch von Reduplikation ab, weshalb sie vorsichtiger als reduplikative Wort-, beziehungsweise Eigennamenbildungen eingeordnet werden sollten.

Nach dieser Einordnung ergründete die Arbeit, wie ICCs ihre Bedeutung erhalten und wandte dazu viele unterschiedliche Theorien und Modelle zur Kompositasemantik auf ICCs an. Dabei stellte sich heraus, dass die reduplikative Struktur der ICCs ein Problem für viele Modelle darstellt, die für Komposita unterschiedliche Konstituenten voraussetzen. Die Modelle, die mit der ungewöhnlichen Struktur von ICCs keine Probleme haben, berücksichtigen die deutlichen formalen Unterschiede zwischen den Bildungen nicht oder ignorieren die Rolle, die das Lexikon bei der Interpretation von ICCs spielt. Beides berücksichtigt der zum Ende der Arbeit entwickelte Ansatz, ICCs im Rahmen der *Relational Morphology* zu beschreiben.

Dieser Ansatz begreift die ICC-Typen als im Lexikon abgespeicherte Schemata, in denen die Bedeutung, beziehungsweise die Funktion, auf unterschiedliche Weise mit einer syntaktischen sowie einer phonologischen, beziehungsweise orthographischen Struktur verbunden ist. Die formalen Unterschiede der ICC-Typen auf phonologischer und orthographischer Ebene helfen den Adressat:innen bei der Entscheidung, auf der Grundlage welchen Schemas ein ICC zu interpretieren ist. Zusätzlich stehen die drei Schemata in Verbindung zueinander, weshalb Sprecher:innen, die ein ICC bilden wollen, eine ambige Bildung antizipieren und entweder zusätzlich kontextuell anreichern oder gänzlich vermeiden, indem sie auf Alternativkonstruktionen ausweichen.

Manche Aspekte von ICCs konnte die Arbeit aber nicht ergründen. Weil die Korpusstudien allein auf schriftlichen Daten beruhen, kann etwa zu der Annahme, dass ICCs erstgliedbetont sind, keine Aussage getroffen werden. Auch wurden

nur Bildungen beschrieben, die im Korpus vorkommen. Es konnte also keine Einschätzung dazu gegeben werden, ob die identifizierten Faktoren Länge, Flexion, Verfugung und Schreibung auch auf die Interpretation anderer ICCs einen Einfluss haben. Auch in der bisher einzigen diesbezüglichen Studie (Finkbeiner 2014) wurden diese Faktoren nicht manipuliert. Es wäre daher lohnenswert, die Studie von Finkbeiner zu replizieren und die in dieser Arbeit beschriebenen Unterschiede zwischen den ICC-Typen daraufhin zu testen, ob sie die Proband:innen in ihrer ICC-Interpretation beeinflussen. Auch die Textfunktion von ICCs konnte in dieser Arbeit nicht behandelt werden, da die Datenquellen keinen Zugriff auf die Urspungstexte ermöglichen, beziehungsweise eine solche Textanalyse aufwendig über das Zurückverfolgen der URLs hätte geschehen müssen. Aus demselben Grund kann auf der Grundlage der präsentierten Korpusdaten auch nichts über die Genese der ICC-Typen ausgesagt werden, etwa hinsichtlich der Frage, welcher ICC-Typ der älteste und welcher der jüngste ist. Aufgrund der Beschränkung auf einmalige Wiederholung der Zeichenkette blieb zudem der Prozess der Triplikation außen vor sowie die Frage, ob die Bildung von Prot-ICCs – wie in der Literatur behauptet – wirklich nicht rekursiv ist. Die vorliegende Arbeit stellt also mitnichten eine vollumfängliche Beschreibung von ICCs im Deutschen dar. Sie hat aber vermutlich dennoch etwas Licht in eine sehr dunkle Forschungsnische eines ansonsten gut erforschten Bereichs der Grammatik gebracht.

Nachwort

Die im Vorwort aufgeführten Zitate preisen allesamt die Verbindung von Identischem. Doch stehen den Zitaten mindestens ebenso viele gegenüber, die das Gegenteil behaupten, etwa *Gegensätze ziehen sich an*. Es gibt zudem zahlreiche Beispiele, etwa aus der Physik, die zeigen, dass sich vor allem Unterschiedliches oder sogar Gegensätzliches verbinden lässt. Auch bei der N+N-Komposition im Deutschen werden ungleiche Stämme sehr viel häufiger verbunden als identische. Die Bildungen, die ich untersucht habe, erscheinen den Sprecher:innen deshalb ungewöhnlich. Statt eines neuen Stammes im zweiten Bestandteil des Kompositums folgt auf den ersten Bestandteil nochmal der gleiche. Derselbe ist es zwar nicht, denn die Stämme unterscheiden sich in der Position und tragen ganz unterschiedlich zu dem gebildeten Wort bei. Dennoch zeigen die Sprecher:innen mit allen Mitteln an, etwa mit doppelten Anführungszeichen, Erklärungen oder metasprachlichen Kommentaren, dass hier etwas besonderes vorliegt. Verbindet sich bei der Komposition Identisches, fällt das offensichtlich sowohl den Sprecher:innen als auch den Rezipient:innen sofort auf. Aber warum ist das so? Müsste sich ein Nominalstamm wie *Sommer*, rein statistisch gesehen, nicht eben auch mal mit *Sommer* verbinden? Müssten nicht Komposita aus identischen Konstituenten als gewöhnliche Komposita wahrgenommen werden? Warum fallen die Sprecher:innen so sehr auf, wenn sie ICCs bilden?

Eine Antwort kann die Verquickung zweier Kleist-Texte bieten, und zwar der Texte „Über die allmähliche Verfertigung der Gedanken beim Reden“ und „Über das Marionettentheater“. Grob vereinfacht besagt der erste Text, dass man beim Sprechen am besten einfach drauf los redet; dann werde die Rede gut. Der zweite Text besagt, noch gröber vereinfacht, dass das Bewusstsein des Menschen dem Instinkt des Tieres unterlegen ist. Kleist beschreibt hier einen Bären, gegen den kein Mensch mit dem Degen bestehen kann, da er jeden Schlag instinktiv pariert. Die Kleist-Texte kann man wie folgt auf ICCs anwenden: Nachdem man den Kompositumsbestandteil *Sommer* gehört oder gelesen hat, ist man auf alles vorbereitet, aber eben nicht auf das Zweitglied *Sommer*. Wiederholt sich ein Wortbestandteil unmittelbar, horcht man auf. Es ist dieses Aufhorchen und in der Folge das Unterbrochen-worden-sein, was dem Sprachfluss der Sprechenden und dem aufmerksamen Folgen der Zuhörenden den Garaus macht. Man ist plötzlich nicht mehr Kleists fechtender Bär, redet nicht mehr allmählich verfertigend ins Blaue hinein, hört bewusst hin und beschäftigt sich mit der Form anstatt mit dem Inhalt. Das fein austarierte Verhältnis von Unterschiedlichem und Gleichem in der Sprache, es wird durch ein ICC gestört. Andersherum ist das übrigens nicht anders. In den vielfach in der Psychologie verwendeten Oddball-Experimenten führt bei-

https://doi.org/10.1515/9783111315416-012

spielsweise ein sehr selten auftretender, abweichender Zielreiz innerhalb einer Reihe identischer Reize zu erhöhter Aktivität des Gehirns. Hier ist die Abweichung der Auslöser für das Aufhorchen. Eben dieser Moment, also der Moment, in dem den Beteiligten bewusst wird, dass gerade gegen eine ungeschriebene Regel verstoßen wurde, verleiht ICCs so eine Kraft. Dass die Sprecher:innen sich dieses Phänomen des menschlichen Geistes mithilfe von ICCs zunutze machen, zeigt einmal mehr, dass auch zur Sprache zwangsläufig beides gehört: die Identität und die Differenz. Bildungen wie *Sommer Sommer* sind also eigentlich gar nicht so ungewöhnlich und zeigen bloß, was auch generell im Leben gilt: Manchmal macht eben gerade die Identität den Unterschied.

Abb. 43: Das mit Abbildung 1 identische Werbedisplay an der Leipziger S-Bahn-Haltestelle „Markt“.

Quellen und Literaturverzeichnis

Korpora

DECOW16
<http://corporafromtheweb.org/>
deTenTen13
<http://www.sketchengine.eu/detenten-german-corpus/>

Forschungsliteratur

Abraham, Werner. 2005. Intensity and diminution triggered by reduplicating morphology: Janus-faced iconicity. Bernhard Hurch (Hrsg.), *Studies on reduplication,* 547–568. Berlin: Mouton de Gruyter.

Abu-Bader, Soleman & Josephine Pryce. 2006. *Using statistical methods in social work practice: A complete SPSS guide.* Chicago: Lyceum Books.

Ackermann, Tanja. 2016. Die Morphosyntax der Personennamen im Deutschen – synchrone und diachrone Perspektiven. Dissertation Freie Universität Berlin.

Albert, Georg. 2013. *Innovative Schriftlichkeit in digitalen Texten: syntaktische Variation und stilistische Differenzierung in Chat und Forum*. Berlin: Akademie-Verlag.

Allen, Margaret Reece. 1978. Morphological investigations. Dissertation University of Connecticut.

Andreou, Marios. 2017. Stereotype negation in Frame Semantics. *Glossa: a journal of general linguistics* 2(1). 1–30.

Arcodia, Giorgio Francesco. 2007. Chinese: A language of compound words. *Selected proceedings of the 5th décembrettes: Morphology in Toulouse*. 79–90.

Aronoff, Mark. 1976. *Word formation in generative grammar.* Cambridge, MA: MIT Press.

Austefjord, Anders. 1979. Zur Vorgeschichte des germanischen starken Präteritums. *Zeitschrift für Indogermanistik und historische Sprachwissenschaft* 84. 208–215.

Aziz, Zulfadli A & Vivi Nolikasari. 2020. Reduplication as a word-formation process in the Jamee Language: A variety of Minang spoken in South Aceh. *Studies in English Language and Education* 7(1). 43–54.

Baayen, R. Harald. 1989. A corpus-based approach to morphological productivity. Dissertation Vrije Universiteit Amsterdam.

Baayen, R. Harald. 1992. Quantitative aspects of morphological productivity. In Geert Booij & Jaap van Marle (Hrsg.), *Yearbook of morphology 1991*, 109–149. Dordrecht: Kluwer Academic Publishers.

Baayen, R. Harald. 2003. Probabilistic approaches to morphology. In Rens Bod, Jennifer Hay & Stefanie Jannedy (Hrsg.), *Probabilistic Linguistics,* 229–287. Cambridge, MA: MIT Press.

Baayen, R. Harald. 2009. 43. Corpus linguistics in morphology: morphological productivity. In Anke Lüdeling & Merja Kytö (Hrsg.), *Corpus linguistics. An international handbook*, 900–919. Berlin & New York: De Gryuter.

Baayen, R. Harald. 2010. The directed compound graph of English: An exploration of lexical connectivity and its processing consequences. *New impulses in word-formation* 17. 383–402.

Backhaus, Klaus, Bernd Erichson & Rolf Weiber. 2015. *Fortgeschrittene multivariate Analysemethoden: eine anwendungsorientierte Einführung.* Berlin: Springer.

https://doi.org/10.1515/9783111315416-013

Bai, Chen, Ina Bornkessel-Schlesewsky, Luming Wang, Yu-Chen Hung, Matthias Schlesewsky & Petra Burghardt. 2008. Semantic composition engenders an N400: evidence from Chinese compounds. *Neuroreport* 19(6). 695–699.

Baroni, Marco, Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta. 2009. The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation* 43(3). 209–226.

Barz, Irmhild. 2009. Die Wortbildung. In Dudenredaktion (Hrsg.), *Duden. Die Grammatik*, 634–762. Mannheim, Wien & Zürich: Dudenverlag.

Barz, Irmhild. 2016. German. In Peter Müller, Ingeborg Ohnheiser, Susan Olsen & Rainer Franz (Hrsg.), *Word–formation. An international handbook of the languages of Europe*, 2387–2410. Berlin & Boston: Walter de Gruyter.

Barz, Irmhild, Marianne Schröder, Karin Hämmer & Hannelore Poethe. 2003. *Wortbildungspraktisch und integrativ: ein Arbeitsbuch*. Frankfurt am Main u.a.: Lang.

Bauer, Laurie. 1978. *The grammar of nominal compounding with special reference to Danish, English, and French*. Odense: Odense University Press.

Bauer, Laurie. 1979. On the need for pragmatics in the study of nominal compounding. *Journal of Pragmatics* 3(1). 45–50.

Bauer, Laurie. 1998. When is a sequence of two nouns a compound in English? *English Language & Linguistics* 2(1). 65–86.

Bauer, Laurie. 2000. System vs. norm: coinage and institutionalization. *Morphologie/Morphology. HSK* 17(1). 832–840.

Bauer, Laurie. 2001a. Compounding. In Martin Haspelmath, Ekkehard König, Wulf Österreicher & Wolfgang Raible (Hrsg.), *Language Typology and Language Universals*, 695–707. Berlin & New York: De Gruyter.

Bauer, Laurie. 2001b. *Morphological productivity*. Cambridge: Cambridge University Press.

Bauer, Laurie. 2003. *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press.

Bauer, Laurie. 2006a. Compound. In Keith Brown (Hrsg.), *Encyclopedia of Language and Linguistics*, 719–726. Boston, MA: Elsevier.

Bauer, Laurie. 2006b. Compounds and minor word-formation types. In Bas Aarts & April McMahon (Hrsg.), *The handbook of English linguistics*, 483–506. Malden, MA: Blackwell.

Bauer, Laurie, Natalia Beliaeva & Elizaveta Tarasova. 2019. Recalibrating productivity: Factors involved. *Zeitschrift für Wortbildung* 3(1). 44–80.

Beck, Klaus. 2006. *Computervermittelte Kommunikation im Internet.* München u.a.: Oldenbourg.

Becker, Thomas. 1992. Compounding in German. *Rivista di Linguistica* 4(1). 5–36.

Beißwenger, Michael & Angelika Storrer. 2008. Corpora of computer-mediated communication. In Anke Lüdeling & Merja Kytö (Hrsg.), *Corpus Linguistics*, 292–309. Berlin: De Gruyter.

Bell, Melanie J. 2012. The English NN construct: Its prosody and structure. Dissertation University of Cambridge.

Bell, Melanie J. & Martin Schäfer. 2013. Semantic transparency: challenges for distributional semantics. *Proceedings of the IWCS 2013 Workshop Towards a Formal Distributional Semantics*. 1–10.

Bell, Melanie J. & Martin Schäfer. 2016. Modelling semantic transparency. *Morphology* 26(2). 157–199.

Benczes, Réka. 2004. Analysing exocentric compounds in English: a case for creativity. *The even yearbook* 6. 11–18.

Benczes, Réka. 2006. *Creative compounding in English*. Amsterdam & Philadelphia: Benjamins.

Benjamin, Brandon Lee. 2018. Identical constituent compounding: a conceptual integration-based model. Masterarbeit Case Western Reserve University.

Berko-Gleason, Jean. 1958. The child's learning of English morphology. *Word* 14. 150–177.

Bisetto, Antonietta & Sergio Scalise. 2005. The classification of compounds. *Lingue e linguaggio* 4(2). 319–332.

Blauth-Henke, Christine. 2008. Reduplikation in Korpora. Zum Zusammenhang von Methodenreflexion und Forschungsgegenstand. In Steffen Buch, Alvaro Ceballos Viro & Christian Gerth (Hrsg.), *Selbstreflexivität: Beiträge zum 23. Forum Junge Romanistik Göttingen, 30. Mai-0.2 Juni 2007*, 35–50. Bonn: Romanistischer Verlag.

Bliß, Anna. 2017. Die syntaktische Kategorisierung von *als*. In Sandra Döring & Jochen Geilfuß-Wolfgang (Hrsg.), *Probleme der syntaktischen Kategorisierung. Einzelgänger, Außenseiter und mehr*,193–218. Tübingen: Stauffenburg.

Bloomfield, Leonard. 1933. *Language*. Chicago: Holt.

Bollée, Annegret. 1978. Reduplikation und Iteration in den romanischen Sprachen. *Archiv für das Studium der neuen Sprachen und Literaturen* 215. 318–336.

Booij, Geert. 2005. *The grammar of words: An introduction to linguistic morphology*. Oxford: Oxford University Press.

Booij, Geert. 2009. Compounding and construction morphology. In Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of compounding*, 322–347. Oxford: Oxford University Press.

Booij, Geert. 2010a. Compound construction: Schemas or analogy? A construction morphology perspective. Sergio Scalise & Irene Vogel (Hrsg.), *Cross-disciplinary issues in compounding*, 93–108. Amsterdam, Philadelphia: Benjamins.

Booij, Geert. 2010b. *Construction morphology*. Oxford: Oxford University Press.

Booij, Geert. 2012. *The grammar of words: An introduction to linguistic morphology*. Oxford: Oxford University Press.

Booij, Geert & Jenny Audring. 2017. Construction Morphology and the Parallel Architecture of Grammar. *Cognitive Science* 41. 277–302.

Borgert, Udo & Charles Anthony Nyhan. 1975. *A German reference grammar*. Sydney: Sydney University Press.

Borgwaldt, Susanne. 2013. Fugenelemente und Bindestriche in neugebildeten NN-Komposita. In Martin Neef & Carmen Scherer (Hrsg.), *Die Schnittstelle von Morphologie und geschriebener Sprache*, 103–134. Berlin & Boston: De Gruyter.

Botha, Rudolf P. 1968. *The function of the lexicon in transformational generative grammar.* Den Haag: Mouton.

Botha, Rudolf P. 1988. *Form and meaning in word formation: a study of Afrikaans reduplication.* Cambridge & New York: Cambridge University Press.

Breindl, Eva, Maria Thurmair. 1992. Der Fürstbischof im Hosenrock: eine Studie zu den nominalen Kopulativkomposita des Deutschen. *Deutsche Sprache: Zeitschrift für Theorie, Praxis, Dokumentation* 20. 32–61.

Brekle, Herbert E. 1986. The production and interpretation of ad hoc nominal compounds in German: A realistic approach. *Acta Linguistica Academiae Scientiarum Hungaricae* 36(1/4). 39–52.

Bross, Fabian & Katherine Fraser. 2020. Contrastive focus reduplication and the modification puzzle. *Glossa: a journal of general linguistics* 5(1). 1–18.

Brown, Penelope. 1999. Repetition. *Journal of Linguistic Anthropology* 9(1–2). 223–226.

Brunner, Annelen, Stefan Engelberg & Katrin Hein. 2021. The distribution of constituent words in nominal compounds and its impact on semantic interpretation: an empirical study. *Zeitschrift für Wortbildung/Journal of Word Formation* 5(1). 6–36.

Buchmann, Franziska. 2015. *Die Wortzeichen im Deutschen*. Heidelberg: Winter.

Busa, Federica. 1997. Compositionality and the Semantics of Nominals. Dissertation Brandeis University.

Busch, Albert & Oliver Stenschke. 2007. *Germanistische Linguistik*. Tübingen: Narr.
Bußmann, Hadumod. 2006. *Routledge dictionary of language and linguistics*. London & New York: Routledge.
Bußmann, Hadumod. 2008. *Lexikon der Sprachwissenschaft.* Stuttgart: Kröner.
Bzdęga, Andrzej Zdzisław. 1965. *Reduplizierte Wortbildung im Deutschen*. Poznań: Praca Wydana z Zasiłku Polskiej Akademii Nauk.
Cardona, Giorgio Raimondo. 1988. *Dizionario di linguistica*. Rom: Armando.
Carston, Robyn. 2002. *Thoughts and utterances: the pragmatics of explicit communication*. Malden, MA. u.a.: Blackwell.
Carter, Ronald. 1999. Common language: Corpus, creativity and cognition. *Language and Literature* 8(3). 195–216.
Carter, Ronald. 2004. *Language and Creativity: The Art of Common Talk*. London & New York: Routledge.
Carter, Ronald. 2007. Creativities across texts and values. Paper presented at an AHRC-funded seminar on *Transitions and Transformations: Exploring Creativity in Everyday and Literary language*. The Open University, Milton Keynes, UK, 16 March.
Chambers, John M., William S. Cleveland, Beat Kleiner & Paul A. Tukey. 1983. *Graphical methods for data analysis.* Belmont, CA: Wadsworth & Brooks.
Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.
Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel J. Keyser (Hrsg.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–51. Cambridge MA: MIT Press.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Cohen, Joshua. 1988. *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.
Consten, Manfred. 2004. *Anaphorisch oder deiktisch? Zu einem integrativen Modell domänengebundener Referenz.* Tübingen: Niemeyer.
Copestake, Ann A, Fabre Lambeau, Aline Villavicencio, Francis Bond, Timothy Baldwin, Ivan A. Sag & Dan Flickinger. 2002. Multiword expressions: linguistic precision and reusability. *Proceedings of the 3rd international conference on Language Resources and Evaluation LREC*. 1–7.
Costello, Fintan J. & Mark T. Keane. 2005. Compositionality and the pragmatics of conceptual combination. In Edouard Machery, Markus Werning & Gerhard Schurz (Hrsg.), *The compositionality of meaning and content. Volume II. Applications to linguistics, psychology and neuroscience*, 203–216. Frankfurt a.M. Ontos.
Coulmas, Florian. 1988. Wörter, Komposita und anaphorische Inseln. *Folia Linguistica* 22(3–4). 315–336.
Coulson, Seana & Gilles Fauconnier. 1999. Fake guns and stone lions: Conceptual blending and privative adjectives. In Barbara A. Fox, Dan Jurafsky & Laura A. Michaelis (Hrsg.), *Cognition and function in language*, 143–158. Cambridge: Cambridge University Press.
Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective.* Oxford: Oxford University Press.
Culicover, Peter W., & Ray Jackendoff. 2012. Same-except: A domain-general cognitive relation and how language expresses it. *Language* 88 (2). 305–340.
Dattalo, Patrick. 2013. *Analysis of multiple dependent variables.* Oxford: Oxford University Press.
Davies, Mark. 2010. The corpus of contemporary American English as the first reliable monitor corpus of English. *Literary and linguistic computing* 25 (4). 447–464.
Davies, Mark. 2012. Comparing the corpus of American soap operas, COCA, and the BNC. https://corpus.byu.edu/coca/compare-bnc.asp.

Debus, Friedhelm. 1977. Soziale Veränderungen und Sprachwandel: Moden im Gebrauch von Personennamen. In Hugo Moser (Hrsg.), *Sprachwandel und Sprachgeschichtsschreibung*, 167–204. Düsseldorf: Schwann.

Diewald, Gabriele. 1997. *Grammatikalisierung: Eine Einführung in Sein und Werden grammatischer Formen.* Berlin: De Gruyter.

Dirven, René & Majorie Verspoor. 1998. *Cognitive Exploration of Language and Linguistics*. Amsterdam: Benjamins.

Donalies, Elke. 2005. *Die Wortbildung des Deutschen.* Tübingen: Narr.

Donalies, Elke. 2011. *Basiswissen Deutsche Wortbildung.* Tübingen u.a.: Francke.

Dornseiff, Franz. 1934. *Der deutsche Wortschatz nach Sachgruppen.* Berlin: De Gruyter.

Dornseiff, Franz, Uwe Quasthoff & Herbert Ernst Wiegand. 2010. *Der deutsche Wortschatz nach Sachgruppen*. Berlin u.a.: De Gruyter.

Downing, Pamela. 1977. On the creation and use of English compound nouns. *Language* 53(4). 810–842.

Dray, Nancy. 1987. Doubles and modifiers in English. Masterarbeit University of Chicago.

Dressler, Wolfgang U. 2006. Compound types. In Gary Libben & Gonia Jarema (Hrsg.), *The representation and processing of compound words*, 23–44. Oxford: Oxford University Press.

Dressler, Wolfgang U., Gary Libben, Jacqueline Stark, Christiane Pons & Gonia Jarema. 2001. The processing of interfixed German compounds. In Geert Booij & Japp van Marle (Hrsg.), *Yearbook of morphology 1999*, 185–220. Amsterdam: Springer.

Dressler, Wolfgang U. & Karlheinz Mörth. 2012. Produktive und weniger produktive Komposition in ihrer Rolle im Text an Hand der Beziehungen zwischen Titel und Text. In Livio Gaeta & Barbara Schlücker (Hrsg.), *Das Deutsche als kompositionsfreudige Sprache. Strukturelle Eigenschaften und systembezogene Aspekte*, 219–234. Berlin & New York: De Gruyter.

Dudenredaktion. 2016. *Das Wörterbuch der sprachlichen Zweifelsfälle: Richtiges und gutes Deutsch, 8., vollständig überarbeitete Auflage.* Berlin: Bibliographisches Institut.

Dürscheid, Christa. 2005. Normabweichendes Schreiben als Mittel zum Zweck. *Muttersprache* 115(1). 40–53.

Eichinger, Ludwig M. 2000. *Deutsche Wortbildung: eine Einführung.* Tübingen: Narr.

Eisenberg, Peter. 2002. Struktur und Akzent komplexer Komposita. In David Restle & Dietmar Zaefferer (Hrsg.), *Sounds and Systems. Studies in Structure and Change. A Festschrift for Theo Vennemann*, 349–365. Berlin & New York: De Gruyter.

Eisenberg, Peter. 2006. *Grundriss der deutschen Grammatik: Band 1: Das Wort*. Stuttgart: Metzler.

Eisenberg, Peter. 2020. *Grundriss der deutschen Grammatik: Band 1: Das Wort*. Stuttgart: Metzler.

Elsen, Hilke. 2014. *Grundzüge der Morphologie des Deutschen.* Berlin & Boston: De Gruyter.

Engel, Ulrich. 1996. *Deutsche Grammatik. 3., korrigierte Auflage.* Heidelberg: Groos.

Erbaşı, Betül. 2018. Turkish doubled verbs as doubled TPs. In Rita Finkbeiner & Ulrike Freywald (Hrsg.), *Exact repetition in grammar and discourse*, 182–199. Berlin & Boston: De Gruyter.

Erben, Johannes. 1981. Neologismen im Spannungsfeld von System und Norm. In Brigitte Schlieben-Lange (Hrsg.), *Geschichte und Architektur der Sprachen*, 35–43. Berlin & New York: De Gruyter.

Erben, Johannes. 1988. Vergleichsurteile und Vergleichsstrukturen im Deutschen. *Sprachwissenschaft* 13 (3). 309–329.

Erman, Britt & Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text & Talk* 20 (1). 29–62.

Evrard, Muriel. 2002. Ageing and lexical access to common and proper names in picture naming. *Brain and Language* 81(1–3). 174–179.

Fahlbusch, Fabian & Damaris Nübling. 2014. *Der Schauinsland – die Mobiliar – das Turm*. Das referentielle Genus bei Eigennamen und seine Genese. *Beiträge zur Namenforschung* 49(3). 245–288.
Fandrych, Christian & Maria Thurmair. 1994. Ein Interpretationsmodell für Nominalkomposita: linguistische und didaktische Überlegungen. *Deutsch als Fremdsprache* 31. 34–45.
Fanselow, Gisbert. 1981. *Zur Syntax und Semantik der Nominalkomposition On the syntax and semantics of nominal compounds*. Tübingen: Niemeyer.
Fernández-Domínguez, Jesús. 2010. N+N compounding in English: Semantic categories and the weight of modifiers. *Brno Studies in English* 36(1). 47–76.
Fiedler, Sabine. 2007. *English phraseology: A coursebook*. Tübingen: Narr.
Finkbeiner, Rita. 2012. *Naja, normal und normal.* Zur Syntax, Semantik und Pragmatik der x-und-x-Konstruktion im Deutschen. *Zeitschrift für Sprachwissenschaft* 31. 1–42.
Finkbeiner, Rita. 2014. Identical constituent compounds in German. *Word Structure* 7(2). 182–213.
Firth, John Rupert. 1957. *Papers in Linguistics 1934–1951.* London: Longmans.
Fleischer, Wolfgang & Irmhild Barz. 1995. *Wortbildung der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
Fleischer, Wolfgang; & Irmhild Barz. 2012. *Wortbildung der deutschen Gegenwartssprache*. Berlin & Boston: De Gruyter.
Foolen, Ad. 2004. Expressive binominal NPs in Germanic and Romance languages. In Günter Radden &, Klaus-Uwe Panther (Hrsg.), *Studies in linguistic motivation*, 75–100. Berlin: De Gruyter.
Frankowsky, Maximilian. 2022. Extravagant expressions denoting quite normal entities: Identical constituent compounds in German. In Dagmar Haumann & Matthias Eitelmann (Hrsg.), *Extravagant Morphology*, 155–179. Amsterdam & Philadelphia: Benjamins.
Frankowsky, Maximilian & Barbara Schlücker (angenommen): Name-based lexical patterns and lexical creativity. In Natalia Filatkina & Sabine Arndt-Lappe (Hrsg.), *Dynamics at the lexicon-syntax interface: creativity and routine in word formation and multi-word expressions*. Berlin & Boston: De Gruyter Mouton.
Freywald, Ulrike. 2015. Total Reduplication as a productive process in German. *Studies in Language* 39(4). 905–945.
Freywald, Ulrike & Rita Finkbeiner. 2018. Exact Repetition or total reduplication? Exploring their boundaries discourse and grammar. Rita Finkbeiner & Ulrike Freywald (Hrsg.), *Exact repetition in grammar and discourse*, 3–28. Berlin & Boston: De Gruyter.
Fuhrhop, Nanna. 1998. *Grenzfälle morphologischer Einheiten.* Tübingen: Stauffenburg.
Fuhrhop, Nanna. 2000. Zeigen Fugenelemente die Morphologisierung von Komposita an? In Rolf Thieroff, Matthias Tamrat, Nanna Fuhrhop & Oliver Teuber (Hrsg.), *Deutsche Grammatik in Theorie und Praxis*, 201–213. Berlin & Boston: Niemeyer.
Fuhrhop, Nanna. 2008. Das graphematische Wort im Deutschen. eine erste Annäherung. *Zeitschrift für Sprachwissenschaft* 27(2). 189–228.
Gaeta, Livio & Davide Ricca. 2009. Composita solvantur: Compounds as lexical units or morphological objects? *Rivista di Linguistica* 21(1), 35–70.
Gagné, Christina. 2002. Lexical and relational influences on the processing of novel compounds. *Brain and Language* 81(1–3). 723–735.
Gagné, Christina L. & Edward J. Shoben. 1997. Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of experimental psychology: Learning, memory, and cognition* 23(1). 71–87.
Gagné, Christina L. & Thomas L. Spalding. 2009. Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language* 60(1). 20–35.

Geiler von Kaysersberg, Johannes. 1518. *Das Buch der Sünden des Munds*. Straßburg: Grüninger.
Ghomeshi, Jila, Ray Jackendoff, Ray, Nicole Rosen & Kevin Russell. 2004. Contrastive focus reduplication in English The salad-salad paper. *Natural Language & Linguistic Theory* 22(2). 307–357.
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. 2006. Attribution in English and the distinction between phrases and compounds. Petr Rösel (Hrsg.), *Englisch in Zeit und Raum – English in Time and Space: Forschungsbericht für Klaus Faiss*, 10–27. Trier: Wissenschaftlicher Verlag.
Giegerich, Heinz J. 2015. *Lexical structures: compounding and the modules of grammar*. Edinburgh: Edinburgh University Press.
Giessner, Steffen R. & Gabriele Jacobs. 2014. Gruppenprozesse: Identität und Prototypikalität. In Jörg Felfe (Hrsg.), *Trends der psychologischen Führungsforschung: Neue Konzepte, Methoden und Erkenntnisse*, 117–128. Göttingen u.a.: Hogrefe.
Gil, David. 2005. From repetition to reduplication in Riau Indonesian. In Bernhard Hurch (Hrsg.), *Studies on reduplication*, 31–64. Berlin u.a.: Mouton de Gruyter.
Givón, Talmy. 1993. Coherence in text, coherence in mind. *Pragmatics & Cognition* 1(2). 171–227.
Glück, Helmut. 2010. *Metzler Lexikon Sprache*. Stuttgart u.a.: Metzler.
Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
Goldhahn, Dirk, Thomas Eckart & Uwe Quasthoff. 2012. Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. *Proceedings of the Eighth International Conference on Language Resources and Evaluation LREC'12.*
Goodwin Gómez, Gale & Hein van der Voort. 2014. Reduplication in South America: An Introduction. Gale Goodwin Gómez & Hein van der Voort (Hrsg.), *Reduplication in Indigenous Languages of South America*, 1–16. Leiden & Boston: Brill.
Grube, Henner. 1976. Die Fugenelemente in neuhochdeutschen appellativischen Komposita. *Sprachwissenschaft* 1. 187–222.
Grzega, Joachim. 2004. Ein *Spitzenpolitiker* ist nicht immer ein *Spitzer-Politiker*. Wie man prosodische Akzente nutzt, um semantische ‘Akzente’ zu setzen. *Muttersprache* 114(4). 321–332.
Gülich, Elisabeth & Thomas Kotschi. 1996. Textherstellungsverfahren in mündlicher Kommunikation. Ein Beitrag am Beispiel des Französischen. In Wolfgang Motsch (Hrsg.), *Ebenen der Textstruktur: sprachliche und kommunikative Prinzipien*, 37–80. Tübingen: Niemeyer.
Gunkel, Lutz & Gisela Zifonun. 2009. Classifying modifiers in common names. *Word Structure* 2(2). 205–218.
Gunkel, Lutz & Gisela Zifonun. 2011. Klassifikatorische Modifikation im Deutschen und Französischen. In Eva Lavric, Wolfgang Pöckl & Florian Schallhart (Hrsg.), *Comparatio delectat: Akten der Vi. internationalen Arbeitstagung zum romanisch-deutschen und innerromanischen Sprachvergleich: Innsbruck, 3.–5. September 2008*, 549–562. Frankfurt a. M. Lang.
Günther, Hartmut. 1979. N+N: Untersuchungen zur Produktivität eines deutschen Wortbildungstyps. In Leonhard Lipka & Hartmut Günther (Hrsg.), *Wortbildung*, 258–280. Darmstadt: Wissenschaftliche Buchgesellschaft.
Haase, Martin. 1989. *Komposition und Derivation: Ein Kontinuum der Grammatikalisierung*. Köln: Universitätsverlag.
Hacken, Pius. 2019. *Word formation in Parallel Architecture: The case for a separate component.* Chad: Springer International Publishing.
Hagège, Claude. 1982. *La structure des langues*. Paris: Presses universitaires de France.

Hale, Ken & Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure.* Cambridge MA.: MIT press.
Halle, Morris. 1973. Prolegomena to a theory of word formation. *Linguistic Inquiry* 4(1). 3–16.
Harris, Zellig S. 1954. Distributional structure. *Word* 10(2-3). 146–162.
Haspelmath, Martin. 1999. Why is grammaticalization irreversible? *Linguistics* 37(6). 1043–1068.
Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold Publishers.
Haspelmath, Martin. 2006. Against markedness and what to replace it with. *Journal of Linguistics* 42(1). 25–70.
Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19. 1–33.
Haspelmath, Martin & Uri Tadmor. 2009. *Loanwords in the world's languages: a comparative handbook.* New York: Mouton de Gruyter.
Hatcher, Anna Granville. 1960. An introduction to the analysis of English noun compounds. *Word* 16(3). 356–373.
Hawkins, John A. 1978. *Definiteness and Indefiniteness: A study in reference and grammaticality prediction.* London: Croom Helm.
Hawkins, John A. 1984. A note on referent identifiability and co-presence. *Journal of Pragmatics* 8(5–6). 649–659.
Hay, Jennifer & Harald Baayen. 2002. Parsing and productivity. In Geert Booij & Jaap van Marle (Hrsg.), *Yearbook of morphology 2001*, 203–235. Dordrecht: Springer.
Heide, Judith. 2010. Warum *vertragen* anders ist als *vergiften* und *vergessen.* Ein Einblick in unser mentales Lexikon. *Spektrum Patholinguistik* 3. 71–88.
Hein, Katrin. 2015. *Phrasenkomposita im Deutschen: empirische Untersuchung und konstruktionsgrammatische Modellierung.* Tübingen: Narr Francke Attempto Verlag.
Hentschel, Elke & Petra M. Vogel. 2009. *Deutsche Morphologie*. Berlin & New York: De Gruyter.
Hentschel, Elke & Harald Weydt. 2013. *Handbuch der deutschen Grammatik, 4., vollständig überarbeitete Auflage.* Berlin & Boston: De Gruyter.
Heringer, Hans-Jürgen. 1984. Wortbildung: Sinn aus dem Chaos. *Deutsche Sprache* 12. 1–13.
von Heusinger, Klaus. 2010. Zur Grammatik indefiniter Eigennamen. *Zeitschrift für germanistische Linguistik* 38 (1). 88–120.
von Heusinger, Klaus. 2012. Referentialität, Spezifizität und Diskursprominenz im Sprachvergleich am Beispiel von indefiniten Demonstrativpronomen. Lutz Gunkel & Gisela Zifonun (Hrsg.), *Deutsch im Sprachvergleich: Grammatische Kontraste und Konvergenzen*, 417–456. Berlin: De Gruyter.
Hogg, Michael A., Louise Cooper-Shaw & David W. Holzworth. 1993. Group prototypically and depersonalized attraction in small interactive groups. *Personality and Social Psychology Bulletin* 19(4). 452–465.
Hohenhaus, Peter. 1996. *Ad-hoc-Wortbildung: Terminologie, Typologie und Theorie kreativer Wortbildung im Englischen.* Frankfurt am Main u.a.: Lang.
Hohenhaus, Peter. 1998. Non-Lexicalizability – As a characteristic feature of nonce-formations in English and German. *Lexicology* 4(2). 241–260.
Hohenhaus, Peter. 2004. Identical constituent compounding – a corpus-based study. *Folia Linguistica* 38(3–4). 297–331.
Hohenhaus, Peter. 2005. Lexicalization and institutionalization. In Pavol Stekauer & Rochelle Lieber (Hrsg.), *Handbook of word-formation*, 353–373. Dordrecht: Springer.
Hohenhaus, Peter. 2007. How to do even more. things with nonce words other than naming. Judith Munat (Hrsg.), *Lexical creativity, texts and contexts*, 15–38. Amsterdam & Philadelphia: Benjamins.

Hohenhaus, Peter. 2015. Anti-naming through non-word-formation. *Skase Journal of Theoretical Linguistics* 12(3). 272–292.
Hopper, Paul J. & Sandra A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60 (4). 703–752.
Hopper, Paul J. & Sandra A. Thompson. 1985. The iconicity of the universal categories 'noun' and 'verb'. In John Haiman (Hrsg.), *Iconicity in syntax*, 151–183. Stanford: John Benjamins.
Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge & New York: Cambridge University Press.
Horn, Laurence. 1993. Economy and redundancy in a dualistic model of natural language. *SKY: 1993 Yearbook of the Linguistic Association of Finland*. 33–72.
Horn, Laurence. 2002. Uncovering the un-word: A study in lexical pragmatics. *Sophia linguistica* 49. 1–64.
Horn, Laurence. 2005. An un-paper for the un-syntactician. In Salikoko Mufwene, Elaine J. Francis & Rebecca S. Wheele (Hrsg.), *Polymorphous Linguistics: Jim McCawley's Legacy*, 329–365. Cambridge, MA: MIT Press.
Horn, Laurence. 2006. Speaker and hearer in neo-Gricean pragmatics. *Journal of Foreign Languages* 164(4). 2–26.
Horn, Laurence. 2018. The lexical clone: Pragmatics, prototypes, productivity. In Rita Finkbeiner & Ulrike Freywald (Hrsg.), *Exact repetition in grammar and discourse*, 233–264. Berlin & Boston: De Gruyter.
Hosmer, David, Stanley Lemeshow & Rodney Sturdivant. 2013. *Applied logistic regression*. Hoboken: John Wiley & Sons.
Huang, Yan. 2015. Lexical cloning in English: A neo-Gricean lexical pragmatic analysis. *Journal of Pragmatics* 86. 80–85.
Hurch, Bernhard. 2005. Graz database on reduplication. http://reduplication.uni-graz.at/redup/.
Hurch, Bernhard, Motomi Kajitani, Veronika Mattes, Ursula Stangel & Ralf Vollmann. 2008. Other reduplication phenomena. Universität Graz: http://reduplication.uni-graz.at
Inkelas, Sharon. 2012. Reduplication. In Jochen Trommer (Hrsg.), *The morphology and phonology of exponence*, 355–378. Oxford: Oxford University Press.
Inkelas, Sharon & Cheryl Zoll. 2005. *Reduplication: doubling in morphology*. Cambridge: Cambridge University Press.
van Jaarsveld, Henk J. & Gilbert E. Rattink. 1988. Frequency effects in the processing of lexicalized and novel nominal compounds. *Journal of Psycholinguistic Research* 17(6). 447–473.
Jackendoff, Ray. 2002. *Foundations of language: brain, meaning, grammar, Evolution*. Oxford: Oxford University Press.
Jackendoff, Ray. 2008. 'Construction after Construction' and its theoretical challenges. *Language* 84(1). 8–28.
Jackendoff, Ray. 2009. Compounding in the parallel architecture and conceptual semantics. In Rochelle Lieber & Pavol Stekauer (Hrsg.), *The Oxford handbook of compounding*, 162–203. Oxford: Oxford University Press.
Jackendoff, Ray. 2010a. The ecology of English noun-noun compounds. In Ray Jackendoff (Hrsg.): *Meaning and the lexicon*, 413–451. Oxford: Oxford University Press.
Jackendoff, Ray. 2010b. *Meaning and the lexicon: The Parallel Architecture 1975-2010*. Oxford: Oxford University Press.
Jackendoff, Ray. 2016. English noun-noun compounds in Conceptual Semantics. In Pius Ten Hacken (Hrsg.), *The semantics of compounding*, 15–37. Cambridge: Cambridge University Press.

Jackendoff, Ray & Jenny Audring. 2020a. Relational Morphology: A cousin of construction grammar. *Frontiers in Psychology* 11(2241). 1–12.

Jackendoff, Ray & Jenny Audring. 2020b. *The texture of the lexicon: relational morphology and the parallel architecture.* Oxford: Oxford University Press, USA.

Jakobson, Roman. 1960. Closing Statement: Linguistics and Poetics. In Thomas A. Sebeok (Hrsg.), *Style in Language*, 350–377. New York: Wiley.

Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý & Vít Suchomel. 2013. The TenTen Corpus Family. *Proceedings of the 7th International Corpus Linguistics Conference.* 125–127.

Jakubíček, Miloš, Adam Kilgarriff, Diana McCarthy & Pavel Rychlý. 2010. Fast syntactic searching in very large corpora for many languages. *PACLIC*. 741–747.

James, William. 1890. *The principles of psychology. Henry Holt and Company Reprint 1950*. New York: Dover Books.

Jespersen, Otto. 1942. *A modern English grammar on historical principles. 6. Morphology*. London, Kopenhagen: Munksgaard.

Johnston, Michael & Federica Busa. 1999. Qualia structure and the compositional interpretation of compounds. In Evelyne Viegas (Hrsg.), *Breadth and depth of semantic lexicons*, 167–187. Dordrecht, Boston: Kluwer.

Kageyama, Taro. 2009. Isolate: Japanese. In Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of compounding*, 512–526. Oxford: Oxford University Press.

Kanngießer, Siegfried. 1987. Kontingenzen der Komposition. In Herbert Ernst Brekle (Hrsg.), *Neuere Forschungen zur Wortbildung und Historiographie der Linguistik. Festgabe für Herbert E. Brekle zum 50. Geburtstag*, 3–30. Tübingen: Narr.

Karttunen, Lauri. 1969. Discourse referents. *Proceedings of the 1969 International Conference on Computational Linguistics COLING. Stockholm: ACL*. 363–385.

Kauffman, Charles A. 2015. *Reduplication reflects uniqueness and innovation in language, thought and culture.* Pennsylvania: York College of Pennsylvania.

Kavka, Stanislav. 2009. Compounding and idiomatology. Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of compounding*, 19–33. Oxford: Oxford University Press.

Kay, Paul & Karl Zimmer. 1978. On the semantics of compounds and genitives in English. Paper presented at the sixth California linguistics Association. San Diego State University. 249–256.

Kazakovskaya, Viktoria V. 2017. Acquisition of nominal compounds in Russian. In Wolfgang U. Dressler, F. Nihan Ketrez & Marianne Kilani-Schoch (Hrsg.), *Nominal Compound Acquisition*, 63–90. Amsterdam: Benjamins.

Keller, Rudi. 1995. Zeichentheorie: zu einer Theorie semiotischen Wissens. Tübingen u.a.: Francke.

Kempf, Luise, Damaris Nübling & Mirjam Schmuck. 2020. Warum eine Linguistik der Eigennamen? In Luise Kempf, Damaris Nübling & Mirjam Schmuck (Hrsg.), *Linguistik der Eigennamen*, 1–18. Berlin & Boston: De Gruyter.

Kentner, Gerrit. 2017. On the emergence of reduplication in German morphophonology. *Zeitschrift für Sprachwissenschaft* 36(2). 234–277.

Kilgarriff, Adam, Pavel Rychly, Pavel Smrz & David Tugwell. 2004. The Sketch Engine. *Proceedings of the XI Euralex International Congress. Lorient: Université de Bretagne Sud*. 105–116.

Kirchner, Jesse Saba. 2010. Minimal Reduplication. Dissertation University of California Santa Cruz.

Kitzinger, Celia. 2013. Repair. In Jack Sidnell & Tanya Stivers (Hrsg.), *The handbook of conversation analysis*, 229–256. Oxford: Wiley-Blackwell.

Klein, Wolfgang & Alexander Geyken. 2010. Das digitale Wörterbuch der deutschen Sprache DWDS. *Lexicographica: International annual for lexicography* 26. 79–96.

Klos, Verena. 2011. *Komposition und Kompositionalität: Möglichkeiten und Grenzen der semantischen Dekodierung von Substantivkomposita*. Berlin & New York: De Gruyter.
Kobele, Gregory Michael. 2006. Generating Copies: An investigation into structural identity in language and grammar. Dissertation University of California, Los Angeles.
Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36. 15–43.
Koch, Peter & Wulf Oesterreicher. 1990. *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch.* Tübingen: Niemeyer.
Kolde, Gottfried. 1992. Zur Referenzsemantik von Eigennamen im gegenwärtigen Deutschen. *Deutsche Sprache* 20. 139–152.
Kolde, Gottfried. 1995. Grammatik der Eigennamen Überblick. In Ernst Eichler (Hrsg.), *Namenforschung. Ein internationales Handbuch zur Onomastik*, 400–408. Berlin: De Gruyter.
Köpcke, Klaus-Michael & David A. Zubin. 2009. Genus. In Elke Hentschel, Elke & Petra M. Vogel (Hrsg.), *Deutsche Morphologie*, 132–153. Berlin & New York: De Gruyter.
Kopf, Kristin. 2017. Fugenelement und Bindestrich in der Compositions-Fuge. In Nanna Fuhrhop, Renata Szczepaniak & Karsten Schmidt (Hrsg.), *Sichtbare und hörbare Morphologie*, 177–204. Berlin & Boston: De Gruyter.
Kopf, Kristin. 2018. *Fugenelemente diachron: eine Korpusuntersuchung zu Entstehung und Ausbreitung der verfugenden N+ N-Komposita.* Berlin & Boston: Walter de Gruyter.
Koptjevskaja-Tamm, Maria. 2013. A Mozart sonata and the Palme murder: The structure and uses of proper-name compounds in Swedish. In Kersti Börjars, David Denison & Alan Scott (Hrsg.), *Morphosyntactic categories and the expression of possession*, 253–290. Amsterdam & Philadelphia: Benjamins.
Koptjevskaja-Tamm, Maria & Anette Rosenbach. 2005. On the fuzziness of nominal determination. University of Stockholm & University of Düsseldorf: http://freelance-haven.weebly.com/uploads/5/0/1/1/5011326/fuzziness_of_nominal_ determination.pdf
Kouwenberg, Silvia & Darlene LaCharité. 2015. Arbitrariness and iconicity in total reduplication: Evidence from Caribbean Creoles. *Studies in Language* 39(4). 971–991.
Krifka, Manfred. 2003. Bare NPs: kind-referring, indefinites, both, or neither? *Semantics and Linguistic Theory* 13. 180–203.
Krifka, Manfred, Francis Jeffry Pelletier, Gregory Carlson, Alice Ter Meulen, Gennaro Chierchia & Godehard Link. 1995. Genericity: an introduction. *The Generic Book*, 1–124. Chicago: University of Chicago Press.
Krott, Andrea, Christina L. Gagné & Elena Nicoladis. 2009. How the parts relate to the whole: Frequency effects on children's interpretations of novel compounds. *Journal of child language* 36(1). 85–112.
Krott, Andrea, Christina L. Gagné & Elena Nicoladis. 2010. Children's preference for HAS and LOCATED relations: A word learning bias for noun–noun compounds. *Journal of child language* 37(2). 373–394.
Kürschner, Wilfried. 1974. *Zur syntaktischen Beschreibung deutscher Nominalkomposita auf der Grundlage generativer Transformationsgrammatiken.* Tübingen: Niemeyer.
Lakoff, George & Mark Johnson. 1980a. Conceptual metaphor in everyday language. *The journal of Philosophy* 77(8). 453–486.
Lakoff, George & Mark Johnson. 1980b. *Metaphors we live by*. Chicago: University of Chicago Press.
Lakoff, Robin. 1974. Remarks on this and that in papers from the tenth regional meeting of the Chicago Linguistic Society. *Chicago: University of Chicago*. 345–356.

Langacker, Ronald W. 2000. *Grammar and Conceptualization*. Berlin: Mouton de Gruyter.
Laudanna, Alessandro & Cristina Burani. 1995. Distributional properties of derivational affixes: Implications for processing. In Laurie Beth Feldman (Hrsg.), *Morphological aspects of language processing*, 345–364. New York: Psychology Press.
Lavric, Eva. 2000. Ein Modell der Referenz determinierter Nominalphrasen. *Zeitschrift für romanische Philologie* 116(1). 20–55.
Lee, Binna. 2007. A focus account for contrastive reduplication: prototypicality and contrastivity. *SNU Working Papers in English Linguistics and Language* 6. 78–90.
Lees, Robert B. 1960. *The grammar of english nominalizations*. Bloomington: Indiana University Press.
Lehmann, Christian. 2007. Motivation in language. Attempt at a systematization. In Peter Gallmann, Christian Lehmann & Rosemarie Lühr (Hrsg.), *Sprachliche Motivation. Zur Interdependenz von Inhalt und Ausdruck*, 105–140. Tübingen: Narr.
Lensch, Anke. 2018. *Fixer-uppers*. Reduplication in the derivation of phrasal verbs. In Rita Finkbeiner, Rita & Ulrike Freywald (Hrsg.), *Exact repetition in grammar and discourse*, 158–181. Berlin & Boston: De Gruyter.
Levi, Judith N. 1978. *The syntax and semantics of complex nominals.* New York: Academic Press.
Levinson, Stephen C. 2000. *Presumptive meanings: the theory of generalized conversational implicature.* Cambridge, MA: MIT Press.
Leys, Odo. 1989. Was ist ein Eigenname. In Friedhelm Debus & Wilfried Seibicke (Hrsg.), *Reader zur Namenkunde: Namentheorie*, 143–165. Hildesheim u.a.: Georg Olms.
Li, Charles N. & Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
Libben, Gary. 2006. Why study compound processing? An overview of the issues.In Gary Libben & Gonia Marema (Hrsg.), *The representation and processing of compound words*, 1–22. Oxford: Oxford University Press.
Libben, Gary, Martha Gibson, Yeo Bom Yoon & Dominiek Sandra. 2003. Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language* 84(1). 50–64.
Lieber, Rochelle. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.
Lieber, Rochelle. 2009a. IE, Germanic: English. In Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of compounding*, 357–369. Oxford: Oxford University Press.
Lieber, Rochelle. 2009b. A lexical semantic approach to compounding. In Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of compounding*, 78–104. Oxford: Oxford University Press.
Lieber, Rochelle & Pavol Štekauer. 2009. *The Oxford handbook of compounding*. Oxford: Oxford University Press.
Löbner, Sebastian. 1985. Definites. *Journal of semantics* 4. 279–326.
Löbner, Sebastian. 2011. Concept types and determination. *Journal of semantics* 28(3). 279–333.
Lohde, Michael. 2006. *Wortbildung des modernen Deutschen: ein Lehr- und Übungsbuch*. Tübingen: Narr.
Lüssy, Heinrich. 1974. *Umlautprobleme im Schweizerdeutschen: Untersuchungen an der Gegenwartssprache*. Frauenfeld: Huber.
Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
Lyons, John. 1977. *Semantics. Volume I*. Cambridge: Cambridge University Press.
Maas, Utz. 2005. Syntactic reduplication in Arabic. In Bernhard Hurch (Hrsg.), *Studies on reduplication*, 395–429. Berlin: Mouton de Gruyter.
Mächler, Patrick. 2017. Zu germ.* ē1 im Präteritum der germanischen starken Verben. In Barbara Sonnenhauser, Caroline Trautmann, Daniel Holl, Aziz Noe & Patrizia Hanna (Hrsg.), *Synchronie und Diachronie*, 33–59. München: Ibykos.

Maguire, Phil, Edward J. Wisniewski & Gert Storms. 2010. A corpus study of semantic patterns in compounding. *Corpus Linguistics & Linguistic Theory* 6. 49–73.
Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13(3). 435–482.
Marchand, Hans. 1969. *The categories and types of Present-day English word formation: a synchronic-diachronic approach*. München: Beck.
Marconi, Diego. 1997. *Lexical competence*. Cambridge, MA: MIT Press.
Mau, Thorsten. 2002. *Form und Funktion sprachlicher Wiederholungen.* Frankfurt a. M.: Peter Lang.
Medici, Mario. 1959. Il tipo 'caffè-caffè'. *Lingua nostra* 20. 84.
Meibauer, Jörg. 2013. Expressive compounds in German. *Word Structure* (1). 21–42.
Meid, Wolfgang. 1983. Bemerkungen zum indoeuropäischen Perfekt und zum germanischen starken Praeteritum. *STUF - Language Typology and Universals.* 329–336.
Mel'čuk, Igor. 1996. *Cours de morphologie générale, Volume 3, Moyens morphologiques. Syntactiques morphologiques.* Montréal Canada, Paris: Presses de l'Université de Montréal et CNRS Editions.
Menn, Lise & Brian MacWhinney. 1984. The repeated morph constraint: Toward an explanation. *Language* 60. 519–541.
Meys, Willem Johannes. 1975. *Compound adjectives in English and the ideal speaker-listener: a study of compounding in a transformational-generative framework*. Amsterdam: North Holland.
Michel, Sascha. 2009. Schaden-0-ersatz vs. Schaden-s-ersatz. Ein Erklärungsansatz synchroner Schwankungsfälle bei der Fugenbildung von N+ N-Komposita. *Deutsche Sprache* 37(4). 334–351.
Mithun, Marianne. 2010. Constraints on compounds and incorporation. In Sergio Scalise & Petra M. Vogel (Hrsg.), *Cross-disciplinary issues in compounding*, 37–56. Amsterdam: Benjamins.
Montermini, Fabio. 2010. Units in compounding. In Sergio Scalise & Petra M. Vogel (Hrsg.), *Cross-disciplinary issues in compounding*, 77–92. Amsterdam: Benjamins.
Moravcsik, Edith A. 1978. Reduplicative Constructions. In Joseph H. Greenberg (Hrsg.), *Universals of Human Language*, 297–334. Stanford: Stanford University Press.
Motsch, Wolfgang. 2004. *Deutsche Wortbildung in Grundzügen.* Berlin u.a.: De Gruyter.
Müller, Gereon. 1997. Beschränkungen für Binomialbildung im Deutschen. Ein Beitrag zur Interaktion von Phraseologie und Grammatik. *Zeitschrift Fur Sprachwissenschaft* 16(2). 5–51.
Murphy, Gregory L. 1988. Comprehending complex concepts. *Cognitive Science* 12(4). 529–562.
Murphy, Gregory L. 1990. Noun phrase interpretation and conceptual combination. *Journal of Memory and Language* 29(3). 259–288.
Nagelkerke, Nico J. D. 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78(3). 691–692.
Nakov, Preslav I. & Marti A. Hearst. 2013. Semantic interpretation of noun compounds using verbal and other paraphrases. *ACM Trans. Speech Lang. Process.* 10(3). 1–51.
Neef, Martin. 2009. Chapter 20: IE, Germanic: German. In Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of compounding*, 610–632. Oxford: Oxford University Press.
Neef, Martin. Borgwaldt, Susanne R. 2012. Fugenelemente in neugebildeten Nominalkomposita. In Livio Gaeta & Barbara Schlücker (Hrsg.), *Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte*, 27–56. Berlin & New York: Walter de Gruyter.
Neuß, Elmar. 1981. Kopulativkomposita. *Sprachwissenschaft* 6(1). 31–68.
Newman, Paul. 1989. Reduplication and tone in Hausa ideophones. *Annual Meeting of the Berkeley Linguistics Society* 15. 248–255.
Newman, Paul. 2000. *The hausa language. An Encyclopedic Reference Grammar.* New Haven: Yale University Press.

Nowak, Jessica & Damaris Nübling. 2017. Schwierige Lexeme und ihre Flexive im Konflikt: Hör- und sichtbare Wortschonungsstrategien. In Nanna Fuhrhop, Renata Szczepaniak & Karsten Schmidt (Hrsg.), *Sichtbare und hörbare Morphologie*, 113–144. Berlin & Boston: De Gruyter.

Nübling, Damaris. 2004. Die prototypische Interjektion: Ein Definitionsvorschlag. *Journal of Pragmalics* 18. 193–207.

Nübling, Damaris. 2005. Zwischen Syntagmatik und Paradigmatik: Grammatische Eigennamenmarker und ihre Typologie. *Zeitschrift für germanistische Linguistik* 33(1). 25–56.

Nübling, Damaris. 2011. Von *in die* über *in'n* und *ins* bis *im*. Die Klitisierung von Präposition und Artikel als 'Grammatikalisierungsbaustelle'. In Torsten Leuschner, Tanja Mortelmans & Sarah Groodt (Hrsg.), *Grammatikalisierung im Deutschen*, 105–132. Berlin & New York: De Gruyter.

Nübling, Damaris. 2012. Auf dem Wege zu Nicht-Flektierbaren: Die Deflexion der deutschen Eigennamen diachron und synchron. Björn Rothstein (Hrsg.), *Nicht-flektierbare und nicht-flektierte Wortarten*, 224–246. Berlin & New York: De Gruyter.

Nübling, Damaris, Fabian Fahlbusch & Rita Heuser. 2015. *Namen: Eine Einführung in die Onomastik*. Tübingen: Narr Francke Attempto.

Nübling, Damaris & Mirjam Schmuck. 2010. Die Entstehung des s-Plurals bei Eigennamen als Reanalyse vom Kasus- zum Numerusmarker. Evidenzen aus der deutschen und niederländischen Dialektologie. *Zeitschrift für Dialektologie und Linguistik* 77(2). 145–182.

Nübling, Damaris & Renata Szczepaniak. 2009. *Religion+s+freiheit*, *Stabilität+s+pakt* und *Subjekt +s+.pronomen*: Fugenelemente als Marker phonologischer Wortgrenzen. In Peter O. Müller (Hrsg.), *Studien zur Fremdwortbildung*, 195–222. Hildesheim: Olms.

Ó Séaghdha, Diarmuid. 2008. *Learning compound noun semantics*. Cambridge: University of Cambridge, Computer Laboratory.

Ogden, Charles Kay, Ivor Armstrong Richards. 1923. *The meaning of meaning: A study of the influence of thought and of the science of symbolism.* London: Kegan Paul, Trench & Trubner.

Olsen, Susan. 1986. *Wortbildung im Deutschen: eine Einführung in die Theorie der Wortstruktur.* Stuttgart: Kröner.

Olsen, Susan. 2012. Semantics of Compounds. In Claudis Maienborn, Klaus von Heusinger & Paul Portner (Hrsg.), *Semantics. An International handbook of natural language meaning*, 2120–2150. Berlin, New York & Boston: De Gruyter.

Olsen, Susan. 2015. 20. Composition. In Peter O. Müller; Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (Hrsg.), *Word-Formation. An International handbook of the languages of Europe*, 364–386. Berlin & Boston: De Gruyter Mouton.

Olsen, Susan. 2019. Semantics of compounds. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (Hrsg.), *Semantics-Interfaces*, 103–142. Berlin & Boston: De Gruyter.

Onysko, Alexander. 2014. Figurative processes in meaning interpretation: A case study of novel English compounds. *Yearbook of the German Cognitive Linguistics Association* 2(1). 69–88.

Ortner, Hanspeter & Lorelies Ortner. 1984. *Zur Theorie und Praxis der Kompositaforschung.* Tübingen: Narr.

Ortner, Lorelies & Elgin Müller-Bollhagen. 1991. *Deutsche Wortbildung. Typen und Tendenzen in der Gegenwart. Vierter Hauptteil: Substantivkomposita.* Berlin & New York: Walter de Gruyter.

Packard, Jerôme L. 2000. 构词法. *The morphology of Chinese. A Linguistic and Cognitive Approach.* Cambridge: Cambridge University Press.

Pallottino, Margherita & Tabea Ihsane. 2015. Characterizing Genericity: how bare plural objects affect the generic interpretation. Vortrag auf der ACED17, Universität Bukarest.

Partee, Barbara. 1984. Compositionality. *Varieties of formal semantics* 3. 281–311.

Pavlov, Vladimir Mikhaĭlovich. 1983. *Zur Ausbildung der Norm der deutschen Literatursprache im Bereich der Wortbildung 1470–1730. von der Wortgruppe zur substantivischen Zusammensetzung.* Berlin: Akademie-Verlag.

Payne, John & Rodney D. Huddleston. 2002. Nouns and noun phrases. In Rodney D. Huddleston & Geoffrey K. Pullum (Hrsg.), *The Cambridge grammar of the English language*, 323–523. Cambridge: Cambridge University Press.

Payne, Thomas Edward. 1997. *Describing morphosyntax: A guide for field linguists.* Cambridge: Cambridge University Press.

Perkuhn, Rainer, Holger Keibel & Marc Kupietz. 2012. *Korpuslinguistik*. Paderborn: Fink.

Peschel, Corinna. 2002. *Zum Zusammenhang von Wortneubildung und Textkonstitution*. Tübingen: Niemeyer.

Pfeifer, Wolfgang & Wilhelm Braun, Akademie der Wissenschaften der DDR. Zentralinstitut für Sprachwissenschaft. 1993. *Etymologisches Wörterbuch des Deutschen*. Berlin: Akademie-Verlag.

Pfeiffer, Wolfgang. 1995. *Etymologisches Wörterbuch des Deutschen. Digitalisierte und von Wolfgang Pfeifer überarbeitete Version im Digitalen Wörterbuch der deutschen Sprache.* München: Deutscher Taschenbuch Verlag.

Pinker, Steven. 1997. Words and rules in the human brain. *Nature* 387(6633). 547–548.

Pinker, Steven. 1998. Words and rules. *Lingua* 106(1–4). 219–242.

Pinker, Steven. 1999. *Words and rules*. New York: Basic Books.

Pittner, Karin & Judith Berman. 2006. *video ist echt schrott aber single ist hammer* – jugendsprachliche Nomen-Adjektiv-Konversion in der Prädikativposition. *Deutsche Sprache* 34(3). 233–251.

Plag, Ingo. 2006. The variability of compound stress in English: structural, semantic, and analogical factors. *English Language & Linguistics* 10(1). 143–172.

Plank, Frans. 1981. *Morphologische Ir-.Regularitäten: Aspekte der Wortstrukturtheorie*. Tübingen: Narr.

Platen, Christoph. 2013. *Ökonymie: Zur Produktnamen-Linguistik im Europäischen Binnenmarkt.* De Gruyter.

Poethe, Hannelor. 2014. Wortbildung im Schnittpunkt von Wortartenbedeutung, Wortbedeutung, Wortbildungsbedeutung und Motivationsbedeutung. In Sascha Michel & József Tóth (Hrsg.), *Wortbildungssemantik zwischen Langue und Parole. Semantische Produktions- und Verarbeitungsporzesse komplexer Wörter*, 23–38. Stuttgart: Ibidem.

Pott, August Friedrich. 1862. *Doppelung Reduplikation, Gemination. als eines der wichtigesten Bildungsmittel der Sprache, beleuchtet aus Sprachen aller Welttheile*. Lemgo: Meyer.

Pustejovsky, James. 1991. The syntax of event structure. *Cognition* 41(1–3). 47–81.

Quasthoff, Uwe & Matthias Richter. 2005. Projekt Deutscher Wortschatz. *Babylonia*. 33–35.

Raimy, Eric. 2000. *The phonology and morphology of reduplication.* Berlin & New York: Mouton de Gruyter.

Récanati, François. 2004. *Literal meaning*. Cambridge: Cambridge University Press.

Rhyne, James R. 1975. A lexical process model of nominal compounding in English. *American Journal of Computational Linguistics* 33. 33–44.

Richling, Julia. 2008. *Die Sprache in Foren und Newsgroups: eine Untersuchung der konzeptionellen Mündlichkeit und Schriftlichkeit im Wandel der Zeit.* Saarbrücken: VDM Publishing.

Rijkhoff, Jan. 2008. Layers, levels and contexts in Functional Discourse Grammar. *Trends in linguistic studies and monographs* 195. 63–116.

Rijkhoff, Jan. 2010. Functional categories in the noun phrase: on jacks-of-all-trades and one-trick-ponies in Danish, Dutch and German. *Deutsche Sprache* 2. 97–124.

Roca, Francesc & Avellina Suñer. 1998. Reduplicación y tipos de cuantificación en español. *Estudi General* 17. 37–66.

Rosch, Eleanor. 1975. Cognitive reference points. *Cognitive Psychology* 7(4). 532–547.
Rosch, Eleanor. 1978. Principles of categorization. Eleanor Rosch & Barbara B. Lloyd (Hrsg.), *Cognition and categorization*, 27–48. Hillsdale, NJ: Lawrence Erbaum Associates.
Rosenbach, Anette. 2007. Emerging variation: Determiner genitives and noun modifiers in English. *English Language and Linguistics* 11(1). 143–189.
Rosenbach, Anette. 2019. On the non-. equivalence of constructions with determiner genitives and noun modifiers in English. *English Language and Linguistics* 23(4). 759–796.
Rossi, Daniela. 2011. Lexical reduplication and affective contents. A pragmatic and experimental perspective. In Gregory Bochner, Philippe de Brabanter, Mikhail Kissine & Daniela Rossi (Hrsg.), *Cognitive and empirical pragmatics. Issues and perspectives*, 148–175. Amsterdam: Belgian Journal of Linguistics.
Rozhanskiy, Fedor Ivanovich. 2015. Two semantic patterns of reduplication: Iconicity revisited. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 39(4). 992–1018.
Rubino, Carl. 2005. Reduplication: form, function and distribution. In Bernhard Hurch (Hrsg.), *Studies on reduplication*, 11–29. Berlin: Mouton de Gruyter.
Rubino, Carl. 2013. Reduplication. In Matthew Dryer & Martin Haspelmath (Hrsg.), *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/feature/27A#2/28.3/149.4
Ryder, Mary Ellen. 1994. *Ordered chaos. The interpretation of English noun-noun compounds.* Berkeley, Los Angeles & London: University of California Press.
Sahel, Said, Guido Nottbusch, Angela Grimm & Rüdiger Weingarten. 2008. Written production of German compounds: Effects of lexical frequency and semantic transparency. *Written Language & Literacy* 11(2). 211–227.
Sandra, Dominiek. 1990. On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology Section A* 42(3). 529–567.
van Santen, Ariane & Geert Booij. 2017. *Morfologie: de woordstructuur van het Nederlands.* Amsterdam: Amsterdam University Press.
Sapir, Edward. 1921. *Language: An Introduction to the study of speech*. New York: Harcourt, Brace & Co.
Sauer, Hans. 2004. Lexicalization and demotivation. In Geert Booij, Christian Lehmann, Joachim Mugdan & Stravros Skopeteas (Hrsg.), *Morphology: An international handbook on inflection and word-formation*, 1625–1636. Berlin: Walter de Gruyter.
Schäfer, Martin. 2018. *The semantic transparency of English compound nouns.* Berlin: Language Science Press.
Schäfer, Martin & Melanie J. Bell. 2020. Constituent polysemy and interpretational diversity in attested English novel compounds. *The Mental Lexicon* 15(1). 42–61.
Schäfer, Roland. 2015. Processing and querying large web corpora with the COW14 architecture. *CMLC-3*. 28–34.
Schäfer, Roland & Felix Bildhauer. 2012. Building large corpora from the web using a new efficient tool chain. *LREC*. 486–493.
Scherer, Carmen. 2012. Vom *Reisezentrum* zum *Reise Zentrum.* Variation in der Schreibung von N+N-Komposita. In Livio Gaeta (Hrsg.), *Das Deutsche als kompositionsfreudige Sprache: strukturelle Eigenschaften und systembezogene Aspekte*, 57–81. Berlin: De Gruyter.
Schindler, Wolfgang. 1991. Reduplizierte Wortbildung im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 44(5). 597–613.

Schlaefer, Michael. 2002. *Lexikologie und Lexikographie. Eine Einführung am Beispiel deutscher Wörterbücher*. Berlin: Schmidt.
Schlechtweg, Marcel. 2018. *Memorization and the compound-phrase distinction: an investigation of complex constructions in German, French and English*. Berlin & Boston: Walter de Gruyter.
Schler, Jonathan, Moshe Koppel, Shlomo Argamon & James W. Pennebaker. 2006. Effects of age and gender on blogging. *Proceedings of the American Association for Artificial Intelligence AAAI. spring symposium: Computational approaches to analyzing weblogs*. 199–205.
Schlobinski, Peter & Torsten Siever. 2000. Kommunikationspraxen im Internet. *Der Deutschunterricht* 52(1). 54–65.
Schlücker, Barbara. 2012. Die deutsche Kompositionsfreudigkeit. Übersicht und Einführung. In Livio Gaeta, Livio & Barbara Schlücker (Hrsg.), *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.
Schlücker, Barbara. 2014. *Grammatik im Lexikon: Adjektiv-Nomen-Verbindungen im Deutschen und Niederländischen.* Berlin & Boston: De Gruyter.
Schlücker, Barbara. 2017. Eigennamenkomposita im Deutschen. In Johannes Helmbrecht, Damaris Nübling & Barbara Schlücker (Hrsg.), *Namengrammatik*, 59–93. Hamburg: Helmut Buske.
Schlücker, Barbara. 2020. Between word-formation and syntax: Cross-linguistic perspectives on an ongoing debate. *Zeitschrift für Wortbildung/Journal of Word Formation* 4(1). 26–74.
Schlücker, Barbara & Tanja Ackermann. 2017. The morphosyntax of proper names: An overview. *Folia Linguistica* 51(2). 309–339.
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.
Schmerling, Susan F. 1971. A stress mess. *Studies in the Linguistic Sciences* 1. 52–66.
Schmid, Helmut & Florian Laws. 2008. Estimation of conditional probabilities with decision trees and an application to fine-grained POS tagging. *Proceedings of the 22nd International Conference on Computational Linguistics Coling 2008*. 777–784.
Schmuck, Mirjam. 2017. Movierung weiblicher Familiennamen im Frühneuhochdeutschen und ihre heutigen Reflexe. In Johannes Helmbrecht, Damaris Nübling & Barbara Schlücker (Hrsg.), *Namengrammatik*, 33–58. Hamburg: Helmut Buske.
Schreuder, Robert & R. Harald Baayen. 1995. Modeling morphological processing. *Morphological aspects of language processing* 2. 257–294.
Schreuder, Robert & R. Harald Baayen 1997. How complex simplex words can be. *Journal of Memory and Language* 37(1).118–139.
Schulte im Walde, Sabine & Susanne R. Borgwaldt. 2015. Association norms for German noun compounds and their constituents. *Behavior research methods* 47(4). 1199–1221.
Schwaiger, Thomas. 2015. Reduplication. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (Hrsg.), *Word-formation. An international handbook of the languages of Europe*, 467–484. Berlin & Boston: Walter de Gruyter.
Schwitalla, Johannes. 1996. Beziehungsdynamik. Kategorien für die Beschreibung der Beziehungsgestaltung sowie der Selbst-und Fremddarstellung in einem Streit- und Schlichtungsgespräch. In Werner Kallmeyer (Hrsg.), *Gesprächsrhetorik: rhetorische Verfahren im Gesprächsprozess*, 279–349. Tübingen: Narr.
Schwitalla, Johannes. 2006. *Gesprochenes Deutsch. Eine Einführung*. Berlin: Erich Schmidt.
Searle, John R. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Seiler, Hansjakob. 1978. *Determination: A functional dimension for interlanguage comparison.* Tübingen: Narr.

Selkirk, Elisabeth O. 1982. *The syntax of words*. Cambridge, MA: MIT Press.

Semenza, Carlo. 2006. Retrieval pathways for common and proper names. *Cortex* 42(6). 884–891.

Shieber, Stuart M. 1986. *An introduction to unification based approaches to grammar*. Stanford, CA: CSLI.

Sick, Bastian. 2009. *Der Dativ ist dem Genitiv sein Tod – Folge 3: Noch mehr Neues aus dem Irrgarten der deutschen Sprache.* Köln: Kiepenheuer & Witsch.

Søgaard, Anders. 2005. Compounding theories and linguistic diversity. In Zygmunt Frajzyngier, Adam Hodges & David S. Rood (Hrsg.), *Linguistic diversity and language theories*, 319–337. Amsterdam & Philadelphia: Benjamins.

Song, Myounghyoun & Chungmin Lee. 2015. CF-reduplication in English: Dynamic Prototypes & Contrastive Focus Effects. *Semantics and Linguistic Theory* 21(10). 444–462.

Spalding, Thomas, Christina Gagné, Allison Mullaly & Hongbo Ji. 2010. Relation-based interpretation of noun-noun phrases: A new theoretical approach. In Susan Olsen (Hrsg.), *New impulses in word-formation*, 283–315. Hamburg: Helmut Buske Verlag.

Spalding, Thomas, Christina Gagné, Kelly Nisbet, Jenna Chamberlain & Gary Libben. 2019. If birds have sesamoid bones, do blackbirds have sesamoid bones? The modification effect with known compound words. *Frontiers in Psychology* 10(1570). 1–16.

Spencer, Andrew. 2003. Does English have productive compounding. Geert E. Booij, Janet De Cesaris, Angela Ralli & Sergio Scalise (Hrsg.), *Topics in morphology. Selected Papers from the Third Mediterranean Morphology Meeting, Barcelona, Sept 20-22, 2001*, 329–341. Barcelona: Institut Universitari de Linguistica Aplicada.

Spencer, Andrew. 2005. Towards a typology of 'mixed categories'. In Orhan C. Orgun & Peter Sells (Hrsg.), *Morphology and the web of grammar. Essays in memory of Steven G. Lapointe*, 95–138. Stanford, CA: CSLI.

Stefanowitsch, Anatol. 2007. Wortwiederholung im Englischen und Deutschen: eine korpuslinguistische Annäherung. In Andreas Ammann & Aina Urdze (Hrsg.), *Wiederholung, Parallelismus, Reduplikation*, 29–45. Bochum: Universitätsverlag Dr. N. Brockmeyer.

Steinhauer, Anja. 2000. *Sprachökonomie durch Kurzwörter: Bildung und Verwendung in der Fachkommunikation.* Tübingen: Narr.

Štekauer, Pavol. 2000. *English word-formation: A history of research, 1960-1995*. Tübingen: Narr.

Štekauer, Pavol. 2005. *Meaning Predictability in Word Formation: Novel, context-free naming units.* Amsterdam & Philadelphia: Benjamins.

Štekauer, Pavol. 2009. Meaning predictability of novel context-free compounds. In Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of compounding*, 431–470. Oxford: Oxford University Press.

Stolz, Thomas. 1988. Markiertheitshierarchie und Merkmalhaftigkeit in Numerussystemen: Über den Dual. *STUF-Language Typology and Universals* 41(1–6). 476–487.

Stolz, Thomas. 2006. (Wort-)Iteration: (k)eine universelle Konstruktion. In Kerstin Fischer & Anatol Stefanowitsch (Hrsg.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie*, 105–132. Tübingen: Stauffenburg.

Stolz, Thomas. 2018. (Non-)Canonical reduplication. In Aina Urdze (Hrsg.), *Non-prototypical reduplication*, 201–276. Berlin & Boston: De Gruyter.

Stolz, Thomas & Nataliya Levkovych. 2018. Function vs Form – On ways of telling repetition and reduplication apart. In Ulrike Freywald & Rita Finkbeiner (Hrsg.), *Exact repetition in grammar and discourse*, 29–66. Berlin & Boston: Walter de Gruyter.

Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2011. *Total reduplication the areal linguistics of a potential universal*. Berlin: Akademie-Verlag.

Storrer, Angelika. 2018. Interaktionsorientiertes Schreiben im Internet. In Arnulf Deppermann & Silke Reineke (Hrsg.), *Sprache im kommunikativen, interaktiven und kulturellen Kontext*, 219–244. Berlin & Boston: De Gruyter.

Sturm, Afra. 2005. Eigennamen als kontextabhängige und inhärent definite Ausdrücke. *Zeitschrift für Sprachwissenschaft* 24(1). 67–91.

Tannen, Deborah. 1989. *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Tarasova, Elizaveta. 2013. Some new insights into the semantics of English N+N compounds. Dissertation Victoria University of Wellington.

Ten Hacken, Pius. 1994. Defining morphology: A principled approach to determining the boundaries of compounding, derivation, and inflection. Dissertation Universität Basel.

Ten Hacken, Pius. 2019. *Word formation in Parallel Architecture: the case for a separate component.* Chad: Springer International Publishing.

Teubert, Wolfgang. 1998. *Neologie und Korpus*. Tübingen: Narr.

Teyssier, Jacques. 1968. Notes on the syntax of the adjective in modern English. *Lingua* 20. 225–249.

Thornton, Anna M. 2009. Italian verb reduplication between syntax and the lexicon. *Italian Journal of Linguistics* 21(1). 235–261.

Vacek, Jaroslav. 1988. Über einen Aspekt der Reduplikation. *Studia Orientalia Pragensia* 17. 111–125.

Vater, Heinz. 1984. Determinantien und Quantoren im Deutschen. *Zeitschrift für Sprachwissenschaft* 3(1). 19–42.

Vater, Heinz. 2010. Sprachspiele: kreativer Umgang mit Sprache. *Linguistische Berichte* 221. 3–36.

Verschueren, Jef. 2012. Metalanguage. In Adam Jaworski, Nikolas Coupland & Dariusz Galasinski (Hrsg.), *Notes on the role of metapragmatic awareness in language use*, 53–74. Berlin & Boston: De Gruyter Mouton.

Vogel, Petra M. 2017. Deonymische Adjektivkomposita „Eigenname + Adjektiv'" vom Typ *goethefreundlich*. In Johannes Helmbrecht, Damaris Nübling & Barbara Schlücker (Hrsg.), *Namengrammatik*, 95–120. Hamburg: Helmut Buske.

von der Heide, Claudia & Susanne Borgwaldt. 2009. Assoziationen zu Unter-, Basis- und Oberbegriffen. Eine explorative Studie. In Ralf Vogel & Said Sahel (Hrsg.), *Proceedings of the 9th Norddeutsches Linguistisches Kolloquium*, 51–74. Bielefeld: Universität Bielefeld.

Wälchli, Bernhard. 2007. Ko-Komposita im Vergleich mit Parallelismus und Reduplikation. In Andreas Ammann & Aina Urdze (Hrsg.), *Wiederholung, Parallelismus. Reduplikation. Strategien der multiplen Strukturanwendung*, 81–107. Bochum: Brockmeyer.

Ward, Gregory, Richard Sproat & Gail McKoon. 1991. A pragmatic analysis of so-called anaphoric islands. *Language* 67(3). 439–474.

Warren, Beatrice. 1978. Semantic patterns of noun-noun compounds. *Acta Universitatis Gothoburgensis. Gothenburg Studies in English Goteborg* 41. 1–266.

Wegener, Heide. 2005. Das Hühnerei vor der Hundehütte: von der Notwendigkeit historischen Wissens in der Grammatikographie des Deutschen. In Elisabeth Berner, Manuela Böhm & Anja Voeste (Hrsg.), *Ein gross vnnd narhafft haffen: Festschrift für Joachim Gessinger*, 175–188. Potsdam: Universitätsverlag.

Wellmann, Hans. 1975. *Deutsche Wortbildung. Typen und Tendenzen in der Gegenwartssprache. Zweiter Hauptteil: Das Substantiv.* Düsseldorf: Schwann.

Wellmann, Hans. 1998. Wortbildung. *Dudenband* 4. 408–557.

Whitney, William Dwight. 1889. *Sanskrit grammar*. Leipzig: Breitkopf & Härtel.

Whitton, Laura. 2006. The semantics of contrastive focus reduplication in English: does the construction mark prototype-prototype. Manuskript Stanford University.
Widlitzki, Bianca. 2016. *Talk talk, not just small talk*. Exploring English contrastive focus reduplication with the help of corpora. *ICAME* 40(1). 119–142.
Wierzbicka, Anna. 1986. Italien reduplication: cross-cultural pragmatics and illocutionary semantics. *Linguistics* 24(2). 287–316.
Wierzbicka, Anna. 1991. *Cross-cultural pragmatics*. Berlin & New York: Mouton de Gruyter.
Wiese, Richard. 1990. Über die Interaktion von Morphologie und Phonologie – Reduplikation im Deutschen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 4. 603–624.
Wilbur, Ronnie. 1973. *The phonology of reduplication*. Bloomington: Indiana University Linguistics Club.
Wippich, Werner. 1995. Namengedächtnis und Namenlernen. In Ernst Eichler, Gerold Hilty, Heinrich Löffler, Hugo Steger & Ladislav Zgusta (Hrsg.), *Namenforschung. Ein internationales Handbuch zur Onomastik*, 489–493. Berlin & New York: De Gruyter.
Wisniewski, Edward J & Gregory L. Murphy. 2005. Frequency of relation type as a determinant of conceptual combination: a reanalysis. *Journal of experimental psychology: Learning, memory, and cognition* 31(1). 169–174.
Wunderlich, Dieter. 1986. Probleme der Wortstruktur. *Zeitschrift Fur Sprachwissenschaft* 5(2). 209–252.
Wüster, Eugen. 1979. *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie.* Wien & New York: Springer.
Yu, Alan C. L. 2008. On iterative infixation. *26th West Coast Conference on Formal Linguistics*. 516–524.
Zepp, Christian, Jens Kleinert & Anja Liebscher. 2013. Inhalte und Strukturen prototypischer Merkmale in Fußballmannschaften. *Sportwissenschaft* 43(4). 283–290.
Zhu, Jinyang, Christine Culp & Karl-Heinz Best. 1995. Formen und Funktionen der Doppelungen im Chinesischen im Vergleich zum Deutschen. *Oriens Extremus* 38(1). 183–208.
Zienkowski, Jan. 2014. Diskursforschung. In Johannes Angermuller, Martin Nonhoff, Eva Herschinger, Felicitas Macgilchrist, Martin Reisigl, Juliette Wedl, Daniel Wrana & Alexander Ziem (Hrsg.), *Kritisches Bewusstsein durch den Gebrauch metapragmatischer Marker*, 1188–1215. Bielefeld: transcript.
Zifonun, Gisela. 2010. Possessive Attribute im Deutschen. *Deutsche Sprache* 38(2). 124–153.
Zimmer, Benjamin & Charles E. Carson. 2018. 'Among the New Words': The prospects and challenges of short-term historical lexicography. *Dictionaries: Journal of the Dictionary Society of North America* 39(1). 59–74.
Zimmer, Karl E. 1972. Appropriateness conditions for nominal compounds. *Working Papers on Language Universals* 8. 3–20.
Zimmer, Karl E. 1971. Some General Observations about Nominal Compounds. *Working Papers on Language Universals* 5. 1–21.
Zinsmeister, Heike. 2013. Corpus-based modeling of the semantic transparency of noun-noun compounds. In Holden Härtl (Hrsg.), *Interfaces of Morphology: A Festschrift for Susan Olsen*, 303–321. Berlin: Akademie Verlag.
Zürn, Constanze. 2016. *Untersuchungen zur Semantik okkasioneller Nominalkomposita.* Norderstedt: GRIN.
Zwitserlood, Pienie. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes* 9(3). 341–368.
Zwitserlood, Pienie. 2018. Processing and representation of morphological complexity in native language comprehension and production. In Geert Booij (Hrsg.), *The construction of words. Studies in morphology*, 583–602. Chad: Springer.

Erratum zum Vorwort

Veröffentlicht in: Maximilian Frankowsky, N+N-Komposita mit identischen Konstituenten im Deutschen, 978-3-11-131496-9

Erratum

Die ursprüngliche Version des Kapitels wurde revidiert: Im Vorwort auf S. V wurde „A0B“ zu „A≠B“ korrigiert.

Das revidierte Originalkapitel ist abrufbar unter: https://doi.org/10.1515/9783111315416-203

https://doi.org/10.1515/9783111315416-014