

Frequenz. Prototyp. Schema.

Ein gebrauchsbasierter Ansatz zur
Entstehung grammatischer Varianten

Eleonore Schmitt

Open Germanic Linguistics 7



Open Germanic Linguistics

Editors: Michael T. Putnam, B. Richard Page, Laura Catharine Smith

In this series:

1. Lemke, Robin. Experimental investigations on the syntax and usage of fragments.
2. Larsson, Ida & Erik Petzell (eds.). Morphosyntactic change in Late Modern Swedish.
3. Hall, Tracy Alan. Velar fronting in German dialects: A study in synchronic and diachronic phonology.
4. Cysouw, Michael. Encyclopaedia of German diatheses.
5. Bacskai-Atkari, Julia. The syntax of functional left peripheries: Clause typing in West Germanic and beyond.
6. Nickel, Grit. Nominale Flexionsmorphologie in den ostoberdeutschen Dialekten Bayerns.
7. Schmitt, Eleonore. Frequenz. Prototyp. Schema.: Ein gebrauchsbasierter Ansatz zur Entstehung grammatischer Varianten.

ISSN (print): 2750-5588

ISSN (electronic): 2750-557X

Frequenz. Prototyp. Schema.

Ein gebrauchsbasierter Ansatz zur
Entstehung grammatischer Varianten

Eleonore Schmitt

Eleonore Schmitt. 2023. *Frequenz. Prototyp. Schema.: Ein gebrauchsbasierter Ansatz zur Entstehung grammatischer Varianten* (Open Germanic Linguistics 7).

Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/406>

© 2023, Eleonore Schmitt

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

Diese Arbeit hat der Fakultät Geistes- und Kulturwissenschaften der Otto-Friedrich-Universität Bamberg als Dissertation vorgelegen.

Gutachter_innen: Prof. Dr. Renata Szczepaniak; Prof. Dr. Alexander Werth; Prof. Dr. Klaus-Michael Köpcke

Tag der mündlichen Prüfung: 21.06.2021.

ISBN: 978-3-96110-423-9 (Digital)

978-3-98554-081-5 (Hardcover)

ISSN (print): 2750-5588

ISSN (electronic): 2750-557X

DOI: 10.5281/zenodo.8434909

Source code available from www.github.com/langsci/406

Errata: paperhive.org/documents/remote?type=langsci&id=406

Cover and concept of design: Ulrike Harbort

Proofreading: Hella Olbertz, Katja Politt, Lea Schäfer, Nicole Benker, Rebecca Madlener, Sophie Ellsaeßer, Tom Bossuyt, Yvonne Treis

Fonts: Libertinus, Arimo, DejaVu Sans Mono

Typesetting software: Xe_{La}TeX

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

<http://langsci-press.org>

Storage and cataloguing done by FU Berlin

Freie Universität  Berlin

Inhaltsverzeichnis

Danksagung	v
1 Einleitung und Zielsetzung	1
1.1 Frequenz, Prototyp und Schema als Einflussfaktoren auf Variation	4
1.2 Gliederung der Arbeit	8
2 Variation aus gebrauchsbasierter Perspektive	9
2.1 Frequenz	9
2.1.1 Typen- und Tokenfrequenz	12
2.1.2 Frequenzeffekte und Korrelation mit anderen Faktoren .	24
2.1.3 Frequenz und statistisches Lernen	30
2.2 Prototypizität	45
2.2.1 Prototypizität sprachlicher Kategorien	46
2.2.2 Prototypeneffekte	55
2.2.3 Prototyp- und exemplarbasierte Kategorien	61
2.3 (Form-)Schematizität	70
2.3.1 Grundlegender Schemabegriff	71
2.3.2 Abgrenzung von Schematizität und Form-Schematizität	73
2.3.3 Form-Schema-Effekte	85
2.4 Zusammenfassung: Variation aus gebrauchsbasierter Perspektive	89
3 Variation in Konjugation, Deklination und Auxiliarselektion	97
3.1 Variation in der Konjugationsklasse	97
3.1.1 Frequenz	100
3.1.2 Prototypizität	108
3.1.3 Form-Schematizität	114
3.1.4 Zusammenfassung und Zusammenwirken der Faktoren	118
3.2 Variation in der Deklinationsklasse	122
3.2.1 Frequenz	124
3.2.2 Prototypizität	128
3.2.3 Form-Schematizität	134
3.2.4 Zusammenfassung und Zusammenwirken der Faktoren	151

Inhaltsverzeichnis

3.3	Variation in der Selektion von <i>haben</i> und <i>sein</i>	154
3.3.1	Frequenz	155
3.3.2	Prototypizität	158
3.3.3	Zusammenfassung und Zusammenwirken der Faktoren	172
4	Psycholinguistische Betrachtung der Einflussfaktoren: Methode	177
4.1	Reaktionsmessverfahren	177
4.1.1	<i>Lexical decision tasks</i>	177
4.1.2	<i>Self-paced reading tasks</i>	180
4.1.3	<i>Sentence maze tasks</i>	184
4.2	Lexical-decision-Studie zu Frequenzeffekten	186
4.2.1	Fragestellung und Hypothesen	187
4.2.2	Materialien	190
4.2.3	Versuchsdesign	194
4.2.4	Versuchsablauf	195
4.2.5	Metadaten	196
4.2.6	Probandenakquise und Proband_innen	197
4.3	Sentence-maze-Studie zu Prototypizitätseffekten	199
4.3.1	Fragestellung und Hypothesen	199
4.3.2	Materialien	201
4.3.3	Versuchsdesign	203
4.3.4	Versuchsablauf	205
4.3.5	Metadaten, Probandenakquise und Proband_innen . . .	205
4.4	Self-paced-reading-Studie zu Form-Schematizitätseffekten . . .	207
4.4.1	Fragestellung und Hypothesen	208
4.4.2	Materialien	209
4.4.3	Versuchsdesign	210
4.4.4	Versuchsablauf	213
4.4.5	Metadaten	214
4.4.6	Probandenakquise und Proband_innen	214
4.5	Lexical-decision- und sentence-maze-Studie zu Form-Schematizitätseffekten	216
4.5.1	Fragestellung und Hypothesen	216
4.5.2	Materialien	218
4.5.3	Versuchsdesign	219
4.5.4	Versuchsablauf, Metadaten, Probandenakquise, Proband_innen	221

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse	225
5.1 (Generalisierte) gemischte lineare Modelle und <i>random forests</i>	226
5.1.1 (Generalisierte) gemischte lineare Modelle	226
5.1.2 <i>Random forests</i>	230
5.2 Frequenz	231
5.2.1 Antwortverhalten	232
5.2.2 Reaktionszeiten	243
5.3 Prototypizität	256
5.3.1 Antwortverhalten	257
5.3.2 Reaktionszeiten	267
5.4 Form-Schematizität	276
5.4.1 Antwortverhalten	276
5.4.2 Reaktionszeiten	299
5.5 Diskussion der Ergebnisse	326
5.5.1 Frequenz	326
5.5.2 Prototypizität	329
5.5.3 Form-Schematizität	332
5.5.4 Zusammenfassung und Evaluation der Reaktionszeitmessverfahren	336
6 Zusammenfassung und Ausblick	339
6.1 Frequenz	339
6.2 Prototypizität	345
6.3 (Form-)Schematizität	347
6.4 Gewichtung der Einflussfaktoren	351
Anhang A: Digitaler Anhang	355
Anhang B: Ablautreihen des Deutschen	357
Anhang C: Frequenz der Präteritalformen in der Frequenzstudie	359
Anhang D: Lexical-decision-Studie	361
D.1 Aufbau der Blöcke in der lexical-decision-Studie	361
D.2 Überblick über die Füllertems	364
Anhang E: Sentence-maze-Studie	365
E.1 Aufbau der Blöcke in der sentence-maze-Studie	365
E.2 Überblick über die Füllsätze	365

Inhaltsverzeichnis

Anhang F: Plausibilitätstest zur sentence-maze-Studie	369
F.1 Ergebnisse des Plausibilitätstests	369
F.2 Distraktorsätze	371
Anhang G: Self-paced-reading-Studie	373
G.1 Texte zu den Testitems	373
G.1.1 Prototyp des Form-Schemas: <i>Schettose</i> und <i>Truntake</i> . .	373
G.1.2 Peripherie des Form-Schemas: <i>Knatt</i>	375
G.1.3 Form-Schema der starken Flexion: <i>Grettel</i>	375
G.2 Texte zu den Distraktoren	376
G.3 Produktionsexperiment	378
Anhang H: Assoziationstest zur self-paced-reading-Studie	379
Anhang I: Ergebnisse der Studien	383
I.1 Ergebnisse der Frequenzstudie	383
I.2 Ergebnisse der Prototypizitätsstudie	385
I.3 Ergebnisse der Form-Schematizitätsstudien	386
I.3.1 Lexical-decision-Studie	386
I.3.2 Sentence-maze-Studie	389
I.3.3 Self-paced-reading-Studie	392
I.4 Vergleich zwischen den Reaktionszeiten der Verben und der Substantive in der lexical-decision-Studie	395
Literaturverzeichnis	397
Register	423
Personenregister	423

Danksagung

Mein Dank gilt allen, die mir während meiner Promotion zur Seite gestanden haben, egal ob mit Rat, Literaturhinweisen oder einer Umarmung und aufmunternden Worten. Ich konnte mich auf eine sehr hohe Frequenz an Rückhalt und Zuspruch stützen. Dabei war sowohl die Typen- als auch die Tokenfrequenz extrem hoch, was aus frequenztheoretischer Sicht ungewöhnlich ist. Gemessen an dieser hohen Frequenz ist das Lemma (*be-*)*danken* in meiner Dissertation relativ selten. Eine Danksagung mit einer möglichst hohen Dichte an Dank soll diesen Missstand ausgleichen.

Besonderer Dank gilt Renata Szczepaniak, die meine Dissertation betreut hat. In ihren Seminaren habe ich früh meine Faszination für Sprachwissenschaft entdeckt. Während der Dissertation konnte ich stets auf ihren fachlichen Rat und ihre Unterstützung zählen. Auch einer thematischen Verschiebung hin zu einer Dissertation mit dezidiert gebrauchsbasiertem Ansatz und ohne Fehler(-diskussion) stimmte sie ohne Bedenken zu.

Besonders möchte ich auch Alexander Werth danken, der meine Arbeit zweitbetreut hat. Alexander fragte mich im Februar 2018, wer eigentlich mein Zweitbetreuer sei. Ich antwortete: „Wenn du magst, du.“ Damit konnte ich mich bereits früh in meiner Promotion auf Rat von zwei Seiten stützen. Ich danke Alexander insbesondere für wertvolle methodische Hinweise und viele Diskussionen zum Aufbau der Reaktionszeitstudien.

Für fachlichen Input und offene Ohren danke ich meinen Hamburger und Bamberger Kolleg_innen. Weiterhin bedanke ich mich bei der Studienstiftung des deutschen Volkes, von deren finanzieller und ideeller Förderung ich während meines Studiums und meiner Promotion profitieren durfte.

Für unendliche emotionale Unterstützung danke ich meinem Freundeskreis und ganz besonders dem Promowendland. Meinen Eltern Jürgen und Margret danke ich für ihr Vertrauen in mich und den unbedingten Rückhalt, den sie mir geben. Genauso danke ich meinen Geschwistern Verena und Maximilian: Es ist schön zu wissen, dass ich immer auf euch zählen kann. Zum Schluss danke ich David – für alles.

Eure Unterstützung hat die Wahrscheinlichkeit für das Gelingen der Dissertation massiv zum Positiven beeinflusst.

1 Einleitung und Zielsetzung

Sprache und Variation sind eng miteinander verwoben – bzw. verwebt. Dies verdeutlichen die verschiedenen sprachlichen Ebenen, auf denen Variation zu beobachten ist: Varianten existieren in der Phonologie ([çe: mi:] vs. [ʃ e: mi:] vs. [ke: mi:]), in der Morphologie (*geglommen* vs. *geglimmt*; *des Grafen* vs. *des Grafs*), in der Syntax (*des Vaters Haus* vs. *das Haus des Vaters*), an der Syntax-Semantik-Schnittstelle (*ein Auto gefahren sein* vs. *ein Auto gefahren haben*) und der Morphosyntax (*wegen des* vs. *wegen dem*). Wo Sprache ist, ist auch Variation.

Aber nicht alles variiert, was variieren könnte: Zwar schwankt *glimmen* zwischen *geglommen* und *geglimmt*, aber *ziehen* nicht zwischen *gezogen* und **gezieht*. Genauso schwankt *Graf* zwischen *des Grafen* und *des Grafs*, aber *Matrose* nicht zwischen *des Matrosen* und *des *Matroses*. Und während *Auto fahren* sowohl *sein* als auch *haben* im Perfekt erlaubt, ist bei *die Kinder nach Hause fahren* nur *haben* möglich und bei *zur Kur fahren* nur *sein*. Die vorliegende Arbeit entwickelt ein Modell, das Vorhersagen darüber ermöglicht, unter welchen Umständen Variation auftritt und unter welchen sie unwahrscheinlich ist.

Zu Variation im Deutschen existieren zahlreiche Studien, die den Sprachgebrauch in Hinblick auf Variation bspw. in der Flexion und Auxiliarselektion synchron und diachron in Korpora untersuchen (z. B. Nowak 2015, Gillmann 2016, Schäfer 2019) oder mithilfe von Produktionsexperimenten überprüfen, welche Faktoren die Variation beeinflussen (z. B. Köpcke 2000b, 2005). Diese fokussieren jedoch jeweils auf einzelne Variationsphänomene und erklären diese post-hoc. Es fehlt daher ein umfassendes Modell, das Variation und deren Entstehung übergreifend modelliert und es damit ermöglicht, Variation zu prognostizieren. Zudem sind bisherige Studien zu Variation derzeit stark auf Korpora fokussiert. Zwar wurde zur Prozessierung von ambigen Strukturen im Vergleich zu eindeutigen Strukturen vereinzelt geforscht, wie bspw. Roehm u. a. (2013) zur Prozessierung der Auxiliare *haben* und *sein* bei *degree achievements* (*Der Pffirsich ist/hat geschimmelt*) und telischen Verben (*Der Pffirsch ist/*hat verschimmelt*), um die Entstehung von Variation umfassend beleuchten zu können, sollte aber die Verarbeitung von Varianten und von Formen, die nicht im Gebrauch schwanken, stärker in den Fokus gerückt werden. Dies hat die vorliegende Arbeit zum Ziel.

1 Einleitung und Zielsetzung

In der Arbeit wird ein gebrauchsbasiertes Modell entwickelt, das psycholinguistisch überprüft wird. Die Arbeit verfolgt damit einen dezidiert gebrauchsbasierten Ansatz. In dem gebrauchsbasierten Modell zu Variation werden drei Faktoren ausgearbeitet, die Einfluss darauf nehmen, ob Variation (z. B. das Verb *ziehen* oder die Konstruktion *Auto fahren*) auftritt oder nicht. Der Einfluss der Faktoren wird psycholinguistisch mithilfe von Reaktionszeitmessungen überprüft.

Ein solches Modell ermöglicht es, Variation unabhängig von einzelnen Variationsphänomenen zu betrachten und Variation nicht nur post hoc zu erklären, sondern zu prognostizieren, wie wahrscheinlich es ist, dass Variation auftreten wird. Die empirische Überprüfung mithilfe von Reaktionszeiten ist ein wichtiger Bestandteil des Modells, da Reaktionszeiten Hinweise auf Variationspotential liefern können, das noch nicht im Sprachgebrauch sichtbar ist. Dies ist bspw. der Fall, wenn im Sprachgebrauch existierende Varianten ähnlich schnell verarbeitet werden wie Formen, die im Sprachgebrauch (noch) keine Variation aufweisen. Die Betrachtung der Prozessierung von Varianten zusätzlich zu ihrer Verwendung ermöglicht somit einen umfassenden Blick auf Variation. Die vorliegende Arbeit ergänzt und erweitert die bestehende Forschung durch den Fokus auf Faktoren, die variationsphänomenübergreifend Einfluss auf Variation nehmen, und die systematische psycholinguistische Überprüfung der Faktoren. Im Rahmen der psycholinguistischen Überprüfung der Faktoren wird ein direkter Vergleich von Varianten im Sprachgebrauch (*geglimmt* vs. *geglommen*) und Formen vorgenommen, die keine Variation aufweisen (*gezogen* vs. **gezieht*). Hierdurch ermöglicht die Arbeit es, die Prozessierung von Varianten besser zu verstehen. Die Arbeit greift auf drei verschiedene Verfahren zur Reaktionszeitmessung zurück (*self-paced reading*, *lexical decision* und *sentence maze tasks*) und bietet somit einen vielfältigen Zugriff auf Reaktionszeiten.

Die Arbeit nutzt den gebrauchsbasierten Ansatz als Grundlage für die Modellierung von Variation. Der gebrauchsbasierte Ansatz sieht Sprache als ein wandelbares und kontextabhängiges Geflecht aus Strukturen, das durch Nutzung entsteht und gefestigt wird. Die kognitive Repräsentation sprachlichen Wissens ist in dem Ansatz daher stets auf Sprachgebrauch zurückzuführen: „The basic premise of Usage-based Theory is that experience with language creates and impacts the cognitive representations for language“ (Bybee 2013: 49). Die Struktur von Sprachsystemen entsteht also durch Gebrauch und ist nicht a priori festgelegt (Bybee 2006: 714). Grammatik lässt sich dabei mit mentalen Repräsentationen im Gehirn und Prozessierung gleichsetzen (Bybee & Hopper 2001: 2–3, Kemmer & Barlow 2000: xii). Mentale Repräsentationen – und damit auch Grammatik – verändern und stabilisieren sich ständig durch Gebrauch (Bybee & Hopper 2001: 2, Bybee 2013: 68) und stellen somit „provisional and temporary states of affair“

(Bybee & Hopper 2001: 2) dar. In der gebrauchsbasierten Perspektive kann ein Sprachsystem daher nur existieren, wenn es immer wieder genutzt wird, und durch die Nutzung wird es ständig verändert. Der gebrauchsbasierte Ansatz eignet sich zur Modellierung von Variation, da er mit seinem Fokus auf die Veränderung von mentalen Repräsentationen durch Nutzung Sprache eng an Variation koppelt. Zudem verdeutlicht der Ansatz die Relevanz, die Prozessierung von Varianten zu erforschen, da sie Rückschlüsse auf die mentalen Repräsentationen der Varianten liefern kann.

Frequenz, Prototyp und Schema werden innerhalb des Modells als grundlegende kognitive Faktoren entwickelt, die Einfluss auf Variation in einer Sprache nehmen. Sie werden anhand von drei Variationsphänomenen in den Blick genommen, die sich zur Untersuchung der Einflussfaktoren eignen: Variation in der Konjugation zwischen starken und schwachen Verben (*geglommen/geglimmt*), Variation in der Deklination zwischen schwachen und starken Maskulina (*des Bären/Bärs*) und Variation in der Auxiliarselektion von *haben* und *sein* (*das Auto gefahren haben/sein*). Mit diesem Zuschnitt fokussiert die Arbeit auf morphologische und semantisch-syntaktische Variation. Die Faktoren, ihr Einfluss auf Variation und ihre Verknüpfung mit den Variationsphänomenen werden in Abschnitt 1.1 näher erläutert.

Auf empirischer Ebene schließt sich die Arbeit dem *open science movement* an (Cumming & Calin-Jageman 2017: 257–292): Für alle empirischen Studien bis auf eine¹ wurden die Hypothesen, das Untersuchungsdesign und die statistische Auswertung vorab online registriert. Im Rahmen der Präregistrierung wurden Prästudien durchgeführt und anhand der Prästudien eine Datensimulation vorgenommen, um sicherzustellen, dass die für die Hauptstudie angestrebte Stichprobengröße ausreichend ist, um die Wahrscheinlichkeit für ein zufällig negatives Ergebnis möglichst gering zu halten. Außerdem werden die erhobenen Daten und die Skripte, mit denen die Daten aufbereitet und ausgewertet wurden, zur Verfügung gestellt. Die Nutzung von frei zugänglichen Programmen für die Datenerhebung und -auswertung erleichtert zudem eine Replikation der Studien und führt dazu, dass die Datenauswertung im Detail nachvollzogen werden kann. Im Folgenden wird auf die in dieser Arbeit modellierten kognitiven Einflussfaktoren auf Variation – Frequenz, Prototyp und Schema – eingegangen und die Gliederung der Arbeit erläutert.

¹Die self-paced-reading-Studie zu Form-Schematizitätseffekten wurde nicht registriert. Sie wurde vor den anderen Studien durchgeführt. In der self-paced-reading-Studie wurde die Notwendigkeit für eine Präregistrierung inklusive einer Prästudie für die weiteren Studien deutlich.

1.1 Frequenz, Prototyp und Schema als Einflussfaktoren auf Variation

Frequenz, Prototyp und Schema sind zentrale Einflussfaktoren auf Variation und Stabilität in einem Sprachsystem. In Hinblick auf Frequenz ist zunächst festzustellen, dass hohe Tokenfrequenz zu gefestigten und leicht aktivierbaren Formen und damit zu Stabilität führt (Bybee & Thompson 1997: 380). Der Einfluss von Frequenz auf Variation ist aber nicht nur auf Tokenfrequenz beschränkt. Typenfrequenz beeinflusst Variation ebenfalls. Dabei ist bei Flexionsklassen ein Zusammenspiel aus Typen- und Tokenfrequenz zu beobachten, das sich anhand der Variation in der Konjugation gut veranschaulichen lässt: Schwache Verben weisen als typenfrequente Flexionsklasse viele und sehr unterschiedliche Mitglieder auf. Die Merkmale der Flexionsklasse treffen deswegen auf viele Verben zu und können daher abstrahiert werden (Bybee & Thompson 1997: 384, Goldberg 2019: 63–67): Wenn *-t* an einen Infinitivstamm gehängt wird, setzt man ein Verb in die Vergangenheit (*lachte*). Bei typeninfrequenten Klassen wie den starken Verben kann das Verhalten dagegen nur eingeschränkt abstrahiert werden, da es nur wenige Mitglieder gibt, die sich zudem nicht uniform verhalten. *Geben* wird bspw. durch den Ablaut /a:/ in die Vergangenheit gesetzt (*gab*), aber *heben* durch /o:/ (*hob*). Theoretisch sind bei *geben* und *heben* auch schwache Formen möglich (**gebte*, **hebte*). Dies verdeutlichen bspw. Übergeneralisierungen der schwachen Formen bei Kindern (Rumelhart & McClelland 1986: 219–220). Die Option der schwachen Formen (**gebte*, **hebte*) wird aber abseits von Übergeneralisierungen nicht genutzt (Goldberg 2019: 75–77), da mit *gab* und *hob* andere, häufigere Formen existieren. Hier zeigt sich der Einfluss von Tokenfrequenz: Nur wenn ein Verb aus einer typeninfrequenten Klasse tokenfrequent ist (wie bspw. *geben*), sind auch seine starken Formen (z. B. *gab*, *gegeben*) häufig. Es existieren somit sehr oft genutzte Formen, sodass die schwachen Formen (**gebte*, **gegeben*) statistisch ausgestochen werden und deren Verwendung daher unwahrscheinlich ist (Goldberg 2019: 83–84). Wird ein Verb dagegen selten genutzt wie bspw. *glimmen*, sind auch seine starken Formen (*glomm*, *geglommen*) nicht häufig. Die schwachen Formen (*glimmte*, *geglimmt*) werden daher weniger stark dominiert. In der Folge weist das tokeninfrequente Verb starke sowie schwache Formen (*geglommen*, *geglimmt*) auf und wechselt schließlich die Konjugationsklasse (Augst 1975: 267–268). Typenfrequenz beeinflusst daher die Schwankungsrichtung (von typeninfrequent zu typenfrequent). Hohe Tokenfrequenz führt zu Stabilität in der typeninfrequenten Klasse und niedrige Tokenfrequenz zu Variation.

1.1 Frequenz, Prototyp und Schema als Einflussfaktoren auf Variation

Um den Einfluss der Frequenz auf Variation psycholinguistisch zu überprüfen, werden im empirischen Teil der Arbeit starke Verben mit unterschiedlicher Tokenfrequenz in starker (*geglommen, geflogen*) und schwacher Flexion (*geglimmt, *gefliagt*) präsentiert und die Reaktionszeiten auf diese Formen gemessen. Die Verben unterteilen sich dabei in drei Gruppen: tokenfrequente Verben (*fliegen*), tokeninfrequente Verben mit Schwankung in Korpora (*glimmen*) und tokeninfrequente Verben ohne Schwankung in Korpora (*fechten*). Auf diese Weise kann der Einfluss von Tokenfrequenz auf die Verarbeitung gemessen werden und gleichzeitig überprüft werden, ob sich Variationspotential in der Prozessierung niederschlägt, bevor Schwankungen beobachtet werden können.

Der Einfluss von Prototypen auf Variation lässt sich anhand der Auxiliarselektion von *haben* und *sein* gut umreißen. Sätze mit einem Bewegungsverb (z. B. *fahren*) weisen je nach Transitivitätsgrad unterschiedliche Auxiliare auf (Gillmann 2016: 316–319): Prototypisch transitive Sätze mit einem Objekt, das als Patiens fungiert (*die Kinder zur Schule fahren*), selektieren *haben* als Auxiliar (*Ich habe/*bin die Kinder zur Schule gefahren*). Prototypisch intransitive Sätze ohne Objekt (*zur Kur fahren*) weisen dagegen *sein* auf (*Ich bin/*habe zur Kur gefahren*). Schwankungen zwischen den Auxiliaren treten nur in Sätzen auf, die im Übergangsbereich zwischen Transitivität und Intransitivität sind. Im Satz *Ich bin/habe das Auto gefahren* kann das Objekt als Patiens gesehen werden, da das Auto der Handlung ausgesetzt ist, aber auch als Instrument, da das Auto die Handlung erst ermöglicht (Gillmann 2016: 251). Da prototypisch transitive Sätze ein Patiens enthalten, führt die Interpretation als Patiens zu hoher Transitivität, die Interpretation als Instrument dagegen zu geringerer Transitivität. Durch die Ambiguität zwischen Patiens und Instrument ist *das Auto fahren* im Übergangsbereich zwischen Transitivität und Intransitivität, weswegen beide Auxiliare (*Ich bin/habe das Auto gefahren*) möglich sind. Prototypische Vertreter zeigen somit ein stabiles Verhalten, während im Übergangsbereich und damit jeweils in der Peripherie der Funktionen Variation auftreten kann.

In Bezug auf Variation in der Flexion ermöglicht Prototypizität außerdem, Flexionsklassen als prototypisch organisiert zu betrachten. Anhand von Flexionseigenschaften können prototypisch starke und prototypisch schwache Vertreter einer Flexionsklasse ermittelt werden. Dies lässt sich anhand der starken und schwachen Verben verdeutlichen: *Geben* ist prototypisch stark, da es alle Eigenschaften starker Verben (Imperativhebung *gib*, Wechselflexion im Präsens *er/sie gibt*, Ablaut *gab, gegeben*) vereint (Bittner 1985: 57, Nowak 2013: 157). *Sagen* ist prototypisch schwach, da es keine Imperativhebung und keine Wechselflexion aufweist und Präterital- sowie Partizip-II-Formen mit Dentalsuffix bildet (*sagte, gesagt*). Ein Verb wie *mahlen* ist ein peripheres Mitglied der starken Flexion, da

1 Einleitung und Zielsetzung

es nur das Partizip II stark bildet (*gemahlen*). Nicht nur die Mitglieder der starken Konjugationsklasse lassen sich als prototypisch und peripher beschreiben, sondern auch die Flexionseigenschaften der starken Konjugation. Der Ablaut im Perfekt ist eine prototypische Eigenschaft starker Flexion, da alle starken Verben sie aufweisen, die Imperativhebung hingegen eine periphere, da nur wenige starke Verben diese Flexionseigenschaft zeigen. Diese Prototypizitätsskala ermöglicht Vorhersagen über Variation: Periphere Eigenschaften wie die Imperativhebung werden zuerst aufgegeben, dann folgen prototypische wie der Ablaut (Bittner 1985: 57).

Der Einfluss der Prototypizität auf Variation wird im empirischen Teil der Arbeit anhand der Prozessierung der Auxiliarselektion von *haben* und *sein* überprüft. Dafür werden Sätze mit Bewegungsverben mit unterschiedlichem Transitivitätsgrad genutzt und die Reaktionszeiten auf die Auxiliare gemessen. Prototypisch transitive Sätze (*die Kinder zur Schule gefahren haben/*sein*) werden prototypisch intransitiven Sätzen (*zur Kur gefahren sein/*haben*) gegenüber gestellt. Zusätzlich werden zwei ambige Strukturen präsentiert: Bewegungsverben mit einem Fortbewegungsmittel als Objekt (*ein Cabrio gefahren haben/sein*) und Bewegungsverben mit einem Abstraktum als Akkusativergänzung, das zu einer Ambiguität zwischen Objekt und Adverbial führt (*die Strecke in Rekordzeit gefahren sein/haben*). Mithilfe dieses Designs kann die Reaktionszeit auf Auxiliare in prototypischen und damit eindeutigen Kontexten mit Reaktionszeiten in peripheren und damit ambigen Kontexten gegenübergestellt werden.

Schemata im Sinne von abstrakten Konstruktionen (Form-Funktions-Paaren) (Bybee 2010: 80) und ihr Einfluss auf Variation wurden bereits anhand der Prototypizität diskutiert: Die Form *haben* und die Funktion Transitivität bilden ein Schema, genauso wie die Form *sein* und die Funktion Intransitivität bei Bewegungsverben (Gillmann 2016: 316–319). Die Funktionen der beiden Schemata stehen in einer prototypischen Beziehung zueinander, wodurch Schwankungen im Übergangsbereich der Funktionen möglich sind.

Neben Schemata in diesem Sinn spielt für Variation eine zweite Form von Schemata eine Rolle. Diese werden in der vorliegenden Arbeit erstmals modelliert. Während bei der ersten Form von Schemata eine Form mit einer Funktion assoziiert ist, trifft bei der zweiten Form eine Funktion auf zwei (oder mehr) Formen, mit denen sie assoziiert sein kann. Dieses Prinzip lässt sich gut anhand der Variation in der Deklination veranschaulichen: So kann die Funktion [+Genitiv] durch starke *-(e)s*; *des Anglers*) und schwache Formen *-(e)n*; *des Matrosen*) ausgedrückt werden. Dies wird in der vorliegenden Arbeit als Form-Schematizität bezeichnet. Die Formen *-(e)s* und *-(e)n* verbinden sich mit verschiedenen Maskulina: Während Formen auf *-(e)s* mit einer Vielzahl von Maskulina kompatibel sind, sind

1.1 Frequenz, Prototyp und Schema als Einflussfaktoren auf Variation

Formen auf $-(e)n$ nur bei bestimmten Maskulina möglich: Sie sind prototypischerweise dreisilbig, auf der Pänultima betont, enden auf Schwa und referieren auf menschliche Entitäten (*Matrose, Geselle*) (Köpcke 1995: 168–170). Es steht somit eine abstrakte und damit hochschematische Form ($X-(e)s$) einer weniger abstrakten und damit teilschematischen Form ($X_{[\text{dreisilbig, Pänultima, Schwa, menschlich}]}-(e)n$) gegenüber. Die Assoziation aus teilschematischer Form und Funktion wird in der vorliegenden Arbeit als Form-Schema bezeichnet. Form-Schematizität führt zu Stabilität und zu Variation: Maskulina, die dem Form-Schema komplett entsprechen (z. B. *Matrose*) werden stabil schwach flektiert (Köpcke 2000a,b). Maskulina in der Peripherie des Form-Schemas schwanken hingegen zwischen starken und schwachen Formen (*des Grafs/Grafen; des Willen/Willens*): *Der Graf* referiert zwar auf eine menschliche Entität, erfüllt aber nicht die prosodisch-phonologischen Eigenschaften des Form-Schemas; *der Wille* entspricht zwar den prosodisch-phonologischen Eigenschaften, aber nicht der semantischen, da er auf ein Abstraktum referiert (Köpcke 2000a: 111). Zudem kann Variation entstehen, wenn Substantive, die nach der hochschematischen Form flektieren, Ähnlichkeiten mit der teilschematischen Form des Form-Schemas aufweisen, wie bspw. bei *Autor*. Hier kann es zu Schwankungen zur teilschematischen Form kommen (*des Autoren* statt *des Autors*) (Köpcke 2005: 78–79).

Der Einfluss von Form-Schematizität auf Variation wird im empirischen Teil der Arbeit anhand der Verarbeitung von schwachen Maskulina überprüft. Es werden Maskulina mit schwacher (*des Matrosen/des Grafen*) und schwacher Flexion (*des *Matroses/des Grafs*) präsentiert und die Reaktionszeiten auf die präsentierten Formen gemessen. Dabei werden starke (*Vogt*) oder schwache Maskulina (*Matrose/Graf*) genutzt. Die schwachen Maskulina entsprechen dem Form-Schema ($X_{[\text{dreisilbig, Pänultima, Schwa, menschlich}]}$) entweder prototypisch (*Matrose*) oder nur peripher (*Graf*). Die Gegenüberstellung dieser Formen ermöglicht es, den Einfluss des Form-Schemas auf Reaktionszeiten zu messen und gleichzeitig zu prüfen, inwiefern sich der prototypische Aufbau des Form-Schemas auf Reaktionszeiten niederschlägt.

Aus den bisherigen Ausführungen wird bereits deutlich, dass die Einflussfaktoren Frequenz, Prototyp und Schema nicht einzeln betrachtet werden können, da sie sich gegenseitig bedingen und somit in hohem Maße miteinander interagieren. Um den Einfluss der einzelnen Faktoren dennoch klar benennen zu können, werden sie in der Arbeit getrennt voneinander modelliert und ihr Zusammenspiel im Anschluss an die getrennte Betrachtung erläutert. Aus demselben Grund umfassen die psycholinguistischen Experimente jeweils nur einen Einflussfaktor. Die Arbeit möchte damit den Grundstein für weitere Forschung zum Zusammenspiel der Faktoren legen.

1.2 **Gliederung der Arbeit**

Die Arbeit gliedert sich in vier große Kapitel, zwei theoretische und zwei empirische: In Kapitel 2 werden Frequenz, Prototypizität und (Form-)Schematizität theoretisch fundiert und modelliert. Daran anschließend wird in Kapitel 3 der Einfluss von Frequenz, Prototypizität und (Form-)Schematizität in Bezug auf Variation in der Konjugation, Deklination und Auxiliarselektion anhand der bestehenden Forschung zu den Variationsphänomenen diskutiert. Die letzten beiden Kapitel überprüfen den Einfluss der vorgeschlagenen Faktoren psycholinguistisch, indem Reaktionszeiten gemessen werden. Dabei werden in Kapitel 4 zunächst die in den Studien genutzten Verfahren zur Elizitation von Reaktionszeiten sowie das Untersuchungsdesign der Studien vorgestellt, anschließend werden in Kapitel 5 die Ergebnisse der Studien präsentiert und diskutiert.

Die vier Kapitel werden durch das abschließende Kapitel 6 abgerundet, das die theoretischen und methodischen Überlegungen zusammenfasst und einen Ausblick auf weitere Forschungsfragen gibt.

Mit diesem Zuschnitt bietet die Arbeit einen neuen theoretischen und methodischen Zugriff auf Variation: Durch den gebrauchsbasierten Ansatz in dieser Arbeit lassen sich Verbindungen zwischen einzelnen Variationsphänomenen aufzeigen und die Mechanismen hinter der Variation besser begreifen. Der Einbezug von Sprachprozessierung erlaubt neue Einblicke in Variation, die die Betrachtung von Sprachproduktion allein nicht ermöglicht. Daher ergänzt die vorliegende Arbeit bestehende Forschung zu Variation um die Perspektive der Sprachprozessierung. Dabei wird mit drei Verfahren zur Reaktionszeitmessung gearbeitet, um einen breiten Zugriff auf Reaktionszeiten zu ermöglichen. Die Arbeit leistet damit einen wichtigen Beitrag zur Theorie und Empirie der variationslinguistischen Forschung.

2 Variation aus gebrauchsbasierter Perspektive

In diesem Kapitel wird sprachliche Variation aus gebrauchsbasierter Perspektive beleuchtet. Dabei steht die Frage im Vordergrund, wie Stabilität und Variation in einer Sprache theoretisch modelliert werden kann. Um diese Frage zu beantworten, werden Frequenz, Prototypizität und (Form-)Schematizität als Faktoren eingeführt, die Einfluss auf Sprachstabilität, -variation und -wandel nehmen. Dabei wird auf grammatische Variation fokussiert. Der Einfluss von Frequenz, Prototypizität und (Form-)Schematizität wird in den folgenden Abschnitten diskutiert.

2.1 Frequenz

Der gebrauchsbasierte Ansatz sieht Frequenz als einen grundlegenden Einflussfaktor für die Evolution und den Erwerb eines Sprachsystems (Kemmer & Barlow 2000: x, Diessel 2017: 1–2; siehe hierzu ausführlich Tomasello 2009 für eine Modellierung des Spracherwerbs aus gebrauchsbasierter Perspektive). Wie bereits in der Einleitung dargelegt wurde, setzt der Ansatz Sprachsysteme mit mentalen Repräsentationen gleich. Diese entstehen aus dem Sprachgebrauch, genauer gesagt aus konkreten Verwendungen von Konstruktionen (Form-Funktions-Paaren): Aus den Verwendungen (z. B. *Die Katzen trinken Milch, die Kinder schlürfen Suppe, die Hasen essen Karotten*) werden gemeinsame Eigenschaften (X_{NP_Nom} X_{VVFIN} X_{NP_Akk} , + Nahrungsaufnahme) der Konstruktionen abstrahiert (Tomasello 1992: 20–29, Ellis 2002a: 168–170, Langacker 2008: 220). Frequenz ist schon für die Grundkonzeption des Ansatzes relevant, denn die mentale Repräsentation einer häufig verwendeten Konstruktion ist gefestigter als die einer selten genutzten (Bybee & Thompson 1997: 380). Zudem ist eine Konstruktion, die Sprecher_innen in vielen verschiedenen Verwendungen begegnet, abstrakter und somit schematischer als eine, die auf wenige Kontexte beschränkt ist (Goldberg 2019: 67–68; Konstruktionen und ihre Abstraktionsgrade werden in Abschnitt 2.3.2 diskutiert). Frequenz stärkt somit die mentale Repräsentation einer Konstruktion und erleichtert dadurch deren Aktivierung, sodass sie einen grundlegenden Einfluss auf Grammatik nimmt (Diessel 2017: 1–2).

2 Variation aus gebrauchsbasierter Perspektive

Die Zipfsche Verteilung verdeutlicht, wie stark Sprache durch Frequenz beeinflusst wird: Wenn man Wörter eines Texts nach ihrer Frequenz ordnet und anhand der Frequenz Ränge verteilt, ist die Wahrscheinlichkeit für das Auftreten eines Worts umgekehrt proportional zu seinem Rang (Zipf 1972 [1949]: 22–27, Ellis 2012: 13). Das zweithäufigste Wort (Rang 2) ist also nur halb so häufig wie das häufigste Wort (Rang 1) und das dritthäufigste Wort (Rang 3) ist nur noch ein Drittel so häufig wie das häufigste Wort. Aus der Zipfschen Verteilung folgt, dass in einer Sprache nur wenige hochfrequente Wörter existieren, aber viele wenig frequente. Abbildung 2.1 zeigt exemplarisch die Frequenz der einzelnen Token in der Einleitung dieser Dissertation, die nach ihrem Rang sortiert sind.¹

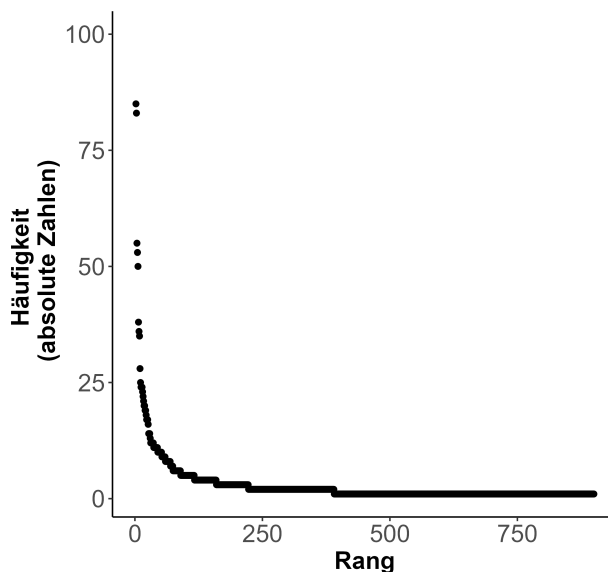


Abbildung 2.1: Frequenz der Token in der Einleitung nach ihrem Rang

Auf Rang 1 und damit das häufigste Token ist *und* (106 Belege), gefolgt von *die* (85 Belege) und *der* (83 Belege) und *auf* (55 Belege). Auf Rang 5 folgt mit 53 Belegen das Substantiv *Variation*, dessen Häufigkeit sich aus dem Thema der Dissertation ergibt. Die Ränge stehen nicht ganz in einem umgekehrt proportionalem Verhältnis zueinander, allerdings ist ein deutlicher Frequenzabfall von Rang 1 zu Rang 5 zu beobachten. Ab Rang 45 weisen die Token nur noch maximal zehn Belege pro Rang auf.

¹Die Skripte zur Erstellung der Grafik sind im digitalen Anhang dokumentiert. Für die Aufbereitung der Daten wurde das Paket *NLTK* (NLTK-Team 2020) genutzt.

Die Zipfsche Verteilung lässt sich nicht nur für einzelne Token beobachten, sondern auch für Phoneme, Konstruktionen² und idiomatische Phrasen (Ellis 2008: 235). Die Frequenzkurve in Abbildung 2.1 illustriert zudem, dass grammatische Elemente häufiger sind als lexikalische (Gervain u. a. 2008: 58–59), denn die häufigsten drei Token in der Kurve stellen grammatische Elemente dar. Frequenz kann somit den Zugriff auf syntaktische Muster ermöglichen: Die Abfolge frequent–infrequent weist bpsw. auf ein Grammem hin, dem ein Lexem folgt (bspw. Determinierer + Bezugsnomen). Gervain u. a. (2008) zeigen mithilfe eines *head turn paradigms*³, dass Kinder mit acht Monaten bereits zwischen frequenten und infrequenten Wörtern unterscheiden und so syntaktische Muster erkennen können. Zusätzlich zur Frequenz nutzen Kinder hierfür prosodische Hinweise im Input (Gervain & Werker 2013). Die frühe Sensitivität für Frequenz von Kindern ist ein Hinweis darauf, dass Spracherwerb unter anderem über statistisches Lernen (*statistical learning*) funktioniert (Saffran 2003: 2003, siehe hierzu auch Abschnitt 2.1.3).

Im Folgenden wird zunächst der Unterschied zwischen Typen- und Tokenfrequenz eingeführt. Typen- und Tokenfrequenz sind grundlegend für diese Arbeit. Sie werden im späteren Verlauf der Arbeit zentral sein, da die Unterscheidung zwischen Typen- und Tokenfrequenz genutzt wird, um den Einfluss der Frequenz auf Variation zu modellieren. Im Anschluss an die Unterscheidung zwischen Typen- und Tokenfrequenz setzt sich Abschnitt 2.1.2 kritisch mit Frequenzeffekten auseinander. Dabei werden Frequenzeffekte in der Prozessierung

²Goldberg (2006: 75) stellt fest, dass einzelne verbale Konstruktionen von der Verwendung mit einem bestimmten Verb dominiert werden: Im Korpus von Bates u. a. (1988), in dem Äußerungen von Müttern an 28 Monate alte Kinder aufgenommen wurden, macht *to put* 38 % der Verwendungen innerhalb der caused-motion-Konstruktion aus (*David puts the cheese into the fridge*).

³Bei dem *head turn paradigm* werden Kindern Stimuli aus Lautsprechern vorgespielt, die sich links und rechts vom Kind befinden. Dabei wird die Zeitspanne gemessen, in der das Kind seinen Kopf in Richtung des Lautsprechers dreht, aus dem der Stimulus zu hören ist (Gervain & Werker 2013: 4). Im Experiment von Gervain u. a. (2008: 68–69) wurde den Kindern in der Eingewöhnungsphase eine artifizielle Sprache vorgespielt, die aus einer Abfolge aus häufigen und seltenen Elementen bestand. Die artifizielle Sprache begann dabei entweder mit einem frequenten (frequent-infrequent) oder einem infrequenten Element (infrequent-frequent). In der Testphase wurde sowohl die bereits bekannte Sprache als auch eine neue vorgespielt, die jeweils die andere Reihenfolge aufwies (Gervain u. a. 2008: 69–70). Dabei wurde getestet, ob es einen Unterschied in der Zeitspanne gibt, in der die Kinder den Kopf zum Stimulus drehen. Dies war bei dem Experiment der Fall: Kinder mit Italienisch als L1 schauten länger zum frequenz-initialen Stimulus, während Kinder mit Japanisch als L1 länger zum frequenz-finalen Stimulus blickten (Gervain u. a. 2008: 70). Die Präferenz für einen Stimulus reflektiert die Reihenfolge von frequenten und infrequenten Elementen im Italienischen bzw. Japanischen.

2 Variation aus gebrauchsbasierter Perspektive

vorgestellt und die Korrelation von Frequenz mit anderen Faktoren (bspw. Kontextdiversität) problematisiert: Durch die Korrelation mit zahlreichen Faktoren kann Frequenz zwar als ein guter Prädiktor für menschliches Verhalten angesehen werden, aber nicht zwingend als Erklärung für dieses Verhalten dienen. Im Anschluss daran wird in Abschnitt 2.1.3 ein Ausweg aus diesem Problem aufgezeigt. Frequenz wird als ein wichtiger Einflussfaktor innerhalb des statistischen Lernens modelliert: Statistisches Lernen beruht auf Wahrscheinlichkeiten und Frequenz nimmt Einfluss auf die Wahrscheinlichkeit für das Auftreten einer Struktur. Auf diese Weise kann der Einfluss von Frequenz samt korrelierenden Faktoren modelliert und anhand der Modellierung Vorhersagen über den Einfluss von Frequenz auf Variation gemacht werden.

2.1.1 Typen- und Tokenfrequenz

Die Zipfsche Verteilung verdeutlicht, wie grundlegend Frequenz den Sprachgebrauch beeinflusst: Es existieren nur wenige hochfrequente Elemente, aber viele niedrigfrequente. Frequenz nimmt daher mittelbar durch den Sprachgebrauch auch Einfluss auf die mentale Repräsentation sprachlicher Einheiten. Frequenz kann dabei mit Divjak & Caldwell-Harris (2015: 54) gefasst werden als „a practical term that was, and still is, used to capture how frequently a stimulus (such as a word or a phrase) is encountered and processed in the environment“. In dieser Definition wird deutlich, dass der zu zählende⁴ Stimulus stets zu definieren ist. Vor diesem Hintergrund stellt sich also die Frage, welche Einheiten gezählt werden: So können bspw. *bist* und *ist* als zwei Zählheiten gefasst werden, da es sich um verschiedene Wortformen handelt, oder als eine einzelne Zählheit, da beide Wortformen Teil des Paradigmas von *sein* sind. Hierbei ist die Unterscheidung zwischen Typen- und Tokenfrequenz grundlegend (Bybee & Thompson 1997: 378).

Tokenfrequenz meint die Häufigkeit, mit der Sprecher_innen einer bestimmten Wortform begegnen wie bspw. *lachte*. Auf dieser Zählweise beruht die Frequenzkurve in Abbildung 2.1. Typenfrequenz blickt hingegen auf die Anzahl verschiedener Elemente, die innerhalb eines Musters, Paradigmas, einer Kategorie

⁴Neben dem Ermitteln der Frequenz durch Zählen eines Elements in einem Korpus oder mehreren Korpora erwähnen Divjak & Caldwell-Harris (2015: 54) die Frequenzeinschätzung von L1-Sprecher_innen. Die Ergebnisse der beiden Methoden stellen sie als objektiv bzw. subjektiv dar. Hierbei ist zu beachten, dass auch die Wahl des Korpus Frequenzen verzerren kann, sodass die Frequenzbestimmung aufgrund eines Korpus nicht notwendigerweise als objektiv gelten kann.

oder Konstruktion⁵ auftreten können (Ellis 2002a: 166). Das Tempusparadigma von *lachen* enthält bspw. die Elemente *lachte*, *lache*, *gelacht*. Formen innerhalb eines Flexionsparadigmas können also Token eines Types sein. Allerdings lassen sich auch schwache Verben (*lachen*, *machen*) als Token des Types SCHWACHES VERB auffassen (zum Einfluss der Typen- und Tokenfrequenz auf die Variation in der Konjugation siehe Abschnitt 3.1.1).⁶ Die Unterscheidung von Typen- und Tokenfrequenz ist für diese Arbeit grundlegend, da sowohl die Typen- als auch die Tokenfrequenz Stabilität und Variation beeinflussen: Wie anhand der Ausführungen ersichtlich werden wird, ist Variation für tokeninfrequente Mitglieder einer typeninfrequenten Kategorie am wahrscheinlichsten, wenn eine typenfrequente Kategorie existiert, der sie sich in ihrem Verhalten anschließen können.

Bybee & Thompson (1997: 384) gehen davon aus, dass die Typenfrequenz einer Kategorie deren Produktivität determiniert. Sie definieren Produktivität dabei als „likelihood that a pattern will be applied to new forms“ (Bybee & Thompson 1997: 384). Die Interaktion von Typenfrequenz und Produktivität machen sie an drei Punkten fest:

1. Je mehr Elemente eine Kategorie hat, desto unwahrscheinlicher ist es, dass die Kategorie nur mit einem einzigen Element assoziiert wird, sondern dass eine generelle Kategorie mit den Elementen geformt wird, die zu der Kategorie gezählt werden. Für die typenfrequenten regelmäßigen Verben ist

⁵Sprachliche Kategorien und Konstruktionen stehen in einem engen Verhältnis: Konstruktionen sind Form-Funktions-Paare. Der Terminus *Kategorie* hebt auf Gemeinsamkeiten zwischen mehreren Konstruktionen oder mehreren Verwendungen einer Konstruktion ab. Die Verbindung der Buchstabenkombination <Katze> mit der Bedeutung ‚vierbeiniges Tier, das miaut‘ ist daher eine Konstruktion, aber auch eine Kategorie, weil unter dem Form-Funktions-Paar <Katze> und ‚vierbeiniges Tier, das miaut‘ mehrere Entitäten zusammengefasst werden, auf die die Funktion zutrifft. Dasselbe gilt für die Verbindung der Buchstabenkombination <Substantiv> mit Wörtern, die bestimmte syntaktische und morphologische Eigenschaften aufweisen. *Konstruktion* wird in der vorliegenden Arbeit genutzt, wenn die Interaktion zwischen Form und Funktion im Vordergrund steht. *Kategorie* findet Verwendung, wenn die gemeinsamen Eigenschaften relevant sind.

⁶Einen Überblick zu weiteren Zählmethoden bietet Lerche (2019: 15–21). Dabei differenziert sie Zählmethoden für die Tokenfrequenz eines Lemmas und eines Basislemmas: Bei der Tokenfrequenz eines Lemmas gelten *abweisen* und *verweisen* als zwei Lemmata; bei der Tokenfrequenz des Basislemmas hingegen als Token des Lemmas *weisen* (Lerche 2019: 15–18). Innerhalb der Typenfrequenz weist sie auf die Notwendigkeit hin, intraparadigmatische Frequenzen zu bestimmen. Dabei gilt es zu ermitteln, a) wie viele Formen eines Paradigmas eine bestimmte Eigenschaft aufweisen (bspw. *e* als Stammvokal) und b) wie häufig eine bestimmte Form innerhalb des Paradigmas (bspw. der Imperativ) genutzt wird. Beide Frequenztypen können Einfluss auf die analoge Anpassung unregelmäßiger Verben hin zum regelmäßigen Paradigma haben (Lerche 2019: 19–21). Sie werden daher in den Abschnitten 3.1 und 3.2 zur Variation in der Konjugation und Deklination wieder aufgegriffen.

2 Variation aus gebrauchsbasierter Perspektive

es also unwahrscheinlich, dass nur das Exemplar *lachen* mit regelmäßiger Flexion durch Dentalsuffix verknüpft ist. Vielmehr ist davon auszugehen, dass eine allgemeine Kategorie geformt wird. Bybee & Thompson (1997: 384) betonen dabei, dass Typenfrequenz nicht diskret einzuteilen, sondern graduell ist, da die Typenfrequenz einer Kategorie von einem einzigen Type (*sein*) bis zu einer hohen Anzahl von Types schwanken kann.

2. Je mehr Elemente eine Kategorie enthält, desto abstrakter muss die Kategorie sein. Daher ist es für eine Kategorie mit vielen Elementen wahrscheinlicher, dass neue Elemente zu der Kategorie hinzu kommen (Bybee & Thompson 1997: 384). Die Konstruktion X_{NP_Nom} *werden* X_{VVP} (*er wird getragen, wir werden gefragt*) lässt bspw. jedes transitive Verb zu. Die Kategorie ist daher abstrakt. Die Konstruktion X_{NP_Nom} *bekommen* X_{VVP} (*Die Kinder bekommen eine Geschichte vorgelesen/Ich bekomme das Geld überwiesen*) ist dagegen semantisch eingeschränkt, bspw. sind keine abstrakten Subjekte möglich (Wegener 1985: 133–134), sodass der Verb-Slot dieser Konstruktion weniger produktiv ist als der Verb-Slot der *werden*-Passiv-Konstruktion.
3. Hohe Typenfrequenz bedeutet auch, dass eine Kategorie häufig genutzt wird. Dementsprechend wird ihre mentale Repräsentation gestärkt, sodass die Kategorie leichter aufzurufen ist und mit neuen Elementen genutzt werden kann. Die mentale Repräsentation wird dabei vor allem durch die tokeninfrequenten Mitglieder einer Kategorie gestärkt (Bybee 2013: 62): Die Prozessierung der tokeninfrequenten Mitglieder setzt voraus, dass die Kategorie aktiviert wird, sodass sie gestärkt wird. Tokenfrequente Mitglieder können hingegen prozessiert werden, ohne die gesamte Kategorie aufzurufen und dadurch zu stärken. Tokenfrequente Mitglieder einer Kategorie können deren mentale Repräsentation somit nur bedingt stärken (Bybee 2013: 62).

Bybee (2013: 62) ergänzt diese Punkte mit dem Hinweis darauf, dass in einem exemplar-basierten Lernmodell (*exemplar based model*, siehe Abschnitt 2.2.3) hohe Typenfrequenz einer Kategorie bedeutet, dass auch mehr Ankerpunkte für neue Vertreter der Kategorie existieren. Auch dies verdeutlicht den eingeschränkten Einfluss von Tokenfrequenz auf Produktivität: Die hohe Tokenfrequenz eines Vertreters führt nicht zu neuen Anknüpfungspunkten. Wie Bybee & Thompson (1997: 384) verknüpft auch Goldberg (2019: 62–63) Typenfrequenz mit Produktivität, indem sie die Abdeckung (*coverage*) einer Kategorie als Einflussfaktor auf

Produktivität vorschlägt. Dabei konzeptualisiert Goldberg (2019: 16) Kategorien über Cluster: Vertreter einer Kategorie weisen Ähnlichkeiten in ihrer mentalen Repräsentation auf. Aufgrund dieser Ähnlichkeiten werden sie geclustert und so als Einheit wahrgenommen (Goldberg 2019: 36–37). Cluster können dabei nicht nur durch semantische Ähnlichkeit, sondern bspw. auch durch phonologische⁷ und kontextuelle Ähnlichkeit entstehen (auf die exemplarbasierte Entstehung der Cluster wird in Abschnitt 2.2.3 näher eingegangen). Goldberg (2019: 16–17) situiert die Cluster daher in einem hyperdimensionalen konzeptuellen Raum (*hyperdimensional conceptual space*).

Mithilfe des Konzepts der Abdeckung lässt sich evaluieren, wie gut potentielle neue Vertreter zu einer Kategorie passen: Für potentielle Vertreter wird überprüft, wie gut die neue Kategorie, die aus den bereits beobachteten Vertretern und dem neuen Vertreter besteht, durch das Cluster im hyperdimensionalen Raum abgedeckt wird: Wenn der Neuzugang zu wenig Ähnlichkeit zu den bereits attestierten Exemplaren bspw. hinsichtlich der phonologischen Form aufweist, ist die neue Kategorie nur lückenhaft abgedeckt, weil zwischen den neuen und den alten Exemplaren verbindende Exemplare fehlen.⁸ Wenn der Neuzugang dagegen viel Ähnlichkeit mit bereits bestehenden Exemplaren aufweist, ist die Kategorie gut abgedeckt. Abbildung 2.2 veranschaulicht das Prinzip anhand der Verben der Ablautreihe 3a und schwachen Verben.



Abbildung 2.2: Beispiel für Kategorien mit lückenhafter und lückenloser Abdeckung

Die Verben der Ablautreihe 3a haben mit /ɪ/ gefolgt von einem Nasal und einem weiteren Konsonanten ähnliche phonologische Eigenschaften und clustern

⁷In exemplarbasierten Modellen wird sowohl von phonetischen als auch phonologischen Clustern ausgegangen (Johnson 2007: 27–29). Phonologische Cluster sind dabei als Generalisierung phonetischer Cluster zu sehen (Johnson 2005, Kirchner u. a. 2010). Im Folgenden wird der Terminus *phonologisch* genutzt, der somit bereits eine Generalisierung beinhaltet.

⁸Das Konzept der Abdeckung funktioniert daher ähnlich wie Familienähnlichkeit (siehe hierzu genauer Abschnitt 2.2.1).

2 Variation aus gebrauchsbasierter Perspektive

daher stark (ausführlich zu den phonologischen Eigenschaften der Verben und dem sich daraus ergebenden Form-Schema siehe Abschnitt 3.1.3). Für ein potentiell neues Verb (*googeln*) ergibt sich daher eine Kategorie, die lückenhaft abgedeckt ist: *Googeln* ähnelt den Verben der Ablautreihe 3a bis auf die Zweisilbigkeit und den gemeinsamen Endkonsonanten *-n* nicht. Dies ist bei den schwachen Verben anders. Da die Vertreter dieser Flexionsklasse ein hochvariables Cluster bilden, kann *googeln* Teil des Clusters werden, ohne dass die Abdeckung der Kategorie sich verschlechtert.

Neben der Typenfrequenz nehmen also weitere Faktoren Einfluss auf die lückenlose Abdeckung der neuen Kategorie und somit auf die Produktivität: die Variabilität der bereits beobachteten Exemplare und die Ähnlichkeit des potentiellen neuen Exemplars mit den alten. Dabei korrelieren Typenfrequenz und die Variabilität einzelner Exemplare einer Kategorie zu einem bestimmten Grad: Mitglieder einer Kategorie, die 1.000 Mitglieder hat, werden wahrscheinlich eine höhere Variabilität aufweisen als Mitglieder einer Kategorie, die nur zehn Mitglieder aufweist (Goldberg 2019: 65). Jedoch kann eine (relativ) typenfrequente Kategorie auch wenig variabel sein, wie bspw. die Verben der Ablautreihe 3a mit der Form [#_I + Nasal + (C)]. Diese kann dann weniger leicht auf Exemplare ausgeweitet werden, die außerhalb der bereits beobachteten Variabilität liegen (Goldberg 2019: 67). Außerdem interagieren Variabilität und Ähnlichkeit miteinander: Je variabler die Mitglieder einer Kategorie sind, desto wahrscheinlicher ist es, dass ein neues Exemplar Ähnlichkeit zu bereits attestierten Exemplaren aufweist.

Als weiteren Einflussfaktor auf Produktivität lässt sich die Existenz einer alternativen Formulierung anführen, die dasselbe ausdrückt und leichter zu aktivieren ist als die potentielle neue Formulierung (Goldberg 2019: 70; siehe hierzu genauer Abschnitt 2.1.3). Produktivität wird somit von zwei Faktoren beeinflusst: Abdeckung der Cluster (Typenfrequenz, Variabilität und Ähnlichkeit) und Wettbewerb in der Aktivierung (Goldberg 2019: 117).

Konstruktionen, deren Form jeweils mit derselben Funktion assoziiert ist,⁹ unterscheiden sich i. d. R. in ihrer Abdeckung voneinander. Dies ist bspw. bei verschiedenen Flexionsklassen der Fall. Die Flexionsklassen haben verschiedene Formen (wie bspw. *lachte* und *sank*), aber die Funktion der Formen ist jeweils gleich (in diesem Fall ‚Vergangenheit‘). Dabei weist eine Flexionsklasse eine hohe Abdeckung auf (schwache Verben), während die andere Flexionsklasse eine niedrige Abdeckung aufweist (starke Verben). Die Flexionsklasse mit hoher Abdeckung hat ein transparentes Flexionsmuster, das für alle Token des Types gilt:

⁹Diese Assoziation von einer Funktion mit mehreren Formen wird in Abschnitt 2.3.2 als Form-Schematizität bezeichnet und näher ausgeführt.

Die Funktion ‚Vergangenheit‘ ist für alle Token mit der Form *-t* verbunden. Das Verhalten innerhalb der typeninrequenten Klasse (z. B. /i:/ und /u:/ oder /a/ und /ʊ/ als Ablaute im Präteritum und Partizip II) ist hingegen nicht verallgemeinerbar, sondern auf spezifische Token (wie z. B. *rufen*, *rief*, *gerufen*) oder Token mit phonologischen Gemeinsamkeiten (Verben der Ablautreihe 3a) beschränkt. Verben mit phonologischen Gemeinsamkeiten, die dasselbe Flexionsverhalten aufweisen, werden in Abschnitt 2.3.2 diskutiert und als Form-Schema bezeichnet.

Die Generalisierbarkeit eines Flexionsmusters korreliert somit mit der Abdeckung der Kategorie (Goldberg 2019: 67–68). Die Abdeckung legt dabei den Inhalt der Generalisierung fest (Goldberg 2019: 136). Im Fall der schwachen Verben ist der hyperdimensionale Raum hinsichtlich der Phonologie und Semantik weit abgedeckt und abgesehen vom Dentalsuffix und der Bedeutung [+Vergangenheit] nicht restringiert. Diese Restriktionen sind der Inhalt der Generalisierung. Eine solch umfassende Generalisierung ist für typeninfrequente Flexionsklassen nicht möglich, da sie nur wenige Types mit geringer Variabilität enthalten. Das Ablautmuster der Verben der Ablautreihe 3a ist bspw. an eine spezifische, prototypisch organisierte phonologische Form gekoppelt und daher nur auf Verben mit dieser Form generalisierbar (Bybee 1985: 132–133; siehe Abschnitt 2.3.2 zu Form-Schemata).

Die hohe Tokenfrequenz starker Verben ermöglicht Unregelmäßigkeit (Ellis 2002a: 166). Durch die ständige Nutzung ist die mentale Repräsentation frequenter Formen *entrenched* (gefestigt): Je häufiger eine Aktivität ausgeführt wird, desto stärker wird sie mental *entrenched*, sodass jede Ausführung die nächste Ausführung erleichtert (Bybee & Thompson 1997: 380, Schneider 2014: 13). Kemmer & Barlow (2000: x) bezeichnen *entrenchment* daher als „cognitive routinization“. Aufgrund der hohen Frequenz sind die Formen also trotz Unregelmäßigkeit leicht zu prozessieren. Unregelmäßige Formen bringen zudem den Vorteil mit sich, dass sie kurz und gut unterscheidbar sind (Werner 1989: 42–43, Nübling 2000: 226–227). Diese Eigenschaften erleichtern wiederum eine ökonomische Prozessierung hochfrequenter Elemente (Nowak 2013: 174). Zudem ist davon auszugehen, dass die Wortformen frequenter Verben weniger stark untereinander verbunden sind als die Wortformen niederfrequenter Verben (Bybee 1985: 117–124): Die Formen *gehen* und *ging* sind also weniger stark miteinander verbunden als *lachen* und *lachte*. Frequente Verben haben in dieser Betrachtung ein weniger eng zusammenhängendes Paradigma als infrequente Verben. Bybee (2006: 715) bezeichnet dies als Autonomie-Effekt: Hochfrequente Elemente einer Konstruktion entwickeln einen autonomen Chunk und aktivieren die Ursprungs konstruktion nicht, weshalb hochfrequente Vertreter die Produktivität einer Kategorie nicht stärken (Bybee 2013: 62). Auch nach dem Konzept der Abdeckung von Goldberg

2 Variation aus gebrauchsbasierter Perspektive

(2019) tragen frequente Elemente generell nur eingeschränkt zur Produktivität einer Konstruktion bei, weil sie die Abdeckung der Kategorie nicht erhöhen. Die Tokenfrequenz stärkt lediglich die mentale Repräsentation des frequenten Elements (Goldberg 2019: 68).¹⁰

Bei schwachen und starken Verben lässt sich der unterschiedliche Kohärenzgrad des Paradigmas auch in der Phonologie erkennen. Die Formen starker Verben weisen weniger phonologische Ähnlichkeiten auf als schwache Verben. Daher kann auf lautlicher Ebene kein kohärentes Cluster entstehen, sodass die einzelnen Wortformen eher als unabhängig voneinander wahrgenommen werden können. Abbildung 2.3 veranschaulicht dies anhand der Verben *sagen* und *biegen*.



Abbildung 2.3: Cluster der Verbformen bei schwachen und starken Verben

Wie aus Abbildung 2.3 deutlich wird, clustern die Verbformen des schwachen Verbs *sagen* nicht nur semantisch, sondern auch lautlich. Dies ist bei dem starken Verb *biegen* nur bedingt der Fall: Durch den Ablaut ändert sich die Form deutlich. Das Dentalsuffix *-t* beeinflusst die Form hingegen nur leicht. Noch deutlicher wird der Unterschied bei hochfrequenten Verben mit Suppletivparadigmen wie bspw. *sein*, bei denen zwar eine semantische Beziehung zwischen den einzelnen Wortformen besteht, aber keine phonologische (Bybee 2007: 171).¹¹ Aufgrund der

¹⁰Allerdings können frequente Formen im Spracherwerb als Analogievorlagen dienen und somit zur Produktivität von Kategorien beitragen (Ellis 2016: 242). Dabei ist davon auszugehen, dass tokenfrequente Elemente im Spracherwerb zur analogischen Erweiterung einer noch rudimentären Kategorie dienen. Je mehr die Kategorie gefestigt und abstrahiert wird, desto geringer ist der Einfluss der Tokenfrequenz einzuschätzen.

¹¹Bybee (1985: 118–129) argumentiert nicht über gebrauchsbasierte Cluster, geht aber prinzipiell ähnlich in ihrer Argumentation vor. Sie geht in ihrem dynamischen Modell lexikalischer Repräsentation davon aus, dass zwischen Wortformen ein morphologisches Verhältnis besteht, das sich durch semantische und phonologische Verbindungen ergibt (Bybee 1985: 118–129). Die semantischen und phonologischen Verbindungen ließen sich auch als Cluster aus ähnlichen Exemplaren fassen. Die semantische Verbindung zwischen Wortformen eines Paradigmas wird

geringen bis nicht vorhandenen phonologischen Beziehung sind hochfrequente Wortformen weniger stark mit ihrer Grundform assoziiert als niedrigfrequente und werden somit auch weniger stark analysiert (Bybee 1985: 118). Bybee (1985: 117–118) beschreibt diese Autonomie hochfrequenter Formen als lexikalische Stärke.

Für die Autonomie hochfrequenter Elemente ist zudem relevant, dass die einzelnen Verwendungen des hochfrequenten Elements in einer Konstruktion ein gesondertes Subcluster bilden (Goldberg 2019: 69–70). Dieses Subcluster kann daher als eigenständiges Cluster angesehen werden. Effekte der Autonomie hochfrequenter Elemente lassen sich auch innerhalb von Konstruktionen beobachten: Elemente, die häufig in einem Slot einer Konstruktion genutzt werden, können die Variabilität in anderen Slots der Konstruktion erhöhen. Als Beispiel nennt Goldberg (2019: 69) *to make* als frequentes Verb in der Resultativ-Konstruktion, das viele Adjektive zulässt (*to make someone/oneself sick/ill/happy/crazy*). Das weniger frequente *eat* ist in dieser Konstruktion hingegen auf die Verwendung mit *ill* beschränkt (*to eat oneself ill*), auch das semantisch verwandte Adjektiv *sick* ist nicht möglich (Goldberg 2019: 69). Die erhöhte Variabilität in einem Slot, die durch ein hochfrequentes Element in einem anderen Slot bedingt ist, führt somit auch zu einer größeren Produktivität.

Der Einfluss der Tokenfrequenz auf mentale Repräsentationen wurde im obigen Absatz bereits angesprochen. Zusätzlich zum Autonomie-Effekt diskutieren Bybee & Thompson (1997: 378–384) den konservierenden Effekt: Die Bedeutung hochfrequenter Wörter verändert sich kaum (Ellis 2012: 12–13), zudem sind hochfrequente Wortformen vor analogischem Ausgleich (*analogical leveling*) geschützt. Daher behalten hochfrequente unregelmäßige Verben ihre Flexion bei (*schlafen, schlief, geschlafen; sleep, slept*), während infrequente unregelmäßige Verben zur regelmäßigen Flexion wechseln (*gequollen/gequellt; wept/weeped*) (Bybee 1985: 119–120, Bybee & Thompson 1997: 280; siehe hierzu auch Abschnitt 3.1.1 zur Variation in der Konjugation). Dabei ist zu betonen, dass

durch die Relevanzskala beeinflusst (Bybee 1991: 79–81): Relevante Kategorien verändern die Semantik stärker als weniger relevante Kategorien. Eine Änderung einer relevanten Kategorie (Tempus) führt somit zu einer weniger engen Verbindung als die Änderung einer weniger relevanten Kategorie (Person). Schwächere semantische Verbindungen können in den phonologischen Verbindungen gespiegelt werden bspw. durch eine Änderung des Stammvokals über Ablautung (Bybee 2007: 170). Auch dies lässt sich durch Cluster fassen: Aufgrund des Ablauts und der großen Änderung in der Semantik ist das Cluster weniger kohärent als Cluster mit gleich bleibendem Vokal und nur geringer Änderung der Semantik. Neben Semantik und Phonologie beeinflusst die Tokenfrequenz das morphologische Verhältnis: Die Verbindung zwischen Wortformen eines Lexems ist bei hochfrequenten Lexemen loser als bei niedrigfrequenten. Ausführlich zum dynamischen Modell lexikalischer Repräsentation siehe Bybee (1985).

2 Variation aus gebrauchsbasierter Perspektive

der Blick auf reine Tokenfrequenz nicht ausreicht, um das Ausbleiben eines Flexionswechsels zu erklären. Wie aus den obigen Erläuterungen deutlich wurde, basiert der Frequenzeffekt auf *entrenchment*: Durch die ständige Nutzung entsteht eine gefestigte mentale Repräsentation, die alle Vorkommnisse eines Tokens clustert. Aber auch die Token selbst können zu einer Kategorie geclustert werden (wie bspw. die Verben der Ablautreihe 3a), die dann wiederum Einfluss auf Variation nehmen kann. Sowohl Tokenfrequenz als auch die Effekte, die durch Cluster von ähnlichen Exemplaren bedingt sind und in Abschnitt 2.3 als Form-Schemata diskutiert werden, haben einen konservierenden Einfluss auf Flexion.

Der konservierende Effekt von Tokenfrequenz ist nicht auf die Morphologie beschränkt, sondern zeigt sich auch in der Syntax: Während Modalverben aufgrund ihrer Frequenz in Fragesätzen weiterhin Inversion aufzeigen (*May I ask you a question?*), ist dies bei Vollverben nicht der Fall (**Read you a book?*) (Bybee 2002a: 619–621). Bybee (2006: 728–729) stellt zudem fest, dass die *no*-Negation (*It looks like nothing to me*), die weniger produktiv ist als die *not*-Negation (*It doesn't look like anything to me*), sich in frequenten Konstruktionen weiterhin hält, wie bspw. in Existenzaussagen mit *be* (*There was no work to do* statt *There wasn't work to do*).

Für die Abgrenzung von Autonomie- und Konservierungseffekten stellt sich die Frage, ab welcher Frequenz welcher Effekt greift. Der Autonomie-Effekt ist auf extrem frequente Elemente begrenzt und greift somit für frequentere Elemente als der Konservierungseffekt (Bybee 2006: 715).¹² Die Verwendungen extrem frequenter Elemente bilden ein größeres Cluster als die Verwendungen weniger frequenter Elemente. Das Cluster hat somit auch eine größere Variabilität, die die Autonomie fördert.

Neben der Autonomie und der Konservierung beobachten Bybee & Thompson (1997: 378–384) Reduktion als Effekt hoher Tokenfrequenz. Der reduzierende Effekt ist bspw. in der Phonologie sichtbar: Häufige Kombinationen werden kontrahiert (*no thing > nothing*, *does not > doesn't*; *zu dem > zum*). Zudem sind Lautwandelphänomene bei hochfrequenten Wörtern früher zu beobachten als bei infrequenten (Fenk-Oczlon 1990: 46–47): So werden /t/ und /d/ nach einem Konsonanten in hochfrequenten englischen Wörtern (*and*, *went*) eher getilgt als in niedrigfrequenten (Bybee 2000: 69–71). Am deutlichsten ist der reduzierende Einfluss der Tokenfrequenz in der Grammatikalisierung zu beobachten: Elemente, die aufgrund von Grammatikalisierung an Frequenz gewinnen (Detges

¹²Bybee (2006) unterscheidet zwar zwischen hochfrequenten und extrem frequenten Elementen, geht aber nicht darauf ein, wie sie diese unterscheidet: „I refer to high and low frequency and to extreme high frequency without specifying exactly what these values mean in numerical terms“ (Bybee 2006: 715).

& Waltereit 2002: 188), verlieren an phonologischer Substanz (*going to* > *gonna*) (Bybee & Thompson 1997: 378–380). Die phonologische Reduktion hat einen weiteren Effekt: Sie führt zu einer geringeren akustischen Salienz dieser Elemente und damit vor allem im L2-Spracherwerb zu Schwierigkeiten in der Lernbarkeit, sodass lexikalische Elemente leichter erworben werden als grammatische (Ellis 2008: 236).¹³

Auch Zipf (1968 [1935]: 20–29) beobachtet, dass Kürze und Frequenz zusammenhängen: Häufig vorkommende Wörter sind sprachübergreifend kürzer als seltene. Wie Bybee & Thompson (1997) sieht Zipf (1968 [1935]: 30) die Kürze frequenter Einheiten als einen Effekt ihrer Frequenz: Da die Einheiten so oft genutzt werden, ist Kürze ein Vorteil, weil Zeit und Aufwand gespart wird (Fenk-Oczlon 1990: 39). Den Kürzungseffekt frequenter Einheiten kann man dabei auch im Kleinen beobachten: Bei der zweiten Verwendung eines Worts in einem Kontext wird dieses kürzer gesprochen als in der ersten Verwendung (Fowler & Housum 1987). Bereits genutzte Wörter können reduziert werden, da sie durch das erstmalige Nutzen bereits aktiviert sind und durch den Kontext gestützt werden (Fowler & Housum 1987: 489; dieser Effekt wird als Rezenzeffekt in Abschnitt 2.1.2 noch einmal aufgegriffen). Hörer_innen sind dabei sensitiv für die Längenunterschiede zwischen erster und zweiter Verwendung (Fowler & Housum 1987: 501). Wiederholungen schneller zu sprechen, ist also nicht nur eine Produktionsstrategie, sondern markiert auch Wiederholungen als solche (Bybee & Hopper 2001: 11–12).

Reduktion lässt sich nicht nur auf phonologischer Ebene, sondern auch in der Syntax beobachten. Elemente, die häufig zusammen verwendet werden, wie bspw. *going to*, werden nicht mehr als Bestandteile einer komplexen Phrase aufgefasst, sondern als ein einzelnes Element prozessiert und inzwischen auch produziert (*gonna*) (Blumenthal-Dramé 2012: 68–69): „Items that are used together fuse together“ (Bybee 2002b: 112). Somit geht die interne syntaktische Struktur von hochfrequenten Phrasen verloren. In diesem Zusammenhang lässt sich die phonologische Reduktion von *going to* zu *gonna* auch als Reaktion auf die Prozessierung hochfrequenter Phrasen als ein einzelnes Element verstehen. Zwischen den Einheiten ist eine „ever-increasing cognitive connection“ (Schneider 2014: 13) entstanden. Bei *going to* und *gonna* kann man zwar noch eine Beziehung von *gonna* zu *going to* herstellen, die Reduktion kann aber soweit fortgeschritten sein, dass ohne historisches Wissen keine Beziehung mehr zwischen Fusion und ursprünglicher syntaktischer Struktur hergestellt werden kann, bspw. bei *oje*, das

¹³Ellis (2002b: 307) geht davon aus, dass grammatische, wenig saliente Elemente durch lexikalische, salientere Elemente blockiert werden. Als Beispiel nennt er *today* und die Endung *-s* (als Marker für die dritte Person Singular Präsens), beide weisen laut Ellis (2002b: 307) auf Präsens hin, aber *today* hat eine höhere Wahrscheinlichkeit, wahrgenommen zu werden.

2 Variation aus gebrauchsbasierter Perspektive

auf *Oh mein Jesus* zurückgeht (Nübling 2001). Dieser Prozess wird als *chunking*¹⁴ bezeichnet (Schneider 2014: 13): Häufig zusammen auftretende Einheiten werden als eine große, hierarchisch gegliederte Einheit verarbeitet. Grundlage für das *chunking* ist *entrenchment* (Schneider 2014: 13). *Entrenchment*¹⁵ wurde bereits in Bezug auf einzelne Formen eingeführt: Eine Einheit kann je nach Häufigkeit des Gebrauchs mehr oder weniger *entrenched* sein. Das gleiche Prinzip lässt sich auf mehrere Einheiten anwenden: Diese sind stärker miteinander assoziiert, wenn sie häufig gemeinsam auftreten.

Chunking hat einen kognitiven Vorteil: Da das Arbeitsgedächtnis nur eine bestimmte Anzahl von Elementen gleichzeitig verarbeiten kann, bietet es sich an, häufig zusammen auftretende Einheiten als eine Einheit zu verarbeiten, sodass mehr Informationen im Arbeitsgedächtnis genutzt werden können (Bybee & Edgington 2006: 328, Bybee 2013: 54–55, Diessel 2017: 15). *Chunking* erleichtert daher Prozessierung: „When you do not have to put together every utterance from scratch, you need a minimum of on-line processing“ (Blumenthal-Dramé 2012: 11). Dieser Vorteil ergibt sich erst durch Wiederholung, sodass Chunks inhärent mit Frequenz verbunden sind (Bybee 2006: 717). *Chunking* zeigt sich auch in der Wahrnehmung von Sprecher_innen: Kollokationen (wie bspw. *absolute silence*) werden als natürlicher wahrgenommen als semantisch gleichwertige Kombinationen, die aber nicht *entrenched* sind (*pure silence*) (Dąbrowska 2014).

Nicht nur *chunking* und phonologische Reduktion, sondern auch die semantische Verblässung, die im Zuge der Grammatikalisierung einsetzt, lässt sich als Frequenzeffekt modellieren (Bybee & Thompson 1997: 379–380): Hochfrequente Wörter können semantisch verblässen (Detges & Waltereit 2002: 174) und aufgrund der Verblässung noch häufiger genutzt werden, sodass ein Kreislauf einsetzt. Ein Beispiel für semantische Verblässung durch Frequenz ist die Entstehung neuer Negationsmarker (Bybee & Thompson 1997: 380): Diese sind zunächst stark expressiv, durch häufige Verwendung schleift sich die Expressivität ab, sodass nur noch die Negation als semantisches Merkmal übrig bleibt. Auch hier zeigt sich, dass Salienz und Frequenz miteinander interagieren. Durch die frequenzbedingte semantische sowie phonologische Reduktion des Negationsmarkers verringert sich dessen kognitive und physiologisch bedingte Salienz (ausführlich zu den Salienzbegriffen siehe Auer 2014): Der Wegfall der Expressivität

¹⁴Diessel (2017: 14–15) verwendet neben *chunking* auch den Terminus *Automatisierung*, ohne einen Bedeutungsunterschied zu machen. Schneider (2014: 13) differenziert zwischen *Automatisierung* und *chunking*: *Automatisierung* setzt für sie erst ein, wenn eine phonologische Reduktion der Chunks zu beobachten ist.

¹⁵Der Terminus *entrenchment* geht auf Langacker (1987: 59–60) zurück. Die Idee von *chunking* aufgrund von Frequenz ist jedoch weit älter: Blumenthal-Dramé (2012: 11) merkt an, dass bereits Saussure frequenzbasierte Chunks diskutiert.

verringert das Potential, sich von anderen Ausdrucksformen zu unterscheiden (kognitive Salienz), und die phonologische Reduktion verringert das Potential, sensorisch zugänglich zu sein (physiologisch bedingte Salienz). Die verringerte Salienz des Negationsmarkers führt dazu, dass der Negationsmarker durch expressive Mittel verstärkt wird, welche wiederum selbst durch häufige Nutzung ihre Expressivität und damit Salienz verlieren (Jespersen 1917, zur Interaktion von Frequenz und Expressivität siehe Detges & Waltereit 2002: 176–186).

Der frequenzbedingte Reduktionseffekt verdeutlicht die fundamentale Rolle, die Frequenz für die Grammatikalisierung spielt. Aufgrund dieser Rolle wird Grammatikalisierung auch als Ritualisierung oder Routinisierung bezeichnet (Bybee 2002a: 603–604, Detges & Waltereit 2002: 189). Ritualisierung ist hier als Konsequenz von Wiederholung gemeint, die zu Gewohnheit, Automatisierung, Reduktion und Emanzipation der Struktur von der ursprünglichen führt (Haiman 1994: 4–8).

Es zeigt sich, dass sowohl die Typen- als auch die Tokenfrequenz einen fundamentalen Einfluss auf mentale Repräsentationen haben. Die Typenfrequenz von Kategorien ist eng mit deren Produktivität verknüpft. Zusätzlich zur Typenfrequenz beeinflussen Variabilität und Ähnlichkeit die Produktivität. Variabilität und Ähnlichkeit sind dabei eng mit Typenfrequenz verzahnt. Die Tokenfrequenz nimmt kaum Einfluss auf die Produktivität einer Kategorie. Hohe Tokenfrequenz kann jedoch zu Autonomie, Konservierung und Reduktion aufgrund von *entrenchment* führen. Für die weitere Arbeit sind vor allem die Generalisierbarkeit von Flexionsverhalten aufgrund von hoher Typenfrequenz und der festigende und damit konservierende Einfluss von hoher Tokenfrequenz relevant. Der Einfluss von Typen- und Tokenfrequenz auf Variation wird anhand der drei Variationsphänomene in Kapitel 3 näher untersucht. Dabei wird vorrangig die Generalisierbarkeit aufgrund von Typenfrequenz sowie *entrenchment* durch hohe Tokenfrequenz relevant sein. Im empirischen Teil der Arbeit wird der konservierende Einfluss von Tokenfrequenz psycholinguistisch überprüft.

Der nächste Abschnitt gibt einen vertiefenden Einblick in Frequenzeffekte. Dabei werden Frequenzeffekte in der Prozessierung und die Interaktion von Frequenz mit Salienz und Rezenz diskutiert. Zusätzlich wird auch die Korrelation von Frequenz mit anderen Faktoren wie der Kontextdiversität und Abfolgewahrscheinlichkeit betrachtet. In diesem Zusammenhang wird die Frage diskutiert, welche Probleme entstehen, wenn Frequenz als Erklärung für Prozessierungsunterschiede herangezogen wird.

2.1.2 Frequenzeffekte und Korrelation mit anderen Faktoren

Frequenz wird als einer der besten Prädiktoren angesehen, um das Verhalten von Menschen vorherzusagen (Divjak & Caldwell-Harris 2015: 55). Da frequente Wörter mental gefestigt sind, können sie schneller und leichter prozessiert als infrequente: Bei *lexical decision tasks* erkennen Proband_innen schneller, dass es sich um ein real existierendes Wort handelt, wenn dieses frequent ist (Rubenstein u. a. 1970: 489–491, Whaley 1978: 150–152).¹⁶ In Eyetrackingstudien lässt sich beobachten, dass frequente Wörter kürzer fokussiert werden als infrequente (Rayner & Duffy 1986). Divjak & Caldwell-Harris (2015: 56) stellen zudem fest, dass die Prozessierung frequenter Wörter weniger störungsanfällig ist als die Prozessierung infrequenter Wörter: Hochfrequente Wörter werden bei Störgeräuschen besser verstanden als infrequente (Pollack u. a. 1959). Der Einfluss der Frequenz zeigt sich auch mittelbar im Design psycholinguistischer Studien: Um einen konfundierenden Frequenzeffekt auszuschließen, wird auf Frequenz kontrolliert (Divjak & Caldwell-Harris 2015: 54).

Die Sensitivität von Menschen für Frequenz zeigt sich auch in Einschätzungsstudien. Proband_innen sind bspw. gut darin, die Frequenz einzuschätzen, mit der Silben und Wörter genutzt werden: Die Einschätzung korreliert jeweils mit der tatsächlichen Frequenz (Shapiro 1969, Rubin 1974). Außerdem können Proband_innen gut einschätzen, wie häufig ihnen in einem Experiment ein bestimmter Stimulus präsentiert wurde (Hasher & Zacks 1984). Flexser & Bower (1975: 323) präsentieren eine Liste mit 117 Pseudowörtern, die in ihrer Tokenfrequenz von eins bis sechs variieren. Die Proband_innen werden im Anschluss gebeten, zu schätzen, wie häufig die einzelnen Pseudowörter in der Liste vorkamen. Dabei korreliert die Schätzung der Proband_innen mit der tatsächlichen Frequenz der Pseudowörter, unabhängig davon, ob die Proband_innen vorab wussten, dass sie am Ende Frequenzschätzungen abgeben sollen (Flexser & Bower 1975: 323–324).

Neben der Sensitivität für die Frequenz von Wörtern und Silben können Frisch u. a. (2001) eine Sensitivität für die Häufigkeit von Lautkombinationen nachweisen: Sie arbeiten mit ausgedachten Wörtern, die häufige und seltene Lautkombinationen einer Sprache enthalten. Lässt man Proband_innen beurteilen, ob die Testwörter theoretisch mögliche Wörter in einer Sprache sind, erzielen die Wörter hohe Werte, die häufige Lautkombinationen einer Sprache enthalten (Frisch

¹⁶Rubenstein u. a. (1970: 488–489) kontrollieren auf Wortlänge, Whaley (1978: 144–148) tut dies nicht, inkludiert aber die Wortlänge in die Analyse. So kann Whaley (1978: 148–152) zeigen, dass sowohl Frequenz als auch Wortlänge Einfluss auf die Reaktionszeiten nehmen. Zudem nutzt Whaley (1978: 146) verschiedene Frequenzmaße, sodass nicht nur die reine Wortwiederholung, sondern auch Kontextdiversität berücksichtigen werden kann.

u. a. 2001: 162–163). Sie stellen dabei interessante individuelle Unterschiede in den Bewertungen fest: Je größer der Wortschatz von Proband_innen, desto eher schätzen Proband_innen Testwörter mit seltenen Lautkombinationen als möglich ein (Frisch u. a. 2001: 166). Die Wahrscheinlichkeit für die Existenz eines Wortes wird also anhand der Häufigkeit von mental repräsentierten Lautkombinationen beurteilt, daher ist sie vom individuellen Wortschatz abhängig (Frisch u. a. 2001: 175). Auf der Idee, dass Sprecher_innen aktuell zu verarbeitenden Input mit Wahrscheinlichkeiten vergleichen, die auf ihren vorherigen Spracherfahrungen basieren, baut der bayesianische Ansatz für Sprachverarbeitungs- und Lernmodelle auf, der im folgenden Abschnitt ausführlich vorgestellt wird.

Sowohl die schnelle Prozessierung als auch die Frequenzeinschätzung von Sprecher_innen verdeutlichen, dass frequente Wörter mental stark gefestigt (*entrenched*) sind. Das *entrenchment* frequenter Wörter zeigt sich aber auch darin, dass sie eine höhere Resilienz haben als infrequente Wörter: Frequente semantische Konzepte, die auf dem Basislevel operieren (bspw. *Hund*, *Pferd*; siehe hierzu genauer Abschnitt 2.2), werden von Patienten mit semantischer Demenz bewahrt und sogar übergeneralisiert (bspw. wird auf alle größeren Tiere mit *Pferd* referiert), wohingegen für weniger frequente Konzepte nur noch ein rudimentäres Konzept zu bestehen scheint (Rogers & McClelland 2003: 12). Diese Beobachtung lässt sich gut mit Goldinger (1998: 255) vereinbaren, der anhand einer Computersimulation zeigt, dass frequente Wörter abstrakter repräsentiert sind als infrequente Wörter. Er testet dies anhand des Vorteils für gleichbleibende Stimmen (*same voice advantage*)¹⁷: Für wenig frequente Wörter lässt sich ein Vorteil für gleichbleibende Stimmen beobachten. Wenn man ein Wort hört, das mehrfach von derselben Person gesagt wird, kann man dieses besser erkennen, als wenn das Wort von verschiedenen Personen gesagt wird. Für frequente Wörter besteht dieser Effekt nicht (Ellis 2002a: 153). Hört man ein neues Wort, koppelt man dieses also zunächst an die Stimme der Person, die es produziert hat. Dieses spezifische Detail verschwindet jedoch, je häufiger man ein Wort hört (Goldinger 1998: 254–255). Der Effekt stützt ein exemplarbasiertes Modell der Kategorienbildung, nach dem Kategorien anhand konkreter Exemplare abstrahiert werden (Ellis 2002a: 153). Diese Idee wurde bereits durch Goldbergs Cluster-Modell in Abschnitt 2.1.1 eingeführt und wird in Abschnitt 2.2.3 zu Kategorienbildung durch exemplarbasiertes Lernen vertieft.

Für exemplarbasiertes Lernen sprechen auch Prozessierungseffekte, die sich für Wörter beobachten lassen, die ähnlich aufgebaut, aber unterschiedlich fre-

¹⁷Nicht nur dieselbe Stimme, sondern auch derselbe Raum bringt Vorteile: Proband_innen fällt es leichter, Wörter von einer Liste wiederzugeben, wenn sie sich im selben Raum befinden, indem sie die Liste zuvor gelesen haben (Goldberg 2019: 14).

2 Variation aus gebrauchsbasierter Perspektive

quent sind: Reaktionszeiten sind für infrequente Wörter bei *lexical decision tasks* erhöht, wenn zuvor ein hochfrequenter orthographischer Nachbar präsentiert wurde, z. B. *blue* vor *blur* (Grainger u. a. 1989, Grainger 1990). Derselbe Effekt lässt sich auch durch Eyetracking beobachten (Grainger u. a. 1989). Dies verdeutlicht einerseits, dass ähnliche Exemplare sich in ihrer Prozessierung beeinflussen und andererseits, dass Wörter nicht entweder aktiviert werden oder nicht, sondern dass Aktivierung graduell ist, da sie von der Frequenz des Wortes sowie der Anzahl ähnlicher Wörter und deren Frequenz abhängt (Ellis 2002a: 150–152).

Der Einfluss orthographischer Ähnlichkeit lässt sich auch in der Flexion beobachten. Englische regelmäßige Verben werden schneller gebildet, wenn keine unregelmäßigen Verben existieren, die orthographisch ähnlich sind: *walk*, *walked* wird also schneller gebildet als *bake*, *baked*, weil *bake* Ähnlichkeit zu den unregelmäßigen Verben *make*, *made* und *take*, *took* aufweist (Seidenberg 1992: 102–104; siehe Daugherty & Seidenberg 1994: 371–387 für eine Studie, in der derselbe Effekt in einer Datensimulation beobachtet wird). Die Unterschiede in der Schnelligkeit der Produktion lassen auf Unterschiede in der Prozessierung schließen: Die Fähigkeit, ein Verb einer Klasse zuzuordnen, ist eingeschränkt aufgrund der Ähnlichkeit zu Verben, die einer anderen Klasse angehören.

Hieran zeigt sich der Einfluss von *cue validity* (Signalgüte) auf Prozessierung: Wenn ähnliche Strukturen miteinander konkurrieren, weisen diese eine geringe *cue validity* auf, weshalb die Strukturen schwer zu differenzieren und damit zu erlernen sind (Ellis 2002a: 153, siehe Abschnitt 2.3 für weitere Ausführungen zu *cue validity*). Die herabgesenkte *cue validity* von regelmäßigen Verben, die unregelmäßigen Verben ähneln, weist zurück auf das Prinzip der Abdeckung nach Goldberg (2019) und voraus auf Form-Schemata (siehe Abschnitt 2.3): Die regelmäßigen Verben ähneln zwar den unregelmäßigen Verben in der Form, gehören aber dennoch einer anderen Flexionsklasse an. Dies kann vereinzelt zu Variation führen, indem die Flexionsklasse der Form angepasst und bspw. statt der regelmäßigen Form *baked* die Form *book* in Analogie zu *took* gebildet wird.

Wie der letzte Abschnitt bereits verdeutlicht, treten Frequenzeffekte nicht isoliert auf, sondern interagieren mit anderen Faktoren. Dies zeigt auch die in Abschnitt 2.1.1 diskutierte Interaktion von Frequenz und Salienz. Pfänder & Behrens (2016: 6–8) weisen darauf hin, dass Frequenz auch mit Rezenz interagiert: Je kürzer der Zeitabschnitt zwischen der Wiederholung ein und desselben Elements ist, desto höher ist die Aktivierung des Elements.¹⁸ Die in Abschnitt 2.1.1

¹⁸In der Psychologie ist der Rezenzeffekt als Priming bekannt (Ellis 2012: 14–15). Um Rezenzeffekte zu messen, wird m. W. allein die Zeit zwischen der Wiederholung eines Elements gemessen und nicht berücksichtigt, wie komplex die Struktur zwischen der Erstnennung und der Wiederholung ist.

erwähnte schnellere Produktion des zweiten Vorkommens eines Elements in einem bestimmten Kontext (Fowler & Housum 1987) lässt sich auch als Rezenzeffekt beschreiben. Pfänder & Behrens (2016: 7) gehen davon aus, dass der Effekt sich direkt aus der Struktur unseres Erinnerungsvermögens ableitet. Für den Effekt spielt neben der Länge des Zeitabschnitts die Häufigkeit des sich wiederholenden Elements in einem bestimmten Zeitabschnitt eine Rolle: Je häufiger ein Element in einem Zeitabschnitt genutzt wird, desto stärker wird es aktiviert (Pfänder & Behrens 2016: 6–8). Außerdem wird Rezenz von der generellen Frequenz des Elements beeinflusst: Der Rezenzeffekt ist für Elemente mit geringer Tokenfrequenz höher, da das Auftreten von Elementen mit geringer Tokenfrequenz überraschender und somit salienter ist als das Auftreten von Elementen mit hoher Tokenfrequenz (Pfänder & Behrens 2016: 6–8). An dieser Stelle zeigt sich, dass auch Salienz mit Rezenz verwoben ist.

Zusätzlich zur Interaktion von Frequenz mit anderen Faktoren ist zu beachten, dass der Einfluss der Frequenz auf Prozessierung mit anderen Eigenschaften korreliert, die mit einem Wort assoziiert sind. Dies sind bspw. das Erwerbssalter von Wörtern sowie deren Vertrautheit¹⁹: Früh erlernte und vertraute Wörter sind in der Regel frequenter als spät erlernte und wenig vertraute Wörter (Gernsbacher 1984, McDonald & Shillcock 2001: 298–299).²⁰ Neben diesen Worteigenschaften korrelieren einige Kontexteigenschaften mit Frequenz, nämlich die Größe der syntaktischen und morphologischen Familie (*syntactic* und *morphological family size*) sowie die Kontextdiversität bzw. Dispersion²¹ (Baayen 2012: 189–191). Die Häufigkeit eines Worts ist zwar der beste einzelne Prädiktor für das

¹⁹Vertrautheit wird ermittelt, indem Proband_innen Wörter nach deren Vertrautheit mit dem Wort bewerten. Es ist somit ein intersubjektives Maß (Gernsbacher 1984: 261, McDonald & Shillcock 2001: 298).

²⁰Morrison & Ellis (1995) und Morrison & Ellis (2000) untersuchen, inwiefern das Erwerbssalter unabhängig von der Frequenz Einfluss auf die Reaktionszeiten in *word naming* und *lexical decision tasks* nimmt. Sie kontrollieren u. a. auf Frequenz und messen dennoch einen Effekt für das Erwerbssalter (Morrison & Ellis 1995: 120–121, 2000: 176). Umgekehrt ist bei kontrolliertem Erwerbssalter kein Frequenzeffekt zu beobachten (Morrison & Ellis 1995: 120–121). Zevin & Seidenberg (2002: 4–10) weisen darauf hin, dass letzteres Ergebnis in Folgestudien nicht repliziert wurde: Auch wenn auf Erwerbssalter kontrolliert wurde, waren Frequenzeffekte zu messen. McDonald & Shillcock (2001: 298) stellen die Notwendigkeit, Frequenz und Erwerbssalter zu trennen, in Frage und weisen auf das Konzept der kumulativen Frequenz von Lewis u. a. (2001) hin, das Erwerbssalter und Frequenz gleichermaßen berücksichtigt. Kumulative Frequenz misst, wie oft ein Wort insgesamt gelesen oder gehört wurde: Ein häufiges Wort lesen und hören Sprecher_innen öfter als seltene Wörter; genauso wurden Wörter mit frühem Erwerbssalter öfter gehört oder gelesen als Wörter mit spätem Erwerbssalter (Lewis u. a. 2001: 191).

²¹Die *syntactic family size* ist die Menge an verschiedenen Wörtern, die auf ein bestimmtes Wort folgen (Baayen 2012: 176). Um die *morphological family size* eines Worts zu bestimmen, wird die Anzahl an verschiedenen Wörtern ermittelt, in denen das Wort als Konstituente genutzt wird,

2 Variation aus gebrauchsbasierter Perspektive

Verhalten in *lexical decision tasks* (Baayen 2012: 178–179), anhand einer Principal-Components-Analyse kann Baayen (2012: 181–191) jedoch zeigen, dass die pure Wortwiederholung nur einen kleinen Anteil der beobachteten Varianz in den Daten erklären kann. Die meiste Varianz in den Daten wurde von einer Komponente erklärt, die vornehmlich Maße für Kontexteigenschaften enthält (wie bspw. Größe der jeweiligen syntaktischen oder morphologischen Familie sowie Dispersion) (Baayen 2012: 191).

Auch McDonald & Shillcock (2001: 308) diskutieren das Verhältnis von Wortfrequenz und Kontext, wobei sie Kontext als kontextuelle Distinktivität konzeptualisieren. Sie stellen fest, dass Wortfrequenz und Kontext negativ korrelieren: Je frequenter ein Wort ist, desto weniger distinkt sind die Kontexte, in denen es genutzt wird. Die Korrelation ist aber nicht perfekt: Es können auch frequente Wörter nur in spezifischen Kontexten genutzt werden, bspw. Zahlwörter. Diese werden meist zusammen mit anderen Zahlwörtern oder Maßeinheiten genutzt (McDonald & Shillcock 2001: 301). In einer *lexical decision task* können McDonald & Shillcock (2001: 310–312) zeigen, dass die kontextuelle Distinktivität Einfluss auf das Antwortverhalten nimmt: Je höher diese ist (in je weniger Kontexten ein Wort genutzt wird), desto höher sind die Reaktionszeiten. Sie stellen zudem fest, dass Frequenzeffekte verschwinden, wenn auf Kontextdistinktivität kontrolliert wird. McDonald & Shillcock (2001: 318–319) sehen in der kontextuellen Distinktivität eine bessere Erklärung für behaviorale Daten als Frequenz, da für die kontextuelle Distinktivität die Umgebung eines Worts explizit einbezogen wird und dieses Maß somit einen realistischeren Zugang zum Sprachgebrauch bietet.

Neben der Kontextdiversität und -distinktivität nimmt auch die Abfolgewahrscheinlichkeit (*transitional probability*) einzelner Elemente Einfluss auf die Prozessierung: Wie Saffran u. a. (1996) zeigen, sind acht Monate alte Kinder bereits im Stande, Abfolgewahrscheinlichkeiten von Silben zu bestimmen. Diese können sie nutzen, um Wortgrenzen festzustellen: Die Wahrscheinlichkeit für die Lautabfolge *klei-ne* ist größer als die Wahrscheinlichkeit für die Lautabfolge *ne-maus*, weswegen Kinder erkennen können, dass in der Äußerung *kleine Maus* sowohl *kleine* als auch *Maus* Wörter sind, aber nicht **nemaus*. McDonald & Shillcock (2003) zeigen zudem, dass Fixationszeiten beim Lesen von Abfolgewahrscheinlichkeiten beeinflusst werden. Es ist somit davon auszugehen, dass Sprecher_in

z. B. *Haus* in *Haustür*, *Mietshaus*, *Hausbau* (Baayen 2012: 177). Die Kontextdiversität bzw. Dispersion gibt an, ob Wörter durchweg in verschiedenen Korpora genutzt werden (Baayen 2012: 175). Durch den Blick auf Kontexte kann somit auch die Registerspezifität (bspw. unabhängig von Registern oder auf formelle Register beschränkt) von Wörtern berücksichtigt werden. Für einen Überblick zu verschiedenen Operationalisierungen von Kontextdiversität siehe Divjak & Caldwell-Harris (2015: 58–60).

nen beim Lesen auf Abfolgewahrscheinlichkeiten zurückgreifen. Abfolgewahrscheinlichkeiten sind dabei inhärent mit Frequenz verbunden, da häufigere Abfolgekombinationen auch eine höhere Wahrscheinlichkeit haben, aufzutreten. Zudem bieten sie eine Möglichkeit, den Einfluss des Kontexts auf Prozessierung zu operationalisieren.

Die hier diskutierten Studien verdeutlichen, dass eine Operationalisierung der Frequenz als pure Wortwiederholung problematisch sein kann, da sie den Blick auf weitere Effekte wie bspw. den Ko- und Kontext versperrt (Divjak & Caldwell-Harris 2015: 58–60). Den grundlegenden Einfluss des Kontexts auf mentale Repräsentation betont auch Goldberg (2019: 17), die davon ausgeht, dass Tokenfrequenz und die Kontexte, in denen ein Wort vorkommen kann, dessen mentale Repräsentation beeinflussen.

Frequenzeffekte lassen sich jedoch auch isolieren: Gardner u. a. (1987) untersuchen in einer *lexical decision task* das Antwortverhalten von Menschen aus verschiedenen sozialen Gruppen (Jurastudent_innen, Krankenpfleger_innen und Ingenieur_innen). In die *lexical decision task* sind jeweils Wörter aus Fachtexten der unterschiedlichen Disziplinen eingeflochten (ausführlich zum methodischen Vorgehen siehe Gardner u. a. 1987: 25–26). Gardner u. a. (1987: 26–28) stellen dabei fest, dass die Proband_innen niedrigere Reaktionszeiten für Wörter aus ihrer Disziplin aufwiesen als für Wörter aus anderen Disziplinen. In diesem Zusammenhang zeigt sich wie bei der Studie zur phonotaktischen Struktur von Wörtern, dass die individuelle Erfahrung von Sprecher_innen einen Einfluss auf die mentale Repräsentation von Sprache nimmt, denn sie beeinflusst die Frequenz, mit der Sprecher_innen bestimmte Konstruktionen wahrnehmen und somit auch, wie stark diese mental gefestigt sind. Dies führt zu dem Schluss, dass individuelle Unterschiede zwischen den Grammatiken existieren, die Sprecher_innen im Spracherwerb konstruieren. Kemmer & Barlow (2000: xviii) sehen darin die Basis für soziolinguistische Variation: Da Sprecher_innen, die viel miteinander agieren, ähnlichere Sprachgebrauchsmuster aufweisen als Sprecher_innen, die selten interagieren, entsteht ein gruppenspezifischer Sprachgebrauch und daher auch gruppenspezifische mentale Repräsentationen.

Es zeigt sich, dass Frequenz ein guter Prädiktor für menschliches Verhalten ist, jedoch aufgrund der Korrelation mit anderen Faktoren (z. B. Kontextdiversität) nicht immer als Erklärung für das Verhalten genutzt werden kann. Dieser Aspekt muss stets berücksichtigt werden, wenn der Einfluss von Frequenz auf Variation betrachtet wird. Um den Einfluss der Frequenz sauber fassen zu können, modelliert der folgende Abschnitt Frequenz als einen grundlegenden Einflussfaktor auf statistisches Lernen: Sie nimmt Einfluss auf die Wahrscheinlichkeit, mit der Menschen einem bestimmten Stimulus begegnen.

2.1.3 Frequenz und statistisches Lernen

Insgesamt stützen die hier beschriebenen Frequenzeffekte ein konnektionistisches Modell (*connectionist model*) des Spracherwerbs. Dieses Modell geht davon aus, dass Sprache erworben wird, indem Unterschiede und Gemeinsamkeiten bspw. zwischen regelmäßigen und unregelmäßigen Verbklassen im Input durch Lernalgorithmen generalisiert werden:

On this [connectionist, E.S.] account, generally instantiated as a feedforward network employing a learning algorithm, differences in behavior between the two regularity classes develop through a single processor's experience with data that are heterogeneous with respect to item frequency, phonological similarity, and other factors. These networks learn inductively, and are driven to develop a set of weighted connections that associates each input-output pair. As a result, they abstract generalizations from the data. Their subsequent behavior is based on the productive application of such generalizations (Hare u. a. 2001: 183).

Das eigene Sprachverhalten ist somit das Resultat von Generalisierungen aus dem sprachlichen Input.²² Als Grundlage dient lediglich eine basale Netzwerkstruktur, die mit Beispielen gefüllt wird. Aus ihnen wird schließlich ein System abstrahiert, das im Idealfall die Muster aus dem Input generieren kann (Kemmer & Barlow 2000: xiii). Kemmer & Barlow (2000: xiii) betonen, dass das Modell somit wie das menschliche Gehirn funktioniert, da sowohl das Modell als auch das Gehirn keine zentrale Prozessierungseinheit haben. Im Gehirn ist jedes Neuron eine Prozessierungseinheit, die Verbindungen zwischen anderen Neuronen aktiviert oder hemmt. Informationen ergeben sich erst durch die Verbindung zwischen den Neuronen, genauso ergeben sich im Modell sprachliche Informationen erst durch Aktivierungsmuster (Kemmer & Barlow 2000: xiii).

Die Grundlage für den konnektionistischen Ansatz legen Rumelhart & McClelland (1986), die den Erwerb der regelmäßigen und unregelmäßigen Konjugation im Englischen simulieren. Der Erwerb der Konjugationsklassen verläuft bei Kindern in einer Kurve, die wie ein U geformt ist (*u-shaped curve*): Zunächst nutzen Kinder nur Präteritalformen von hochfrequenten und damit i. d. R. unregelmäßigen Verben (Rumelhart & McClelland 1986: 219–220). Im Anschluss an diese Phase erwerben Kinder die regelmäßige Konjugation. Sie konjugieren dabei nicht nur bekannte Wörter nach der regelmäßigen Konjugation, sondern wenden sie produktiv auf unbekannte Verben an. Zudem übergeneralisieren sie die

²²Siehe Elman (1992) für eine ausführliche Diskussion zum Nutzen des konnektionistischen Ansatzes für die Modellierung kognitiver Prozesse.

regelmäßige Konjugation (auch auf Formen, die sie zuvor unregelmäßig gebildet haben), sodass in dieser Phase kaum unregelmäßige Verbformen zu beobachten sind. Im dritten Stadium koexistieren schließlich beide Konjugationen (Rumelhart & McClelland 1986: 219–220). Diesen Verlauf können Rumelhart & McClelland (1986) gut mit ihrem Modell simulieren. Zunächst werden zehn hochfrequente, vornehmlich unregelmäßige Verben in das Netzwerk eingespeist. Aufgrund der großen Unterschiede zwischen den unregelmäßigen Verben wird die Beziehung zwischen Infinitiv- und Präteritalform im Netzwerk für jedes Verb separat betrachtet. In der zweiten Phase werden 420 mittelfrequente, vornehmlich regelmäßige Verben eingespeist. Aufgrund der Gemeinsamkeiten zwischen den regelmäßigen Verben kann das Netzwerk nun Verbindungen zwischen den Verben ziehen. Da zwischen den meisten eingespeisten Verben Verbindungen herstellbar sind, übergeneralisiert das Modell die regelmäßige Konjugation auf unregelmäßige Verben. Schließlich eliminiert das Modell die Übergeneralisierung und zeigt für die unregelmäßigen Verben auch unregelmäßige Formen (Rumelhart & McClelland 1986: 230–231, 240–246).²³ Das Modell verdeutlicht, dass Typenfrequenz und Regelmäßigkeit Hand in Hand gehen, denn erst die Typenfrequenz ermöglicht es, ein Muster zu abstrahieren (siehe Abschnitt 2.1.1).

Pinker & Prince (1991) gehen im Gegensatz zum konnektionistischem Ansatz davon aus, dass regelmäßige und unregelmäßige Verben unterschiedlich prozessiert werden. Diese Annahme fußt auf der Beobachtung, dass nur bei unregelmäßigen Verben die Frequenz der Vergangenheitsformen einen Einfluss auf die Prozessierung hat, nicht aber bei regelmäßigen Verben (Pinker & Prince 1991: 232–233). Pinker & Prince (1991: 233–234) schlagen deshalb vor, von zwei verschiedenen Mechanismen auszugehen (*dual mechanism model*): Für unregelmäßige Verbformen halten sie einen konnektionistischen Ansatz für angemessen, aber nicht für regelmäßige Verben. Für diese gehen sie vom klassischen Ansatz der regelbasierten Grammatik aus: Flektierte Formen werden regelbasiert gebildet und die durch die Regel generierten Formen nicht mental gefestigt, weshalb bei regelmäßigen Verben keine Frequenzeffekte für einzelne Verbformen zu beobachten sind.

²³Für eine ausführliche Erläuterung des Modells siehe Rumelhart & McClelland (1986: 220–240). In dem Modell wurden auch phonologische Gemeinsamkeiten zwischen unregelmäßigen Verben berücksichtigt, siehe hierfür Rumelhart & McClelland (1986: 245–266). Zu Kritik an dem Modell siehe Pinker & Prince (1988) und Pinker & Prince (1991: 232–233). Siehe Plunkett & Marchman (1991) sowie Daugherty & Seidenberg (1994) für weitere Ansätze, um regelmäßige und unregelmäßige Konjugation mithilfe eines konnektionistischen Modells zu simulieren. Bybee (1988: 136–138) diskutiert Unterschiede und Gemeinsamkeiten zwischen diesem Modell und ihrem Modell der lexikalischen Verbindungen und Stärke (siehe Abschnitt 2.1.1).

2 Variation aus gebrauchsbasierter Perspektive

Der (fragliche) Einfluss der Frequenz von Präteritalformen regelmäßiger Verben auf die Prozessierung wurde im Rahmen der sogenannten *past tense debate* diskutiert und überprüft. Für das Deutsche lassen sich – in Einklang mit den Vorhersagen des *dual mechanism*-Ansatzes – frequenzbedingte Unterschiede für starke Verbformen, aber nicht für schwache feststellen (Clahsen u. a. 1997, 2001, 2004, die Studien werden in Abschnitt 3.1.1 zu Variation in der Konjugation erneut aufgegriffen). Für das Englische stellen Alegre & Gordon (1999) Frequenzeffekte für regelmäßige Verben ab einer Frequenz von sechs pro Million Token fest und plädieren daher für eine schwache Auslegung des *dual-mechanism*-Modells, während Hare u. a. (2001) keinen Unterschied zwischen regelmäßigen und unregelmäßigen Verben messen.

Als Evidenz für das *dual mechanism model* werden zudem Ergebnisse von EEG- (Elektroenzephalografie), PET- (*positron-emission tomography*) und fMRI- (*functional magnetic resonance imaging*) Studien gesehen, die Unterschiede in der Verarbeitung von regelmäßigen und unregelmäßigen Formen nahelegen: Unregelmäßige Verbformen zeigen insgesamt eine größere Aktivierung als regelmäßige, zudem wurden Unterschiede in der Aktivierung der Hirnregionen festgestellt; wobei die Ergebnisse der Studien hierbei kein eindeutiges Bild ergeben (siehe Jaeger u. a. (1996); Indefrey u. a. (1997); Penke u. a. (1997); Beretta u. a. (2003); Krott & Lebib (2013) zu unregelmäßigen und regelmäßigen Verben im Deutschen und Englischen und Weyerts u. a. (1997); Bartke u. a. (2005) zu regelmäßigen und unregelmäßigen Pluralformen²⁴ sowie Bornkessel-Schlesewsky & Schlesewsky

²⁴Hierbei wird von *-s* als regelmäßiger Flexion ausgegangen, da Kinder diese Flexion übergeneralisieren, Erwachsene den Plural von Namen und neuer unbekannter Wörter mit *-s* bilden und *-s* anders als andere Pluralformen keine Frequenzeffekte hervorruft (Weyerts u. a. 1997: 958). Das Verhalten von Erwachsenen lässt sich aber auch als Wortkörperschonung für periphere Substantive sehen (Ackermann & Zimmer 2017: 150–154). Auch Köpcke & Wecker (2017: 79) sehen den *s*-Plural aufgrund seiner Verwendung in spezifischen Kontexten nicht als Default an. Zudem zeigen Köpcke u. a. (2021: 19–21) in einem Experiment, dass bei der Beurteilung der Pluralität von Pseudowörtern kein Unterschied zwischen Versuchstems auf *-n* und auf *-s* feststellbar sind. Sie sehen *-s* daher als einen Notplural an, da *-s* (bis auf den Ausschluss von Substantiven auf *-s*) fast keine phonetischen Anforderungen an das Substantiv stellt. Zudem stellt aufgrund der komplexen Pluralbildung im Deutschen die Opposition von regelmäßig (*-s*) und unregelmäßig (alle anderen Formen) eine starke Vereinfachung dar. Bartke u. a. (2005: 33–35) weisen darauf hin, dass sich auch für andere Pluralformen Regelmäßigkeiten beobachten lassen, wie etwa bei *-e*, das den vorherrschenden Pluralmarker bei Maskulina und Neutra darstellt. Die Ergebnisse ihrer EEG-Studie zeigen Unterschiede zwischen *-e* als vorherrschenden Pluralmarker für Maskulina und Neutra und *-n* als typeninfrequenten Marker: Das typenfrequente Suffix *-e* ruft bei Maskulina und Neutra eine weniger starke Aktivierung hervor als das typeninfrequente Suffix *-n* (Bartke u. a. 2005: 40–47). Dies entspricht den Studien zu Konjugation, in denen für unregelmäßige Verben eine stärkere Aktivierung festgestellt wurde als für regelmäßige.

(2009: 47–68) für einen Überblick zu Prozessierungsunterschieden zwischen regelmäßiger und unregelmäßiger Flexion). Zudem unterscheiden sich die Flexionsklassen in Bezug auf Priming-Effekte: Während sich für die regelmäßige Klasse Priming-Effekte nachweisen lassen (*tanzen* kann *getanzt* primen), ist dies bei unregelmäßigen Paradigmen nicht der Fall (Bornkessel-Schlesewsky & Schlesewsky 2009: 47–68).

M. E. lassen die hier vorgestellten Unterschiede in der Verarbeitung nicht zwingend auf das *dual mechanism model* schließen. So schließt der konnektionistische Ansatz Unterschiede zwischen regelmäßiger und unregelmäßiger Flexion nicht aus (siehe die oben zitierte Definition von Hare u. a. 2001), sondern geht davon aus, dass sich diese „through a single processor’s experience“ (Hare u. a. 2001: 183) entwickeln. Unterschiede in der Prozessierung können also auch Ergebnis desselben Mechanismus sein. So lassen sich bspw. die ausbleibenden Priming-Effekte durch Cluster erklären: Diese hängen bei unregelmäßigen Verben aufgrund der starken phonologischen Veränderungen weniger stark zusammen, weswegen die Formen auch weniger stark geprimt werden können (siehe Abschnitt 2.1.1). Hierzu passt, dass Smolka u. a. (2013) graduelle Unterschiede im Priming von regelmäßigen und unregelmäßigen Verben feststellen. Smolka u. a. (2013: 6–9) unterscheiden zwischen regelmäßigen Verben mit Dentalsuffix (*lachen* - *gelacht*), unregelmäßigen Verben mit Ablaut, der dem Infinitivstammvokal gleicht (*fahren* - *gefahren*), sowie unregelmäßigen Verben mit Ablaut, der sich vom Infinitivstammvokal unterscheidet (*schwimmen* - *geschwommen*).²⁵ Für die Verben stellen Smolka u. a. (2013: 11–13) unterschiedliche Aktivierungsmuster fest: Während die regelmäßigen Verben eine Aktivierung in Frontal-, Temporal- und Parietallappen zeigten, war die Aktivierung bei den Verben mit *n*-Suffix und Vokalidentität zwischen Ablaut und Infinitivstammvokal geschwächt. Bei *n*-Suffix und anderem Ablaut als dem Infinitivstammvokal war keine Aktivierung zu erkennen. Dies lässt darauf schließen, dass die Vokalveränderung Priming-Effekte blockiert.

Auch die ausbleibenden Frequenzeffekte bei den Präteritalformen regelmäßiger Verben lassen sich ohne Rückgriff auf das *dual mechanism model* erklären. Ellis & Schmidt (1998) gehen davon aus, dass das Ausbleiben auf das „power law of practice“ zurückgeht (Ellis & Schmidt 1998: 307): Je häufiger eine mentale Aufgabe erledigt werden muss, desto schneller wird diese auch erledigt. Die

²⁵Smolka u. a. (2013: 6–9) sehen Verben mit Identität zwischen Infinitivstammvokal und Ablaut als weniger unregelmäßig an als Verben, deren Ablaut sich vom Infinitivstammvokal unterscheidet. Aus typenfrequentieller Sicht sind jedoch die Verben mit gleichem Ablaut wie Infinitivstammvokal (Ablautalternanz ABA) unregelmäßiger als die Verben mit ungleichem Ablaut und Infinitivstammvokal (Ablautalternanzen ABB und ABC), da mehr Verben die Ablautalternanzen ABB oder ABC aufweisen, siehe Nowak (2018: 159–161) sowie Abschnitt 3.1.1 zu Frequenzeffekten bei der Variation in der Konjugation.

meisten Verben werden regelmäßig konjugiert,²⁶ deswegen kann die Vergangenheit von regelmäßigen Verben schnell gebildet werden, unabhängig davon, wie häufig die Vergangenheitsform eines speziellen regelmäßigen Verbs ist. Ellis & Schmidt (1998: 307) setzen in diesem Zusammenhang Regelmäßigkeit mit Typenfrequenz gleich: „Regularity is just [type, E.S.] frequency with another name“ (Ellis & Schmidt 1998: 307). Die Unterschiede zwischen regelmäßigen und unregelmäßigen Verben lassen sich nach Ellis & Schmidt (1998: 309) somit durchaus mit dem konnektionistischen Modell vereinen. Um den Einfluss des *power law of practice* zu testen, lassen Ellis & Schmidt (1998: 311–316) Proband_innen regelmäßige und unregelmäßige Pluralformen in einer artifiziellen Sprache bilden. Am Anfang des Spracherwerbs zeigen sich Frequenzeffekte für regelmäßige und unregelmäßige Formen: Frequente Formen werden in beiden Flexionsklassen akkurater und schneller gebildet. Mit der Zeit verschwindet der Frequenzeffekt jedoch für die regelmäßigen Formen (Ellis & Schmidt 1998: 313–315).

Insgesamt lassen sich die empirischen Befunde, die als Beleg für das *dual mechanism model* ausgelegt wurden, also durchaus mit dem konnektionistischen Ansatz verbinden. Gegen die Annahme zweier separater Mechanismen sprechen zudem die in Abschnitt 2.1.2 vorgestellten Ergebnisse von Seidenberg (1992), nach denen die Produktion regelmäßiger Verbformen durch die Existenz unregelmäßiger, ähnlich klingender Verbformen beeinflusst wird.

Der hier vorgestellte konnektionistische Ansatz lässt sich mit bayesianischen Lernmodellen verbinden (Norris 2006: 334). Diese statistischen Modelle basieren auf bedingten Wahrscheinlichkeiten und fragen somit danach, wie wahrscheinlich bspw. ein Ereignis ist unter der Voraussetzung, dass ein anderes Ereignis eingetreten ist; bspw. gegeben, dass ein Metalldetektor bei einer Person ausschlägt, wie wahrscheinlich ist es, dass die Person tatsächlich metallische Gegenstände bei sich trägt? Das Beispiel ist angelehnt an Norris (2006: 330–331) und Carstensen u. a. (2010: 120–123). In Bezug auf Sprache (und automatisierte Spracherkennung) lässt sich bspw. fragen: Wenn ein bestimmter Input gegeben ist, welches Wort ist am wahrscheinlichsten? Jemand hört/sieht also einen bestimmten Input und möchte herausfinden, welches Wort bzw. welche mentale Repräsentation hierzu am besten passt (Norris 2006: 330). Diese abhängige Wahrscheinlichkeit lässt sich nach Bayes' Theorem wie folgt berechnen (Carstensen u. a. 2010: 123):

²⁶Interessanterweise gehen bspw. Clahsen & Rothweiler (1993: 6) und Beretta u. a. (2003: 72) davon aus, dass regelmäßige Verben im Deutschen weniger typenfrequent seien als unregelmäßige. Die dieser Annahme zugrunde liegende Zählung kritisiert Bybee (1995: 435–438): Darin wurde die Anzahl an starken Verben erhöht, indem nicht die Basislemmafrequenz, sondern die Lemmafrequenz ermittelt wurde und so Partikelverben auf Basis von *schreiben* als unterschiedliche Types gelten.

$$P_{A|I} = \frac{P_{I|A} * P_A}{P_{I|A} * P_A + P_{I|\neg A} * P_{\neg A}}$$

$P_{A|I}$ ist die Wahrscheinlichkeit für ein bestimmtes Wort A gegeben den Input I . Im Zähler stehen zwei Wahrscheinlichkeiten. $P_{I|A}$ ist die umgekehrte Wahrscheinlichkeit zu $P_{A|I}$: Dieser Wert gibt an, wie wahrscheinlich es ist, dass ein Input auf das Wort A passt. P_A ist die Wahrscheinlichkeit für Wort A unabhängig vom Input. Der Nenner benennt die Wahrscheinlichkeit für alle weiteren Szenarios, die unter dem gegebenen Input möglich sind: Zum einen findet sich hier noch einmal die Wahrscheinlichkeit des Inputs in Abhängigkeit von Wort A multipliziert mit der Wahrscheinlichkeit für Wort A . Zum anderen wird aber auch ein anderes Szenario in Betracht gezogen: Hierbei wird die Wahrscheinlichkeit für den Input betrachtet unter der Voraussetzung, dass ein anderes Wort als Wort A verwendet wurde ($= P_{I|\neg A}$). Diese wird wiederum multipliziert mit der Wahrscheinlichkeit für ein anderes Wort als Wort A , unabhängig vom Input ($P_{\neg A}$) (Norris 2006: 330).

Die Formel wird verständlicher, wenn wir uns das Metalldetektorbeispiel einmal näher ansehen. Nehmen wir an, dass der Metalldetektor bei 80 % der Personen, die metallische Gegenstände mit sich tragen, ausschlägt. Außerdem schlägt der Detektor bei 5 % der Personen fälschlicherweise aus, die keine metallischen Gegenstände mit sich tragen. Die Wahrscheinlichkeit für den Ausschlag gegeben, dass die Person Metall mit sich trägt ($= P_{A|M}$), beträgt also 80 %. Die Wahrscheinlichkeit für einen Ausschlag gegeben, dass die Person kein Metall mit sich trägt ($= P_{A|\neg M}$), liegt bei 5 %. Zudem nehmen wir an, dass von 100 Personen tatsächlich nur zehn Metall mit sich tragen, wenn sie durch den Metalldetektor gehen. Bevor jemand durch den Metalldetektor geht, liegt die Wahrscheinlichkeit dafür, dass er_sie Metall bei sich trägt, also bei 10 % ($= P_M$). Die Wahrscheinlichkeit, dass eine Person kein Metall bei sich hat ($P_{\neg M}$), ist dementsprechend $1 - P_M = 90$ %. Wie wahrscheinlich ist es nun, dass eine Person Metall bei sich trägt, nachdem der Metalldetektor ausgeschlagen hat?

$$P_{M|A} = \frac{P_{A|M} * P_M}{P_{A|M} * P_M + P_{A|\neg M} * P_{\neg M}} = \frac{0,8 * 0,1}{0,8 * 0,1 + 0,05 * 0,9} = 0,64$$

Die Wahrscheinlichkeit dafür, dass eine Person tatsächlich Metall bei sich trägt, wenn der Metalldetektor ausschlägt, ist 64 %. Durch die neue Evidenz (in diesem Fall über den Metalldetektor) korrigieren wir die vorherige (priore) Wahrscheinlichkeit (10 %) also deutlich nach oben und gehen von einer neuen, posterioren Wahrscheinlichkeit aus.

2 Variation aus gebrauchsbasierter Perspektive

Nehmen wir an, dass nur eine Person von 1.000 Personen Metall bei sich trägt. Die priore Wahrscheinlichkeit P_M ist nun also mit 0,1 % weitaus geringer als in dem oben diskutierten Beispiel. Wie wahrscheinlich ist es nun, dass jemand Metall mit sich trägt, gegeben, dass der Metalldetektor ausschlägt?

$$P_{M|A} = \frac{0,8 * 0,001}{0,8 * 0,001 + 0,05 * 0,999} = 0,015$$

Die Wahrscheinlichkeit dafür, dass eine Person tatsächlich Metall bei sich trägt, wenn der Detektor ausschlägt, beträgt nur noch 1,5 %. Obwohl der Detektor also reliabel arbeitet (zumindest werden die Schwellenwerte 5 % und 80 % für fälschlicherweise positiv bzw. negativ eingestufte Ergebnisse bspw. in Field u. a. 2012: 53–59 genutzt und empfohlen), ist die Wahrscheinlichkeit dafür, dass jemand Metall bei sich trägt, wenn der Detektor ausschlägt, sehr gering. Das liegt daran, dass die priore Wahrscheinlichkeit dafür, dass jemand Metall bei sich trägt, mit 0,1 % bereits gering war. Die posteriore Wahrscheinlichkeit (in diesem Fall die bedingte Wahrscheinlichkeit für Metall gegeben den Ausschlag) wird also direkt beeinflusst von der prioren Wahrscheinlichkeit (Mitchel 1997: 156–157). Bayes' Theorem erlaubt es daher, die priore Wahrscheinlichkeit mit neuen Informationen abzugleichen und daraus eine posteriore Wahrscheinlichkeit zu ermitteln. Dabei berücksichtigt die posteriore Wahrscheinlichkeit für eine Hypothese die vorherige Wahrscheinlichkeit für die Hypothese zusammen mit dem Wahrscheinlichkeitsverhältnis (*likelihood ratio*), dass die neue Information mit den vorherigen Hypothesen ($P_{A|M}/P_{A|\neg M}$) konsistent ist (Norris 2006: 330).

An dieser Stelle wird auch deutlich, weshalb bayesianische Modelle als Sprachlern- und Sprachverarbeitungsmodelle²⁷ nützlich sind: Wenn wir kommunizieren, müssen wir ständig den Input, den wir bekommen, mit unserem Wissen abgleichen. Dabei muss die priore Wahrscheinlichkeit bspw. für ein Wort mit beobachteten Daten abgeglichen und in Reaktion darauf die posteriore Wahrscheinlichkeit ermittelt werden. Zudem erlauben es bayesianische Modelle, Frequenz als Teil der prioren Wahrscheinlichkeit für eine bestimmte Struktur zu modellieren. Nehmen wir an, eine Person hört ein Wort und möchte herausfinden, zu welcher mentalen Repräsentation das gesprochene Wort am besten passt. Es stellt sich also die Frage danach, welches Wort am wahrscheinlichsten ist unter einem bestimmten Input ($P_{W|I}$). Um diese Wahrscheinlichkeit zu berechnen,

²⁷Für eine ausführliche Erläuterung eines bayesianischen Ansatzes zur Worterkennung siehe Jurafsky (1996) sowie Norris (2006). Darin werden auch Berechnungen vorgeschlagen, die in Hinblick auf ein gegebenes Wort die Wahrscheinlichkeit für einen bestimmten Input ermitteln. Baayen u. a. (2011) schlagen einen modifizierten Ansatz vor, der ebenfalls als bayesianisch angesehen werden kann (Baayen u. a. 2011: 70).

benötigen wir die umgekehrte bedingte Wahrscheinlichkeit $P_{I|W}$ für den Input gegeben das Wort, die Wahrscheinlichkeit $P_{I|\neg W}$ für den Input gegeben andere Wörter und die Wahrscheinlichkeit für das Auftreten des Worts unabhängig vom Input (P_W) sowie die Wahrscheinlichkeit für das Auftreten eines anderen Worts unabhängig vom Input ($P_{\neg W}$), also die prioren Wahrscheinlichkeiten.

P_W und $P_{\neg W}$ sind abhängig von der Tokenfrequenz: Je häufiger ein Wort ist, desto wahrscheinlicher ist seine Nutzung (Norris 2006: 330). Die priore Wahrscheinlichkeit hängt aber nicht nur von Frequenz ab, sondern auch von anderen Faktoren. So kann bspw. syntaktisches Wissen die Wahrscheinlichkeit für Wörter in einem Satz beeinflussen (Norris 2006: 331). Zudem zeigen McDonald & Shillcock (2003), dass Abfolgewahrscheinlichkeiten Einfluss auf das Leseverhalten nehmen und somit ebenfalls als Teil der prioren Wahrscheinlichkeit angesehen werden können. Die umgekehrte bedingte Wahrscheinlichkeit $P_{I|W}$ wird bspw. dadurch beeinflusst, wie eindeutig der Input auf das gesuchte Wort passt (Norris 2006: 330). Wenn bspw. der Anfang des Inputwortes nicht zu verstehen ist oder es noch nicht zu Ende gesprochen wurde, ergeben sich neben dem tatsächlichen Wort andere mögliche Wörter, die auf das Inputwort passen könnten. Diese stellen die Wahrscheinlichkeit für den Input gegeben andere Wörter dar ($P_{I|\neg W}$). Kann das Inputwort hingegen zweifelsfrei erkannt und andere konkurrierende Wortformen ausgeschlossen werden, ist die Wahrscheinlichkeit für das Inputwort gegeben das Wort 100 % und dementsprechend die Wahrscheinlichkeit für das Inputwort gegeben ein anderes Wort 0 %. Dann beeinflussen die priore Wahrscheinlichkeiten (und damit die Frequenz der Wörter) die posteriore Wahrscheinlichkeit nicht mehr, da sie sich verkürzen bzw. mit 0 multipliziert werden (Norris 2006: 331):

$$P_{W|I} = \frac{P_{I|W} * P_W}{P_{I|W} * P_W + P_{I|\neg W} * P_{\neg W}} = \frac{1 * P_W}{1 * P_W + 0 * P_{\neg W}} = \frac{1 * P_W}{1 * P_W} = 1$$

Das Modell blockiert infrequente Strukturen daher nicht generell, sondern Frequenz beeinflusst nur ambige Strukturen, die bspw. noch nicht komplett prozessiert wurden: Da in dem probabilistischen Ansatz die Rolle der Frequenz als priore Wahrscheinlichkeit integriert ist, nimmt sie nicht alleine Einfluss auf die Sprachprozessierung. Nehmen wir an, dass ein_e Sprecher_in kein Wissen über die Frequenz des Wortes hat, so wird lediglich die Wahrscheinlichkeit dafür, dass der Input mit dem Wort konsistent ist, betrachtet und das Wort gewählt, das dem Input am besten entspricht (Norris 2006: 330). Die Frequenz verändert als priore Wahrscheinlichkeit die posteriore Wahrscheinlichkeit dahingehend, dass bei ambigem Input frequentere Wörter wahrscheinlicher sind als infrequente Wör-

2 Variation aus gebrauchsbasierter Perspektive

ter (Jurafsky 1996: 157, Norris 2006: 331). Je weniger frequent ein Wort ist, desto mehr Input wird benötigt, damit $P_{I|W}$ groß genug wird, um das Wort zu erkennen (Norris 2006: 334). Neben der Frequenz können weitere Faktoren wie bspw. der Kontext als Teil der prioren Wahrscheinlichkeit Einfluss auf die posteriore Wahrscheinlichkeit nehmen.

Der bayesianische Ansatz kann zur Modellierung des Einflusses der Typen- und Tokenfrequenz bspw. auf Konjugation herangezogen werden: Auch hier ist die Typen- bzw. die Tokenfrequenz die priore Wahrscheinlichkeit für das Auftreten einer Struktur. Je höher die Typen- oder Tokenfrequenz ist, desto höher die priore Wahrscheinlichkeit für eine bestimmte Konjugationsklasse oder eine bestimmte Flexionsform. In einem bayesianischen Modell kann der Einfluss der Frequenz dabei nicht nur beobachtet, sondern auch erklärt werden: Sie ist relevant, da sie Einfluss auf die priore Wahrscheinlichkeit für das Auftreten einer bestimmten Struktur nimmt (Jurafsky 1996: 157).

Auch Prozessierungsprobleme lassen sich mithilfe des bayesianischen Ansatzes erklären. Beispielsweise lösen Verletzungen in der Kongruenz von Numerus, Kasus und Genus Probleme in der Prozessierung aus (Osterhout u. a. 1994, Kanan & Swaab 2003, Gouvea u. a. 2010: 150–151). Diese Schwierigkeiten zeigen sich in EEG-Studien im P600-Effekt, einem ereigniskorrelierten Potential (EKP), das einen positiven Ausschlag rund 600 ms nach Onset des Stimulus benennt (Kanan u. a. 2000: 169–172, Gouvea u. a. 2010: 155–156). Kongruenzverletzungen lösen zusätzlich einen LAN-Effekt (*left anterior negativity*) aus (Roehm u. a. 2005: 878). Dieser tritt nicht nur bei Verletzungen von Verbkongruenzen (*Every morning he *mow the lawn*) auf, sondern auch bei Abweichungen von der Phrasenstruktur (*The students enjoyed Bill's *of review of play*) (Hahne & Friederici 1999, Gouvea u. a. 2010: 150–151). Auch semantisch auffällige Sätze können Schwierigkeiten in der Prozessierung evozieren, die sich bspw. im N400-Effekt zeigen. Dieser Effekt wird bspw. durch Sätze hervorgerufen, deren Bedeutung nicht zu unserem Weltwissen passt, wie bspw. *Die Pizzen sitzen auf dem Baum* (Kutas & Hillyard 1983: 540). Die Prozessierungsprobleme lassen sich in einem bayesianischen Modell gut fassen: Die erwähnten Strukturen haben gemein, dass sie Abweichungen von kanonischen Verwendungen darstellen. Somit ist die priore Wahrscheinlichkeit für diese Strukturen gering, da sie so gut wie nie im sprachlichen Input vorkommen, und gleichzeitig existiert eine andere, konkurrierende Struktur mit hoher priorer Wahrscheinlichkeit.

Wie neue Evidenz die Prozessierung sprachlicher Strukturen beeinflusst, lässt sich anhand von Holzwegsätzen (*garden path sentences*) gut zeigen. Holzwegsätze lösen einen P600-Effekt aus und führen außerdem zu einer Erhöhung der

Lesezeit (Tabor & Galantucci 2000). Sie sind semantisch ambig und verleiten Leser_innen zunächst zu einer Interpretation, die sich im Verlauf des Satzes als semantisch unsinnig oder ungrammatisch herausstellt, sodass der Satz neu prozessiert werden muss. Fine & Jaeger (2013) präsentieren Proband_innen Holzwegsätze, in denen die Ambiguität durch einen reduzierten Nebensatz ausgelöst wird (siehe Beispiel 1, das Fine & Jaeger 2013: 2283 entnommen wurde). *Warned* in Beispiel 1 könnte sowohl das finite Verb des Hauptsatzes (Beispiel 1a) als auch des Relativsatzes sein (Beispiel 1b). Dabei wird die Interpretation als Verb des Hauptsatzes präferiert, bis der_die Leser_in zum eigentlichen finiten Verb des Hauptsatzes (*conducted*) kommt. Die ursprüngliche Interpretation der ambigen Struktur geht an dieser Stelle nicht mehr auf, sodass die Interpretation korrigiert werden muss. Dementsprechend sind an dieser Stelle hohe Lesezeiten messbar (Fine & Jaeger 2013: 2279–2280).

- (1) *The experienced soldiers warned about the dangers conducted the midnight raid.*
 - a. *The experienced soldiers warned about the dangers *conducted the midnight raid.*
→ präferierte Lesart ohne Relativsatz.
 - b. *The experienced soldiers [warned about the dangers] conducted the midnight raid.*
→ dispräferierte Lesart mit reduziertem Relativsatz.
 - c. *The experienced soldiers [who were warned about the dangers] conducted the midnight raid.*
→ disambiguerter Satz mit nicht-reduziertem Relativsatz.

Die Präferenz für eine Lesart ergibt sich aus der Frequenz der Konstruktionen: *Warned* dient häufiger als finites Verb des Hauptsatzes als des Nebensatzes (Fine & Jaeger 2013: 2280). Die priore Wahrscheinlichkeit für *warned* in einem Hauptsatz ist also höher als für *warned* in einem Nebensatz. Wenn Sprecher_innen eine ambige Struktur prozessieren, gehen sie nach dem bayesianischen Ansatz zunächst von der wahrscheinlicheren Konstruktion aus. Diese Strategie geht bei Holzwegsätzen nicht auf, weswegen Prozessierungsschwierigkeiten entstehen.

Fine & Jaeger (2013) überprüfen, ob sie durch Verschiebungen in der Frequenz der Konstruktionen die Präferenz für eine Lesart ändern können. Dafür lassen Fine & Jaeger (2013: 2280–2281) Proband_innen Holzwegsätze sowie disambiguierte Sätze wie in Beispiel 1c lesen. Dabei nimmt die anfänglich erhöhte Lesezeit für Holzwegsätze im Lauf des Experiments ab. Die Proband_innen haben

2 Variation aus gebrauchsbasierter Perspektive

also die neue Evidenz aus dem sprachlichen Input genutzt, um die priore Wahrscheinlichkeit für *warn* in einem Relativsatz zu aktualisieren. In einem weiteren Experiment zeigen Fine & Jaeger (2013: 2281–2283), dass nun umgekehrt längere Lesezeiten für Sätze entstehen, in denen *warn* im Hauptsatz verwendet wird: Der Satz *The experienced soldiers warned about the dangers before the midnight raid* evoziert also längere Lesezeiten ab *before*, da die priore Wahrscheinlichkeit für eine Verwendung von *warned* im Relativsatz nun höher ist als für eine Verwendung von *warned* im Hauptsatz. Basierend auf der prioren Wahrscheinlichkeit erwarten die Proband_innen ein finites Verb vor *before the midnight raid* und müssen nun korrigieren, was sich in hohen Lesezeiten widerspiegelt. Die Ergebnisse deuten also darauf hin, dass die Proband_innen die neuen statistischen Verhältnisse im Input nutzen, um ihre Erwartungen an den weiteren Input anzupassen.

Dies zeigt auch eine Studie von Hanulíková u. a. (2012): Wird ein nicht-kanonisches Genus (bspw. **die Tisch*) von L1-Sprecher_innen des Deutschen produziert, führt dies bei Hörer_innen zu einem P600-Effekt. Dieser Effekt bleibt jedoch aus, wenn dieselbe Abweichung von L2-Sprecher_innen produziert wird. Auch hier kann bayesianisch argumentiert werden: Hören L1-Sprecher_innen des Deutschen Input von L2-Sprecher_innen ist die priore Wahrscheinlichkeit für Genusabweichungen höher als wenn sie Input von L1-Sprecher_innen hören (Hanulíková u. a. 2012: 879–880). Einen vergleichbaren Effekt messen Hahne & Friederici (1999). In ihrer EEG-Studie variieren sie die Anzahl an Abweichungen von der kanonischen Phrasenstruktur. Bei den abweichenden Sätzen folgte in einem Passivsatz auf eine Präposition statt der zu erwartenden Nominalphrase ein Vollverb im Partizip II (*Die Gans wurde im gefüttert*). Eine der beiden Konditionen enthält mit 20 % nur wenige Abweichungen, die andere mit 80 % viele. Hahne & Friederici (1999: 196–199) konnten nur in der Kondition mit wenigen ungrammatischen Stimuli einen P600-Effekt messen, unabhängig von der Kondition war ein LAN-Effekt zu beobachten. Wenn also die priore Wahrscheinlichkeit für ungrammatische Sätze hoch war, riefen diese keinen P600-Effekt mehr hervor. Die Prozessierung sprachlicher Strukturen scheint somit stark von prioren Wahrscheinlichkeiten geprägt zu sein.

Die Interpretation des P600-Effekts als eine Reaktion darauf, dass eine geläufigere und somit wahrscheinlichere Konstruktion existiert als die prozessierte, teilt auch Goldberg (2019: 83–84): Sie weist darauf hin, dass auch Rechenfehler ($2+2=5$) und Rechtschreibfehler (*The toddler is too years old*) P600-Effekte evozieren. In beiden Fällen existiert eine präferierte Alternative; dasselbe gilt für die oben beschriebenen strukturellen Auffälligkeiten, die P600-Effekte auslösen.

P600-Effekte geben einen Hinweis auf das Wirken eines statistischen Vorkaufsrechts (*statistical preemption*) in der Prozessierung (Goldberg 2019: 83–84):

Wenn bereits eine konventionelle Weise existiert, um etwas auszusagen, dann ist die Wahrscheinlichkeit höher, dass diese genutzt wird als die unkonventionelle Weise (Goldberg 2019: 86). Auf diese Art können Kinder Übergeneralisierungen verlernen: Wenn ein Kind die Wortform *Ball* nicht nur mit der Bedeutung ‚Ball‘, sondern aufgrund der Eigenschaft [+rund] auch mit der Bedeutung ‚Knopf‘ verknüpft hat, kann es diese Assoziation verlernen, indem es die Wortform *Knopf* in Kontexten hört, in denen statt der Bedeutung ‚Ball‘ die Bedeutung ‚Knopf‘ gemeint ist. Begegnet dem Kind die Wortform *Knopf* in diesen Kontexten immer wieder, kann die Wortform *Knopf* die Wortform *Ball* statistisch ausstechen, denn *Knopf* ist in diesen Kontexten häufig und *Ball* selten bis inexistent. Derselbe Mechanismus ist auf Konstruktionsebene zu beobachten.

Das statistische Vorkaufsrecht ist ein wichtiger Aspekt des fehlerbasierten Lernens (*error-driven learning*) (Goldberg 2019: 94). Ein weiterer Aspekt des fehlerbasierten Lernens ergibt sich dadurch, dass wir versuchen, Vorhersagen darüber zu treffen, was als Nächstes gesagt werden wird, wenn wir Sprache prozessieren (Goldberg 2019: 91). Entspricht unsere Erwartung nicht dem Inhalt (das Kind sieht einen Knopf und erwartet, dass *Ball* gesagt wird, aber die Bezugsperson sagt *Knopf*), gibt es ein Fehlersignal (Goldberg 2019: 91). Aufgrund des Fehlersignals haben überraschende, unvorhergesehene und somit saliente Informationen einen großen Einfluss auf unsere Erinnerung (Goldberg 2019: 116). Das Fehlersignal kann daher genutzt werden, um zukünftige Vorhersagen zu verbessern und zukünftige Fehlersignale gering zu halten.²⁸ Die Wahrscheinlichkeit für ein Wort in einem bestimmten Kontext wird immer wieder überprüft und (wenn es ein Fehlersignal gibt) aktualisiert. Zusätzlich wird das statistische Vorkaufsrecht und damit auch fehlerbasiertes Lernen durch Selbstsicherheit (*confidence*) beeinflusst (Goldberg 2019: 88). Je häufiger ein_e Sprecher_in ein Verb in einer bestimmten Konstruktion, aber nicht in einer anderen, hört (bspw. *explain this to me* statt *explain me this*), desto sicherer wird der_die Sprecher_in, dass für das jeweilige Verb eine Konstruktion die andere aussticht. Dies führt dazu, dass die unkonventionelle Form vermieden wird.

Die Wahrscheinlichkeit dafür, dass eine Konstruktion die andere statistisch aussticht und so das Vorkaufsrecht erhält, hängt laut Goldberg (2019: 85) von dem Vorkommen der Konstruktion in Kontexten ab, in denen auch eine ande-

²⁸L2-Sprecher_innen können Formen schlechter vorhersagen als L1-Sprecher_innen (Goldberg 2019: 115–116). Zudem funktioniert das statistische Vorkaufsrecht zwischen Konstruktionen bei L2-Sprecher_innen nicht so gut wie bei L1-Sprecher_innen: Neue Strukturen werden von L2-Sprecher_innen unabhängig davon, ob die Struktur im Wettbewerb zu einer anderen steht, weniger akzeptiert als bekannte Strukturen (Goldberg 2019: 116). L2-Sprecher_innen können daher weniger von fehlerbasiertem Lernen profitieren als L1-Sprecher_innen.

re Konstruktion möglich wäre. Der Beispielsatz *?explain me this* hat danach eine geringe Wahrscheinlichkeit, die *to*-Dativ-Konstruktion auszustechen, weil die meisten Verwendungen von *explain* in der *to*-Dativ-Konstruktion (*explain this to me*) stehen und sie somit gebräuchlicher ist als die ditransitive Konstruktion. Für andere Verben sind beide Konstruktionen gleich häufig (*tell the boy a story/tell a story to the boy*), sodass die Konstruktionen sich nicht statistisch ausstechen können. Eine Konstruktion kann die andere jedoch kontextabhängig ausstechen: Wenn der Rezipient durch ein Pronomen ausgedrückt wird, sticht die ditransitive Konstruktion (*tell him a story*) die *to*-Dativ-Konstruktion statistisch aus (*tell a story to him*), da sie in diesem Diskurskontext häufiger vorkommt (Goldberg 2019: 86). Die Häufigkeit lässt sich dabei aus der Informationsstruktur ableiten: Der Rezipient (*him*) ist bereits im Kontext bekannt und steht daher idealerweise vor der neuen Information (*a story*) (Goldberg 2019: 42–43).

Im Gegensatz zum reinen Blick auf Frequenz mit der Annahme, dass wenig frequente Konstruktionen weniger akzeptabel sind, erlaubt diese statistische Perspektive einen genaueren Blick auf Akzeptabilität. Wird durch die infrequente Form eine andere kontextuelle Funktion ausgedrückt, kann kein Abgleich mit existierenden Konstruktionen stattfinden (Goldberg 2019: 75–76). Deshalb ist die transitive Verwendung von *to sneeze* in *She sneezed the tissue over the table* akzeptabel, obwohl *sneeze* i. d. R. nicht transitiv genutzt wird. Die intransitive Konstruktion (*She sneezed*) hat eine andere Bedeutung, weswegen sie nicht mit der anderen Konstruktion konkurriert (Goldberg 2019: 76). Evidenz für dieses Prinzip liefern Perek & Goldberg (2017): In einem Experiment wurden Proband_innen zwei Szenen gezeigt und mithilfe von zwei verschiedenen Konstruktionen beschrieben. Die Konstruktionen waren jeweils mit einer der Szenen assoziiert (Szene A: ein Hase schlägt eine Katze, die nur leicht weghüpft; Szene B: ein Hase schlägt eine Katze, die aus dem Bild springt). Perek & Goldberg (2017: 283–286) ließen Proband_innen sechs Verben lernen: drei Verben wurden nur in Konstruktion A und in der mit ihr assoziierten Szene genutzt, zwei nur in Konstruktion B in der mit ihr assoziierten Szene. Das letzte Verb wurde immer in derselben Konstruktion genutzt, aber in beiden Szenen präsentiert. Anschließend sollten die Proband_innen neue Beschreibungen der Szenen bewerten und selbst produzieren. Dabei zeigte sich, dass die Proband_innen eine mentale Verbindung zwischen den Szenen und den Konstruktionen geschaffen hatten. Unabhängig vom Verb wurde also Konstruktion A nur in der mit ihr assoziierten Szene genutzt, dasselbe galt für Konstruktion B (Perek & Goldberg 2017: 286–312). Obwohl die Verben also zuvor nur mit einer Konstruktion präsentiert wurden, scheuten die Proband_innen sich nicht davor, die Verben für die andere Szene in der mit der Szene assoziierten Konstruktion zu nutzen. Nur bei dem Verb, das unabhängig

von der Szene immer in derselben Konstruktion verwendet wurde, beharrten die Proband_innen auf dieser Konstruktion unabhängig von der Szene. Das Verb wurde also von den Proband_innen nur in der Konstruktion genutzt, in der es ihnen zuvor präsentiert wurde. Dies lässt sich mit dem statistischen Vorkaufsrecht erklären: Im Fall der Verben, die nur auf eine Szene beschränkt waren, greift das statistische Vorkaufsrecht nicht, da die Proband_innen nur Evidenz für die Verwendung der Verben in einer mit der Szene assoziierten Konstruktion hatten. Bei dem Verb, das unabhängig von der Szene in einer bestimmten Konstruktion genutzt wurde, greift das Vorkaufsrecht hingegen: Unabhängig von der Szene wird im Input eine Konstruktion bevorzugt, sodass diese die andere aussticht (Perek & Goldberg 2017: 286–312).²⁹

Insgesamt verdeutlicht die Diskussion, dass Frequenz einen grundlegenden Einfluss auf den Sprachgebrauch, die Prozessierung und somit auf die mentale Repräsentation von Sprache hat. Frequenz gilt dabei als einer der besten Prädiktoren für menschliches Verhalten (Divjak & Caldwell-Harris 2015: 55). Allerdings korreliert Frequenz mit anderen Faktoren, sodass sie zwar ein guter Prädiktor ist, jedoch keine gute Erklärung für die Distribution der Daten liefern kann, wie die Studie von Baayen (2012) verdeutlicht. Frequenzeffekte sind deshalb nicht als pure Wortwiederholung und somit als isoliert zu betrachten, stattdessen sollten zusätzlich bspw. Verwendungskontexte berücksichtigt werden. So sind Fachbegriffe wie *Typenfrequenz* und *Tokenfrequenz* zwar in der Regel seltene Wörter, fallen sie aber in einem linguistischen Kontext, ist eine schnellere Verarbeitung zu erwarten als würden sie in einem alltäglichen Gespräch genutzt werden. Zudem ist stets zwischen Typen- und Tokenfrequenz zu unterscheiden: Hohe Typenfrequenz hängt direkt mit der Produktivität einer Kategorie zusammen. Mitglieder typenfrequenter Kategorien haben i. d. R. eine höhere Variabilität als Mitglieder typeninfrequenter Kategorien, weshalb typenfrequente Kategorien eine größere Abdeckung haben als typeninfrequente Kategorien. Potentielle Neuzugänge können daher eine größere Ähnlichkeit zu bereits bestehenden Exemplaren aufweisen. Dies lässt sich bspw. bei Flexionsklassen beobachten: Aufgrund der vielen Mitglieder und der Variabilität der Mitglieder lassen typenfrequente Flexionsklassen eine Generalisierung zu, die Sprecher_innen dazu führt, ein regelmäßiges Verhalten zu erkennen. Dies ist bei typeninrequenten Klassen nur bedingt der Fall, sodass diese als unregelmäßig wahrgenommen werden.

Die typeninfrequente Klasse (bspw. unregelmäßige Verben) enthält vornehmlich tokenfrequente Mitglieder. Die Unregelmäßigkeit typeninfrequenter Klassen ergibt sich dabei aus der Tokenfrequenz der wenigen Mitglieder: Tokenfrequente Formen sind stark *entrenched* und daher leicht zu aktivieren (Bybee &

²⁹Für weitere empirische Studien zum statistischen Vorkaufsrecht siehe Goldberg (2019: 77–84).

Thompson 1997: 380, Schneider 2014: 13). Deswegen stellt Unregelmäßigkeit für frequente Verben keinen Nachteil, sondern einen Prozessierungsvorteil dar, denn sie führt zu kurzen und gut differenzierten Formen (Nowak 2013: 174). Zudem ergibt sich aus hoher Tokenfrequenz ein Autonomie-Effekt, sodass die Formen hochfrequenter Verben weniger stark verbunden sind als die Formen niedrigfrequenter Verben (Bybee 1985: 117–124, 2006: 715). Der Autonomie-Effekt zeigt sich auch darin, dass eine hohe Tokenfrequenz zu einem Flexionsklassenwechsel führen kann (Nübling 2000: 17–18, siehe Abschnitt 3.1.1). Tokenfrequenz hat zudem einen konservierenden Effekt: Tokenfrequente Mitglieder einer typeninrequenten Flexionsklasse halten sich in dieser Klasse, während die tokeninfrequenten zur typenfrequenten Klasse abwandern (Bybee & Thompson 1997: 280, siehe Abschnitt 3.1.1). Neben Autonomie und Konservierung kann hohe Tokenfrequenz zu Reduktion und Fusion führen: Hochfrequente Elemente, die oft zusammen genutzt werden, werden als eine Einheit wahrgenommen und verschmolzen.

Die diskutierten Frequenzeffekte lassen auf ein konnektionistisches Sprachmodell schließen: So nimmt nicht nur die reine Worthäufigkeit Einfluss auf Reaktionszeiten, sondern auch die Existenz orthographischer Nachbarn. Zudem zeigt sich, dass Sprecher_innen sensitiv für Abfolgewahrscheinlichkeiten sind. In einem probabilistischen Ansatz wie dem bayesianischen kann Frequenz als Faktor modelliert werden, der unter anderen Faktoren Einfluss auf die priore Wahrscheinlichkeit einer bestimmten Struktur nimmt, die anhand neuer Evidenz angepasst werden kann. Die Modellierung als priore Wahrscheinlichkeit liefert eine Erklärung für den Einfluss der Frequenz: Frequenz kann bei ambigen Strukturen als priore Wahrscheinlichkeit Einfluss auf die Prozessierung nehmen. Ihr Einfluss kann aber auch ausgehebelt werden. Zudem lässt sich das Modell nutzen, um einen statistischen Wettbewerb zwischen ähnlichen Konstruktionen zu modellieren: Gibt es einen frequenteren Weg, um einen Inhalt auszudrücken, sticht dieser den weniger frequenten Weg aus. Eine wenig frequente Konstruktion kann aber dennoch genutzt werden, wenn sie einen anderen Verwendungskontext hat als eine frequentere Konstruktion, sodass sie nicht von dieser ausgestochen werden kann.

Im weiteren Verlauf der Arbeit wird Frequenz als ein Faktor verstanden, der Einfluss auf die priore Wahrscheinlichkeit und damit auf das statistische Vorkaufsrecht von Konstruktionen nimmt. Auf diese Weise ist Frequenz kein leicht messbarer Prädiktor, sondern ein Faktor, dessen Einfluss auf Prozessierung erklärbar ist. Im weiteren Verlauf der Arbeit wird das Spiel zwischen Typen- und Tokenfrequenz und das sich daraus ergebende statistische Vorkaufsrecht sowie *entrenchment* durch Tokenfrequenz relevant sein. Im empirischen Teil wird der konservierende Einfluss der Tokenfrequenz psycholinguistisch überprüft.

Im folgenden Abschnitt wird Prototypizität als Faktor vorgestellt, der Einfluss auf Variation nimmt. Hierfür wird zunächst diskutiert, wie Prototypizität die Bildung von Kategorien und die Organisation mentaler Repräsentationen beeinflusst.

2.2 Prototypizität

Neben der Frequenz ist die Prototypizität kognitiver Kategorien ein grundlegendes Konzept im gebrauchsbasierten Ansatz. Prototypisch organisierte Kategorien basieren auf einzelnen, ähnlichen Exemplaren, die im sprachlichen Input wahrgenommen und zu einer Kategorie abstrahiert werden (Ellis u. a. 2014: 56). Wie der Abschnitt zeigen wird, nimmt Prototypizität Einfluss auf Variation und Stabilität in einem Sprachsystem.

Das Konzept der Prototypizität stammt ursprünglich aus der Kognitionspsychologie und wurde dort vornehmlich von Rosch entwickelt (siehe Rosch 1973, 1975a,b, Rosch & Mervis 1975, Rosch, Mervis u. a. 1976, Rosch, Simpson u. a. 1976, Rosch 2004 [1978]). Im Fokus standen zunächst vornehmlich lexikalische Kategorien.³⁰ Dieser Fokus zeigt sich auch in der Linguistik: Mit der Entwicklung der Prototypensemantik stehen auch hier lexikalische Kategorien und ihre Semantik im Vordergrund. Prototypizität lässt sich aber als generelles Ordnungsprinzip von kognitiven Kategorien ausmachen (Lakoff 1987: 57–58) und greift daher auch für grammatische Kategorien (Bybee & Moder 1983, Taylor 1995: 173–238, Bybee 2010: 78–80).

Der Abschnitt verfolgt zwei Ziele: Einerseits soll Prototypizität definiert und andererseits ihr genereller Einfluss auf Variation und Stabilität erläutert werden. Abschnitt 2.2.1 stellt Prototypizität zunächst als Ordnungsprinzip sprachlicher Kategorien vor, dabei werden auch die Grundannahmen der Prototypensemantik diskutiert. In diesem Abschnitt steht zudem die Frage im Vordergrund, wie weit Prototypizität den Aufbau sprachlicher Kategorien erklären kann. Im anschließenden Abschnitt 2.2.2 werden Prototypeneffekte bei lexikalischen und grammatischen Kategorien diskutiert, hieraus wird der Einfluss der Prototypizität auf Variation und Stabilität in einem Sprachsystem modelliert. Abschließend werden in Abschnitt 2.2.3 prototyp- und exemplarbasierte Kategorisierungsansätze verglichen und prototypisch organisierte Kategorien als Resultat einer exemplarbasierten Kategorisierung modelliert. Diese Modellierung ist die Grundlage für die

³⁰Um die Vielfalt der untersuchten Kategorien zu verdeutlichen, folgt hier eine beispielhafte Liste: Farben und Formen (Rosch 1973); Früchte, Gemüse, Möbel, Fahrzeuge, Waffen, Werkzeuge, Vögel, Sportarten, Spielzeug und Kleidung (Rosch 1975b, Rosch & Mervis 1975, Rosch, Simpson u. a. 1976); Tassen, Schüsseln und Vasen (Labov 2004 [1973]).

weitere Betrachtung prototypisch organisierter Kategorien in dieser Arbeit: Kategorien sind in dieser Betrachtung graduell aufgebaut mit prototypischen und peripheren Vertretern. Dabei werden prototypische Vertreter anders verarbeitet als periphere. Die Gradualität lässt sich nutzen, um die Reihenfolge von Variationsfällen zu prognostizieren. Die Unterschiede in der Verarbeitung dienen als Erklärungsansatz für Variation. Während Variation für prototypische Vertreter unwahrscheinlich ist, ist sie für periphere wahrscheinlich.

2.2.1 Prototypizität sprachlicher Kategorien

Die Prototypizität sprachlicher Kategorien ist eine grundlegende Annahme der Prototypensemantik. Die Standardversion der Prototypensemantik sieht Prototypen als „typische, gute und hervorragende Vertreter einer Kategorie“ (Schmid 2000: 33). So ist bspw. ein Gefäß aus Porzellan, das so hoch wie breit ist, einen Henkel hat und zum Kaffee- oder Teetrinken verwendet wird, ein prototypischer Vertreter der Kategorie TASSE. Die Bezeichnung *Prototyp* geht dabei auf das Exemplar zurück, das vor einer Serienproduktion entwickelt wird. Der Prototyp ist hier Vorbild für die folgenden hergestellten Exemplare (Kleiber 1993: 31). Genau so bildet der Prototyp einer Kategorie in der Prototypensemantik das Zentrum derselben, während die anderen Vertreter der Kategorie sich um den Prototyp herum gruppieren.

Die zentrale Idee der Prototypensemantik ist nicht allein die Annahme eines Prototyps, sondern die Idee, dass die Zugehörigkeit zu einer Kategorie graduell ist (Mangasser-Wahl 2000: 15). Langackers Definition eines Prototyps trägt dem Prinzip der Gradualität explizit Rechnung: „A prototype is a typical instance of a category, and other elements are assimilated to the category on the basis of their perceived resemblance to the prototype; there are degrees of membership based on degrees of similarity“ (Langacker 1987: 371).³¹ Es können also randständige Vertreter einer Kategorie existieren, wie bspw. Tassen ohne Henkel.

³¹Langacker (1987: 371) nutzt diese Definition, um Prototypen von der Kategorisierung durch Schemata abzugrenzen. Langacker (1987) definiert ein Schema als eine Charakterisierung, die auf alle Mitglieder der von ihr definierten Kategorie zutrifft. Hier ist somit keine Gradualität vorgesehen (Taylor 1990: 530). Schemata sind dabei hierarchisch organisiert. Diese Organisation ergibt sich aus Weiterentwicklungen des Prototyps (Taylor 1995: 66): Wenn wir eine Eiche, eine Buche und einen Ahorn sehen, können wir daraus das Schema *Baum*₁ abstrahieren. Dies dient nun als Prototyp für die Kategorie BAUM. Eine Tanne lässt sich aufgrund der graduellen Zugehörigkeit als Baum ansehen, ist aber nicht kompatibel mit dem Schema *Baum*₁, da Tannen keine Blätter haben. Da aber Tannen aufgrund der Ähnlichkeit zum Prototyp mit Bäumen assoziiert sind, extrahieren wir ein zweites Schema *Baum*₂, das Gemeinsamkeiten von *Baum*₁ und Tannen enthält: Ein Baum hat einen Stamm und Äste. Dies kann zu einem noch abstrakteren Schema *Baum*₃ führen, in dem bspw. auch Syntaxbäume enthalten sind, da hier

Wie lexikalische Kategorien lassen sich auch grammatische Kategorien als prototypisch organisiert auffassen: Lakoff (1987: 59–67) sieht phonologische, morphologische und syntaktische Kategorien als prototypisch organisiert an.³² Taylor (1995: 175–182) diskutiert Wörter und Affixe als prototypisch organisierte Kategorien. Zudem wendet er die Prototypentheorie auf Phoneme (Taylor 1995: 223–229) sowie auf morphologische und syntaktische Konstruktionen an (Taylor 1998: 192–198, 2015: 570–575; genauer zu Konstruktionen siehe Abschnitt 2.3.2).

Die Prototypizität grammatischer Kategorien lässt sich anhand von Wortarten verdeutlichen (Taylor 1998: 181–184, Hundt 2000: 8–9). Prototypische Eigenschaften der Wortart Adjektiv sind die Flexion, die Steigerbarkeit und die Möglichkeit, es attributiv, prädikativ und adverbial zu verwenden. Adjektive, die diesen Eigenschaften voll entsprechen, sind bspw. *schön*, *schnell* und *kreativ*. Adjektive wie *schwanger* verfügen hingegen nicht über alle Eigenschaften, da sie nicht gesteigert werden können. Adverbien wie *heute* lassen sich nach dieser Auffassung als Adjektive auffassen, die mit anderen Adjektiven nur die Eigenschaft teilen, dass sie adverbial verwendbar sind.³³ Vergleichbar argumentiert Taylor (2015: 569–570) für eine prototypische Organisation von Substantiven: So kann bspw. *Musik* als ein unprototypisches Substantiv gelten, da *Musik* nicht pluralisierbar ist. Wie bei lexikalischen Kategorien lässt sich somit auch für grammatische Kategorien eine graduelle Zugehörigkeit der Vertreter postulieren.

Prototypisch organisierte Kategorien zeigen sich auch in der Flexionsmorphologie. Hierbei lässt sich einerseits nach der prototypischen phonologischen Form von Wörtern fragen, die ein bestimmtes Flexionsverhalten aufweisen (Bybee 1985: 111–135; siehe hierzu auch Abschnitt 2.3.2 zu Form-Schemata) und andererseits nach den Flexionseigenschaften, die verschiedene Flexionsklassen voneinander trennen: Je mehr Eigenschaften der einen Flexionsklasse aufgegeben werden, desto mehr Eigenschaften der anderen Flexionsklasse kommen hinzu.

nur noch das Konzept des Verästeln relevant ist. Auf die gleiche Weise können auch konkretere Subschemata entstehen wie bspw. das Schema *Fruchtbaum* (Taylor 1995: 66). Taylor (1995: 66) sieht Kategorisierung durch Schemata und Kategorisierung durch Prototypen als Aspekte des gleichen Phänomens an: „In the former case [schemas, E.S.], an entity happens to be fully compatible with an abstract representation, in the latter case [prototypes, E.S.], it is only partially compatible.“ Im Folgenden wird allein auf das Konzept der Prototypen näher eingegangen. Schemata spielen im Verlauf der Arbeit nur als abstrakte Form-Funktion-Paare (siehe Abschnitt 2.3) eine Rolle.

³²Lakoff (1987: 59–61) weist darauf hin, dass die Idee der Markiertheit von Formen als Prototypeneffekt gedacht werden kann. Dabei dient ein Vertreter oder eine Subkategorie als Referenzpunkt (z. B. Singular). Aufgrund des Status als Referenzpunkt gilt der Vertreter bzw. die Subkategorie als unmarkierter Default.

³³Für eine Analyse von Quantoren und Adjektiven als prototypische Kategorien siehe Hundt (2000).

2 Variation aus gebrauchsbasierter Perspektive

Dies ist anhand der schwachen und gemischt-deklinierten Maskulina leicht zu verdeutlichen: *-n* im Plural ist ein Merkmal der schwachen Deklination. Somit können Substantive mit gemischter Flexion als periphere Vertreter der schwachen Flexion modelliert werden, da sie nur im Plural schwache Flexionsmerkmale aufweisen (siehe hierzu ausführlich Abschnitt 3.2.2). Dieselbe Logik greift für Konjugationsklassen: Hier kann bspw. *mahlen* als Verb gesehen werden, dessen Flexionsverhalten nur peripher der starken Konjugation zuzurechnen ist, da *mahlen* mit dem Ablaut im Perfekt (*gemahlen*) nur ein starkes Konjugationsmerkmal aufweist, aber ansonsten schwache Flexionsmerkmale wie bspw. ein Präteritum mit Dentalsuffix (*mahlte*) (Bittner 1996: 80, Nowak 2013: 153–156; siehe hierzu ausführlich Abschnitt 3.1.2). Auch in Bezug auf Flexionsklassen lässt sich somit eine graduelle Zugehörigkeit beobachten.

Die Idee der graduellen Zugehörigkeit zu einer Kategorie steht im Kontrast zum aristotelischen Kategorienbegriff, bei dem notwendige und hinreichende Bedingungen die Zugehörigkeit zu einer Kategorie klar definieren. Nach dieser Logik kann eine Entität entweder Vertreter einer Kategorie sein oder nicht: Ein Adjektiv muss bspw. steigerbar sein. Eine ungerade Zahl muss eine Zahl und ungerade sein. Die einzelnen notwendigen Bedingungen sind dabei hinreichend dafür, die Kategorie zu definieren (Kleiber 1993: 11–13; ausführlich zum aristotelischen Kategorienbegriff siehe Kleiber 1993: 11–28 sowie Taylor 1995: 21–29). Prototypisch aufgebaute Kategorien sind dagegen flexibel und bieten daher einen Vorteil in der Kognition: Man kann ein Wort als Adjektiv bezeichnen, das attributiv genutzt werden kann, obwohl es nicht steigerbar ist. Die Kategorie selbst muss deshalb nicht infrage gestellt werden (Taylor 1995: 53–54). Prototypisch organisierte Kategorien können somit einerseits dem Wunsch nach klaren Kriterien nachkommen, da der Prototyp viele Eigenschaften der betreffenden Kategorie auf sich vereint, andererseits erlauben sie Flexibilität, da auch Entitäten zur Kategorie zählen können, die nur einige Eigenschaften mit dem Prototyp teilen (Taylor 1995: 54).

Aus der Gradualität prototypisch organisierter Kategorien ergibt sich zum einen die Unschärfe prototypisch organisierter Kategorien, da sie nicht erfassen, ob eine Entität ein Vertreter der Kategorie ist, sondern inwiefern die Entität ein Vertreter der Kategorie ist (Kleiber 1993: 106, Schmid 2000: 33). Zum anderen ist Typizität eine Konsequenz aus der Gradualität. Die Typizität gibt an, wie typisch ein Vertreter für eine Kategorie ist (Schmid 2000: 33). Der Prototyp ist dabei immer der Vertreter mit der höchsten Typizität. Typizität und Unschärfe sind zwar ähnliche Konzepte, da beide aus der Gradualität der prototypisch organisierten Kategorien erwachsen, dennoch meinen die Ausdrücke nicht dasselbe: Ein Vertreter einer Kategorie kann untypisch sein, dieser Kategorie aber dennoch klar zugeordnet werden. So sind Pinguine untypische Vögel, deren Status als Vogel

dadurch jedoch nicht infrage gestellt wird (Taylor 2015: 563). Niedrige Typizität und Unschärfe einer Kategorie können allerdings korrelieren, da bei untypischen Vertretern einer Kategorie unklar sein kann, ob man sie noch der Kategorie zurechnen sollte. Aus der Unschärfe prototypisch organisierter Kategorien ergibt sich auch, dass Vertreter zwei Kategorien angehören können: So kann ein rundes Gefäß aus Porzellan ohne Henkel, das etwas größer ist als eine Tasse, ein Becher oder eine (kleine) Vase sein.³⁴

Die Gradualität von Kategorien wird durch Familienähnlichkeit gestützt. Das Konzept der Familienähnlichkeit nach Wittgenstein (2004 [1953/1968]) geht davon aus, dass zwei Vertreter einer Kategorie nicht dieselben Eigenschaften teilen müssen, solange sie Eigenschaften mit einem anderen Vertreter der Kategorie teilen (Taylor 1995: 38–40). Palmen und Syntaxbäume teilen keine Eigenschaften miteinander, aber beide teilen Eigenschaften mit einer Eiche: Die Palme hat einen Stamm und Syntaxbäume verästeln sich wie eine Eiche. Somit sind beide über Familienähnlichkeit verbunden. Die Familienähnlichkeit der Kategorien ergibt sich Bybee (2010: 90–91) zufolge aus der Kategoriererweiterung mittels Analogie: Aus lokalen Analogien werden Analogieketten.³⁵ Ähnlich funktioniert das Konzept der Abdeckung von Goldberg, das in Abschnitt 2.1.1 diskutiert wurde. Die Analogieketten basieren auf Prototypen, die als kognitiver Bezugspunkt dienen (Kleiber 1993: 38–39, Goldberg 2006: 89, Ellis 2016: 242). Prototypische Vertreter sind aufgrund ihrer Tokenfrequenz und Kontextdiversität mental stark gefestigt.³⁶ Sie eignen sich daher als Analogievorlagen, die ein abstraktes Muster ermöglichen (Bybee 2010: 90–91).

Die grundsätzliche Ablehnung kategorialer Grenzen in der Standardversion der Prototypensemantik kann nicht vollständig aufrecht erhalten werden. Anhand der Familienähnlichkeit kann man bspw. die Behauptung aufstellen, dass Mäuse eine Art von Vögeln sind: Mäuse teilen sich Eigenschaften mit Fledermäusen, denn beide sind klein und haben Fell. Fledermäuse teilen sich wiederum eine Eigenschaft mit Vögeln, da sie fliegen können. Mäuse und Fledermäuse werden aber dennoch nicht unter die Kategorie VOGEL gefasst. Somit muss es trotz Familienähnlichkeit eine Grenze geben, an der eine Kategorie endet. Dieses Problem

³⁴Der Kontext kann hier desambiguieren (siehe die Erläuterungen unten zum Experiment von Labov 2004 [1973]). Fehlt der Kontext jedoch, kann auf das Objekt sowohl mit *Becher* als auch mit *Vase* referiert werden.

³⁵Wenn einzelne Exemplare einer Konstruktion, die durch Analogieketten verbunden sind, eine breite Semantik haben, führt dies zu einer hohen Schematizität der Konstruktion (siehe Abschnitt 2.3.2 zu Schemata in der Konstruktionsgrammatik), sodass neue Mitglieder leicht aufgenommen werden können.

³⁶Neben dem hohen Grad an *entrenchment* trägt auch die Salienz prototypischer Vertreter dazu bei, dass sie leicht aktivierbar sind und somit als Analogievorlagen genutzt werden können (Ellis 2016: 242).

kann leicht gelöst werden, indem man Prototypizität selbst als prototypisch organisiert betrachtet (Kleiber 1993: 109–110, Taylor 1995: 65): Prototypische prototypisch organisierte Kategorien gehen somit ineinander über wie bspw. Tassen und Schüsseln (Labov 2004 [1973]), bei peripheren prototypisch organisierten Kategorien ist dies nicht der Fall, wie bspw. bei Vögeln und Mäusen (Lakoff 1987: 56).

Genauso ist die grundsätzliche Ablehnung notwendiger Bedingungen in Frage zu stellen, denn es ist eine notwendige Bedingung für die Kategorie KATZE, Teil der Kategorie TIER zu sein (Kleiber 1993: 89–91). Dies ist jedoch nicht problematisch, wenn man notwendige, aber keine hinreichenden Bedingungen ansetzt (Kleiber 1993: 90–91): Die notwendigen Bedingungen führen also nicht zwangsläufig dazu, dass eine Entität zu einer Kategorie gehört.³⁷ Jackendoff (2004: 128) geht davon aus, dass für die Kategorisierung sowohl notwendige Bedingungen (eine Tasse muss ein Gefäß sein) als auch graduelle (die Höhe/Breite einer Tasse) sowie typische (Tassen sind aus Porzellan) eine Rolle spielen. Die graduellen Bedingungen haben dabei einen zentralen Punkt, um den sie sich gruppieren.³⁸ Die typischen Bedingungen einer Kategorie können im Gegensatz zu den notwendigen Bedingungen auch Ausnahmen erlauben. Sie führen dazu, dass Kategorien nach Familienähnlichkeit gestaffelt sind (Jackendoff 2004: 128). Für einzelne Kategorien stehen jeweils unterschiedliche Bedingungen im Mittelpunkt: Bei der Kategorie ROT sind dies bspw. die graduellen Bedingungen, bei der Kategorie TASSE dagegen die typischen (Jackendoff 2004: 128).

Ein weiteres Problem der Standardversion der Prototypensemantik ist die viel diskutierte Frage, was einen Prototyp zum Prototyp macht (Bärenfänger 2002: 13). Definiert man den Prototyp in Hinblick auf die Typizität,³⁹ ist der Prototyp

³⁷Allerdings existieren Kategorien, bei denen dies der Fall ist, wie bspw. ungeraden Zahlen: Hier sind zwei notwendige Bedingungen hinreichend für die Zugehörigkeit zur Kategorie.

³⁸Aufgrund dieser Eigenschaft nennt Jackendoff (2004: 128) die graduellen Bedingungen Zentralitätsbedingungen (*centrality conditions*).

³⁹Neben diesem Ansatz werden Vertrautheit, Frequenz und das Erwerbsalter als Einflussfaktoren diskutiert (Kleiber 1993: 41–43; 51–54, Taylor 1995: 52–53, Rosch 2004 [1978]: 100): Hierbei wäre immer das Exemplar mit dem höchsten Grad an Vertrautheit, der höchsten Frequenz und dem geringsten Erwerbsalter der Prototyp. Offensichtlich korrelieren die Faktoren, da Vertrautheit von Kleiber (1993: 42) definiert wird als „das Exemplar [...], das den Sprechern am vertrautesten ist; d. h. dem sie am häufigsten begegnen (können)“, hier spielt also auch Frequenz eine Rolle. Vor allem die Gleichsetzung des frequentesten Vertreters mit dem prototypischen Vertreter wird kritisch diskutiert, dabei wird davor gewarnt, Prototypizität mit Frequenz gleichzusetzen (siehe Kleiber 1993: 42, Taylor 1995: 52, Schmid 2000: 50 und Taylor 2015: 567–568). Aufgrund dieses kritischen Umgangs der Standardversion der Prototypensemantik mit Frequenz bezeichnet Fenk-Oczlon (1991: 365) die Prototypensemantik als frequenzfeindlich. Generell wirft die Modellierung von Vertrautheit, Frequenz und Erwerbsalter als singuläre Einflussfaktoren auf Prototypizität Probleme auf, weswegen hier nur die Typizität als Einflussfaktor ausführlich

der Prototyp, da er die typischen Eigenschaften einer Kategorie auf sich vereint (Kleiber 1993: 42). Diese Definition ist auf den ersten Blick wenig hilfreich, da sich nun die Frage stellt, was typische Eigenschaften sind (Kleiber 1993: 42). Zudem kann damit zirkulär argumentiert werden: Der Prototyp ist prototypisch, weil er die meisten typischen Eigenschaften aufweist, und die typischen Eigenschaften sind typisch, weil der Prototyp sie aufweist.

Der Fokus auf Typizität kann prototypisch organisierte Kategorien dennoch präzisieren, da er zu einer Begriffsverschiebung führt (Kleiber 1993: 43): Es wird nicht mehr das beste Exemplar gesucht, sondern die Eigenschaften, die als typisch für eine Kategorie gelten. Der Prototyp wird somit zu einer „aus typischen Attributen zusammengesetzten Entität“ (Kleiber 1993: 43) und ist mit dieser Definition auch nicht mehr mit einer Entität gleichzusetzen, sondern „a schematic representation of the conceptual core of a category“ (Taylor 1995: 59). Diese Sichtweise favorisieren sowohl Kleiber (1993: 43) als auch Taylor (1995: 59–60) gegenüber der Definition des Prototyps als bestes Exemplar: Selbst wenn man den Prototyp als bestes Exemplar sieht, muss die Repräsentation des Prototyps ein Stück weit abstrakt sein, sodass der Prototyp in verschiedenen Situationen erkannt werden kann (Taylor 1995: 60).

Als Antwort auf die Frage, welche Eigenschaften typisch sind, lässt sich der Einflussfaktor Frequenz⁴⁰ anführen: Hierbei ist von einer Kombination aus Typen- und Tokenfrequenz auszugehen. So stellen Ellis u. a. (2014: 85) fest, dass der prototypische Vertreter einer Konstruktion die höchste Frequenz (Tokenfrequenz) aufweist und die meisten Gemeinsamkeiten (Typenfrequenz) mit anderen Mitgliedern der Kategorie auf sich vereint. Der prototypische Vertreter wird daher schneller prozessiert, denn „like Rome, all ways lead to it“ (Ellis u. a. 2014: 85). Auch Goldberg (2006: 88) und Bybee (2010: 79–80) sehen den prototypischen Vertreter einer Kategorie als den häufigsten Vertreter an.⁴¹ So geht Goldberg (2006: 88) davon aus, dass Konstruktionen sich um ein oder zwei Wörter zentrieren

vorgestellt wird. Definiert man Prototypen über Typizität, können die singulären Einflussfaktoren jedoch dennoch eine Rolle spielen, da sie Einfluss auf die Typizität nehmen. Dies ist insbesondere für die Frequenz der Fall, wie aus den folgenden Erläuterungen deutlich werden wird.

⁴⁰Wie in Fußnote 39 bereits angesprochen, stehen Vertreter_innen der Standardversion der Prototypensemantik Frequenz skeptisch gegenüber, da befürchtet wird, dass Frequenz und Prototyp gleichgesetzt werden. Modelliert man Frequenz als Einflussfaktor auf Typizität, verschiebt sich die Argumentation jedoch dahingehend, dass typische Eigenschaften einer Kategorie frequent sind. Diese Modellierung ist bereits nah an dem exemplarbasierten Ansatz des Kategorienlernens, der in Abschnitt 2.2.3 vorgestellt wird.

⁴¹Diese Annahme ist aus Sicht einer rein prototypbasierten Kategorisierung problematisch, da Frequenz und Prototypizität eng verquickt werden (siehe hierzu Kleiber 1993: 42, Taylor 1995: 52, Schmid 2000: 50 und Taylor 2015: 567–568). Sie ist aber mit exemplarbasiertem Lernen vereinbar (siehe Abschnitt 2.2.3).

(bspw. um *put* als Prototyp der caused-motion-Konstruktion, siehe Fußnote 2 in Abschnitt 2.1), die sie als semantischen Prototyp bezeichnet.

Die frequenten Vertreter einer Kategorie bedingen, dass auch die Eigenschaften der Vertreter frequent vorkommen. Die frequent in einer Kategorie vorkommenden Eigenschaften sind dann auch die typischen Eigenschaften der Kategorie (Kleiber 1993: 52). Prototypizität wird somit durch tokenfrequente Vertreter und durch die Tokenfrequenz der Eigenschaften der frequenten Vertreter gestützt. Daher kommt zusätzlich die Typenfrequenz ins Spiel: Für die Kategorisierung ist nicht nur eine einzige Eigenschaft der frequenten Vertreter relevant, sondern mehrere. Prototypische Vertreter werden daher doppelt durch Frequenz gestützt: Tokenfrequenz stützt die Vertreter und ihre Eigenschaften. Zusätzlich stützt Typenfrequenz prototypische Vertreter dadurch, dass sie viele Eigenschaften der Kategorie auf sich vereinen. Daneben ist auch die Typenfrequenz der Kategorie relevant: Erst wenn sich viele Vertreter bestimmte Eigenschaften teilen, kann daraus eine Kategorie mit prototypischer Struktur abstrahiert werden.

Frequenz alleine ist jedoch zu ungenau, um das Entstehen einer Kategorie zu modellieren, denn nur weil eine Eigenschaft in einer Kategorie häufig ist, bedeutet dies noch nicht, dass diese Eigenschaft die Kategorie gut von einer anderen unterscheidet: [+belebt] kommt sowohl in der Kategorie VOGEL als auch in der Kategorie FISCH häufig vor. Das Konzept der *cue validity*, das in Abschnitt 2.1.2 kurz diskutiert wurde, berücksichtigt dies, indem *cue validity* die Frequenz einer Eigenschaft in einer Kategorie mit der Frequenz einer Eigenschaft in anderen Kategorien vergleicht (Rosch & Mervis 1975: 575, Kleiber 1993: 52–53). *Cue validity* ist somit ein Maß für die Distinktivität einer Eigenschaft: [+Flügel] hat eine höhere *cue validity* als [+belebt], da (fast) nur Vögel Flügel haben. Die Eigenschaft [+Kiemen] hat eine noch höhere *cue validity*, da diese Eigenschaft nur Fische⁴² aufweisen (Lakoff 1987: 52–53). Eine hohe *cue validity* kann dazu führen, dass eine Eigenschaft als salient wahrgenommen wird, da sie nur in einer bestimmten Kategorie erscheint.⁴³ Typenfrequenz und *cue validity* lassen sich dabei kom-

⁴²Interessanterweise existiert in der Biologie zwar die Kategorie FISCH, diese stellt aber anders als die Kategorie VÖGEL kein Taxon dar, da Fische keinen gemeinsamen Ursprung haben (Helfman 2009: 3).

⁴³*Cue validity* kann außerdem dazu genutzt werden, die Distinktivität einer Kategorie zu evaluieren. Kategorien auf der Basisebene (z. B. Stuhl) haben generell eine höhere *cue validity* als Kategorien, die der Basisebene übergeordnet (z. B. Möbel) oder untergeordnet (z. B. Klappstuhl) sind (Rosch, Mervis u. a. 1976: 428). Für weitere Merkmale der Basisebene und der übergeordneten Ebene siehe Lakoff (1987: 51–52). Wie prototypische Vertreter einer Kategorie verfügt die Basisebene über die höchste *cue validity*: Die Basiskategorie hat Vertreter, die viele gemeinsame Attribute aufweisen und gleichzeitig die wenigsten Attribute mit Vertretern anderer Kategorien gemein haben (Kleiber 1993: 64). Die Basisebene ist somit distinktiver als die anderen Ebenen. Deshalb wird die Basisebene auch als primäre Kategorie gesehen, auf deren

binieren: Prototypische Vertreter einer Kategorie teilen die meisten Eigenschaften mit anderen Vertretern dieser Kategorie, haben aber umgekehrt nur wenige Eigenschaften mit Vertretern anderer Kategorien gemein. Anders ausgedrückt maximieren prototypische Vertreter ihre *cue validity*, indem sie die meisten Eigenschaften mit der größten *cue validity* auf sich vereinen (Taylor 2015: 564).

Auf Basis der oben dargelegten Überlegungen zur Typizität wurde die Prototypensemantik revidiert (Kleiber 1993: 111–114). Während die Standardversion der Prototypensemantik davon ausgeht, dass der Prototyp eine Kategorie konstituiert und strukturiert, lehnt die erweiterte Version dies ab und geht davon aus, dass es sich beim Prototyp um einen Effekt handelt (Kleiber 1993: 111–114): Der Prototyp ist also kein Exemplar, sondern durch typische Eigenschaften ergeben sich Prototypeneffekte, wie bspw. eine schnellere Prozessierung (weitere Effekte werden in Abschnitt 2.2.2 vorgestellt). Eine Kategorie wird in dieser Perspektive durch die Anzahl der typischen Eigenschaften und durch die *cue validity* dieser Eigenschaften bestimmt (Kleiber 1993: 53). Zudem geht die erweiterte Version der Prototypensemantik davon aus, dass auch prototypisch organisierte Kategorien prototypisch gegliedert sind, es also mehr oder weniger prototypisch organisierte Kategorien gibt. Somit wird auch die Annahme der generell unscharfen Grenzen aufgegeben: Prototypische prototypisch organisierte Kategorien haben unscharfe Grenzen (wie bspw. die Kategorien ADJEKTIV und ADVERB), aber unprototypische prototypisch organisierte Kategorien weisen scharfe Grenzen auf (wie bspw. die Kategorie VOGEL). Mit der Standardversion hat die erweiterte Version weiterhin Prototypeneffekte und Familienähnlichkeit als grundlegende Prinzipien gemeinsam. Anders als in der Standardversion sind die Prototypeneffekte jedoch eine Folge der Kategorienstruktur, die nach dem Prinzip der Familienähnlichkeit organisiert ist. Der Prototyp ist also eine Art Epiphänomen und nicht die Ursache dieser Effekte (Lakoff 1987: 56, Kleiber 1993: 113, Blutner 1995: 254–256). Dies betont auch Rosch (2004 [1978]: 101) selbst: „To speak of a prototype at all is simply a convenient grammatical fiction; what is really referred to are judgments of degree of prototypicality.“

Abbildung 2.4 fasst die Unterschiede zwischen Standard- und erweiterter Version der Prototypensemantik zusammen.

Basis die anderen Kategorien abgeleitet werden (Lakoff 1987: 13). Die Distinktivität einer Kategorie hängt eng mit ihrer Lernbarkeit zusammen (Ellis 2002a: 153). Für empirische Evidenz siehe Taraban & Roark (1996), die die Distinktivität von Substantiven mit bestimmten Genera manipulieren. Je distinktiver die Eigenschaften von Substantiven mit einem bestimmten Genus von den Eigenschaften der Substantive mit einem anderen Genus sind, desto leichter ist die Kategorie zu erlernen, unabhängig davon, wie ähnlich sich die Substantive eines Genus untereinander sind.

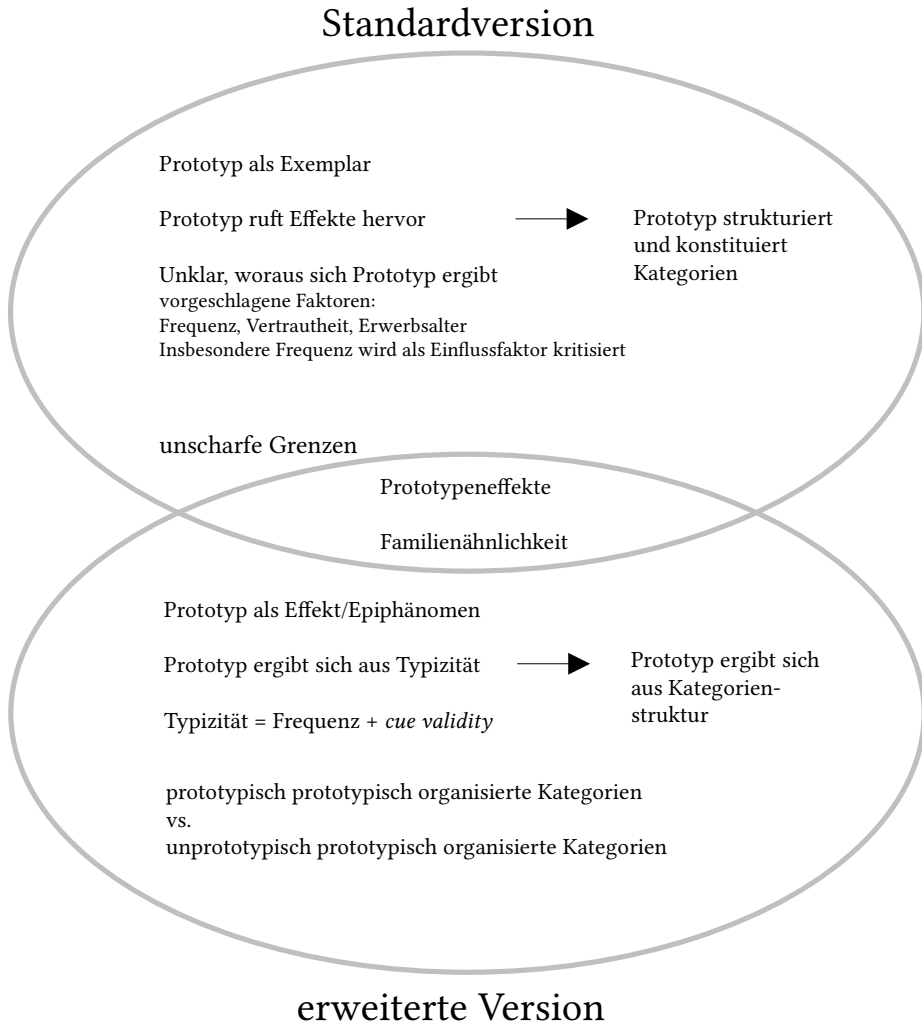


Abbildung 2.4: Vergleich zwischen Standardversion und erweiterter Version der Prototypensemantik

Im weiteren Verlauf der Arbeit wird Prototypizität im Sinne der erweiterten Version der Prototypensemantik relevant sein: Dabei ist im Kern wichtig, dass Kategorien graduell organisiert sind. Einige Mitglieder einer Kategorie vereinigen viele typische Merkmale der Kategorie auf sich und stellen daher prototypische Vertreter der Kategorie dar. Andere Mitglieder vereinigen nur wenige Eigenschaften der Kategorie auf sich und sind daher periphere Mitglieder der Kategorie. Dabei lässt sich für prototypische Vertreter ein Prozessierungsvorteil ausmachen. Dieser Effekt wird im folgenden Abschnitt näher erläutert. Dabei wird herausgearbeitet, inwiefern Prototypizität Einfluss auf Variation nimmt: Variation ist bei prototypischen Vertretern unwahrscheinlich, da diese mental gefestigt sind, in der Peripherie ist sie dagegen aufgrund der geringeren Festigung wahrscheinlich. Diese Hypothese wird im empirischen Teil der Arbeit psycholinguistisch überprüft. Um den Einfluss der Prototypizität auf die Prozessierung genauer modellieren zu können, werden im folgenden Abschnitt die empirische Evidenz und Erklärungen für Prototypeneffekte näher betrachtet.

2.2.2 Prototypeneffekte

Empirische Evidenz für die Existenz von Prototypeneffekten liefern zahlreiche Experimente. So werden prototypische Vertreter einer Kategorie am häufigsten genannt, wenn man Proband_innen dazu auffordert, einen Vertreter einer Kategorie zu nennen (bspw. *Nenne einen Vogel!*) (Kleiber 1993: 38–39, Taylor 1995: 45). Proband_innen sind in der Evaluation von An-X-is-an-Y-Aussagen (bspw. *Eine Amsel ist ein Vogel*) schneller, wenn ein prototypischer Vertreter einer Kategorie genannt wird (bspw. *Amsel* für die Kategorie *VOGEL*), als wenn es sich um einen peripheren Vertreter (bspw. *Pinguin*) handelt (Rosch 2004 [1978]: 100). Zudem existiert empirische Evidenz dafür, dass Kinder prototypische Vertreter einer Kategorie früher erlernen als randständige (Rosch 2004 [1978]: 100). Prototypische Vertreter einer Kategorie sind zudem salienter als randständige Vertreter (Taylor 2015: 566).

Wie Lakoff (1973) zeigt, lässt sich der Status von Entitäten innerhalb einer Kategorie als prototypisch oder peripher durch Heckenausdrücke verbalisieren, die ausdrücken, zu welchem Grad eine Entität Teil einer Kategorie ist (*This is a typical cup/This is basically a cup*). Zudem nimmt Prototypizität Einfluss auf die Wortabfolge im Satz (Kelly & Bock 1986). Prototypische Vertreter werden vor peripheren genannt: Proband_innen tauschen in Testsätzen wie *The child's errand was to buy a lemon and apple at the fruit stand* die Wortfolge, sodass der prototypische Vertreter *apple* vor dem peripheren Vertreter *lemon* genannt wird. Wenn

2 Variation aus gebrauchsbasierter Perspektive

hingegen der prototypische Vertreter zuerst genannt wird, behalten Proband_innen die Reihenfolge bei (Kelly & Bock 1986: 66–67). Zudem werden die Sätze als natürlicher bewertet, in denen der prototypische Vertreter zuerst genannt wird (Kelly & Bock 1986: 68).

Rosch (1975b: 197–199) bestimmt mithilfe des Goodness-of-example-Designs,⁴⁴ welche Vertreter als prototypisch wahrgenommen werden und welche als peripher: Dabei werden amerikanischen Student_innen verschiedene Kategorien (u. a. FRÜCHTE, VÖGEL und GEMÜSE) präsentiert und zusätzlich mögliche Vertreter⁴⁵ der Kategorie. Die Proband_innen werden gebeten, auf einer Skala von 1 (gutes Beispiel) bis 7 (schlechtes Beispiel) zu bewerten, wie sehr ein Vertreter in eine Kategorie gehört. Auf diese Weise möchte Rosch intersubjektive Prototypen ausmachen. In neun von zehn Kategorien wurde das Exemplar, das durchschnittlich als bester Vertreter gewertet wurde, von 95 % der Proband_innen mit der Bestzahl 1 bewertet (Rosch 1975b: 198). Die intersubjektiven Prototypen nutzt Rosch (1975b: 199–206), um zu überprüfen, ob prototypische Vertreter einer Kategorie anders wahrgenommen werden als andere Vertreter. So kann sie mithilfe einer same-different-match-Aufgabe⁴⁶ zeigen, dass prototypische Vertreter generell schneller kategorisiert werden und dass Kategoriennamen den prototypischen Vertreter der Kategorie primen, nicht aber periphere Vertreter.

Nicht nur für lexikalische, sondern auch für grammatische Kategorien lassen sich Prototypeneffekte beobachten: Elemente mit prototypischen Eigenschaften einer Kategorie werden stabil genutzt. Die stabile Verwendung des Prototyps spiegelt sich im frühen Erwerb: Als erstes Exemplar wird das prototypische und

⁴⁴Schmid (2000: 34–37) kritisiert das Versuchsdesign, da bereits durch die Aufgabe präsupponiert wird, dass typische und weniger typische Vertreter einer Kategorie existieren. Zudem unterstellt Rosch (1975b: 198) in der Aufgabenstellung, dass es eine intersubjektive Vorstellung davon gibt, was der prototypische Vertreter einer Kategorie ist: „Think of dogs. You all have some notion of what a ‘real dog’, a ‘doggy dog’ is. To me a retriever or a German shepherd is a very doggy dog while a Pekinese is a less doggy dog“ (Rosch 1975b: 198). Außerdem kritisiert Poitou (2004a: 74), dass in Roschs Arbeiten implizit auf die denotative Bedeutung fokussiert wird. Dies wird in einer Anweisung deutlich, in der konnotative Komponenten explizit ausgeschlossen werden: „But try not to just free associate – for example, if bicycles just happen to remind you of your father, don’t write down father“ (Rosch & Mervis 1975: 578).

⁴⁵Diese wurden anhand einer Studie von Battig & Montague (1969) ausgewählt: In dieser Studie sollen Proband_innen Vertreter von Kategorien benennen. Rosch (1975b: 197–199) übernahm pro Kategorie die Vertreter, die von über zehn Versuchspersonen genannt worden waren; zusätzlich wurden zufällig weitere Vertreter ausgesucht, die von weniger als zehn Personen genannt wurden.

⁴⁶In einer same-different-match-Aufgabe werden Proband_innen gebeten, zu bewerten, ob zwei Wörter oder Bilder der gleichen Kategorie angehören oder nicht, z. B. *Apfel* und *Banane* oder *Apfel* und *Stuhl*. Zusätzlich zu dieser Aufgabe wird bei einigen der Paaren mit einem Prime gearbeitet, um zu testen, inwiefern der Kategoriename Vertreter der Kategorie primt. Ausführlich zum methodischen Vorgehen und zu den Ergebnissen siehe Rosch (1975b).

frequenteste Verb in einer Konstruktion (bspw. *give* in der ditransitiven Konstruktion) erworben (Ellis & Ferreira-Junior 2009, Ellis u. a. 2014: 58). In der Peripherie grammatischer Kategorien sind hingegen Schwankungen möglich (Ágel 2008: 67–68). Diese lassen sich bspw. in der Peripherie von Form-Schemata beobachten, siehe hierzu Abschnitt 2.3.3. Anders als prototypische Elemente weisen periphere Elemente weniger gemeinsame Eigenschaften mit anderen Vertretern der Kategorie auf und sind somit auch weniger stark mit der Kategorie assoziiert.

Schwankungen in der Peripherie entstehen zum einen dadurch, dass periphere Vertreter den prototypischen Eigenschaften der Kategorie angepasst werden; so wird bspw. das ursprünglich nicht-flektierbare Farbadjektiv *lila* flektiert (*eine lilane Hose*) (Ágel 2008: 67–68). Zum anderen entstehen Schwankungen, weil die Peripherie zum Anziehungsbereich von zwei Kategorien zählt. Dies ist bspw. bei *brauchen* der Fall, das in der Peripherie zwischen Voll- und Modalverb steht und daher zwischen Infinitivformen mit und ohne *zu* schwankt (*Das brauchst du nicht (zu) machen*) (Eisenberg & Voigt 1990: 11). Dasselbe gilt für die Präposition *wegen*: *Wegen* ist eine Sekundärpräposition, die jedoch Eigenschaften von Primärpräpositionen angenommen hat (Schmitt 2019b: 100–101, Vieregge 2019: 74). Der periphere Status zwischen zwei Kategorien zeigt sich unter anderem darin, dass *wegen* sowohl mit Genitiv (als typischen Kasus für Sekundärpräpositionen) als auch mit Dativ (als typischen Kasus für Primärpräpositionen) stehen kann. Die beiden Kasus evozieren vergleichbare Lesezeiten (Schmitt 2019b: 115–120), wohingegen *wegen* + Nominativ zu hohen Lesezeiten führt: *Wegen* + Nominativ wurde in der Studie als ungrammatische Kontrollkondition genutzt, da der Nominativ als Präpositionalkasus ausgeschlossen ist.⁴⁷

Modelliert man die Wahrscheinlichkeit der Kasus bei *wegen* in einem bayesischen Modell, ist die priore Wahrscheinlichkeit für Genitiv und Dativ aufgrund des peripheren Status von *wegen* zwischen Primär- und Sekundärpräposition vergleichbar groß, während sie für den Nominativ gering ist, da der Nominativ prototypischerweise zur Markierung des Subjekts verwendet wird. Hier ist das statistische Vorkaufsrecht nach Goldberg (2019: 74–94) relevant: Der Nominativ kommt häufig vor, um Subjekte zu markieren, wird aber nie als Rektionskasus genutzt. Genitiv und Dativ werden hingegen häufig als Rektionskasus genutzt, weshalb sie den Nominativ statistisch ausstechen (ausführlich zum statistischen Vorkaufsrecht siehe Abschnitt 2.1.3). Da *wegen* sowohl Eigenschaften von Sekundär- als auch von Primärpräpositionen aufweist und sich somit jeweils in der Peripherie der Präpositionsarten befindet, kann bei *wegen* keiner der Rektionskasus den anderen ausstechen, sodass sowohl Genitiv als auch Dativ möglich sind.

⁴⁷Hierbei ist zu beachten, dass Präpositionen ohne Kasusforderung wie *außer* durchaus Nominativformen aufweisen können, die sich aber nicht aus der Rektion der Präposition, sondern aufgrund von Verbkongruenz ergeben: *Alle essen Koriander außer ich* (Di Meola 2000: 46–47).

2 Variation aus gebrauchsbasierter Perspektive

Taylor (1998: 198) stellt fest, dass periphere Vertreter einer Konstruktion eher idiomatische Einschränkungen aufweisen als prototypische Vertreter einer Konstruktion. Konstruktionen sind in der Peripherie somit weniger produktiv, wobei periphere Verwendungen einer Konstruktion nicht zwangsläufig weniger akzeptabel sind als prototypische Verwendungen (Taylor 1998: 198). So führt Tomasello (1998: xvii) einen großen Anteil der kreativen Verwendung von Sprache darauf zurück, dass Wörter in einer Konstruktion genutzt werden, die keinen prototypischen Verwendungskontext für das Wort darstellt. Als Beispiel nennt er *Mary kicked John the football*, da *to kick* normalerweise nicht für Besitzwechsel genutzt wird und daher auch nicht in der ditransitiven Konstruktion. In der vorliegenden Verwendung kann aber ein Besitzwechsel konstruiert werden, sodass eine Verwendung in der Konstruktion möglich ist (Tomasello 1998: xvii).⁴⁸

Die eingeschränkte Produktivität in der Peripherie von Konstruktionen zeigt sich in P600-Effekten bei EEG-Studien. So können Sätze P600-Effekte auslösen, in denen das Subjekt kein prototypisches Agens ist (Gouvea u. a. 2010: 157), wie bspw. *the hearty meal was devouring the kids*, in dem mit dem nicht-belebten Substantiv *meal* ein untypisches Agens verwendet wird (zur prototypischen Organisation der Agensrolle siehe Lakoff 1987: 64–66). Auch dieser Effekt kann bayesisch modelliert werden: Das Substantiv *meal* ist als Agens weniger semantisch erwartbar als bspw. ein Substantiv mit belebtem Referenten, sodass die priore Wahrscheinlichkeit für *meal* als Agens vergleichsweise gering ist.

Die Integration der empirisch nachweisbaren Prototypeneffekte in die Theorie prototypisch organisierter Kategorien stellt eine Herausforderung dar. Einen Versuch legt Rosch (1975b: 198) vor, die für ihre Modellierung lexikalischer Kategorien das Ordnungsprinzip der Typizität nutzt, indem sie anhand der Bewertungen aus dem oben erwähnten goodness-of-example-Experiment Typizitätsskalen erstellt, die die Distanz von einzelnen Vertretern einer Kategorie zum Prototyp angeben: Ein Rotkehlchen hat den Wert 1,02 und stellt somit den prototypischen Vertreter der Kategorie VOGEL dar, während ein peripherer Vertreter wie bspw. ein Pinguin einen Wert von 4,53 aufweist.⁴⁹ Poitou (2004b: 7–9) kritisiert Typizitätsskalen, da die Typizität bei verschiedenen Kategorien auf unterschiedliche

⁴⁸Die Interpretation als Übergabe ergibt sich durch die Verwendung des Verbs in der ditransitiven Konstruktion, sodass die Konstruktion selbst Bedeutung trägt (Tomasello 1998: xviii; in Abschnitt 2.3.2 werden Konstruktionen ausführlich diskutiert). Konstruktionen können somit Elemente in ihre Slots zwingen (Taylor 1998: 198, Goldberg 2019: 37), daher diskutieren Audring & Booi (2016) Effekte dieser Art unter dem Stichwort *coercion* (Zwang).

⁴⁹Lakoff (1973: 458–459) setzt diese unterschiedlichen Werte der Typizität mit Zugehörigkeit gleich: Pinguine sind für ihn also weniger vogelig als Rotkehlchen. Diese Gleichsetzung wird in der erweiterten Version der Prototypensemantik aufgegeben (Kleiber 1993: 112).

Eigenschaften zurückzuführen ist. So sind in der Kategorie VOGEL die Vögel prototypisch, die man frequenterweise antrifft. Umgekehrt ist in der Kategorie EDELSTEIN der Stein prototypisch, der am seltensten ist. Poitou (2004b: 9) nennt als Beispiel einen lupenreinen Smaragd, der prototypischer ist als ein Smaragd mit Einschlüssen. Allerdings ist davon auszugehen, dass lupenreine Edelsteine im Diskurs einen größeren Stellenwert einnehmen als fehlerhafte Edelsteine und daher darin auch häufiger verhandelt werden. Ein anderer Kritikpunkt Poitous an Typizitätsskalen ist zentraler: Es ist problematisch, dass Typizitätsskalen Kategorien auf einer einzigen Skala abbilden wollen (Poitou 2004b: 1–3). Wenn Proband_innen aufgefordert werden, Mitglieder einer Kategorie zu nennen, beeinflusst nicht nur die Typizität die Reihenfolge der Nennungen, sondern bspw. auch die Lautstruktur der vorher genannten Vertreter einer Kategorie. So kann es sein, dass auf *Nektarine* aufgrund des ähnlichen Auslauts *Apfelsine* und *Clementine* genannt werden: Die Substantive bilden ein lautliches Cluster, das aktiviert wird, wenn ein Vertreter des Clusters genutzt wird (Goldberg 2019: 62–65). Typizitätsskalen stellen somit eine Vereinfachung der Kategorienstruktur dar und können nur einen ersten Einblick in die Struktur von Kategorien bieten (Poitou 2004b: 2–3).

Die von Poitou (2004b) angesprochene Problematik zeigt sich auch darin, dass die durch das goodness-of-example-Design erarbeiteten Typizitätsskalen nicht universell sind: So werden bei Rosch (1975b: 232) Blauhäher (*blue jay*) und Hüttensänger (*bluebird*) mit Werten von 1,29 und 1,31 als sehr typische Vertreter der Kategorie VOGEL identifiziert. Diese Vögel sind in Nordamerika beheimatet; es ist somit davon auszugehen, dass das Experiment andere Ergebnisse hervorbringt, wenn es mit Proband_innen durchgeführt wird, die in anderen Erdteilen sozialisiert wurden. Zudem sind Experimente dieser Art kontextabhängig: Wenn der Kontext spezifiziert wird (etwa Vögel auf dem Bauernhof), werden andere Prototypen genannt als ohne spezifischen Kontext (Schmid 2000: 35–36).

Der Einfluss des Kontexts auf die Kategorisierung wird auch in einem der bekanntesten Experimente zur prototypischen Struktur lexikalischer Kategorien von Labov (2004 [1973]) deutlich. Labov (2004 [1973]) zeigt Proband_innen Zeichnungen von tassenähnlichen Objekten und bittet sie, die darin dargestellten Objekte zu benennen. Die Objekte variieren u. a. in Breite und Höhe (für eine genauere Beschreibung des methodischen Vorgehens siehe Labov 2004 [1973]: 76–78). Je breiter das dargestellte Objekt ist, desto eher benennen die Proband_innen dieses als *bowl* statt als *cup*; je länglicher ein Objekt ist, desto wahrscheinlicher benennen die Proband_innen das Objekt zunächst als *mug* statt als *cup* und schließlich als *vase* (Labov 2004 [1973]: 78–79). Zusätzlich wird der Kontext manipuliert, in dem die Objekte bewertet werden sollen: Im neutralen Kontext

2 Variation aus gebrauchsbasierter Perspektive

wird lediglich nach dem Objekt gefragt, im Essen-Kontext sollen die Proband_innen sich vorstellen, dass das Objekt mit Kartoffelbrei gefüllt ist, und im Blumen-Kontext, dass das Objekt Blumen enthält und auf einem Fensterbrett steht. Der Kontext beeinflusst die Benennung der Proband_innen: Im Vergleich zum neutralen Kontext sagen die Proband_innen im Essen-Kontext bereits bei schmalen Objekten, dass es sich dabei um eine Schüssel und nicht um eine Tasse handelt (Labov 2004 [1973]: 78–79). Im Blumen-Kontext benennen die Proband_innen Objekte als *vase*, die im neutralen Kontext *cup* genannt werden (Labov 2004 [1973]: 79–80). Zudem benennen die Proband_innen das Objekt in diesem Kontext unabhängig von seiner Gestalt nicht als *mug*.

Das Experiment verdeutlicht, dass für die Zuordnung zu einer Kategorie nicht nur die Form eines Gegenstands relevant ist, sondern auch dessen Zweck. Zur Tasse gehört bspw. der Zweck, Kaffee oder Tee daraus zu trinken; eine Suppe daraus zu essen, ist hingegen eine randständige Funktion. Die hier skizzierten Interaktionen zwischen Kontext, Semantik und Phonologie können durch Typizitätsskalen nicht modelliert werden (Poitou 2004b: 2–3). Somit ist ein bloßer Blick auf Typizität nicht ausreichend, um Kategorien zu umfassen. Um einen Kategorien umfassend modellieren zu können, ist ein exemplarbasierter Ansatz nötig. Dieser Ansatz geht davon aus, dass alle Eigenschaften (inklusive dem Äußerungskontext) eines Exemplars mental repräsentiert werden und Eigenschaften einer Kategorie erst durch die Überlappung von Eigenschaften einzelner Exemplare der Kategorie entstehen (Goldberg 2019: 62–73).

Der Abschnitt zeigt, dass prototypische Vertreter einer Kategorie anders verarbeitet werden als periphere. Dies wird für die Betrachtung der Variationsphänomene essentiell sein: Prototypizität kann damit als Erklärung für Variation genutzt werden. Es ist davon auszugehen, dass prototypische Vertreter einer Kategorie wenig Variation aufweisen, während periphere zugänglich für Variation sind (Ágel 2008: 67–68). Dieser Einfluss wird für die einzelnen Variationsphänomene näher in den Blick genommen und im empirischen Teil der Arbeit psycholinguistisch überprüft. Der Einfluss von Prototypizität kann genauso wie der Einfluss von Frequenz bayesianisch modelliert und somit über Wahrscheinlichkeiten gefasst werden. Gleichzeitig zeigen sich die Grenzen von prototypisch organisierten Kategorien: Kategorien sind vielfältig gestaltet, sodass sie sich nicht anhand von einer einzigen Dimension modellieren lassen. Daher wird im folgenden Abschnitt der exemplarbasierte Ansatz zur Kategorienentstehung in den Blick genommen und mit prototypbasierten Kategorien verglichen. Das Ergebnis der Diskussion ist eine Modellierung von prototypisch organisierten Kategorien als Ergebnis exemplarbasierter Lernens. Diese Definition wird dem Rest der Arbeit zugrunde gelegt.

2.2.3 Prototyp- und exemplarbasierte Kategorien

Im gebrauchsbasierten Ansatz wird statt der prototypbasierten Kategorisierung eine exemplarbasierte Kategorisierung als grundlegendes Konzept angesetzt (Bybee 2013: 68). Die exemplarbasierte Kategorisierung bildet einen Gegensatz zur Standardversion der Prototypensemantik, in der davon ausgegangen wird, dass eine Kategorie durch den Prototyp repräsentiert wird, der die zentralen Tendenzen der Kategorie auf sich vereint (siehe Abschnitt 2.2.1). Wie dieser Abschnitt zeigen wird, lässt sich die exemplarbasierte Kategorisierung jedoch gut mit der erweiterten Version der Prototypensemantik vereinen, die Vertreter nach ihrer Typizität unterscheidet und nicht von einem Prototyp als Exemplar ausgeht, sondern ihn als einen Effekt sieht. So wurden in Abschnitt 2.2.1 bereits einige Grundprinzipien des exemplarbasierten Lernens in Hinblick auf die erweiterte Version der Prototypensemantik diskutiert.

In der Standardversion der Prototypensemantik findet Kategorisierung im Vergleich zum Prototyp statt (Ross & Makin 1999: 208) und der Prototyp wird daher als einzige zusammenfassende Repräsentation einer Kategorie konzeptualisiert. Dies ist jedoch eine zu einseitige Sicht auf Kategorisierung. Kategorisierung ist nicht allein vom Prototyp abhängig, sondern bspw. auch von der Ähnlichkeit der zu kategorisierenden Entität mit kurz zuvor erlernten Vertretern der Kategorie (Whittlesea 1987, Ross & Makin 1999: 208, Goldberg 2006: 46).

Das menschliche Erinnerungsvermögen basiert zudem nicht allein auf prototypischen Eigenschaften, sondern ist assoziativ (Goldberg 2019: 52). Dies zeigt bspw. der Vorteil für gleichbleibende Stimmen (Goldinger 1998: 255): Wörter werden nicht nur für sich mental repräsentiert, sondern mit dem Menschen assoziiert, der sie gesprochen hat (siehe Abschnitt 2.1.2 zum Vorteil gleichbleibender Stimmen). Außerdem umfasst das menschliche Erinnerungsvermögen nicht nur einzelne Eigenschaften, sondern auch, wie viele Werte eine Eigenschaft hat: Erdbeeren sind rot, daher kann die Eigenschaft FARBE in Bezug auf Erdbeeren nur einen Wert haben. Äpfel können dagegen auch grün sein, die Eigenschaft FARBE hat hier also mindestens zwei Werte. Zudem verfügen wir über ein implizites Wissen darüber, wie die Werte verschiedener Eigenschaften untereinander korrelieren: Paprikaschoten können süß und bitter sein sowie rot, orange, gelb oder grün. Dabei korreliert die Farbe mit dem Grad der Bitterkeit (Ross & Makin 1999: 211–212). Diese Korrelation wird in der Standardversion der Prototypensemantik nicht erfasst. Die erweiterte Version der Prototypensemantik könnte sie berücksichtigen, da der Blick auf typische Eigenschaften auch die Korrelation zwischen Eigenschaften ermöglicht. Sie beschränkt sich jedoch auf Typizitätsskalen, deren Problematik bereits in Abschnitt 2.2.2 diskutiert wurde.

2 Variation aus gebrauchsbasierter Perspektive

Im Gegensatz zur Standardversion der Prototypensemantik geht der exemplar-basierte Ansatz davon aus, dass eine Kategorie durch Token geformt wird, die als gleich angesehen werden (Bybee 2013: 53): Hört man eine Wortform, werden die Eigenschaften dieser Wortform (Form, Funktion, Kontext) mental repräsentiert.⁵⁰ Hört man die Wortform noch einmal, wird die neue Repräsentation ergänzt. Die neue Repräsentation überlappt mit der alten, aber sie ist nicht dieselbe; z. B. könnte die Wortform von einer anderen Person gesagt werden und daher leicht anders klingen als die zuerst gehörte Wortform. Die Repräsentationen der Wortformen ergeben schließlich ein emergentes Cluster, „which constitutes what we think of as a coherent word meaning (or lemma)“ (Goldberg 2019: 16). Im exemplarbasierten Ansatz hat somit jedes Exemplar (Token) Einfluss auf die mentale Repräsentation einer Kategorie (Type): Immer wenn eine mentale Repräsentation durch Kodierung oder Dekodierung aktiviert wird, ändert sich die Repräsentation (Bybee 2006: 722, 2013: 52, Goldberg 2019: 51–52). Abbildung 2.5, die an die Abbildung in Goldberg (2019: 16) angelehnt ist, veranschaulicht diesen Prozess.

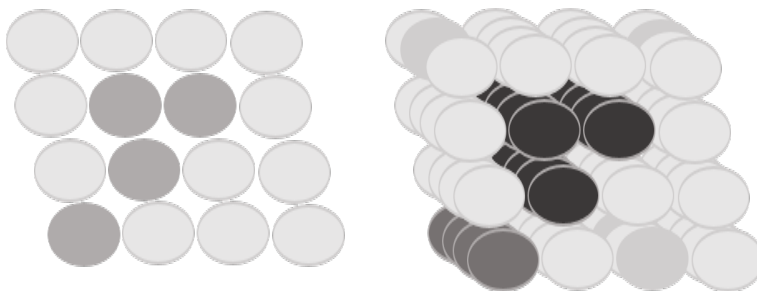


Abbildung 2.5: Stärkung der mentalen Repräsentation eines Wortes durch Wiederholung

Der linke Teil der Abbildung zeigt die mentale Repräsentation einer Wortform: Informationen, die als informativ⁵¹ für den Gebrauch der Wortform wahrgenom-

⁵⁰Dabei hinterlässt bereits die erste Verarbeitung eines Exemplars eine Erinnerungsspur, da ansonsten jede Verarbeitung als die erste gelten müsste (Bybee 2006: 722–723, 2013: 56, Goldberg 2019: 13–14, 54).

⁵¹Was dabei als informativ wahrgenommen wird, hängt von zahlreichen Einflussfaktoren innerhalb der Verwendungssituation ab: Wenn ein Elternteil bspw. *Ball* sagt und dabei auf einen Ball zeigt, wird durch die Zeigegeste klar, dass der Ball wichtig ist und nicht das daneben liegende Buch. Wenn das Elternteil *heiß* sagt und das Kind gleichzeitig das heiße Wasser spürt, kann der Temperaturwechsel als relevant empfunden und mit dem Wort verknüpft werden. Dabei können auch Dinge als relevant eingestuft werden, die nicht relevant sind, und so z.B. *heiß* mit Wasser anstelle der Temperatur verknüpft werden. Da eine Repräsentation aber nicht anhand von einer einzigen Gebrauchssituation gebildet werden, sondern anhand von vielen, werden falsche Verbindungen i.d.R. nicht verstärkt und somit im Lauf der Zeit als unwichtig empfunden.

men werden, sind dabei dunkel hervorgehoben (Goldberg 2019: 16). Die Repräsentation enthält detaillierte Informationen über Form, Bedeutung und Kontext. Wenn das Wort immer wieder gehört wird, werden die wiederkehrenden Aspekte in der Repräsentation gestärkt, wohingegen singuläre Aspekte nicht gestärkt werden. Hieraus ergibt sich ein Cluster, das im rechten Teil der Abbildung dargestellt ist. Nach demselben Prinzip werden Exemplare, die ähnlich sind, miteinander assoziiert. Sie bilden ein Cluster (eine Kategorie), das es ermöglicht, gemeinsame Informationen der einzelnen Exemplare zu abstrahieren und strukturiert zu repräsentieren (Bybee 2013: 54).

Im exemplarbasierten Ansatz funktioniert Kategorisierung auf Basis wiederkehrender Ähnlichkeit der zu klassifizierenden Entität zu den bestehenden Exemplaren einer Kategorie (Ross & Makin 1999: 212–213, Goldberg 2019: 51–52): Um bspw. als Frucht eingestuft werden zu können, muss eine Entität wiederkehrend Ähnlichkeit zu bereits existierenden Exemplaren der Kategorie FRUCHT aufweisen. Die Kategorisierung wird dabei vorrangig von den Vertretern einer Kategorie mit der größten Ähnlichkeit zueinander beeinflusst. Das Kriterium der Ähnlichkeit zwischen Exemplaren legt bereits an, dass sich die Struktur von Kategorien im exemplarbasierten Ansatz – genau wie im prototypbasierten Ansatz – durch Familienähnlichkeit ergibt (Bybee & Eddington 2006: 327). Zudem beeinflusst Variabilität die Kategorisierung: Die Variabilität der bereits bestehenden Exemplare einer Kategorie nimmt Einfluss darauf, wie groß die Ähnlichkeit des zu kategorisierenden Exemplars zu den bereits existierenden sein muss (Goldberg 2019: 62–64, siehe Abschnitt 2.1.1 für weitere Ausführungen zu Goldbergs Konzept der Abdeckung, das die Einflussfaktoren Typenfrequenz, Ähnlichkeit und Variabilität beinhaltet).

Aufgrund der Familienähnlichkeit sind auch im exemplarbasierten Ansatz Prototypeneffekte möglich: Einige Vertreter der Kategorie sind typischer als andere, weil typische Exemplare mehr Ähnlichkeiten zu anderen Exemplaren der Kategorie aufweisen als untypische (Ross & Makin 1999: 213). Da eine Amsel also mehr Ähnlichkeiten zu anderen Vögeln aufweist, ist sie typisch; der Pinguin ist dagegen untypisch, weil er weniger Ähnlichkeiten zu anderen Vögeln aufweist (Ross & Makin 1999: 213). Dies ist auch die exemplarbasierte Erklärung dafür, dass typische Vertreter einer Kategorie schneller prozessiert werden als untypische.

Für die Kategorisierung werden im exemplarbasierten Ansatz aber nicht nur semantische Ähnlichkeiten einbezogen, sondern bspw. auch kontextuelle und phonologische Ähnlichkeiten (Bybee 2006: 716). Goldberg (2019: 15) verdeutlicht die Notwendigkeit eines so umfassenden Ansatzes anhand eines Beispiels: Beim Verb *schreiben* ist es egal, ob hastig geschrieben wird oder welche Wörter geschrieben werden. Dennoch ist davon auszugehen, dass diese Informationen in

der mentalen Repräsentation einzelner Verwendungen des Verbs *schreiben* berücksichtigt werden. Für *schreiben* sind die Informationen zwar irrelevant, aber für Hyponyme von *schreiben* essentiell: So ist die Art des Schreibens bzw. das geschriebene Wort Teil der Bedeutung von *krakeln* („hässlich schreiben“) und *unterschreiben* („den eigenen Namen schreiben“). Zudem kann das Verhältnis einzelner Informationen zueinander nur dann bestimmt werden, wenn diese in die mentale Repräsentation des Vertreters eingeflossen sind (Ross & Makin 1999: 214–215): Die Farbe einer Paprika und ihr Bitterkeitsgrad können nur verknüpft werden, wenn die Informationen auch mental repräsentiert sind. Wir verfügen dementsprechend implizit über Korrelationen zwischen Eigenschaften von Vertretern einer Kategorie. Dies kann die Kategorisierung beeinflussen, auch wenn sich Sprecher_innen der Korrelation nicht bewusst sind (Ross & Makin 1999: 214–215).

Die Modellierung von Kategorien als Cluster verschiedener mentaler Repräsentationen ist ein Vorteil des exemplarbasierten Ansatzes. Durch die Annahme von mehrfachen Repräsentationen kann der selektive Zugang zu Wissen über Kategorien berücksichtigt werden. Wenn wir eine Kategorie aufrufen, werden nicht alle, sondern nur die im gegebenen Kontext relevanten Eigenschaften der Kategorie aktiviert. Dies wird auch im obigen Beispiel zu *schreiben*, *unterschreiben* und *krakeln* deutlich, denn auch für das Verb *schreiben* können Kontexte existieren, in denen relevant ist, wie geschrieben wird. Unser Wissen über Kategorien ist daher sensitiv für Inhalte (Goldberg 2019: 52). Zudem werden im exemplarbasierten Ansatz Eigenschaften nicht einfach unabhängig voneinander aufsummiert, sondern relational genutzt (Ross & Makin 1999: 216). Dies ermöglicht es, dass Informationen über die Vertreter der Kategorie und ihr Verhältnis zueinander abgerufen werden können, wenn es nötig ist.

Wie bereits in Abschnitt 2.2.1 erläutert, wird Tokenfrequenz als Einflussfaktor auf Prototypizität von Vertreter_innen der Standardversion der Prototypensemantik kritisch betrachtet (siehe Kleiber 1993: 42, Taylor 1995: 52, Schmid 2000: 50 und Taylor 2015: 567–568), weswegen der prototypbasierte Ansatz als frequenzfeindlich bezeichnet wird (Fenk-Oczlon 1991: 365). Diese Frequenzfeindlichkeit wird erst durch den Fokus auf typische Eigenschaften relativiert. Im exemplarbasierten Ansatz ist Frequenz ein grundlegender Einflussfaktor auf Kategorisierung: Kategorien werden als analogische Sets ähnlicher Exemplare mental repräsentiert, weil die Exemplare häufig genutzt werden (Bybee & Eddington 2006: 326, Diessel 2017: 13). Informationen über Kategorien können daher verloren gehen, wenn diese nicht durch Wiederholung gestärkt werden (Bybee 2013: 54).

Exemplare mit hoher Tokenfrequenz sind im exemplarbasierten Ansatz die zentralen Einheiten einer Kategorie (Bybee 2013: 61): Sie formen den semantischen Slot innerhalb einer Konstruktion (zu Konstruktionen siehe Abschnitt 2.3.2). Zudem sind typenfrequente und produktive Kategorien um einen frequenten Vertreter gruppiert. Evidenz für diese Annahme sieht Bybee (2006: 727) in Kategorien mit geringer Typenfrequenz und geringer Produktivität, die keinen hochfrequenten Vertreter haben. Für diese These spricht außerdem, dass frequente und prototypische Verben in Verb-Argument-Konstruktionen als Erstes erworben werden (Ellis 2013: 5).

Zudem konnte in einem Experiment, das mit Mustern aus Punkten und somit nicht sprachbasiert arbeitet, der prototypische Vertreter besser als solcher erkannt werden, wenn er häufiger als andere Muster gezeigt wurde (Goldberg 2006: 87–88). Daher ist davon auszugehen, dass das Lernen von Kategorien durch einen Vertreter mit prototypischen Eigenschaften erleichtert wird, der häufig im Input erscheint (Goldberg 2006: 87). Tokenfrequenz scheint die Produktivität einer Kategorie somit zumindest mittelbar zu beeinflussen, da sie bei der Entstehung einer Kategorie eine Analogievorlage ermöglicht, die zur Erweiterung und zur Abstraktion der Kategorie führt (ausführlich zum Einfluss der Tokenfrequenz auf Produktivität siehe Abschnitt 2.1.1).⁵² Zudem scheinen frequente Vertreter einer Kategorie und Vertreter mit semantischer Ähnlichkeit zum frequenten Vertreter höhere Akzeptanzwerte aufzuweisen als infrequente und dem frequenten Vertreter unähnliche Exemplare (Bybee 2006: 727–728, Bybee & Eddington 2006: 325–327). Vertreter mit semantischer Ähnlichkeit zum frequenten Vertreter sind somit akzeptabler als Vertreter ohne semantische Ähnlichkeit. Dies verdeutlicht, dass die Ausweitung einer zu erlernenden Kategorie nicht allein an tokenfrequenten Vertretern hängt: Auch die Variabilität der Kategorie und die Ähnlichkeit des potentiellen neuen Vertreters zu bereits existierenden Exemplaren spielen eine Rolle. Goldberg (2019: 62–65) fasst dies unter dem Konzept der Abdeckung, siehe hierzu Abschnitt 2.1.1.

Die Einflussfaktoren Ähnlichkeit und Frequenz ermöglichen Hierarchien in der Kategorisierung (Bybee 2006: 716): Exemplarbündel mit ähnlicher phonologischer Form und ähnlicher Bedeutung werden als Wort oder Phrase geclustert, z. B. das Verb *schicken* oder die Phrase *Ich schenke dir ein Buch*. Je nachdem, wie variabel das Cluster ist, können neue Exemplare als dem Cluster angehörig wahrgenommen werden (Goldberg 2019: 62). Hieraus kann eine abstraktere

⁵²Die Annahme eines frequenten Mitglieds als Zentrum einer Kategorie scheint der Verbindung zwischen Tokenfrequenz und Autonomie zu widersprechen (siehe Abschnitt 2.1.1). Die Autonomie beschränkt Bybee (2006: 715) allerdings auf extrem frequente Elemente. Eine klare Grenze zwischen hoch- und extrem frequenten Exemplaren nennt sie dabei nicht.

2 Variation aus gebrauchsbasierter Perspektive

Konstruktion entstehen, wenn wiederum einzelne Wörter oder Phrasen bspw. mit ähnlicher Bedeutung als Cluster repräsentiert werden, z. B. Verben mit der Bedeutung ‚Übergabe‘. Daraus kann dann mit der ditransitiven Konstruktion ein noch abstrakteres Cluster abstrahiert werden. Die Abstraktheit der Konstruktion ergibt sich, weil nur die geteilten Eigenschaften einzelner Exemplare verstärkt werden. In der Folge werden auch nur bestimmte Aspekte der Exemplare als Teil der Konstruktion gesehen (Bybee 2013: 65). Für Verben mit der Bedeutung ‚Übergabe‘ sind die Teilbedeutungen ‚dauerhaft‘ und ‚umsonst‘ von *schchenken* bspw. irrelevant.

Mit diesem Ansatz lässt sich das Entstehen neuer Konstruktionen aus bestehenden Konstruktionen gut modellieren: Ein Vertreter einer Konstruktion wird mit einer für die Konstruktion besonderen Bedeutung genutzt. Wird derselbe oder ein ähnlicher Vertreter der Konstruktion wiederholt mit dieser Bedeutung verwendet, formt sich daraus ein Cluster, sodass sich eine neue Konstruktion entwickelt (Bybee 2013: 63). Die Entstehung neuer Konstruktionen ist somit vornehmlich durch Frequenz bedingt: Wenn ein Vertreter einer Konstruktion frequenter verwendet wird als andere, kann der Vertreter als eine eigene Einheit wahrgenommen werden und sich somit von der Ursprungs konstruktion lösen (Bybee 2006: 720, Goldberg 2019: 68–70; das Phänomen wurde in Abschnitt 2.1.1 als Autonomie-Effekt diskutiert). In der Folge weist der Vertreter weniger Ähnlichkeit zur Ursprungs konstruktion auf und bildet ein eigenes Cluster. Bybee (2006: 719–721) nennt *going to* als Beispiel, das zu *gonna* reduziert wird und so seine Verbindung zum Verb *go* verliert. Zudem geht die semantische Verbindung zum Vollverb *go* verloren, da keine Bewegung mehr ausgedrückt wird, sondern eine Intention bzw. Futur.

Wie bereits durch die Einführung des exemplarbasierten Ansatzes deutlich wurde, stellen sich die oben skizzierten Probleme der prototypbasierten Ansätze für den exemplarbasierten Ansatz nicht: Da er von Exemplaren ausgeht, ist eine Klassifizierung von potentiellen neuen Exemplaren aufgrund der Ähnlichkeit zu einem bestimmten, in der Kategorie vorhandenen Exemplar abgedeckt (Ross & Makin 1999: 214). Das bereits vorhandene Exemplar muss dabei nicht unbedingt prototypische Eigenschaften der Kategorie aufweisen. Auf diese Weise können auch marginale Eigenschaften als Vorlage für Erweiterungen der Kategorie dienen und spezifische Informationen über einzelne Kategorievertreter Teil der Repräsentation werden (Bybee 2006: 717). Kategorisierung hängt in diesem Ansatz von individuellen Erinnerungen ab (Diessel 2017: 13) und nicht wie beim prototypbasierten Lernen von Generalisierungen. Der exemplarbasierte Ansatz kann damit experimentelle Daten besser erklären als ein streng ausgelegter prototypischer Ansatz (Ross & Makin 1999: 215).

Problematisch für den exemplarbasierten Ansatz ist, dass er nicht erklären kann, warum bestimmte Einheiten als Vertreter einer Kategorie angesehen werden (Goldberg 2006: 47). Anders als der prototypbasierte Ansatz geht der exemplarbasierte Ansatz zudem davon aus, dass Abstraktionen (wie bspw. *Vögel haben Flügel*) nicht für die Kategorisierung genutzt werden, was die Frage offen lässt, warum Menschen dann die Fähigkeit haben, Eigenschaften zu abstrahieren (Ross & Makin 1999: 215). Goldberg (2006: 46) weist jedoch darauf hin, dass auch Exemplare eine Abstraktion beinhalten: Die Eigenschaften des Exemplars werden gewichtet, sodass die Eigenschaften, die in einem bestimmten Kontext relevanter sind als andere, mehr Beachtung finden. Eigenschaften, die nicht beachtet werden, werden nicht mental repräsentiert, sodass die Repräsentation immer abstrakter ist als das konkrete Exemplar. Zudem merkt Goldberg (2006: 46) an, dass Eigenschaften auch wieder vergessen werden können, sodass auch deshalb die Repräsentation des Exemplars immer abstrakter ist als das Exemplar selbst. Somit ist auch im exemplarbasierten Ansatz Abstraktion angelegt.

Abbildung 2.6 zeigt Gemeinsamkeiten und Unterschiede von exemplar- und prototypbasierten Kategorien. Es zeigt sich, dass die erweiterte Version der Prototypensemantik bereits einige Eigenschaften mit dem exemplarbasierten Ansatz teilt. Der exemplarbasierte Ansatz zeichnet sich im Gegensatz zu den prototypbasierten Ansätzen durch seine Mehrdimensionalität und den Fokus auf einzelne Exemplare aus.

Die hier diskutierten Ansätze können miteinander kombiniert werden, sodass die Vorteile beider Ansätze (Abstraktion und prototypische Staffelung der prototypischen Kategorien; Mehrdimensionalität der exemplarbasierten Kategorien) zum Tragen kommen: Eine Kategorie ist exemplarbasiert, wenn sie nur wenige Vertreter enthält. Wenn weitere Mitglieder hinzukommen, werden Abstraktionen über die Eigenschaften der Mitglieder geformt, die dann als abstrakte Repräsentation der Kategorie genutzt werden (siehe Ross & Makin 1999: 216–222 für empirische Evidenz zu kombinierten Ansätzen). Prototypisch organisierte Kategorien können somit als ein Resultat einer exemplarbasierten Kategorisierung gesehen werden. In diesem Ansatz ist der Prototyp ein Effekt (wie bereits in der erweiterten Version der Prototypensemantik in Abschnitt 2.2.1), der sich aus den Mitgliedern der Kategorie ergibt, die sich am ähnlichsten sind. Diese dienen als kognitiver Referenzpunkt, der eine Kategorisierung neuer Entitäten erlaubt (Diessel 2017: 13). Die Organisation der Kategorie ist aber nicht allein durch die prototypische Struktur beeinflusst, da bestimmte Eigenschaften je nach deren kontextueller Relevanz berücksichtigt werden können.

Betrachtet man prototypische Effekte auf diese Weise, können sie auch als ein Resultat von Frequenz betrachtet werden. Einerseits spielt Tokenfrequenz ei-

2 Variation aus gebrauchsbasierter Perspektive

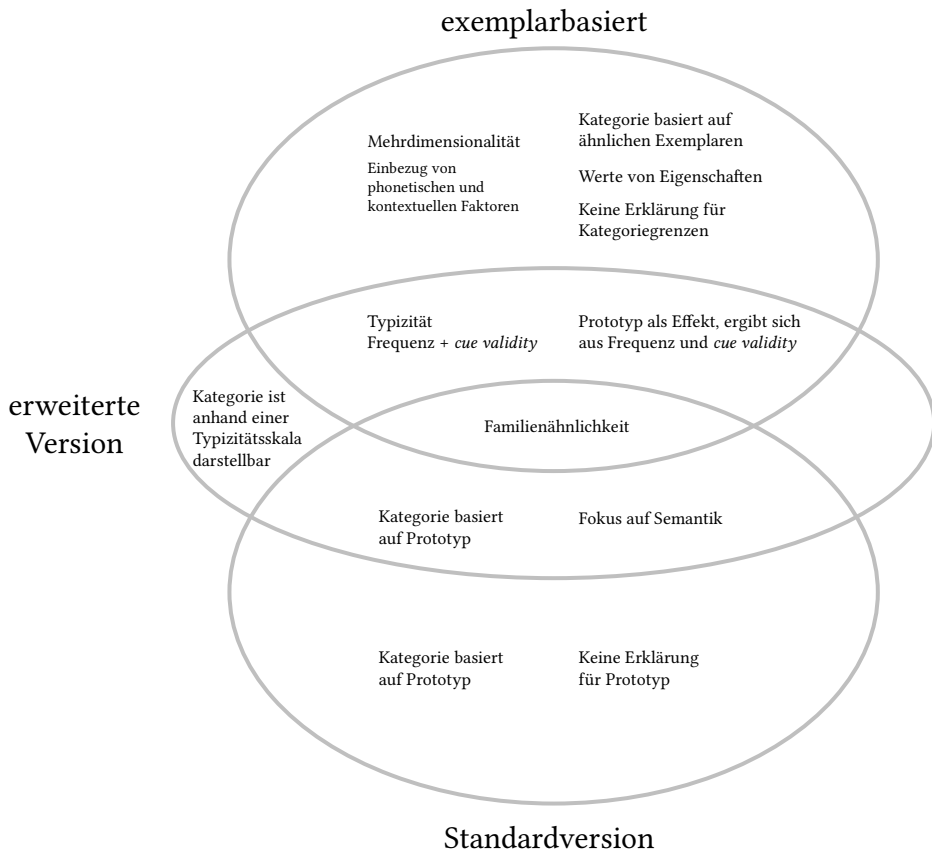


Abbildung 2.6: Vergleich zwischen exemplarbasierten und prototypbasierten Kategorien

ne Rolle, da häufig auftretende Vertreter die Kategorisierung beeinflussen: Die Frequenz, mit der ein Vertreter auftritt, und die Wahrscheinlichkeit, dass dieser Vertreter als Prototyp wahrgenommen wird, korrelieren stark (Goldberg 2006: 85). Andererseits spielt Typenfrequenz eine Rolle, da Vertreter mit vielen ähnlichen Eigenschaften die Kategorisierung beeinflussen.⁵³ Vertreter mit prototypischen Eigenschaften werden somit in zweifacher Weise mental gestärkt: Zum einen ist der Vertreter mit prototypischen Eigenschaften i. d. R. tokenfrequent, zum anderen teilt er viele Eigenschaften mit anderen Vertretern der Kategorie und ist daher qua seiner Eigenschaften typenfrequent (Ellis u. a. 2014: 85). Die

⁵³Den Einfluss modelliert Goldberg (2019: 51–73) unter dem Konzept der Abdeckung.

Kombination aus Token- und Typenfrequenz bildet Lernpräferenzen ab: Elio & Anderson (1984) zeigen in einem Experiment, dass Kategorien besser gelernt werden, wenn zunächst nur die frequenten, prototypischen Vertreter berücksichtigt werden und erst im Anschluss die Kategorie erweitert wird.

Der Einfluss der Frequenz wird ergänzt durch die Ähnlichkeit einzelner Exemplare zueinander. Die Ähnlichkeit zwischen den Exemplaren kann zu typischen Eigenschaften abstrahiert werden. Neben der Ähnlichkeit zwischen einzelnen Exemplaren ist die *cue validity* der Eigenschaften relevant: Je distinkter die ähnlichen Exemplare in ihren Eigenschaften gegenüber anderen Exemplaren sind, desto eher werden sie auch als gesonderte Gruppe angesehen. Bei der Kategorisierung spielen daher Typen- und Tokenfrequenz sowie die Ähnlichkeit der Vertreter zueinander und deren *cue validity* eine Rolle. Daher sind Eigenschaften prototypisch, die von vielen frequenten Vertretern der Kategorie geteilt werden und distinkt gegenüber Eigenschaften anderer Kategorien sind.

Im Folgenden werden prototypisch organisierte Kategorien als Resultat eines exemplarbasierten Lernens gesehen. Als prototypisch gelten dabei die Vertreter einer Kategorie, die häufig auftreten sowie die größte Ähnlichkeit zu anderen Vertretern der Kategorie aufweisen und somit eine hohe *cue validity* haben: Sie teilen die meisten distinktiven Eigenschaften mit weiteren Vertretern der Kategorie und die wenigsten distinktiven Eigenschaften mit Vertretern anderer Kategorien.

Prototypizität wird im weiteren Verlauf der Arbeit in zweifacher Weise relevant sein. Zum einen wird Prototypizität als Erklärungsansatz für Variation genutzt, denn Variation tritt vorrangig in der Peripherie einer Kategorie oder im Übergangsbereich zwischen zwei Kategorien auf (Ágel 2008: 67–68). So können die Funktionen zweier Konstruktionen in einem prototypisch organisierten Verhältnis zueinander stehen. Dies ist bspw. bei der Selektion von *haben* und *sein* der Fall: *Haben* ist mit transitiven und atelischen Sätzen ohne Bewegungsverbsemantik assoziiert, *sein* mit intransitiven Sätzen, die telisch sind und/oder Bewegungssemantik aufweisen (Gillmann 2016: 316–319). Im Übergangsbereich der Funktionen sind Schwankungen zwischen *haben* und *sein* zu beobachten, wie bspw. in Sätzen mit Bewegungsverb, in denen das Objekt als Patiens oder als Instrument interpretiert werden kann (*Ich bin/habe das Auto gefahren*) (siehe Abschnitt 3.3.2 zum Einfluss der Prototypizität auf die Auxiliarselektion für ausführliche Erläuterungen). Im empirischen Teil der Arbeit wird anhand der Selektion von *haben* und *sein* überprüft, ob sich Unterschiede in der Verarbeitung von prototypischen und peripheren Strukturen messen lassen. Zudem sind Form-Schemata prototypisch organisiert. So bildet die Form [Dreisilbigkeit, Pänultimabetonung, Schwandung, + menschlich] ein Form-Schema innerhalb der Maskulina (Köpcke 1995:

168–176). In der Peripherie des Form-Schemas sind Schwankungen möglich: Das Substantiv *Bär* schwankt bspw. zwischen schwachen und starken Formen (*des Bären/Bärs*), da das Substantiv nicht den prosodisch-phonologischen Eigenschaften des Form-Schemas entspricht (Köpcke 2005: 69). Form-Schemata und deren Prototypizität werden im folgenden Abschnitt 2.3 vorgestellt. Der Einfluss von Form-Schemata auf die Variation in der Konjugation und Deklination wird in den Abschnitten 3.1.3 und 3.2.3 diskutiert. Im empirischen Teil der Arbeit wird der Einfluss des Form-Schemas schwacher Maskulina auf die Prozessierung überprüft.

Zum anderen wird Prototypizität genutzt, um die Reihenfolge von Schwankungsfällen bei Flexionsklassenwechseln zu erklären und zu prognostizieren. Dies baut auf der Beobachtung auf, dass die Eigenschaften von Flexionsklassen prototypisch gestaffelt sind: Je mehr Eigenschaften der einen Flexionsklasse aufgegeben werden, desto mehr Eigenschaften der anderen Flexionsklasse kommen hinzu. Somit ist zu erwarten, dass Eigenschaften einer Flexionsklasse, die nur wenige Mitglieder aufweisen (wie bspw. Imperativhebung), zuerst aufgegeben werden, während Eigenschaften, die viele Mitglieder aufweisen (Ablaut im Perfekt) als letztes aufgegeben werden. Zudem lassen sich anhand dieses Prinzips Verben und Substantive danach klassifizieren, ob sie ein für ihre Flexionsklasse prototypisches oder unprototypisches Flexionsverhalten aufweisen. Die Prototypizität von Flexionsklassen wird in den Abschnitten 3.1.2 und 3.2.2 relevant sein, die die prototypisch organisierten Flexionseigenschaften starker und schwacher Verben sowie schwacher und starker Maskulina behandeln.

2.3 (Form-)Schematizität

Der Einfluss von (Form-)Schemata auf Variation ist bereits im letzten Abschnitt zu Prototypizität angeklungen. Im folgenden Abschnitt wird der Einfluss von (Form-)Schemata auf Variation ausführlich diskutiert. Hierfür wird zunächst der dieser Arbeit zugrundeliegende Schemabegriff definiert.

Das Konzept der Schemata wird in verschiedenen wissenschaftlichen Disziplinen wie etwa der Psychologie, Kognitionspsychologie und Linguistik (Bücker 2015: 447–450) genutzt. Dabei wird der Terminus aber nicht deckungsgleich verwendet. Als allgemeine Grundlage für den in dieser Arbeit relevanten Schemabegriff wird in Abschnitt 2.3.1 ein Ansatz aus der Kognitionspsychologie diskutiert, der Schemata als Wissensseinheiten auffasst. Darauf aufbauend wird in Abschnitt 2.3.2 der für diese Arbeit relevante Schemabegriff der Konstruktionsgrammatik vorgestellt und erweitert. Der Schemabegriff der Konstruktionsgrammatik

verknüpft Schemata und Konstruktionen, indem Schemata als Konstruktionen gesehen werden, die eine abstrakte und somit hochschematische Form aufweisen (Goldberg 2006). Neben Schemata in diesem Sinn werden die Termini *Form-Schematizität* und *Form-Schemata* neu eingeführt und von Schemata und Schematizität abgegrenzt. Unter *Form-Schematizität* werden Konstruktionen gefasst, bei denen eine Funktion mit zwei (oder mehr) Formen verknüpft ist (bspw. [+Vergangenheit] mit den Formen Dentalsuffix, [#_aŋk] wie in *trank* und [#_it] wie in *ritt*). Dabei ist eine der Formen hochschematisch (Dentalsuffix), während die anderen nicht-schematisch oder teilschematisch sind. Die Verbindung aus teilschematischer Form und Funktion wird dabei unter dem Terminus *Form-Schema* gefasst.

(Form)-Schemata bilden damit den dritten Faktor, der Einfluss auf Variation Einfluss nimmt. Dabei nimmt (Form-)Schematizität in verschiedener Weise Einfluss auf die hier diskutierten Variationsphänomene: In Hinblick auf *haben* und *sein* konkurrieren zwei Schemata miteinander, da *haben* und *sein* mit verschiedenen Funktionen verknüpft sind, die prototypisch organisiert sind. Bei der Variation in der Deklination und Konjugation ist hingegen eine Funktion mit verschiedenen Formen verknüpft. Diese Formen können teilschematisch sein und damit Form-Schemata bilden. Die Form-Schemata erhöhen die Wahrscheinlichkeit, dass die mit dem Form-Schema assoziierte Form genutzt wird. Der grundlegende Einfluss von Form-Schemata auf Variation wird in Abschnitt 2.3.3 näher diskutiert.

2.3.1 Grundlegender Schemabegriff

Innerhalb der Kognitionspsychologie wird mit einer sehr breiten Definition des Terminus *Schema* gearbeitet (Rumelhart 1980, Barlow & Kemmer 1994: 25–26): Rumelhart (1980: 33) sieht Schemata als „building blocks of cognition“, dabei geht er davon aus, dass Schemata Wissenseinheiten darstellen und in allen Wissensdomänen relevant sind. Rumelhart (1980) diskutiert beispielsweise das Schema EINKAUFEN, zu dem die Variablen WARE, KÄUFER_INNEN und VERKÄUFER_INNEN gehören. Eine ähnliche Auffassung von Schemata findet sich im Scenes and Frames-Modell nach Fillmore (1977) (Bücker 2015: 448–449). Schemata enthalten aber nicht nur Wissen an sich, sondern auch Anwendungsinformationen über dieses Wissen (Rumelhart 1980: 34). So gehört zum Schema EINKAUFEN das Wissen, dass dieses Schema bspw. im Supermarkt angewandt werden kann. Als grundlegende Eigenschaften von Schemata sieht Rumelhart (1980: 34–41) sechs Aspekte an:

2 Variation aus gebrauchsbasierter Perspektive

1. Schemata haben Variablen
2. Schemata betten ein
3. Schemata repräsentieren Wissen auf allen Ebenen
4. Schemata repräsentieren Wissen, keine Definitionen
5. Schemata sind aktive Prozesse
6. Schemata sind Erkennungsvorrichtungen, deren Prozessierung evaluiert, wie gut sie zu den prozessierten Daten passen (*goodness of fit*)
(Übersicht nach Barlow & Kemmer 1994: 26)

Die Variablen eines Schemas sind nach Rumelhart (1980: 35) die mit dem Schema assoziierten Entitäten. So lässt das Schema EINKAUFEN eine_n Verkäufer_in und eine Kasse erwarten. Schemata können dabei selbst wiederum Schemata enthalten (Rumelhart 1980: 37–40). So lässt sich etwa das Schema BEZAHLEN innerhalb des Schemas EINKAUFEN ausmachen (Rumelhart 1980: 39–40). Schemata betten also ein. Sie sind nicht auf eine bestimmte Wissensebene begrenzt, sondern strukturieren auf allen Ebenen. Zudem hebt Rumelhart (1980: 40) hervor, dass es sich bei Schemata um Wissen, nicht um Definitionen handelt. Schemata stellen aktive Prozesse dar, d. h. sie sind keine festgeschriebenen und unveränderlichen Muster, sondern existieren aufgrund ihrer Anwendung. Ein Schema ist in diesem Sinne „a procedure whose function it is to determine whether, and to what degree, it accounts for the pattern of observations“ (Rumelhart 1980: 39). Der Hinweis auf die Passungsgüte (*goodness of fit*), anhand derer Schemata evaluiert werden, ist ebenfalls grundlegend: Schemata basieren nicht auf einer distinkten Einteilung, sondern sind probabilistisch aufgebaut. Je ähnlicher eine Situation dem Schema ist, desto wahrscheinlicher ist die Aktivierung des Schemas (Rumelhart 1980: 39). Schemata lassen sich somit gut mit einem bayesianischen Modell vereinen (siehe Abschnitt 2.1.3).

Rumelhart (1980: 38) geht davon aus, dass sich Schemata verselbstständigen können, sodass wir in einer Situation nicht mehr wissen, ob unsere Annahmen über die Situation aufgrund unserer Sinneswahrnehmung getroffen werden oder ob sie auf das Schema zurückgehen, von dem wir ausgehen, dass es in dieser Situation greift.⁵⁴ So kann bspw. Bar (2004) zeigen, dass Proband_innen dasselbe Objekt je nach Bildkontext eindeutig als Föhn oder als Bohrer erkennen. Der

⁵⁴Für die Anwendung des Schemabegriffs nach Rumelhart auf sein in Abschnitt 2.1.3 diskutiertes konnektionistisches *parallel distributed processing model* siehe Rumelhart & McClelland (1986).

Einfluss von Schemata wird auch in falschen Erinnerungen deutlich: Wenn Proband_innen Wörter wie *nurse*, *sick*, *hospital* hören, erinnern sich einige fälschlicherweise daran, auch *doctor* gehört zu haben (Brainerd u. a. 2002, Goldberg 2019: 100). Die Wörter sind Teil eines Schemas (etwa KRANKHEIT). Die falsche Erinnerung deutet darauf hin, dass das Schema aufgerufen wurde und daher weitere Vertreter des Schemas aktiviert wurden.

Der kognitionspsychologische Zugriff auf Schemata kann für die grundlegende Ausrichtung des Schemabegriffs in dieser Arbeit fruchtbar gemacht werden. So werden die Variablen der (Form-)Schemata von starken Verben und schwachen Maskulina in den Abschnitten 3.1.3 und 3.2.3 erarbeitet. Die Einbettung von Schemata zeigt, dass es nicht ausreicht, nur von einem einzigen Schema auszugehen, sondern dass Schemata Teil eines Makro-Schemas sein können (Neset 2008: 14–21, Booij 2012: 344–345). Dieser Gedanke wird für Form-Schemata im Allgemeinen sowie für das Form-Schema schwacher Maskulina in Abschnitt 3.2.3 relevant sein. Da in der vorliegenden Arbeit sprachliche Variationsphänomene behandelt werden, ist lediglich der dritte Punkt (Schemata repräsentieren Wissen auf allen Ebenen) auf sprachliches Wissen einzuschränken (Barlow & Kemmer 1994: 26).

Die Idee der aktiven Prozesse deckt sich mit dem gebrauchsbasierten Ansatz, der dieser Arbeit zugrunde liegt: Schemata werden als Teil der allgemeinen kognitiven Fähigkeiten von Menschen aufgefasst, mithilfe derer Kategorien strukturiert werden. Der probabilistische Ansatz von Schemata verankert die Möglichkeit von Variation bereits in der Grundlage des Schemabegriffs: Schemata können einerseits Variation hemmen, wenn eine Struktur komplett zum Schema passt, aber andererseits auch zu Variation führen, wenn eine Struktur nur teilweise zum Schema passt oder Eigenschaften zweier Schemata aufweist. Hier zeigt sich, dass auch Prototypizität, die in Abschnitt 2.2 vorgestellt wurde, für Schemata relevant ist. Im folgenden Abschnitt wird der Schema-Begriff geschärft, indem Schemata im Sinne der Konstruktionsgrammatik vorgestellt werden und die Unterscheidung zwischen Schematizität und Form-Schematizität eingeführt wird. Diese Unterscheidung wird auch verdeutlichen, warum eine separate Betrachtung von Prototypizität und Schematizität notwendig ist, um Variation umfassend modellieren zu können.

2.3.2 Abgrenzung von Schematizität und Form-Schematizität

Bevor Schemata aus Sicht der Konstruktionsgrammatik vorgestellt und um Form-Schemata erweitert werden können, müssen zunächst Konstruktionen näher in

2 Variation aus gebrauchsbasierter Perspektive

den Blick genommen werden. Konstruktionen sind Form-Funktions-Paare (Lakoff 1987: 467). Mit dem sprachgebrauchsbasierten Ansatz ist anzunehmen, dass diese aus dem Gebrauch entstehen und daher erlernte Form-Funktions-Paare sind (Goldberg 2006: 5).⁵⁵ Dabei ist die Idee essentiell, dass Konstruktionen selbst Bedeutung tragen (Goldberg 1998: 203–205, Ellis u. a. 2014: 56–57, Goldberg 2019: 34–35). Wie die Lexeme einer Sprache ein Lexikon formen, formen Konstruktionen ein Konstruktikon (Goldberg 2019: 36).⁵⁶ Dass Konstruktionen Bedeutung tragen, kann man anhand von Pseudowörtern leicht nachvollziehen: Das Verb *kaschnitzen* existiert nicht, dennoch kann im Satz *Ich kaschnitze im Raum herum*

⁵⁵Lakoff (1987: 467) sowie Goldberg (1998: 205) knüpfen Konstruktionen an Nicht-Vorhersagbarkeit: Nur Form-Funktions-Paare, deren Form oder Funktion nicht vorhersagbar ist, sind Konstruktionen. *Esser* ist in dieser Definition keine Konstruktion, da sich Form und Funktion anhand der Konstruktion *ess* als Verbstamm und der Konstruktion *-er* als Derivationsuffix vorhersagen lassen. *Trinker* ist hingegen eine Konstruktion, da die Bedeutung ‚abhängig von Alkohol sein‘ sich nicht aus dem Verbstamm *trink* und dem Derivationsuffix *-er* ergibt (Stefanowitsch 2011: 185–186). Goldberg (2006: 5) grenzt sich hiervon ab: „[P]atterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency“. Diese Definition von Konstruktionen wird als Teil der gebrauchsbasierten kognitiven Konstruktionsgrammatik gesehen (Boas 2013: 17). Aus einer gebrauchsbasierten Perspektive ist davon auszugehen, dass auch ein Form-Funktions-Paar wie *Esser* mental repräsentiert wird. Dies zeigt sich bspw. daran, dass Kollokationen wie *ein guter Esser* existieren, in denen *Esser* nicht ausgetauscht werden kann (?*ein guter Essender*) (Stefanowitsch 2011: 185–186). Anhand der Tokenfrequenz wäre *Esser* somit als Konstruktion zu bezeichnen. Allerdings bleibt der Einfluss der Tokenfrequenz mit „sufficient frequency“ bei Goldberg (2006: 5) ähnlich wie bereits bei Bybee (2006: 715) in Bezug auf Konservierung und Autonomie unspezifiziert (siehe Abschnitt 2.1.1). Nicht nur Tokenfrequenz gibt einen Hinweis darauf, dass auch vorhersagbare Form-Funktions-Paare mental repräsentiert werden. Dies ist bereits im gebrauchsbasierten Ansatz angelegt: Um zu erkennen, dass ein Form-Funktions-Paar anhand der Bestandteile vorhersagbar ist, müssen zunächst mehrere mentale Repräsentationen existieren, von denen diese Erkenntnis abgeleitet werden kann. Um den konzeptionellen Unterschied zwischen vorhersagbaren und nicht-vorhersagbaren Form-Funktions-Paaren terminologisch fassen zu können, schlägt Stefanowitsch (2009: 569) vor, nur Form-Funktions-Paare mit nicht-vorhersagbarer Komponente als *Konstruktion* zu bezeichnen, vorhersagbare Form-Funktions-Paare hingegen als *Muster*. Da diese Unterscheidung in dieser Arbeit jedoch nicht vordergründig verhandelt wird, werden mit Goldberg (2006: 5) sowohl vorhersagbare als auch nicht-vorhersagbare Form-Funktions-Paare als Konstruktionen bezeichnet.

⁵⁶Dabei stellen Lexeme ebenfalls Konstruktionen dar, da sie Form-Funktions-Paare sind. Genau so wie Lexeme können auch komplexere Konstruktionen Priming-Effekte auslösen (Johnson & Goldberg 2013, Goldberg 2019: 33): Proband_innen sind in einer *lexical decision task* bei der Beurteilung von Verben wie *give* schneller, wenn ihnen zuvor eine semantisch kongruente Konstruktion mit Pseudowörtern (z. B. *He daxed her the norp*) präsentiert wurde. Die Reaktionszeiten sind dagegen verlangsamt, wenn zuvor eine semantisch inkongruente Konstruktion präsentiert wurde. Dies ist unabhängig von der Frequenz des Verbs der Fall (Johnson & Goldberg 2013: 1447–1451).

aufgrund der Konstruktion ($X_{NP_Nom} X_{VVFIN}$ im Raum herum) eine Bewegungssemantik erkannt werden. Die Semantik ergibt sich in diesem Fall also erst durch die Verwendung in der Konstruktion. Wie bereits in Abschnitt 2.2.2 deutlich wurde, können Konstruktionen auch real existierende Wörter in ihre Slots zwingen (Taylor 1998: 198, Goldberg 2019: 37), sodass ein eigentlich intransitives Verb transitiv genutzt werden kann (wie bspw. *niesen* in *Ich habe die Serviette vom Tisch geniest*). In dieser Hinsicht weisen Konstruktionen Ähnlichkeiten zu Schemata in der Kognitionspsychologie auf (siehe Abschnitt 2.3.1).

Konstruktionen und Schematizität sind miteinander verbunden, da die Bestandteile von Konstruktionen mehr oder weniger schematisch sein können: „On the low end of the schematic scale, positions in constructions can be completely fixed; higher schematicity is a function of the range of variation within the category“ (Bybee 2010: 80). Schematizität stellt dabei die substantielle Bedeutung einer Kategorie dar und kann sich auf semantische und phonologische Eigenschaften sowie „more holistic patterns“ beziehen (Bybee 2010: 80). Eine Redewendung wie *Da brat mir doch einer 'nen Storch* ist demnach nicht schematisch, da die einzelnen Bestandteile der Konstruktion fest definiert sind. Dagegen ist die ditransitive Konstruktion hochematisch, da sie sich aus variablen Bestandteilen zusammensetzt: *Ich gebe dir ein Buch, Er klaut der Katze die Milch* und *Ich schenke dir Essen* sind Instanzen der Konstruktion, deren Form sich mit $X_{NP_Nom} X_{VVFIN} X_{NP_Dat} X_{NP_Akk}$ formalisieren lässt. Hochematische Konstruktionen wie bspw. die ditransitive Konstruktion werden als Schemata bezeichnet (Booij 2010: 5, Bybee 2010: 80). Schemata sind somit eine Teilmenge von Konstruktionen.

Zwischen nicht- und hochematischen Konstruktionen lassen sich teilschematische Konstruktionen ansiedeln. Hierunter fallen Konstruktionen mit offenen Slots, die aber Restriktionen aufweisen wie bspw. *Das ist mir [egal, Wurst, Jacke wie Hose, schnurz]*. Die Produktivität eines Slots mit Restriktionen hängt dabei davon ab, wie ähnlich das potentielle neue Exemplar zu bestehenden Exemplaren im Slot ist (Blumenthal-Dramé 2012: 8, Diessel 2017: 17).⁵⁷ Dieses Prinzip zeigt sich bspw. beim *am*-Progressiv, der u. a. mit *activities* (*Ich bin am Arbeiten*), aber (derzeit) nicht mit *states* genutzt werden kann (**Ich bin am Wissen*) (Flick 2016: 181–184).

⁵⁷Diese Beobachtung lässt sich gut mit exemplarbasierten Kategorien verbinden: Bei Konstruktionen mit festen und offenen Slots enthält jedes Exemplar die Eigenschaften des festen Slots, dieser wird somit durch Tokenfrequenz gefestigt und hierdurch zum festen Slot (Bybee 2013: 59). Beim offenen Slot herrscht hingegen Variation. Wenn zwischen den Exemplaren, die in diesem offenen Slot auftreten können, bspw. semantische Ähnlichkeiten existieren, wird hieraus eine Kategorie geformt (Bybee 2006: 718). In der Folge hat auch der offene Slot semantische Restriktionen.

2 Variation aus gebrauchsbasierter Perspektive

Die Schematizität von Konstruktionen entsteht durch Gebrauch: Durch die häufige Verwendung werden die gemeinsamen Eigenschaften von Elementen in der Konstruktion gestärkt, als Cluster gebündelt und abstrahiert, weshalb ein Schema entsteht (Kemmer & Barlow 2000: xxiii, Langacker 2014: 218–220). Kemmer & Barlow (2000: xxiii) fassen Schemata daher als „cognitive representation[s] comprising a generalization over perceived similarities among instances of usage“. Durch die Generalisierung sind hochschematische Konstruktionen produktiv. Teilschematischen Konstruktionen kann hingegen nur bedingt Produktivität zugesprochen werden, da potentielle Neuzugänge der teilschematischen Form entsprechen müssen und daher restringiert sind (Bybee 2010: 67–69).

Bybee (2013: 62) schlägt vor, Schematizität und Produktivität als separate Einflussfaktoren zu betrachten, da Schematizität und Produktivität zwar miteinander korrelieren, aber teilschematische Konstruktionen dennoch innerhalb der Restriktionen produktiv sein können (Bybee 2013: 62). Als Beispiel nennt Bybee (2013: 62) die *to drive someone crazy*-Konstruktion, in der der *crazy*-Slot produktiv ist, so lange Synonyme von *crazy* genutzt werden (*to drive someone mad*). Diesbezüglich lässt sich gut mit Goldbergs exemplarbasiertem Modell der Abdeckung argumentieren (siehe Abschnitt 2.1.1 und 2.2.3): Da die Adjektive *crazy* und *mad* die Eigenschaft [+verrückt] teilen, sind beide kompatibel mit der Konstruktion. Durch die Überlappung der Eigenschaft wird der Adjektiv-Slot zudem spezifiziert und auf Adjektive mit der Eigenschaft [+verrückt] beschränkt (Goldberg 2019: 72). Dementsprechend können Cluster mit geringer Variabilität produktiv sein, solange die neuen Mitglieder sich innerhalb der beobachteten Variabilität befinden, wie im Beispiel der *to drive someone crazy*-Konstruktion (Bybee 2010: 67–69, Goldberg 2019: 65). Der Einfluss von Typenfrequenz und Ähnlichkeit auf Produktivität, der bereits in Abschnitt 2.1.1 diskutiert wurde, zeigt sich an dieser Stelle erneut: Die Offenheit eines Slots ist einerseits abhängig von seiner Typenfrequenz, d. h. von der Anzahl an verschiedenen Elementen, die in dem Slot verwendet werden können, und andererseits von der Ähnlichkeit der im Slot genutzten Elemente (Blumenthal-Dramé 2012: 8). Ähnlich argumentiert Diessel (2017: 16), der als Bedingung für die analogische Ausweitung von Schemata einerseits die Stärke der Aktivierung des Schemas (die von der Typenfrequenz abhängig ist) ansetzt und andererseits die Ähnlichkeit der Formen, die im Schema auftreten.

In diesem Abschnitt wurden Konstruktionen bislang in Hinblick auf Syntax diskutiert, aber die konstruktionsgrammatische Idee ist nicht darauf beschränkt, sondern umfasst „all levels of grammatical analysis“ (Goldberg 2006: 5). Sie lässt sich daher auch auf Morphologie anwenden. So kann das Konjugationsverhalten unregelmäßiger Verben als nicht- sowie teilschematische Konstruktionen

modelliert werden, da nur Verben mit einer bestimmten phonologischen Form ein bestimmtes (unregelmäßiges) Flexionsverhalten aufweisen können (Bybee & Slobin 1982, Bybee & Moder 1983, Bybee 2010: 66–69). Die phonologische Form kann vollständig spezifiziert sein wie bspw. bei einem unregelmäßigen Verb wie *sein*, dann liegt eine nicht-schematische Form vor. Die Form kann aber auch nur teilweise spezifiziert sein wie bei den Verben der Ablautreihe 3a und somit teilschematisch sein (Bücker 2015: 450). Die regelmäßige Konjugation ist hingegen eine hochschematische Konstruktion, da jedes Verb unabhängig von seiner phonologischen Form regelmäßig flektiert werden kann (Bybee 2010: 67). Dieser Blick auf Konjugation wird von Bybee (1991: 86) vorbereitet: Sie diskutiert die Teilschematizität der unregelmäßigen Konjugation und schlägt vor, (scheinbar) regelgeleitete Morphologie als schematische Konstruktion zu betrachten.

Hierbei ist hervorzuheben, dass das Verhältnis zwischen Konstruktionen und Schematizität in Bezug auf Flexionsklassen anders gelagert ist als bei den oben diskutierten syntaktischen Konstruktionen. Während bei den oben diskutierten Konstruktionen eine Funktion mit einer mehr oder weniger abstrakten Form verbunden ist, ist bei den Flexionsklassen eine Funktion (z. B. [+Vergangenheit]) mit zwei oder mehr Formen verbunden.⁵⁸ Eine der Formen ist hochschematisch und daher kaum gefüllt ([Dentalsuffix]),⁵⁹ die anderen Formen sind nicht- oder teilschematisch, da sie komplett (z. B. [war]) oder weitgehend festgelegt sind (z. B. [#_ a + η + (C)] wie in *sank* in Bezug auf Verben der Ablautreihe 3a). Passt ein Verb nicht oder nur teilweise zu den nicht- oder teilschematischen Formen, kann es die Funktion [+Vergangenheit] stets mit der schematischen Form [Dentalsuffix] ausdrücken.

Sieht man Konstruktionen als Resultat exemplarbasierten Lernens (siehe Abschnitt 2.1.3), ergibt es Sinn, dass nicht immer nur eine, sondern auch zwei (oder mehr) Formen mit einem Inhalt assoziiert sein können: Regelmäßiges Konjugationsverhalten kann anhand der Vielzahl und phonologischen Vielfalt von Verben, die dieses Verhalten aufweisen, generalisiert werden (Goldberg 2019: 67–68; siehe 2.1.1). Aber auch ein Flexionsverhalten, das vom Default-Verhalten abweicht, kann bis zu einem bestimmten Grad generalisiert werden: „[C]lustering of ‘exceptions’ leads speakers to generalize the behavior of exceptional cases“ (Goldberg 2019: 93). Die Generalisierung der Ausnahmen entsteht dabei aufgrund der partiellen Überschneidung der Eigenschaften einiger Exemplare (Bybee 2007: 171,

⁵⁸Die Formen sowie die Funktion werden jeweils durch eckige Klammern ([]) als Teil des Schemas gekennzeichnet.

⁵⁹Auch schwache Verben können teilschematisch motiviert sein (Goldberg 2019: 135): So sind Verben, die auf /f/ enden, im Englischen regelmäßig (*coughed, laughed, stuffed*); in Abschnitt 3.2.3 wird diese Idee in Bezug auf starke und schwache Maskulina aufgegriffen.

2 Variation aus gebrauchsbasierter Perspektive

Goldberg 2019: 121). So lässt sich aufgrund der Gemeinsamkeiten der Verben der Ablautreihe 3a die teilschematische Form [#_ a + η + (C)] für den Inhalt [+Vergangenheit] abstrahieren. Die Teilschematizität schränkt das Abstraktionspotential und damit die Produktivität zwar ein, gleichzeitig führt sie aber auch zu einem Gangeffekt: „[A] high concentration of verbs sharing phonetic properties is more likely to attract new members than one of comparable numbers without clear phonetic definition“ (Bybee 2010: 69). Die Generalisierung teilschematischer Formen trägt daher zur „Stabilität, Produktivität und Binnendifferenzierung [einer] Funktionsklasse“ bei (Bücker 2015: 449).

Da bei Flexionsklassen zwei (oder mehr) Formen mit einer Funktion verknüpft sind, können einzelne Verben nicht in eine Form gezwungen werden. Dies ist nur dann möglich, wenn lediglich eine Form mit einer Funktion verknüpft ist, sodass Elemente, die normalerweise nicht in einer bestimmten Form vorkommen, dennoch in diese gezwungen werden können, wenn sie zur Funktion passt (Taylor 1998: 198, Goldberg 2019: 37).

In Hinblick auf Variation sind bezüglich der Assoziation aus mehreren Formen und einer Funktion insbesondere die teilschematischen Formen relevant: Die Beziehung der teilschematischen Formen zur Funktion ist durch Familienähnlichkeit strukturiert (Bybee & Moder 1983: 263). Es besteht also wie bei Schemata in der Kognitionspsychologie keine distinkte Einteilung, stattdessen stellt sich die Frage nach der Passungsgüte (*goodness of fit*): Je besser eine Verbform zur Form [#_ a + η + (C)] passt, desto wahrscheinlicher ist die Assoziation der Verbform mit [+Vergangenheit]. Passt ein Verb nur peripher zu dieser teilschematischen Form, kann es zu Schwankungen zwischen der teilschematischen und der schematischen Form mit Dentalsuffix kommen (z. B. *sann/sinnte*) (siehe Abschnitt 2.3.3 zu Form-Schema-Effekten und Abschnitt 3.1.3 und 3.2.3 zu Form-Schematizität als Einflussfaktor auf die Konjugation und Deklination). Die Assoziation zwischen einer teilschematischen Form und der Funktion ist somit prototypisch organisiert und probabilistisch (Rumelhart 1980: 39, Bybee & Moder 1983: 262–264). Abbildung 2.7 stellt am Beispiel der teilschematischen Form [#_ a + η + (C)] das Verhältnis von hoch- und teilschematischer Form zur Funktion dar.

Die Assoziation zwischen Funktion und Formen lässt sich bayesianisch fassen, indem bspw. gefragt wird, wie wahrscheinlich die Assoziation mit [+Vergangenheit] ist, wenn eine bestimmte Wortform gegeben ist. Dabei ist eine Assoziation mit der schematischen Form immer möglich. Die Assoziation mit der teilschematischen Form hat eine höhere priore Wahrscheinlichkeit als die Assoziation mit der hochschematischen Form, wenn die gegebene Wortform die meisten Eigenschaften der teilschematischen Form mit hoher *cue validity* aufweist und daher einen prototypischer Vertreter der teilschematischen Form darstellt. Diese

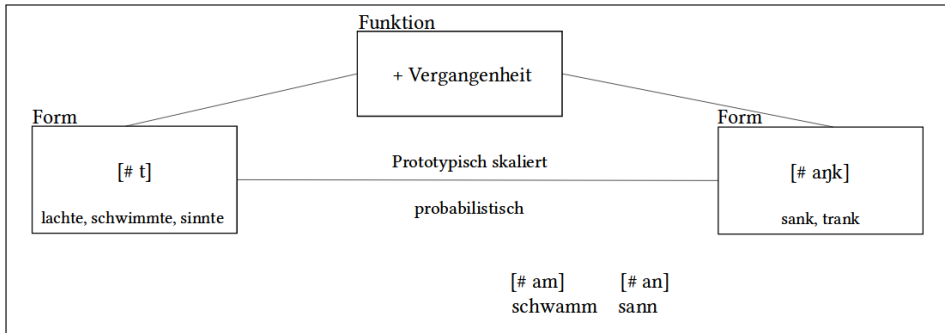


Abbildung 2.7: Verhältnis von hoch- und teilschematischer Form zur Funktion

Definition deckt sich mit der Definition eines Prototyps in Abschnitt 2.2.3. Die Assoziation zwischen einer teilschematischen Form und der Funktion [+Vergangenheit] blockiert die Assoziation der hochschematischen Form mit der Funktion nicht komplett, da man bspw. **sankte* als Präteritalform erkennen kann, jedoch führt die Assoziation dazu, dass eine konventionelle Form existiert, die die unkonventionelle Form statistisch aussticht (zum statistischen Vorkaufsrecht siehe Goldberg 2019: 74–94 sowie Abschnitt 2.1.3). Erfüllt die Form nur einige Eigenschaften der teilschematischen Form, ist die Assoziation mit der Funktion geschwächt, sodass auch die hochschematische Form genutzt werden kann (*sann* > *sinnte*), da die teilschematische Form die schematische nicht mehr so leicht dominieren kann. An dieser Stelle ist also Variation möglich.

Hinsichtlich der Assoziation von mehreren Formen mit einer Funktion ist von ineinander verschachtelten Schemata auszugehen. Die Assoziation einer teilschematischen Form wie bspw. $[\# _ a + \eta + (C)]$ mit der Funktion [+Vergangenheit] führt zu einer Assoziation der dazugehörigen Wortformen mit einem bestimmten grammatischen Verhalten: Da die Form $[\# _ a + \eta + (C)]$ mit [+Vergangenheit] assoziiert ist, ist auch die Form $[\# _ \text{I} + \eta + (C)]$ mit dem Konjugationsverhalten der Verben der Ablautreihe 3a verbunden. Wir haben es somit mit einem Makro-Schema zu tun: Die Assoziation einer teilschematischen Form mit einer Funktion (z. B. $[\# _ a + \eta + (C)]$ mit [+Vergangenheit]) ermöglicht die Assoziation des gesamten Lexems mit einem bestimmten Flexionsverhalten. Das Makro-Schema (Lexem und Flexionsverhalten) lässt sich stets dahingehend herunterbrechen, dass die Flexionsformen mit einer bestimmten Funktion (z. B. [+Vergangenheit], [+Partizip II]) verbunden sind (Schmitt 2019a: 161), weswegen die hochschematische Form statistisch ausgestochen wird (Goldberg 2019: 83–84). Das Form-Schema der Verben der Ablautreihe 3a wird in Abschnitt 3.1.3 näher erläutert.

2 Variation aus gebrauchsbasierter Perspektive

Form-Schemata können auch komplexer sein, als es bei den Verben der Ablautreihe 3a der Fall ist. Sie können auch semantische Merkmale enthalten, wie die Diskussion um das Form-Schema schwacher Maskulina in Abschnitt 3.2.3 zeigt. Hierbei ist von ineinander verschachtelten Schemata auszugehen: Die phonologischen Eigenschaften Dreisilbigkeit, Pänultimabetonung, Schwa sind mit hoher Belebtheit assoziiert und dieses Schema wiederum ist mit schwacher Deklination verknüpft und bildet damit ein Form-Schema. Dieses spezielle Form-Schema wird ausführlich in Abschnitt 3.2.3 vorgestellt. Darin wird auch diskutiert, welche Eigenschaften des Form-Schemas Einfluss auf die Deklination nehmen.

Makro-Schemata können nicht nur in Bezug auf die teilschematische, sondern auch in Bezug auf nicht- und hochschematische Formen angesetzt werden. Die nicht-schematische Form *war* ist mit [+Vergangenheit] assoziiert, auch hier sticht die Assoziation andere Formen statistisch aus und assoziiert *sein* mit unregelmäßiger Flexion. Genauso führt die Verknüpfung der hochschematischen Form [Dentalsuffix] mit [+Vergangenheit] zur Assoziation eines Verbs mit regelmäßiger Flexion. Makro-Schemata lassen sich in diesem Hinblick als Resultat exemplarbasierten Lernens fassen: Zunächst werden eine Wortform und eine Funktion verknüpft, dann das Lexem und ein bestimmtes grammatisches Verhalten. An beiden Punkten ist eine Erweiterung auf Wörter mit ähnlichen Eigenschaften möglich: Dadurch, dass Wörter mit ähnlichen Eigenschaften wiederholt dieselbe Funktion aufweisen, kann eine (teil-)schematische Repräsentation der Form abstrahiert werden (Schmitt 2019a: 161). Nessel (2008: 18–21) verfolgt mit *second-order* Schemata einen ähnlichen Ansatz: Schemata wie die oben genannten Verbindungen aus Form und Funktion bei Vergangenheit und Partizip II, die Nessel (2008: 18–21) *first-order* Schemata nennt, können zu einem *second-order* Schema zusammengeführt werden, sodass paradigmatische Beziehungen entstehen. Köpcke & Wecker (2017) und Köpcke u. a. (2021) zeigen in einem Experiment zur Wahrnehmung von Pseudowörtern als Singular- oder Pluralform, dass Proband_innen sich in der Bewertung sowohl auf *first-order* als auch auf *second-order* Schemata im Sinne Nessels verlassen. Kinder mit Deutsch als L1 verlassen sich dabei zunächst verstärkt auf *first-order* Schemata, bevor sie sich auf *second-order* Schemata stützen. Dies spricht dafür, dass *second-order* Schemata auf Basis von *first-order* Schemata abstrahiert werden.

Nicht nur in Bezug auf zwei Formen, die mit einer Funktion assoziiert sind, kann es zu Variation kommen (*sann/sinnte*). Auch in Hinblick auf Konstruktionen, bei denen nur eine Form mit einer Funktion assoziiert ist, ist dies möglich. So können die Funktionen zweier Schemata prototypisch gestaffelt und die jeweiligen Pole der Funktionen daher mit unterschiedlichen Formen assoziiert

sein. Dies lässt sich bspw. bei der Auxiliarselektion von *haben* und *sein* beobachten, die u. a. durch Transitivität beeinflusst wird (siehe hierzu ausführlich Abschnitt 3.3.2). Transitive Sätze selektieren *haben*, intransitive (in telischen Sätzen und Sätzen mit Bewegungssemantik) *sein* (Gillmann 2016: 316–319). Wir haben es also mit zwei separaten Schemata zu tun, die einem Makro-Schema ([+Perfekt]) angehören. Variation tritt auf, wenn die Funktion nicht eindeutig transitiv oder intransitiv ist, wie bspw. bei einem Objekt, das sowohl als Patiens als auch als Instrument interpretiert werden kann (*Ich habe/bin den Porsche gefahren*; Gillmann 2016: 282). Abbildung 2.8 stellt die prototypisch organisierte Beziehung zwischen den beiden Funktionen schematisch dar, die zu Variation in der Form führen.

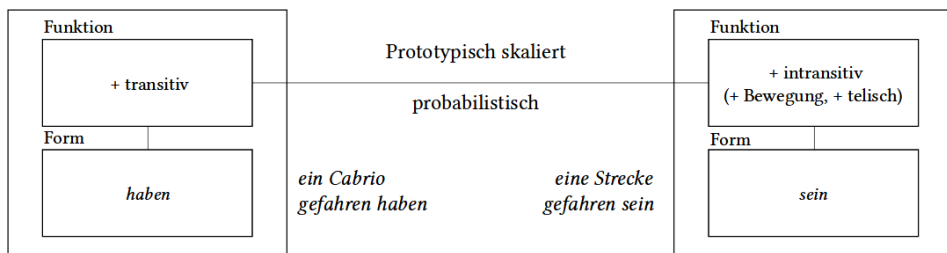


Abbildung 2.8: Prototypisch organisierte Beziehung zwischen zwei Funktionen separater Schemata

Im Folgenden werden zwei Arten von Schematizität unterschieden. Der Terminus *Schematizität* wird genutzt, um auf die Abstraktheit der Form bei Konstruktionen zu referieren; als *Schema* werden dabei hochschematische, also abstrakte Konstruktionen bezeichnet (Booij 2010: 5, Bybee 2010: 80). Wenn zwei (oder mehr) Formen, die sich durch ihren Schematizitätsgrad unterscheiden, um die Assoziation mit einer Funktion konkurrieren, wird dies unter dem Terminus *Form-Schematizität* gefasst. Dabei werden teilschematische Konstruktionen in Abgrenzung zur hochschematischen Konstruktion (sowie von nicht-schematischen Konstruktionen wie *war* und [+Vergangenheit]) *Form-Schema* genannt. Bei der Form-Schematizität stehen sich somit Schema (hochschematische Form) und Form-Schema (teilschematische Form) gegenüber.⁶⁰

⁶⁰Die Benennung der teilschematischen Konstruktion als (*Form-*)*Schema* lehnt sich an Bybee (1985: 129) an. Sie führt den Terminus *Schema* in Bezug auf Flexionsklassen ein, um teilschematische Konstruktionen zu bezeichnen. Dies ist gegenläufig zur Benennung hochschematischer Konstruktionen als *Schema* (Booij 2010: 5, Bybee 2010: 80) und m. E. darauf zurückzuführen, dass die Assoziation zwischen Form und Funktion zunächst nur für unregelmäßige Verben diskutiert wurde (bspw. von Bybee & Moder 1983). Die Idee, regelmäßige Verben als Assoziation zwischen einer abstrakten Form und einer Funktion zu sehen, entstand meines Wissens erst später und findet sich bspw. bei Bybee (1991: 86).

2 Variation aus gebrauchsbasierter Perspektive

Schemata und Form-Schemata sind eng mit Prototypizität verknüpft. Schemata können in prototypischer Beziehung zueinander stehen, wie bei der Selektion von *haben* und *sein*. Genauso zeigt sich bei Form-Schemata eine prototypische Organisation. Dennoch ist es wichtig, Prototypizität und (Form-)Schematizität zu unterscheiden: (Form)-Schemata sind Form-Funktionspaare, während Prototypizität ein Ordnungsprinzip darstellt, das sich auf (Form)-Schemata anwenden lässt. Wenn zwei separate Schemata durch die prototypische Beziehung zwischen zwei Funktionen verbunden sind und daher im Übergangsbereich der Funktionen Schwankungen zwischen den Formen der Schemata zu beobachten sind, wird dies in der vorliegenden Arbeit als prototypisch bedingte Konkurrenz zwischen zwei ansonsten separaten Schemata gesehen und als Effekt der Prototypizität verhandelt. Bei Form-Schemata geht es dagegen um die prototypisch organisierte Assoziation zwischen einer teilschematischen Form, die mit einer anderen Form um eine Funktion konkurriert.

Schemata wurden anhand der Schematizität von Konstruktionen bereits ausführlich vorgestellt und diskutiert; im weiteren Verlauf dieses Abschnitts sowie in Abschnitt 2.3.3 wird daher auf Form-Schemata fokussiert.

Wichtig bei der Betrachtung von Schemata und Form-Schemata ist der unterschiedliche Status von Typenfrequenz. Für Form-Schemata ist typenfrequenz essentiell, da nur eine relativ typenfrequente Form innerhalb der typeninrequenten Flexion stark genug mit der jeweiligen Funktion assoziiert sein kann, um die typenfrequente konkurrierende Form statistisch auszustechen. Bei Schemata spielt Typenfrequenz hingegen eine untergeordnete Rolle: Da hier zwei Formen jeweils mit zwei separaten Funktionen assoziiert sind, kann die typenfrequente die typeninfrequente nicht so leicht statistisch ausstechen. Die typenfrequente Form ist mit einer anderen Funktion assoziiert und damit greift das statistische Vorkaufsrecht nicht. Genau wie Schemata können auch Form-Schemata komplex sein und selber ein Schema enthalten. Die Unterscheidung, ob separate Schemata oder Form-Schemata vorliegen, hängt immer daran, ob eine Funktion oder mehrere Funktionen mit den verschiedenen Formen verbunden sind. Bei *haben* und *sein* haben wir es mit zwei Funktionen zu tun, was man leicht daran erkennen kann, dass Verben in das Schema gezwungen werden können, wenn sie transitiv bzw. intransitiv verwendet werden, auch wenn sie eigentlich mit dem jeweils anderen Schema assoziiert sind. Dies zeigt sich bspw. bei *heizen*: *Heizen* ist zwar intransitiv, aber weder ein Bewegungsverb noch telisch, daher bildet es das Perfekt mit *haben* (*Ich habe geheizt*). Wird es jedoch metaphorisch verwendet, um auf schnelles Fahren zu verweisen, wechselt das Auxiliar aufgrund der Bewegungssemantik zu *sein* (*Ich bin durch die Gegend geheizt*). Dasselbe Prinzip zeigt sich bei Verben, die aufgrund ihrer Telizität normalerweise mit *sein* verknüpft

sind wie bspw. *verderben*. Sobald diese transitiv genutzt werden, weisen diese *haben* als Auxiliar auf (*Das Salz hat die Suppe verderben*). Im Gegensatz dazu können weder einzelne Verben noch einzelne Maskulina können in ein Form-Schema gezwungen werden. Das liegt daran, dass nicht zwei Funktionen wie z.B. transitiv und intransitiv gegenüber stehen, sondern teilschematische Formen mit einer hochschematischen um eine Funktion konkurrieren.

Köpcke (1993) beschäftigt sich mit Form-Schematizität in der Pluralbildung des Deutschen. Die Pluralbildung im Deutschen lässt sich als Form-Schematizität auffassen, da auch hier eine Funktion [+Plural] mit mehreren Formen (z. B. *-e, -s*) assoziiert ist (*Autos, Tische*). Köpcke (1993) arbeitet nicht mit der Unterscheidung von Schema und Form-Schema, sondern sieht jeweils die Verbindung aus Form und Funktion als Schema, die er als „ausdrucksseitige Gestalt, der eine spezifische Regelmäßigkeit in dem Sinne anhaftet, daß sie ein bestimmtes Konzept [...] wiederholt ausdrucksseitig repräsentiert“ definiert (Köpcke 1993: 72).

Köpcke (1993: 82–83) geht davon aus, dass den verschiedenen Formen, die mit einer Funktion assoziiert sein können, eine unterschiedliche Signalstärke zukommt. Formen mit höherer Signalstärke geben einen besseren Hinweis auf die Funktion als Formen mit niedriger Signalstärke. Indem Köpcke (1993: 82–83) die Signalstärke von Typen- und Tokenfrequenz, akustischer Salienz und *cue validity* beeinflusst sieht,⁶¹ nimmt er implizit die in Abschnitt 2.2.1 diskutierten Einflussfaktoren auf Typizität auf.

Köpcke (1993: 82) selbst erläutert den Einfluss der Typen- und Tokenfrequenz nicht, allerdings führt er sie in späteren Publikationen aus. So argumentiert er, dass *-en* als Pluralmarker typen- und tokenfrequenter und damit prototypischer ist als andere Pluralmarker (Köpcke & Wecker 2017: 83). Auch aus Abschnitt 2.1.1 wird deutlich, dass frequente Exemplare sowie Exemplare mit frequent auftretenden Eigenschaften als prototypisch (für Form-Schemata) gelten können. So schreiben Bybee (2010: 67) und Poitou (2004a: 75) Form-Schemata⁶² mit vielen Vertretern eine höhere Produktivität und Anziehungskraft zu als Form-Schemata

⁶¹Zusätzlich zu diesen Eigenschaften geht Köpcke (1993: 83) davon aus, dass Form-Schematizität durch Ikonizität befördert werden. Hierbei greift er auf die Natürlichkeitsmorphologie nach Wurzel (1984) zurück, indem er davon ausgeht, dass Plural durch ein Mehr an Form ikonisch dargestellt wird (Köpcke 1993: 83). Diese Eigenschaft ist jedoch m. E. nicht auf Form-Schematizität im Allgemeinen übertragbar, weswegen der Verbindung zwischen Ikonizität und Signalstärke hier nicht weiter auf den Grund gegangen wird.

⁶²Weder Bybee (2010) noch Poitou (2004a) nutzen den Terminus *Form-Schema*. Poitou (2004a) verwendet auch den Terminus *Schema* nicht, sondern *Prototyp*. Beide referieren jedoch auf Form-Funktions-Paare mit teilschematischer Form, die in dieser Arbeit als Form-Schema modelliert werden.

2 Variation aus gebrauchsbasierter Perspektive

mit wenigen Vertretern. Akustische Saliienz definiert Köpcke (1993: 82) relativ unspezifisch mit „Bestimmung des Ausmaßes, mit dem eine morphologische Markierung vom Hörer identifizierbar ist“. Als Beispiel nennt er segmentierbare Pluralmarkierungen wie *-s* oder *-er*, denen er aufgrund ihrer Segmentierbarkeit sowie der Positionierung am Ende des Worts eine höhere Saliienz zuschreibt als dem Umlaut, der lediglich ein Begleitphänomen ist (Köpcke 1993: 82).

Die *cue validity* benutzt Köpcke (1993: 82) deckungsgleich wie in Abschnitt 2.1.1 erläutert. Je zuverlässiger eine Form eine Funktion repräsentiert, desto größer ist deren *cue validity*: „A category has high cue validity if the features associated with it frequently occur with members of the category, and rarely occur with members of other categories“ (Bybee & Moder 1983: 266). So ist die *cue validity* von *-s* geringer als die *cue validity* von *-er*, da *-s* auch den Genitiv markiert (Köpcke 1993: 82–83).⁶³ In Bezug auf unregelmäßige Verben im Englischen stellen Bybee & Moder (1983: 266) fest, dass / $\Delta\eta$ / und / $\Delta\eta k$ / eine hohe *cue validity* für Vergangenheit aufweisen, da nur zwei Verben in ihrer Grundform auf / $\Delta\eta k$ / (*flunk* und *dunk*) enden. / $\Delta\eta$ / hat dabei eine noch größere *cue validity*, da keine Verben und nur wenige Substantive und Adjektive (z. B. *tongue*, *lung*, *dung*) / $\Delta\eta$ / aufweisen.

Der Abschnitt hat Schemata und Form-Schemata als verwandte Konzepte herausgearbeitet. Schemata sind abstrakte und damit hochschematische Konstruktionen. Konkurrieren zwei (oder mehr) Formen um die Assoziation mit einer Funktion, wird dies als Form-Schematizität bezeichnet. Eine der Formen ist hochschematisch, die anderen teil- oder nicht-schematisch. Die Verbindung aus teilschematischer Form und Funktion wird dabei als Form-Schema bezeichnet. Form-Schemata lassen sich somit als probabilistische Assoziationen zwischen einer teilschematischen Form und einer Funktion fassen, die eine Assoziation einer hochschematischen Form mit derselben Funktion statistisch aussticht (zum statistischen Vorkaufsrecht siehe Goldberg 2019: 74–94 sowie Abschnitt 2.1.3). Der teilschematischen Form kommt dabei eine hohe *cue validity* zu, sodass die Form eng mit der Funktion assoziiert ist.

Form-Schemata nehmen Einfluss auf Variation: Prototypische Vertreter des Form-Schemas weisen keine Variation auf, da sie die hochschematische Form statistisch ausstechen können. In der Peripherie der teilschematischen Form kann es dagegen zu Schwankungen kommen, da das statistische Vorkaufsrecht hier

⁶³Köpcke (1993: 83) geht jedoch davon aus, dass die Genitivmarkierung nur einen geringen Störfaktor darstellt, da der Artikel die Konstruktion i. d. R. desambiguiert (*des Autos* vs. *die Autos*). In der Diskussion um *cue validity* scheint Köpcke (1993) nur auf Konkurrenzformen innerhalb der Flexion einzugehen. Betrachtet man auch Wortbildung, stellt sich die Frage, ob die *cue validity* von *-er* ebenfalls herabgesenkt ist, weil es auch als Wortbildungselement genutzt wird.

nicht mehr vollständig greift. Dieser Effekt wird bei der Betrachtung der Variation in der Deklination und Konjugation eine zentrale Rolle spielen. Im folgenden Abschnitt werden Effekte genauer vorgestellt, die durch Form-Schemata ausgelöst werden.

2.3.3 Form-Schema-Effekte

Bybee & Moder (1983) überprüfen den Einfluss des Form-Schemas [#_Λ + velarer und/oder nasaler Konsonant] und [+Vergangenheit] auf die Flexion im Englischen. In ihrer Studie sollen Proband_innen Vergangenheitsformen von Pseudoverben bilden. Dabei nehmen Bybee & Moder (1983) an, dass der Ablaut auf /Λ/ für Pseudoverben am wahrscheinlichsten ist, die in ihrer Grundform /s/CC am Silbenanfang, /ɪ/ als Stammvokal und einen velaren Nasal in der Silbencoda haben (wie das Verb *string*). Diese spezifische Form lässt sich durch die Verben der Klasse II⁶⁴ erklären (Bybee & Moder 1983: 257–258): Die meisten Verben der Klasse II haben s-haltige Konsonantencluster, /ɪ/ als Stammvokal und /ŋ/.⁶⁵ Die Ergebnisse des Pseudowortexperiments bestätigen die Hypothese: Pseudoverben mit /s/CC am Silbenanfang, /ɪ/ als Stammvokal und einem velaren Nasal in der Silbencoda weisen im Experiment mit 44 % den größten Anteil an Ablautungen auf /Λ/ auf.⁶⁶ Die Produktivität des Form-Schemas kann zudem diachron nachvollzogen werden (Bybee & Moder 1983: 263, 265–266): Anders als andere Klassen unregelmäßiger Verben des Englischen hat die Klasse II Verben angezogen (bspw. *to hang*).

Die einzelnen Slots der Form (Silbenanfang, Vokal, Silbencoda) haben einen unterschiedlich großen Einfluss auf die Kompatibilität eines Verbs mit dem Form-Schema. So hat der velare Nasal einen größeren Einfluss als das Anfangscluster:

⁶⁴Bybee & Moder (1983: 252) arbeiten mit zwei Verbklassen: Klasse I enthält Verben mit dem Ablautmuster x - æ - Λ, Klasse II Verben mit dem Ablautmuster x - Λ - Λ. Diese Klasseneinteilung beruht nicht auf Ablautreihen. Beide Klassen gehen m. E. auf die Ablautreihe 3a zurück. Dies verdeutlichen Verben der Ablautreihe 3a wie *singen* und *spinnen*. Sie sind kognat mit den Verben *sing* und *spin*, die Klasse I bzw. Klasse II angehören.

⁶⁵Neun von achtzehn Verben der Klasse II beginnen mit einem s-haltigen Konsonantencluster. Bybee & Moder (1983) gehen nicht explizit darauf ein, warum sie /ɪ/ als Stammvokal annehmen, allerdings findet sich dieser Vokal auffällig häufig bei real existierenden Verben der Klasse II (Bybee & Moder 1983: 252). Auch hinsichtlich der Silbencoda ist davon auszugehen, dass /ŋ/ besser in das Form-Schema passt als /ŋk/, da /ŋ/ häufiger in Klasse II zu beobachten ist (Bybee & Moder 1983: 257).

⁶⁶Leider schlüsseln Bybee & Moder (1983) nicht auf, auf welche Flexionsformen die restlichen Prozente entfallen. Aufgrund der Aufschlüsselungen zu einzelnen Bestandteilen des Form-Schemas ist davon auszugehen, dass sich diese auf das Dentalsuffix *-ed* und den Ablaut /æ/ verteilen.

2 Variation aus gebrauchsbasierter Perspektive

Wird statt des velaren Nasals ein anderer (weder velarer noch nasaler) Konsonant (/v/, /t/, /d/, /p/ oder /d/) verwendet, fällt die /ʌ/-Ablaurate auf 4 %, während sie sich bei Ersetzung des Anfangsclusters durch einen einfachen Konsonanten bei 24 % hält (Bybee & Moder 1983: 261). Die Silbencoda scheint das Flexionsverhalten also stärker zu beeinflussen als der Anfang. Als Erklärung hierfür führen Bybee & Moder (1983: 264) an, dass Flexion im Englischen durch Suffixe ausgedrückt wird und die Silbencoda in der Flexion somit mehr im Fokus steht als der Silbenanfang.

Einen noch geringeren Einfluss als der Silbenanfang hat allerdings der Vokal: /æ/ als Stammvokal evoziert ähnlich häufig den Ablaut auf /ʌ/ wie der als ideal angenommene Vokal /ɪ/. Bybee & Moder (1983: 255–256) gehen davon aus, dass die Zielform (*strung*) für das Form-Schema eine größere Rolle als die Ursprungsform (*string*) spielt. Aufgrund der Ablautung beeinflusst der Stammvokal das Form-Schema daher kaum.⁶⁷

Den geringen Einfluss des Vokals begründen Bybee & Moder (1983) auch diachron, da ursprünglich nur Verben mit dem Stammvokal /ɪ/ den Ablaut auf /ʌ/ bildeten, inzwischen aber auch andere Stammvokale möglich sind wie bspw. /i:/ und /aɪ/ (*to sneak, to strike*) (Bybee & Moder 1983: 255). Zudem zeigen Konsonantenveränderungen, die Proband_innen innerhalb des Pseudowortexperiments von Bybee & Moder (1983) vornahmen, dass die Zielform für das Form-Schema wichtiger zu sein scheint als die Ursprungsform: So wurden zu dem Stimulus *to spriv* die Vergangenheitsformen *sprung* und *sprug* gebildet. Hier wurde die Silbencoda von /v/ zu /ŋ/ bzw. /g/ verändert und somit dem Form-Schema angepasst (Bybee & Moder 1983: 263).

Neben dem unterschiedlichen Einfluss einzelner Slots (Silbenanfang, Stammvokal, Silbencoda) zeigt sich, dass die Slots selbst prototypisch organisiert sind (Bybee & Moder 1983: 260). In der Silbencoda ist der velare Nasal /ŋ/ am besten kompatibel mit dem Form-Schema, da er zu 37 % Ablautung auf /ʌ/ hervorruft.

⁶⁷Bybee & Moder (1983: 255–256) nennen diese Modifikation daher *product oriented*, weil allein das Produkt im Vordergrund steht. *Product oriented* wird dabei als Gegensatz zu *source oriented* verwendet. Bei dieser Modifikation steht die paradigmatische Beziehung von zwei Formen zueinander im Vordergrund (Bybee 2007: 171). Hierbei kommt wiederum Nessets Unterscheidung von *first-* und *second-order* Schemata zutragen: *First-order* Schemata sind *product oriented*, da allein die Beziehung von Funktion und Form im Vordergrund steht und nicht die Beziehung der Ursprungsform zur flektierten Form. *Second-order* Schemata sind dagegen *source oriented*, weil hier die Beziehung zwischen den Wortformen relevant ist. *Second-order* Schemata entstehen aus *first-order* Schemata (Köpcke u. a. 2021: 77), starke und schwache Konjugation gehen somit nicht auf unterschiedliche Modifikationen zurück, sondern auf denselben Prozess, wobei sich schwache Formen lediglich leichter zu *second-order* Schemata generalisieren lassen als starke Formen.

Wenn bspw. mit der Nasalität eine Eigenschaft entfällt, ist der Konsonant weniger geeignet: /k/ und /g/ führen zu weniger Ablautungen (14 %). Ähnlich verhalten sich die nicht-velaren Nasale /n/ und /m/ mit 15 % Ablautung auf /ʌ/. Entfällt sowohl die Eigenschaft [+velar] als auch [+nasal], ist die Kompatibilität weiter herabgesetzt: /p/, /b/ und /v/ rufen nur zu 2 % Ablautung auf /ʌ/ hervor (Bybee & Moder 1983: 260).

Die Passungsgüte innerhalb einzelner Slots lässt sich also sehr exakt anhand von Lauteigenschaften (wie bspw. Nasalität) bestimmen. Dabei nehmen die Slots an sich und die Passungsgüte innerhalb der Slots Einfluss auf die Kompatibilität eines Verbs mit dem Form-Schema. Zudem haben die Slots nicht nur einzeln einen Einfluss auf die Wahrscheinlichkeit der Ablautung auf /ʌ/, sondern wirken sich auch gemeinsam auf die Ablautung auf /ʌ/ aus: Je mehr prototypische Vertreter der einzelnen Slots vorhanden sind, desto wahrscheinlicher wird die Ablautung auf /ʌ/ (Bybee & Moder 1983: 261–261). Auch hier zeigt sich die prototypische Organisation von (Form-)Schemata (Bybee 1988: 135).

Köpcke u. a. (2021) zeigen in Pseudowortstudien die Form-Schematizität von Pluralformen im Deutschen auf. Dabei stellen sie zum einen fest, dass den unterschiedlichen Pluralformen unterschiedliche Signalvalidität zukommt. So ist *-en* ein besserer Marker für Plural als ein Nullmarker, weil die meisten Substantive im Deutschen eine Pluralmarkierung aufweisen. Pseudowörter auf *-en* werden daher eher pluralisch interpretiert als solche ohne Pluralmarkierung (Köpcke u. a. 2021: 10–19). Zudem stellen sie fest, dass ihre Proband_innen nicht nur auf *first-order* Schemata zurückgreifen, sondern auch auf *second-order* Schemata. Werden die Versuchswörter mit dem Artikel *die* präsentiert, werden Substantive auf *-e* seltener als Plural angesehen, als wenn die Versuchswörter ohne *die* präsentiert werden: Die Kombination aus *die* und Substantiv fördert die Assoziation mit einem Femininum. Für ein Femininum ist die Opposition aus auf Schwa endenden Singular und einer Pluralform auf *-en* wahrscheinlich, daher wird die Pluralinterpretation für *-e* bei *die* + Pseudosubstantiv dispräferiert (Köpcke u. a. 2021: 18). Die Studie zeigt zudem, dass bei ambigen Formen (z.B. Formen auf *-e* und mit Umlaut und *-e*) die Reaktionszeiten im Vergleich zu stark mit Singular oder Plural assoziierten Formen (*-en*, *-s* für Plural, Nullmarkierung für Singular) erhöht sind (Köpcke u. a. 2021: 19).

Aufgrund der prototypischen Struktur haben Form-Schemata sowohl einen variationshemmenden als auch einen variationstreibenden Effekt in der Flexionsmorphologie. Einerseits hemmen Form-Schemata Variation, weil Vertreter eines Form-Schemas sich gegenseitig stützen. Für tokeninfrequente Verben des Englischen mit /s/CC am Silbenanfang, /ɪ/ als Stammvokal und einem velaren Nasal in der Silbencoda ist Variation hin zur regelmäßigen Konjugation auf *-ed* somit

2 Variation aus gebrauchsbasierter Perspektive

weniger wahrscheinlich als für andere unregelmäßige Verben mit ähnlicher Frequenz, die aber keinem Form-Schema angehören. In diesem Sinne sind Form-Schemata eine Art Gegenspieler zur Typenfrequenz (siehe Abschnitt 2.1), welcher die typeninfrequente Flexionsklasse vor dem Abbau bewahrt. Gleichzeitig kann man Form-Schemata als typenfrequentes Cluster innerhalb einer typenin-frequenten Klasse sehen: Erst wenn viele Vertreter einer typenin-frequenten Klasse eine bestimmte Form aufweisen, können sie sich gegenseitig stützen und neue Mitglieder anziehen.

Form-Schemata können andererseits Variation befeuern: Zum einen kann eine Form nur teilweise dem Form-Schema entsprechen. Die Assoziation zwischen Form und Funktion ist in diesem Fall weniger stark, weswegen die Funktion auch mithilfe der hochschematischen Form (also bspw. durch Dentalsuffix) ausgedrückt werden kann, da sie nicht mehr statistisch ausgestochen wird. Zum anderen können Formen existieren, die zwar formal zum Form-Schema passen, allerdings dennoch die Funktion mithilfe der hochschematischen Form ausdrücken. In diesem Fall kann es aufgrund der Ähnlichkeit zu den Vertretern des Form-Schemas zu Schwankungen in der Form kommen, wie bspw. bei *winken*, das eigentlich schwach flektiert, aber aufgrund der Ähnlichkeit zu Verben der Ablautreihe 3a mit *gewunken* auch ein starkes Partizip aufweist (siehe hierzu genauer Abschnitt 3.1.3). Neben der Entwicklung hin zur teilschematischen Form besteht auch die Möglichkeit, die Form so zu verändern, dass die Passungsgüte zu dem anziehenden Form-Schema verringert wird. Dies ist bspw. bei schwachen Maskulina wie *Friede* der Fall, die einen geringen Belebtheitsgrad aufweisen und somit nicht zum Form-Schema passen, aber bezüglich der prosodisch-phonologischen Eigenschaften dem Form-Schema schwacher Maskulina entsprechen. Sie nehmen *-n* im Nominativ an (*der Frieden*) (Köpcke 1995: 173) und werden hierdurch dem Form-Schema unähnlicher, da sie nicht mehr auf Schwa enden (siehe hierzu ausführlich Abschnitt 3.2.3).⁶⁸

Der Abschnitt hat Schemata und ihren Einfluss auf Variation betrachtet. Dabei wurden zwei Arten von Schemata ausgearbeitet: Der erste Typus von Schemata sind abstrakte Konstruktionen. Hier kann es aufgrund der prototypischen

⁶⁸Als weiteres Beispiel lassen sich die Pluralmarker des Deutschen nennen (Köpcke 1993: 119–120). Im Frühneuhochdeutschen existierten zahlreiche Feminina, deren Nominativ Singular auf *-n* endete (*die kuchen, die lügen, die reden*). Da *-n* sich als Pluralmarkierung für Feminina durchsetzte, führt die *n*-haltige Nominativform im Singular zur irreführenden Assoziation mit der Funktion [+Plural]. Daher wurde *-n* in den Singularformen zum Neuhochdeutschen hin getilgt (*die Küche, die Lüge, die Rede*), sodass die Singularform nicht mehr einer Form gleicht, die mit Plural assoziiert ist (Köpcke 1993: 119). Auf diese Weise wurde auch die *cue validity* von *-n* als Marker für Plural gestärkt.

2.4 Zusammenfassung: Variation aus gebrauchsbasierter Perspektive

Organisation zu Variation kommen, bspw. wenn die Funktionen zweier Schemata prototypisch organisiert sind, wie bei der Auxiliarselektion von *haben* und *sein*. Dies wird in der Arbeit als Prototypeneffekt gefasst, da die Schemata abseits der prototypischen Verbindung der Funktionen separat sind. Für den Abschnitt war daher vornehmlich der zweite Typus von Schemata relevant: Form-Schemata. Form-Schemata sind Verbindungen einer Funktion (z. B. [+Vergangenheit]) mit einer teilschematischen Form (z. B. [#_ a + η + (C)]). Dabei ist die Funktion jedoch nicht nur mit der teilschematischen Form, sondern mit mehreren Formen mit unterschiedlichem Schematizitätsgrad verbunden, im Fall der Funktion [+Vergangenheit] bspw. auch mit der hochschematischen Form [Dentalsuffix]. Es liegt somit Form-Schematizität vor. Form-Schemata nehmen vorrangig innerhalb von typeninrequenten Flexionsklassen Einfluss auf Variation: Sie stabilisieren die Flexion der typeninrequenten Klasse, da sie eine Typenfrequenz innerhalb der Typeninfrequenz darstellen. Der Einfluss von Form-Schemata zeigt sich dabei auch bei Pseudowörtern. Aufgrund ihres probabilistischen Aufbaus können Form-Schemata auch zu Variation führen, bspw. wenn ein Verb nur teilweise zur Form des Form-Schemas passt. Hier kann es zu Schwankungen hin zur hochschematischen Form kommen (*sann* > *sinnte*). Außerdem kann das Form-Schema neue Mitglieder anziehen, wie starke Partizip-II-Formen des ursprünglich schwachen Verbs *winken* zeigen, dessen phonologische Eigenschaften zum Form-Schema der Verben der Ablautreihe 3a passen. Der Einfluss von Form-Schemata auf die Variation in der Konjugation und Deklination wird in den Abschnitten 3.1 und 3.2 näher erläutert. Im empirischen Teil der Arbeit wird der Einfluss von Form-Schemata auf die Variation überprüft, indem Substantive mit periphere und prototypischer Form des Form-Schemas schwacher Maskulina gegenübergestellt werden. Der folgende Abschnitt fasst den Einfluss von Frequenz, Prototypizität und Schematizität auf Variation zusammen und diskutiert, wie die Einflussfaktoren interagieren.

2.4 Zusammenfassung: Variation aus gebrauchsbasierter Perspektive

Im gebrauchsbasierten Ansatz sind mentale Repräsentationen und damit auch Sprachsysteme eine Konsequenz des Sprachgebrauchs, sodass Frequenz einen grundlegenden Einfluss auf Variation hat. Wie fundamental Frequenz den Sprachgebrauch formt, wird bspw. anhand der Zipfschen Verteilung deutlich: Im Sprachgebrauch existieren wenige hochfrequente Elemente und viele niedrigfrequente (Zipf 1972 [1949]: 22–27, Ellis 2012: 13). Der grundlegende Einfluss der

2 Variation aus gebrauchsbasierter Perspektive

Frequenz zeigt sich außerdem darin, dass sie den Einfluss von Prototypizität und (Form-)Schematizität auf Variation erst ermöglicht.

Frequenz kann Variation fördern und hemmen. Hierbei ist zunächst festzustellen, dass tokenfrequente Elemente generell gefestigter und damit leichter zu aktivieren sind als tokeninfrequente (Bybee & Thompson 1997: 380, Schneider 2014: 13). Hohe Tokenfrequenz führt also zu Stabilität, geringe kann zu Variation führen. Existieren mehrere Kategorien (bspw. Flexionsklassen) mit unterschiedlicher Typenfrequenz⁶⁹, ist zudem ein Zusammenspiel aus Token- und Typenfrequenz zu beobachten. Die typenfrequenteren Kategorie ist produktiv⁷⁰ (Bybee & Thompson 1997: 384), weist ein regelmäßiges Verhalten auf und die Mitglieder der Klasse sind i. d. R. nicht tokenfrequent. Im Gegensatz dazu sind typeninfrequente Kategorien i. d. R. nicht produktiv⁷¹ und flektieren unregelmäßig. Die Mitglieder von typeninfrequenten Kategorien weisen zudem eine hohe Tokenfrequenz auf (Bybee & Thompson 1997: 380, Ellis 2002a: 166). Diese fördert Unregelmäßigkeit: Durch die hohe Tokenfrequenz sind die unregelmäßigen Formen mental stark gefestigt. Zudem sind die Wortformen eines hochfrequenten Worts weniger stark miteinander verknüpft als die Wortformen eines infrequenten Worts (Bybee 1985: 117–118, Bybee 2006: 715). Gleichzeitig führt die hohe Tokenfrequenz aufgrund des *entrenchments* zu einem konservierenden Effekt und macht einen Wechsel in die typenfrequente Klasse unwahrscheinlich (Bybee 1985: 117, 2002a: 619–621).

Verliert ein Vertreter der typeninfrequenten Kategorie an Frequenz, ist die mentale Repräsentation der unregelmäßigen Formen weniger stark gefestigt und ein Wechsel in die typenfrequente Klasse wird wahrscheinlicher. Typenfrequenz gibt somit bei konkurrierenden Flexionsklassen die Variationsrichtung vor: Mitglieder der typeninfrequenten Klasse wechseln zur typenfrequenten. Tokenfrequenz kann innerhalb der typeninfrequenten Klasse einen konservierenden Effekt haben: Mitglieder mit hoher Tokenfrequenz bleiben in der typeninfrequenten Klasse. Tokenfrequenz kann neben der Konservierung auch zu Autonomie und Reduktion führen: Häufige Elemente können sich von der Ursprungskonstruktion lösen und häufig zusammen auftretende Elemente können als ein ein-

⁶⁹Im Sinne der Form-Schematizität lässt sich diese Konstellation auch als Assoziation zwischen mehreren Formen mit unterschiedlichem Schematizitätsgrad und einer Funktion fassen.

⁷⁰Produktivität hängt dabei aber nicht allein an Typenfrequenz, sondern ist zusätzlich von der Variabilität der Vertreter einer Kategorie und der Ähnlichkeit eines potentiellen neuen Vertreters zu bereits bestehenden Exemplaren in der Kategorie abhängig. Alle drei Faktoren (Typenfrequenz, Variabilität, Ähnlichkeit) fasst Goldberg (2019: 62–63) unter dem Prinzip der Abdeckung.

⁷¹Mit Ausnahme von Form-Schemata, die produktiv sein können, wenn der potentielle Neuzugang den bereits vorhandenen Exemplaren ähnelt (Bybee 2010: 6, Goldberg 2019: 62–63).

2.4 Zusammenfassung: Variation aus gebrauchsbasierter Perspektive

ziges Element wahrgenommen (*chunking*) und in der Folge phonologisch verschmolzen werden (Bybee & Thompson 1997: 378–384).

Die hier beschriebenen Frequenzeffekte lassen sich gut in einem konnektionistisches Modell vereinen, in dem Wortformen untereinander verbunden sind. Hierbei liegt exemplarbasiertes Lernen zugrunde. Regelmäßiges Flexionsverhalten kann somit als eine Generalisierung gesehen werden, die anhand von vielen Exemplaren mit hoher Variabilität vollzogen wird. Den Inhalt der Generalisierung stellen dabei die wiederkehrenden Eigenschaften der Exemplare dar, wie bspw. die Form [Dentalsuffix] oder die Funktion [+Vergangenheit] (Goldberg 2019: 67–68). Das konnektionistische Modell verdeutlicht zudem, dass eine Operationalisierung von Tokenfrequenz als reine Wortwiederholung nicht ausreicht, um deren Einfluss zu modellieren (Ellis 2002a: 150–152, Divjak & Caldwell-Harris 2015: 57–58), sondern dass bspw. auch die Frequenz anderer ähnlicher Wörter betrachtet werden muss und die Frequenz, mit der Wörter gemeinsam auftreten.

Frequenzeffekte lassen sich als priore Wahrscheinlichkeit in ein bayesianisches Modell integrieren. Dies verdeutlicht das bayesianisch konzipierte statistische Vorkaufsrecht nach Goldberg (2019: 74–94): Frequente Strukturen haben eine hohe priore Wahrscheinlichkeit aufzutreten und können daher konkurrierende Strukturen mit geringerer Frequenz statistisch ausstechen (Goldberg 2019: 74–94).⁷² Das statistische Vorkaufsrecht kann dabei als Erklärung für das Zusammenspiel aus Typen- und Tokenfrequenz bei Konjugations- und Deklinationsklassen herangezogen werden: Die Formen der typenfrequenten, regelmäßigen Klasse sind theoretisch für jedes Verb bzw. Substantiv möglich. Gehört ein Verb oder Substantiv der typeninrequenten Klasse an und ist tokenfrequent, haben die unregelmäßigen Formen aber aufgrund der Tokenfrequenz eine höhere priore Wahrscheinlichkeit als regelmäßige Formen und können diese statistisch ausstechen. Bei geringer Tokenfrequenz greift das Vorkaufsrecht nur noch eingeschränkt, da die priore Wahrscheinlichkeit für die unregelmäßigen Formen aufgrund der geringen Frequenz herabgesenkt ist. Im empirischen Teil der Arbeit wird der Einfluss von Tokenfrequenz auf Variation anhand von starken Verben mit unterschiedlicher Tokenfrequenz überprüft.

Neben der Frequenz beeinflusst Prototypizität kognitive Kategorien und damit auch Variation. Prototypisch organisierte Kategorien sind Ergebnis eines exemplarbasierten Lernens: Zunächst werden einzelne Exemplare einer Kategorie

⁷²Dies ist nur der Fall, wenn beide Strukturen mit gleicher Funktion und im gleichen Kontext genutzt werden. Ist die infrequente Struktur mit einer anderen Funktion oder einem anderen Kontext assoziiert als die frequente Struktur, greift das statistische Vorkaufsrecht nicht, da die Strukturen nicht miteinander konkurrieren (Goldberg 2019: 74–94).

2 Variation aus gebrauchsbasierter Perspektive

durch Gebrauch erlernt, Verbindungen zwischen ihnen gezogen und die Kategorie anschließend anhand der Gemeinsamkeiten der Exemplare abstrahiert (Goldberg 1998: 206–209, Ross & Makin 1999: 216–222). Als Ankerpunkt für eine analogische Ausweitung der Kategorie dienen vorrangig prototypische Exemplare: Exemplare mit ähnlichen Eigenschaften wie die prototypischen Exemplare werden als der Kategorie zugehörig betrachtet (Goldberg 2006: 89, Ellis 2016: 242, Goldberg 2019: 51–73).⁷³ Hierdurch entsteht die Familienähnlichkeit prototypisch organisierter Kategorien. Aufgrund der Familienähnlichkeit ergeben sich Prototypeneffekte: Vertreter mit typischen Eigenschaften werden bspw. schneller prozessiert, als Erste genannt und lassen sich durch den Kategoriennamen primen (Rosch 1975b, Kleiber 1993: 38–39, Taylor 1995: 45, Ellis u. a. 2014: 85).

Prototypizität fördert und hemmt Variation: Prototypische Vertreter sind in ihrer Verwendung stabil, während bei peripheren Vertretern Variation wahrscheinlicher ist. Dies liegt an der probabilistischen Struktur prototypisch organisierter Kategorien: Je mehr Eigenschaften ein Element hat, die als typisch für die Kategorie wahrgenommen werden, desto wahrscheinlicher ist es, dass das Element als Teil der Kategorie wahrgenommen wird. Daher ist für ein Element auch eine stabile Verwendung wahrscheinlicher, wenn es mehr Eigenschaften der Kategorie erfüllt. Hier wird wiederum der Einfluss von Frequenz deutlich. Dieser zeigt sich auch darin, dass Vertreter mit typischen Eigenschaften einer Kategorie frequenter sind als Vertreter mit peripheren Eigenschaften (Ellis u. a. 2014: 85). Tokenfrequenz scheint also die Vertreter mit typischen Eigenschaften zu stützen. Außerdem werden die typischen Eigenschaften durch Typenfrequenz hervorgebracht: Erst wenn sich viele Vertreter bestimmte, distinkte Eigenschaften teilen, kann daraus eine Kategorie mit prototypischer Struktur abstrahiert werden. Dabei ist die *cue validity* der Eigenschaften relevant: Je mehr Eigenschaften eines Elements nur in einer Kategorie vorkommen, aber nicht in einer anderen, desto eher kann man das Element eindeutig einer Kategorie zuordnen (Rosch & Mervis 1975: 575, Kleiber 1993: 52).

Die prototypische Struktur von Kategorien findet sich auch in Flexionsklassen: Das Verhalten von Flexionsklassen ist prototypisch und graduell organisiert. Es lässt sich daher nutzen, um Vorhersagen über die Reihenfolge zu machen, in der Schwankungen auftreten. Dabei werden zunächst periphere Eigenschaften der Flexion der typeninfrequenten Klasse aufgegeben, dann prototypische. Auch hier zeigt sich, dass Variation in der Peripherie wahrscheinlicher ist als für Prototy-

⁷³Dies schließt eine Erweiterung der Kategorie anhand von anderen Vertretern nicht aus, die prototypischen Vertreter haben nur eine höhere Wahrscheinlichkeit, als Ausgangspunkt für Analogien zu fungieren.

2.4 Zusammenfassung: Variation aus gebrauchsbasierter Perspektive

pen. Der Einfluss von Prototypizität auf Variation wird im empirischen Teil der Arbeit anhand der Selektion von *haben* und *sein* überprüft.

Neben Frequenz und Prototypizität nimmt (Form-)Schematizität Einfluss auf Variation. Dabei ist zwischen zwei Arten von Schema und Schematizität zu unterscheiden. Der Terminus *Schematizität* meint den Abstraktheitsgrad von Konstruktionen. Als Schemata werden Konstruktionen bezeichnet, die eine hochschematische Form aufweisen (Booij 2010: 5, Bybee 2010: 80). Unter dem Terminus *Form-Schematizität* werden Konstruktionen verhandelt, bei denen zwei (oder mehr) Formen mit einer Funktion verknüpft sind. Eine Form ist dabei hochschematisch, die anderen teilschematisch oder nicht-schematisch.⁷⁴ Eine solche Verteilung lässt sich bei Flexionsklassen beobachten: Die hochschematische Form stellt dabei die typenfrequente, regelmäßige Klasse dar. Aufgrund der vielen und vielfältigen Mitglieder kann ein regelmäßiges Flexionverhalten abgeleitet werden (Goldberg 2019: 67–68). Die Assoziation aus einer teilschematischen Form und einer Funktion wird als Form-Schema bezeichnet. Form-Schemata sind typisch für typeninfrequente, unregelmäßige Klassen.⁷⁵ Sie entstehen wie die hochschematische Form durch exemplarbasiertes Lernen. Da das Cluster jedoch aus sehr ähnlichen Lexemen besteht, sind die über die Lexeme hinweg beobachtbaren Eigenschaften Teil der Generalisierung, sodass eine teilschematische Form abstrahiert wird (Goldberg 2019: 121). Ergebnis des exemplarbasierten Lernens ist die prototypisch organisierte Form von Form-Schemata: Je mehr Eigenschaften ein Element mit der Form des Form-Schemas gemein hat, desto wahrscheinlicher ist die Assoziation mit der Funktion und desto eher kann das Form-Schema die hochschematische Form statistisch ausstechen (Rumelhart 1980, Bybee & Moder 1983, Köpcke 1993). Das Prinzip der prototypisch aufgebauten Kategorien zeigt sich somit auch bei Form-Schemata. Aufgrund der prototypisch organisierten Form sind Form-Schemata probabilistisch und können deshalb gut bayesianisch modelliert werden. Zudem lassen sich Form-Schemata als Effekt der Typenfrequenz fassen, da viele Vertreter benötigt werden, um die teilschematische Form zu abstrahieren (Bybee 1995: 430, 2010: 67–69, Goldberg 2019: 67).

Form-Schemata haben einen variationshemmenden Effekt. Sie stechen die konkurrierende, hochschematische Form statistisch aus und stützen daher Mitglieder typeninfrequenter Flexionsklassen. Dieser Effekt wird im empirischen

⁷⁴Hierbei ist jeweils von einem Makro-Schema auszugehen: Eine Verbform (z. B. Dentalsuffix) ist mit der Funktion [+Vergangenheit] verknüpft und in der Folge das gesamte Lexem mit einem bestimmten (in diesem Beispiel regelmäßigen) Flexionsverhalten.

⁷⁵Form-Schemata sind für die typenfrequente Klasse jedoch nicht ausgeschlossen, siehe hierzu Abschnitt 3.2.3, in dem Form-Schemata starker Maskulina diskutiert werden.

2 Variation aus gebrauchsbasierter Perspektive

Teil der Arbeit anhand des Form-Schemas schwacher Maskulina überprüft. Umgekehrt kann ein Form-Schema Variation auslösen, wenn die Verknüpfung von hochschematischer Form und Funktion für Elemente gilt, die der teilschematischen Form des Form-Schemas ähneln wie bei *winken* (Köpcke 1999). In diesem Fall kann es zu Schwankungen weg von der hochschematischen Form und hin zu der teilschematischen kommen (*gewinkt* > *gewunken*).

Nicht nur Form-Schemata beeinflussen Variation, sondern auch separate Schemata, deren Funktionen prototypisch organisiert sind (z. B. [*haben*] und [+Transitivität, –Telizität, –Bewegungsemantik] sowie [*sein*] und [–Transitivität, +Telizität oder +Bewegungsemantik], siehe hierzu ausführlich Abschnitt 3.3.2). An den prototypischen Enden der Funktionen ist die Assoziation eindeutig, zu Schwankungen kann es jedoch im Übergangsbereich zwischen den Funktionen kommen. Hier könnten beide Schemata greifen und somit beide Formen genutzt werden. Dies wird in der vorliegenden Arbeit als Prototypizitätseffekt verhandelt, da die Schwankung nicht zwischen zwei Formen, die mit einer Funktion verbunden sind, stattfindet, sondern zwischen zwei Schemata.

Insgesamt zeigt das Kapitel, dass Frequenz, Prototypizität und (Form-)Schematizität grundlegenden Einfluss auf Variation ausüben. Der Einfluss der Frequenz ist dabei am prominentesten. Er zeigt sich deutlich im *entrenchment* tokenfrequenter Elemente, das zu einer stabilen mentalen Repräsentation und Verwendung der Elemente führt. Bei Flexionsklassen reicht ein Blick auf Tokenfrequenzen nicht aus, da hier Typen- und Tokenfrequenz relevant sind: Variation ist bei tokeninfrequenten Mitgliedern der typeninfrequenten Klasse wahrscheinlich, da die Formen der typeninfrequenten Klasse aufgrund der geringen Tokenfrequenz weniger gefestigt sind und daher die konkurrierenden Formen der typenfrequenten Klasse nicht mehr statistisch ausstechen können.

Prototypizität beeinflusst Variation, da prototypische Vertreter einer Kategorie i. d. R. stabil genutzt werden, während periphere Mitglieder Variation aufweisen können. Dies zeigt sich bspw. bei Schemata wie der Auxiliarselektion von *haben* und *sein*, deren Funktionen prototypisch verknüpft sind. Zudem ermöglicht es Prototypizität, Flexionsklassen graduell mit peripheren und prototypischen Eigenschaften zu modellieren. Auf diese Weise können Schwankungen von der typeninfrequenten zur typenfrequenten Flexionsklasse prognostiziert werden. Diese beginnen mit der Aufgabe von peripheren Eigenschaften und können bis hin zur Aufgabe von prototypischen Eigenschaften reichen.

Der Einfluss der Prototypizität auf Variation ist auch essentiell für Form-Schemata. Form-Schemata stützen als relativ typenfrequente, wenig variable Cluster in der typeninfrequenten Klasse das Flexionsverhalten der typeninfrequenten Klasse. Auf diese Weise kann die hochschematische Form der typenfrequenten

2.4 Zusammenfassung: Variation aus gebrauchsbasierter Perspektive

Klasse auch bei relativ geringer Tokenfrequenz statistisch ausgestochen werden. Form-Schemata sind probabilistisch aufgebaut, sodass prototypische Vertreter eines Form-Schemas stabil flektieren, während Schwankungen bei peripheren Vertretern hin zum regelmäßigen, typenfrequenten Flexionsverhalten wahrscheinlich sind. Der Einfluss der Frequenz, Prototypizität und Form-Schematizität und ihr Zusammenwirken wird im nächsten Kapitel in Hinblick auf Variation in der Konjugation, Deklination und der Auxiliarselektion bei *haben* und *sein* näher betrachtet.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Im folgenden Kapitel wird der Einfluss der Faktoren Frequenz, Prototypizität und Form-Schematizität auf Variation in der Konjugationsklasse (Abschnitt 3.1), in der Deklinationsklasse (Abschnitt 3.2) und in der Auxiliarselektion (Abschnitt 3.3) diskutiert. In allen Abschnitten wird das Variationsphänomen zunächst vorgestellt und anschließend auf die Einflussfaktoren hin untersucht. Innerhalb des empirischen Teils der Arbeit wird jeweils ein Einflussfaktor anhand eines Variationsphänomens untersucht. Der Einfluss von Frequenz wird anhand der Variation in der Konjugation überprüft, der Einfluss von Form-Schematizität anhand der Variation in der Deklination und der Einfluss von Prototypizität anhand der Variation in der Selektion von *haben* und *sein*. Dennoch werden auch die anderen Faktoren umfassend diskutiert, damit die Relevanz jedes einzelnen Faktors sowie das Zusammenspiel der Faktoren herausgearbeitet werden können.

3.1 Variation in der Konjugationsklasse

Die deutschen Verben teilen sich in zwei große Konjugationsklassen: schwach und stark. Die schwache Konjugation mit dem Dentalsuffix *-t* ist dabei jünger als die starke und weist deutlich mehr Mitglieder auf (Szczepaniak 2011a: 111–112, siehe Abschnitt 3.1.1). Die schwache Konjugationsklasse weist ein regelmäßiges und additives Verhalten auf. Die Vergangenheitstempora aller schwachen Verben werden mit dem Dentalsuffix *-t* markiert (*lachte, gelacht*) (Bittner 1985: 54–55, Nowak 2013: 178). Im Gegensatz dazu zeichnet sich die starke Konjugation durch ein modulatorisches Verfahren aus: Die Vergangenheitstempora werden markiert, indem der Stammvokal der starken Verben abgelautet wird. Das modulatorische Verfahren schließt aber Addition nicht aus, so wird das Partizip II mithilfe einer Kombination aus additivem (*ge...-en*) und i. d. R. modulatorischem Verfahren (Ablaut)¹ gebildet. Hinsichtlich der Ablaute ist hervorzuheben, dass

¹Verben der Ablautreihen 5 bis 7 weisen im Partizip II keinen Ablaut auf (*fahren, fuhr, gefahren*) bzw. ist der Ablaut im Partizip II identisch mit dem Infinitivstammvokal.

3 Variation in Konjugation, Deklination und Auxiliarselektion

nicht alle starken Verben nach dem gleichen Muster ablauten, vielmehr bestehen verschiedene Ablautmuster wie bspw. *singen*, *sang*, *gesungen* und *schlafen*, *schliefe*, *geschlafen*. Je nach Zählung existieren im Deutschen 39 bis 52 Vokalalternanzen.² Neben der Dichotomie zwischen Ablaut und Dentalsuffix bestehen in den Vergangenheitstempora nur kleine Unterschiede zwischen den Verbklassen. So weisen die starken Verben im Partizip II das Zirkumfix *ge-...-en* auf, während die schwachen das Partizip II mithilfe von *ge-...-t* markieren.

Nicht nur hinsichtlich der Präterital- und Perfekt-II-Formen existieren Unterschiede zwischen der starken und schwachen Konjugation, sondern auch in der Bildung des Imperativs und im Präsensparadigma: Während schwache Verben den Imperativ auf *-e* oder endungslos markieren (*lache!/lach!*), kann der Stammvokal bei *e*-haltigen starken Verben zu /ɪ/ bzw. /i:/ gehoben werden (*gib!/lies!*). Zudem können starke Verben Wechselflexion im Präsens aufweisen (*ich gebe*, *du gibst*, *er/sie gibt*). Imperativhebung und Wechselflexion sind aber nicht für alle starken Verben zu beobachten, dies weist bereits auf eine prototypische Staffellung der Flexionseigenschaften starker und schwacher Verben hin, die in Abschnitt 3.1.2 näher erläutert werden. Einen weiteren Unterschied zwischen starker und schwacher Konjugation macht die Bildung des Konjunktivs II aus. Dieser wird bei starken Verben synthetisch mithilfe von Umlautung des Präteritalstamms geformt (*ich gäbe*), bei schwachen Verben lauten Konjunktiv II und die Indikativformen im Präteritum dagegen gleich (*ich lachte*) (Bittner 1985: 54–55). In beiden Verbklassen kann der Konjunktiv II jedoch auch analytisch gebildet werden (*ich würde lachen*, *ich würde geben*).

Die Unterschiede zwischen den Flexionsparadigmen prototypisch schwacher und prototypisch starker Verben sind in Tabelle 3.1 zusammengefasst.

Neben den starken und schwachen Verben existieren unregelmäßige³ Konjugationsklassen: Hierzu zählen bspw. Verben mit suppletivem Paradigma wie *sein* und *gehen*. Zudem sind Präterito-Präsentia sowie die sogenannten Rückumlautverben unregelmäßige Verben. Die Präterito-Präsentia weisen aufgrund des Bedeutungswechsels von [+Vergangenheit] zu [+Präsens] Ablautbeziehungen im

²Unter die 52 Alternanzen werden auch konsonantische Veränderungen bspw. aufgrund des grammatischen Wechsels (*ziehen*, *zog*, *gezogen*) subsumiert (Nübling 1998: 192, Dammel 2008: 13).

³In diesem Abschnitt wird im Gegensatz zu Kapitel 2 zwischen starker und unregelmäßiger Konjugation unterschieden. Die starke Konjugation kann zwar im Gegensatz zur schwachen Konjugation als unregelmäßig beschrieben werden, im Vergleich zu den hier vorgestellten unregelmäßigen Verben ist das Konjugationsverhalten starker Verben aber als regelmäßig einzuordnen, da es auf mehr Verben generalisiert werden kann als das Verhalten einzelner unregelmäßiger Verben. Daher wird in diesem Abschnitt mit der Unterscheidung zwischen schwach, stark und unregelmäßig gearbeitet.

3.1 Variation in der Konjugationsklasse

Tabelle 3.1: Unterschiede in der Flexion schwacher und starker Verben

	schwach <i>lachen</i>	stark <i>geben</i>
Imperativ	<i>lache!</i>	<i>gib!</i>
Präsensparadigma	<i>ich lache</i>	<i>ich gebe</i>
	<i>du lachst</i>	<i>du gibst</i>
	<i>er/sie/es lacht</i>	<i>er/sie/es gibt</i>
Präteritum	<i>ich lachte</i>	<i>ich gab</i>
Konjunktiv II	<i>ich lachte</i>	<i>ich gäbe</i>
Partizip II	<i>gelacht</i>	<i>gegeben</i>

Präsensparadigma (*wissen, ich weiß*) auf (Bergmann u. a. 2016: 99–101), die sogenannten Rückumlautverben zeichnen sich durch einen Umlaut im Präsensparadigma aus, der jedoch nicht im Präteritum und Perfekt zu finden ist (*brennen, brannte, gebrannt*) (Bittner 1985: 60–61, Bergmann u. a. 2016: 98–99).

Im Folgenden wird die Variation in der Konjugation zwischen starkem und schwachem Konjugationsverhalten im Vordergrund stehen. Hierbei zeigen sich in allen Tempora und Modi Tendenzen, starke Konjugationseigenschaften zugunsten schwacher Eigenschaften abzubauen: Die Imperativhebung wird aufgegeben (*tret(e)!* statt *tritt!*) und die Wechselflexion im Präsens (*er flechtet* statt *er flicht*). Zudem weicht das modulatorische Verfahren (Ablaut) im Präteritum dem additiven (Dentalsuffix) (*ich backte* statt *ich buk*). Dasselbe geschieht im Partizip II, hier wird zudem *-en* durch *-t* ersetzt (*gegärt* statt *gegoren*). Neben dem Wechsel von modulatorisch zu additiv ist ein Wechsel des ursprünglichen Präteritalvokals hin zu /o:/ bzw. /ɔ/ zu beobachten: *ich schor* statt *ich schar*, *ich schwomm* statt *ich schwamm*. Abseits des Wechsels von stark zu schwach ist zudem vereinzelt auch die umgekehrte Richtung, ein Wechsel von schwach zu stark, zu beobachten. Dies ist bspw. bei *winken* der Fall, das das Partizip II ursprünglich schwach (*gewinkt*) gebildet hat und nun neben der schwachen die starke Form *gewunken* aufweist (Duden 2020).

In den folgenden Abschnitten wird der Einfluss von Frequenz, Prototypizität und Form-Schematizität auf die Variation in der Konjugation diskutiert. Abschnitt 3.1.1 untersucht den Einfluss der Typen- und Tokenfrequenz, Abschnitt 3.1.2 macht anhand der prototypischen Staffelung von starken Konjugationseigenschaften Prognosen über die Reihenfolge von Schwankungen, die beim

3 Variation in Konjugation, Deklination und Auxiliarselektion

Wechsel von der starken zur schwachen Flexionsklasse beobachtet werden können. Abschnitt 3.1.3 diskutiert Form-Schemata, die sich stabilisierend auf die starke Flexion auswirken. Abschließend beleuchtet Abschnitt 3.1.4, wie die Faktoren sich gemeinsam auf die Variation in der Konjugation auswirken. Im empirischen Teil der Arbeit wird dann der Einfluss von Tokenfrequenz auf die Variation zwischen starker und schwacher Flexion psycholinguistisch untersucht.

Für die theoretische Diskussion der Einflussfaktoren werden Ergebnisse von Korpusanalysen und psycholinguistischen Studien betrachtet. Die Psycholinguistik hat sich in Hinblick auf starke und schwache Flexion vorrangig mit Frequenzeffekten befasst. Dabei wurde aber nicht auf eine mögliche Variation fokussiert, sondern untersucht, ob Frequenzeffekte sowohl bei schwachen als auch bei starken Verben auftreten. Letzteres wird nicht in diesem Abschnitt erörtert, siehe hierfür die Diskussion um die *past tense debate* in Abschnitt 2.1.3. In diesem Abschnitt werden nur psycholinguistische Studien berücksichtigt, die Rückschlüsse auf den Einfluss von Frequenz auf Variation in der Konjugation zulassen.

3.1.1 Frequenz

Starke und schwache Verben unterscheiden sich in ihrer Typen- und Tokenfrequenz: Rund 170 Verben⁴ weisen ein starkes Konjugationsverhalten auf (Augst 1975: 258, Dammel & Nowak 2014: 310), während rund 4.000 Verben schwach konjugiert werden (Nübling 1998: 195). Die schwachen Verben stellen somit mit 96 % eindeutig die typenfrequenterere Klasse dar.⁵ Das schwache Konjugationsverfahren ist außerdem produktiv: Neu entlehnte Verben werden schwach konjugiert (*googeln, ich googelte*) (Nübling 1998: 195). Zudem wird die schwache Konjugation im Spracherwerb übergeneralisiert (Rumelhart & McClelland 1986: 219–220). Das starke Konjugationsverfahren ist hingegen nicht mehr produktiv (siehe jedoch Neuzugänge aufgrund von Form-Schemata in Abschnitt 3.1.3) und hat diachron Mitglieder verloren: Im Germanischen existierten mit 381 mehr als doppelt so viele starke Verben wie im Neuhochdeutschen (Dammel & Nowak 2014: 310). Der Abbau der starken Verben ist dabei auch in anderen germanischen Sprachen wie bspw. Schwedisch, Niederländisch, Friesisch und Luxemburgisch zu beobachten (siehe hierzu Dammel u. a. 2010, Nowak 2010a, Schmuck 2010, Dammel & Nowak 2011, 2014, Nowak 2015). Die Produktivität der schwachen Flexion ergibt sich aus der Typenfrequenz: Aufgrund der vielen variablen Mitglieder lässt sich eine

⁴Die genauen Zahlen schwanken hier. Augst (1975: 258) zählt 169 starke Verben, Dammel & Nowak (2014: 310) nennen 40 Jahre später noch 165 starke Verben.

⁵Die unregelmäßigen Verben machen nach Augst (1975: 258) 0,5 % des Wortschatzes aus und sind somit in Hinblick auf die Typenfrequenz zu vernachlässigen.

simple Regel ableiten (Anhängen von *-t*), um Tempus anzuzeigen. Aufgrund der leichten Generalisierbarkeit wird dieses Verfahren auf neue Verben angewandt (Goldberg 2019: 67–68, siehe hierzu Abschnitt 2.1.3 sowie Abschnitt 2.4).

Auf Ebene der Token sind die Frequenzverhältnisse umgekehrt: Starke Verben sind tokenfrequenter als schwache Verben (Augst 1975: 257). Dies ist nicht verwunderlich, da starke Verben Grundtätigkeiten ausdrücken: 63 % der Verben aus dem Grundwortschatz werden stark flektiert (Augst 1975: 258). Im Althochdeutschen gehörten 70 % der Verben aus dem Grundwortschatz der starken Flexion an (Augst 1975: 258). Damit ist der Anteil starker Verben innerhalb des Grundwortschatzes vom Althochdeutschen (Ahd.) bis zum Neuhochdeutschen (Nhd.) stabil geblieben, obwohl sich die Anzahl der starken Verben um die Hälfte reduziert hat. Hieraus lässt sich ableiten, dass nur die infrequenten starken Verben zur schwachen Klasse wechseln (Augst 1975, Nübling 1998: 195, Carroll u. a. 2012⁶). Darunter ist bspw. *bellen*, das ursprünglich auch ‚röhren‘ oder ‚grunzen‘ bedeutete (Berlin-Brandenburgische Akademie der Wissenschaften 2019, DWDS) und vermutlich aufgrund der Bedeutungsverengung zu ‚Geräusche produzieren, die ein Hund von sich gibt‘ seltener gebraucht wird.

Die aktuellen Schwankungsfälle zur schwachen Flexion sind ebenfalls niedrigfrequent (Nowak 2016: 149): Dazu zählen bspw. *fechten* (*fichtst/fechtest*), *melken* (*molk/melkte*) und *gären* (*gegärt/gegoren*). Die Schwankungen hin zur schwachen Flexion treten dabei zunächst im Imperativ, dann im Präsens, gefolgt vom Präteritum und schließlich im Partizip II auf (Bittner 1996: 78–79, siehe hierzu genauer Abschnitt 3.1.2 zur Prototypizität der starken Konjugation).

Der Zusammenhang zwischen Tokenfrequenz und stabiler starker bzw. unregelmäßiger Flexion zeigt sich nicht nur in Schwankungsfällen, sondern auch darin, dass nur vier der 25 gebräuchlichsten⁷ Verben schwach⁸ konjugiert werden (Harnisch 1988: 430). Zudem verdeutlicht die Irregularisierung frequenter regelmäßiger Verben, dass Frequenz ein Indiz für Unregelmäßigkeit sein kann.

⁶Carroll u. a. (2012) stellen in ihrer Untersuchung fest, dass im Englischen wie im Deutschen infrequente starke Verben zur schwachen Flexion wechseln. Dabei ist im Englischen eine lineare Entwicklung zu beobachten. Dies ist im Deutschen nicht der Fall. Hier ist ab dem Frühneuhochdeutschen ein deutlicher Anstieg an Verben zu verzeichnen, die zur schwachen Flexion wechseln (Carroll u. a. 2012: 168–169).

⁷Die Gebrauchshäufigkeit basiert auf dem Häufigkeitswörterbuch zur gesprochenen Sprache von Ruoff (1981). Die Daten beziehen sich auf Südwestdeutschland.

⁸Harnisch (1988) nutzt den Terminus *schwach* nicht, sondern unterscheidet reguläre und irreguläre Konjugation. Dabei definiert er nicht, was er unter *irregulär* versteht. Da er vor dem Hintergrund der Natürlichkeitsmorphologie argumentiert, ist davon auszugehen, dass er das für die Natürlichkeitsmorphologie reguläre schwache Konjugationsverfahren mit im Sinne der Natürlichkeitsmorphologie irregulären Konjugationsverfahren kontrastiert und nicht weiter zwischen starken und unregelmäßigen Verbklassen unterscheidet.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Haben war ursprünglich regelmäßig und weist nun ein Suppletivparadigma auf, da es mit der Grammatikalisierung zum Auxiliar an Frequenz gewonnen hat: ahd. *habēn, habēta* > nhd. *haben, hatte* (Nübling 2000: 17–18, Nowak 2013: 174).

Der Einfluss der Tokenfrequenz auf Flexion wird neben dem Klassenwechsel infrequenter starker Verben zu schwachen auch innerhalb der starken Verben deutlich. Nowak (2013) zeigt in einer diachronen Korpusuntersuchung, dass einige infrequente Verben einen Vokalwechsel von einem ursprünglichen Präteritalvokal zu /o:/ bzw. /ɔ/ (*schar* > *schor*, *quall* > *quoll*) vollziehen. Der Vokalwechsel führt zu dem Ablautmuster *x-o-o*, in dem ein beliebiger Stammvokal in den Vergangenheitstempora auf /o:/ bzw. /ɔ/ abgelautet wird. Nowak (2016: 127–128) zufolge bildet sich das Ablautmuster *x-o-o* im Zuge des präteritalen Numerusausgleichs⁹ im Frühneuhochdeutschen (Frnhd.) basierend auf der Ablautreihe 2¹⁰ (*biegen - bog - gebogen*) aus (Nowak 2013: 170), in der sich ebenfalls /o:/ in den Vergangenheitstempora findet. Anders als beim Ablautmuster *x-o-o*, das auch als achte Ablautreihe bezeichnet wird, ist in der Ablautreihe 2 jedoch /i:/ als Stammvokal im Infinitiv vorgesehen. Die Analogiebildung wurde vermutlich dadurch erleichtert, dass viele der Verben, die *x-o-o* angenommen haben, phonologische Ähnlichkeiten zu Verben der Ablautreihe 2 aufweisen; so enden *weben* und *heben* wie *schieben* auf *-ben* (Nowak 2018: 166, siehe dazu genauer Abschnitt 3.1.3 zu Form-Schemata bei starken Verben).

Das Ablautmuster *x-o-o* kann einerseits als Sammelbecken für infrequente starke Verben fungieren (Nowak 2013: 183). So weisen *scheren* und *heben* bis heute stabil *x-o-o* auf. Andererseits kann *x-o-o* eine Übergangsstation zur schwachen Flexion darstellen: *Pflegen* und *bellen* wechselten zunächst zu *x-o-o*, bevor sie zur schwachen Flexion übergingen. Derzeit sind Schwankungen zwischen *x-o-o* und schwacher Flexion bspw. bei *glimmen* und *weben* zu beobachten (*glomm/glimmte; geglommen/geglimmt*).¹¹

Um den Einfluss der Tokenfrequenz auf das Vorkommen des Vokalmusters *x-o-o* zu untersuchen, vergleicht Nowak (2013: 176–178) die Flexion starker Ver-

⁹Ausführlich zum Numerusausgleich siehe Nübling (1998), die im Numerusausgleich ein gemeinsames Wirken von Tokenfrequenz und Relevanzprinzipien nach Bybee (1985) sieht: Der Numerus ist eine weniger relevante Verbkategorie als Tempus, sodass zu erwarten ist, dass eine gesonderte Markierung abgebaut wird. Hier spielt jedoch auch Tokenfrequenz eine Rolle: So wiesen tokenfrequente Verben wie bspw. *werden* noch lange Formen ohne Numerusausgleich auf (Nübling 1998: 198).

¹⁰Eine Übersicht über die Ablautreihen im Deutschen findet sich in Tabelle B.1.

¹¹Bei einigen Verbvarianten lässt sich eine Tendenz zur semantischen Ausdifferenzierung beobachten. So scheint bei *glimmen* der Gegensatz zwischen konkreter (*die Asche glimmte*) und übertragener Verwendung (*es glomm noch etwas Hoffnung ihn ihr*) mit starker bzw. schwacher Konjugation assoziiert zu sein; siehe hierzu genauer Nowak (2011).

3.1 Variation in der Konjugationsklasse

ben mit unterschiedlichen Tokenfrequenzen vom Mittel- bis zum Neuhochdeutschen. *Gewinnen* war im Mittelhochdeutschen (Mhd.) und im Frnhd. frequent und ist es bis heute (325 Belege pro Million Token). Wenig überraschend weist *gewinnen* ein stabiles Ablautverhalten auf. *Glimmen* und *klimmen* waren hingegen schon im Mhd. und Frnhd. niedrigfrequent und sind es bis heute (unter einem bzw. vier Belegen pro Million Token). Beide Verben haben im Frnhd. zu *x-o-o* gewechselt. *Rinnen*, *schwimmen*, *sinnen* und *spinnen* sind heute mit zwei bis 17 Belegen pro Million Token zwar teilweise ähnlich infrequent wie *glimmen* und *klimmen*, waren im Frnhd. mit 12 bis 56 Belegen pro Million Token aber weit frequenter. Diese Verben zeigen gegenwärtig Schwankungen im Ablaut (*rann/ronn;schwamm/schwomm;sann/sonn;spann/sponn*) (Nowak 2013: 176–178).

Generell stellt Nowak (2016: 138) für Verben, die zu *x-o-o* wechseln, eine geringe Tokenfrequenz fest: Verben der Ablautreihen 3 bis 6 mit Tendenz zu *x-o-o* haben im Mhd. eine Tokenfrequenz von 631 Belegen pro Million Token, während Verben dieser Ablautreihen mit stabilem ursprünglichen Ablaut mit 3.274 Belegen pro Million Token weit frequenter sind. Dasselbe Bild ergibt sich für das Neuhochdeutsche: Verben mit Tendenz zu *x-o-o* sind mit 1.144 Belegen pro Million Token weit weniger frequent als Verben mit stabilem ursprünglichen Ablaut mit 32.089 Belegen pro Million Token (Nowak 2016: 138).

Zusätzlich zu den heutigen Frequenzverhältnissen verdeutlichen Verben den Einfluss der Tokenfrequenz, die zum Nhd. hin starke Frequenzverluste zu verzeichnen haben. Hierzu zählen *heben* und *schwören*. Beide Verben gehören ursprünglich der Ablautreihe 6 an, sind aber im Frnhd. zu *x-o-o* gewechselt (*heben*, *hob*, *gehoben*). *Heben* macht 0,4 % des mhd. Korpus (Mittelhochdeutsche Begriffsdatenbank) aus, aber nur 0,04 % des nhd. Korpus (Projekt Deutscher Wortschatz). Noch drastischer ist der Frequenzverlust bei *schwören*, das von 0,02 % auf 0,0006 % fällt (Nowak 2016: 139–140). Auch der durchschnittliche Anteil an *x-o-o*-Verben ist im nhd. Korpus mit 0,0005 % verschwindend gering. Starke Verben mit stabilem ursprünglichen Ablaut kommen dagegen immerhin auf 0,2 %.

Nowak (2013: 178) sieht in der Herausbildung des Ablautmusters *x-o-o* eine Regularisierung der starken Verben: *x-o-o* ähnelt trotz des modulatorischen Verfahrens dem schwachen Konjugationsverfahren, da es wie schwache Verben Präsens und Vergangenheitstempora kontrastiert. Zudem weisen *x-o-o*-Verben keine Imperativhebung und keine Wechselflexion auf, sodass sie sich nur in der Bildung von Präterital- und Perfekt-II-Formen von den schwachen Verben unterscheiden und damit nur in der Kategorie TEMPUS, der in der Relevanzskala nach Bybee (1985: 33–37) eine höhere Relevanz zukommt als die Kategorien MODUS und PERSON (Nübling 1998: 192).

3 Variation in Konjugation, Deklination und Auxiliarselektion

Die Vokale /o:/ und /ɔ/ bieten sich zur Markierung der Vergangenheitstempora an, da von den starken Verben nur *stoßen* und *kommen* /o:/ bzw. /ɔ/ als Infinitivstammvokal enthalten (Nowak 2016: 136). Zudem sind viele Vokale mit dem Ablaut auf /o:/ bzw. /ɔ/ kompatibel: Nowak (2016: 136) zeigt, dass – bis auf /o:/ bzw. /ɔ/ selbst – nur die Infinitivstammvokale /u:/ und /a:/ nicht mit dem Ablaut auf /o:/ oder /ɔ/ kombiniert werden. *Rufen* ist das einzige starke Verb mit /u:/ als Infinitivstammvokal. /a:/ ist Teil des typenfrequenten Musters *ei-i-i* (/a:ɪ-ɪ-ɪ/), das vermutlich einem Wandel zu /o:/ bzw. /ɔ/ im Wege steht (siehe hierzu genauer Abschnitt 3.1.3). Somit eignet sich /o:/ bzw. /ɔ/ als Kontrast zu fast allen Präsensvokalen. /o:/ bzw. /ɔ/ rückt daher in die Nähe eines transparenten und uniformen Tempusmarkers (Nowak 2016: 136–137). Zudem bietet sich /o:/ bzw. /ɔ/ als Präteritalvokal an, da /o:/ bzw. /ɔ/ im Partizip II bereits in den Ablautreihen 2, 3b, 4 und bei einigen Verben der Ablautreihe 3a vorgesehen ist (Nowak 2018: 167).

Neben der Opposition zwischen Präsens und Vergangenheitstempora sieht Nowak (2013: 178–182) in *x-o-o* weitere Regularisierungen auf intra- sowie interparadigmatischer Ebene. Intraparadigmatisch bringt der Vokalwechsel eine Vereinfachung des Ablautmusters mit sich: *schwimmen* gehört ursprünglich der Ablautreihe 3a an und weist somit drei unterschiedliche Vokale auf (*schwimmen*, *schwamm*, *geschwommen*). Der Wechsel von /a/ im Präteritum zu /ɔ/ vereinfacht somit die Ablautalternanz, weil sie von ABC zu ABB reduziert wird. Auch bei Verben mit der Ablautalternanz ABA (*weben*, *wab*, *geweben*)¹² hat der Wechsel des Ablautvokals einen Vorteil, da so Tempus profiliert werden kann: *x-o-o* bietet einen klaren Kontrast zwischen Präsens und Vergangenheitstempora, den die Ablautalternanz ABA nicht bieten kann, da der Präsensstammvokal und der Partizip-II-Vokal identisch sind (Nowak 2013: 179).

Interparadigmatisch führt *x-o-o* ebenfalls zu einer Regularisierung in typenfrequentieller Hinsicht. 28 % der starken Verben weisen im Nhd. das Ablautmuster *x-o-o* auf (Nowak 2013: 181). Damit stellt *x-o-o* das typenfrequenteste Ablautmuster dar. Es hat seine Typenfrequenz vom Mhd., in dem es nur durch die Ablautreihe 2 repräsentiert wurde, zum Nhd. hin verdoppelt. Typenfrequentiell betrachtet ist *x-o-o* somit innerhalb der starken Verben ein reguläres, da häufiges Muster. Ähnlich frequent ist das Ablautmuster *ei-i-i* der Ablautreihe 1 mit 23 % der starken Verben. Es folgt das Ablautmuster *i-a-u* (/ɪ-a-ʊ/) der Ablautreihe 3a mit 11 % (Nowak 2013: 181).¹³ Die Ablautmuster *ei-i-i* und *i-a-u* bilden Form-Schemata, die die

¹²Offensichtlich hat hier nicht nur der Ablaut im Präteritum, sondern auch der Ablaut im Partizip II zu /o:/ gewechselt. Siehe hierzu ausführlich Nowak (2015: 253) sowie Nowak (2018: 166–167), die zeigt, dass das Partizip II vor dem Präteritum /o:/ annahm, da /o:/ im Partizip II frequenter ist als im Präteritum und somit besser als Analogievorlage dienen konnte.

¹³Das Ablautmuster *i-a-u* findet sich nicht nur innerhalb der Flexion von Verben, sondern auch in Reduplikationen wie *Rabimmel-Rabammel-Rabumm*, *Tri-Tra-Trulalla* und *Schnick-Schnack-Schnuck* (Nübling u. a. 2013: 269).

3.1 Variation in der Konjugationsklasse

starke Flexion stützen (siehe hierzu genauer Abschnitt 3.1.3 zu Form-Schemata bei starken Verben).

Fasst man das Ablautmuster *x-o-o* als Teil der Ablautalternanz ABB, werden die typenfrequentiellen Verhältnisse noch deutlicher. ABB hat die meisten Mitglieder innerhalb der starken Verben: 52 % der starken Verben weisen ABB (*biegen, bog, gebogen*) auf, 29 % ABC (*sinken, sank, gesunken*) und nur 19 % ABA (*schlafen, schlief, geschlafen*).¹⁴ Der Wechsel von ABC zu ABB, den bspw. *schwimmen* durchläuft, stellt somit nicht nur eine Vereinfachung des Ablautmusters dar, sondern auch einen Anschluss an die typenfrequente Alternanz ABB.¹⁵

Auch innerhalb der starken Verben stehen Typen- und Tokenfrequenz invers zueinander: ABB ist die typenfrequenteste Alternanz, die darin enthaltenen Verben sind mit einer durchschnittlichen Tokenfrequenz von 70 Token pro Million Wortformen vergleichsweise infrequent. Die Alternanz ABC kommt hingegen auf eine relative Frequenz von 140 und ABA auf 340 Token pro Million Wortformen (Nowak 2015: 165–166). In dieses Bild passt, dass die Tokenfrequenz singulärer Ablautmuster wie *sitzen, saß, gesessen* und *rufen, rief, gerufen* jeweils höher ist als die durchschnittliche Frequenz der Ablautalternanzen, der die Verben angehören (ABC *sitzen, saß, gesessen* bzw. ABA *rufen, rief, gerufen*) (Nowak 2018: 161–162). Die hohe Tokenfrequenz singulärer Ablautmuster verdeutlicht deren stabilisierende Wirkung. Im Gegensatz dazu können Verben, die einem typenfrequenten Ablautmuster (z. B. *x-o-o, ei-i-i*) angehören, eine geringere Tokenfrequenz aufweisen, weil sie durch die Typenfrequenz der Ablautmuster gestützt werden.

Da ABA die typeninrequenteste Ablautalternanz ist, ist davon auszugehen, dass es für tokeninfrequente ABA-Verben wahrscheinlicher ist, zur schwachen Flexion überzugehen als bspw. für tokeninfrequente ABB-Verben, da ABA-Verben weniger gut durch Typenfrequenz in der starken Klasse gehalten werden können (Nowak 2018: 161). Dies ist auch diachron zu beobachten: 56 % der mhd. Verben mit Vokalalternanz ABA sind zum Nhd. schwach geworden, bei ABB und ABC sind es nur 22 % (Solms 1984: 323–324). Dies lässt sich auch dadurch erklären,

¹⁴Hinsichtlich der Ablautmuster ist ein Blick auf weitere germanische Sprachen interessant: Wie das Deutsche stärken auch das Niederländische und das Englische das ABB-Muster. Das Schwedische stärkt hingegen das ABA-Muster, ABB wird gar nicht genutzt (Dammel u. a. 2010: 340–342). Da das Schwedische den aspektuellen Unterschied zwischen Präteritum und Perfekt bewahrt hat, wird das Präteritum frequent genutzt und der Ablautvokal im Präteritum daher verfestigt, siehe hierzu genauer Dammel u. a. (2010).

¹⁵Interessanterweise spielt das Muster *x-o-o* auch im Niederländischen und Jiddischen eine wichtige Rolle in der Variation zwischen starken und schwachen Verben (Nowak 2010a: 437–443, Nowak 2010b, Schäfer 2023: 235–237).

3 Variation in Konjugation, Deklination und Auxiliarselektion

dass ABA-Verben durch wenige Eingriffe in das Paradigma zu schwachen Verben gemacht werden können (Nowak 2016: 141). Aufgrund des Präteritalschwundes im Deutschen reduziert sich die Alternanz ABA auf Vokalidentität zwischen Präsens- und Partizip-II-Vokal (Nowak 2016: 141). Somit unterscheiden sich Verben der Vokalalternanz A(B)A im Partizip II nur noch durch *-en* vom schwachen Konjugationsverfahren. Der Wechsel vom Nasal- zum Dentialsuffix stellt keine große Veränderung dar, weshalb die Flexion direkt von stark zu schwach umgewandelt werden kann (Nowak 2016: 141). Es ist daher zu erwarten, dass ABA-Verben *x-o-o* meiden. Dies scheint auch der Fall zu sein, denn nur vier Verben, die sich *x-o-o* angeschlossen haben, wiesen ursprünglich ABA als Ablautalternanz auf (*gären, wägen, bewegen* und *weben*) (Nowak 2016: 132–133). Die Mehrheit der Verben mit Anschluss an *x-o-o* sind ABC-Verben. Hier bietet sich ein Anschluss an *x-o-o* an, da die Verben bereits /o:/ bzw. /ɔ/ im Perfekt enthalten (Nowak 2018: 167).

Der Einfluss der Tokenfrequenz auf die starke Flexion lässt sich nicht nur diachron erkennen, sondern auch in *lexical decision tasks* und Produktionsexperimenten beobachten. Clahsen u. a. (1997: 224–227) nutzen Partizip-II-Formen starker und schwacher Verben mit geringer und hoher Frequenz in einer *lexical decision task*.¹⁶ Sie stellen fest, dass die Reaktionszeiten für frequente Partizipien starker Verben niedriger sind als für infrequente. Bei den schwachen Verben beeinflusste die Partizipfrequenz die Reaktionszeiten hingegen nicht. Clahsen u. a. (1997: 227) interpretieren dies als Evidenz für das *dual mechanism model*. Wie in Abschnitt 2.1.3 bereits argumentiert wurde, lässt sich der Unterschied aber auch auf das *power law of practice* (Ellis & Schmidt 1998: 307) zurückführen: Da schwache Verben immer gleich gebildet werden, hat die Tokenfrequenz einzelner Partizip-II-Formen keinen Einfluss auf die Reaktionszeit.

Ähnliche Ergebnisse wie Clahsen u. a. (1997) erzielen Clahsen u. a. (2004: 692–964), die Erwachsene sowie Kinder starke und schwache Verben mit niedriger und hoher Partizip-II-Frequenz konjugieren lassen. Dabei war stets der Verbstamm vorgegeben (*Der Frosch hat die Fliege fress*). Im Beispielsatz müsste aus dem Stamm *fress* also das Partizip II *gefressen* geformt werden. Bei dieser Aufgabe waren Kinder anfälliger für Fehler in der Produktion, je weniger frequent die starke Partizip-II-Form war (Clahsen u. a. 2004: 695–696). Wenig überraschend war dabei, dass die Übergeneralisierung der schwachen Flexion die häufigste produzierte Abweichung war. Auch die Reaktionszeiten hingen von der Frequenz des Partizips II ab: Bei starken Verben waren Erwachsene sowie Kinder schneller,

¹⁶Die Partizip-II-Frequenz der Testverben wurde anhand der Datenbank CELEX ermittelt. Nur korrekte Antworten wurden in die Datenanalyse aufgenommen.

wenn das Partizip II frequent war (Clahsen u. a. 2004: 698–700).¹⁷ Für schwache Verben ließ sich kein Effekt für die Erwachsenen feststellen. Für Kinder und eine Subgruppe der Erwachsenen, die generell ein langsames Reaktionsverhalten aufwiesen, stellen Clahsen u. a. (2004: 701) einen umgekehrten Frequenzeffekt (niedrigere Reaktionszeiten für die infrequenten Verben) fest. Dieser Effekt war jedoch nicht in der *lexical decision task* von Clahsen u. a. (1997) sichtbar. Es lässt sich daher festhalten, dass Reaktionszeiten durch die Frequenz starker Verben beeinflusst werden: Frequente starke Verben rufen i. d. R. niedrigere Reaktionszeiten hervor als infrequente. Die Ergebnisse verdeutlichen den grundlegenden Einfluss, den hohe Tokenfrequenz auf starke Verben hat: Durch ihre Tokenfrequenz sind sie mental stark gefestigt und daher leicht zu aktivieren. Verlieren die starken Verben jedoch an Frequenz, sind sie weniger leicht zu aktivieren, wie die erhöhten Reaktionszeiten zeigen.

Im empirischen Teil der Arbeit wird die bisherige psycholinguistische Forschung zur Verarbeitung von starken und schwachen Verben um die Perspektive auf Variation ergänzt. Dafür werden starke und schwache Partizip-II-Formen von Verben mit unterschiedlicher Tokenfrequenz sowie attestierter und nicht-attestierter Schwankung in Korpora präsentiert und die Reaktionszeiten auf diese Formen miteinander verglichen. Auf diese Weise ist es möglich, Reaktionszeiten starker und schwacher Formen in Abhängigkeit von Tokenfrequenz zu vergleichen.

Insgesamt wird deutlich, dass Typen- und Tokenfrequenz Einfluss auf die Variation in der Konjugation bei starken und schwachen Verben nehmen. Dabei ist das aus Abschnitt 2.1 bekannte Zusammenspiel aus Typen- und Tokenfrequenz zu beobachten: Tokenfrequente Verben finden sich in typeninrequenten Klassen (starke und unregelmäßige Verben) bzw. typeninrequenten Ablautalternanzen (ABC und ABA) (Werner 1989: 40–41). Aufgrund der hohen Tokenfrequenz sind sie mental verfestigt und leicht zu aktivieren (Bybee & Thompson 1997: 380). Tokenfrequente starke Verben rufen daher kurze Reaktionszeiten hervor. Die Irregularität stellt somit keinen Performanznachteil dar, sondern einen Vorteil, da sie zu differenzierten und kurzen Formen führt (Werner 1989: 41–43, Nübling 2000: 256, Nowak 2013: 174). Die hohe Tokenfrequenz führt außerdem dazu, dass

¹⁷Da in der Studie die Frequenz der Partizip-II-Formen und die generelle Frequenz des Verbs kovariieren, könnten die Ergebnisse auch auf einen generellen Frequenzeffekt hinweisen. Dasselbe könnte bei der oben diskutierten Studie von Clahsen u. a. (1997: 224–227) der Fall sein, da hier nicht spezifiziert wird, ob Kovarianz vorliegt. Gegen die Interpretation als generellen Frequenzeffekt sprechen die Studien von Clahsen u. a. (1997: 227–231) sowie von Clahsen u. a. (2001), diese werden in Abschnitt 3.1.2 zur Prototypizität der starken Konjugation diskutiert, da sie Hinweise auf eine prototypische Staffelung der Eigenschaften starker Flexion liefern.

die starken Formen eine hohe priore Wahrscheinlichkeit haben und die schwachen daher statistisch ausstechen können (zum statistischen Vorkaufsrecht siehe Goldberg 2019: 74–94 sowie Abschnitt 2.1.3).

Für tokeninfrequente Verben ist dagegen ein typenfrequentes und transparentes Muster von Vorteil. Tokeninfrequente Verben, die zur typeninfrequenten starken Klasse gehören, schließen sich deswegen der typenfrequenten schwachen Klasse an, da die schwachen Formen durch die geringe Tokenfrequenz nicht mehr statistisch ausgestochen werden. Dieses Prinzip wird im empirischen Teil der Arbeit anhand von starken und schwachen Partizipformen von starken Verben mit unterschiedlicher Tokenfrequenz untersucht.

Auch innerhalb der starken Klasse zeigt sich das Zusammenspiel aus Typen- und Tokenfrequenz: Gehört ein tokeninfrequentes Verb der typeninfrequenten Ablautalternanz ABC an, kann es sich zunächst der typenfrequenten Ablautalternanz ABB anschließen, in Form des Ablautmusters *x-o-o*. Diese Anpassung führt entweder zu einer stabilen Flexion oder stellt eine Zwischenstation zur schwachen Flexion dar. Im Folgenden wird der Einfluss der Prototypizität auf die Variation in der Konjugation schwacher und starker Verben diskutiert.

3.1.2 Prototypizität

Die Eigenschaften der starken und schwachen Konjugation sind prototypisch organisiert und daher graduell. Dementsprechend sind die oben genannten Eigenschaften starker Verben (Imperativhebung, Wechselflexion im Präsens, Ablaut in den Vergangenheitstempora, Umlautung des Präteritalstamms im Konjunktiv II) als prototypisch für die starke Flexion zu sehen (Nowak 2016: 130–131). Die einzelnen Eigenschaften stehen in einem Implikationsverhältnis: Weist ein Verb eine Imperativhebung auf, dann sind auch Wechselflexion im Präsens sowie Ablaute in den Vergangenheitstempora vorhanden (Bittner 1985: 54–57, Dammel & Nowak 2011: 26). Die Implikationsbeziehung ist dabei gerichtet. Aus einem Ablaut in den Vergangenheitstempora lässt sich keine Imperativhebung erschließen: Die Form *molk* impliziert also nicht den Imperativ *milk*, aber die Imperativhebung (*gib!*) den Ablaut im Präteritum (*gab*) (Dammel & Nowak 2011: 26).¹⁸ Bittner (1996: 80)¹⁹ stellt Implikationsbeziehungen auf, die in Abbildung 3.1 dargestellt sind.

Die Implikationsbeziehungen in Abbildung 3.1 lassen sich auch als Prototypizitätsskala nutzen (Bittner 1985: 57, Nowak 2015: 157): Prototypisch starke Verben

¹⁸Ursprünglich wies das Verb *melken* aber einen Imperativ mit Hebung auf (Nowak 2018: 156).

¹⁹Der Konjunktiv II fehlt bewusst in der Darstellung. Er wird von Bittner (1996: 80) zwischen dem Ablaut im Präteritum und dem Ablaut im Partizip II angesiedelt.

3.1 Variation in der Konjugationsklasse

Imperativ Singular Hebung)	2./3. Sg. Präsens Wechselflexion)	Indikativ Präteritum Ablaut)	Partizip II Ablaut
<i>gib!</i>		<i>du gibst/er,sie,es gibt</i>		<i>ich gab</i>		<i>gegeben</i>

Abbildung 3.1: Implikationsskala nach Bittner (1996: 80)

wie *geben* weisen die Imperativhebung auf und damit alle weiteren Eigenschaften starker Verben, da die Imperativhebung Wechselflexion und Ablaute in den Vergangenheitstempora impliziert. Einen Schritt vom Prototyp entfernt sind Verben wie *fangen*, die keine Imperativhebung aufweisen, aber Wechselflexion im Präsens. Damit sind auch Ablaute in den Vergangenheitstempora impliziert. Da die Implikation gerichtet ist, impliziert die Wechselflexion nur Ablaute in den Vergangenheitstempora, aber nicht die Imperativhebung. Am weitesten vom Prototyp entfernt sind dementsprechend Verben, die nur das Partizip II stark bilden (*salzen, salzte, gesalzen*). Auch die Eigenschaften starker Flexion lassen sich als prototypisch gestaffelt sehen: Die Imperativhebung und die Wechselflexion sind periphere Eigenschaften starker Flexion, da nur einige starke Verben sie aufweisen, der Ablaut stellt hingegen eine prototypische Eigenschaft dar.

Die Implikationsbeziehungen sind mit der Relevanzskala nach Bybee (1985) und mit intraparadigmatischer Frequenz verknüpft. Die Kategorie TEMPUS ist relevanter als die Kategorie PERSON, da das Tempus die Semantik eines Verbs stärker verändert als die Person (Nübling 1998: 192). Die Relevanz steigt also von der linken Seite (Imperativhebung, Wechselflexion im Präsens) zur rechten Seite der Prototypizitätsskala (Ablaute in den Vergangenheitstempora). Zudem sind die Kategorien auf der linken Seite der Skala i. d. R. weniger frequent als die auf der rechten Seite. So ist der Imperativ seltener als die 3. Person Singular (Lerche 2019: 117–118).²⁰ Die 3. Person Singular stellt jedoch die frequenteste Form dar (Dammel 2011: 61), sodass hier die geringe Relevanz wichtiger zu sein scheint als der Einfluss der intraparadigmatischen Frequenz. In Bezug auf die Ablaute in Präteritum und Partizip II lässt sich der Einfluss der intraparadigmatischen Frequenz erkennen, da das Präteritum aufgrund des Präteritalschwundes (Solms 1984: 311) immer seltener genutzt wird.²¹

²⁰Die 2. Person Singular Präsens ist Lerche (2019: 117–118) zufolge ähnlich frequent wie der Imperativ. Lerche (2019: 112–113) geht auf Basis der *paradigm structure hypothesis* nach Nessel & Janda (2010: 712–713) jedoch davon aus, dass die 2. Person Singular als eine prototypischere Verbform wahrgenommen wird als der Imperativ, sodass sich die Wechselflexion trotz relativ niedriger Frequenz besser in der 2. Person Singular Präsens hält als die Hebung im Imperativ.

²¹Bereits im 17. Jahrhundert liegt die Gebrauchsfrequenz des Präteritums mit 25 % deutlich unter der des Perfekts (Solms 1984: 311). Im 14. Jahrhundert war das Präteritum mit 61 % dagegen noch weitaus frequenter als das Perfekt.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Indizien für den Einfluss der intraparadigmatischen Frequenz auf die Konjugation liefert auch psycholinguistische Forschung: So werden in *lexical decision tasks* bei gleicher Lemmafrequenz frequente starke Partizip-II-Formen schneller verarbeitet als infrequente starke Partizip-II-Formen (Clahsen u. a. 1997: 227–231), dasselbe gilt für Präteritalformen (Clahsen u. a. 2001: 527–528). Clahsen u. a. (2001: 528–530) testen zudem ABB-Verben mit unterschiedlicher Frequenz der Präteritalformen. Dabei wird sichergestellt, dass die Frequenz des Ablauts in Präterital- und Partizip-II-Formen für alle Testverben vergleichbar ist. Die Testverben mit geringer Frequenz der Präteritalformen haben also eine höhere Frequenz hinsichtlich der Partizip-II-Formen, sodass der Ablaut insgesamt gleich häufig ist wie für Verben mit hoher Frequenz der Präteritalformen. Das Design erlaubt es daher, zu testen, ob die Frequenz des Ablauts an sich Einfluss auf die Prozessierung nimmt (dann müssten auch tokeninfrequente Präteritalformen schnell prozessiert werden) oder nur die Frequenz des Ablauts in Bezug auf Präteritalformen (dann müsste zwischen tokenfrequenten und -infrequenten Präteritalformen ein Unterschied zu erkennen sein). Letzteres war der Fall: Frequente Präteritalformen werden schneller verarbeitet als infrequente (Clahsen u. a. 2001: 530). Die Ergebnisse deuten somit darauf hin, dass vorrangig die Frequenz der jeweiligen Form im Paradigma ausschlaggebend für die Prozessierung ist.

Die Implikationsbeziehungen lassen sich für Vorhersagen darüber nutzen, in welcher Reihenfolge Schwankungen auftreten, wenn starke Verben in die schwache Flexion übergehen. Dabei lässt sich ein Einfluss von Frequenz und Relevanz festmachen: Infrequente Verben tendieren dazu, die Eigenschaften starker Flexion zunächst in den weniger frequenten und weniger relevanten Kategorien und schließlich in den frequenten und relevanten Kategorien abzubauen.²² Es ist davon auszugehen, dass der Übergang jeweils unumkehrbar ist, sobald sich die schwache Form in einer Kategorie verfestigt hat. Die Skala nach Bittner (1996: 80) ist dabei nicht so zu verstehen, dass alle Verben diese komplett durchlaufen. Vielmehr können Verben auch auf einer Stufe stehen bleiben. Dies nimmt Nowak (2013: 182–183) bspw. für *fechten* und *flechten* an, die sich dem Muster *x-o-o*²³ angeschlossen haben und bislang keine Tendenzen zu schwachen Präterital- und Partizip-II-Formen aufweisen.

Den Implikationsbeziehungen folgend stellt die Aufgabe der Imperativhebung den ersten und die Aufgabe des starken Partizips II den letzten Schritt zur schwachen Flexion dar (Nowak 2013: 182–183). Schwankungen im Imperativ lassen sich

²²Nübling (1998: 197–200) sieht Relevanz und Frequenz auch als entscheidende Faktoren, die den Numerusausgleich im Frnhd. beeinflusst haben.

²³Nowak (2013: 182–183) schlägt *x-o-o* als weitere fakultative Stufe in der Prototypizitätsskala vor, dies wird unten näher erläutert.

daher nicht nur bei infrequenten, sondern auch bei frequenten Verben beobachten (*gib/geb, wirf/werf*). Frequente Verben konservieren ihre Imperativhebung jedoch nach wie vor eher als infrequentere Verben: So neigen *geben* und *werfen* eher zu Imperativformen mit Hebung als *treten*, das weniger frequent ist (Lerche 2019: 124–125). Lerche (2019: 160–162) kann auch einen frequenzbedingten Einfluss auf Lesezeiten von Imperativformen feststellen: Bei frequenten Verben wird die Form mit Imperativhebung schneller gelesen, bei infrequenten hingegen die Form ohne Imperativhebung. Dieser Effekt ist jedoch abhängig von den jeweiligen Proband_innen: Während junge Proband_innen (geboren zwischen 1985 und 1997) aus Baden-Württemberg bereits bei frequenteren Verben die Form ohne Hebung schneller lesen, setzt der Effekt bei älteren (geboren zwischen 1955 und 1970) erst in einer niedrigeren Frequenzklasse ein. Lerche sieht dies als einen Hinweis auf einen voranschreitenden Wandel zu Formen ohne Imperativhebung. Für eine weitere Testgruppe aus Sachsen²⁴ kann sie hingegen keinen Einfluss der Frequenz feststellen: Hier wird stets die Form ohne Imperativhebung langsamer gelesen als die Form mit Hebung.

Aufgrund der geringen intraparadigmatischen Frequenz des Imperativs kann man die Schwankungen von frequenten Verben wie *geben* im Imperativ nicht unbedingt als ein erstes Zeichen für ihren Anschluss an die schwache Flexion deuten: Da der Imperativ an sich keine frequente Kategorie darstellt, könnte allein der Imperativ von der Schwankung betroffen sein. Zudem betrifft die Imperativhebung generell nur wenige starke Verben, da sie nur Verben offen steht, die /e:/ als Präsensstammvokal haben, weil nur diese historisch eine Hebung vorgenommen haben können (Braune & Heidermanns 2018: 365). Verben mit Imperativhebung stellen daher innerhalb der typeninrequenten starken Klasse eine noch typeninrequentere Gruppe dar. Aus dieser Perspektive wäre es nicht verwunderlich, wenn die Imperativhebung über infrequente hin zu frequenten Verben generell aufgegeben würde und sich daher starke und schwache Verben im Imperativ angleichen. Schwache Verben weisen zwar in der Regel ein Schwa im Imperativ auf (*sage!*). Die Schwa-Apokope (*sag* statt *sage!*) führt jedoch dazu, dass die Imperativformen starker und schwacher Verben sich formal gleichen (Dammel 2011: 99–101), wenn keine Imperativhebung erfolgt (*geb!*; *sag!*).

In der bislang diskutierten Prototypizitätsskala fehlen zwei Eigenschaften der starken Flexion: Dies ist zum einen die Markierung des Konjunktivs II mit Um-

²⁴Lerche (2019: 132–133) akquirierte Proband_innen aus Baden-Württemberg sowie Sachsen mit der Annahme, dass die Proband_innen aus Sachsen höhere Reaktionszeiten für die Formen ohne Imperativhebung aufweisen als die Proband_innen aus Baden-Württemberg, da Formen ohne Hebung bei niedrigfrequenten Verben im Niederalemannischen verbreiteter sind als im Obersächsischen. Diese Annahme wird in der Untersuchung bestätigt (Lerche 2019: 142–143).

3 Variation in Konjugation, Deklination und Auxiliarselektion

laut und zum anderen die Vokaländerung starker Verben hin zum Muster *x-o-o*. Bittner (1985: 56–57) positioniert den Umlaut im Konjunktiv auf seiner Skala zwischen dem Ablaut im Präteritum und dem Ablaut im Partizip II. Dies ist zunächst nachvollziehbar, da alle Verben, die im Präteritum auf einen umlautfähigen Vokal ablauten, auch den synthetischen Konjunktiv mithilfe der Umlautung des Ablauts bilden können. Allerdings ist die gerichtete Implikation hier gebrochen: Auch bei einem Umlaut im Konjunktiv II wird ein Ablaut im Präteritum benötigt, da dieser umgelautes werden soll (Nowak 2016: 129).²⁵ Bittner (1985: 56) führt *schinden*, *schleißen* und *melken* als Beispiele für Verben an, die keinen Ablaut im Präteritum aufweisen (*schindete*, *schleißte*, *melkte*), aber einen Umlaut im Konjunktiv II (*schünde*, *schlisse*, *mölke* neben *melkte*). Er bietet dafür jedoch keine empirischen Belege, sondern bezieht sich auf Grammatiken und Wörterbücher.

Nowak (2013: 179–180) geht davon aus, dass die Angleichung der Ablautvokale an *x-o-o* nach der Aufgabe der Wechselflexion im Präsens erfolgt, siehe Abbildung 3.2.²⁶

Imperativ Singular Hebung)	2./3. Sg. Präsens Wechselflexion)	Ablautmuster x-o-o)	Indikativ Präteritum Ablaut)	Partizip II Ablaut
<i>bill!</i>		<i>du billst/er,sie,es bellt</i>				<i>ich ball/boll</i>		<i>gebollen</i>
> <i>bell!</i>		> <i>du bellst/er,sie,es billt</i>				> <i>ich bellte</i>		> <i>gebellt</i>
				<i>ich ball</i>				
				> <i>ich boll</i>				

Abbildung 3.2: Abbaustufen der starken Flexion nach Nowak (2013: 183)

Nowak (2013: 182–183) schlägt deshalb die Änderung des Ablautmusters zu *x-o-o* als weitere fakultative Entwicklungsstufe vor der Aufgabe des Ablauts im Präteritum vor. Dabei kann *x-o-o* aber gleichzeitig mit Wechselflexion auftreten:

²⁵Dies lässt sich gut mit Relevanz in Verbindung bringen: Tempus ist relevanter als Modus, somit legt das Tempus den Vokal fest, während der Modus diesen nur modifiziert (Nübling & Dammel 2004: 195).

²⁶Wie in Abbildung 3.1 wird der Konjunktiv II hier nicht berücksichtigt (*begänne* > *begönne*). Abbildung 3.2 ließe sich zudem noch spezifizieren, da der Wandel zu *x-o-o* nicht nur das Präteritum, sondern auch das Partizip II betreffen kann. Hierbei ergibt sich aber ein Konflikt zwischen Entwicklungsrichtung und Implikationsbeziehung: Nach der Entwicklungsrichtung müsste in der Abbildung Partizip II vor dem Präteritum stehen, da Nowak (2018: 166) zufolge der Ablautvokal im Partizip II vor dem Vokal im Präteritum zu /o:/ bzw. /ɔ/ modelliert wird. Den Implikationsbeziehungen zufolge müsste das Präteritum jedoch vor dem Partizip II stehen: /o:/ bzw. /ɔ/ im Präteritum impliziert /o:/ bzw. /ɔ/ im Partizip II, da meines Wissens kein Verb existiert, das /o:/ bzw. /ɔ/ im Präteritum als Ablaut hat, aber nicht im Partizip II. Umgekehrt impliziert /o:/ bzw. /ɔ/ im Partizip II nicht /o:/ bzw. /ɔ/ im Präteritum (*stehlen*, *stahl*, *gestohlen*).

3.1 Variation in der Konjugationsklasse

Dreschen, das ursprünglich der Ablautreihe 5 angehörte (*dreschen, drasch, gedroschen*) (Nowak 2013: 176), hat sich *x-o-o* angeschlossen (*dreschen, drosch, gedroschen*), weist aber nach wie vor Wechselflexion im Präsens auf (*er, sie drischt*). Allerdings postuliert Nowak (2016: 133) auf Grundlage einer Googlerecherche Abbautendenzen (*drischt > drescht*). Zudem weist Nowak (2016: 33) darauf hin, dass nur wenige Verben sowohl *x-o-o* als auch Wechselflexion aufweisen. Einige dieser Verben etablierten bzw. etablieren hierbei interessanterweise einen neuen Infinitiv: Diese Verben erfordern kein belebtes Subjekt (wie *schmelzen* und *erlischen*) oder stellen unpersönliche Verben dar (z. B. *ziemen*, das bereits komplett schwach flektiert), weshalb die 3. Person Singular häufiger genutzt wird als andere Personen, die keine Wechselflexion aufweisen. Daher wird der Wechselflexionsvokal als Infinitivstammvokal generalisiert (*schmelzen > schmilzen, erlöschen > erlischen, zemen > ziemen*) (Bittner 2010: 249).

Nowak (2016: 132) weist auf eine Einschränkung der hier vorgestellten Prototypizitätsskala hin. Sie kann nur eine Dimension der starken Flexion abdecken, da sie keine Hinweise über die Vokalalternanzen enthält.²⁷ So stehen Verben mit einem singulären Ablautmuster wie *rufen, rief, gerufen* auf einer Stufe mit Verben, die einem typenfrequenten Muster wie *i-a-u* angehören (*singen, sang, gesungen*). Basierend auf der Typenfrequenz ist das singuläre Ablautmuster eher als stark einzustufen als typenfrequente Ablautmuster, weil sich hier keine Gemeinsamkeiten zwischen mehreren Verben ableiten lassen. Das Gleiche gilt für die Ablautalternanzen: ABC und ABA sind prototypischer stark als die Vokalalternanz ABB, da ABB einerseits weit typenfrequenter ist als die anderen Alternanzen und andererseits das schwache Konjugationsverfahren spiegelt, in dem Präsens mit den Vergangenheitstempora kontrastiert wird (Nowak 2016: 135–136; siehe hierzu ausführlich Abschnitt 3.1.1). Dennoch ist die Skala nützlich, um den Wandel von starker zu schwacher Konjugation schematisch darzustellen.

Insgesamt ist festzuhalten, dass die Eigenschaften starker Verben prototypisch organisiert sind. Diese Organisation lässt sich durch Implikationsbeziehungen fassen und ist durch die Relevanzhierarchie sowie intraparadigmatische Frequenz gestützt. Die Implikationsbeziehungen führen zu peripheren und prototypischen Eigenschaften starker Flexion: Die peripheren Eigenschaften (Imperativhebung) weisen nur einige Verben auf, die prototypischen (Ablaut im Partizip II) dagegen alle. Die prototypisch gestaffelten Eigenschaften starker Verben lassen sich nutzen, um die Reihenfolge der Schwankungen zu prognostizieren: Periphere Eigenschaften, die eine wenig relevante sowie seltene Kategorie (bspw. Imperativ) betreffen, werden von niedrigfrequenten Verben zuerst abgebaut. Dies

²⁷Wie in Abschnitt 2.2.2 erläutert, ist dies ein generelles Problem von Typizitätsskalen.

3 Variation in Konjugation, Deklination und Auxiliarselektion

sind die Imperativhebung sowie die Wechselflexion im Präsens. Der Ablaut betrifft die relevante Kategorie Tempus und wird dementsprechend zuletzt abgebaut. Da das Präteritum seltener ist als das Perfekt, ist der Abbau des Ablauts im Präteritum vor dem Abbau im Perfekt zu beobachten. Zudem kann der Ablaut zu *x-o-o* modifiziert werden, bevor das Verb ganz in die schwache Flexion übergeht. Im folgenden Abschnitt wird der Einfluss der Form-Schematizität auf die Variation zwischen starker und schwacher Konjugation in den Blick genommen.

3.1.3 Form-Schematizität

Infrequente Verben, deren phonologische Eigenschaften mit starkem Konjugationsverhalten assoziiert sind und die somit zu einem Form-Schema gehören, halten sich eher in der starken Flexion als Verben, die vergleichbar infrequent sind, aber keinem Form-Schema angehören (zur Unterscheidung von Schemata und Form-Schemata siehe Abschnitt 2.3). Dabei beeinflussen zwei Form-Schemata die starke Konjugation: Verben mit der phonologischen Form [#_aɪ + bilabialer Obstruent] (*reiben*) sowie [#_ɪ + ŋ + (C)] (*sinken*) scheinen sich eher in der starken Flexion zu halten als starke Verben mit anderen phonologischen Eigenschaften (Köpcke 1999: 52–27).

Köpcke (1999) untersucht 171 starke Verben²⁸ auf Cluster in ihren phonologischen Eigenschaften. Dabei beobachtet er, dass die Infinitivstammvokale /i:/ bzw. /ɪ/ bei 53 und /e:/, /ɛ:/ bzw. /ɛ/ bei 47 Verben²⁹ vorkommen und damit deutlich häufiger sind als andere Vokale: /ɑ:/ bzw. /a/ ist mit 18 Verben auch noch relativ häufig, während /y:/ bzw. /ʏ/, /u:/ bzw. /ʊ/ und /o:/ bzw. /ɔ/ nur zwei- bis dreimal vorkommen.³⁰ Auch bei den Diphthongen ergibt sich eine klare Verteilung: /aɪ/ ist mit 39 Verben sehr viel häufiger als /aʊ/ mit nur fünf. Köpcke (1999: 53) schließt hieraus, dass vornehmlich vordere Vokale mit der starken Flexion assoziiert sind. Dabei betont er jedoch, dass der Vokalismus alleine noch kein Form-Schema konstituiert, da /i:/, /ɪ/ und /aɪ/ nicht ausschließlich bei starken Verben auftreten und die Häufung dieser Infinitivstammvokale bei den starken Verben noch keine Präferenz für starke Konjugation indiziert.

Köpcke (1999: 53–56) betrachtet deshalb auch den Konsonantismus bei Verben mit den Infinitivstammvokalen /i:/ bzw. /ɪ/ und /aɪ/. Hierbei zählt er aus, wie

²⁸Dabei folgt er der Liste von Fabricius-Hansen (1977: 197–198).

²⁹Dies ist vor dem Hintergrund, dass *x-o-o*-Verben oft /e:/, /ɛ:/ oder /ɛ/ als Infinitivstammvokal enthalten (Nowak 2018: 169), nicht verwunderlich. Im weiteren Verlauf des Abschnitts wird auf *x-o-o* als Form-Schema eingegangen.

³⁰Köpcke (1999: 53) differenziert hierbei nicht nach Vokallänge bzw. -gespanntheit.

3.1 Variation in der Konjugationsklasse

viele starke und schwache Verben eine bestimmte phonologische Form aufweisen.³¹ Für /aɪ/ stellt Köpcke (1999: 54) zunächst fest, dass viele starke Verben die phonologische Form [aɪ + Obstruent] aufweisen. Nur ein Verb weist [aɪ + Nasal] (*scheinen*) auf und bei zehn Verben bleibt die postvokalische Position leer (*gedeihen*).

Aufgrund dieser Verteilung ermittelt Köpcke (1999: 54–55) für Plosive und Frikative, wie häufig starke und schwache Verben die Kombination aus /aɪ/ und dem jeweiligen Konsonant aufweisen. Für die Form [aɪ + Plosiv] steigt der Anteil an starken Verben, je weiter vorne im Mundraum der Konsonant gebildet wird: Verben mit [aɪ + k/g] sind zu 20 % stark, Verben mit der Form [aɪ + t/d] zu 30 % und Verben mit [aɪ + p/b] zu 60 %. Ähnliche Ergebnisse erzielt Köpcke (1999: 55) für die Kombination aus [aɪ + Frikativ]: Auch hier steigt der Anteil starker Verben im Vergleich zu schwachen, je weiter vorne der Frikativ produziert wird. Für die Kombination [aɪ + f/v] ermittelt Köpcke (1999: 55) einen Anteil starker Verben von 60 %.

Der Vokal /ɪ/ wird bei starken Verben ausschließlich mit Nasalen oder einer Kombination aus Nasal und anderem Konsonant genutzt (Köpcke 1999: 55).³² Auch hier ist entscheidend, wo der Nasal im Mundraum gebildet wird. Je weiter hinten im Mundraum der Nasal gebildet wird, desto höher ist der Anteil an starker Flexion im Vergleich zur schwachen: Bei [m (+ C)] liegt er bei 40 %, bei [n (+ C)] bei 50 % und bei [ŋ (+ C)] bei 90 %. Die Form [#_ɪ + ŋ (+ C)] ist also eindeutig mit der starken Flexion assoziiert, für [m (+ C)] und [n (+ C)] ist hingegen von einer weniger starken Assoziation mit der starken Flexion auszugehen.

Es lässt sich festhalten, dass die phonologischen Formen [aɪ + bilabialer Obstruent] sowie [#_ɪ + ŋ (+ C)] mit starker Konjugation assoziiert sind und somit Form-Schemata bilden. Die Abstufungen der Form-Schemata zeigen, dass diese probabilistisch sind und einen Prototyp aufweisen (ausführlich zu Prototypizität und den prototypisch organisierten Eigenschaften von (Form-)Schemata siehe Abschnitt 2.2 und 2.3.2). Die hier postulierten Form-Schemata wurden für das Deutsche nur anhand von real existierenden Verben diskutiert. Für das Englische haben Bybee & Moder (1983) allerdings die Existenz eines ähnlich gelagerten Form-Schemas (*string, strung, strung*) auch anhand von Pseudoverben nachgewiesen (siehe Abschnitt 2.3.3 für genauere Erläuterungen).

Wie in Abschnitt 2.3.3 argumentiert wurde, lassen sich Form-Schemata einerseits als Gegenspieler von Typenfrequenz begreifen, da sie starke Verben in der

³¹Als Basis für diese Zählung dient die Liste der Grundverben von Mater (1967).

³²/i:/ wird dagegen innerhalb der starken Verben nie mit Nasalen kombiniert, sondern mit Frikativen (*triefen*), Plosiven (*bieten*) oder Liquiden (*frieren*) (Köpcke 1999: 55).

3 Variation in Konjugation, Deklination und Auxiliarselektion

Flexionsklasse halten und die Variationsrichtung sogar umdrehen können. Andererseits stellen sie eine typenfrequente Gruppe innerhalb der typeninrequenten Klasse dar: Verben mit ähnlichen phonologischen Eigenschaften bilden ein Cluster und stützen sich auf diese Weise gegenseitig in ihrer Flexion. In diesem Sinne lassen sich *senken* und *trinken* als Token des Types [#_I + ŋ (+ C)] ansehen. Hierbei ist von ineinander verschachtelten Form-Schemata auszugehen: *trank* ist mit [+Vergangenheit] verbunden und aufgrund dieser Assoziation ist das Verb *trinken* mit starker Flexion assoziiert bzw. – abstrakter gefasst – Verben mit der Form [#_I + ŋ (+ C)] (zur Verschachtelung von Form-Schemata siehe Abschnitt 2.3.2). Daher werden schwache Formen für Verben mit dieser Form statistisch ausgestochen und in der Folge dispräferiert (**trinkte*). Es ist davon auszugehen, dass die Assoziation von Form und Funktion für [#_I + ŋ (+ C)] stärker ist als für [#_aɪ + bilabialer Konsonant], da der Anteil starker Verben im Vergleich zu schwachen Verben für diese Form höher ist und die schwache Flexion somit besser statistisch ausgestochen werden kann.

Neben einem variationshemmenden Effekt können die beiden Form-Schemata auch einen variationstreibenden Effekt haben: Die Form *gewunken*, die sich inzwischen neben der ursprünglichen Form *gewinkt* standardsprachlich behauptet (Duden 2020), ist auf das Wirken des Form-Schemas [#_I + ŋ + (C)] zurückzuführen. *Winken* entspricht diesem Form-Schema und wird daher im Partizip II wie ein starkes Verb behandelt. Der Anschluss von *winken* an die starke Flexion findet dabei genau in entgegengesetzter Reihenfolge zum Anschluss starker Verben an die schwache Flexion statt. Bisher ist nur das Perfekt von *winken* stark, die anderen Formen werden schwach flektiert. Allerdings ist das Potential für *winken*, starke Flexionsformen aufzuweisen, damit fast ausgeschöpft. Lediglich das Präteritum könnte noch stark gebildet werden (*wank*), Wechselflexion im Präsens und Imperativhebung sind aufgrund des Infinitivstammvokals /i:/ jedoch ausgeschlossen. Köpcke (1999: 56) verweist neben *winken* auf Verben wie *bleichen*, *kneifen* und *pfeifen* sowie *dingen* und *zinden*, die im Frnhd. stark geworden sind und diese Flexion teilweise bis ins Nhd. bewahrt haben. Aber auch innerhalb der starken Verben können Form-Schemata Variation hervorrufen. So scheint das Form-Schema [#_aɪ + bilabialer Konsonant] dazu zu führen, dass Verben der Ablautreihe 7, die /aɪ/ als Infinitivstammvokal aufweisen (*heißen*, *hieβ*, *geheißen*), zu /i:/ im Partizip II schwanken (*gehießen* statt *geheißen*) (Nowak 2018: 171). Form-Schemata führen jedoch nicht zwangsläufig zu Schwankungen: So behaupten sich *hinken*, *blinken* und *schminken* trotz einzelner starker Formen weiterhin in der schwachen Flexion.

Auch das in Abschnitt 3.1.1 diskutierte Ablautmuster *x-o-o* lässt sich als ein Form-Schema auffassen. Nowak (2018: 169) geht davon aus, dass *x-o-o* eine pro-

3.1 Variation in der Konjugationsklasse

totypisch organisierte Form aufweist: Sie schlägt [vorderer Konsonant/Liquid + e:/ε/ε: + vorderer Konsonant/l] (*schwellen, quellen, melken*) als Prototyp vor. Um den Prototyp gruppieren sich Verben mit den Eigenschaften [vorderer Konsonant/Liquid + e:/ε/ε:] (*pflegen, erwägen*) sowie [e:/ε/ε: + vorderer Konsonant/Liquid] (*verhehlen*). Neben /e:/, /ε/ bzw. /ε:/ sind auch /ɪ/ und /ø:/ als Infinitivstammvokal möglich (*glimmen, schwören*). Trotz der phonologischen Gemeinsamkeiten der Mitglieder ist das *x-o-o*-Form-Schema insgesamt weitaus abstrakter als die Form-Schemata [#_aɪ + bilabialer Konsonant] und [#_ɪ + ɲ + (C)], was sich bspw. daran zeigt, dass die Konsonanten des Form-Schemas vergleichsweise wenig definiert sind.

Wie in Abschnitt 3.1.1 bereits erläutert wurde, ist davon auszugehen, dass sich das *x-o-o*-Form-Schema in Analogie zur Ablautreihe 2 ausgebildet hat (Nowak 2018: 166). So lassen sich lautliche Ähnlichkeiten zwischen Verben der Ablautreihe 2 und *x-o-o*-Verben erkennen, bspw. haben *schieben* und *heben* denselben stamfinalen konsonantischen Auslaut. Dass die Abweichung im Infinitivstammvokal (/e:/, /ε/ bzw. /ε:/ statt /i:/) die Analogiebildung nicht behindert hat, erklärt Nowak (2018: 165–166) damit, dass die Ablautreihe 2 im Mhd. phonologisch nicht stark restringiert war. Neben /i:/ als Infinitivstammvokal existierten im Mhd. auch /u:/ (*sûgen* ‚saugen‘), /y:/ (*triuwen* ‚zutrauen‘) und /ɤɤ/ (*soufen* als Nebenform zu *sûfen* ‚saufen‘) (Nowak 2015: 233). Erst zum Nhd. verfestigt sich der Infinitivstammvokal der Ablautreihe 2 auf /i:/, nur die Verben *lügen, trügen, saufen, saugen* und *schnauben* weichen bis heute hiervon ab (Nowak 2015: 233, 2018: 169).

Die Verfestigung des Form-Schemas *x-o-o* auf /e:/, /ε/ und /ε:/ lässt sich dadurch begründen, dass Verben, die bereits /o:/ bzw. /ɔ/ im Partizip-II-Stamm hatten und daher den Ausgangspunkt für *x-o-o* darstellen, /e:/ (*heben*), /ε/ (*schwellen*) oder /ε:/ (*gären*) als Infinitivstammvokal haben (Nowak 2015: 266–267). Sie bilden daher ein Cluster und konnten als Analogievorlage wahrgenommen werden. *a*-haltige Verben scheuen *x-o-o* hingegen, sodass das Muster *a-o-o* mit *erschallen* singularär ist. Dazu passt, dass *erschallen* trotz Anschluss an *x-o-o* inzwischen schwache Formen aufweist (Nowak 2015: 272).

Die Assoziation von /e:/, /ε/ bzw. /ε:/ mit *x-o-o* liefert eine Erklärung dafür, warum Verben der Ablautreihe 1 und 7 *x-o-o* nicht annehmen: Keine der Ablautreihen enthält Verben mit /e:/, /ε/ oder /ε:/ als Infinitivstammvokal. Ablautreihe 1 enthält nur Verben mit /aɪ/, Ablautreihe 7 ist zwar variabel im Infinitivstammvokal, enthält aber historisch keine Verben mit dem Infinitivstammvokal *e* (Braune & Heidermanns 2018: 399–404). Ein Blick auf Ablautalternanzen liefert weitere Gründe gegen *x-o-o*: Verben der Ablautreihe 1 weisen bereits die Vokalalternanz ABB auf, sodass der Anschluss an *x-o-o* keine Vereinfachung der Vokalalternanz mit sich

3 Variation in Konjugation, Deklination und Auxiliarselektion

bringt. Zudem werden die Verben aufgrund des Form-Schemas [#_aɪ + bilabialer Konsonant] gestützt (Nowak 2018: 171). Verben der Ablautreihe 7 weisen ABA als Vokalalternanz auf, die laut Nowak (2016: 141) aufgrund des Präteritalschwundes in die Nähe schwacher Verben rückt.

Wegen der geringen formalen Restriktion kann *x-o-o* als Auffangbecken für Verben agieren, die (aufgrund zu geringer Tokenfrequenz) nicht mehr in ihrem ursprünglichen Ablautmuster gehalten werden. Dabei kann auch ein Wechsel vom ursprünglichen Form-Schema zu *x-o-o* vollzogen werden, wie bei *rinnen* (*rann* > *ronn*) und *glimmen* (*glamm* > *glomm*). Hierbei scheinen Verben eher zu *x-o-o* zu neigen, wenn sie periphere Vertreter eines Form-Schemas sind (Nowak 2010a: 435): Die Verben *rinnen* und *sinnen* sind niedrigfrequent und passen zudem nicht perfekt in das ursprüngliche Form-Schema [#_ɪ + ɪ + (C)], da die Nasale vorne gebildet werden. Das abstraktere Form-Schema *x-o-o* springt somit ein, wenn die konkreteren Form-Schemata aufgrund von geringer Tokenfrequenz und/oder mangelnder Passfähigkeit der Verben nicht mehr greifen. Die Existenz des *x-o-o*-Form-Schemas müsste noch eingehender empirisch überprüft werden, bspw. anhand von Auszählungen wie von Köpcke (1999) oder Experimenten mit Pseudowörtern wie von Bybee & Moder (1983).

Insgesamt zeigt sich, dass Form-Schemata Einfluss auf die Variation in der Konjugation nehmen. Dabei existieren Form-Schemata wie bspw. [#_ɪ + ɪ + (C)], die ein kohärentes phonologisches Cluster bilden, und abstraktere wie das phonologisch eher lose definierte Form-Schema für das Ablautmuster *x-o-o*. Die Form-Schemata haben einerseits einen flexionsbewahrenden Effekt, da sie niedrigfrequente Verben in der starken Flexion halten, andererseits können sie Variation evozieren, indem sie schwache Verben mit form-schemakompatibler Form mit der starken Flexion assoziieren. Bei der Stabilisation der starken Flexion scheint eine doppelte Sicherung zu existieren: Zunächst bewahrt das konkrete Form-Schema (z. B. [#_ɪ + ɪ + (C)]) vor dem Wechsel in die schwache Flexion. Wenn dies nicht mehr ausreicht, springt das generelle Form-Schema *x-o-o* ein.³³ Im folgenden Abschnitt wird ein abschließender Blick auf den Einfluss von Frequenz, Prototypizität und (Form-)Schematizität auf die Variation in der Konjugation geworfen.

3.1.4 Zusammenfassung und Zusammenwirken der Faktoren

Frequenz nimmt grundlegenden Einfluss auf die Variation in der Konjugation. Hierbei interagieren Typen- und Tokenfrequenz. Die starke Flexion stellt die ty-

³³Bei Verben, die keinem Form-Schema angehören und tokeninfrequent sind, kann *x-o-o* selbstverständlich auch greifen.

peninfrequente Flexion dar. Somit ist zu erwarten, dass tokeninfrequente Verben zur typenfrequenten schwachen Flexion übergehen. Dies lässt sich theoretisch über das statistische Vorkaufsrecht (Goldberg 2019: 74–94; zur Modellierung in Bezug auf Flexionsklassen siehe Abschnitt 2.1.3) gut fundieren und wurde empirisch in zahlreichen Studien bestätigt (Augst 1975, Bittner 1996, Nowak 2015). Der Abbau der starken Flexion geschieht entlang einer Prototypizitätsskala. Diese leitet sich aus Flexionseigenschaften starker Verben ab, die in Implikationsbeziehungen zueinander stehen (Dammel & Nowak 2011: 26–27). Die Imperativhebung und Wechselflexion sind prototypisch starken Verben vorbehalten, während periphere Mitglieder der starken Flexion lediglich Ablaute im Präteritum und Partizip II aufweisen. Auch die Flexionseigenschaften lassen sich als peripher und prototypisch für die starke Flexion modellieren. Dabei sind die Flexionseigenschaften prototypisch, die alle starken Verben aufweisen (Ablaut), und die Eigenschaften peripher, die nur die prototypischen starken Verben zeigen (Imperativhebung). Der Wechsel von stark zu schwach beginnt bei peripheren und endet bei prototypischen Eigenschaften: Die erste Stufe stellt den Wegfall der Imperativhebung dar, anschließend wird die Wechselflexion abgebaut, es folgt die Aufgabe des Ablauts im Präteritum und schließlich des Ablauts im Partizip II samt Wechsel von *ge-...-n* zu *ge-...-t*. An dieser Stelle zeigen sich auch unterschiedliche Grade der Relevanz (Nowak 2016: 133): Die Markierung der für Verben weniger relevanten Kategorien Modus und Person wird vor der Markierung der relevanten Kategorie Tempus abgebaut. Zudem weisen die Stufen unterschiedliche intraparadigmatische Frequenzen auf: So ist bspw. das Perfekt frequenter als das Präteritum und kann sich daher länger in der starken Flexion halten.

Zusätzlich zu diesen Stufen schlägt Nowak (2013: 182–183) eine weitere fakultative Stufe vor der Aufgabe des Ablauts im Präteritum vor: den Wechsel der ursprünglichen Vokalalternanz hin zu *x-o-o*. Das Ablautmuster *x-o-o* hat den Vorteil, dass Präsens und Vergangenheitstempora auf die gleiche, invariante Weise miteinander kontrastiert werden können; es ähnelt somit dem schwachen Distinktionsmuster. Zudem reduziert *x-o-o* die Komplexität der Ablautalternanz ABC. Auch für die Ablautalternanz ABA ergibt sich durch *x-o-o* eine Komplexitätsreduktion, da der identische Vokal in Präsens und Partizip II keine Funktionsäquivalenz hat (Nowak 2013: 178–179). Allerdings weist Nowak (2016: 141) darauf hin, dass Verben mit der Ablautalternanz ABA selten *x-o-o* annehmen. Dies erklärt sie mit dem Präteritalschwund, der dazu führt, dass das Präteritum eine untergeordnete Rolle spielt und die Ablautalternanz A(B)A daher in die Nähe der schwachen Flexion rückt. Durch die Vokalidentität im Präsens und Partizip II unterscheidet sich die Ablautalternanz A(B)A nur durch das Suffix *-n* im

3 Variation in Konjugation, Deklination und Auxiliarselektion

Partizip II von der schwachen Flexion. Es ist daher davon auszugehen, dass Verben der Ablautalternanz ABA i. d. R. ohne den Weg über *x-o-o* schwach werden.

Die Änderung zu *x-o-o* führt nicht nur zu einer Vereinfachung der Ablautalternanz, sondern auch zu einem typenfrequentiellen Vorteil: Viele Verben weisen *x-o-o* als Ablautmuster auf (Nowak 2013: 180–181). Zudem ist die Ablautalternanz ABB, zu der *x-o-o* gehört, die typenfrequenteste Ablautalternanz. Typenfrequentiell lässt sich beim Übergang von der starken zur schwachen Flexion somit ein Zweischnitt ausmachen, wie Abbildung 3.3 zeigt. Die Angaben in der Abbildung sind Nowak (2013: 182) und Nowak (2015: 166) entnommen.

Schwache Verben ca. 4000	Starke Verben ca. 170	
Diese Fläche mal 5	ABB 52 %	ABC 29 %
	TF: 70	TF: 140
		ABA 19 % TF: 340

Abbildung 3.3: Typenfrequenz schwacher und starker Verben nach Ablautalternanz. TF: Tokenfrequenz pro Millionen Token.

Die schwachen Verben machen die größte Gruppe innerhalb der Verben aus. Innerhalb der starken Verben ist ABB die mitgliederstärkste Gruppe (Nowak 2016: 136). Wie bei starken und schwachen Verben stehen auch innerhalb der starken Verben die typenfrequentiellen Verhältnisse invers zur Tokenfrequenz: Die Ablautalternanz ABB hat viele Mitglieder, die aber nicht tokenfrequent sind. Umgekehrt weisen die typeninrequenten Ablautalternanzen hohe Tokenfrequenzen auf (Nowak 2016: 141). Somit erscheint es logisch, dass ein tokeninquentes Verb sich zunächst der typenfrequenteren Ablautalternanz ABB anschließt,

indem es *x-o-o* annimmt, und erst im Anschluss daran zu schwachen Präterital- und Perfekt-II-Formen übergeht.

Neben dem Anschluss an ein typenfrequentes Ablautmuster können Form-Schemata infrequente starke Verben in der starken Flexion halten. Dabei spielen zwei Form-Schemata eine Rolle: Verben mit der Form [#_aɪ + bilabialer Konsonant] und [#_ɪ + ŋ + (C)] scheinen mit der starken Flexion assoziiert zu sein (Köpcke 1999). Dies wird besonders bei [#_ɪ + ŋ + (C)] deutlich: 90 % der Verben, die diese Form aufweisen, flektieren stark. Somit halten sich Verben mit dieser Form auch eher in der starken Flexion. Außerdem werden schwache Verben mit form-schemakompatibler Form in die starke Flexion gezogen (Köpcke 1999: 56).

Nowak (2018: 166–170) sieht auch *x-o-o* als Form-Schema an, dessen Form jedoch abstrakter ist als die der anderen beiden Form-Schemata. So kann *x-o-o* Verben in der starken Flexion halten, wenn Tokenfrequenz und die anderen Form-Schemata die starke Flexion nicht mehr ausreichend stützen.

Aus dem Zusammenwirken der drei Einflussfaktoren lassen sich verschiedene Vorhersagen über mögliche Variationsfälle machen: Zunächst ist davon auszugehen, dass die Schwankungsrichtung generell von stark zu schwach geht. Eine umgekehrte Schwankungsrichtung ist (mit Ausnahme der Schwankung durch Form-Schemata) hingegen unwahrscheinlich. Zudem ist anzunehmen, dass der Übergang von stark zu schwach unumkehrbar ist. Die prototypisch gestaffelten Eigenschaften der starken Flexion geben zudem einen Einblick in die Abfolge der Schwankungen. Hierbei ist davon auszugehen, dass die Eigenschaften in der von Bittner (1996) vorgeschlagenen Reihenfolge aufgegeben werden. Dabei kann es jedoch gut sein, dass Schwankungen auf mehreren Stufen gleichzeitig zu beobachten sind (*er bäckt/backt; er buk/backte*). Allerdings ist nicht zu erwarten, dass bspw. der Ablaut im Präteritum aufgegeben wird, wenn ein Verb eine stabile Imperativhebung aufweist. Zudem muss die Wechselflexion je nach Frequenz der einzelnen Personen nicht zwingend zugunsten des Infinitivstammvokals nivelliert werden: Bei Verben mit frequenter 3. Person kann auch der Wechselflexionsvokal als Infinitivstammvokal reanalysiert werden (*schmelzen* > *schmilzen*).

Im empirischen Teil der Arbeit wird der Einfluss der Frequenz auf Variation in der Konjugation psycholinguistisch überprüft. Dabei werden starke und schwache Formen von starken Verben mit unterschiedlicher Partizip-II-Frequenz in einer *lexical decision task* gegenübergestellt (*gezogen/*gezieht, geglimmen/geglimmt*) (siehe Abschnitt 4.1.1 für ausführliche Erläuterungen zu *lexical decision tasks*). Anhand der Partizip-II-Frequenz werden frequente und infrequente Verben kontrastiert. Zudem wird innerhalb der infrequenten Verben zwischen Verben unterschieden, bei denen eine Schwankung zwischen starken und schwa-

chen Partizip-II-Formen in Korpora attestiert ist (*geglommen/geglimmt*),³⁴ und Verben, bei denen das (noch) nicht der Fall ist (*gefochten/gefechtet*). Es wird erwartet, dass starke und schwache Partizip-II-Formen tokenfrequenter Verben schneller akzeptiert bzw. abgelehnt werden als Partizip-II-Formen tokeninfrequenter Verben, da die starke Flexion die schwache aufgrund des statistischen Vorkaufsrechts klar ausstechen kann. Bei den Verben mit tokeninfrequenten Partizip-II-Formen ist das aufgrund der niedrigen Frequenz nicht der Fall. Der Blick auf die attestierten und nicht-attestierten Formen ermöglicht es, zu evaluieren, ob Formen bereits als Varianten prozessiert werden, bevor sie in Korpora attestiert sind: Da starke Partizip-II-Formen infrequenter Verben mental nicht mehr so stark gefestigt sind, könnten die Reaktionszeiten von Verben ohne attestierte Schwankung in Korpora bereits vergleichbar sein mit den Reaktionszeiten von den Verben, bei denen Variation bereits in Korpora attestiert ist. Dies würde dafür sprechen, dass Reaktionszeiten bereits einen Hinweis auf Variation liefern, bevor sie in Korpora sichtbar wird. Das Versuchsdesign und die Hypothesen werden ausführlich in Abschnitt 4.2 erläutert, Abschnitt 5.2 stellt die Ergebnisse vor. Der nächste Abschnitt nimmt die Variation in der Deklinationsklasse anhand von schwachen und starken Maskulina in den Blick.

3.2 Variation in der Deklinationsklasse

Die Substantive im Deutschen lassen sich in verschiedene Deklinationsklassen einteilen. Die bestehenden Klasseneinteilungen variieren dabei in ihrer Feingliedrigkeit, je nachdem, ob einzelne Pluralallomorphe zu einer neuen Deklinationsklasse führen oder nicht (Krischke 2012: 57). Da im Folgenden vor allem die Deklination im Singular im Fokus stehen wird, folgt dieser Abschnitt mit Nübling (2008) einer relativ basalen Einteilung in Deklinationsklassen.

Nübling (2008: 283) unterscheidet zwischen drei Deklinationsklassen: schwach, stark und gemischt.

1. schwache Flexion: Genitiv Singular und Plural auf *-en*
2. starke Flexion: weder Genitiv Singular noch Plural auf *-en*
3. gemischte Flexion: Plural mit *-en*, Genitiv Singular nicht mit *-en*
(Übersicht nach Nübling 2008: 283)

³⁴Nicht zu beobachtende Schwankungen in Korpora sind nicht mit inexistenten Schwankungen gleichzusetzen, da Korpora nicht die gesamte Sprachwirklichkeit abdecken können. Dennoch können Korpora als Indiz dafür genutzt werden, ob und wie stark ein Verb bereits zwischen starken und schwachen Formen schwankt.

Sie beschränkt sich somit auf die Markierung des Genitivs Singular sowie der Pluralformen als Kriterien zur Unterscheidung verschiedener Flexionsklassen.³⁵ Dabei kontrastiert sie Formen auf *-en* mit anderen Formen. In diesem Abschnitt werden zusätzlich zwei weitere Deklinationseigenschaften relevant sein: Im Gegensatz zu der schwachen Klasse zeichnen sich starke und gemischt-deklinierte Klassen durch Endungslosigkeit des Dativs und Akkusativs Singular aus (Eisenberg 2013: 153–155). Die starke Klasse markiert den Dativ Plural gesondert auf *-n*, dies ist bei schwachen und gemischt-deklinierten Klassen nicht zu beobachten (Eisenberg 2013: 153–155).³⁶ Anders als in anderen Kategorisierungsansätzen wird Genus in dieser Arbeit nicht als Kriterium für die Unterscheidung von Deklinationsklassen genutzt, da sich Flexionsklassen am Substantiv direkt zeigen, während Genus ein Kongruenzphänomen ist (Nübling 2008: 283).³⁷

Im Folgenden wird die Deklination der Maskulina im Fokus stehen. Für diese ergeben sich folgende Deklinationseigenschaften:

1. schwache Flexion: *der Junge*

- alle Kasus *n*-haltig, außer Nominativ Singular
- Pluralsuffix *-(e)n*

2. starke Flexion: *der Tisch*

- alle Kasus endungslos, außer Genitiv Singular auf *-s*
- Pluralsuffix nicht *-(e)n*
- Dativ Plural gesondert durch *-n* markiert

3. gemischte Flexion: *der Autor*

- alle Kasus endungslos, außer Genitiv Singular auf *-s*
- Pluralsuffix *-en*

³⁵ Anders als in anderen Klassifikationen werden Feminina, die den Plural auf *-(e)n* bilden (*die Blume, die Blumen*) in dieser Einteilung aufgrund der *n*-Losigkeit im Genitiv Singular nicht zu der schwachen, sondern zu der gemischten Flexion gezählt (Nübling 2016: 164).

³⁶ Nübling (2008: 283) bezieht die Bildung des Dativs Plural auf *-n* (*die Tische, den Tischen*) explizit nicht als Kriterium für starke Flexion ein, da sie davon ausgeht, dass die Realisierung dieser Eigenschaft phonologisch konditioniert ist.

³⁷ Einen Ansatz, der Genus und Deklinationsklasse mischt, verfolgt z. B. Eisenberg (2013: 152–156). Auf diese Weise unterscheidet er vier Deklinationsklassen: Starke Maskulina und Neutra, schwache Maskulina, gemischte Maskulina und Neutra sowie Feminina. Innerhalb der Typen setzt er Unterkategorien an, die sich teilweise aufgrund von Pluralallomorphen, teilweise durch silbische und nicht-silbische Suffixe ergeben.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Zwischen diesen Deklinationsklassen lassen sich zahlreiche Grenzgänger beobachten: So wird die Dativ- und Akkusativendung schwacher Maskulina teilweise unterdrückt (*dem/den BärØ* statt *dem/den Bären*), der Genitiv auf *-s* oder *-ns* statt auf *-n* gebildet (*des Bär/s/Bärens* statt *des Bären*) und auch die Nominativform einiger schwacher Maskulina schwankt (*der Friede/n, Wille/n*) (Eisenberg 2013: 153–154). Bei der gemischten Flexion lassen sich Schwankungen hin zur schwachen Flexion beobachten: *des Autoren* statt *des Autors* (Thieroff 2003: 105–106, Köpcke 2005: 78).

In den folgenden Abschnitten wird die Variation in der Deklinationsklasse anhand von Maskulina mithilfe der Einflussfaktoren Frequenz, Prototypizität und Form-Schematizität näher beleuchtet. Dabei werden schwache Maskulina und die oben erwähnten Schwankungsfälle zur starken Flexion im Vordergrund stehen, da hier die meiste Variation zu beobachten ist.

Zunächst wird in Abschnitt 3.2.1 der Einfluss der Token- und Typenfrequenz auf die Variation diskutiert. Hierbei werden theoretische Überlegungen und korpuslinguistische Studien berücksichtigt. Meines Wissens existieren bis auf Schmitt (2019a) keine psycholinguistischen Studien zu Variation in Deklinationsklassen. Studien zur Prozessierung von Pluralallomorphen haben sich vorrangig mit der Frage befasst, welches Pluralallomorph im Deutschen als regelmäßig gelten kann und inwiefern die Verarbeitung von regelmäßigen und unregelmäßigen Pluralformen unterschiedlich verläuft (siehe hierzu Abschnitt 2.1.1).

Abschnitt 3.2.2 geht auf das prototypische Flexionsverhalten starker und schwacher Maskulina ein und formuliert anhand dessen Prognosen darüber, in welcher Abfolge Schwankungen im Paradigma auftreten. Abschnitt 3.2.3 stellt Form-Schemata als deklinationsstabilisierenden Einfluss vor und diskutiert empirische Evidenz hierfür. Abschließend wird in Abschnitt 3.2.4 darauf eingegangen, wie Frequenz, Prototypizität und Form-Schematizität gemeinsam Einfluss auf das Flexionsverhalten der Maskulina im Deutschen nehmen. Im empirischen Teil der Arbeit wird dann der Einfluss von Form-Schemata auf die Variation in der Deklination von Maskulina psycholinguistisch überprüft.

3.2.1 Frequenz

Zum Einfluss der Typen- und Tokenfrequenz bei Maskulina existiert bislang nur wenig empirische Forschung. Schäfer (2019: 385) geht von 450 schwach flektierenden maskulinen Simplicia im Deutschen aus und sieht schwache Maskulina daher als kleine Klasse an, ohne jedoch Mitgliederzahlen starker Maskulina gegenüberzustellen. Pavlov (1995: 45) ermittelt basierend auf 9.685 Einträgen im

3.2 Variation in der Deklinationsklasse

Wörterbuch Wahrig,³⁸ dass schwache Maskulina 17,4 % der insgesamt 2.633 in der Stichprobe enthaltenen Maskulina ausmachen.³⁹ Von den 17,4 % entfallen 1,6 % auf native Wörter und 15,8 % auf Fremdwörter. Auch wenn Ergebnisse genauerer Zählverfahren fehlen, kann die Deklinationsklasse der schwachen Maskulina daher m. E. als typeninrequent bezeichnet werden. Für eine exaktere Bestimmung der Klassengröße wäre es interessant, die Type-Token-Ratio der Substantivklassen zu berücksichtigen, da bei einer typeninrequenten Klasse zu erwarten ist, dass diese geringer ausfällt als bei einer typenfrequenten Klasse.

Die Klasse der schwachen Maskulina scheint anders als die Klasse der starken Maskulina i. d. R. nicht produktiv zu sein. Als Ausnahmefall diskutiert Bittner (2003 [1991]: 92–93) die schwache Adjektivflexion, die dazu führt, dass substantivierte Adjektive schwach flektieren und somit zur Produktivität und Typenfrequenz der Deklinationsklasse beitragen: *schön > der Schöne, des Schönen*. Zudem sind gemischt-deklinierte Lexeme auf *-or (Autor)* als Ausnahmen anzusehen, die zu schwachen Flexionsendungen im Singular schwanken: *des Autoren, dem/den Autoren* (Köpcke 1995: 175–176).⁴⁰

Schäfer (2019) stellt in einer Korpusuntersuchung fest, dass starke und schwache Maskulina unterschiedliche Tokenfrequenzen aufweisen: Schwache Maskulina sind tokenfrequenter als starke (Schäfer 2019: 401–402). Auch bei den Maskulina zeigt sich also das Wechselverhältnis von Typen- und Tokenfrequenz: Hohe Typenfrequenz ist mit geringer Tokenfrequenz gepaart und umgekehrt. Zudem weisen frequente schwache Maskulina eine stabilere schwache Flexion auf als infrequente (Schäfer 2019: 404–405). So flektiert das mit 221.210 Belegen hochfrequente Lexem *Mensch* in Schäfers Korpus stabil schwach, während das mit 51 Belegen weit weniger frequente Lexem *Fink* zwischen starken und schwachen Formen schwankt (Schäfer 2019: 410–411). Aufgrund der geringeren Tokenfrequenz können die schwachen Formen die starken nicht mehr so deutlich dominieren und daher greift das statistische Vorkaufsrecht für schwache Formen nur noch eingeschränkt (zum statistischen Vorkaufsrecht siehe Goldberg 2019: 74–94 sowie Abschnitt 2.1.3).

Die Schwankungen von der schwachen zur starken Flexion lassen sich somit als Anpassung an die typenfrequente Klasse und daher als analogischen Aus-

³⁸Leider gibt Pavlov (1995: 45) nicht an, um welche Ausgabe es sich handelt.

³⁹Pavlov (1995: 45) zufolge machen die gemischt-deklinierten Substantive 3 % der Maskulina aus.

⁴⁰Obwohl die Schwankung nicht von der typenfrequenten starken Klasse, sondern von der noch infrequenteren gemischten Klasse ausgeht, ist dies eine Ausnahme, da aus typenfrequentieller Sicht ein Wechsel zur typenfrequenten und nicht zu einer anderen typeninrequenten Klasse zu erwarten ist. Der Wechsel kann mithilfe von Form-Schemata erklärt werden (siehe Abschnitt 3.2.3).

3 Variation in Konjugation, Deklination und Auxiliarselektion

gleich (Bybee 2010: 66) fassen. In der Aufgabe der schwachen Dativ- und Akkusativendungen im Singular sowie dem Wechsel von *-n* zu *-s* im Genitiv Singular zeigen sich jedoch auch weitere typenfrequentielle Einflüsse. Die Markierung von Dativ und Akkusativ Singular mithilfe des *n*-Suffixes stellt eine Besonderheit dar: Schwache Maskulina weisen als einzige Flexionsklasse Endungen im Dativ und Akkusativ Singular auf (zur Funktion der Markierung siehe Abschnitt 3.2.2).⁴¹ Da schwache Maskulina typeninfrequent sind, hat auch die Markierung von Dativ und Akkusativ eine geringe Typenfrequenz. Eine Deklinationsklasse mit wenigen Types leistet sich also eine Sonderdeklination, die nicht von den anderen Klassen geteilt wird (Thieroff 2003: 113–114, Krischke 2012: 57). Wurzel (1984: 96) bezeichnet die Markierung von Dativ- und Akkusativformen im Singular daher aus Sicht der Natürlichkeitsmorphologie sogar als „systemunangemessen“ (zum Einfluss der Frequenz auf die Natürlichkeit sprachlicher Formen siehe Wurzel 1984: 86) und sieht die Tilgung der Dativ- und Akkusativendung bei schwachen Maskulina (*dem Bär∅/den Bär∅*) daher als logische Anpassung.

In Hinblick auf das *(e)n*-Suffix im Genitiv Singular kann eine analoge Argumentation verfolgt werden: Da alle nicht-femininen Deklinationsklassen bis auf die schwachen Maskulina den Genitiv auf *-(e)s* bilden, ist der Genitiv auf *-(e)n* eine infrequente Flexionseigenschaft. Der Wechsel von *-(e)n* zu *-(e)s* stellt daher nicht nur die Anpassung an die typenfrequente starke Flexion dar, sondern insgesamt an die typenfrequente Flexionseigenschaft *-s* (Thieroff 2003: 113–114). Genauso lassen sich Genitivformen auf *-(e)ns* als Anschluss an die typenfrequente Flexionseigenschaft interpretieren, weil die Formen *s*-haltig sind (Thieroff 2003: 113–114). Krischke (2012: 58–59) nennt die Genitivform auf *-(e)ns* kasusaugmentierend, weil *-s* zusätzlich zu *-(e)n* genutzt wird. Die Form ist aber nur dann kasusaugmentierend, wenn gleichzeitig Dativ- und Akkusativendungen im Paradigma abgebaut werden. Da in diesem Fall alle anderen Kasus endungslos sind, kann *-(e)n* allein genutzt werden, um den Genitiv zu markieren. Ansonsten markiert *-(e)n* lediglich oblique Kasus als solche, nicht aber den Genitiv.

Genitive auf *-(e)ns* sowie endungslose Dativ- und Akkusativformen sind vereinzelt parallel zu beobachten, genauso wie Genitivformen auf *-(e)ns* und *-(e)n*

⁴¹Feminina gehören entweder der starken (*die Wand, der Wand, die Wände*) oder der gemischten (*die Frau, der Frau, die Frauen*) Klasse an und zeichnen sich dementsprechend durch Nullmarkierung der Dativ- und Akkusativendung im Singular aus. Für Neutra ist die schwache Flexion bis auf fossilisierte Formen (*Es tut mir im Herzen weh*) ausgeschlossen und somit sind auch Dativ- und Akkusativendungen im Singular nicht möglich. Neutra flektieren wie Feminina entweder stark (*das Haus, des Hauses, der Häuser*) oder gemischt (*das Auge, des Auges, die Augen*) (Nübling 2008: 298–299). Zwar ist die Endung auf *-e* im Dativ bei starken Maskulina und Neutra vereinzelt noch zu beobachten (*im Laufe, am Tore*), sie ist aber auf archaische Kontexte und feste Wendungen beschränkt, sodass sie kaum ins Gewicht fällt (Eisenberg 2013: 153).

(Krischke 2012: 71). Dies ist aus einer gebrauchsbasierten Perspektive nicht verwunderlich: Da die schwankenden Substantive selten sind, sind die Formen nicht stark gefestigt und parallel auftretende starke und schwache Formen somit zu erwarten, auch wenn diese zusammen ein von den anderen Deklinationsklassen abweichendes Paradigma ergeben würden.

Krischke (2012) ermittelt anhand einer Google-Recherche und einer Korpusuntersuchung im Deutschen Referenzkorpus, wie häufig die Genitivformen auf *-(e)n*, *-(e)ns* und *-(e)s* bei ursprünglich schwachen Maskulina vorkommen. Fünf Substantive (*Bauer*, *Narr*, *Planet*, *Elefant*, *Vorfahr*) weisen in seiner Google-Recherche mehr Genitivformen auf *-(e)ns* als auf *-(e)s* auf (Krischke 2012: 65). Insgesamt sind beide Genitivformen (*-(e)s* und *-(e)ns*) aber infrequent, denn die Genitive auf *-(e)ns* machen in den Google-Ergebnissen von Krischke (2012: 65) maximal 8 % aus, die Genitive auf *-s* maximal 14 %.⁴² Im Deutschen Referenzkorpus, das mit Zeitungstexten vorrangig redigierte Texte enthält, sind die Genitivformen auf *-(e)ns* erwartungsgemäß noch weniger frequent als in den Google-Belegen: Bei den meisten Substantiven liegen die Formen auf *-(e)ns* und *-(e)s* bei unter 2 % (Krischke 2012: 65). Deutlich häufiger wird *des Nachbars* (26 %, 45.000 Belege) sowie *des Magnets* (8 %, 20.000 Belege) gebildet. Der Genitiv auf *-s* wird bei *Nachbar* somit im standardnahen Korpus häufiger verwendet als bei den Google-Belegen (12 %) (Krischke 2012: 65). Aus Krischkes Untersuchung lässt sich festhalten, dass im Genitiv Schwankungen von *-(e)n* zu *-(e)s* und *-(e)ns* beobachtet werden können, wobei die Formen auf *-(e)ns* bei den untersuchten Substantiven seltener zu sein scheinen als die Formen auf *-(e)s*. Im folgenden Abschnitt 3.2.2 wird die Markierung von Genitivformen auf *-(e)ns* als erster Schritt weg von der schwachen hin zur starken Flexion modelliert.

Die Betrachtung der Deklination von Maskulina in Hinblick auf Frequenz verdeutlicht zunächst, dass diesbezüglich noch viel Forschungsbedarf besteht. Es lässt sich jedoch festhalten, dass schwache sowie gemischt-flektierende Maskulina typeninfrequent sind. Zudem zeichnen sich die schwachen Maskulina durch eine Markierung von Dativ- und Akkusativformen im Singular aus, die sich in keiner anderen Deklinationsklasse und in keinem anderen Genus findet. Daher ist zu erwarten und auch zu beobachten, dass infrequente schwache Maskulina zur starken Flexion schwanken (Schäfer 2019: 404–405).

Aufgrund der geringen Typenfrequenz der Klasse und der Flexionseigenschaften sind schwache Formen generell mental nicht stark gefestigt. Somit kann nur eine hohe Tokenfrequenz einzelner Mitglieder der Klasse zu mental gefestigten Repräsentationen der schwachen Flexionsformen führen. Schwankungen

⁴²Die restlichen Prozente lässt Krischke (2012: 65) unerwähnt. Es ist davon auszugehen, dass sie auf die Genitivformen auf *-(e)n* entfallen.

3 Variation in Konjugation, Deklination und Auxiliarselektion

gemischt-deklinierter Substantive wie *Autor* hin zur schwachen Flexion weisen auf Form-Schemata hin, die die Flexion zusätzlich beeinflussen. Zudem deutet die Vielfalt an Varianten (Genitiv auf *-n*, *-ns*, *-s*) darauf hin, dass mehrere Entwicklungslinien für Substantive hin zur starken Flexion existieren. Diese werden im folgenden Abschnitt in den Blick genommen, der den Einfluss von Prototypizität auf die Flexion bei Maskulina diskutiert.

3.2.2 Prototypizität

Das Flexionsverhalten schwacher und starker Maskulina ist prototypisch organisiert. Das prototypische Flexionsverhalten der schwachen Maskulina sieht eine klare Abgrenzung des Nominativs Singular von den Kasus obliqui vor (siehe Tabelle 3.2), denn der Nominativ Singular ist der einzige Kasus der schwachen Deklination ohne Kasusendung. Alle anderen Kasus sind mit *-(e)n* markiert. Somit gilt die Opposition zwischen Nominativ und obliquen Kasus nur im Singular.

Tabelle 3.2: Überblick über die Flexionsparadigmen schwacher und starker Maskulina

	schwache Flexion		starke Flexion	
	Singular	Plural	Singular	Plural
Nom.	der Junge∅	die Jungen	der Stein∅	die Steine
Gen.	des Jungen	der Jungen	des Steins	der Steine
Dat.	dem Jungen	den Jungen	dem Stein∅	den Steinen
Akk.	den Jungen	die Jungen	den Stein∅	die Steine

Im Gegensatz dazu zeichnen sich starke Maskulina durch eine klare Numerusopposition aus, während die meisten Kasus nicht voneinander zu unterscheiden sind. Die einzigen Kasus, die noch systematisch markiert werden, sind Genitiv Singular (*-(e)s*) sowie Dativ Plural (*-n*).

Das Deutsche hat im Lauf(e) seiner Geschichte die Numerusopposition ausgebaut, wohingegen es die Kasusopposition abgebaut hat (Hotzenköcherle 1962: 326–330, Sonderegger 1979, Wegera & Solms 2000: 304–319). So ist bspw. der Dativ auf *-e* nur noch in fossilisierter Verwendung zu finden, z. B. bei Tertiärpräpositionen (*am Rande*, *im Laufe*) (Szczepaniak 2011b: 101). Die Entwicklung ist mit der Relevanzskala nach Bybee (1985) zu erklären: Der Numerus ist für die Flexion relevanter als der Kasus. Der Numerus affiziert die Entitäten, auf die das Substantiv referiert, da er die Anzahl markiert. Durch einen Kasuswechsel wird

eine Entität nicht affiziert, denn der Kasus markiert nur die Beziehung, in der die Entität zu anderen Entitäten steht (Bybee 1985: 34). Der Numerus zählt daher zur inhärenten Morphologie, der Kasus zur kontextuellen. Daher ist es sinnvoll, den Numerus direkt am Substantiv zu markieren, während der Kasus auch syntaktisch markiert werden kann (Dammel & Nübling 2006: 100). Somit weist das Deutsche eine relevanzgesteuerte Markierung der Deklinationsklassen auf (Nübling 2016: 158–159): Die Flexionsklasse wird in der relevanteren Kategorie Plural direkt am Substantiv markiert, im Singular werden die Kasus durch den Artikel markiert, sodass das Genus als Kongruenzphänomen greift (Nübling 2016: 165–167).⁴³ Dies entspricht dem Prinzip der Ikonizität in der Natürlichkeitstheorie, da ein Plus an Bedeutung (Plural) mit einem Plus an Form (Markierung der Deklinationsklasse) einhergeht (Wurzel 1984: 23).⁴⁴

Maskulina lassen sich aufgrund ihrer Flexionseigenschaften dem schwachen oder dem starken Pol zuordnen. Der schwache Pol zeichnet sich durch die Opposition zwischen Nominativ und obliquen Kasus aus, der starke Pol durch die Numerusdichotomie und die Abgrenzung von Genitiv Singular und anderen Kasus (Thieroff 2003: 113–114). In Abbildung 3.4 wird der schwache Pol durch *Junge* veranschaulicht, der starke durch *Stern*.

Nah an der prototypisch schwachen Deklination ist das Substantiv *Gedanke*. Zwar wird der Genitiv mithilfe von *-s* markiert und somit von den anderen *n*-haltigen Kasus abgehoben, davon ist jedoch die deutliche Opposition von Nominativ Singular (\emptyset -Endung) und anderen Kasus (*n*-haltige Endung) nur leicht betroffen. Nur die obliquen Kasus erscheinen aufgrund der neuen Genitivendung nicht mehr so einheitlich wie in der prototypisch schwachen Flexion. Die Nicht-Profilierung der Genitivform lässt sich daher als periphere Eigenschaft der schwachen Flexion ansehen.

Das Genitivsuffix *-s* kann als überstabiler Marker gesehen werden: Überstabile Marker sind Endungen einer stabilen Deklinationsklasse (in diesem Fall die der starken Deklination), die sich auf andere Flexionsklassen ausbreiten (Wurzel

⁴³Nübling (2016: 158–159) sieht darin keinen Rückzug der Deklinationsklassen, sondern eine Verfestigung, da sich die Deklinationsklassen in relevanteren Kategorien zeigen und somit stabilisiert werden.

⁴⁴Zudem spielt semantische Markiertheit eine Rolle: Die semantische Markiertheit eines Numerus bzw. Kasus wird durch den Grad bestimmt, in dem diese Kategorie „prototypische Sprecher-eigenschaften spiegelt“ (Wurzel 1984: 21). Der Plural ist aus Perspektive der Markiertheitstheorie somit eine markierte Kategorie, der Singular hingegen eine unmarkierte, da Sprecher_innen normalerweise einzeln sprechen (Wurzel 1984: 22). Innerhalb der Kasus ist der Nominativ als unmarkiert anzusehen und die anderen Kasus als markiert, da der Nominativ prototypischerweise für das Subjekt genutzt wird (Wurzel 1984: 22).

3 Variation in Konjugation, Deklination und Auxiliarselektion

Nom. Sg. -Ø Gen Sg. -n Dat/Akk.Sg. -n Plural -n Dat .Pl. -Ø <i>der Junge</i> <i>des Jungen</i> <i>dem/den Jungen</i> <i>die Jungen</i> <i>den Jungen</i>	Nom. Sg. -Ø Gen. Sg. -ns Dat/Akk. Sg. -n Plural -n Dat .Pl. -Ø <i>der Gedanke</i> <i>des Gedankens</i> <i>dem/den Gedanken</i> <i>die Gedanken</i> <i>den Gedanken</i>	Nom. Sg. -Ø Gen. Sg. -s Dat/Akk. Sg. -Ø Plural -Ø Dat .Pl. -Ø <i>der Schatten</i> <i>des Schattens</i> <i>dem/den Schatten</i> <i>die Schatten</i> <i>den Schatten</i>	Nom. Sg. -Ø Gen. Sg. -n Dat/Akk. Sg. -Ø Plural -n Dat .Pl. -Ø <i>der Bär</i> <i>des Bären</i> <i>dem/den Bär</i> <i>die Bären</i> <i>den Bären</i>	Nom. Sg. -Ø Gen. Sg. -s Dat/Akk. Sg. -Ø Plural Umlaut,-n, -s Dat .Pl. -Ø <i>der Schaden</i> <i>des Schadens</i> <i>dem/den Schaden</i> <i>die Schäden</i> <i>den Schäden</i> <i>der Schmerz</i> <i>des Schmerzes</i> <i>dem/den Schmerz</i> <i>die Schmerzen</i> <i>den Schmerzen</i>	Nom. Sg. -Ø Gen. Sg. -s Dat/Akk. Sg. -Ø Plural -e,-er Dat. Pl. -n <i>der Stern</i> <i>des Sterns</i> <i>dem/den Stern</i> <i>die Sterne</i> <i>den Sternen</i>
--	--	--	--	--	--

N ≠ D = A	N ≠ D = A	N = D = A
G ≠ N G = D = A	G ≠ N G ≠ D = A	
Singular = Plural		Singular ≠ Plural

prototypisch schwach

prototypisch stark

Abbildung 3.4: Prototypisch schwache und prototypisch starke Deklination

1984: 139).⁴⁵ Wurzel (1984: 140) sieht Indizien für einen überstabilen Marker darin, dass doppelte Endungen zugelassen werden. Es ist jedoch fraglich, ob dies bei dem Genitiv auf *-ns* tatsächlich der Fall ist, da erst *-s* eine klare Unterscheidung von Genitiv und anderen obliquen Kasus ermöglicht.

Einen Schritt weiter entfernt vom Prototyp der schwachen Deklination besteht keine klare Dichotomie zwischen Nominativ und obliquen Kasus, dafür eine klare Dichotomie zwischen Genitiv und anderen Kasus: Da der Nominativ wie die anderen Kasus *n*-haltig ist, ist er nicht von ihnen zu unterscheiden. Der Genitiv *des Schattens* ist dagegen klar profiliert, da er als einziger Kasus im Singular eine Markierung aufweist. Eine eindeutige Numerusdichotomie hat *der Schatten* aber nicht, er teilt daher Eigenschaften mit der schwachen Flexion.

Eine eindeutig markierte Genitivform sowie eine klare Numerusdichotomie weisen monosyllabische Substantive auf, die die schwache Flexion zugunsten der starken aufgeben: Dies lässt sich bei *Bär* beobachten. *Bär* weist derzeit noch schwache Formen auf, aber auch Apokopen der *n*-haltigen Endung im Dativ und

⁴⁵Dammel & Nübling (2006: 100–101) sehen überstabile Marker als Zeichen für Flexionsschwäche an, da diese nur einen geringen Grad an Fusion aufweisen und sich lediglich in wenig relevanten Kategorien (Singular) zeigen. Allomorphe sind Gegenspieler der überstabilen Marker; diese können einen hohen Grad an Fusion aufweisen (wie bspw. die Pluralmarkierung durch Umlaut) und markieren vor allem relevante Kategorien (Plural).

Akkusativ: *dem/den Bär*∅ statt *dem/den Bären* (Köpcke 2005: 69). Die Apokopen führen zu einer gesonderten Markierung der Genitiv- und Pluralformen auf *-en*. Der Genitiv und die Pluralformen sind jedoch formgleich.

Nur einen Schritt von der prototypisch starken Flexion entfernt sind Maskulina, die zwar den Genitiv auf *-s* markieren und auch eine Numerusdichotomie aufweisen, allerdings den Dativ im Plural nicht gesondert markieren. Dies ist bspw. der Fall bei *Schmerz*.⁴⁶ Da *Schmerz* im Singular nur starke Flexion aufweist, im Plural dagegen mit *-en* schwache Flexion, ist dieses Substantiv der gemischten Deklination zuzuordnen (Nübling 2008: 283). Die gemischte Deklination erscheint somit als Übergangsbereich zwischen schwachen und starken Flexionseigenschaften. Dies wird auch in der Benennung der Deklinationsklasse deutlich. Substantive der gemischten Flexion weisen nur Pluralformen auf *-en* als schwache Flexionseigenschaft auf. Pluralformen auf *-en* können daher als prototypisch für die schwache Deklination gelten, da auch Substantive mit peripher schwachem Flexionsverhalten diese Eigenschaft aufweisen. Substantive mit Umlaut als Pluralmarkierung sind ebenfalls nah an der prototypisch starken Flexion: *Schaden* kann mithilfe der Umlautung eindeutig Plural vom Singular abgrenzen (*die Schäden*). Da der Plural *n*-haltig ist, kann der Dativ Plural nicht gesondert markiert werden. Auch Substantive mit Plural auf *-s* können den Dativ Plural nicht gesondert markieren. Der *s*-Plural findet sich vorrangig bei Entlehnungen und Kurzwörtern wie bspw. *der Blog - die Blogs* oder *der Azubi - die Azubis*.⁴⁷ Einen Sonderfall bilden Substantive wie *Löffel* und *Hammer*. Diese weisen zwar keine generelle Numerusdichotomie auf (*ein Hammer/Löffel, viele Hammer/Löffel*), allerdings bilden sie den Dativ-Plural auf *-n*. Hier ist davon auszugehen, dass phonologische Faktoren die Bildung des Plurals auf *-e* blockieren: Durch Hinzufügen des *-e* würde einen Dreisilbler mit zwei reduzierten Silben entstehen (**die Hammere,* die Löffele*), diese Konstruktionen werden im Deutschen generell gemieden (Eisenberg 1991: 47–49).

Die Prototypizitätsskala lässt sich nutzen, um die Reihenfolge von Schwankungen hin zur starken Deklination zu prognostizieren. Beim Übergang von der schwachen in die starke Flexion wird zunächst die periphere Eigenschaft der Nicht-Profilierung des Genitivs aufgegeben und die Genitivmarkierung⁴⁸ pro-

⁴⁶Der Dativ Plural kann hier nicht gesondert markiert werden, da alle Pluralsuffixe *n*-haltig sind.

⁴⁷Im Bereich der nicht-nativen Substantive sowie der Kurzwörter sind auch Nullendungen im Genitiv möglich (*des Blog*∅, *des Azubi*∅). Da diese Substantive in die Peripherie der deutschen Substantive einzuordnen sind, greift hier, ähnlich wie bei Eigennamen, das Prinzip der Morphemkonstanz: Um periphere Strukturen besser verarbeiten zu können, sollen sie möglichst wenig verändert werden (Ackermann & Zimmer 2017: 151–154).

⁴⁸Die Schwankungen beginnen daher in der Kategorie mit der niedrigsten Relevanz. Dammel & Nübling (2006: 100–101) sehen auch überstabile Marker wie *-(e)s* im Genitiv als Zeichen für die geringe Relevanz einer Kategorie.

filiert. Dies kann entweder durch Tilgung des *n*-haltigen Suffixes im Dativ und Akkusativ Singular passieren oder durch Genitivformen auf *-(e)ns*. Das erste Verfahren profiliert nicht nur den Genitiv Singular, sondern löst gleichzeitig die Dichotomie zwischen Nominativ und obliquen Kasus auf und stärkt die Numerusdichotomie, da das Singularparadigma nun bis auf den Genitiv endungslos ist. Dies geschieht bei den Formen auf *-(e)ns* erst in weiteren Schritten: Zunächst wird der Nominativ als *n*-haltig reanalysiert. *-n* kann somit als Teil des Stamms und nicht mehr als Flexionsendung angesehen werden (Eisenberg 2013: 154–155). Anschließend wird der Plural durch Umlautung markiert (Nübling u. a. 2013: 70).

Welcher Entwicklungsweg eingeschlagen wird, hängt von den phonologischen Eigenschaften des Substantivs ab: Monosyllabische Substantive verlieren ihre Dativ- und Akkusativendungen und profilieren auf diese Weise den Genitiv, der nun als einziger Kasus suffigiert wird. Dieser Weg lässt sich bspw. für die frnhd. Einsilbler *Stern* und *Mond* nachvollziehen (Köpcke 2000a: 120–121). Mit dem Verlust der Kasusendung im Singular wird der Plural profiliert. Der Genitiv auf *-en* passt sich im nächsten Schritt der starken Deklination an und wird stark gebildet: *des Sternens* wird zu *des Sterns*.⁴⁹

Synchron lassen sich parallele Schwankungen zwischen starken und schwachen Genitiv-, Dativ- und Akkusativformen beobachten: *des Bärs/Bären, dem/den Bär∅* (Krischke 2012: 70–72). Schäfer (2019: 402) stellt in seinem Korpus jedoch fest, dass bei monosyllabischen Substantiven Dativ und Akkusativ eher zu starken Formen neigen als der Genitiv. Zudem kann der Genitiv bereits durch wegfallende Dativ- und Akkusativendung profiliert werden. Daher scheint der Wegfall der Dativ- und Akkusativendungen vor der Genitivbildung auf *-s* einzusetzen (Schäfer 2019: 396). Dennoch ist nicht von einer strikten Abfolge der hier postulierten Stufen auszugehen, da sich gleichzeitig Schwankungen auf verschiedenen Stufen beobachten lassen.

Substantive auf Schwa schlagen den zweiten Entwicklungsweg ein: Substantive wie frnhd. *der Schade* nahmen zunächst *-s* im Genitiv an (*des Schadens* statt *des Schaden*), anschließend wurde *-n* als Teil des Stamms reanalysiert und deswegen auch im Nominativ genutzt: *der Schade* > *der Schaden* (Köpcke 1995: 172, 2000a: 113–114). Durch die Reanalyse wird der Singular dem Plural ähnlicher gemacht,

⁴⁹Poitou (2004a: 77) geht hiervon abweichend davon aus, dass der Genitiv wie Dativ und Akkusativ endungslos wird. Dies macht er an einer Google-Untersuchung fest, in der er für schwache Maskulina mehr endungslose Genitive als Genitivformen auf *-s* feststellte. Er nennt dabei nur das Substantiv *Mensch* als Beispiel, bei dem der endungslose Genitiv häufiger auftrat als der Genitiv auf *-s*. Hierbei ist anzumerken, dass die Kombination /ʃ s/ phonotaktisch ungünstig ist. Zudem scheint Poitou (2004a) keine Genitive auf *-en* zu berücksichtigen. Krischke (2012: 64) kann in seinem Korpus keine nennenswerte Anzahl von endungslosen Formen feststellen. Die belegten endungslosen Formen sind zudem oft als Eigennamen analysierbar: *Vorgaben des Paragraph 301 SGB V* (Krischke 2012: 64).

da die Nominativformen keine Numerusinformation mehr enthalten.⁵⁰ Die Numerusdichotomie wurde durch Umlautung im Plural profiliert: *der Schaden, die Schaden > der Schaden, die Schäden*. (Nübling u. a. 2013: 70).

Das Ergebnis beider Entwicklungswege ist die Aufgabe der Kasusopposition (Nominativ vs. oblique Kasus) zugunsten einer Numerusopposition. Als einziger Kasus im Singular wird der Genitiv gesondert markiert, ihm kommt eine Sonderrolle zu, da er durch den überstabilen Marker *-(e)s* gekennzeichnet wird (Wurzel 1984: 140). Es ist zudem festzuhalten, dass die Relevanzhierarchie die Schwankungen beeinflusst: Die Schwankungsfälle betreffen zunächst nur die Kasus im Singular, erst im Anschluss daran treten Schwankungen im Plural auf. Es ist daher anzunehmen, dass die Pluralmarkierung auf *-n* als Eigenschaft der schwachen Flexion am längsten beibehalten und erst spät aufgegeben wird. Köpcke (1993: 138) nennt als Schwankungsfälle im Nominativ Plural *die Greife/n*, und *die Spatze/n*.⁵¹ Die *n*-losen Formen müssten sich auf die weiteren Kasus ausdehnen, um eine Profilierung des Dativs zu ermöglichen (*das Singen der Spatze, Ich füttere die Spatze*). Nübling (2008: 283) sieht die Markierung des Dativs Plural allerdings nicht als Deklinationsklasseneigenschaft, da sie davon ausgeht, dass die Markierung aus phonologischen Gründen ausbleibt, wenn das Pluralsuffix auf *-n* endet.⁵²

Der Blick auf die prototypischen Eigenschaften schwacher und starker Flexion bei Maskulina ermöglicht einen differenzierteren Blick auf Deklinationsklassen als die Einteilung in schwach, stark und gemischt. Zudem ermöglicht er eine Prognose der Richtung und Reihenfolge von Schwankungen. Aufgrund der geringen Typenfrequenz schwacher Maskulina ist von Schwankungen hin zur starken Flexion auszugehen. Eine Ausnahme von dieser Entwicklungsrichtung ist das gemischt-deklinierte Substantiv *der Autor*, das vereinzelt schwache Formen im Singular aufweist (siehe hierzu genauer Abschnitt 3.2.3 zu Form-Schemata der schwachen Maskulina).

⁵⁰Nübling (2008: 305) nennt diese Deklinationsweise gemischt, da starke und schwache Deklinationsklasseneigenschaften kombiniert werden. Substantive können diese Flexionseigenschaften dauerhaft aufweisen oder sie als Übergang nutzen (Nübling 2008: 305). Allerdings handelt es sich nicht um eine gemischte Deklination im Sinne der oben eingeführten. Da der Nominativ Singular bereits auf *-en* endet, kann *-en* in diesem Fall nicht als Pluralmarker interpretiert werden, wie es bei gemischt-deklinierten Substantiven wie *Autor* der Fall ist.

⁵¹Woran Köpcke (1993: 138) festmacht, dass es sich um Schwankungsfälle handelt, erläutert er leider nicht.

⁵²Die gesonderte Markierung des Dativs Plural steht zudem im Widerspruch zur Aufgabe der Numerusopposition der Deklinationsklassen. Es zeigen sich bereits erste Abbautendenzen des Dativs Plural: Auf Hinweis- und Preisschildern ist vereinzelt eine *n*-lose Flexion zu beobachten (*Zutritt ab 18 Jahre, Risotto mit Pilze*). Hierbei könnte es sich allerdings auch um unflektierte Substantive handeln, da die Phrasen determiniererlos sind (Ágel 2008: 71–72).

3 Variation in Konjugation, Deklination und Auxiliarselektion

Hinsichtlich der Schwankungen zur starken Flexion lassen sich zwei Entwicklungslinien ansetzen: Die erste Entwicklungslinie betrifft ursprünglich schwach flektierende Substantive auf Schwa. Diese bilden zunächst Genitive auf *-ns* (*des Schadens*) und nehmen anschließend *n*-haltige Nominativformen an (*der Schaden*). Die Profilierung der Numerusopposition erfolgt in einem separaten Schritt durch Umlautung (*die Schäden*). Somit ist ein *n*-haltiger Nominativ unwahrscheinlich, solange der Genitiv stabil auf *-n* gebildet wird (bspw. *der *Matrosen, des Matrosen*), weil ansonsten eine Schwankungsstufe übersprungen wäre. Der Genitiv muss aber nicht stabil auf *-ns* gebildet werden, bevor der Nominativ Schwankungen aufweist.

Die zweite Entwicklungslinie betrifft monosyllabische Substantive. Im Gegensatz zu den mehrsilbigen Substantiven verlieren diese Substantive zunächst die Dativ-/Akkusativendung (*dem/den Bär*∅). Damit geben sie die Dichotomie zwischen Nominativ und Dativ sowie Akkusativ auf und profilieren gleichzeitig den Genitiv. In einem zweiten Schritt wird der Genitiv auf *-s* gebildet. Bei beiden Entwicklungslinien ist von irreversiblen Prozessen auszugehen, da sie weg von der typeninfrequenten und hin zur typenfrequenten Klasse führen.

Wie dieser Abschnitt bereits zeigte, beeinflussen phonologische Eigenschaften den Wandel von schwachen Maskulina zu starken. Der folgende Abschnitt erläutert die phonologischen sowie prosodischen und semantischen Eigenschaften schwacher Maskulina und deren Einfluss auf die Flexion.

3.2.3 Form-Schematizität

Köpcke (1993, 1995, 2000a,b, 2005) zeigt, dass die Klasse der schwachen Maskulina durch Form-Schemata definiert ist (zur Abgrenzung von Schemata und Form-Schemata siehe Abschnitt 2.3). Schwache Maskulina zeichnen sich ihm zufolge durch folgende Eigenschaften aus: Dreisilbigkeit, Pänultimabetonung⁵³, Schwachendung und die semantische Eigenschaft [+menschlich] wie bspw. bei *Matröse* und *Geselle* (Köpcke 1995: 168–170). Neben diesem Form-Schema setzt Köpcke (1995: 174) ein zweites an, das sich aus dem ersten ableitet: Das zweite Form-Schema endet anders als das erste Form-Schema nicht auf Schwa und ist ultimabetont (*Spekulánt*). Köpcke (1995: 174) wertet es als Abspaltung vom ersten

⁵³Köpcke geht in keiner seiner Veröffentlichungen darauf ein, warum die Pänultimabetonung für das Form-Schema relevant ist. Zum einen könnte die Relevanz der Pänultimabetonung auf Frequenz zurückgehen, da alle dreisilbigen Substantive, die schwach flektieren, auch auf der Pänultima betont sind. Zum anderen führt die Dreisilbigkeit zusammen mit der Pänultimabetonung dazu, dass die schwachen Maskulina sich von anderen Substantiven abheben, da sie nicht dem trochäischen Betonungsmuster folgen. Sie haben daher eine hohe *cue validity*.

Form-Schema, da „das Vorhandensein bzw. Fehlen des auslautenden Schwas genau genommen der einzige Unterschied zwischen den Prototypen ist“ (Köpcke 1995: 174). Die Ultimabetonung folgt lediglich aus dem Fehlen des Schwa (Köpcke 1995: 175).⁵⁴

Da die Klasse der schwachen Maskulina weitgehend auf Entitäten mit hohem Belebtheitsgrad beschränkt ist (Bittner 2003 [1991]: 74, Köpcke 1995: 169–170), stellt sie in semantischer Sicht einen Sonderfall dar, da die anderen Klassen nicht semantisch definiert sind (Nübling 2016: 154–155). Belebtheit ist ein kognitives Konzept, das sich auf die Lexik und Grammatik von Sprachen auswirken kann. Hierbei werden Entitäten in Bezug auf Eigenschaften des Menschen in Kategorien geteilt (Enger & Nessel 2011: 208). Da es sich um ein anthropozentrisches Konzept handelt, befindet sich der Mensch weit oben auf der Belebtheitskala. In ihrer einfachsten Form werden nur die Kategorien MENSCHLICH, BELEBT und UNBELEBT angesetzt (Silverstein 1976: 122–123, Comrie 1989: 185). Bei der feineren Untergliederung in Belebtheitsstufen spielt *embodiment* eine Rolle: Der Körper von Menschen ist anders aufgebaut als der von Tieren, somit werden Tiere anders wahrgenommen als Menschen (Enger & Nessel 2011: 208, Flick 2020: 94–96). Tiere sind zudem weniger agentiv als Menschen (Szczeponiak 2011a: 344–345), haben also weniger Handlungsmöglichkeiten. Pflanzen sind zwar belebt, aber dem Menschen in ihrem Aufbau noch unähnlicher als Tiere und sie können sich nur eingeschränkt bewegen (bspw. zur Sonne ausrichten), weshalb sie weniger agentiv und damit weniger belebt sind als Tiere. Es folgen Konkreta, da sie nicht belebt, aber materiell sind. Abstrakta machen den Gegenpol zu Menschen aus, da sie weder belebt noch materiell sind (Szczeponiak 2011a: 344–345).

Der Belebtheitsgrad einer Entität hängt eng mit ihrem Individuiertheitsgrad zusammen (Flick 2020: 104). Das Konzept der Individuiertheit beschreibt, wie gut sich eine Entität von anderen abgrenzt, und bestimmt somit das Potential der Entität, als Individuum wahrgenommen zu werden (Hopper & Thompson 1980: 253, Szczeponiak 2011a: 344–346, Flick 2020: 103). Die Individuiertheit hängt zum Teil von Körperlichkeit ab: Ist etwas klar konturiert, kann die Entität leichter abgegrenzt werden als eine unklar konturierte Entität wie bspw. Massennomina wie *Wasser* und *Milch* (Szczeponiak 2011a: 345). Zusätzlich hängt Individuiertheit an der Pluralität sowie der Zählbarkeit einer Entität: Entitäten im Singular sind individuierter als Entitäten im Plural, das Gleiche gilt für zählbare Entitäten im

⁵⁴Mit der Unterscheidung von zwei Form-Schemata, die mit schwacher Flexion assoziiert sind, baut Köpcke (1995: 1995) auf Bittner (2003 [1991]: 113) auf, die ebenfalls zwei Form-Schemata für schwache Maskulina ansetzt: Substantive auf Schwa mit belebten Referenten und endbetonte Substantive mit nicht-nativen Suffixen, die prototypischerweise belebte Referenten aufweisen, aber nicht zwingend. Sie gliedert die Form-Schemata jedoch nicht weiter auf.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Vergleich zu unzählbaren (für eine Übersicht der Einflussfaktoren auf Individuiertheit siehe Hopper & Thompson 1980: 253, Flick 2020: 104).

Menschen kommt schon allein deshalb ein hoher Individuiertheitsgrad zu, weil sie über große Agentivität verfügen. Sie sind somit kognitiv auffällig, weswegen sie als Individuen wahrgenommen werden (Epstein 1994: 71, Flick 2020: 104). Ein hoher Grad an Belebtheit korreliert also mit einem hohen Grad an Individuiertheit (Szczepaniak 2011a: 345, Flick 2020: 104).⁵⁵ Individuiertheit und *embodiment* können genutzt werden, um die Belebtheitskategorien weiter aufzufächern, siehe Abbildung 3.5.

Menschen	Säugetiere	Vögel	Fische	Reptilien	Insekten	Konkreta	Abstrakta
maximal belebt	maximal unbelebt

Abbildung 3.5: Belebtheitsskala nach Köpcke (2000a: 115)

Mit einem Hund hat der Mensch mehr Gemeinsamkeiten als mit einer Biene, die statt vier sechs Extremitäten hat (Flick 2020: 94). Zudem sind einzelne Bienen weniger gut unterscheidbar als einzelne Hunde. Da Bienen als Schwarmtiere prototypischerweise in der Mehrzahl auftreten, ist die Individuiertheit außerdem aufgrund der Pluralität herabgesetzt. Daher werden Hunde eher als Individuen wahrgenommen als Bienen.

Der hohe Belebtheitsgrad von schwachen Maskulina kann als Erklärung für die unterschiedliche Markierung von Nominativ und obliquen Kasus in der Deklination schwacher Maskulina genutzt werden (Köpcke 2000a: 117–118). Der Nominativ markiert prototypischerweise das Subjekt und diese syntaktische Funktion korreliert meist mit der semantischen Rolle des Agens (Köpcke 2000a: 118). Diese Rolle ist prototypischerweise mit Entitäten besetzt, die handlungsfähig sind und somit einen hohen Belebtheitsgrad aufweisen (Langacker 1991: 248, Primus 2012: 16–17, Flick 2020: 103–104).⁵⁶ Aufgrund der Korrelation zwischen Belebtheit, Agens und Nominativ ist es für Substantive mit hohem Belebtheitsgrad sinnvoll, die Agensrolle im Gegensatz zu den anderen semantischen Rollen unmarkiert zu lassen (Köpcke 2000a: 118): Da Entitäten mit hohem Belebtheitsgrad potentiell handlungsfähig sind, ist es generell wahrscheinlich, dass sie Agens einer

⁵⁵Jedoch können sich Belebtheit und Individuiertheit auch orthogonal zueinander verhalten: Zwar verfügen Abstrakta über einen geringen Belebtheits- sowie Individuiertheitsgrad, ein Abstraktum kann jedoch bspw. mithilfe eines Determinierers individuiert werden (*das Glück, dein Verstand*). Entitäten können daher unabhängig vom Belebtheitsgrad individuiert sein (Flick 2020: 111)

⁵⁶Die Rolle des Agens kann auch mit nicht-belebten Entitäten besetzt sein, dies ist aber nicht die prototypische Besetzung derselben, siehe hierzu genauer Flick (2020: 114–115).

Handlung sind. Daraus folgt, dass die semantischen Rollen gesondert markiert werden, bei denen diesen Entitäten unerwarteterweise keine handlungstragende Funktion zukommt. Die auffällige Markierung der Dativ- und Akkusativformen ist aus dieser Perspektive eine Konsequenz der Belebtheit schwacher Maskulina. Diese Hypothese überprüfen Binanzer u. a. (2020): Sie präsentieren in einer Eye-Tracking-Studie Sätze in SVO- und OVS-Abfolge mit starken und schwachen Maskulina. In der Studie wurden die Sätze auditiv präsentiert und Bilder gezeigt, die in Beziehung zum Sprachmaterial standen (Bei *Welcher Arzt beobachtet die Frau?* und *Welchen Arzt beobachtet die Frau?* waren Figuren zu sehen, die man leicht als Arzt bzw. als eine weibliche Person identifizieren konnte). Erwachsene Proband_innen konnten die Sätze und das Bildmaterial erfolgreicher verknüpfen, wenn schwache Maskulina anstatt starker als Patiens genutzt wurden, bei Kindern zeigte sich dies nur in der SVO-Abfolge (Binanzer u. a. 2020: 399). Obwohl die Markierung durch den Determinierer erfolgt, scheint somit die gesonderte Markierung der Patiensrolle zu einem Prozessierungsvorteil zu führen. Hier lässt sich wiederum das Konzept der Relevanz anführen: Für belebte Entitäten ist die Opposition zwischen Nominativ und obliquen Kasus aufgrund der möglichen semantischen Rollen relevanter, weshalb diese direkt am Substantiv markiert werden.

Wie bei den starken und schwachen Verben stehen sich auch bei schwachen und starken Substantive eine schlecht generalisierbare, weil auf wenige und wenig variable Substantive beschränkte, und eine hochschematische Deklinationsklasse gegenüber. Dies lässt sich vorrangig für den Singular postulieren. Für den Plural ist die Lage komplexer, da nicht nur *-(e)n* und *-e* möglich sind, sondern auch weitere Suffixe. Die folgenden Ausführungen fokussieren auf den Singular. Anders als bei den Form-Schemata der starken Verben sind bei dem Form-Schema der schwachen Maskulina nicht nur prosodisch-phonologische Eigenschaften (Dreisilbigkeit, Pänultimabetonung, Schwa) mit einem bestimmten Flexionsverhalten verknüpft, sondern auch eine semantische (Belebtheit). Dies lässt sich gut durch einen Zweischritt modellieren, wie Abbildung 3.6 zeigt.

Zunächst ist von einer teilschematischen Konstruktion auszugehen, die die Form [dreisilbig, pänultimabetont, Schwa-Endung] mit der Funktion [+menschlich] assoziiert. Diese Konstruktion ist Teil des Form-Schemas: Die Eigenschaften [dreisilbig, pänultimabetont, Schwa-Endung, +menschlich] sind mit schwachem Flexionsverhalten verknüpft. Das Form-Schema ist in diesem Fall also komplex verschachtelt, da nicht nur die Verknüpfung von einzelnen Flexionsformen mit einzelnen Funktionen gemeinsam eine Assoziation von Form auf Lexemebene und Flexionsverhalten ergeben (siehe hierzu genauer Abschnitt 2.3), sondern die Form zusätzlich eine teilschematische Konstruktion darstellt.

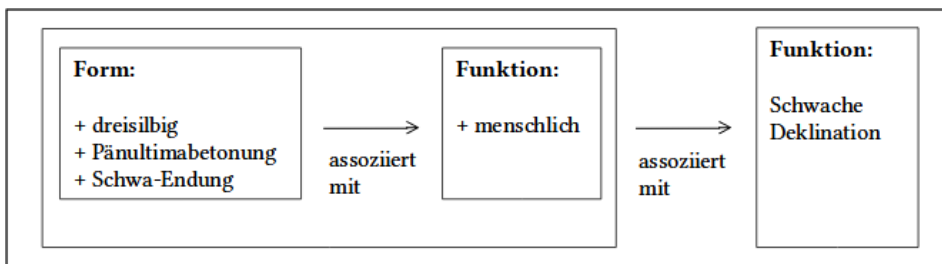


Abbildung 3.6: Verschachtelung des Form-Schemas bei schwachen Maskulina

Insgesamt lässt sich festhalten, dass die Klasse der schwachen Maskulina komplex konditioniert ist, da semantische, prosodische sowie phonologische Faktoren⁵⁷ die Klassenzugehörigkeit beeinflussen (Nübling 2016: 172). In den folgenden Abschnitten 3.2.3.1 und 3.2.3.2 werden die prototypische Organisation und der Einfluss der Form-Schemata auf die Flexion von Maskulina näher in den Blick genommen. Dabei wird mit *Form-Schema I* auf das Form-Schema mit den Eigenschaften [Dreisilbigkeit, Pänultimabetonung, Schwa-Endung, + menschlich] referiert, mit *Form-Schema II* auf das Form-Schema mit den Eigenschaften [Dreisilbigkeit, Ultimabetonung, + menschlich].

3.2.3.1 Form-Schema I

Der phonologischen Form [Dreisilbigkeit, Pänultimabetonung, Schwa-Endung] entsprechen gut 80 native Substantive, hinzu kommen etwa 200 nicht-native Substantive (Köpcke 1995: 171). Die meisten Substantive mit dieser Form haben einen hohen Belebtheitsgrad: Nur ein Substantiv (*Gedanke*) referiert auf eine abstrakte Entität und weitere fünf auf belebte, aber nicht-menschliche Entitäten (bspw. *Schimpanse*). *Gedanke* zeigt allerdings bereits Übergangstendenzen zur starken Deklinationsklasse: Der Genitiv wird auf *-ns* gebildet (*des Gedankens*) und laut Köpcke (1995: 171) existiert auch die *n*-haltige Nominativform (*der Gedanken*).⁵⁸ Die restlichen dreisilbigen Substantive auf Schwa referieren auf menschliche Entitäten. Dreisilbigkeit, Pänultimabetonung und Schwa-Endung scheinen also stark mit einem hohen Belebtheitsgrad assoziiert zu sein und damit eine hohe *cue validity* zu haben (zum Konzept der *cue validity* siehe Abschnitt 2.2.1).

⁵⁷Wenn man das finale Schwa als Belebtheitsmarker interpretiert, ist die Klassenzugehörigkeit zusätzlich morphologisch konditioniert (Nübling 2016: 172).

⁵⁸Woran er diese Aussage festmacht, erläutert Köpcke (1995) jedoch nicht.

3.2 Variation in der Deklinationsklasse

Schäfer (2019: 395–396) geht davon aus, dass die prosodisch-phonologischen Eigenschaften eine höhere Salienz haben und somit stärker zur Profilierung der schwachen Maskulina beitragen als die semantischen, da ein hoher Belebtheitsgrad nicht automatisch zur schwachen Flexion führt (**des Fischern*). Zudem weist er darauf hin, dass die dreisilbigen Substantive mit Pänultimabetonung von den typischen prosodisch-phonologischen Eigenschaften deutscher Substantive (monosyllabisch oder trochäisch) abweichen und daher salient sind.

Form-Schema I ist produktiv: So werden Einwohnerbezeichnungen nach diesem Form-Schema gebildet und schwach flektiert, wie bspw. *der Kasache* (Köpcke 2000a: 112). Ein weiteres Indiz für die Produktivität des Form-Schemas liefern Substantive auf *-or* (*der Autor*): Diese gehören der gemischten Klasse an, da sie den Singular stark bilden, den Plural jedoch auf *-en*. Die Pluralformen auf *-en* machen aus der Anfangsbetonung (*Áutor*) eine Pänultimabetonung (*Autóren*). Im Plural entsprechen Bildungen auf *-or* also der phonologischen Form des Form-Schemas I. Zudem passt *der Autor* aufgrund seines hohen Belebtheitsgrads in das Form-Schema I. Köpcke (2005: 78) sieht darin den Grund dafür, dass dieses Substantiv vereinzelt auch im Singular schwach dekliniert wird (*des Autoren* statt *des Autors*). In seiner Google-Untersuchung liegen die schwachen Belege allerdings lediglich bei 1 % bis 5 %. Nur bei *Juror* sind sie mit 17 % häufiger vertreten. Die empirischen Befunde scheinen somit nur auf einen kleinen Effekt hinzudeuten. Schäfer (2019: 406–407) kann in seiner Korpusstudie zeigen, dass Belebtheit dabei eine Rolle spielt: Substantive auf *-or* werden mit menschlichen Referenten eher schwach dekliniert als mit unbelebten Referenten. Interessanterweise wird der Genitiv bei diesen Substantiven laut Schäfer häufiger schwach dekliniert als Dativ und Akkusativ. Die Substantive auf *-or* verhalten sich also invers zu den schwachen Maskulina, die zur starken Flexion schwanken (siehe hierzu genauer Abschnitt 3.2.2).

Köpcke (1995: 169) sieht in Entlehnungen, die dem Form-Schema I der schwachen Maskulina ähnlicher gemacht wurden, ein weiteres Indiz für dessen Produktivität: Bei *Mormone* und *Schamane* wurde der ursprüngliche Wortakzent weg von der Erstsilben- hin zur Pänultimabetonung verschoben, außerdem haben die Substantive Schwa angenommen: *mórmon* > *Mormóne* (Köpcke 1995: 169). Auch native Substantive nehmen Schwa an, um in der schwachen Flexion zu bleiben: *der Vor-/Nachfahr* > *der Vor-/Nachfahre* (Köpcke 1993: 137–138), allerdings führt dies nicht zu einer Akzentverschiebung (*der Vórfahre*). Laut Kusová (2014: 113–115) ist bei *Vorfahr(e)* die Form auf Schwa im Deutschen Referenzkorpus mit 67 % frequenter als die schwa-lose. Bei *Exot(e)* stellt sie hingegen eine klare Tendenz zur Form ohne Schwa fest. Durch die Ultimabetonung (*Exót*) passt das Substantiv

3 Variation in Konjugation, Deklination und Auxiliarselektion

bereits zu Form-Schema II, daher ist ein Annehmen von Schwa wie bei *Mormone* nicht nötig, um mit schwachem Flexionsverhalten assoziiert zu werden.

Die Abkopplung des Form-Schemas II für schwache Maskulina durch Schwa-Apokope zeigt, dass neben der Annahme von Schwa auch der Wegfall möglich ist. Dies lässt sich bei einigen Entlehnungen nachverfolgen: *Bojare* > *Bojar* (Köpcke 1995: 168). Auch hier entsteht eine Ultimabetonung, sodass das Substantiv zu Form-Schema II passt.

Form-Schema I ist prototypisch organisiert. Während nur knapp 300 Substantive die prototypische Form aufweisen, zählen inklusive der Peripherie ca. 500 Substantive zum Form-Schema I (Bittner 2003 [1991]: 98–99). Dabei existieren sowohl ein prosodisch-phonologischer Prototyp (Dreisilbigkeit, Pänultimabetonung und Schwa-Endung) und eine prosodisch-phonologische Peripherie (Monosyllabizität) als auch ein semantischer Prototyp (Mensch) und eine semantische Peripherie (Unbelebtheit). Abbildung 3.7 stellt die Abstufungen schematisch dar.

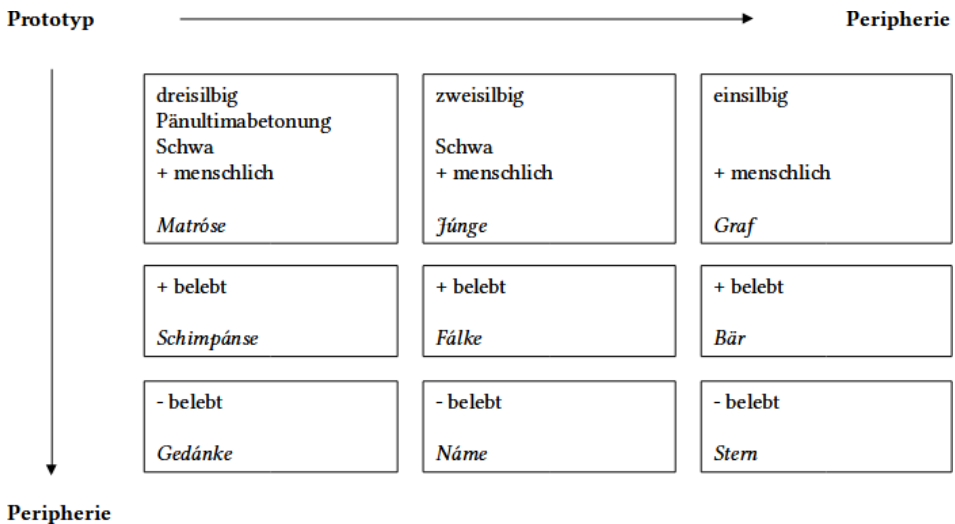


Abbildung 3.7: Abstufungen des Form-Schemas dreisilbig, pänultimabetont, Schwa-Endung, menschlich; angelehnt an Köpcke (1995: 170)

Einen Schritt von der prototypischen phonologischen Form [dreisilbig, pänultimabetont, Schwa-Endung] entfernt sind Substantive, die zweisilbig sind und auf Schwa auslauten: *Júnge*, *Bóte*.⁵⁹ Köpcke (1995: 171) schreibt der zweisilbigen

⁵⁹Da sich die Pänultimabetonung aus der Zweisilbigkeit und der Schwa-Endung ergibt, wird sie nicht als gesonderte Eigenschaft gesehen, auch wenn sie bei Köpcke (2000a: 113) als Eigenschaft aufgeführt wird.

Form nur bedingte Produktivität zu und nennt als Neuzugang lediglich *der Kurde*. Zudem führt die phonologische Form [Zweisilbigkeit und Schwa-Endung] auch in Kombination mit einem hohen Grad an Belebtheit nicht zwangsläufig zur schwachen Deklination: Köpcke (1995: 172) nennt *Piefke*, *Steppke* und *Vize* als potentielle Vertreter des Form-Schemas I, die dennoch starke Flexion aufweisen. Für *Piefke* und *Steppke* nimmt Köpcke (1995: 172) an, dass die pejorative Bedeutung die Deklination beeinflusst.⁶⁰

Dennoch stellt sich generell die Frage, ob es gerechtfertigt ist, für drei- und zweisilbige Substantive einen Unterschied in der Typizität anzunehmen. M. E. ist die Abstufung aus theoretischer Sicht sinnvoll: Da nur wenige Substantive mit den phonologisch-prosodischen Eigenschaften [dreisilbig, pänultimabetont und Endung auf Schwa] auf Entitäten referieren, die nicht menschlich sind, kommt dieser Form eine höhere *cue validity* zu als der Form [zweisilbig und Endung auf Schwa]. Allerdings scheint dies keine praktischen Auswirkungen zu haben. Abbautendenzen der schwachen Flexion zeigen sich sowohl bei den dreisilbigen als auch bei den zweisilbigen Substantiven nur in Kombination mit geringer Belebtheit. Daher sind Schwankungen bei Abstrakta, Konkreta und Substantiven, die auf Vögel, Fische, Reptilien und Insekten referieren, zu beobachten und diachron nachzuvollziehen (Köpcke 2000a: 112–117). Eine Ausnahme von der belebtheitsgesteuerten Klassenzugehörigkeit ist *die Waise*, ein Substantiv, dessen Referent einen hohen Belebtheitsgrad aufweist, das aber dennoch von der schwachen maskulinen (*der Waise*) in die feminine Deklination gewechselt ist (Köpcke 2000a: 116–117). Auch *der Falke*, der sich bislang in der schwachen Flexion halten konnte, erscheint auf den ersten Blick als Ausnahme, da er auf einen Vogel referiert. Falken kam jedoch im Mittelalter ein hoher kultureller Stellenwert und damit ein höherer Belebtheitsgrad als anderen Vögeln zu (Köpcke 2000a: 112–113).

Für zweisilbige Substantive mit geringem Belebtheitsgrad lassen sich zwei Entwicklungslinien festhalten: Wie dreisilbige Substantive mit geringem Belebtheitsgrad nehmen sie im Genitiv *-ns* und in der Folge *-n* an (ausführlich hierzu siehe Abschnitt 3.2.2), zudem können sie das Genus wechseln. *Der Krake* zeigt momentan beide Entwicklungstendenzen: Einerseits schwankt das Substantiv im Genitiv zwischen *-n* (*des Kraken*) und *-ns* (*des Krakens*) und tendiert dazu, im Nominativ *-n* anzunehmen (*der Kraken*) und somit in die starke Deklination überzugehen, andererseits finden sich feminine Formen (*die Krake*). Köpcke & Zubin (2018: 224–227) stellen hierbei Bedeutungs differenzierungen fest: *die Krake* wird

⁶⁰Zusätzlich könnte das Diminutivsuffix *-ke* bei *Piefke* und *Steppke* eine Assoziation mit dem Form-Schema I erschweren. *Vize* ist ein Kurzwort, das sich erst im 19. Jahrhundert aus Komposita wie *Vizekanzler* herausbildet (Pfeifer 2000: 1520). Die späte Entstehung und der Status als Kurzwort könnten die schwache Flexion blockieren.

3 Variation in Konjugation, Deklination und Auxiliarselektion

vorrangig metaphorisch verwendet (*die Datenkrake*), während *der Krake* vorrangig auf das Tier verweist.

Tendenzen zur Annahme von *-n* im Nominativ lassen sich außerdem bei *der Name/n*, *der Wille/n* und *der Buchstabe/n* beobachten (Köpcke 1995: 173). Bei *der Friede/n* sieht Köpcke (1995: 173) den Prozess als abgeschlossen an. Auch diachron zeigen sich Beispiele von *n*-Annahmen im Nominativ: Köpcke (2000a: 113–115, 120) führt 72 Maskulina auf wie bspw. *Rücken*, *Hopfen* und *Weizen*, die ursprünglich auf Schwa auslauteten. Schwankungen in der Nominativform können ebenfalls mit einer Bedeutungs differenzierung einhergehen: So ist *der Drache* ein feuerspeiendes Wesen, während *der Drachen* steigen gelassen werden kann (Köpcke 1995: 173). Genauso werden für *Funke/n* Bedeutungsunterschiede diskutiert: Die Bedeutung ‚geringes Ausmaß‘ (*ein Funken Hoffnung*) ist laut Sandberg (1998: 403) auf die *n*-haltige Nominativform beschränkt. Joeres (1996: 316–317) sieht indes keinen semantischen Unterschied zwischen *n*-haltigen und *n*-losen Formen bei *Funke*. Für *der Friede/n* werden ebenfalls Bedeutungs differenzierungen diskutiert: In der Bedeutung ‚Nicht-Kriegszustand‘ wird laut Duden (2016: 336) die *n*-haltige Nominativform bevorzugt, während *der Friede* in der religiösen Bedeutung ‚Geborgenheit in Gott‘ genutzt wird. Auf eine mögliche Bedeutungs differenzierung geht Joeres (1996: 313–316) nicht ein, er merkt lediglich an, dass viele *n*-lose Belege Redensarten aus der Bibel darstellen. Auch der umgekehrte Weg zur *n*-Annahme im Nominativ – die Tilgung von *-n* – lässt sich diachron feststellen. Diese führte dazu, dass ehemals starke Substantive zu schwachen wurden. Dies ist nur bei Substantiven mit hohem Belebtheitsgrad zu beobachten wie bspw. bei *Schöffe* < mhd. *scheffen* und *Heide* < mhd. *heiden* (Köpcke 2000a: 112).

Der Genuswechsel von schwachen Maskulina zu Feminina lässt sich durch Form-Schemata erklären: Die Form [Zweisilbigkeit und Schwa] deckt sich im Nominativ Singular mit den phonologischen Eigenschaften schwacher Feminina (*die Wölke*). Zudem ist die Pluralmarkierung auf *-(e)n* prototypisch für gemischt-deklinierte Feminina (Köpcke 2000a: 115). Es ist daher ein alternatives Form-Schema verfügbar, an das sich die Substantive anschließen können, deren Referenten einen geringen Belebtheitsgrad aufweisen. Der Einfluss des finalen Schwas auf die Genuswahl wird auch anhand von Minimalpaaren deutlich, die sich allein in der Existenz des Schwas unterscheiden: *die Quelle* vs. *der Quell*, *die Zehe* vs. *der Zeh* (Nübling 2016: 168). Köpcke (1993: 128) sieht die Schwa-Endung daher als Genusindikator. Hinsichtlich der Schwa-Endung wird zudem angenommen, dass sie Belebtheit markieren kann (Köpcke 2005: 75). Inwiefern sich diese Funktionen des Schwas widersprechen und inwiefern sie vereinbar sind (Schwa als Marker für Feminina, Schwa bei Maskulina dagegen als Belebtheitsmarker) wurde in der Forschung meines Wissens nicht diskutiert.

Unabhängig vom Status des Schwa kann festgehalten werden, dass zweisilbige Substantive auf Schwa zur Form zweier Form-Schemata passen. Auch hier ist von zwei Ebenen auszugehen: Zweisilbigkeit und Schwa-Endung können sowohl mit menschlichen als auch mit nicht-menschlichen Referenten assoziiert sein. Auf der übergeordneten Ebene entscheidet daher die Belebtheit darüber, welches Form-Schema greift und welches Genus genutzt wird, siehe Abbildung 3.8.⁶¹

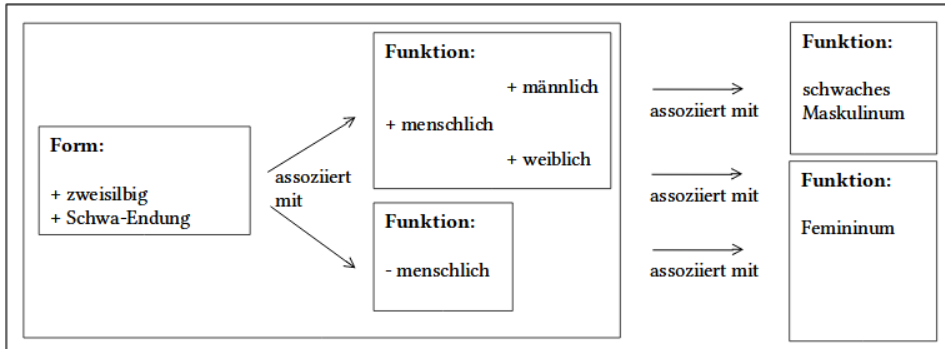


Abbildung 3.8: Verschachtelung der Form-Schemata von schwachen Maskulina und Feminina

Der Genuswechsel von zweisilbigen schwachen Maskulina auf Schwa mit geringer Belebtheit hin zu den Feminina lässt sich als Profilierung der Form-Schemata werten: Im Fall von *Krake* korreliert eine geringe Belebtheit mit der Funktion [+maskulin]. Durch den Genuswechsel zu [+feminin] wird die *cue validity* des Form-Schemas der schwachen Maskulina sowie der gemischt-deklinierten Feminina gestärkt und beide Form-Schemata daher profiliert.

Eine vergleichbare Profilierung lässt sich für den Wechsel zur starken Deklination durch Annahme von *-n* ansetzen: Zwei- und dreisilbige Substantive mit geringem Belebtheitsgrad sind nur wenig mit schwacher Flexion assoziiert. Durch die Annahme von *-n* wird die Assoziation weiter gesenkt, da nun auch die prosodisch-phonologischen Eigenschaften der Form aufgrund des fehlenden Schwa-Auslauts nicht mehr erfüllt werden. Hierdurch sind Substantive wie *der Kraken* weder mit den semantischen noch mit den prosodisch-phonologischen Eigenschaften des Form-Schemas kompatibel und werden daher auch nicht mit schwacher Flexion assoziiert.

⁶¹Innerhalb der Belebtheit beeinflusst Sexus die Assoziation mit schwachen Maskulina bzw. Feminina. So gehören Substantive wie *die Tante* und *die Nichte* aufgrund des Sexus zu den Feminina und nicht zu den schwachen Maskulina.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Welche Faktoren dazu führen, dass zweisilbige Substantive mit geringem Belebtheitsgrad die Deklinationsklasse oder das Genus wechseln, ist meines Wissens bislang nur in Ansätzen erforscht. Einerseits könnte Pluralität eine Rolle spielen: Für Entitäten, auf die häufig im Plural referiert wird, ist eine Pluralopposition relevant. Diese wird durch den Genuswechsel ermöglicht (Nübling 2008: 305). Die Numerusprofilierung kann daher als Erklärung für den Genuswechsel von Schwarmtieren wie Heuschrecken herangezogen werden sowie von paarigen Körperteilen wie *die Niere*, die aufgrund ihrer Paarigkeit häufiger im Plural verwendet werden als im Singular (Nübling 2008: 305).

Andererseits können semantische Rollen für die Entwicklung relevant sein. Wird ein Substantiv häufig als Subjekt in Passivsätzen und der Nominativ somit in der Patiensrolle gebraucht, könnte dies dazu führen, dass die *n*-haltige Form, die auf Kasus in Nicht-Agens-Rollen beschränkt ist, auf den Nominativ ausgeweitet wird.⁶² Aus *der Weize* wird also *der Weizen*, weil *Weizen* nun einmal geerntet wird und nicht selbst erntet. Dieser Erklärungsansatz deckt sich mit der diachronen Entwicklung, in der vornehmlich Konkreta und Abstrakta *n*-haltige Nominative annahmen (Köpcke 2000a: 120). Konkreta und Abstrakta sind aufgrund ihres geringen Belebtheits- sowie Individuiertheitsgrads nur eingeschränkt als Agens nutzbar. Im Einzelfall lässt sich dennoch nicht immer nachvollziehen, warum welcher Weg eingeschlagen wurde. So hat *die Hirse* das Genus gewechselt, während andere Getreidesorten und getreideähnliche Pflanzen wie bspw. *Roggen*, *Weizen* und *Hopfen* *n*-haltig geworden sind (Köpcke 2000a: 120–121).

Variation zwischen *n*-haltigen und endungslosen Dativ- und Akkusativformen (*dem/den Krake/n*) oder zwischen Genitivformen auf *-n* und *-s* (*des Kraken/s*) lässt sich für drei- und zweisilbige Substantive auf Schwa nicht feststellen. Die *n*-haltigen Nominativformen und die Feminina stechen diese Entwicklungsmöglichkeit statistisch aus, da mehr Substantive mit ähnlichen phonologischen Eigenschaften und ähnlicher Semantik existieren, die entweder *-n* im Nominativ enthalten oder Feminina sind. Dies zeigt sich bspw. bei den Genitivformen: Es existieren viele zweisilbige Substantive auf Schwa mit Genitiv auf *-ns* bspw. aufgrund von substantivierten Verben (*des Tragens*, *des Lachens*), aber nur wenige mit Genitiv auf *-s* (*des Käses*) (Köpcke 2005: 71). Formen auf *-ns* sind somit bei zwei- und dreisilbigen Maskulina typenfrequenter und daher konventioneller.

Monosyllabische Substantive⁶³ sind aufgrund ihrer prosodisch-phonologischen Eigenschaften nicht mit schwacher Flexion assoziiert, da viele Einsilber stark

⁶²Für diesen Hinweis danke ich Melitta Gillmann.

⁶³Krischke (2012: 60–61) setzt vor den monosyllabischen Substantiven eine weitere Abstufung des Form-Schemas an: zweisilbige Substantive mit Pänultimabetonung, aber ohne Schwa. Als Beispiele nennt er *Bauer*, *Nachfahr* und *Vorfahr*. Diese Substantive schwanken in ihrer Ver-

flektieren wie bspw. *der Stuhl* und *der Stein*. Allein ein hoher Belebtheitsgrad kann hier die schwache Flexion bewahren. Aber auch Substantive mit hohem Belebtheitsgrad (*Prinz, Graf*) zeigen Schwankungen hin zur starken Flexion: Im Dativ und Akkusativ finden sich endungslose Formen (*den/dem Graf*∅), im Genitiv sind Formen auf *-s* (*des Graf*s) sowie vereinzelt Formen auf *-(e)ns* (*des Graf*ens) zu beobachten.⁶⁴ Bei monosyllabischen Substantiven, die auf Unbelebtes referieren (bspw. *Stern, März*), ist der Prozess abgeschlossen. Die schwachen Formen sind nur noch in Komposita und Redewendungen fossilisiert (*Sternenhimmel, Im Märzen der Bauer*).

Köpcke (2000a: 114–115) beschreibt eine belebtheitsgesteuerte Entwicklung der ursprünglich schwach deklinierten monosyllabischen Substantive: Während 27 % der Substantive mit menschlichen Referenten die schwache Flexion beibehielten (*der Herr, der Fürst*), sind es bei tierischen Referenten nur rund 13 bis 17 % (*der Fink*); Entitäten mit geringerem Belebtheitsgrad sind komplett in die starke Flexion gewechselt (*der Mai*). Köpcke (2000a: 114–115) betrachtet nur Substantive, die das stammfinale Schwa vom Mhd. zum Frnhd. apokopierten und hierdurch in die Peripherie der schwachen Substantive gelangt sind.⁶⁵

Auch für das Neuhochdeutsche kann der Einfluss der prosodisch-phonologischen und semantischen Eigenschaften des Form-Schemas I auf die Flexion nachgewiesen werden. In einem Produktionsexperiment zeigt Köpcke (2005: 77–78), dass Schüler_innen der Jahrgangsstufe 6 vor allem bei monosyllabischen

wendung: *Bauer* schwankt zur starken Flexion (Duden 2016: 132–133), bei *Nach-* und *Vorfahr* finden sich wie oben diskutiert schwa-haltige Nominativformen (*der Nachfahre*) (Köpcke 1993: 137–138). Zwar ist *Bauer* nicht schwa-haltig, es enthält aber mit dem vokalisiertem R einen Reduktionsvokal. Bei *nach* und *vor* in *Nach-* bzw. *Vorfahr* handelt es sich um Präfixe. Anders als *Junge* und *Bauer* sind *Vor-* und *Nachfah* somit komplexe morphologische Wörter. Der Stamm (*-fah*) ist nicht zweisilbig, sondern lediglich das Wortbildungsprodukt.

⁶⁴Krischke (2012: 76–77) beobachtet, dass Genitivformen auf *-(e)ns* bei monosyllabischen Substantiven bereits im 13. Jahrhundert unabhängig vom Belebtheitsgrad der Referenten auftauchen: *des Affens, Botens, Prinzens*. Allerdings scheinen sich diese Formen bei keinem der von Krischke (2012: 76–77) genannten Beispiele durchgesetzt zu haben.

⁶⁵Köpcke (2000a: 120–121) beruft sich in seiner Auflistung auf die Grammatik von Paul (1917). Köpckes Daten sind jedoch mit Vorsicht zu genießen, denn in Köpckes Liste sind auch Substantive wie *Apt* und *Dieb*, die im Mittelhochdeutschen Handwörterbuch von Lexer (1872–1878) lediglich als starke Maskulina geführt werden. Die aufgeführten Formen (u.a. *abbet, abt, diup, diep*) sind zudem allesamt nicht schwa-haltig. Mhd. *karl(e)* (nhd. *Kerl*) wird als schwaches sowie starkes Maskulinum ausgewiesen. Außerdem sind in der Grammatik von Paul (1917) nicht alle von Köpcke (2000a) untersuchten Substantive als schwach flektierend aufgeführt. Einige sind als Substantive klassifiziert „von denen sw. Formen vorkommen“ (Paul 1917: 46). Dazu zählen auch *Apt* und *Dieb*. Die im Fließtext angeführten Beispiele wurden dahingehend überprüft, ob sie im Mittelhochdeutschen Handwörterbuch als auf Schwa auslautendes schwaches Maskulinum geführt werden.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Substantiven die starke Deklination bevorzugen. Bei zweisilbigen Substantiven auf Schwa überwiegt hingegen die schwache Flexion, das dreisilbige Substantiv wird exklusiv schwach dekliniert. Dabei spielt auch Belebtheit eine Rolle, denn *der Bote* wird zu 100 % schwach flektiert, *der Falke* nur zu 75 %. Ähnliche Belebtheitseffekte lassen sich für die monosyllabischen Substantive feststellen (Köpcke 2005: 77). Insgesamt ist die schwache Flexion bei den zweisilbigen Formen aber unabhängig von der Belebtheit stabiler als bei den monosyllabischen Substantiven. Dies bestätigt die Hypothese von Schäfer (2019: 394–396), nach der prosodisch-phonologische Eigenschaften die Flexion besser stützen als die semantische. Auch Schäfer (2019) kann im Rahmen seiner Korpusuntersuchung mithilfe einer logistischen Regression das Wirken des von Köpcke (1995) postulierten prosodisch-phonologisch und semantisch motivierten Form-Schemas bestätigen.

Weitere Hinweise auf das Wirken des Form-Schemas I liefert ein Produktionsexperiment von Köpcke (2000b). 31 Proband_innen⁶⁶ wurden hierfür Pseudosubstantive präsentiert, zu denen ihnen nur das Genus und die Bedeutung genannt wurden. Die Proband_innen wurden gebeten, die Testsubstantive in den Genitiv Singular sowie Nominativ Plural zu setzen (Köpcke 2000b: 159). Die Substantive sind dreisilbig mit Pänultimabetonung und Schwa (z. B. *Schéttose*), zweisilbig mit Schwa (z. B. *Zirfe*) oder monosyllabisch (z. B. *Knatt*), sodass sie in ihren prosodisch-phonologischen Eigenschaften gestaffelt und damit zu unterschiedlichen Graden mit schwachen Maskulina assoziiert sind. Die ein- und zweisilbigen Testitems sind zusätzlich mit unterschiedlichen Belebtheitsgraden versehen: *Der Zirfe* wird bspw. als hoher türkischer Würdenträger eingeführt, während *der Stisse* als ein dem Zebra ähnliches Tier definiert wird (Köpcke 2000b: 168). Zusätzlich nutzt Köpcke (2000b) Substantive auf *-or* sowie Substantive auf *-el* bzw. *-er* als Testwörter.

Mit diesem Design kann Köpcke (2000b: 159–166) den Einfluss des Form-Schemas I auf die Deklination von Maskulina beobachten. Die Häufigkeit der schwachen Flexion nimmt von der dreisilbigen hin zur einsilbigen Form ab: Während die dreisilbigen Substantive zu 76 % *-(e)n* im Genitiv aufweisen, sind es bei den monosyllabischen 9 % (unbelebt) bis 39 % (menschlich).⁶⁷ Die Abstufungen innerhalb der einsilbigen Testsubstantive verdeutlichen den Einfluss der Belebtheit. Dieser zeigt sich auch bei zweisilbigen Substantiven: Korreliert Zweisilbigkeit

⁶⁶Bei allen Proband_innen handelte es sich um Germanistikstudent_innen am Anfang des Studiums (Köpcke 2000b: 159).

⁶⁷Ronneberger-Sibold (2020: 416–418) kann in einer Replikationsstudie mit Studierenden aus Bayern, deren Sprachverhalten stark dialektal geprägt ist, jedoch keine Präferenz für *-s* oder *-n* bei *Schettose* ablesen. Sie erklärt diesen Befund damit, dass der Genitiv in Dialekten gemieden wird und sich somit das Form-Schema schwacher Maskulina nicht festigen kann.

mit einem hohen Grad an Belebtheit, nimmt die Tendenz zur schwachen Flexion im Vergleich zur dreisilbigen Form nur leicht ab (61 %). Wird dagegen auf tierische Entitäten referiert, fällt die Anzahl schwach gebildeter Genitive auf 37 % und gleicht damit dem Anteil schwacher Maskulina bei einsilbigen Substantiven mit hohem Belebtheitsgrad.

Die Substantive auf *-or* weisen mit 27 % einen ähnlichen Anteil an schwachen Genitivformen auf wie die Einsilbler mit tierischen Referenten. Aber auch hier ist der Einfluss der Belebtheit zu erkennen: Während das Pseudosubstantiv *Frátor*, das als [+menschlich] eingeführt wurde, zu 32 % im Genitiv schwach flektiert wird, wird *Tréitor* mit nicht-belebter Referenz nur zu 22 % schwach flektiert. Wie bei dem real existierenden Substantiv *Autor* zeigt sich hier die Anziehungskraft des Form-Schemas I.

Köpcke (2000b: 159–161) stellt die Hypothese auf, dass der Anteil der schwachen Deklination bei Substantiven auf *-el* und *-er* am geringsten ist, da Substantive auf *-el* und *-er* beinahe ausnahmslos stark deklinieren.⁶⁸ Die Hypothese bestätigt sich: Die Substantive auf *-el* oder *-er* weisen mit 87 % vornehmlich starke Genitivformen auf (Köpcke 2000b: 161). Substantive auf *-er* und *-el* lassen sich daher als Form-Schemata der starken Deklination betrachten.⁶⁹

Schmitt (2019a) nutzt die Pseudosubstantive aus Köpcke (2000b) für eine self-paced-reading-Studie. In die Studie sind drei Pseudosubstantive eingeflochten: ein dreisilbiges (*Schettose*), ein monosyllabisches (*Knatt*) und ein Substantiv auf *-el* (*Grettel*). Die Substantive werden als [+menschlich] eingeführt und jeweils mit *-s* und *-n* im Genitiv präsentiert (*des Schettoses/Schettosen*). Die Reihenfolge der Endungen ist dabei pseudorandomisiert. Zudem wird den Proband_innen am Ende der Studie ein Produktionsexperiment vorgelegt, in dem sie gebeten werden, den Genitiv der Testsubstantive zu bilden. Das Produktionsexperiment bestätigt Köpckes Ergebnisse: Das Substantiv auf *-el* wird vorrangig stark flektiert (*des Gretteln*), das dreisilbige Substantiv hingegen schwach (*des Schettosen*), das monosyllabische schwankt in seiner Verwendung (*des Knatts/Knatten*) (Schmitt 2019a: 167–169).

Der Einfluss des Form-Schemas I zeigt sich jedoch nicht in den Lesezeiten: Nur für das Substantiv auf *-el* sind die Lesezeiten für *-s* im Vergleich zu *-n* verringert, die anderen Testsubstantive werden unabhängig von der präsentierten

⁶⁸Köpcke (2000b: 160–161) schätzt, dass 3 % der Substantive auf *-el* und *-er* schwach deklinieren. Bei dieser Schätzung sind durchsichtige Derivationen auf *-er* (*fischen* > *der Fischer*) allerdings ausgeschlossen, sodass der tatsächliche Anteil sogar geringer ist.

⁶⁹Das Form-Schema der starken Maskulina ist vergleichbar mit dem Form-Schema regelmäßiger Verben im Englischen: Englische Verben auf /f/ flektieren regelmäßig (*to laugh, laughed*) (Goldberg 2019: 135). Auch innerhalb eines hochvariablen Clusters können Subgruppen festgestellt werden und somit Form-Schemata existieren.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Endung gleich schnell gelesen (Schmitt 2019a: 169–171). Dieses Ergebnis könnte durch die Pseudosubstantive bedingt sein: Zwar haben die Sprecher_innen aufgrund des Form-Schemas eine Intuition darüber, wie sie sie flektieren würden, aber da sie die Substantive nicht kennen, kann das statistische Vorkaufsrecht nur bedingt greifen, sodass beide Formen möglich erscheinen. Zudem ist es möglich, dass die Methode zu wenig sensitiv ist, um Prozessierungsunterschiede zu messen. Im empirischen Teil der Arbeit wird der Einfluss des Form-Schemas I auf die Variation in der Deklination von Maskulina überprüft. Hierfür werden Maskulina mit starken und schwachen Genitivformen präsentiert, die entweder zur starken Flexion gehören, dem Form-Schema peripher angehören oder ihm komplett entsprechen.

Insgesamt zeigt sich, dass die prosodisch-phonologischen und semantischen Eigenschaften des Form-Schemas I einen umfassenden Einfluss auf die Flexion von Maskulina nehmen. In Hinblick auf Variation sind dabei jeweils die Kombinationen aus Prototyp und Peripherie interessant: Die Kombination aus prosodisch-phonologischem Prototyp und semantischer Peripherie (*der Gedanke, der Wille*) macht einen Wechsel der Deklinationsklasse durch Annahme von *-ns* im Genitiv und im Anschluss daran von *-n* im Nominativ wahrscheinlich. Zweisilbigen Substantiven mit niedrigem Belebtheitsgrad steht zudem offen, zu den Feminina zu wechseln (*der Krake > die Krake*). Im Gegensatz dazu wechselt die Kombination aus prosodisch-phonologischer Peripherie und semantischem Prototyp (*der Bär, der Graf*) die Deklinationsklasse durch Aufgabe der Dativ- und Akkusativendung sowie durch Wechsel von Genitiv *-n* zu *-s*.

3.2.3.2 Form-Schema II

Das Form-Schema II [Dreisilbigkeit, Ultimabetonung und + menschlich] ist weit weniger gut erforscht als das schwa-haltige Pendant.⁷⁰ Zu dem Form-Schema zählen ca. 900 Substantive (Bittner 2003 [1991]: 98–99), es hat also mehr Types als das Form-Schema I.

Wie beim Form-Schema I lassen sich auch für das Form-Schema II prosodisch-phonologische sowie semantische Abstufungen ausmachen (Krischke 2012: 61–62): Dreisilbige Substantive sind prototypische Vertreter des Schemas (*der Polizist*), dann folgen zweisilbige (*der Pirat*). Die weiteren Abstufungen ergeben sich durch Belebtheit (siehe Abbildung 3.9).

⁷⁰Köpcke (1995: 176) sieht trotz dieser schwa-losen Variante des Form-Schemas das Schwa als wichtigen Teil der phonologischen Eigenschaften schwacher Maskulina an. Dies leitet er aus der Produktivität des Form-Schemas I bei Einwohnerbezeichnungen sowie aus der Hinzufügung von Schwa bei Entlehnungen ab (siehe Abschnitt 3.2.3.1). Zudem geht Köpcke (2005: 75) davon aus, dass die semantische Eigenschaft [+menschlich] bei *Matrose* und *Junge* ein phonologisches Korrelat durch die Endung auf Schwa findet. Daraus zieht er den Schluss, dass Schwa als ein Belebtheitsmarker fungieren könnte.

3.2 Variation in der Deklinationsklasse

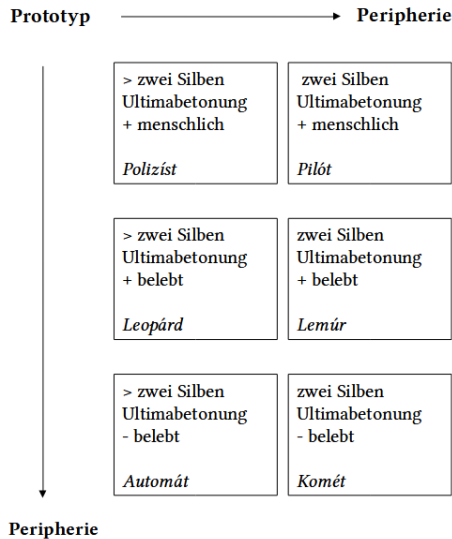


Abbildung 3.9: Abstufungen des Form-Schemas dreisilbig, ultimabetont, menschlich

Köpcke (1995: 175) nutzt ähnliche Abstufungen wie Krischke (2012: 61–62), verzichtet aber auf phonologische Abstufungen. Keiner der beiden begründet die Abstufungen, sodass der Eindruck entsteht, dass diese vorrangig in Anlehnung an das Form-Schema I angenommen werden. Zudem überprüfen weder Köpcke (1995) noch Krischke (2012), ob in der phonologischen und semantischen Peripherie des Form-Schemas mehr Schwankungen stattfinden als beim Prototyp. Thieroff (2003: 113) merkt an, dass Substantive, die dem Form-Schema II völlig entsprechen, dennoch starke Dativ-, Akkusativ- und Genitivformen aufweisen können. Mit konkreten Zahlen belegt er seine Ausführungen allerdings nicht. Auch Bittner (2003 [1991]: 105) beobachtet für einige Substantive mit Endbetonung Schwankungen zur gemischten Flexion (*des Partisans* statt *des Partisanen*).

Es ist festzuhalten, dass das Form-Schema II im Gegensatz zum Form-Schema I sehr heterogen ist (Poitou 2004a: 78–79) und viele nicht-native Suffixe enthält (Bittner 2003 [1991]: 99). Auf Basis der möglichen Suffixe des Form-Schemas II erstellt Poitou (2004a: 78–79) Unterklassen, für die er jedoch zu dem Ergebnis kommt, dass die Flexion der Substantive weitgehend belebtheitsgesteuert ist. Damit bestätigt er Köpcke (1995: 165–176), der zählt, wie viele Substantiven auf *-ist*, *-at*, *-(i)ent*, *-ant*, *-it* und *-graph* schwach bzw. stark flektieren. Köpcke (1995: 175) stellt fest, dass Substantive mit menschlichen Referenten unabhängig vom Suffix stets schwach flektieren. Bei Substantiven mit nicht-belebten Referenten las-

3 Variation in Konjugation, Deklination und Auxiliarselektion

sen sich hingegen beide Deklinationen beobachten. Eine Ausnahme hiervon bilden die Substantive auf *-graph*, die unabhängig vom Belebtheitsgrad stets schwach flektiert werden. Poitou (2004a: 78–79) beobachtet zudem für Substantive auf *-end* aus der Mathematik (*Dividend*) eine stabile schwache Flexion. Zumindest für den Einfluss der Belebtheit und den Einfluss einzelner Suffixe auf schwache Flexion scheint es somit empirische Evidenz zu geben. Insgesamt wäre jedoch weitere Forschung nötig, um das Form-Schema II und dessen Verhältnis zum Form-Schema I genauer beschreiben zu können. So ließe sich das Form-Schema II auch als weitere Abstufung innerhalb des Form-Schemas I modellieren. Zudem ist davon auszugehen, dass das Form-Schema II aufgrund seiner höheren Typenfrequenz auch eine größere Variabilität aufweist als das Form-Schema I. Die Variabilität zeigt sich auch in den nicht-nativen Suffixen, die für das Form-Schema II zu beobachten sind.

Insgesamt zeigt sich, dass schwache Maskulina durch Form-Schemata konditioniert sind, die prosodisch-phonologische sowie semantische Eigenschaften aufweisen. Form-Schema I beinhaltet dreisilbigen Maskulina mit Pänultimabetonung auf Schwa und menschlichen Referenten. Ein zweites Form-Schema lässt sich mit dreisilbigen ultimabetonten Maskulina ohne Schwa-Endung und menschlichen Referenten ansetzen, dieses ist jedoch weit weniger gut untersucht als das Form-Schema I.

Prototypisch prosodisch-phonologische und semantische Eigenschaften führen zu stabiler schwacher Flexion, während periphere prosodisch-phonologische und semantische Eigenschaften zu stabiler starker Flexion führen. Schwankungen treten in Kombinationen aus Peripherie und Prototyp auf: Monosyllabische Substantive mit hohem Belebtheitsgrad schwanken von *n*-haltigen zu endungslosen Dativ- und Akkusativformen (*dem/n Grafen* > *dem/den Graf*) und von Genitivendungen auf *-(e)n* zu *-(e)s* (*des Grafen* > *des Grafes*). Mehrsilbige Substantive mit geringem Belebtheitsgrad schwanken zu Genitiven auf *-ns* (*des Willen* > *des Willens*) und *n*-haltigen Nominativformen (*der Glaube* > *der Glauben*). Zweisilbige Substantive können außerdem das Genus wechseln (*der Krake* > *die Krake*). Der Einfluss des Form-Schemas I auf die Variation in der Deklination wird im empirischen Teil der Arbeit überprüft, indem starke und schwache Genitivformen von Maskulina verglichen werden, die dem Form-Schema komplett, peripher oder gar nicht entsprechen.

Der Einfluss der Belebtheit auf die Flexion lässt sich damit erklären, dass Referenten mit hohem Belebtheitsgrad potentiell die Agensrolle erfüllen, die häufig im Nominativ ausgedrückt wird. Der Nominativ muss in dieser Sichtweise nicht gesondert markiert werden. Dieser Forderung kommen schwache Maskulina nach, da sie im Singular allein den Nominativ unmarkiert lassen. Hierbei

ist zunächst die phonologische Form [Dreisilbigkeit, Pänultimabetonung, Schwa-Endung] mit [+menschlich] assoziiert und diese Assoziation selbst ist wiederum mit schwacher Flexion verknüpft. Daher wirken prosodisch-phonologische und semantische Faktoren auf die Assoziation mit schwacher Flexion. Die Schwankungsfälle, die derzeit zu beobachten sind, führen dazu, dass einerseits die Assoziation zwischen den prosodisch-phonologischen und den semantischen Eigenschaften weiter gestärkt wird und fördern andererseits die Assoziation dieser Eigenschaften mit schwacher Flexion. Somit werden die schwachen Maskulina als prosodisch-phonologisch und semantisch konditionierte Deklinationsklasse geschärft. Im folgenden Abschnitt werden die Einflussfaktoren zusammengefasst und ihr Zusammenwirken diskutiert.

3.2.4 Zusammenfassung und Zusammenwirken der Faktoren

Frequenz, Prototypizität und Form-Schematizität nehmen Einfluss auf die Variation in der Deklination der Maskulina. Dabei ist der Einfluss der Frequenz als grundlegend zu betrachten: Aufgrund der geringen Typenfrequenz von schwachen Maskulina sind Schwankungen in Kombination mit geringer Tokenfrequenz zur starken Flexion zu erwarten und auch zu beobachten (Schäfer 2019: 404–405), da das statistische Vorkaufsrecht für schwache Formen aufgrund der geringen Tokenfrequenz nur noch eingeschränkt greift. Die Entwicklungsrichtung lässt sich auch dadurch begründen, dass komplex markierte Deklinationsklassen wie die schwache zur Schließung neigen (Nübling 2016: 173). Neuzugänge sind aber nicht ausgeschlossen, wie Substantive auf *-or* zeigen (*des Autors* > *des Autoren*). Bezüglich der Typenfrequenz von Maskulina besteht insgesamt noch Forschungsbedarf, so fehlen bspw. genaue Angaben über die Typenfrequenz der einzelnen Deklinationsklassen.

Die Typenfrequenz steuert die Variation grundlegend, indem sie die Schwankungsrichtung vorgibt. Darauf aufbauend beeinflussen neben der Tokenfrequenz vorrangig Form-Schemata die Deklination. Schwache Maskulina sind prosodisch-phonologisch und semantisch motiviert: Dabei stellen dreisilbige, pänultimabetonte Substantive auf Schwa mit Referenz auf Entitäten mit hohem Belebtheitsgrad den Prototyp des Form-Schemas I dar. Das Form-Schema I ergibt sich durch eine doppelte Assoziation: Die prosodisch-phonologischen Eigenschaften sind mit einem hohen Belebtheitsgrad assoziiert und diese Kombination mit schwacher Flexion. Schwankungen zwischen starker und schwacher Flexion sind jeweils für die Kombination aus Peripherie und Prototyp des Form-Schemas zu beobachten. Neben dem Form-Schema I ist ein zweites anzusetzen [Dreisilbigkeit, Ultimabetonung, + menschlich], das weniger gut erforscht ist.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Der Wandel von schwacher zu starker Flexion kann mithilfe von prototypischen Flexionseigenschaften starker und schwacher Maskulina systematisiert und prognostiziert werden. Schwache Maskulina profilieren die Dichotomie zwischen Nominativ Singular und anderen Kasus, während starke Maskulina vorrangig Singular und Plural unterscheiden und eine klare Opposition zwischen Genitiv Singular bzw. Dativ Plural und anderen Kasus aufweisen. Die Flexionseigenschaften von Substantiven in der gemischten Klasse lassen sich zwischen diesen Polen einordnen. Die Nicht-Profilierung der Genitivendung lässt sich in dieser Betrachtung als periphere Eigenschaft der schwachen Flexion betrachten, die Pluralendungen auf *-en* als prototypische Eigenschaft.

Schwankungen zur starken Flexion beginnen mit der Profilierung des Genitivs gegenüber den anderen obliquen Kasus und werden dann auf die Profilierung von Singular und Plural ausgedehnt. Hierbei sind zwei Entwicklungswege⁷¹ festzustellen, die sich durch die unterschiedlichen Kombinationen aus Prototyp und Peripherie des Form-Schemas I ergeben:

- (1) phonologischer Prototyp + semantische Peripherie
schwach > Genitiv auf *-ns* > *n*-haltige Nominativform
der Gedanke > *des Gedankens* > *der Gedanken*
- (2) phonologische Peripherie + semantischer Prototyp
schwach > Dativ/Akkusativ auf \emptyset > Genitiv auf *-s*
der Bär > *dem/den Bär* \emptyset > *des Bärs*

Zweisilbigen Substantiven können neben der Annahme von *-n* im Nominativ auch das Genus wechseln wie bspw. *der Krake* > *die Krake* (Köpcke 2000a: 121).

Interessant ist ein Blick auf mögliche Interaktionen der Einflussfaktoren besonders in Hinblick auf Tokenfrequenz und Form-Schematizität: Hinsichtlich der Tokenfrequenz ist zu erwarten, dass das frequente Lexem *Mensch* stabil schwach flektiert. Da *Mensch* nur zur semantischen Eigenschaft des Form-Schemas I passt, aber nicht zu den prosodisch-phonologischen, sind aus Sicht der Form-Schematizität Schwankungen zu erwarten. Somit wäre es nicht verwunderlich, wenn das Substantiv trotz der hohen Tokenfrequenz schwankt. Jedoch ist hier aufgrund des *entrenchments* tokenfrequenter Formen weniger Variation zu erwarten als bei einem wenig frequenten Substantiv wie bspw. *Prinz*, das ebenfalls nur hinsichtlich der Semantik mit dem Form-Schema I kompatibel ist. Zudem ist es wahrscheinlich, dass für *Mensch* aufgrund der mangelnden Form-Schematizität

⁷¹Da die Entwicklung im Plural für den zweiten Entwicklungsweg bislang kaum untersucht wurde (siehe Abschnitt 3.2.2), wird hier nur die Entwicklung im Singular beschrieben.

3.2 Variation in der Deklinationsklasse

vereinzelte Genitivformen auf *-ns* (*des Menschens*) sowie vereinzelte endungslose Dativ- und Akkusativendungen zu beobachten sind,⁷² jedoch keine Genitive auf *-s* (*des *Menschs*), da die schwache Flexion aufgrund der Tokenfrequenz recht stabil ist und Kombinationen aus /ʃ/ und /s/ phonotaktisch dispräferiert sind. Auch *n*-haltige Nominativformen sind trotz möglicher Genitivformen auf *-ns* unwahrscheinlich, da *Mensch* monosyllabisch ist.

Im empirischen Teil der Arbeit wird der Einfluss der Form-Schematizität auf die Variation in der Deklination überprüft. Hierbei wird auf die prosodisch-phonologischen Eigenschaften des Form-Schemas I fokussiert; die Belebtheit der Testsubstantive wird konstant (menschlich) gehalten. Der Einfluss der prosodisch-phonologischen Eigenschaften des Form-Schemas I wird mithilfe von drei Studien getestet. In einer *self-paced reading task* werden Pseudosubstantive genutzt (zu *self-paced reading tasks* siehe Abschnitt 4.1.2), deren prosodisch-phonologische Form unterschiedlich stark mit schwacher Deklination assoziiert ist: Hierfür werden die Lesezeiten von schwachen und starken Formen von einem dreisilbigen Substantiv mit Pänultimabetonung und Schwa (*des Schettosen/*Schettoses*), einem einsilbigen Substantiv (*des Knatten/Knatts*) sowie einem Substantiv auf *-el* (*des *Gretteln/Grettels*) miteinander verglichen. Anders als in der Studie von Schmitt (2019a) wird die Aufmerksamkeit der Proband_innen auf Wortendungen gelenkt, indem sie nach dem Aufbau der gelesenen Wörter gefragt werden (Haben Sie *des Schettosen* oder *des Schettoses* gelesen?). Dabei wird erwartet, dass die schwachen Formen des dreisilbigen Substantivs niedrigere Lesezeiten hervorrufen als die starken. Das Gegenteil wird für das Substantiv auf *-el* erwartet, da Substantive auf *-el* dem Form-Schema starker Maskulina angehören. Für das einsilbige Substantiv werden vergleichbare Lesezeiten für beide Formen angenommen, da hier Peripherie und Prototyp des Form-Schemas I aufeinandertreffen. Nach der *self-paced reading task* wird den Proband_innen ein Produktionsexperiment vorgelegt, in dem sie die Testsubstantive flektieren. So können die Lesezeiten mit dem Verhalten im Produktionsexperiment kontrastiert werden.

In einer *lexical decision task* sowie einer *sentence maze task* wird mit real existierenden zwei- oder dreisilbigen schwachen Maskulina auf Schwa (z. B. *Kollege*) und einsilbigen schwachen Maskulina (z. B. *Graf*) gearbeitet (zu *lexical decision* und *sentence maze tasks* siehe Abschnitt 4.1.1 und 4.1.3). Zusätzlich dazu werden einsilbige Maskulina (z. B. *Vogt*) genutzt, die stark flektieren. Die Substantive

⁷²Bei determiniererlosen Nominalphrasen wie sie bspw. in Zeitungsüberschriften genutzt werden, ist sogar vermehrt mit endungslosen Dativ- und Akkusativformen zu rechnen, da auf diese Weise trotz fehlendem Determinierer der Singular vom Plural unterschieden werden kann (Thieroff 2003: 116–117).

werden jeweils sowohl mit schwacher als auch mit starker Genitivendung präsentiert (*des Kollegen/*Kolleges*⁷³, *des *Vogten/Vogts*). Es wird davon ausgegangen, dass die starken und schwachen Formen von mehrsilbigen schwachen und einsilbigen starken Substantiven schnell bewertet werden. Zudem wird erwartet, dass die schwachen Formen von mehrsilbigen Substantiven bekannt sind und bevorzugt gewählt werden, die starken hingegen nicht. Das Gegenteil wird für die starken einsilbigen Substantive angenommen. Hinsichtlich der monosyllabischen schwachen Substantive wird erwartet, dass diese insgesamt höhere Reaktionszeiten und ein inkonsistentes Antwortverhalten hervorrufen, da sie der Peripherie des Form-Schemas I angehören. Das methodische Vorgehen und die den Studien zugrunde liegenden Hypothesen⁷⁴ werden ausführlich in den Abschnitten 4.4 und 4.5 erläutert, die Ergebnisse werden in Abschnitt 5.4 vorgestellt. Im folgenden Abschnitt wird die Variation in der Selektion von *haben* und *sein* diskutiert.

3.3 Variation in der Selektion von *haben* und *sein*

Die Selektion der Auxiliare *haben* und *sein* zur Bildung von Perfekt- und Plusquamperfektformen (*ich bin/war gegangen*; *ich habe/hatte gelacht*) unterliegt komplexen Einflussfaktoren: Prototypisch transitive Sätze selektieren *haben*, wohingegen prototypisch intransitive Sätze, die telisch sind oder Bewegungssemantik aufweisen, *sein* selektieren (Shannon 1992, 1995, Gillmann 2016: 316–319). Es handelt sich somit um zwei hochschematische Konstruktionen,⁷⁵ die hinsichtlich ihrer Funktionen in prototypischer Beziehung zueinander stehen.⁷⁶ Im Übergangsbereich zwischen intransitiv und transitiv, telisch und atelisch sowie zwischen Sätzen mit Bewegungssemantik und ohne Bewegungssemantik lassen sich Schwankungsfälle zwischen *sein* und *haben* beobachten. Neben der prototypisch organisierten Beziehung zwischen den beiden Funktionen nimmt Frequenz Einfluss auf die Selektion der Auxiliare (Gillmann 2016: 253–255; 265–268). Zunächst

⁷³Da der Genitiv mit Schwa präsentiert wird (*Kolleges*) ist eine Beeinflussung durch den Genitiv von *Kolleg* (*des Kollegs*) unwahrscheinlich.

⁷⁴Die der lexical-decision- und sentence-maze-Studie zugrunde liegenden Hypothesen wurden anhand der Prästudien angepasst, die angepassten Hypothesen werden in Abschnitt 4.5.1 erläutert.

⁷⁵Diese lassen sich wiederum als Teil einer hochschematischen Makro-Konstruktion sehen, die die Form $X_{NP} X_{VAFIN} ge-V-t/en$ und die Funktion [+Perfekt] bzw. [+Plusquamperfekt] aufweist (Gillmann 2018: 240).

⁷⁶Weil zwei Schemata in prototypischer Beziehung zueinander stehen und nicht mehrere Formen, die mit derselben Funktion assoziiert sind wie bei Flexionsklassen, wird die hierdurch bedingte Variation unter Prototypizität verhandelt und nicht unter Form-Schematizität (zur Unterscheidung von Schematizität und Form-Schematizität siehe Abschnitt 2.3.2).

wird in Abschnitt 3.3.1 der Einfluss der Frequenz auf die Auxiliarselektion betrachtet. Im Anschluss daran wird in Abschnitt 3.3.2 die Prototypizität als Einflussfaktor diskutiert, dabei wird auf Transitivität, Telizität und Bewegungssemantik eingegangen. Im empirischen Teil der Arbeit wird dann der Einfluss der Prototypizität auf die Variation in der Selektion von *haben* und *sein* psycholinguistisch überprüft. Dafür werden Sätze mit Bewegungsverben in unterschiedlichen Transitivitätsgraden verglichen (*nach Rom gefahren sein* vs. *die Kinder zur Schule gefahren haben*).

3.3.1 Frequenz

Ähnlich wie bei den morphologischen Schwankungsfällen, die in den Abschnitten 3.1 und 3.2 diskutiert wurden, zeigen sich auch bei der Selektion von *haben* und *sein* große Unterschiede in der Typenfrequenz. Die *haben*-Selektion gilt als Normalfall (Duden 2016: § 659), da die meisten Verben ihre Perfekt- und Plusquamperfektformen mit *haben* bilden. Die typeninfrequente *sein*-Selektion ist wie Form-Schemata innerhalb einer typeninrequenten Flexionsklasse klar definiert: Nur intransitive Sätze, die telisch sind und/oder Bewegungssemantik aufweisen, selektieren *sein*. Im Unterschied zu den morphologischen Schwankungsfällen sind bei der Selektion von *haben* und *sein* jedoch zwei Schemata anzusetzen, weil zwei Funktionen ([+Transitivität, –Telizität und/oder –Bewegungssemantik] vs. [–Transitivität, +Telizität oder +Bewegungssemantik]) vorliegen. Daher führen die Unterschiede in der Typenfrequenz auch nicht zum Abbau der typeninrequenten Selektion bei geringer Tokenfrequenz: Da die Formen nicht in Konkurrenz zueinander stehen, greift das Zusammenspiel aus Typen- und Tokenfrequenz hier nicht. Dennoch könnte die Typenfrequenz in ambigen Fällen Einfluss auf die Selektion von *haben* und *sein* nehmen: Da das *haben*-Schema typenfrequenter ist, könnte *haben* bei ambigen Sätzen als Default bevorzugt werden.

Der Einfluss der Tokenfrequenz auf die Selektion von *haben* und *sein* wird von Gillmann (2016) untersucht. Hierbei überprüft sie anhand des Archivs W im Deutschen Referenzkorpus,⁷⁷ wie Tokenfrequenz sich auf die Selektion von *haben* und *sein* bei *degree achievements* (*wuchern, wachsen*) und Bewegungsverben (*laufen, springen*) auswirkt (Gillmann 2016: 244–252). Beide Verbklassen können Variation in der Selektion von *haben* und *sein* aufweisen. Auf die Verbklassen und ihren Einfluss auf die Auxiliarselektion wird in Abschnitt 3.3.2 näher eingegangen.

⁷⁷Innerhalb des Archivs sind Artikel und Diskussionen auf Wikipedia ausgeschlossen worden, da die Urheberschaft hier nicht eindeutig ist (Gillmann 2016: 245).

3 Variation in Konjugation, Deklination und Auxiliarselektion

Sowohl bei *degree achievements* als auch bei den Bewegungsverben führt hohe Tokenfrequenz zu einer stabilen Auxiliarselektion (Gillmann 2016: 253–255; 265–268). Dies ist aus einer gebrauchsbasierten Perspektive zu erwarten, da hohe Tokenfrequenz zu mental gefestigten, leicht aktivierbaren Formen und daher zu Stabilität führt. Zunächst wird der Einfluss der Tokenfrequenz auf *degree achievements* näher betrachtet. Wie Tabelle 3.3 zeigt, ist der Einfluss der Tokenfrequenz vor allem für *altern* zu erkennen: *Altern* weist eine klare Tendenz zu *sein* auf und ist mit 769 Perfektformen weitaus frequenter als die anderen untersuchten *degree achievements* *wuchern* (79 Perfektformen), *rosten* (49 Perfektformen), *faulen* (38 Perfektformen) und *schimmeln* (12 Perfektformen). Die wenigen *haben*-Belege für *altern* stammen zudem allesamt aus der Schweiz, sodass hier regionale Einflüsse eine Rolle zu spielen scheinen (Gillmann 2016: 255).

Tabelle 3.3: Auxiliarselektion der *degree achievements* aus Gillmann (2016: 254)

Partizip II	<i>haben</i>		<i>sein</i>		gesamt
	absolut	relativ	absolut	relativ	
<i>geschwitzt</i>	888	88,8 %	112	11,2 %	1000
<i>gealtert</i>	17	2,2 %	752	97,8 %	769
<i>gewuchert</i>	21	26,6 %	58	73,4 %	79
<i>gerostet</i>	14	28,6 %	35	71,4 %	49
<i>gefault</i>	5	13,2 %	33	86,8 %	38
<i>geschimmelt</i>	11	91,7 %	1	8,3 %	12

Frequenter als *altern* ist in Gillmanns (2016) Korpus nur *schwitzen* mit 1.000 Perfektformen. *Schwitzen* zeigt eine weniger eindeutige Verteilung als *altern*, die jedoch der funktionalen Verteilung (resultative vs. nicht-resultative Verwendung) geschuldet ist (Gillmann 2016: 255, siehe hierzu ausführlich Abschnitt 3.3.2.2). *Altern* selektiert hingegen verwendungsunabhängig *sein*. Für die weniger frequenten Verben zeigt sich eine weniger eindeutige Verteilung, auch wenn alle Verben klare Tendenzen zu *haben* oder *sein* aufweisen. Die scheinbar klare Tendenz von *schimmeln* zu *haben* könnte sich allerdings aufgrund der geringen Belegmenge zufällig ergeben haben.

Auch bei den Bewegungsverben kann Gillmann (2016: 265–268) einen klaren Einfluss der Tokenfrequenz nachweisen. Tabelle 3.4⁷⁸ gibt einen Überblick über

⁷⁸Die doppelten Belegzahlangaben bei *reiten* sind der idiomatischen Wendung *X_{NP} hat der Teufel geritten* geschuldet. Diese wird stets mit *haben* gebildet, da sie transitiv ist. Die eingeklammerten

3.3 Variation in der Selektion von *haben* und *sein*

die Auxiliarselektion der Bewegungsverben, *tanzen* wurde von Gillmann (2016: 265–268) als Kontrollverb genutzt, da *tanzen* eine Aktivität, aber keine Fortbewegung bezeichnet und daher *haben* selegiert (*Wir haben gestern viel getanzt*), wenn sich aus dem Kontext keine Fortbewegung ergibt.

Tabelle 3.4: Auxiliarselektion der Bewegungsverben aus Gillmann (2016: 267)

Verb	<i>haben</i>		<i>sein</i>	
	absolut	relativ	absolut	relativ
<i>gegangen</i>	0	0 %	14.633	100 %
<i>gefahren</i>	951	9,5 %	9012	90,5 %
<i>gelaufen</i>	68	2,8 %	2376	97,2 %
<i>geflogen</i>	154	7,3 %	1964	92,7 %
<i>geschwommen</i>	20	4,5 %	425	95,5 %
<i>geritten</i>	53	22,8 %	179	77,2 %
	(183)	(50,6 %)	(179)	(49,4 %)
<i>gejoggt</i>	5	11,4 %	39	88,6 %
<i>gewalkt</i>	2	8,7 %	21	91,3 %
<i>geskatet</i>	3	13,6 %	19	86,4 %
<i>gebik(e)t</i>	0	0 %	4	100 %
<i>gebouldert</i>	1	25 %	3	75 %
<i>getanzt</i>	704	95,9 %	30	4,1 %

Es zeigt sich eine generelle Tendenz zur *sein*-Selektion, die nur vom Kontrollverb *tanzen* nicht geteilt wird. Zudem fällt auf, dass die frequenten Bewegungsverben *gehen*, *fahren*, *laufen* und *fliegen* eine eindeutige *sein*-Präferenz aufweisen. Die wenigen *haben*-Belege bei *fahren*, *laufen* und *fliegen* sind transitiven Verwendungskontexten geschuldet: Werden nur intransitive Verwendungskontexte berücksichtigt, beobachtet Gillmann (2016: 268) keine *haben*-Selektion (siehe hierzu genauer Abschnitt 3.3.2.3).⁷⁹

Die mittelfrequenten Bewegungsverben *schwimmen*, *reiten* und *joggen* weisen dagegen leichte Schwankungen auf, wobei diese bei *schwimmen* mit 95,5 % *sein*-Selektion weniger stark ins Gewicht fallen als bei *reiten* (72 % *sein*-Selektion)

ten Zahlen zeigen die Belegzahlen mit der idiomatischen Wendung. Bei allen Verben bis auf *walken*, *skaten*, *biken* und *bouldern* hat Gillmann (2016: 266) nur eine Stichprobe des Korpus annotiert.

⁷⁹Für *gefahren* findet Gillmann (2016: 268) einen einzigen Beleg, für die anderen Verben keine.

3 Variation in Konjugation, Deklination und Auxiliarselektion

und *joggen* (88,6 % *sein*-Selektion). Auch die niedrigfrequenten entlehnten Bewegungsverben (*bouldern*, *biken*) weisen eine klare Tendenz zu *sein* auf, auch wenn diese Verteilung aufgrund der geringen Stichprobengröße zufällig sein könnte. Die Ergebnisse deuten darauf hin, dass die Bewegungsverben ein semantisches Cluster bilden, das auf neue Bewegungsverben ausgeweitet wird, sodass diese ebenfalls *sein*-selegieren (Gillmann 2018: 250)

Die Betrachtung des Einflusses von Frequenz auf die Selektion von *haben* und *sein* zeigt, dass die *haben*-Selektion typenfrequenter ist als die *sein*-Selektion. Da hier separate Schemata vorliegen, ist der potentielle Einfluss der Typenfrequenz auf ambige Strukturen beschränkt: Nur hier ist es möglich, dass die *haben*-Selektion aufgrund der hohen Typenfrequenz als Default greift und daher bevorzugt wird.

Die Tokenfrequenz hat einen stabilisierenden Einfluss auf die Auxiliarselektion. Dies zeigt sich bei Verbklassen, bei denen *haben* und *sein* möglich sind und somit Variationspotential in der Selektion zwischen *haben* und *sein* besteht. Sowohl bei den *degree achievements* als auch bei den Bewegungsverben sind tokenfrequente Verben stabiler in ihrer Selektion als tokeninfrequente. Die Tokenfrequenz führt zu mental gefestigten Formen, die konkurrierende Formen statistisch ausstechen (zum statistischen Vorkaufsrecht siehe Goldberg 2019: 74–94 sowie Abschnitt 2.1.3). Zwar variieren auch tokenfrequente Verben in ihrer Auxiliarselektion, dies ist jedoch funktional bspw. durch Resultativität bei *schwitzen* oder durch transitive Verwendungen bei den Bewegungsverben begründet. Im folgenden Abschnitt wird der Einfluss von Prototypizität auf die Auxiliarselektion diskutiert.

3.3.2 Prototypizität

Die Auxiliare *haben* und *sein* sind mit zwei Funktionen verbunden. Sie stellen somit zwei Konstruktionen dar. Die Funktionen, die mit *haben* oder *sein* verbunden sind, lassen sich als drei Gegensatzpaare sehen, die in prototypischer Beziehung zueinander stehen: Transitivität vs. Intransitivität, Atelizität vs. Telizität und nicht-vorliegende vs. vorliegende Bewegungssemantik (Shannon 1992, 1995). Die Bestandteile der Funktionen werden im Folgenden vorgestellt. Da die meisten Verben des Deutschen *haben* selegieren und die Selektion dieses Auxiliars daher als Normalfall zu betrachten ist (Duden 2016: § 659), liegt der Fokus in diesem Abschnitt darauf, die Faktoren zu beschreiben, die zu einer Selektion von *sein* führen, die typenfrequentiell betrachtet den Sonderfall darstellt.

3.3.2.1 Transitivität

Die Transitivität von Sätzen kann als Haupteinflussfaktor für die Selektion von *haben* und *sein* betrachtet werden. Transitive Verben selektieren *haben*: *Ich habe dir das Buch gegeben* (Gillmann 2016: 87–95). Allerdings führt Intransitivität nicht zwingend zur *sein*-Selektion, so bilden auch intransitive Verben wie *lachen* oder *weinen* das Perfekt mit *haben* (*Ich habe gelacht/geweint*).

Transitivität ist ein graduelles Konzept und wird mit Hopper & Thompson (1980: 251) als „global property of an entire clause, such that an activity is ‘carried over’ or ‘transferred’ from an agent to a patient“ verstanden. Sie erschöpft sich daher nicht in der Anwesenheit eines Subjekts und eines Akkusativobjekts, sondern wird durch weitere Eigenschaften bestimmt (Gillmann 2016: 46). Die Transitivität eines Satzes ist bspw. gesteigert, wenn die darin beschriebene Handlung volitional ausgeführt wird: *Ich schreibe deinen Namen* ist transitiver als *Ich vergesse deinen Namen* (Beispiele entnommen aus Gillmann 2016: 47). Neben der Volitionalität führen bspw. Punktualität und Telizität zu einer gesteigerten Transitivität (Gillmann 2016: 46–47): Der Satz *Ich schaukle das Kind in den Schlaf* ist punktuell und telisch und daher transitiver als der Satz *Ich schaukle das Kind den ganzen Vormittag*. Außerdem spielt Affiziertheit eine Rolle: So ist *Ich schlage die Vase kaputt* transitiver als *Ich schlage auf die Vase*. Das Verb im ersten Satz ist telisch und die Affiziertheit höher, weil das Subjekt einen größeren Einfluss auf das Objekt nimmt als im zweiten Satz.

Insgesamt identifizieren Hopper & Thompson (1980: 252) zehn Faktoren, die Einfluss auf die Transitivität eines Satzes nehmen, siehe Tabelle 3.5.⁸⁰

Der Zusatz *Agens und Patiens* beim Faktor PARTIZIPANT verdeutlicht, dass mehrere Partizipanten allein nicht zu einer höheren Transitivität führen (Hopper & Thompson 1980: 254–255). Beispiel 3 ist daher trotz nur eines Partizipanten transitiver als Beispiel 4, denn Beispiel 3 enthält eine Handlung, die telisch, volitional und punktuell ist.

- (3) *Ich gehe weg.*
- (4) *Ich mag Hafermilch.*

Beispiele nach Hopper & Thompson (1980: 254)

Beispiel 4 enthält zwar zwei Partizipanten, allerdings referiert das Verb nicht auf eine Handlung, sondern auf einen *state* (siehe Abschnitt 3.3.2.2 zu Ereignis-

⁸⁰Hopper & Thompson (1980: 252) nutzen statt *Patiens* den Terminus *Objekt*. Da sie aber selbst darauf eingehen, dass das Objekt ein *Patiens* sein muss, um die Transitivität zu erhöhen, wird in der Tabelle *Patiens* genutzt.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Tabelle 3.5: Einflussfaktoren auf Transitivität nach Hopper & Thompson (1980: 252)

Einflussfaktor	hohe Transitivität	niedrige Transitivität
Partizipanten	zwei oder mehr, Agens und Patiens	einer
Kinese	Handlung	Nicht-Handlung
Aspekt	telisch	atelisch
Punktualität	punktuell	nicht-punktuell
Volitionalität	volitional	nicht-volitional
Affirmation	affirmiert	negiert
Modus	Realis	Irrealis
Agentivität	hoch	niedrig
Affiziertheit des Objekts	hoch	nicht gegeben
Individuiertheit des Objekts	hoch	gering

klassen von Verben). Der zweite Partizipant kann daher nicht als Patiens eingestuft werden (Hopper & Thompson 1980: 254). Zudem sind Telizität, Volitionalität und Punktualität nicht gegeben und Beispiel 4 daher intransitiver als Beispiel 3 (Hopper & Thompson 1980: 254).⁸¹

Shannon (z. B. Shannon 1992, 1995) entwickelt aufbauend auf dem Transitivitätsmodell von Hopper & Thompson (1980) ein Prototypenmodell, um die Selektion von *haben* und *sein* vorherzusagen. Dabei führt er als Gegenpol zur Transitivität das Konzept der Mutativität ein und geht davon aus, dass prototypisch transitive Sätze *haben* selegieren, während prototypisch mutative Sätze *sein* selegieren. Er definiert Mutativität als „effective change in the patient subject“ (Shannon 1995: 133) und damit als Gegenstück zur Definition von Transitivität als „effective carrying over of an activity from an A to a patient“ nach Hopper & Thompson (1980: 279).

Ein hochmutativer Satz hat nach Shannon (1995: 133) im Gegensatz zu einem hochtransitiven Satz nur einen Partizipanten, der sich durch fehlende Agentivität und Volitionalität auszeichnet. Dabei führt eine geringe Transitivität nicht automatisch zu hoher Mutativität (Shannon 1995: 133). Stattdessen teilen sich Transitivität und Mutativität einige Eigenschaften wie bspw. die Telizität (Shannon 1995:

⁸¹Diese Argumentation lässt sich damit verdeutlichen, dass bspw. im Spanischen Sätze mit Experiencer wie in Beispielsatz 4 nicht mit Akkusativobjekt gebildet werden (Hopper & Thompson 1980: 254). Auch im Deutschen ist eine solche Struktur möglich: *Mir schmeckt Hafermilch.*

132–134, Gillmann 2016: 119–120): Wie in einem hochtransitiven Satz erfährt der Partizipant auch in einem hochmutativen Satz einen Zustandswechsel wie bspw. *Pfirsich* im Satz *der Pfirsich schimmelt*. Der im folgenden Abschnitt 3.3.2.2 diskutierte Einfluss der Telizität auf die Selektion von *haben* und *sein* lässt sich somit gut in das Modell von Shannon (1995) integrieren. Allerdings liefert sein Modell keine Antwort auf den Sonderstatus der Bewegungsverben im Deutschen, denn auch atelisch verwendete Bewegungsverben selegieren *sein* (*Ich habe/*bin im Kreis geschwommen*).

Die Betrachtung von Transitivität zeigt, dass diese grundlegenden Einfluss auf die Selektion von *haben* und *sein* nimmt. Die vielen Faktoren, die Einfluss auf die Transitivität eines Satzes nehmen, verdeutlichen, dass Transitivität prototypisch organisiert ist. Prototypisch transitive Sätze selegieren *haben*, prototypisch intransitive selegieren bei Telizität oder Bewegungssemantik *sein*. Der Einfluss der Prototypizität wird im empirischen Teil der Arbeit anhand von Sätzen mit Bewegungssemantik in unterschiedlichen Transitivitätsgraden untersucht.

3.3.2.2 Telizität

Bereits in der Diskussion um Transitivität bzw. Mutativität ist angeklungen, dass Telizität Einfluss auf die Auxiliarwahl nimmt. Telische Sätze sind nach Shannon (1995) mutativer als atelische. Daher selegieren intransitive telische Sätze *sein* (*Ich bin eingeschlafen*), während intransitive atelische Sätze *haben* selegieren (*Ich habe geschlafen*). Telizität ist eine Aktionsart, die grenzbezogene (telische) von nicht grenzbezogenen (atelischen) Ereignissen trennt. Dementsprechend können Verben, die diese Ereignisse benennen, als telisch bzw. atelisch klassifiziert werden (Shannon 1995: 132, Gillmann 2016: 33).

Auch in der *auxiliar selection hierarchy* (ASH) nach Sorace (2000), die Auxiliarselektion sprachübergreifend als graduelles System mit implikativen Stufen erklärt, führt Telizität zu *sein*-Selektion: Telische Bewegungsverben (*ankommen*) und Verben, die einen Zustandswechsel beschreiben (*einschlafen*), sind in der ASH mit der *sein*-Selektion assoziiert (Sorace 2000: 863–867). Die ASH ist auf das Deutsche jedoch nur bedingt anwendbar, da nach der ASH Bewegungsverben *haben* selegieren müssten (Gillmann 2016: 108–116).

Da telische Verben grenzbezogen sind, gehen sie im Gegensatz zu atelischen Verben mit einem Zustandswechsel einher. Das beschriebene Ereignis kann somit in verschiedene Phasen geteilt werden. Dies ist bei atelischen Verben nicht der Fall (Teuber 2005: 102–104, Gillmann 2016: 32–35): Während das atelische Ereignis *schlafen* sich nicht in Phasen unterteilen lässt, kann *einschlafen* in eine Präphase (kein Schlaf), den Zustandswechsel (*einschlafen*) und die Postphase

3 Variation in Konjugation, Deklination und Auxiliarselektion

(Schlaf) geteilt werden (Teuber 2005: 103, Gillmann 2016: 33), siehe Abbildung 3.10.



Abbildung 3.10: Semantik atelischer und telischer Verben nach Teuber (2005: 103)

Die telischen und atelischen Verben lassen sich verschiedenen Ereignisklassen zuordnen: Zu den telischen Verben zählen *accomplishments* (*einschlafen*) sowie *achievements* (*ankommen*). Atelisch sind hingegen *activities* (*lachen*) sowie *states* (*wissen*) (Gillmann 2016: 35–36; zur Unterscheidung von Verben nach Ereignisklassen siehe Vendler 1974: 97–121).

Bei *accomplishments* (*einschlafen*) und *achievements* (*ankommen*) ist die Telizität dem Verb inhärent, allerdings kann Telizität auch durch die syntaktische Verwendung entstehen. Dies ist bei *degree achievements* (*schimmeln*, *wachsen*) der Fall. Diese Verben beschreiben eine graduelle Veränderung, implizieren dabei jedoch nicht zwangsläufig einen Endpunkt (Dowty 1979: 88–90, Gillmann 2016: 37). Die Verben changieren deshalb zwischen einer telischen Lesart als *achievement* (mit Fokus auf der Zustandsänderung) und einer atelischen als *activity* (mit Fokus auf der Handlung): So kann bspw. *schimmeln* atelisch verwendet werden (*Der Pffirsich schimmelt schon etwas vor sich hin*), aber auch telisch (*Der Pffirsich schimmelt bis zum Kern*). Dies hat Einfluss auf die Auxiliarselektion: Während bei der telischen Lesart *sein* genutzt wird (*Der Pffirsich ist/*hat bis zum Kern geschimmelt*), selegiert die atelische Lesart *haben*: *Der Pffirsich *ist/hat schon etwas vor sich hin geschimmelt* (Gillmann 2016: 261).

Sätze, die hinsichtlich ihrer Klassifikation als *achievement* oder *activity* ambig sind, lassen beide Auxiliare zu: *Der Pffirsich ist/hat geschimmelt*. Die Auxiliarselektion kann hierbei mit verschiedenen Lesarten einhergehen: Während bei *hat geschimmelt* der Vorgang fokussiert wird, wird bei *ist geschimmelt* der Endzustand fokussiert. Somit eröffnen sich zwei Interpretationen: Einerseits kann die Konstruktion unabhängig von der Auxiliarwahl als Tempus gelesen werden, andererseits kann die Wahl des Auxiliars zur aspektuellen Differenzierung dienen, indem sie auf das Ereignis (Tempus) oder den Endzustand (Resultativ) eines Ereignisses fokussiert.

Die resultative Lesart wird durch *sein* begünstigt (Gillmann 2016: 131), eine temporale Lesart ist für *sein* aber nicht ausgeschlossen: Beispiel 5 und 6 enthalten ambige Sätze, in denen für *sein* sowohl die resultative als auch die temporale

Lesart möglich ist, Beispiel 7 zeigt einen Satz, der nur temporal gelesen werden kann.

- (5) *Der Pfirsich ist geschimmelt.*
→ resultative sowie temporale Lesart möglich
- (6) *Der Pfirsich ist geschimmelt, sodass ich ihn nun nicht mehr essen kann.*
→ resultative sowie temporale Lesart möglich
- (7) *Der Pfirsich ist letzte Woche in dem warmen Zimmer geschimmelt, sodass ich ihn nicht mehr essen konnte.*
→ nur temporale Lesart möglich

Beispiel 5 und 6 können resultativ und temporal gelesen werden (Gillmann 2016: 66–67). Die resultative Lesart zeigt sich nach Gillmann (2016: 66–67) bspw. daran, dass Temporaladverbiale mit Gegenwartsbezug wie *nun* in Beispiel 6 möglich sind. Zudem ist eine Modifikation mit *immer (noch)* möglich (*Der Pfirsich ist immer noch geschimmelt*),⁸² die als Test für die Resultativlesart gilt (Bybee u. a. 1994: 65, Gillmann 2016: 66). In Beispielsatz 7 wird mit *letzte Woche* hingegen ein Adverbial mit Vergangenheitsbezug genutzt, sodass nur die temporale Lesart möglich ist.

Da das Resultativ die Nachphase eines Ereignisses fokussiert, können nur telische Verben diese Interpretation ermöglichen, da atelische Verben nicht in Phasen geteilt werden können (Teuber 2005: 103).⁸³ Die resultative Lesart stellt sich vorrangig bei *sein* + Partizip II ein, ist aber für *haben* nicht ausgeschlossen.⁸⁴

⁸²Da *schimmeln* nicht reversibel ist, ist die Verwendung mit *immer noch* semantisch fragwürdig. Im Satz *Das Kind ist immer noch nass geschwitzt, du solltest es baden* ist die Nutzung m. E. sinnvoller.

⁸³Die resultative Lesart stellt den ersten Grammatikalisierungsschritt hin zu einer Vergangenheitsbedeutung dar: Der Fokus verschob sich im Lauf der Grammatikalisierung vom Resultat zum Ereignis, das dem Resultat vorausgeht (siehe hierzu ausführlich Gillmann 2011: 206, 2016: 60–68). Wie oben gezeigt, kann die Konstruktion *haben/sein* + Partizip II aber auch im Neuhochdeutschen eine resultative Bedeutung haben.

⁸⁴Leiss (1992: 182–186) sowie Teuber (2005: 122–138) stellen die These auf, dass *sein* nur resultative Lesarten hervorrufen könne und somit kein Perfekt bilde. *Sein* fungiere somit nicht als Auxiliar, sondern als Kopulaverb. Das Zustandspassiv (*Die Tür ist geschlossen*) und das *sein*-Perfekt sind nach dieser These somit nicht auseinanderzuhalten. Gillmann (2016: 126–132) argumentiert schlüssig gegen diese These, da bspw. auch *haben* eine resultative Lesart ermöglicht wie im Satz *Der Hund hat die Augen gerade noch geschlossen*. Die resultative Lesart stellt sich bei *sein* lediglich leichter ein als bei *haben* (Gillmann 2016: 131). Zudem sind auch bei *sein* + Partizip II nicht-resultative Lesarten möglich, da bspw. Temporaladverbiale mit Vergangenheitsbezug möglich sind: *Der Pfirsich ist gestern geschimmelt*.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Sowohl Resultativität als auch Telizität nehmen Einfluss auf die Selektion von *haben* und *sein* bei *degree achievements* (Gillmann 2016: 252–264).⁸⁵ *Sein* wird bei resultativer Lesart genutzt (siehe Beispiel 8) und bei einer nicht-resultativen, aber telischen Lesart (siehe Beispiel 9). *Haben* wird dagegen selegiert, wenn weder eine telische noch resultative Lesart erzielt werden soll (siehe Beispiel 10).

- (8) *Sieh mal, der Pfirsich ist jetzt verschimmelt/bis zum Kern geschimmelt.*
[+telisch] [+resultativ] -> *sein*-Selektion
- (9) *Der Pfirsich ist letzte Woche verschimmelt/bis zum Kern geschimmelt.*
[+telisch] [- resultativ] -> *sein*-Selektion
- (10) *Der Pfirsich hat schon etwas vor sich hin geschimmelt, ist aber noch essbar.*
[- telisch] [- resultativ] -> *haben*-Selektion

Den Einfluss der Resultativität und Telizität kann Gillmann (2016: 252–265) in ihrer Korpusuntersuchung weitgehend bestätigen, allerdings zeigen sich lexemspezifische Unterschiede zwischen den untersuchten Verben. Gillmann (2016: 252–265) untersucht die *degree achievements* *schwitzen*, *altern*, *wuchern*, *rosten*, *faulen* sowie *schimmeln*. Nur für *schwitzen* stellt sie einen eindeutigen Einfluss der Resultativität fest. *Geschwitzt sein* tritt in allen 112 Belegen in resultativer Bedeutung auf (*Er ist pitschnass geschwitzt*). Bei *faulen* und *rosten* zeigen sich Tendenzen zur Markierung von Resultativität (Gillmann 2016: 257–258). So wird *gefault/gerostet sein* in den insgesamt 35 bzw. 37 Belegen zehnmal resultativ verwendet. Allerdings wird die Konstruktion auch in perfektischer Bedeutung genutzt: Bei *gefault sein* ist die perfektische Bedeutung mit neun Belegen ähnlich häufig wie die resultative. Bei *gerostet sein* ist die perfektische Nutzung mit 18 Belegen sogar häufiger als die resultative (Gillmann 2016: 256). Zudem beobachtet Gillmann (2016: 256–258) 14 ambige Verwendungen für *gefault sein* sowie sieben für *gerostet sein*, die implizieren, dass *gefault/gerostet sein* zwar eine resultative Lesart erlauben, aber nicht erzwingen. Für *gewuchert* und *geschimmelt sein* kann Gillmann (2016: 256–258) keine resultative Verwendung feststellen.

Bezüglich der Telizität erzielt Gillmann (2016: 259–262) eindeutigere Ergebnisse: Bei telischen Prozessen wird *sein*, bei atelischen *haben* selegiert. Von dieser Verteilung weichen nur drei Belege ab: *Haben* tritt zweimal bei telischen Prozessen auf (*hat zu einem matten Rot gerostet*, *hat in die Lorzeebene gewuchert*), *sein* einmal bei einem atelischen Prozess (*vor sich hin gerostet war*) (Gillmann 2016: 261). Die Telizität eines Ereignisses macht Gillmann (2016: 259) dabei anhand von Direktionalergänzungen (*bis zum Ende/bis zur Mitte*) sowie an Konsekutivsätzen

⁸⁵Dabei ist zu beachten, dass Resultativität zwar Telizität voraussetzt, Telizität aber nicht zwangsläufig zu einer resultativen Lesart führt. So sind in Beispiel 9 Temporaladverbiale mit Vergangenheitsbezug möglich, die eine resultative Lesart ausschließen.

fest, die eine Grenze des Ereignisses implizieren (*Das Schloss ist inzwischen so gerostet, dass der Schlüssel feststeckt*).

In ambigen Kontexten ist davon auszugehen, dass die Auxiliarselektion eine telische bzw. atelische Interpretation begünstigt (Gillmann 2016: 260):

- (11) *Die Milch hat in der letzten Woche stark geschimmelt.*
→ atelisch, Fokus auf den unabgeschlossenen Vorgang
- (12) *Die Milch ist in der letzten Woche stark geschimmelt.*
→ telisch, Fokus auf Endzustand
(Beispiele angelehnt an Gillmann (2016: 260))

Die Bedeutungsnuancen sind jedoch metonymisch verbunden, sodass die Bedeutungsunterschiede sehr fein sind. Gillmann (2016: 260) nimmt daher an, dass die Auxiliare in ambigen Kontexten austauschbar sind.

Hier ist der Einfluss der Tokenfrequenz auf *degree achievements*, der in Abschnitt 3.3.1 bereits diskutiert wurde, relevant: Dadurch, dass *degree achievements* oft ambig zwischen bspw. telischer und atelischer Verwendung sind, kann die Tokenfrequenz als stützender Faktor Einfluss nehmen, da die Auxiliare lexemspezifisch durch die häufige Verwendung gefestigt werden und in der Folge das konkurrierende Auxiliar statistisch ausstechen können.

Der Einfluss von Telizität auf die Auxiliarselektion konnte auch neurolinguistisch nachgewiesen werden. Roehm u. a. (2013) führen eine EEG-Studie durch, in der sie Sätze mit telischen (Bewegungs)-Verben sowie mit *degree achievements* präsentieren. Sie stützen ihre Hypothesen auf die ASH von Sorace (2000) (Roehm u. a. 2013: 1246–1248), weshalb sie Bewegungsverben nicht gesondert betrachten, sondern allein Telizität als Einflussfaktor berücksichtigen. Die Testsätze sind identisch aufgebaut. Zusätzlich zu den EEG-Daten werden Akzeptabilitätswerte für die Testsätze erhoben.

Bei telischen Verben wird *sein* zu 88 % bis 93 % akzeptiert, bei den atelischen lediglich zu 17 % (Roehm u. a. 2013: 1250). Zudem weisen die ereigniskorrelierten Potentiale (EKP) N400 und P600 darauf hin, dass das Auxiliar *haben* bei telischen (Bewegungs)-Verben (*aufsteigen, verrosten*) überraschend ist (Roehm u. a. 2013: 1251–1255). Die priore Wahrscheinlichkeit für die Selektion von *haben* ist bei telischen (Bewegungs-)Verben gering, da mit *sein* eine andere konventionellere Form existiert, die *haben* in diesem Kontext statistisch aussticht. Umgekehrt evoziert *sein* bei atelischen Verben wie *lachen* N400- und P600-Effekte.⁸⁶

⁸⁶Laut Roehm u. a. (2013: 1266–1267) deuten die ereigniskorrelierten Potentiale auf verschiedene Prozesse im Gehirn hin. Den zu beobachtenden N400-Effekt sehen sie als semantischen Effekt: Die Verbklasse der telischen Verben ist mit *sein*-Selektion assoziiert. Tritt ein telisches Verb mit *haben* auf, ist dies semantisch unerwartet, weshalb es einen N400-Effekt auslöst. Den P600-Effekt deuten Roehm u. a. (2013: 1266–1267) als Effekt, der reflektiert, dass eine Form als ungrammatisch eingestuft wurde.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Interessant ist ein Blick auf die Prozessierung der *degree achievements* (*rosten*), da diese Variation erlauben. Für *degree achievements* beobachten Roehm u. a. (2013: 1257–1258) in der globalen Analyse der Daten keine Effekte. Auch die Akzeptabilitätsraten der Proband_innen deuten mit durchschnittlich 60 % bis 65 % Zustimmung für beide Auxiliare darauf hin, dass die *degree achievements* sowohl *haben* als auch *sein* zulassen.

Roehm u. a. (2013: 1257–1264) beobachten allerdings Effekte, wenn die behavioralen Daten und die EKP miteinander korreliert werden. So zeigt sich beim Auxiliar *haben* eine negative Korrelation von N400-Effekten und Akzeptabilität für die verschiedenen Testitems: Wenn ein Verb also hohe Akzeptabilitätswerte für *haben* aufweist, dann sind weniger N400-Effekte für dieses Verb in Verbindung mit *haben* gemessen worden (Roehm u. a. 2013: 1268). Roehm u. a. (2013: 1268) interpretieren dieses Ergebnis als Indiz dafür, dass auch Zustandsänderungen einen unterschiedlichen Grad an Telizität aufweisen, der sich in der (In-)Akzeptanz der *haben*-Variante spiegelt: Die „telischeren“ *degree achievements* erzielen bei der dispräferierten *haben*-Selektion eine geringe Akzeptanz, die auch mit einem N400-Effekt korreliert. Das Auxiliar *haben* wird hier also nicht antizipiert (da die Verbklasse mit *sein* assoziiert ist) und löst deswegen einen N400-Effekt aus. Bei den weniger telischen Verben bleibt der Effekt (sowie die niedrigen Akzeptabilitätsraten) hingegen aus, da *haben* hier eine größere priore Wahrscheinlichkeit hat und somit antizipiert wird.

Hinsichtlich des P600-Effekts zeigen sich komplexere Korrelationen als bei den N400-Effekten. Zunächst lässt sich eine zu erwartende negative Korrelation zwischen P600-Effekt und einer hohen Akzeptanz für beide Auxiliare beobachten: Je stärker akzeptiert das Auxiliar ist, desto seltener ist ein P600-Effekt messbar (Roehm u. a. 2013: 1269). Allerdings zeigt sich eine interessante Korrelation für das Auxiliar *haben*, wenn die Akzeptabilität von verschiedenen Proband_innen in Betracht gezogen wird. Der P600-Effekt ist größer, je stärker die Proband_innen die *haben*-Variante akzeptieren. In diesem Fall widersprechen sich somit Akzeptanz und Prozessierung. Roehm u. a. (2013: 1269) erklären diesen Widerspruch als pragmatische Anreicherung: Sie gehen davon aus, dass einige Proband_innen ein telisches Verb als *activity* und somit atelisch uminterpretieren, weswegen die Akzeptabilitätswerte für *haben* steigen. Dies wird jedoch nicht in der Prozessierung gespiegelt. Da die pragmatische Anreicherung nur personenbezogen beobachtbar ist, gehen Roehm u. a. (2013: 1269) davon aus, dass sie eine individuelle Komponente in der Prozessierung darstellt.

Insgesamt bestätigt die Studie von Roehm u. a. (2013) die theoretischen Überlegungen: Prototypische Telizität bzw. Atelizität ist in intransitiven Sätzen mit *sein*- bzw. *haben*-Selektion verbunden. Das jeweils andere Auxiliar wird daher

statistisch ausgestochen und löst P600- und N400-Effekte aus, wenn es verwendet wird. Im Übergangsbereich zwischen Telizität und Atelizität ist dagegen Variation zu beobachten: Das verdeutlichen die *degree achievements*, die sowohl eine telische als auch eine atelische Interpretationen ermöglichen. Zudem deuten die Ergebnisse darauf hin, dass einige *degree achievements* als telischer wahrgenommen werden als andere. Der Einfluss der Telizität bei *degree achievements* zeigt sich auch in der Korpusuntersuchung von Gillmann (2016): Sätze mit *degree achievements*, die bspw. aufgrund von Direktionaladverbialen klar als telisch interpretierbar sind, selektieren *sein*. Gillmann (2016) beobachtet zudem, dass bei einzelnen Verben wie *schwitzen* Resultativität die Auxiliarselektion beeinflusst.

3.3.2.3 Bewegungsemantik

Bewegungsverben wie *fahren*, *gehen* und *fliegen* selektieren ausnahmslos *sein*, wenn sie nicht in transitiven Kontexten gebraucht werden. Die *sein*-Selektion ist für Bewegungsverben zudem produktiv, wie Gillmann (2016: 274–280) zeigt: Neu entlehnte Bewegungsverben wie *bouldern* und *biken* schwanken zwischen *haben* und *sein* (*Ich bin/habe gestern gebouldert*), während neu entlehnte Verben ohne Bewegungsemantik (*googeln*, *twittern*) ausnahmslos *haben* aufweisen (*Ich habe/*bin getwittert*). Dass *bouldern* und *biken* trotz Bewegungsemantik keine eindeutige Präferenz für *sein* aufweisen, sondern in ihrer Verwendung schwanken, liegt an der geringen Tokenfrequenz dieser Verben (Gillmann 2016: 272–274, siehe Abschnitt 3.3.1). Neben der geringen Tokenfrequenz ist zu berücksichtigen, dass die neuen Verben Sportarten benennen: Es kann somit entweder die Bewegung an sich oder die Tätigkeit im Sinne einer Sportart im Vordergrund stehen und daher *sein* bzw. *haben* selektiert werden. Wenn jedoch durch den Kontext eine Bewegungsemantik begünstigt wird (*Ich bin durch die Gegend geskatet*), wird *sein* selektiert. Bei den *haben*-Belegen, die Gillmann (2016: 277) aufführt, ist hingegen eine Interpretation als Tätigkeit möglich (*ich habe nie geskatet, bin über 40, Unternehmer und Arbeitgeber*).

Neben neuen Bewegungsverben liefern Verben wie *tanzen* Hinweise auf die Produktivität des *sein*-Schemas bei Bewegungsemantik: *Tanzen* bildet das Perfekt normalerweise mit *haben* (*ich habe/*bin getanzt*), wenn aber durch den Kontext eine Fortbewegung impliziert wird (*Sie sind/*haben durch den Raum getanzt*), wird *sein* selektiert (Randall u. a. 2004: 334–335, Gillmann 2016: 275–276). Die Selektion von *sein* hängt im Beispielsatz *Sie sind durch den Raum getanzt* nicht von Telizität ab, da *durch* zwar eine Richtung impliziert, aber keinen Endpunkt.

Gillmann (2018: 246–247) zeigt, dass auch bei Bewegungsverben die Selektion von *sein* nicht an Telizität gebunden ist, sondern sich aus der Bewegungseman-

3 Variation in Konjugation, Deklination und Auxiliarselektion

tik ergibt (*Ich bin/*habe im Kreis gelaufen*), denn unabhängig von der Telizität des Satzes überwiegt bei Bewegungsverben die *sein*-Selektion mit 80 bis 100 % deutlich.⁸⁷ Die Schwankungen ergeben sich aus Unterschieden in der Tokenfrequenz: Die infrequenten Bewegungsverben weisen mehr *haben*-Selektion auf als die frequenten.

Gillmann (2018: 246) beobachtet zudem, dass Bewegungsverben häufiger telisch als atelisch genutzt werden. Somit lässt sich die Hypothese aufstellen, dass die hohe Tokenfrequenz der *sein*-Selektion bei Bewegungsverben in telischen Kontexten dazu geführt hat, dass Bewegungsverben auch in atelischen Kontexten *sein* selegieren (Gillmann 2018: 247). In diesem Fall scheint also hohe Tokenfrequenz zur Produktivität einer Konstruktion beizutragen.

Die einzige Ausnahme von der *sein*-Selektion bei Bewegungsverben sind Sätze mit hoher Transitivität (*Ich habe/*bin die Kinder in die Schule gefahren*): Gillmann (2016: 281–286) stellt in ihrer Korpusuntersuchung in hochtransitiven Kontexten zu 100 % *haben*-Selektion bei Bewegungsverben fest. Der Einfluss der Transitivität ist dabei graduell. Je transitiver ein Satz ist, desto wahrscheinlicher wird *haben*. Dabei nimmt die Belebtheit des Objekts Einfluss auf die Transitivität und somit auf die Auxiliarselektion (Gillmann 2016: 280–302). Sätze mit Akkusativobjekten, die auf eine belebte Entität wie bspw. *Kinder* referieren, selegieren allesamt *haben*, da das Objekt eindeutig als Patiens interpretiert werden kann.

Nur bei *reiten* kann die Interpretation eines Objekts mit Referenz auf eine belebte Entität ambig sein: Im Satz *Ich reite ein Pferd* kann *Pferd* sowohl als Patiens als auch als Instrument gesehen werden, da das Pferd als Fortbewegungsmittel genutzt wird. Die beiden Interpretationen sind mit unterschiedlichen Transitivitätsgraden verknüpft. Daher sind hier beide Auxiliare möglich: *Ich bin/habe das Pferd geritten*. Bei Objekten mit Referenz auf Konkreta ergibt sich ein ähnliches Problem in Hinblick auf die Verben *fahren* und *fliegen*: Fortbewegungsmittel wie

⁸⁷Die *sein*-Selektion von atelischen Bewegungsverben mit agentivem Subjekt (*Ich bin im Kreis geschwommen*) sieht Gillmann (2016: 105–107) als Indiz dafür, dass die Vorhersagen der Unakkusativitätshypothese zur Auxiliarselektion für das Deutsche nicht haltbar sind. Die Unakkusativitätshypothese geht davon aus, dass sich intransitive Verben in unakkusative und unergative Verben teilen lassen. Das Argument bei unergativen Verben (*ich lache*) entspricht dabei dem Subjekt eines transitiven Satzes; bei unakkusativen Verben (*ich falle*) hingegen dem Objekt eines transitiven Satzes (Gillmann 2016: 95–96). Nach der Unakkusativitätshypothese sind Unakkusativität und Unergativität sprachübergreifend mit der Auxiliarselektion verknüpft: Unakkusative Verben selegieren *sein*, unergative dagegen *haben* (Gillmann 2016: 100–101). Dies lässt sich für einige Sprachen wie bspw. für das Italienische auch nachweisen (Gillmann 2016: 102–103). Für das Deutsche funktioniert dies jedoch nicht: Aufgrund der Atelizität und der thematischen Rolle des Subjekts sind Bewegungsverben wie *schwimmen* als unergativ einzustufen und müssten somit *haben* selegieren (Gillmann 2016: 105–107). Für einen Überblick über die Unakkusativitätshypothese siehe Alexiadou u. a. (2004).

3.3 Variation in der Selektion von *haben* und *sein*

Auto, *Transporter* und *Helikopter* können sowohl *Patiens* als auch *Instrument* sein. Auch hier sind daher beide Auxiliare möglich. Insgesamt überwiegt bei *fahren* und determinierten Konkreta, die auf Fortbewegungsmittel referieren, die *haben*-Selektion mit 80 %, bei *fliegen* liegt sie nur bei 50 % (Gillmann 2016: 286).

In Sätzen, in denen das Konkretum eindeutig als *Patiens* interpretiert werden kann, wird aufgrund der hohen Transitivität nur *haben* selegiert (Gillmann 2016: 284–285): Dies ist bspw. der Fall, wenn eine *Direktionalergänzung* (*von A nach B*, *bis nach B*) genutzt wird (*Janis hat den Transporter nach Bamberg gefahren*). Die Transitivität des Satzes ergibt sich dabei nicht nur aus der *Patiens*-rolle. Der Satz mit *Direktionalergänzung* ist auch deswegen transitiver als ein Satz ohne *Direktionalergänzung*, weil die *Direktionalergänzung* einen Zielpunkt setzt und hierdurch die *Telizität* des Satzes erhöht. Zusätzlich trennt die *Direktionalergänzung* das *Akkusativobjekt* und das *Verb* voneinander, sodass eine Interpretation als *inkorporiertes Objekt* ausgeschlossen ist.

Gillmann (2016: 286–291) überprüft, ob die *Definitheit* des *Akkusativobjekts* einen Einfluss auf die *Auxiliarselektion* bei Konkreta nimmt. Dies scheint jedoch nicht der Fall zu sein. Allerdings zeigt sich, dass bei *inkorporierten Objekten* (*Ich bin schon einmal Auto gefahren*) fast ausnahmslos *sein*-Selektion vorliegt (Gillmann 2016: 290–291). *Inkorporierte Objekte* ordnet Gillmann (2016: 290) dem *intransitiven Pol* zu, da *Objekt* und *Verbalhandlung* aufgrund der generischen Verwendung des *Substantivs* nicht klar voneinander abgegrenzt werden können.

Stellt die *Akkusativergänzung* im Satz ein *Abstraktum* dar, neigen *Bewegungsverben* stark zur *Perfektbildung* mit *sein* (*eine Strecke gelaufen sein*). Dies ist nicht verwunderlich, da die von Gillmann (2016: 291–292) untersuchten *Abstrakta* nicht als *Objekte*, sondern als *Adverbiale* fungieren. Somit können die Sätze mit *Abstraktum* als *Akkusativergänzung* allesamt als *intransitiv* gewertet werden. Auch für *Abstrakta* hat Gillmann (2016: 292–295) den Einfluss der *Determiniertheit* überprüft. Bei *definiten Objekten* (*das Rennen*) wird eher *haben* selegiert als bei *indefiniten* (*ein Rennen*). Ist das *Abstraktum* in eine *Verbalhandlung* *inkorporiert*, liegt die *sein*-Selektion bei 100 % (Gillmann 2016: 295). Die *haben*-Selektion ist in Sätzen mit *inkorporierten Objekten* somit (unabhängig von deren *Belebtheitsgrad*) generell unwahrscheinlich. Zudem stellt Gillmann (2016: 299–301) fest, dass bei *Abstrakta*, die auf einen *Pfad* referieren (*eine Strecke laufen*), fast ausnahmslos *sein* selegiert wird. *Abstrakta*, die auf ein *Ergebnis* referieren (*die schnellste Zeit*), tendieren hingegen stärker zur Selektion von *haben* als andere *Abstrakta* (Gillmann 2016: 295–296). Insgesamt dominiert aber auch bei *Abstrakta* mit *Ergebnisreferenz* *sein*.

Abbildung 3.11 gibt einen Überblick über die Selektion von *haben* und *sein* bei *Bewegungsverben*.

3 Variation in Konjugation, Deklination und Auxiliarselektion

Tätigkeit Früher habe ich gerne geschwommen		Ich bin/habe geschwommen		Bewegung Ich bin durch den See geschwommen	
transitiv Ich habe die Kinder in die Schule gefahren		Ich bin/habe den Mercedes gefahren		intransitiv Ich bin Auto gefahren	
Ich bin/habe den Marathon gelaufen		Ich bin/habe einen Marathon gelaufen		Ich bin Marathon gelaufen	
haben		haben/sein		sein	

Abbildung 3.11: Selektion von *haben* und *sein* bei Bewegungsverben

Weitere empirische Evidenz für den Einfluss der Bewegungssemantik in Abhängigkeit von der Transitivität des Satzes liefern Hinze & Köpcke (2007). Hinze & Köpcke (2007) führen ein Produktionsexperiment mit 114 Grundschüler_innen durch. Dabei werden Bewegungsverben in Testsätzen mit unterschiedlichen Transitivitätsgraden abgefragt. Zusätzlich werden Bewegungsverben getestet, die auch als Aktivität gedeutet werden können (*schwimmen*), sowie Verben, die eine Bewegung lediglich implizieren (*tanzen*, *humpeln*).⁸⁸

Hinze & Köpcke (2007: 127) beobachten einen klaren Einfluss der Transitivität auf die Auxiliarselektion. Sätze ohne Patiens (*Uwe ____ nach Hause gefahren*) zeigen zu über 94 % *sein*-Selektion, während Sätze mit Patiens (*Die Busfahrerin ____ 24 Reisegäste sicher nach Köln gefahren*) zu über 89 % *haben*-Selektion aufweisen. Zwischen diesen Polen bewegen sich Sätze, die Ambiguitäten aufweisen: Sätze mit einem Abstraktum als Akkusativergänzung (*Der Pilot ____ wegen des schlechten Wetters einen Umweg geflogen*) weisen zu 80 % *sein*-Selektion auf. Bei Konkreta ohne Determinierer (*Der Rennfahrer ____ fünf Jahre lang Ferrari gefahren*) ist der Anteil für *sein* mit 60 % deutlich geringer und bei Konkreta mit Determinierer (*Der Millionär ____ einen roten Ferrari gefahren*) liegt der *sein*-Anteil nur noch bei 50 %. Die Ergebnisse von Hinze & Köpcke (2007) decken sich somit mit den Ergebnissen von Gillmann (2016).

Hinsichtlich der Bewegungssemantik lassen sich die Ergebnisse von Hinze & Köpcke (2007) weniger systematisch betrachten, da Hinze & Köpcke (2007: 105) Bewegungssemantik und Telizität verschränken. Dennoch ergeben sich interessante Ergebnisse: Für *schwimmen* zeigen sich mit 67 % *sein*-Selektion Schwankungen in der Auxiliarwahl, wenn im Satz nicht zwischen Bewegung und Aktivität

⁸⁸Hinze & Köpcke (2007: 102–103) berücksichtigen auch Telizität (bei ihnen Durativität/Resultativität) als Einflussfaktor auf die Auxiliarselektion bei Bewegungsverben. Hierbei gehen sie aber problematischerweise davon aus, dass Bewegungsverben inhärent telisch seien, dies aber im Kontext nicht immer zum Tragen käme (Hinze & Köpcke 2007: 102–103). Für die Bewegungsverben ist in ihrem Experiment m. E. kein Einfluss der Telizität feststellbar. So wird *sein* sowohl im telischen Satz (*Peter ____ von Hannover nach Münster gefahren*) als auch im atelischen Satz (*Der LKW-Fahrer ____ 12 Stunden ununterbrochen gefahren.*) zu über 90 % gewählt (Hinze & Köpcke 2007: 112).

3.3 Variation in der Selektion von *haben* und *sein*

desambiguiert wird (*Melina ____ früher jeden Tag eine Stunde lang geschwommen*). Ist jedoch eine Fortbewegung impliziert, liegt der *sein*-Anteil bei 95 % (*Janis ____ von einem Beckenrand zum anderen geschwommen*) (Hinze & Köpcke 2007: 112–113).

Ähnliches lässt sich für *humpeln* und *tanzen* feststellen: Bei klarer Bewegungssemantik liegt die *sein*-Selektion bei 99 % (*Irene ____ aus dem Zimmer gehumpelt*); ist die Bewegungssemantik jedoch nicht eindeutig, liegt die *sein*-Selektion nur bei 49 % (*Nach seinem Unfall ____ er zwei Monate gehumpelt*). Bei *tanzen* überwiegt mit 92 % klar *haben* (*Patrick ____ früher sehr gern getanzt*), bei eindeutiger Bewegungssemantik fällt die *haben*-Selektion allerdings auf 25 % (*Patrick ____ mit seiner Partnerin durch den ganzen Saal getanzt*) (Hinze & Köpcke 2007: 112–113).

Interessant ist ein Blick auf die Sätze, in denen ein Akkusativobjekt verwendet wurde: Bei *schwimmen* kann das Akkusativobjekt eindeutig als Patiens bestimmt werden, sodass der hohe *haben*-Anteil von 83 % erwartbar ist (*Franziska ____ ihre Konkurrenten müde geschwommen*). Bei *humpeln* (*Der Großvater ____ mehrere Kilometer gehumpelt*) kann die Akkusativergänzung *mehrere Kilometer* als Adverbial gedeutet werden, sodass der Satz weniger transitiv ist als der Testsatz *Franziska ____ ihre Konkurrenten müde geschwommen*. Dementsprechend wird *sein* zu 74 % gewählt. Bei *tanzen* (*Patrick ____ mit seiner Partnerin durch den ganzen Saal Tango getanzt*) zeigt sich mit 54 % *sein*-Selektion keine eindeutige Tendenz zu einem Auxiliar (Hinze & Köpcke 2007: 112–113). *Tango* könnte im Satz als eigenständiges oder inkorporiertes Objekt interpretiert werden.

Der Einfluss von Prototypizität wird im empirischen Teil der Arbeit untersucht, in dem Sätze mit Bewegungsverbren in unterschiedlichen Transitivitätsgraden präsentiert werden. Die Sätze sind eindeutig transitiv (*Kinder zur Schule gefahren haben/sein*) oder eindeutig intransitiv (*zur Kur gefahren sein/haben*). Zudem werden ambige Sätze präsentiert, bei denen das Objekt ein Fortbewegungsmittel darstellt (*ein Cabrio gefahren haben/sein*) oder die Akkusativergänzung ambig zwischen Objekt und Adverbial ist (*eine Strecke gefahren sein/haben*).

Insgesamt zeigt die Betrachtung des Einflusses von Bewegungssemantik auf die Auxiliarselektion, dass Sätze mit Bewegungssemantik *sein* selektieren. Die Bewegungssemantik kann sich dabei durch Bewegungsverbren ergeben oder durch Kontexte, die eine Bewegung implizieren. Nur Transitivität führt in Sätzen mit Bewegungssemantik zur Selektion von *haben*. Im Übergangsbereich zwischen Transitivität und Intransitivität ist Variation beobachtbar, etwa wenn ein Satz mit Bewegungsverb ein Konkretum als Akkusativobjekt aufweist, das auf ein Fortbewegungsmittel referiert. In diesen Sätzen kann das Objekt als Patiens oder Instrument interpretiert werden, sodass die Sätze hinsichtlich ihrer Transitivität

3 Variation in Konjugation, Deklination und Auxiliarselektion

ambig sind. Auch der Einfluss der Bewegungssemantik ist als graduell zu betrachten. Dies zeigen Verben, die entweder als Tätigkeit (Sportart) oder als Bewegung interpretiert werden können.⁸⁹ Sowohl der Einfluss der Bewegungssemantik als auch der Einfluss der Transitivität ist daher prototypisch gestaffelt. Der nächste Abschnitt fasst den Einfluss der Frequenz und der Prototypizität, die sich durch die prototypisch organisierten Einflussfaktoren Transitivität, Telizität und Bewegungssemantik ergibt, auf die Auxiliarselektion zusammen.

3.3.3 Zusammenfassung und Zusammenwirken der Faktoren

Sowohl Prototypizität als auch Frequenz nehmen Einfluss auf die Auxiliarselektion von *haben* und *sein*. Hinsichtlich der Prototypizität konnten Transitivität, Telizität und die Bewegungssemantik als prototypisch organisierte Einflussfaktoren auf die Auxiliarselektion bei *haben* und *sein* ausgemacht werden. Aus diesen Faktoren lassen sich zwei Schemata ableiten: *Sein* ist mit intransitiven Sätzen assoziiert, die telisch sind und/oder Bewegungssemantik aufweisen. *Haben* ist dagegen mit transitiven und atelischen Sätzen sowie Sätzen ohne Bewegungssemantik verknüpft. Abbildung 3.12 fasst die Einflussfaktoren zusammen.

intransitiv	Bewegungsverben mit Akkusativergänzung	transitiv <i>geben</i> <i>schenken</i>
+ telisch <i>verschimmeln</i> <i>verrosten</i>	<i>schimmeln</i> <i>rosten</i>	- telisch <i>lachen</i> <i>weinen</i>
+ Bewegung <i>laufen</i> <i>gehen</i>	<i>schwimmen</i> <i>biken</i>	- Bewegung <i>lachen</i> <i>weinen</i>
<i>sein</i>	<i>sein/haben</i>	<i>haben</i>

Abbildung 3.12: Prototypisch organisierte Selektion von *haben* und *sein*

Die Prototypizität der einzelnen Einflussfaktoren eröffnet zahlreiche Variationsmöglichkeiten: So führt Telizität bei *degree achievements* (*schimmeln*, *rosten*)

⁸⁹Bei hochfrequenten Bewegungsverben ist jedoch von einer Lexikalisierung auszugehen: Auch in idiomatisierten Phrasen ohne Bewegungssemantik (*Das ist dumm gelaufen*), wird *sein* selektiert (Gillmann 2016: 303).

3.3 Variation in der Selektion von *haben* und *sein*

zu Schwankungen, da diese telisch und atelisch gebraucht werden können. Hier kann erst der Kontext desambiguieren: *Die Dose hat lange vor sich hin gerostet* (atelisch, *haben*-Selektion) vs. *Die Dose ist vor Jahren von oben bis unten gerostet* (telisch, *sein*-Selektion). Zudem kann es hinsichtlich der Bewegungssemantik zu Variationsfällen kommen: Diese ergeben sich entweder durch Ambiguitäten in der Transitivität (Sätze mit Akkusativergänzung) oder in der Bewegungssemantik selbst (Interpretation als Sportart oder Bewegung).

Zusätzlich zur Prototypizität nimmt die Tokenfrequenz Einfluss auf Variation. So zeigen frequente *degree achievements* eine klare Präferenz für ein Auxiliar, während infrequente Schwankungen aufweisen. Auch bei Bewegungsverben führt hohe Tokenfrequenz zu Stabilität. Durch die Tokenfrequenz sind die Auxiliare jeweils mental gefestigt und stechen die konkurrierenden Formen daher aus. Da *haben* und *sein* mit unterschiedlichen Funktionen assoziiert sind und daher nicht direkt miteinander konkurrieren, spielt die Typenfrequenz der Schemata nur eingeschränkt eine Rolle. Nur bei ambigen Verwendungen könnte *haben* aufgrund der höheren Typenfrequenz als Default greifen.

Das hier entwickelte Modell weist Ausnahmen auf: So selegieren bspw. *beginnen*, *ab-* und *zunehmen* das Auxiliar *haben*, obwohl es sich um telische Verben handelt (Gillmann 2016: 111). Außerdem zeigen sich regionale Unterschiede in der Verteilung der Auxiliare. Diese stellt Gillmann (2016: 262–264) bei *altern* fest, welches nur in der Schweiz *haben* selegiert. Auch bei anderen *degree achievements* tendiert der Schweizer Sprachraum eher zu *haben* als Deutschland und Österreich. Zudem sind bei den Positionsverben regionale Differenzen festzustellen. Während Positionsverben im Norden *haben* selegieren, selegieren sie im Süden (südlich der Mainlinie) *sein* (Gillmann 2011: 211–212, Duden 2016: 908).⁹⁰

Dennoch lässt sich das Modell nutzen, um Schwankungsfälle zu prognostizieren. Bei den prototypischen Funktionen, die mit *haben* und *sein* assoziiert sind, ist Variation unwahrscheinlich. Dies konnte in Bezug auf Telizität bereits durch die EEG-Studie von Roehm u. a. (2013) psycholinguistisch nachvollzogen werden: *Haben* ist bei telischen Verben unerwartet und umgekehrt *sein* bei atelischen Verben. Dies zeigt sich in der Prozessierung durch N400- und P600-Effekte. Variation ist bezüglich der Telizität somit nur bei den *degree achievements* wahrscheinlich,

⁹⁰Positionsverben waren im Mhd. in ihrer Bedeutung ambig, da sie sowohl auf die Position als auch auf die Bewegung hin zu dieser Position referieren konnten. Je nachdem, ob die Position (atelisch) oder die Bewegung zur Position (telisch) gemeint war, wurde *haben* (atelisch) oder *sein* (telisch) selegiert. Die Interpretation als telisch wurde durch reflexive Verben (*sich setzen*) verdrängt, sodass *sitzen* nur noch auf die Position referieren kann (Gillmann 2011: 211–212). Die areale Distribution der Auxiliarselektion bei Positionsverben ist das einzige Überbleibsel der ursprünglich funktionalen Verteilung.

3 Variation in Konjugation, Deklination und Auxiliarselektion

wenn der Kontext nicht desambiguiert, ob eine telische oder atelische Verwendung vorliegt. Auch dies deckt sich mit den Ergebnissen von Roehm u. a. (2013). Bei frequenten *degree achievements* (*altern*) besteht jedoch die Möglichkeit, dass diese bereits ein Auxiliar lexikalisiert haben. In Bezug auf *altern* wäre somit davon auszugehen, dass *haben* generell Probleme in der Prozessierung hervorruft. Für frequente *degree achievements* ist eine funktionale Verteilung jedoch nicht ausgeschlossen, wie das Verb *schwitzen* in der Untersuchung von Gillmann (2016: 256–257) zeigt.

Hinsichtlich der Transitivität ergibt sich für Bewegungsverben ein Schwankungspotential, da Bewegungsverben wie *fahren* sowohl transitiv (*Ich fahre die Kinder zur Schule*) als auch intransitiv (*Ich fahre zur Arbeit*) gebraucht werden können. Dabei scheint vor allem der unklare Status der Akkusativergänzung als Patiens oder Instrument Variation zu verursachen (*Ich habe/bin ein Auto gefahren*). Dies ist bei Sätzen der Fall, in denen Konkreta (genauer: Fortbewegungsmittel) bzw. beim Verb *reiten* Tiere als Objekt fungieren. Nur wenn ein Direktional (*nach X, bis X*) genutzt wird, ist eine Interpretation als Instrument ausgeschlossen (*Mathias hat den Wagen nach Hamburg gefahren*) und daher nur *haben*-Selektion zu erwarten und auch zu beobachten (Gillmann 2016: 284–286).

Zusätzlich nimmt Objektinkorporation Einfluss auf Transitivität und damit auf die Auxiliarselektion: Bei Sätzen mit Objektinkorporation (*Auto fahren*) liegt für das Verb *fahren* beinahe ausnahmslos (98,5 %) *sein*-Selektion vor (Gillmann 2016: 290). Bei in die Verbalhandlung inkorporierten Abstrakta (*Weltrekord laufen*) liegt die *sein*-Selektion sogar bei 100 % (Gillmann 2016: 294). Hinsichtlich der Auxiliarselektion bei Abstrakta überwiegt zwar generell *sein*, da die Akkusativergänzungen keine Objekte, sondern Adverbiale darstellen, jedoch scheint auch hier die Determiniertheit Einfluss auf die Auxiliarselektion zu nehmen sowie die Semantik der Abstrakta. Auch bei Abstrakta ist daher Variation möglich, allerdings im geringeren Maße als bei den Konkreta, bei denen die *sein*-Selektion generell überwiegt.

Aufgrund der Bewegungssemantik ergeben sich Schwankungsfälle, wenn das Verb nicht nur auf eine Bewegung sondern zusätzlich auf eine Sportart referiert (*schwimmen, reiten*). Dies scheint jedoch mit niedriger Frequenz gepaart zu sein. So könnte auch *laufen* als Sportart begriffen werden (*Ich gehe später noch laufen*), dennoch würde man von einem *r* früheren Marathonläufer *in* nicht sagen, dass er *sie* früher *gelaufen hat*. Zusätzlich sind Schwankungen bei Verben möglich, die eine Bewegung nur implizieren (*humpeln, tanzen*). Sowohl bei den ambigen Bewegungsverben als auch bei Bewegung implizierenden Verben sind Schwankungen nur zu beobachten, wenn der Kontext zwei Lesarten (Bewegung, Nicht-Bewegung) zulässt. Sobald der Kontext desambiguiert wird, ist dagegen Stabilität

3.3 Variation in der Selektion von *haben* und *sein*

in der Auxiliarselektion zu erwarten. Zudem ist davon auszugehen, dass die Auxiliarwahl selbst desambiguieren kann: Wird *sein* gewählt ist eine Bewegungssemantik intendiert, bei *haben* hingegen nicht.

Im empirischen Teil der Arbeit wird der Einfluss der Prototypizität auf die Selektion von *haben* und *sein* mithilfe einer *sentence maze task* untersucht (siehe Abschnitt 4.1.3 für ausführliche Erläuterungen zu *sentence maze tasks*). In der Studie werden Sätze mit Bewegungsverben genutzt, denen ein unterschiedlicher Transitivitätsgrad zukommt: Die Sätze enthalten entweder ein Objekt, das als Patiens dient (*die Kinder zur Schule fahren*, prototypisch transitiv), oder kein Objekt (*zur Kur fahren*, prototypisch intransitiv). Zudem wird mit ambigen Sätzen gearbeitet, die entweder ein Objekt, das als Patiens oder Instrument interpretiert werden kann (*ein Cabrio fahren*), oder eine Akkusativergänzung enthalten, die als Adverbial interpretiert werden kann (*die Strecke fahren*). Die Sätze sind als Nebensätze konzipiert, sodass das Auxiliar am Ende des Satzes steht. Die Proband_innen wählen in der Studie stets zwischen den Auxiliaren *haben* und *sein*. Dabei werden geringe Reaktionszeiten und ein klar verteiltes Antwortverhalten (*haben* für transitiv, *sein* für intransitiv) in den prototypisch transitiven und prototypisch intransitiven Sätzen erwartet, aber erhöhte Reaktionszeiten und Variation im Antwortverhalten in den ambigen Sätzen. Abschnitt 4.3 erläutert die der Studie zugrunde liegenden Hypothesen und das methodische Vorgehen genauer, die Ergebnisse werden in Abschnitt 5.3 vorgestellt. Das folgende Kapitel führt zunächst in die Verfahren zur Reaktionszeitmessung ein, die im empirischen Teil der Arbeit genutzt werden. Im Anschluss daran werden die Hypothesen, die den einzelnen Studien zugrunde liegen, erläutert und das methodische Vorgehen in den einzelnen Studien vorgestellt.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

Der Einfluss von Frequenz, Prototypizität und Form-Schematizität auf Variation wird anhand von Reaktionszeiten untersucht, indem überprüft wird, inwiefern eine Manipulation der Einflussfaktoren Reaktionszeiten verändert. In diesem Kapitel wird das methodische Vorgehen für die einzelnen Reaktionszeitstudien vorgestellt. Hierfür wird in Abschnitt 4.1 zunächst allgemein auf die Verfahren eingegangen, mithilfe derer die Reaktionszeiten in den Studien erhoben wurden. Im Anschluss daran wird das methodische Vorgehen in den Studien zum Einfluss der Frequenz (Abschnitt 4.2), der Prototypizität (Abschnitt 4.3) und der Form-Schematizität (Abschnitt 4.4 und 4.5) auf Reaktionszeiten beschrieben. Der Einfluss der Frequenz wird mithilfe von starken Verben untersucht, der Einfluss der Prototypizität mithilfe der Selektion von *haben* und *sein* und der Einfluss der Form-Schematizität mithilfe des Form-Schemas schwacher Maskulina. Mit dem Einfluss der Form-Schematizität wird auch der Einfluss der Prototypizität überprüft, da Form-Schemata prototypisch organisiert sind (siehe Abschnitt 2.3.2).

4.1 Reaktionsmessverfahren

In den hier vorgestellten Studien wird mit drei verschiedenen Verfahren gearbeitet, um Reaktionszeiten zu evozieren: *lexical decision*, *self-paced reading* und *sentence maze tasks*. Der Einfluss der Frequenz auf Reaktionszeiten wird anhand einer *lexical decision task* überprüft, der Einfluss der Prototypizität anhand einer *sentence maze task*. Der Einfluss der Form-Schematizität wird mithilfe aller hier vorgestellten Verfahren überprüft. Im Folgenden werden die Verfahren näher erläutert.

4.1.1 *Lexical decision tasks*

In *lexical decision tasks* werden den Proband_innen Wörter und Pseudowörter präsentiert. Sie erhalten die Aufgabe, möglichst schnell und präzise zu entscheiden, ob der Stimulus ein Wort ist oder nicht (Baayen 2014: 95). Die Stimuli werden

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

visuell an einem Bildschirm oder auditiv über Lautsprecher präsentiert. Vor den visuellen Stimuli erscheint i. d. R. ein Fixationskreuz, um die Aufmerksamkeit der Proband_innen auf die Mitte des Bildschirms zu lenken. Dabei wird die Reaktionszeit zwischen Stimuluspräsentation und Antwort gemessen und die Antworten der Proband_innen werden gespeichert. Auf diese Weise kann ermittelt werden, wie präzise die Proband_innen antworten, d. h. wie groß der Anteil der Stimuli ist, die korrekt als Pseudowort bzw. korrekt als existierendes Wort eingeschätzt wurden. Die Reaktionszeiten geben bspw. Aufschluss über Worterkennungseffekte: So hängt die Schnelligkeit, mit der Pseudowörter abgelehnt werden, davon ab, wie ähnlich sie real existierenden Wörtern sind (Barca & Pezzulo 2012).

Die Reaktionszeiten aus *lexical decision tasks* wurden meines Wissens noch nicht zur Untersuchung von Variation genutzt. Sie eignen sich aber dafür, da Wortformen wie bspw. **gezieht* vs. *gezogen* und *geglimmt* vs. *geglommen* gezielt kontrastiert werden können. Zudem weisen bisherige Studien darauf hin, dass der Einfluss von Frequenz und Form-Schematizität auf Reaktionszeiten durch *lexical decision tasks* erfasst werden kann. Frequente Wörter werden in *lexical decision tasks* schneller erkannt als infrequente (Rubenstein u. a. 1970: 489–491, Whaley 1978: 150–152, siehe hierzu auch Abschnitt 2.1.2). Zudem sind Proband_innen in *lexical decision tasks* für phonologische Eigenschaften der Stimuli sensibel (Keuleers u. a. 2010: 6–7), sodass auch Form-Schematizität¹ mithilfe der Methode erfassbar zu sein scheint.

Die *lexical decision task* ist eine der am häufigsten genutzten Methoden in der Psycholinguistik (Diependaele u. a. 2012: 1, Baayen 2014: 95). Dies lässt sich vorrangig mit der Praktikabilität der Methode erklären: Sie ist leicht zu implementieren und durchzuführen (Baayen 2014: 96). Die erhobenen Daten sind zudem gut zu interpretieren und die Aufgabe kann leicht vermittelt werden.²

Baayen (2014: 96) weist darauf hin, dass *lexical decision tasks* keinen Einblick in den zeitlichen Verlauf von Prozessierung geben können, da sie nur die späten Phasen der Prozessierung messen. Ein weiterer Nachteil ist, dass Stimuli kontextlos präsentiert werden (Baayen 2014: 96). Zudem wird die Beurteilungskomponente, also die Beurteilung der Bekanntheit des Stimulus, als problematisch gesehen (Haberlandt 1994: 30–31). Die Kritik fußt darauf, dass sowohl die Kontextlosigkeit als auch die Beurteilung einzelner Wörter kein Pendant im alltägli-

¹Der Einfluss der Prototypizität wird mithilfe eines syntaktischen Variationsphänomens überprüft, daher eignet sich die *lexical decision task* hierfür nicht. Da Prototypizität aber auch innerhalb der Form-Schematizität eine Rolle spielt, ist der Einflussfaktor nicht völlig ausgeklammert.

²Außerdem bieten sich *lexical decision tasks* für Priming-Experimente an, weil vor den Stimulus ein Prime geschaltet werden kann (Jones & Estes 2012: 46).

chen Sprachgebrauch haben. Dies führt laut Baayen (2014: 96) bspw. dazu, dass die Reaktionszeiten in lexical-decision-Studien und die Länge der Fixationszeiten in Eyetracking-Studien zwar korrelieren, aber diese Korrelation weniger stark ausfällt als die Korrelation der Reaktionszeiten zwischen lexical-decision- und word-naming-Studien, die ebenfalls kontextlos sind (Kuperman u. a. 2013: 21).³

Ein weiterer Kritikpunkt an *lexical decision tasks* ist die Annahme eines konsistenten Antwortverhaltens einzelner Personen. Ein Stimulus, der abgelehnt wurde, sollte also auch bei mehrmaligem Bewerten abgelehnt werden. Dies ist jedoch laut Diependaele u. a. (2012) nicht der Fall: Bewerten Proband_innen die Stimuli in einer Aufgabe doppelt, ist die Bewertung nicht zwingend konsistent (Diependaele u. a. 2012: 5–8). Dies gilt insbesondere für Substantive mit geringer Frequenz. Diependaele u. a. (2012: 8) konkludieren, dass *lexical decision* eine „noisy task“ ist, wenn man von einem konstanten Antwortverhalten ausgeht. Sie empfehlen, sicherzustellen, dass die Proband_innen die real existierenden Wörter kennen und Antworten mit Vorsicht zu behandeln, die aus einer Kondition stammen, bei der mehr als 10 % der Stimuli falsch beurteilt wurden.

Baayen (2014: 117) sieht wegen der genannten Kritikpunkte *lexical decision tasks* eher als metalinguistische Beurteilung denn als Einblick in die Prozessierung. Wie er selbst schreibt, erlauben sie aber durchaus Rückschlüsse auf Prozessierung (Baayen 2014: 96). Die Schnelligkeit, mit der die Aufgabe erfüllt wird, gibt Aufschluss darüber, wie schnell die mentale Repräsentation von Wörtern aktiviert werden kann. Da die Reaktionszeit durch isoliert präsentierte Stimuli evoziert wird, ist der gewonnene Einblick in die Prozessierung zwar weniger detailliert als bei anderen Methoden, dennoch kann er genutzt werden, um Rückschlüsse auf Prozessierungsunterschiede zwischen Stimuli zu ziehen. Man sollte sich bei der Datenauswertung über die eingeschränkte Interpretierbarkeit der Reaktionszeiten in *lexical decision tasks* bewusst sein. Allerdings ist diese Notwendigkeit nicht auf *lexical decision tasks* beschränkt, sondern gilt für jede Methode.

Generell kritisiert Baayen (2014: 99) die Auswahl der Stimuli für psycholinguistische Experimente: Wenn bspw. Testwörter allein anhand ihrer Tokenfrequenz in verschiedene Gruppen geteilt werden, kann dies problematisch sein, da

³ Ein weiteres Problem kann sich durch die Konstruktion der Pseudowortstimuli ergeben: Pseudowörter, die mit einem komplexen Verfahren erstellt wurden (bspw. durch Rekombinieren von Silben) erzielen in einer Studie von Keuleers u. a. (2010: 6–7) geringere Reaktionszeiten als Pseudowörter, die durch Änderung eines Buchstabens erstellt wurden, da das Rekombinieren von Silben zu Pseudowörtern führt, die real existierenden Wörtern ähnlich sind. Ähnliche Effekte wurden bereits in Abschnitt 2.1.2 diskutiert: Sie weisen auf exemplarbasierte Cluster und damit auf eine generelle Eigenschaft der Prozessierung von Sprache hin und können daher m. E. nicht als spezielles Problem der Methode gesehen werden. Zudem kann der Einfluss durch ein gutes Design der Stimuli minimiert werden.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

Einflussfaktoren ignoriert werden, die mit Frequenz einhergehen. So haben frequente Stimuli i. d. R. breitere Verwendungskontexte und eine breitere Semantik als infrequente (siehe Abschnitt 2.1.2 für weitere Erläuterungen). Zudem ist es aus statistischer Sicht problematisch, dass durch die Auswahl der Stimuli keine zufällige Stichprobe gezogen wird (Baayen 2014: 99). Dieses Problem stellt sich jedoch nicht nur für *lexical decision tasks*, sondern generell für Designs von Experimenten.

Der Einfluss von Frequenz und Form-Schematizität wird mithilfe einer *lexical decision task* überprüft (siehe Abschnitt 4.2 und 4.5 für Erläuterungen zum Studiendesign). Die beschriebenen Probleme der *lexical decision task* gelten nur eingeschränkt für die Studien. In den Studien werden jeweils zwei korrespondierende Formen gegenübergestellt (siehe Abschnitt 4.2 sowie 4.4), bspw. starke und schwache Partizip-II-Formen von *ziehen*. Zwar wird nicht ein und derselbe Stimulus wiederholt, dennoch lässt sich überprüfen, inwiefern die Antworten konsistent sind (siehe Abschnitt 5.2.1 und 5.4.1.2). Dabei wird angenommen, dass die Antworten bei Varianten weniger konsistent sind als bei Wörtern, bei denen nur eine Form mental gefestigt ist (siehe Abschnitt 4.2.1 sowie 4.5.1). Zudem werden Proband_innen ausgeschlossen, die mehr als 10 % der Stimuli, die die Konzentration der Proband_innen evaluieren, falsch beantworten. Die Kontextlosigkeit bleibt eine Einschränkung des Studiendesigns. Um die Validität der Ergebnisse überprüfen zu können, wird der Einfluss der Form-Schematizität nicht nur durch eine *lexical decision task*, sondern auch durch eine *self-paced reading-* und eine *sentence maze task* überprüft.

4.1.2 *Self-paced reading tasks*

Self-paced reading tasks werden genutzt, um Lesezeiten zu evozieren: Bei dieser Methode ist der zu lesende Text zunächst verdeckt (bspw. durch Rauten repräsentiert) und wird erst per Tastendruck von Proband_innen Stück für Stück aufgedeckt (McDonough & Trofimovich 2012: 118). Wie *lexical decision tasks* wurde auch *self-paced reading* meines Wissens noch nicht genutzt, um Variation zu untersuchen. Durch die Segmentierung erlaubt es die Methode, die Lesezeiten von Wortformen (z. B. *des Kollegen* und **des Kolleges*) zu vergleichen. Auf diese Weise kann überprüft werden, ob Reaktionszeitunterschiede zwischen Varianten und Formen, die keine Varianten darstellen, bestehen: Überraschende Formen werden langsamer gelesen als im Kontext erwartbare, siehe bspw. Tabor & Galantucci (2000) mit einer Studie zu Holzwegsätzen. Bei einer nicht zu erwartenden Form wie **des Kolleges* sind also höhere Reaktionszeiten zu erwarten als für *des*

Kollegen. Stellen die beiden Wortformen hingegen Varianten dar (z.B. Genitivvarianten wie bei *des Graf's/des Grafen*), ist davon auszugehen, dass sich keine Unterschiede in der Lesezeit ergeben: Hier koexistieren zwei Formen mit gleicher Funktion, somit sind beide Formen im Kontext erwartbar.

Self-paced reading tasks wurden in den 1970er Jahren entwickelt und sind inzwischen in der Psycholinguistik etabliert (Mitchel 2013: 18–19, Jegerski 2014: 20–21). Sie können auf verschiedene Weisen gestaltet sein (für einen Überblick siehe McDonough & Trofimovich 2012: 118–119, Mitchel 2013: 18–19). In der vorliegenden Studie wird mit einem *moving window*, d. h. mit einem sich bewegenden Sichtfenster, gearbeitet: Der zu lesende Text ist zunächst durch Rauten verdeckt und wird durch Tastendruck Stück für Stück aufgedeckt. Ein aufgedecktes Element bleibt aber nicht die ganze Zeit sichtbar, sondern wird beim nächsten Tastendruck wieder verdeckt. Beispiel 1 gibt einen Eindruck in das Lesen von *self-paced reading* mit *moving window*:

- (1) #####
 1. Tastendruck Die Katzen #####
 2. Tastendruck ##### sitzen #####
 3. Tastendruck ##### auf dem Baum.

Die Zeit zwischen zwei Tastaturanschlägen wird gemessen. Da während des Lesens jeweils nur ein Element aufgedeckt ist, wird die Zeit zwischen zwei Tastaturanschlägen mit der Lesezeit des jeweiligen Elements gleichgesetzt (McDonough & Trofimovich 2012: 118, Jegerski 2014: 21–22).⁴ Um sicherzustellen, dass Proband_innen den Text konzentriert lesen, werden in regelmäßigen Abständen Verständnisfragen gestellt (Forster 2010: 349).

Die Methode erlaubt es, die Lesezeiten verschiedener sprachlicher Strukturen zu messen und miteinander zu vergleichen. Sie ist zudem kostengünstig, effizient, leicht zu implementieren und nicht an einen bestimmten Raum gebunden, da außer einem Computer (und ggf. einem transportablen Response Pad, das eine genauere Reaktionszeitauflösung ermöglicht als herkömmliche Tastaturen) keine technischen Anforderungen gestellt werden (Mitchel 2013: 24, Jegerski 2014: 43). Wie Keller u. a. (2009) zeigen, kann *self-paced reading* auch ohne Verzerrungen der Ergebnisse online implementiert werden.

Der Nachteil der Methode ist, dass sie stark in die Lesegewohnheiten von Proband_innen eingreift: Anders als bei Eyetracking-Studien müssen Proband_innen vorgefertigte Einheiten lesen. Zudem erlauben *self-paced reading tasks* kei-

⁴Diese Annahme ist problematisch (Forster 2010: 349), wie sich auch in den späteren Ausführungen dieses Abschnitts zeigen wird.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

ne Regressionen (Jegerski 2014: 44).⁵ Aufgrund dieser Unnatürlichkeit könnte *self-paced reading* einen zusätzlichen Prozessierungsaufwand generieren (Jegerski 2014: 44). Für diese Hypothese spricht, dass die gemessenen Lesezeiten bei *self-paced reading* verlangsamt sind (Mitchel 2013: 25). Die Methode ist daher nur für geübte Leser_innen geeignet (Jegerski 2014: 43).

Witzel u. a. (2012: 107–108) sehen einen Vorteil in der Segmentierung in *self-paced reading tasks*: Hierdurch werden individuelle Lesestrategien weitgehend blockiert, weswegen standardisiertere Ergebnisse als bei Eyetracking-Studien erzielt werden können. Allerdings weisen Witzel u. a. (2012: 107–108) darauf hin, dass auch bei *self-paced reading* individuelle Lesestrategien zum Tragen kommen können wie bspw. *tapping*, bei dem Proband_innen mit konstant geringer Reaktionszeit durch die Elemente klicken (Forster 2010: 349). Die Segmentierung selbst kann aber auch Einfluss auf die Lesezeit nehmen: So ist davon auszugehen, dass Segmente, die keine Phrasen darstellen wie bspw. *am Eingang der*, schwieriger zu prozessieren sind und somit höhere Lesezeiten evozieren als Segmente, die Phrasen darstellen wie *am Eingang der Universität*. Mitchel (2013: 25–26) konstatiert allerdings, dass ihm keine Fälle bekannt seien, in denen die gemessenen Effekte eindeutig allein auf die Segmentierung innerhalb des Experiments zurückzuführen sind.

Forster (2010: 349) problematisiert die Annahme, dass Tastaturanschlag und Verarbeitung synchron verlaufen: Proband_innen von *self-paced reading tasks* berichten bspw., dass sie in einem konstanten Tempo durch das Experiment klicken und bei Prozessierungsproblemen das Tempo verringern. Die Anpassung des Tempos könnte somit erst nach dem Element passieren, das die Probleme aufgelöst hat. Aber auch unabhängig vom individuellen Verhalten von Proband_innen können spill-over-Effekte entstehen: Die Prozessierung eines Element ist nicht zwingend abgeschlossen, wenn das nächste Element aufgedeckt wird. Somit kann es sein, dass erhöhte Lesezeiten erst nach dem Element auftreten, das die erhöhten Lesezeiten verursacht hat (Mitchel 2013: 24–25). Spill-over-Effekte sind jedoch nicht nur in *self-paced-reading*-Studien zu beobachten, sondern auch in Eyetracking- und EKP-Studien. Mitchel (2013: 25) sieht sie daher als Teil der Satzprozessierung. In *self-paced-reading*-Studien können spill-over-Effekte verringert werden, indem die Länge der Elemente vor den Testitems konstant gehalten wird und Wortgruppen statt einzelner Wörter angezeigt werden (Mitchel 2013: 25).

⁵Neben diesen Eigenschaften wurde das konstante Tastaturanschlagen der Proband_innen als unnatürlicher Faktor diskutiert (Jegerski 2014: 44). Dieser Faktor wurde insbesondere in Kontrast zu gedruckten Texten stark gemacht und ist nach Jegerski (2014: 44) heute zu vernachlässigen, da Lesen an Bildschirmen an Normalität gewonnen hat.

Mitchel (2013: 25) weist darauf hin, dass die Ergebnisse von *self-paced-reading*-Studien in *Eye-tracking*-Studien repliziert werden können. Die Methode scheint also generell geeignet zu sein, um Unterschiede in Lesezeiten festzustellen (Mitchel 2013: 25). Auch Witzel u. a. (2012: 123–126) zeigen, dass *self-paced reading tasks* und *Eye-tracking*-Studien generell ähnliche Ergebnisse erzielen: In ihrer Studie vergleichen sie unter anderem die Ergebnisse von *self-paced reading tasks* und *eye-tracking* anhand von drei verschiedenen syntaktischen Phänomene, bei denen Prozessierungsschwierigkeiten zu erwarten sind. Dies waren bspw. Ambiguitäten, die durch Koordination ausgelöst und erst im Verlauf des Satzes aufgelöst werden: Im Satz *The robber shot the jeweler and the salesman reported the crime to the police* könnte *the salesmann* ein koordiniertes Objekt oder Subjekt eines koordinierten Satzes sein. Die Ambiguität wird erst bei *reported* aufgelöst (Witzel u. a. 2012: 111). Dieser Effekt wird vermieden, wenn ein Komma vor *and* gesetzt wird (*The robber shot the jeweler, and the salesman reported the crime to the police*). Es ist daher zu erwarten, dass bei der ambigen Struktur längere Reaktions- bzw. Fixationszeiten zu messen sind als bei dem desambiguierten Satz. Dies war nur bei *eye-tracking* zu beobachten, die Sätze mit und ohne Ambiguität lösten bei *self-paced reading* aber keine Unterschiede in der Lesezeit aus. Die anderen in der Studie untersuchten syntaktischen Phänomene⁶ zeigten hingegen sowohl bei *eye-tracking* als auch bei *self-paced reading* dieselben Ergebnisse.

Der Einfluss der Form-Schematizität auf Reaktionszeiten soll mithilfe einer *self-paced reading task* überprüft werden (siehe Abschnitt 4.4). Hierfür sind die Nachteile von *self-paced reading* m. E. nicht schwerwiegend: Nur Lesezeiten einzelner Items sind relevant, sodass davon auszugehen ist, dass der Eingriff in die Lesegewohnheiten keine großen Auswirkungen auf die Ergebnisse hat. Zudem werden Lesezeiten miteinander verglichen, die in demselben unnatürlichen Setting gemessen wurden. Eventuelle Einflüsse des Settings lassen sich also bei jedem gemessenen Item finden, sodass Unterschiede zwischen den Items nicht auf das Untersuchungsdesign zurückzuführen sein sollten. Dies bedeutet, dass die Lesezeiten nur in Bezug auf eine relativ unnatürliche Lesesituation interpretiert

⁶Dies waren Sätze, die Relativsätze und Adverbiale mit ambigem Skopus enthielten. Der Relativsatz kann sich, wenn er eingeleitet wird, entweder auf das Subjekt oder auf das Attribut zum Subjekt beziehen (*The son of the actress who shot herself/himself on the set was under investigation*), die Ambiguität wird dann durch das Reflexivpronomen *himself/herself* aufgelöst. Das Adverbial kann sich auf zwei Verbalphrasen im Satz beziehen, hierbei desambiguiert das Adverb durch den Zeitbezug (*Susan bought the wine she will drink next week/this week, but she didn't buy any cheese*). Geringere Reaktions- und Fixationszeiten werden dabei von dem Relativsatz *who shot himself* mit Bezug auf das Subjekt *the son of the actress* bzw. von dem Adverbial *next week* mit Bezug auf die zweite Verbalphrase *she will drink* ausgelöst (Witzel u. a. 2012: 115–120).

werden können. Spill-over-Effekte werden vermieden, indem der Text in Wortgruppen statt einzelne Wörter segmentiert wird. Der Gefahr einer Verzerrung der Lesezeiten durch Segmentierung wird begegnet, indem eine möglichst unauffällige Einteilung der Sätze vorgenommen wird, die sich an Phrasenstrukturen orientiert. Die Aufmerksamkeit der Proband_innen wird über Fragen zum Aufbau der Wörter sichergestellt (z. B. Haben Sie *des Schettosen* oder *des Schettoses* gelesen?, siehe hierzu genauer Abschnitt 4.4). Um die Ergebnisse der *self-paced reading task* evaluieren zu können, wird der Effekt der Form-Schematizität zusätzlich mithilfe einer *lexical decision* und einer *sentence maze task* überprüft.

4.1.3 Sentence maze tasks

Sentence maze tasks lassen sich als eine Mischung aus *self-paced reading* und *lexical decision tasks* betrachten: Wie in einer *self-paced reading task* lesen Proband_innen Sätze Stück für Stück. Wie in einer *lexical decision task* bewerten Proband_innen Stimuli (Forster 2010: 347). Der Anfang der Sätze ist dabei vorgegeben und die Proband_innen haben im Anschluss daran die Wahl zwischen zwei Möglichkeiten, siehe Beispiel 2, das an ein Beispiel von Forster u. a. (2009: 163–164) angelehnt ist.

- (2) Kein ...
- a. Mensch geht
 - b. Tür ist
 - c. illegal wackelt

Nur eine der Kombinationsmöglichkeiten in Beispiel 2 führt zu einem grammatischen Satz: *Kein Mensch ist illegal*. Die anderen Kombinationsmöglichkeiten sind hingegen ungrammatisch (z. B. **Kein geht Tür illegal/*Kein Mensch ist wackelt*). Es handelt sich somit um eine *grammaticality maze (G-maze) task*, da die Alternativen zu den präferierten Möglichkeiten jeweils zu einem ungrammatischen Satz führen.

Wie die anderen Verfahren zur Reaktionszeitmessung wurden auch *sentence maze tasks* meines Wissens bislang noch nicht zur Erforschung von Variation genutzt. Dabei bietet sich die Methode dafür an, da Proband_innen zwischen zwei Formen wählen müssen. So können die Reaktionszeiten für die Wahl zwischen Formen, die Varianten darstellen (*Auto gefahren sein/haben*), und Formen, bei denen dies nicht der Fall ist (*zur Kur gefahren sein/*haben*), gegenüber gestellt werden.

Neben *G-maze tasks* existieren auch *lexicality maze (L-maze) tasks*, bei denen die Wahl zwischen real existierenden und Pseudowörtern besteht (Witzel u. a. 2012: 108). Die *G-maze task* wird dabei als die komplexere gesehen, da Proband_innen den Satz inkrementell prozessieren müssen, um sie zu lösen. Das ist bei einer *L-maze task* nicht der Fall, da Proband_innen jeweils unabhängig von den anderen Stimuli entscheiden können, ob sie den zu bewertenden Stimulus kennen (Witzel u. a. 2012: 109).

Anders als in *self-paced reading tasks* wird in *sentence maze tasks* i. d. R. mit einzelnen Wörtern und nicht mit Wortgruppen gearbeitet (Forster u. a. 2009: 163). Ähnlich wie bei *self-paced reading* mit *moving window* sehen die Proband_innen immer nur die Möglichkeiten, zwischen denen sie sich entscheiden sollen. Sie müssen sich daher die Satzteile merken, die sie gewählt haben. Das Abfragen des Textverständnisses entfällt bei *sentence maze tasks*, da die Aufmerksamkeit der Proband_innen über die Wahl der Wörter evaluiert werden kann. Dabei kann der aktuelle Versuch abgebrochen und zum nächsten Satz gesprungen werden, wenn Proband_innen einen Fehler machen (Forster u. a. 2009: 163). Eine Wiederholung des fehlerhaften Versuchs ist dabei nicht vorgesehen.

Ein Vorteil der *sentence maze task* ist, dass Proband_innen zu einem inkrementellen Lesen gezwungen werden (Forster u. a. 2009: 164). Zudem wird der Leseprozess verlangsamt, sodass die Synchronisierung von Tastendruck und Lesegeschwindigkeit erleichtert wird und spill-over-Effekte vermieden werden (Forster u. a. 2009: 164). Ein weiterer Vorteil besteht darin, dass die Aufgabe der Proband_innen klar definiert ist, da sie anders als beim *self-paced reading* nicht weiter klicken, wenn sie meinen, ein Wort verstanden zu haben, sondern aktiv zwischen zwei Möglichkeiten wählen.⁷

An der *sentence maze task* wird kritisiert, dass zwei anstelle von einem Wort prozessiert werden müssen und Reaktionszeiten somit verzerrt sein könnten (Forster 2010: 353). Zudem spiegelt die Methode offensichtlich keine natürliche Lesesituation wider (Forster u. a. 2009: 350, Forster 2010: 164, Witzel u. a. 2012: 107). Witzel u. a. (2012: 108–109) sehen darin aber auch Vorteile: Dadurch, dass die Methode die Proband_innen zwingt, Sätze Wort für Wort zu prozessieren, werden individuelle Lesestrategien verhindert, sodass gezielt gemessen werden kann, welchen Prozessierungsaufwand das Integrieren von Wörtern in einen Satz hat. Zudem berichten Proband_innen von *sentence maze tasks*, dass sie das Gefühl haben, die Sätze relativ natürlich zu lesen (Forster u. a. 2009: 164).

⁷Inwiefern die Reihenfolge der Möglichkeiten (grammatisch-intendierte Form links oder rechts von der ungrammatisch-intendierten Form) eine Rolle in der Auswahl spielt, wurde meines Wissens noch nicht erforscht.

Witzel u. a. (2012) vergleichen die Resultate von *eyetracking*, *self-paced reading* und *sentence maze tasks* anhand drei verschiedener syntaktischer Phänomene (siehe Abschnitt 4.1.2 für weitere Erläuterungen). Die *sentence maze task* zeigt genauso wie die *self-paced reading task* bei zwei der drei Phänomene die erwarteten Ergebnisse: Wie bei der *self-paced reading task* konnten auch in der *sentence maze task* keine Effekte von Ambiguität in der Koordination zwischen Objekten und Sätzen festgestellt werden (Witzel u. a. 2012: 121–124). Forster (2010: 167–170) zeigt, dass die *sentence maze task* sensitiv für Holzwegsatz- und Frequenzeffekte ist (siehe Abschnitt 2.1.3 für eine bayesianische Modellierung der Effekte).

Der Einfluss der Prototypizität und der Form-Schematizität auf Reaktionszeiten wird mithilfe einer *sentence maze task* überprüft (siehe Abschnitt 4.3 und 4.5). Da das Interesse der Studie nicht das Lesen an sich, sondern die Prozessierung einzelner Formen ist, bietet sich diese Methode an, da sie es ermöglicht, die Reaktionszeit für die Wahl verschiedener Formen gezielt zu vergleichen. Dabei wird die *sentence maze task* leicht abgewandelt: Die vorliegende Studie arbeitet nicht mit einzelnen Wörtern, sondern mit Wortgruppen. Bereits für den Satzanfang wählen die Proband_innen aus zwei Möglichkeiten. Die alternativen Möglichkeiten zur im Design präferierten Möglichkeit stellen grammatische und orthographische Abweichungen sowie Anschlüsse dar, die im Kontext semantisch unsinnig sind. Dabei wird neben klaren Abweichungen (**Schuhle*) auch mit Varianten (*Fahrradfahren/Fahrrad fahren*) gearbeitet. Daher führen Fehler nicht zu einem Abbruch des Versuchs. Die Konzentration der Proband_innen wird anhand des Antwortverhaltens bei Filleritems evaluiert. Wie in der lexical-decision-Studie werden Proband_innen ab 10 % falscher Antworten aus der Stichprobe ausgeschlossen. Der Einfluss der Form-Schematizität wird nicht nur in der *sentence maze task*, sondern auch mithilfe einer *lexical decision* und einer *self-paced reading task* überprüft, sodass die verschiedenen Verfahren zur Reaktionszeitmessung miteinander verglichen und evaluiert werden können. In den folgenden Abschnitten wird das methodische Design der einzelnen Studien erläutert.

4.2 Lexical-decision-Studie zu Frequenzeffekten

Der Einfluss der Frequenz auf Variation wird anhand der Variation in der Konjugation mithilfe einer *lexical decision task* untersucht. In die Aufgabe sind sowohl starke als auch schwache Formen starker Verben mit unterschiedlicher Tokenfrequenz eingeflochten. Die Proband_innen beurteilen die Formen stets danach, ob sie sie kennen. Im Folgenden werden zunächst die Fragestellungen und Hypothesen vorgestellt. Anschließend wird das methodische Vorgehen (Materialien, Versuchsdesign und -ablauf, Metadaten, Probandenakquise und Proband_innen) näher erläutert.

4.2.1 Fragestellung und Hypothesen

Die Untersuchung geht der Frage nach, welchen Einfluss Tokenfrequenz auf die Verarbeitung und die Bekanntheit starker und schwacher Partizip-II-Formen von starken Verben hat (zum Einfluss der Tokenfrequenz auf die Konjugation siehe Abschnitt 3.1.1). Zudem wird untersucht, ob schwache Partizip-II-Formen starker Verben anders wahrgenommen und verarbeitet werden, wenn sie bereits nachgewiesenermaßen im Sprachgebrauch genutzt werden. Dies wird überprüft, indem Verben genutzt werden, bei denen Schwankungen in Korpora attestiert bzw. nicht attestiert sind. In der Studie wird daher mit der Unterscheidung zwischen frequenten und infrequenten Verben gearbeitet und innerhalb der infrequenten Verben zwischen attestierter und nicht-attestierter Schwankung unterschieden. Die frequenten Verben weisen aufgrund der hohen Tokenfrequenz allesamt keine attestierte Variation auf (siehe Abschnitt 4.2.2 für Erläuterungen zu den Testverben).

In Hinblick auf Reaktionszeiten und Antwortverhalten der Proband_innen werden folgende Hypothesen aufgestellt:

4.2.1.1 Reaktionszeiten

1. Die Reaktionszeiten für frequente Verben sind niedriger als für infrequente Verben mit und ohne Schwankung in Korpora.
 - 1.1 Die Unterschiede in den Reaktionszeiten lassen sich sowohl für starke als auch für schwache Formen feststellen.
 - 1.2 Die Unterschiede in den Reaktionszeiten lassen sich sowohl für als bekannt bewertete Formen als auch für als unbekannt bewertete Formen feststellen.

Sollten sich die Hypothesen bestätigen, können die Unterschiede in den Reaktionszeiten als ein Hinweis auf Variationspotential gedeutet werden. Zwar ist auch aus rein frequentieller Sicht zu erwarten, dass die Reaktionszeiten bei frequenten Verben kürzer sind als bei infrequenten (Clahsen u. a. 2001), jedoch kann über die schwachen Formen der generelle Einfluss der Frequenz eingegrenzt werden: Die schwachen Partizip-II-Formen von tokenfrequenten Verben (bspw. **gezieht* statt *gezogen*) sind zwar theoretisch mögliche Partizip-II-Formen, die Formen an sich sind aber nicht frequent. Somit kann ein schnelles Ablehnen dieser Formen nicht mit der Frequenz der Formen erklärt werden, da diese praktisch inexistent sind. Ein Unterschied in den Reaktionszeiten zwischen frequenten und infrequenten Verben kann daher als Hinweis auf Variationspotential interpretiert werden:

Nicht nur starke Formen, sondern auch schwache Formen von frequenten starken Verben sollten schnell beurteilt werden, weil eine der Formen stark gefestigt ist. Bei infrequenten Verben sind dagegen unabhängig von Schwankungen in Korpora erhöhte Reaktionszeiten für beide Formen zu erwarten, da die starken Formen nicht so stark gefestigt sind und daher deren statistisches Vorkaufsrecht (Goldberg 2019: 74–94; siehe Abschnitt 2.1.3 für weitere Erläuterungen) nur noch bedingt greifen kann. Erhöhte Reaktionszeiten würden also darauf hindeuten, dass das statistische Vorkaufsrecht bei infrequenten starken Verben nicht mehr vollständig greift, und damit auf Variationspotential hinweisen, da mit dem schwindenden Einfluss des statistischen Vorkaufsrechts für starke Formen schwache Formen möglich werden.

2. Zwischen den infrequenten Verben mit attestierter Schwankung und ohne attestierte Schwankung in Korpora lassen sich keine systematischen Unterschiede in den Reaktionszeiten messen.
 - 2.1 Die Reaktionszeiten bei infrequenten Verben mit und ohne Schwankung werden nicht durch starke oder schwache Verbformen beeinflusst.
 - 2.2 Die Reaktionszeiten bei infrequenten Verben mit und ohne Schwankung werden nicht durch die Bewertung der Formen als bekannt oder unbekannt beeinflusst.

Da die Verben mit und ohne attestierte Schwankung infrequent sind, sind auch die starken Formen dieser Verben seltener und sie können die schwachen Formen daher im Vorkommen nicht so stark dominieren. Das statistische Vorkaufsrecht frequenter Formen kann daher nicht mehr vollständig greifen, weshalb sowohl Verben mit attestierter Schwankung als auch Verben ohne attestierte Schwankung vergleichbare Reaktionszeiten hervorrufen sollten. Diese Hypothese hat sich in der Prästudie bestätigt, die Prästudie ist im digitalen Anhang zu finden. Neben dieser Hypothese lässt sich eine Alternativhypothese aufstellen: Kleine Unterschiede in der Reaktionszeit sind zwischen Verben mit attestierter und nicht-attestierter Schwankung erwartbar, weil schwache Formen bei Verben ohne attestierte Schwankung weniger frequent sind als bei Verben mit attestierter Schwankung. Diese Unterschiede sollten jedoch geringer sein als der Unterschied in den Reaktionszeiten zwischen frequenten und infrequenten Verben. Das Studiendesign erlaubt es, beide Hypothesen zu überprüfen. Sollten keine Unterschiede zwischen Verben mit attestierter und ohne attestierte Schwankung in Korpora gemessen werden, kann das als Hinweis darauf

gedeutet werden, dass die Formen von Verben ohne Schwankung in Korpora ähnlich gefestigt sind wie die Formen von Verben mit Schwankung in Korpora. Vergleichbare Reaktionszeiten würden daher auf ein Variationspotential hinweisen, auch wenn Variation noch nicht in Korpora zu beobachten ist.

4.2.1.2 Antwortverhalten

1. Starke Formen werden bei infrequenten Verben eher als unbekannt bewertet als bei den frequenten. Trotz dieser Unterschiede ist die Bekanntheit der starken Formen bei allen Testverben auf einem hohen Niveau.
2. Die Bekanntheit der schwachen Formen ist von der Tokenfrequenz abhängig: Schwache Formen frequenter Verben sind unbekannt, schwache Formen infrequenter Verben weisen hingegen einen höheren Bekanntheitsgrad auf.
3. Schwache Formen von infrequenten Verben mit attestierter Schwankung in Korpora sind eher bekannt als schwache Formen von Verben ohne attestierte Schwankung in Korpora.

Schwache Formen haben eine größere Wahrscheinlichkeit bekannt zu sein, wenn sie bereits in Gebrauch, d. h. in einem Korpus attestiert sind. Auch wenn Korpora nur einen Ausschnitt der Sprachwirklichkeit abbilden, die Sprecher_innen rezipieren, ist bei Formen, die bereits in einem Korpus attestiert sind, die Wahrscheinlichkeit groß, dass Proband_innen die Form schon einmal rezipiert oder sogar selbst gebildet haben. Bei Verben, bei denen die schwache Form (noch) nicht attestiert ist, ist die Wahrscheinlichkeit hingegen geringer, dass die schwache Form bereits rezipiert wurde. Die Ergebnisse der Prästudie stützen diese Hypothese. Die Prästudie ist im digitalen Anhang zu finden. Wie bei den Reaktionszeiten lässt sich auch in Bezug auf das Antwortverhalten eine alternative Hypothese aufstellen: Da die Verben gleichermaßen infrequent sind, könnten die starken und schwachen Partizip-II-Formen der Verben unabhängig von ihrem attestierten Vorkommen in Korpora ähnlich bewertet werden. Das Design erlaubt es, beide Hypothesen zu überprüfen.

4. Für frequente Verben ist ein konsistentes Antwortverhalten zu beobachten (überwiegend Unbekanntheit für schwache Formen, überwiegend Bekanntheit für starke Formen), bei infrequenten Verben sind auch inkonsistente Antworten zu finden (Bekanntheit/Unbekanntheit für beide Formen).

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

Die hier vorgestellten Hypothesen sowie das methodische Vorgehen inklusive der geplanten statistischen Auswertung wurden vor der Datenerhebung registriert. Die Registrierung lässt sich unter <https://osf.io/ynvb5> aufrufen. Die statistische Auswertung wird in Abschnitt 5.2 erläutert, in dem auch die Ergebnisse vorgestellt werden.

4.2.2 Materialien

Für die Studie werden Partizip-II-Formen starker Verben mit unterschiedlicher Frequenz genutzt. Der Einfluss der Frequenz auf starke Verben wird anhand von Partizip-II-Formen getestet, da alle starken Verben – anders als bei der Wechselflexion im Präsens und der Imperativhebung – starke Partizip-II-Formen aufweisen und das Partizip II frequenter ist als das Präteritum (Fischer 2018: 161).

Die Ermittlung der Frequenz erfolgte anhand der Wortverlaufskurven im Digitalen Wörterbuch der Deutschen Sprache (DWDS, Berlin-Brandenburgische Akademie der Wissenschaften 2019). Basis der Wortverlaufskurven ist das DWDS-Zeitungskorpus. Um die Unterscheidung zwischen Verben mit attestierter und ohne attestierte Schwankung zu operationalisieren, wurde die Ratio zwischen starken und schwachen Partizip-II-Formen im Deutschen Referenzkorpus ermittelt (DeReKo, Leibniz-Institut für Deutsche Sprache 2019b). Die Abfragen im DWDS und im DeReKo wurden am 5.11.2019 durchgeführt.

Tabelle 4.1⁸ gibt einen Überblick über die Testverben. Darin wird die relative Frequenz des starken Partizips II der Verben (Vorkommen pro Million Token) in den Wortverlaufskurven im Jahr 2018 des DWDS⁹, die absolute Frequenz der starken und schwachen Formen im DeReKo sowie die Ratio zwischen starken und schwachen Formen im DeReKo angegeben. Zwischen Verwendungen der Partizip-II-Formen als Perfekt- bzw. Plusquamperfektform und Prädikatsnomen wird nicht unterschieden. Die Verben sind absteigend nach ihrer Ratio zwischen starken und schwachen Formen sortiert.

⁸Die Ergebnisse wurden nur in Hinblick auf schwache Formen für frequente und infrequente Verben ohne Schwankung bereinigt. Dabei wurden dialektale und metasprachliche Belege sowie Fehlschreibungen (*getragt* statt *gefragt*) ausgeschlossen. Belege, die darauf schließen lassen, dass ein Kind spricht oder nachgeahmt wird, wurden nicht ausgeschlossen, da nicht immer eindeutig ersichtlich war, ob dies der Fall ist. Bei *gefahren* wurden groß geschriebene Belege ausgeschlossen, um Fehltreffer aufgrund des homographen Substantivs (*die Gefahren*) zu vermeiden.

⁹Die relative Frequenz der schwachen Partizip-II-Formen ist zu vernachlässigen: Die höchste relative Frequenz im Jahr 2018 weisen *gesinnt* (0,47) und *gewebt* (0,31) auf. Es folgen *eingesaugt* (0,06) und *gehaut* (0,03). Die Verben entstammen alle der Ausprägung INFREQUENT MIT SCHWANKUNG. Bei allen anderen Testverben liegt die relative Frequenz der schwachen Partizip-II-Formen bei 0.

4.2 Lexical-decision-Studie zu Frequenzeffekten

Tabelle 4.1: Überblick über die Testverben (RF: Relative Frequenz)

Verb	RF	stark	absolut	schwach	absolut	Ratio
frequent						
<i>sprechen</i>	53	<i>gesprochen</i>	506.078	<i>gesprecht</i>	1	506.000/1
<i>fahren</i>	26	<i>gefahren</i>	308.377	<i>gefahren</i>	0	308.00/0
<i>tragen</i>	21	<i>getragen</i>	283.025	<i>getragt</i>	1	283.000/1
<i>ziehen</i>	44	<i>gezogen</i>	481.702	<i>gezieht</i>	2	240.000/1
<i>halten</i>	46	<i>gehalten</i>	537.019	<i>gehalten</i>	5	107.000/1
<i>schreiben</i>	61	<i>geschrieben</i>	664.701	<i>geschrieben</i>	10	66.000/1
<i>sinken</i>	22	<i>gesunken</i>	178.562	<i>gesunken</i>	3	60.000/1
<i>fliegen</i>	13	<i>geflogen</i>	113.074	<i>geflogen</i>	4	28.000/1
infrequent ohne Schwankung						
<i>flechten</i>	0,4	<i>geflochten</i>	4.847	<i>geflechtet</i>	6	807/1
<i>spinnen</i>	0,38	<i>gesponnen</i>	5.712	<i>gespinn</i>	11	519/1
<i>schmelzen</i>	0,93	<i>geschmolzen</i>	9.278	<i>geschmelzt</i>	19	488/1
<i>anschwellen</i>	0,69	<i>angeschwollen</i>	4.910	<i>angeschwellt</i>	11	446/1
<i>kneifen</i>	0,16	<i>gekniffen</i>	2.070	<i>gekneift</i>	6	345/1
<i>fechten</i>	0,23	<i>gefochten</i>	3.449	<i>gefechtet</i>	13	265/1
<i>melken</i>	0,34	<i>gemolken</i>	4.441	<i>gemelkt</i>	25	177/1
<i>dreschen</i>	0,16	<i>gedroschen</i>	2.967	<i>gedrescht</i>	24	123/1
infrequent mit Schwankung						
<i>salzen</i>	0,31	<i>gesalzen</i>	2.955	<i>gesalzt</i>	49	60/1
<i>hauen</i>	1,45	<i>gehauen</i>	14.999	<i>gehaut</i>	593	25/1
<i>quellen</i>	0,03	<i>gequollen</i>	268	<i>gequollt</i>	29	9/1
<i>einsaugen</i>	0,11	<i>ingesogen</i>	1.012	<i>ingesaugt</i>	446	2,2/1
<i>weben</i>	0,16	<i>gewoben</i>	2.242	<i>gewebt</i>	2.946	0,8/1
<i>glimmen</i>	0	<i>geglommen</i>	25	<i>geglimmt</i>	49	0,5/1
<i>gären</i>	0	<i>gegoren</i>	130	<i>gegärt</i>	280	0,46/1
<i>sinnen</i>	0,15	<i>gesonnen</i>	3.285	<i>gesinnt</i>	5.611	0,59/1

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

Ausschlaggebend für die Einteilung der Testverben in Tokenfrequenzausprägungen ist die Frequenz der starken Partizip-II-Form (z. B. @gezogen) innerhalb der Wortverlaufskurve im Jahr 2018: Frequent sind Verben mit einer Partizip-II-Frequenz ab 13 Belegen pro Million Token, infrequent sind alle Verben mit bis zu zwei Belegen pro Million Token.

Neben der Einteilung in Frequenzausprägungen wurden die infrequenten Verben in Verben mit attestierter Schwankung und ohne attestierte Schwankung in Korpora geteilt. Um die Unterscheidung zwischen Verben mit und ohne Schwankungen in Korpora zu operationalisieren, wurde nach schwachen und starken Partizip-II-Formen der Verben im Archiv W der geschriebenen Sprache (Korpus W-öffentlich) im DeReKo über Cosmas II (Leibniz-Institut für Deutsche Sprache 2019a) gesucht und die Ratio zwischen starken und schwachen Formen bestimmt. Ab einer Ratio von über 100 starken Formen zu einer schwachen gilt die Schwankung als nicht attestiert, bei weniger oder genau 100 starken Formen zu einer schwachen Form gilt sie als attestiert.

Die Testverben gehören unterschiedlichen Ablautreihen an. Dies hat zur Folge, dass bspw. Effekte von Form-Schemata mit der Tokenfrequenz interagieren und Verben wie *spinnen* sich eher in der starken Flexion halten könnten als ähnlich frequente Verben, die nicht dem Form-Schema [#_1 + η + (C)] angehören (zu Form-Schemata bei starken Verben siehe Abschnitt 3.1.3). Allerdings ist *spinnen* aufgrund des alveolaren Nasals in der Peripherie des Form-Schemas. Daher ist es nicht überraschend, dass *spinnen* im Präteritum zwischen dem ursprünglichen Ablaut /a/, dem neuen Ablaut /ɔ/ und der schwachen Form schwankt (Nowak 2013: 172–175). Neben Form-Schemata kann die Ablautalternanz Einfluss auf die Reaktionszeit nehmen: Bei der Ablautalternanz ABA ist bspw. zu erwarten, dass der Ablaut im Partizip II keine große kognitive Herausforderung darstellt, da er identisch mit dem Präsensvokal ist (zu Ablautalternanzen und Frequenz siehe Abschnitt 3.1.1). Daher könnten Verben mit der Ablautalternanz ABA niedrigere Reaktionszeiten evozieren als Verben mit den Ablautalternanzen ABC oder ABB. In der Auswertung werden die Lemmata der Testverben als zufällige Effekte in das statistische Modell eingeflochten, sodass einzelne Lemmata das Ergebnis nicht verzerren.

Die Auswahl der infrequenten Verben stellte eine Herausforderung dar, da die meisten starken Verben tokenfrequent sind (siehe Abschnitt 3.1.1). Für die Auswahl musste zusätzlich zur geringen Tokenfrequenz darauf geachtet werden, dass die schwachen Partizip-II-Formen der Testverben nicht mit anderen Formen im Paradigma zusammenfallen. Daher können Präfixverben wie *erklimmen* nicht genutzt werden, da *erklimmt*¹⁰ auch die 3. Ps. Sg. Präs. Ind. sein könnte

¹⁰Aufgrund der Ähnlichkeit von *geklimmen* zu *beklimmen* wurde *klimmen* nicht als Simplex in die Untersuchung eingeflochten.

(*Sie erklimmt den Berg*). Zudem musste darauf geachtet werden, dass die schwache Form des starken Verbs nicht mit der Form eines anderen schwachen Verbs zusammenfällt, wie bei *gebackt* im Sinne von ‚festkleben‘ (Duden 2020), oder als orthographische Abweichung des Partizips II eines schwachen Verbs wahrgenommen werden könnte wie bei *gemahlt*, das als eine Abweichung von *gemalt* angesehen werden könnte.

Bis auf *anschwellen* und *einsaugen* sind alle Testverben Simplizia. *Anschwellen* wurde statt *schwellen* genutzt, um eine transitive Lesart auszuschließen, die nur schwache Formen zulässt (*Der Affe schwellte die Brust vor Stolz*) (Duden 2020). *Einsaugen* wurde genutzt, um die Assoziation mit *staubsaugen* zu stören, da *staubsaugen* nur schwache Formen zeigt, während *saugen* im Sinne von ‚eine Flüssigkeit aufnehmen‘ weiterhin zwischen stark und schwach schwankt (Nowak 2011: 316). Bei *anschwellen* und *einsaugen* handelt es sich um Partikelverben, weshalb eine Verwechslung der schwachen Partizipform (*eingesaugt/angeschwellt*) mit der 3. Ps. Sg. Ind. (*er/sie/es saugt ein/schwellt an*) ausgeschlossen ist.

Wenig überraschend gehören bis auf *salzen*, *kneifen* und *hauen* alle infrequenten Verben dem Ablautmuster *x-o-o* an, haben also den ursprünglichen Ablaut im Präteritum gewechselt und /o:/ bzw. /ɔ/ angenommen.¹¹ *Salzen* bildet das Präteritum schwach (Duden 2020). *Kneifen* zeigt im Präteritum Schwankungen zwischen stark und schwach: Im Archiv W des DeReKo stehen 15 schwache Formen 1704 starken Formen gegenüber (Ratio 75/1). Bei *hauen* ist die Schwankung im Präteritum deutlich zu beobachten: 5.844 starke Belege stehen 5.526 schwachen Belegen gegenüber (Ratio 1,05/1).¹² Somit zeigen alle infrequenten Testverben Schwankungen im Präteritum zu *x-o-o* und/oder schwachen Formen, was eine Schwankung im Partizip II wahrscheinlicher macht (siehe Tabelle C.1 in Anhang C für einen Überblick über die Frequenz starker und schwacher Präteritalformen der infrequenten Testverben).

¹¹Dies gilt nicht für *saugen*, das bereits im Mhd. zur Ablautreihe 2 zählte und daher bereits ursprünglich auf /o:/ ablautete (Nowak 2015: 233). Wie oben erwähnt zeigen sich für *saugen* im Partizip II Schwankungen.

¹²Bei der Suche nach *kniff*, *hieb* und *haute* wurden großgeschriebene Substantive ausgeschlossen, um Fehltreffer aufgrund von Konversion (*der Kniff*, *der Hieb*) und dem homographen französischen Adjektiv *haute* (wie in *Haute Couture*) zu vermeiden. Für die kleingeschriebenen Belege von *haute* wurde eine Stichprobe von 1.000 Belegen durchgesehen: Darin waren 30 Fehltreffer enthalten, weshalb die Trefferanzahl um 3 % nach unten korrigiert wurde, unbereinigt ergeben sich 5.697 Belege.

4.2.3 Versuchsdesign

Die Studie nutzt ein within-subject-Design mit zwei Faktoren: Frequenz (mit drei Ausprägungen: FREQUENT, INFREQUENT OHNE SCHWANKUNG, INFREQUENT MIT SCHWANKUNG) und Partizip-II-Form (mit zwei Ausprägungen: stark, schwach). Die Ausprägungen enthalten jeweils acht Testwörter. Die Proband_innen müssen entscheiden, ob sie das angezeigte Wort kennen. Dabei können sie *ja* oder *nein* antworten. Die Tastaturbelegung für *ja* und *nein* ist *f* für *nein* und *j* für *ja*. Um keine Unruhe in das Experiment zu bringen, wird diese Tastaturbelegung das gesamte Experiment beibehalten. Vor dem Versuchswort erscheint jeweils zwei Sekunden lang eine Fixation (###), um spill-over-Effekte zu vermeiden. Die Antwort sowie die Reaktionszeit werden gespeichert.

Als Filler dienen schwache und starke Maskulina mit schwacher bzw. starker Flexion (*des Kollegen/des Kolleges*). Diese stellen die Testwörter der lexical-decision-Studie zu Form-Schematizitätseffekten bei schwachen Maskulina dar (siehe Abschnitt 4.5). Zudem werden starke Maskulina und Neutra mit langer Genitivform als Filler genutzt, die lange Endungen phonotaktisch nicht zulassen (z. B. *des *Bodenes*) (Szczepaniak 2010: 107–108) (siehe Anhang D.2 für eine Aufschlüsselung der Filleritems).

Als Test, ob das Design funktioniert, dient der Kontrast zwischen langer und kurzer Genitivendung bei zwei Substantiven, die nur die kurze Endung zulassen (*des Ultimatums/*Ultimatumes, des Internets/*Internetes*), sowie zwischen unregelmäßiger und regelmäßiger Konjugation bei zwei unregelmäßigen Verben (*gegangen/*geht, gebracht/*bringt*). Diese Items sind nur relevant, wenn für die Testitems keine Unterschiede in den Reaktionszeiten festzustellen sind. Für die kurzen Endungen bzw. unregelmäßigen Formen sind geringere Reaktionszeiten zu erwarten als für die langen bzw. regelmäßigen Formen. Zudem ist ein klares Antwortverhalten zu erwarten (kurze bzw. unregelmäßige Formen sind bekannt, lange bzw. regelmäßige unbekannt). Die Items eignen sich daher, um zu prüfen, ob das Versuchsdesign generell Unterschiede in Reaktionszeiten hervorruft.

Insgesamt werden 109 Versuchitems präsentiert. Dabei entfallen 48 auf die Testitems, 49 auf Filler (30 starke/schwache Maskulina, 19 lange Genitivformen), acht auf Items, die das Design testen, sowie vier auf Items im Beispielblock. Bei 34 Items sind affirmierende Antworten auf die Frage nach der Bekanntheit der Formen zu erwarten (bspw. bei der Form *gegangen*), bei 49 verneinende Antworten (bspw. bei der Form *geht*). Es verbleiben 26 Items, bei denen theoretisch beide Formen bekannt sein können. Der vermeintliche Überschuss an verneinenden Antworten ergibt sich aus Fillern für die Experimentblöcke mit langen Genitivformen. Hier wird nur die lange Genitivform, nicht aber die kurze abge-

fragt. Der vermeintliche Überschuss ist dabei beabsichtigt, um ein Gegengewicht zu den 26 Items zu erhalten, bei denen beide Formen bekannt sein können. Dies ist bspw. bei den infrequenten Verben mit attestierter Schwankung (*gewebt* und *gewoben*) der Fall. Berechnet man diese Items als affirmierend in das Verhältnis von affirmierenden und verneinenden Antworten ein, ist davon auszugehen, dass mehr Items affirmierend als verneinend beantwortet werden. Deswegen werden durch die Filler zusätzlich verneinende Antworten evoziert, um ein Ungleichgewicht zwischen affirmierenden und verneinenden Antworten zu vermeiden. Bei den frequenten Testverben sowie den infrequenten Testverben ohne attestierte Schwankung wird im Gegensatz zu den infrequenten Verben mit attestierter Schwankung erwartet, dass die starke Form jeweils bekannt, die schwache hingegen nicht bekannt ist.

Die Testitems sind in vier Experimentblöcken organisiert, die jeweils 25 Items enthalten. Die Blöcke erscheinen in randomisierter Reihenfolge, aber die Reihenfolge der Items innerhalb der Blöcke variiert nicht. Zudem existiert ein Beispielblock mit vier Items und ein Eingewöhnungsblock mit fünf Items (siehe Tabelle D.1 in Anhang D.1). In den Experimentblöcken sind jeweils zwölf Testverben (vier Testverben pro Frequenzprägung) enthalten. Dabei sind die starken und schwachen Formen eines Testitems stets auf zwei Blöcke verteilt. Die restlichen Items sind Filler. Jeder Block beginnt mit einem Filler, um das zufällige Aufeinandertreffen von zwei Testwörtern zwischen zwei Blöcken zu verhindern. Innerhalb der Blöcke konnte nicht vermieden werden, dass zwei Testverben hintereinander präsentiert werden. Es wurde jedoch darauf geachtet, dass dies maximal zweimal pro Experimentblock geschieht und dass die Testverben jeweils aus unterschiedlichen Frequenzprägungen stammen. Pro Experimentblock sind elf Items enthalten, bei denen eine verneinende Antwort antizipiert wird, und 14, bei denen eine affirmierende Antwort antizipiert wird. Bei sieben bis acht der Items mit affirmierender Antwort (bspw. *gezogen*) ist die Wahrscheinlichkeit hoch, dass die korrespondierende Form unbekannt ist und daher verneinende Antworten evoziert (bspw. **gezieht*). Die restlichen Items gehören zu den Items, bei denen beide Formen bekannt sein könnten (bspw. bei *gewebt/gewoben*). Innerhalb der Blöcke wurde darauf geachtet, dass keine Muster in den zu erwartenden Antworten entstehen.

4.2.4 Versuchsablauf

Die Studie wurde online mithilfe der Plattform Pavlovia (PsychoPy Team 2019) durchgeführt. In der Studie erhalten die Teilnehmer_innen alle Anweisungen per Bildschirm. Zunächst werden die Teilnehmer_innen in einem gesonderten

Browser-Fenster begrüßt und darauf hingewiesen, dass Ihre Metadaten anonymisiert gespeichert werden. In diesem Browser-Fenster werden auch die Metadaten abgefragt. Dann werden im ursprünglichen Browser-Fenster allgemeine Informationen gegeben mit dem Hinweis, dass die Proband_innen mit der Teilnahme der anonymisierten Speicherung ihrer Daten zustimmen. Zudem werden die Instruktionen zum Experiment angezeigt. Die Proband_innen werden anhand eines Beispiels an die Aufgabe gewöhnt. Die Proband_innen werden dazu aufgefordert, die angezeigten Wörter bzw. Wortabfolgen danach zu bewerten, ob sie sie kennen (Kennen Sie dieses Wort bzw. diese Wortabfolge?). Im Beispielblock erscheinen sowohl Formen, die wahrscheinlich bekannt sind (*die Tische, du siehst*), als auch Formen, die wahrscheinlich unbekannt sind (*die *Autoe, du *gebst*). Im Anschluss an das Beispiel folgt zunächst ein Eingewöhnungsblock, der fließend und somit von den Proband_innen unbemerkt in die Experimentblöcke übergeht. Für die Erstellung des Experiments wurde PsychoPy3 genutzt (Peirce u. a. 2019), das eine präzise Messung von Reaktionszeiten ermöglicht. Die Skripte, die für die Datenerhebung verwendet wurden, sind im digitalen Anhang dokumentiert.

4.2.5 Metadaten

Folgende Metadaten wurden abgefragt. Die Abfrage erfolgte jeweils über Freitextfelder.

1. Geburtsjahr
2. Geschlecht
3. Legasthenie
4. höchster Abschluss
5. L1
6. Händigkeit
7. Vornehmlich genutzte Varietät
8. Dialektkenntnis
9. Herkunft (Ort, Bundesland)
10. Wohnort
11. Leseintensität (täglich, wöchentlich, monatlich).

4.2 Lexical-decision-Studie zu Frequenzeffekten

Die Metadaten dienen vornehmlich dazu, sicherzustellen, dass die untersuchten Gruppen möglichst homogen sind. Die Händigkeit der Proband_innen wird abgefragt, da diese Einfluss auf das Antwortverhalten der Personen haben kann: Je nachdem, ob für die Beantwortung von Fragen die rechte oder linke Hand genutzt wird, sind Rechts- bzw. Linkshänder schneller (Höhle 2010: 215). Da es sich um ein within-subject-Design handelt und somit alle Proband_innen Daten zu jedem Versuchitem liefern, sollte die Händigkeit jedoch nur bedingt Einfluss auf die Ergebnisse nehmen. Zudem werden die Proband_innen als zufälliger Effekt in die statistische Auswertung aufgenommen, sodass individuelle Präferenzen die Ergebnisse nicht verzerren sollten.

Die L1 ist als die Sprache(n) definiert, die vor dem 12. Lebensjahr außerhalb der Schule erworben wurde(n). Die Frage wird genutzt, um Proband_innen auszuschließen, die nicht Deutsch als L1 sprechen oder bilingual aufgewachsen sind. In der Studie wird nicht der Terminus *L1*, sondern *Muttersprache* genutzt. Die Altersgrenze wurde gewählt, da die Pubertät als obere Grenze der *critical period* angesetzt wird, innerhalb derer ein Vorteil beim Erlernen einer Sprache gegenüber Sprecher_innen besteht, deren Spracherwerb erst später einsetzt (Fromkin u. a. 2014: 434).

Die vornehmlich genutzte Varietät wurde abgefragt, um besser einschätzen zu können, inwiefern die Proband_innen einen ähnlichen Sprachgebrauch aufweisen. In der Frage wurden Standard und Dialekt gegenübergestellt (Sprechen Sie in Ihrem familiären Umfeld vornehmlich Hochdeutsch oder Dialekt?). Dabei wurde in der Studie der Terminus *Hochdeutsch* genutzt, da davon auszugehen ist, dass dieser Terminus den Proband_innen geläufiger ist als *Standard*. Die Spezifizierung auf das familiäre Umfeld wurde genutzt, um in den Daten eher eine Verzerrung hin zum Dialekt als zum Standard zu erreichen. Auf diese Weise ist es unwahrscheinlicher, dass die Homogenität zwischen den Sprecher_innen überschätzt wird. Im Anschluss an die Frage nach der präferierten Varietät wurde die Dialektkenntnis der Proband_innen erhoben. Dabei wurde gefragt, ob die Proband_innen einen Dialekt sprechen und wenn ja, welchen. Die Dialektkenntnis wird somit als reine Selbsteinschätzung abgefragt. Alle Fragen zu den Metadaten sind im digitalen Anhang aufgelistet.

4.2.6 Probandenakquise und Proband_innen

Die Proband_innen wurden akquiriert, indem der Link zur Studie an Studierende in Germanistikveranstaltungen an der Universität Bamberg sowie der Universität Mainz¹³ geschickt wurde. Zusätzlich habe ich den Link in meinem persönli-

¹³Ich danke Jessica Nowak, Renata Szczepaniak, Anette Kremer und Jonas Fehn für das Weiterleiten des Studienaufrufs.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

chen Umfeld verbreitet. Die Germanistikveranstaltungen, in denen Proband_innen akquiriert wurden, werden üblicherweise am Anfang des Studiums besucht, sodass die linguistische Vorbildung der Proband_innen begrenzt sein sollte. Es wurde eine Stichprobengröße von 50 Proband_innen angestrebt. Um diese zu erreichen, wurden insgesamt 100 Datensätze erhoben. Für die Studie werden nur Proband_innen berücksichtigt, die Deutsch als L1 sprechen, monolingual aufgewachsen sind und keine Legasthenie haben. Proband_innen, die keine Angaben zu L1 oder Legasthenie machten, wurden ausgeschlossen. Mit der Art der Probandenakquise geht einher, dass vornehmlich Menschen im Alter zwischen 20 und 30 getestet wurden. Diese Limitation gilt für alle Studien, die innerhalb des Dissertationsvorhabens durchgeführt wurden.

Zudem wurden Proband_innen ausgeschlossen, deren Antwortverhalten auf Unkonzentriertheit schließen lässt. Als Entscheidungskriterium hierfür dient das Antwortverhalten bei Fillern und Items, die im Beispielblock genutzt wurden bzw. als Test für das Design dienen. Bei 31 dieser Items ist von einem eindeutigen Antwortverhalten auszugehen. Proband_innen, die drei oder mehr der 31 Items anders beantworteten als erwartet, werden von der Studie ausgeschlossen. Dies entspricht einem Ausschluss bei einer Fehlerquote ab 10 %. Außerdem werden Proband_innen ausgeschlossen, deren Reaktionszeiten mehr als drei Standardabweichungen über oder unter dem Mittelwert eines Testitems liegen.

Die Stichprobengröße von 50 wurde basierend auf einer Prästudie¹⁴ bestimmt. Die Ergebnisse der Prästudie deuten darauf hin, dass frequente Testverben schneller prozessiert werden als infrequente. Die Werte aus der Prästudie wurden als Grundlage für eine Datensimulation genutzt, um die Stichprobengröße zu bestimmen, die für eine Teststärke von über 0,9 (für den Schätzwert der Ausprägung *INFREQUENT MIT SCHWANKUNG* mit schwachen Formen im Vergleich zu den Schätzwerten für dieselbe Ausprägung mit starken Formen sowie für Formen der Ausprägung *FREQUENT*) bei einem α -Level von 0,01 benötigt wird. Um die Zielgröße von 50 Proband_innen zu erreichen, wurden doppelt so viele Daten erhoben, da in der Prästudie mit 13 von 32 Proband_innen 40,6 % der Proband_innen ausgeschlossen werden mussten.

Die Daten wurden vom 9.1.2020 bis zum 27.1.2020 erhoben. Nach Bereinigung der Daten blieb eine Stichprobe von 53 Proband_innen. Zu einigen Metadaten

¹⁴Die Stichprobengröße in der Prästudie betrug nach Ausschluss von Proband_innen anhand der oben genannten Kriterien 19 Proband_innen. Die Prästudie hat denselben Versuchsablauf wie die Hauptstudie. Nur der aktuelle Wohnort wurde in der Prästudie nicht abgefragt. Die Proband_innen der Prästudie erfüllen dieselben Kriterien wie in der Hauptstudie. Die Probandenakquise lief über mein persönliches Umfeld. Die Metadaten der Proband_innen und die Ergebnisse der Prästudie sind im digitalen Anhang aufgeschlüsselt.

4.3 Sentence-maze-Studie zu Prototypizitätseffekten

fehlen Angaben, daher ergeben die nun folgenden Angaben zu den Proband_innen aufaddiert nicht immer 53. Alle Metadaten sind im digitalen Anhang auch tabellarisch aufgeschlüsselt.

Die Proband_innen sind im Mittel 26,5 Jahre alt mit einer Standardabweichung von sechs Jahren. Zwölf identifizieren sich als männlich, 38 als weiblich; niemand gab ein anderes Geschlecht an. 18 der Proband_innen kommen aus Bayern, sechs aus Rheinland-Pfalz, jeweils vier aus Baden-Württemberg und Niedersachsen. Die anderen Bundesländer sind maximal mit zwei Proband_innen vertreten. 38 Proband_innen leben derzeit in Bayern, acht in Rheinland-Pfalz, zwei in Hessen und jeweils ein_e in Baden-Württemberg und Niedersachsen. 45 Proband_innen gaben an, täglich zu lesen, fünf lesen wöchentlich, eine Person gab an, gar nicht zu lesen, verhielt sich aber in der Aufgabe ähnlich wie die anderen Proband_innen. Die restlichen Proband_innen machten keine Angabe zu ihrem Leseverhalten. 36 Proband_innen gaben an, in ihrem familiären Umfeld Standard zu sprechen, elf sprechen dialektal, zwei nach eigenen Angaben eine Mischform. Dabei erläutern sie jedoch nicht, wie sie zu der Einschätzung als Mischform kommen. Die angegebenen Dialekte sind im digitalen Anhang aufgeschlüsselt. 31 der Proband_innen haben Abitur, vier einen Bachelor- und zehn einen Masterabschluss. 45 Proband_innen verwenden vornehmlich die rechte Hand, fünf die linke.

Die Ergebnisse der Studie zu Frequenzeffekten werden in Abschnitt 5.2 vorgestellt. Im folgenden Abschnitt wird das methodische Vorgehen in der sentence-maze-Studie zu Prototypizitätseffekten erläutert.

4.3 Sentence-maze-Studie zu Prototypizitätseffekten

Der Einfluss der Prototypizität auf Variation wird anhand der Auxiliarselektion von *haben* und *sein* mithilfe einer *sentence maze task* untersucht. Darin wählen Proband_innen in Sätzen mit unterschiedlichem Transitivitätsgrad zwischen *hat* und *ist*. Im Abschnitt werden zunächst die der Studie zugrunde liegenden Fragestellungen und Hypothesen erläutert und anschließend das methodische Vorgehen (Materialien, Versuchsdesign und -ablauf, Metadaten, Probandenakquise und Proband_innen).

4.3.1 Fragestellung und Hypothesen

In der Studie wird der Frage nachgegangen, inwiefern Prototypizität sich in den Reaktionszeiten bei der Selektion von *haben* und *sein* zeigt und wie sie die Selektion beeinflusst. Dabei wird mit verschiedenen Transitivitätsgraden gearbeitet: Prototypisch transitive sowie intransitive Sätze werden gegenübergestellt.

Transitive Sätze selektieren *haben*. Bei intransitiven Sätzen beeinflussen Telizität und Bewegungssemantik, ob *sein* selektiert wird (Gillmann 2016: 316–319; siehe Abschnitt 3.3.2.1 für Erläuterungen zu Transitivität). Da die Testverben Bewegungsverben sind, können zwei Prototypen kontrastiert werden: Transitivität mit *haben*-Selektion und Intransitivität mit *sein*-Selektion. In den prototypisch transitiven Sätzen ist das Objekt ein Patiens, in den prototypisch intransitiven Sätzen ist kein Objekt vorhanden. Im Übergangsbereich zwischen Transitivität und Intransitivität werden Sätze getestet, deren Objekt ambig zwischen Patiens und Instrument ist (AMBIG I). Zudem werden Sätze mit Satzgliedern getestet, die ein Adverbial darstellen (AMBIG II), siehe Abschnitt 4.3.2 für genauere Erläuterungen zu den Testsätzen. Es werden folgende Hypothesen überprüft:

4.3.1.1 Reaktionszeiten

1. Die Reaktionszeiten sind bei intransitiven Sätzen niedriger als bei den anderen Sätzen.
2. Die transitiven Sätze rufen vergleichbare Reaktionszeiten hervor wie AMBIG II.
3. AMBIG I evoziert die höchsten Reaktionszeiten.

Diese Annahmen basieren auf den Ergebnissen der Prästudie. Diese sind im digitalen Anhang dokumentiert. Ursprünglich wurde angenommen, dass die Reaktionszeiten für AMBIG I und II höher sind als für die Prototypen der beiden Schemata (eindeutige Transitivität und eindeutige Intransitivität), denn durch die Ambiguität sind theoretisch beide Funktionen (Transitivität und Intransitivität) möglich, weshalb die Entscheidung zwischen *haben* und *sein* länger dauern sollte als bei den prototypisch intransitiven und transitiven Sätzen. Bei den Prototypen der Schemata existiert hingegen nur eine gefestigte Form, sodass die Wahl zwischen *haben* und *sein* schnell erfolgen sollte. Allerdings sind für transitive und ambige Sätze auch etwas höhere Reaktionszeiten möglich als für intransitive Sätze, da diese komplexer sind als die intransitiven Sätze. In der Prästudie hat INTRANSITIV die niedrigsten Reaktionszeiten evoziert, AMBIG II hat vergleichbare Reaktionszeiten hervorgerufen wie die transitiven Sätze, während AMBIG I die höchsten Reaktionszeiten evoziert hat. Sollte sich dies in der Hauptstudie bestätigen, wären die Reaktionszeiten auch durch die Komplexität der Sätze beeinflusst. Zudem könnten die geringen Reaktionszeiten von AMBIG II im Vergleich zu AMBIG I einen Hinweis darauf geben, dass AMBIG II nur theoretisch ambig ist, sich

aber tatsächlich nur eine Form mental gefestigt hat und daher vergleichbare Reaktionszeiten wie bei der Ausprägung TRANSITIV evoziert werden, bei der nur eine Form gefestigt ist.

4.3.1.2 Antwortverhalten

1. Bei intransitiven Sätzen wird *sein* bevorzugt, bei transitiven *haben*.
2. In den ambigen Sätzen zeigen Proband_innen eine klare Tendenz zu *sein*, dennoch wird hier mit mehr *haben*-Antworten gerechnet als für die intransitiven Sätze.
3. AMBIG II zeigt einen höheren Anteil an *sein*-Antworten als AMBIG I.
4. Für alle Sätze ist mit itemspezifischen Präferenzen zu rechnen (bspw. bei AMBIG I *sein*-Antworten für *reiten*, aber *haben*-Antworten für *fahren*).

Die Hypothesen zu den ambigen Sätzen basieren auf den Ergebnissen der Prästudie. Ursprünglich wurde Variation zwischen *haben* und *sein* für die ambigen Sätze angenommen. Für AMBIG I wurde eine kleine Tendenz zu *haben* erwartet, da die *haben*-Selektion für *fahren* in Sätzen mit determinierten Konkreta, die auch in AMBIG I enthalten sind, mit 80 % deutlich frequenter ist als die *sein*-Selektion und für *fliegen* beide Auxiliare immerhin etwa gleich häufig vorkommen (Gillmann 2016: 288–290). Die *sein*-Präferenz für AMBIG II war vorab erwartet worden: Die verwendeten Abstrakta stellen Pfad-Akkusative dar (siehe Abschnitt 4.3.2). Für diese stellt Gillmann (2016: 299–301) korpuslinguistisch fest, dass sie mit über 90 % stark zur *sein*-Selektion tendieren. Die itemspezifischen Präferenzen sind vor allem für das Testverb *reiten* zu erwarten, das in der Prästudie ein anderes Antwortverhalten aufwies als *fahren* und *fliegen*.

Die hier vorgestellten Hypothesen sowie das methodische Vorgehen inklusive der geplanten statistischen Auswertung wurden vor der Datenerhebung registriert. Die Registrierung lässt sich unter <https://osf.io/6vy8k> aufrufen. Die statistische Auswertung wird zusammen mit den Ergebnissen in Abschnitt 5.3 vorgestellt.

4.3.2 Materialien

Als Testitems wurden die Bewegungsverben *fahren*, *reiten* und *fliegen* gewählt, da diese unterschiedliche Transitivitätsgrade zulassen. Sie wurden in folgende Sätze integriert:

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

- (3) transitiv (Objekt dient als Patiens)
- Er sah, wie die Mutter die Kinder zur Schule gefahren **hat/*ist**.
 - Ich nehme an, dass der Pilot die Passagiere zum nächstgelegenen Flughafen geflogen **hat/*ist**.
 - Ich gehe davon aus, dass das Mädchen das Pferd zur Koppel geritten **hat/*ist**.
- (4) ambig I (Objekt könnte Patiens oder ein Instrument sein)
- Er erzählte mir, dass seine Tante eine Zeit lang ein Cabrio gefahren **hat/ist**.
 - Ich kann kaum glauben, dass meine Schwester neulich einen Helikopter geflogen **hat/ist**.
 - Mir wurde erzählt, dass mein Onkel früher ein schwarzes Pferd geritten **hat/ist**.
- (5) ambig II (Objekt vs. Adverbial)
- In den Nachrichten sagten sie, dass der neue Fahrer die Strecke in Rekordzeit gefahren **ist/hat**.
 - In der Zeitung stand, dass der Pilot die Distanz in Rekordzeit geflogen **ist/hat**.
 - Im Radio hieß es, dass der Reiter den Parcours fehlerlos geritten **ist/hat**.
- (6) intransitiv (kein Objekt)
- Ich ging davon aus, dass die Oma am Sonntag zur Kur gefahren **ist/*hat**.
 - Ich wusste nicht, dass mein Bruder am Mittwoch in die USA geflogen **ist/*hat**.
 - Meine Oma erzählte, dass sie in den Siebzigern regelmäßig durch den Schwarzwald geritten **ist/*hat**.

Die Sätze wurden als Nebensätze konzipiert, damit das Auxiliar das letzte Element im Satz darstellt und somit die Transitivität des Satzes beim Prozessieren des Auxiliars eindeutig ist. Die Testsätze wurden vorab auf ihre semantische Plausibilität geprüft, indem sie 58 Proband_innen vorgelegt wurden. Alle Sätze wurden dabei als semantisch plausibel bewertet.¹⁵ Erläuterungen zum Design des Plausibilitätstests und die Plausibilitätsbewertungen für die einzelnen Testsätze sind in Anhang F zu finden.

¹⁵Die Sätze erzielten durchschnittlich mindestens eine Bewertung von 3 auf einer Skala von 1 (nicht plausibel) bis 5 (plausibel). Nur ein Satz musste leicht modifiziert werden, da er mit 2,7 unter dem Grenzwert lag. Ursprünglich lautete der Satz *Mir wurde erzählt, dass mein Onkel während des Studiums ein schwarzes Pferd geritten hat. Während des Studiums wurde durch früher ersetzt, sodass der Satz mehr den anderen Sätzen mit reiten ähnelt. Zusätzlich musste eine Kleinigkeit korrigiert werden: Parkour wurde zu Parcours geändert, da Parkour norma-*

4.3.3 Versuchsdesign

Die Studie nutzt ein within-subject-Design mit zwei Faktoren: Transitivität (mit vier Ausprägungen: TRANSITIV, AMBIG I, AMBIG II, INTRANSITIV) und Auxiliariumwahl (mit zwei Ausprägungen: *hat*, *ist*). Die Ausprägungen enthalten jeweils drei Testwörter (*fahren*, *fliegen*, *reiten*). Die Aufgabe der Proband_innen ist es, Satzstücke für Stück zusammensetzen. Dabei werden den Proband_innen zwei Möglichkeiten gegeben, die links und rechts auf dem Bildschirm angezeigt werden. Die Möglichkeiten können ausgewählt werden, indem die Taste *f* für das linke Element oder die Taste *j* für das rechte gedrückt wird. Die Aufgabe der Proband_innen besteht darin, die Möglichkeit zu wählen, die den bisher gelesenen Satz am besten fortsetzt. Beim letzten Element der Testsätze wählen die Proband_innen jeweils zwischen *hat* und *ist*. Dabei wird die Antwort der Proband_innen sowie die Reaktionszeit gespeichert.

Zwischen der Tastatureingabe der Proband_innen und dem Erscheinen des nächsten Satzelements ist eine Verzögerung von 0,5 Sekunden vorgesehen. Dies soll spill-over-Effekte vermeiden. Um die Prozessierung des Satzes nicht durch eine lange Pause zu stören und da die Gefahr von spill-over-Effekten bei *sentence maze tasks* relativ gering ist (Forster u. a. 2009: 164), ist die Pause mit 0,5 Sekunden sehr kurz gehalten.

Der Matrixsatz der Testsätze wird stets als Ganzes präsentiert. Dabei werden zwei Möglichkeiten angeboten. Die Alternative zum Matrixsatz, der in den Testsätzen vorgesehen ist, wird durch orthographische und grammatische Abweichungen eindeutig als nicht zu favorisieren gekennzeichnet. Die Nebensätze sind in Wörter und Wortgruppen geteilt. Als Alternativen zu den beabsichtigten Wörtern und Wortgruppen werden orthographische (**Schuhle* statt *Schule*) und grammatische (**meine Neffe* statt *mein Neffe*) Abweichungen angeboten sowie Wörter, die im Kontext semantisch unpassend sind (*singen* statt *betrachten*). Tabelle E.1 in Anhang E.1 führt alle Sätze in der Studie samt der Einteilung in Wörter und Wortgruppen sowie den Alternativen zu den intendierten Wörtern und Wortgruppen auf. Die Sätze in der Tabelle sind nach ihrem Vorkommen in den Blöcken sortiert.

Das Ende des Matrixsatzes wird jeweils durch Kommata angezeigt, das Ende des Satzgefüges jedoch nicht durch Punkte. Auf diese Weise können wrap-up-Effekte am Ende eines Satzes (Rayner u. a. 2000) minimiert und außerdem verhindert werden, dass die Proband_innen das letzte Element des Satzes unkonzentriert wählen, da die Proband_innen beim Prozessieren des letzten Items in

lerweise nicht in Bezug auf Hindernisreiten genutzt wird. Zudem war bei AMBIG II mit dem Testverb *fliegen* ursprünglich *Bestzeit* vorgesehen, was in *Rekordzeit* geändert wurde. Dies war ein Übertragungsfehler, der die Plausibilität des Satzes jedoch nicht verändern sollte.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

einem Satz noch nicht wissen, ob der Satz noch fortgeführt wird oder nicht.

In den Testsätzen haben Proband_innen bei dem Partizip II vor dem Auxiliär stets die Wahl zwischen dem Testverb und einem anderen, semantisch unplausiblen Bewegungsverb (bspw. *gefahren* oder *geschwommen* im Satz *Er sah, wie die Mutter die Kinder zur Schule gefahren hat*). So kann ausgeschlossen werden, dass die Alternative zum intendierten Vollverb die Wahl des Auxiliärs beeinflusst. Dies könnte passieren, wenn statt *geschwommen* als semantisch unplausible Alternative *gelacht* zur Auswahl stünde, das immer *haben* selegiert.

Als Filler dienen Sätze, die Maskulina mit schwacher bzw. starker Flexion enthalten, sowie Sätze mit ähnlichem Inhalt wie die Testsätze. Die Maskulina sind Versuchitems der sentence-maze-Studie zum Einfluss der Form-Schematizität auf Deklination (siehe Abschnitt 4.5). Die Füllsätze sind wie die Testsätze Satzgefüge. Zusätzlich zu den Füllsätzen lenkt das gesamte Untersuchungsdesign vom Zweck der Studie ab, da die Proband_innen semantische, grammatische sowie orthographische Abweichungen vermeiden müssen. Als Test, ob das Design funktioniert, dienen Genusfehler (*der Kurier/*die Lieferant*), Rechtschreibfehler (*Schule/*Schuhle*) sowie Varianten (*Fahrrad fahren/fahrradfahren; fehlerlos/fehlerfrei*). Diese Items sind nur relevant, wenn für die Testitems kein Unterschied in den Reaktionszeiten festzustellen ist. Bei Genus- und Rechtschreibfehlern ist ein klares und schnelles Antwortverhalten zu erwarten, bei den Varianten jedoch nicht. Somit eignen sich die Items, um zu überprüfen, ob das Design generell Unterschiede in Reaktionszeiten hervorruft.

Die zu bevorzugende Möglichkeit erscheint gleich häufig links und rechts (jeweils 86-mal), um keine Verzerrung der Reaktionszeiten durch die Antwortverteilung hervorzurufen. Bei 28 der Items sind die Möglichkeiten Varianten, sodass hier kein Antwortverhalten antizipiert werden kann. Dabei sind auch Möglichkeiten als Variante kodiert, die keine Variante darstellen, aber häufig verwechselt werden wie bspw. *dass* und *das* sowie *seit* und *seid*.

Auch innerhalb der Blöcke, in die die Test- und Füllersätze geteilt sind, ist die zu bevorzugende Möglichkeit gleich häufig links und rechts mit maximaler Abweichung um eins. Davon abgesehen ist die Verteilung der Antwortmöglichkeiten innerhalb der Blöcke zufällig, sodass keine Muster zu antizipieren sind. Die Proband_innen müssen nicht nur in den Testsätzen zwischen *hat* und *ist* wählen, sondern auch in vier Füllsätzen, sodass die Auxiliärwahl nicht mit den Testsätzen assoziiert wird. Damit die Auxiliäre nicht immer nur am Ende der Sätze erscheinen, werden zwei der Füllsätze nach den Auxiliären fortgeführt.

Insgesamt werden vier Experimentblöcke mit jeweils sechs Sätzen präsentiert. Die Experimentblöcke werden in randomisierter Reihenfolge angezeigt, aber die

4.3 Sentence-maze-Studie zu Prototypizitätseffekten

Reihenfolge der Sätze innerhalb der Blöcke ist fest. In den Blöcken werden jeweils drei Testsätze genutzt. Diese entstammen jeweils unterschiedlichen Prototypizitätsausprägungen und enthalten unterschiedliche Testverben. Die anderen drei Sätze sind Füllsätze. In den Blöcken wechseln sich Füllsätze und Testsätze regelmäßig ab (siehe Tabelle E.1 in Anhang E.1 für einen Überblick über den Aufbau der Blöcke). Um Priming-Effekte zu vermeiden, wird in den Sätzen vor den Testsätzen der Ausprägungen AMBIG I und AMBIG II jeweils kein Auxiliar verwendet. In Sätzen vor den transitiven Testsätzen kommt stets *ist* vor, während in Sätzen vor den intransitiven Testsätzen stets *hat* oder *habe* steht. So kann ein möglicher Priming-Effekt die Auxiliarselektion nicht überlagern, da sich der Effekt auf das gegenteilige Auxiliar auswirkt.

Zusätzlich zu den Experimentblöcken lesen die Proband_innen einen Beispielblock mit einem Beispielsatz sowie einen Eingewöhnungsblock mit drei Füllsätzen, der den Einstieg in das Experiment erleichtern soll.

4.3.4 Versuchsablauf

Der Versuchsablauf ist identisch mit dem Ablauf in der lexical-decision-Studie zu Frequenzeffekten (siehe Abschnitt 4.2.4). Nur die Aufgabe der Proband_innen weicht logischerweise von der lexical-decision-Studie ab: Wählen Sie die Möglichkeit, die für Sie am besten zu dem bisher gelesenen Satz passt. Genau wie die lexical-decision-Studie wurde auch die sentence-maze-Studie mit PsychoPy3 (Peirce u. a. 2019) erstellt und online mithilfe der Plattform Pavlovia (PsychoPy Team 2019) durchgeführt.

4.3.5 Metadaten, Probandenakquise und Proband_innen

Es wurden dieselben Metadaten erhoben wie für die lexical-decision-Studie zu Frequenzeffekten (siehe Abschnitt 4.2.5) und dieselben Ausschlusskriterien angewandt.¹⁶ Die angestrebte Stichprobengröße von 50 Proband_innen basiert auf einer Datensimulation, die mithilfe der Daten aus der Prästudie¹⁷ durchgeführt

¹⁶Die Konzentriertheit der Proband_innen wurde anhand von 132 Filleritems bestimmt. Für diese 132 Items ist die dispräferierte Antwortmöglichkeit leicht als solche zu erkennen. Dies ist bspw. der Fall bei orthographischen und grammatischen Abweichungen. Bei den orthographischen Abweichungen wurden nur auffällige Abweichungen (**Schuhle*) für die Evaluation des Antwortverhaltens genutzt. Zusammen- und Getrennschreibung wurde daher nicht berücksichtigt. Auch eindeutig geregelte Fälle wie *dass/das* wurden nicht genutzt, da der Diskurs um *dass/das* darauf schließen lässt, dass die Unterscheidung für Schreiber_innen problematisch sein kann. Ab 10 % (13) falsch beantworteter Items wurden Proband_innen ausgeschlossen.

¹⁷Die Stichprobengröße der Prästudie betrug 26 Proband_innen nach Ausschluss von Proband_innen, die die Kriterien nicht erfüllten. Wie in der lexical-decision-Studie zu Frequenzeffekten

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

wurde. Die Prästudie deutet darauf hin, dass das Auxiliar in intransitiven Sätzen schneller gewählt wird als in den anderen Konditionen und dass die Ausprägungen TRANSITIV sowie AMBIG II niedrigere Reaktionszeiten hervorrufen als AMBIG I. Unter der Annahme, dass diese Schätzung robust ist, reichen 50 Proband_innen bei einem α -Level von 0,01 für eine Teststärke von mindestens 0,89 (für den Schätzwert der Ausprägung AMBIG I im Vergleich zu den Schätzwerten der anderen Ausprägungen) aus. Eine Dokumentation des Codes für die Datensimulation ist im digitalen Anhang zu finden.

Um die Stichprobengröße von 50 Proband_innen zu erreichen, wurden insgesamt 170 Personen akquiriert. Zwar mussten in der Prästudie mit sechs von 32 Proband_innen nur 19 % ausgeschlossen werden, da mit der Prototypizitätsstudie jedoch auch die Daten für die sentence-maze-Studie zu Form-Schematizität (siehe Abschnitt 4.5) erhoben wurden, in der 125 Proband_innen benötigt werden, mussten deutlich mehr Proband_innen akquiriert werden als für die Prototypizitätsstudie nötig sind.

Abweichend von der Frequenzstudie wurden für die Prototypizitätsstudie nur Proband_innen berücksichtigt, die nicht in Bayern, Baden-Württemberg, Rheinland-Pfalz oder Saarland aufgewachsen sind oder dort derzeit leben. In diesen Bundesländern werden Positionsverben mit *sein* genutzt (Scherr 2018). Um eine Interferenz von Positions- auf Bewegungsverben auszuschließen, werden Proband_innen aus diesen Bundesländern daher ausgeschlossen. Die Datenerhebung lief vom 9.1.2020 bis zum 2.2.2020. Um Proband_innen zu akquirieren, wurde der Link zur Studie an Dozent_innen sowie Vertreter_innen der Studierendenschaft¹⁸ an deutschen Universitäten versendet mit der Bitte, Studierende, die sich am Anfang und in der Mitte des Bachelorstudium befinden, zur Teilnahme aufzurufen. Zudem wurden Teilnehmer_innen über mein persönliches Umfeld akquiriert.

Nach Bereinigung der Daten bleiben 107 Proband_innen. Wie in der lexical-decision-Studie zu Frequenzeffekten war die Angabe von Metadaten freiwillig, weshalb nicht zu allen Metadaten Angaben vorliegen. Eine tabellarische Aufschlüsselung aller Metadaten ist im digitalen Anhang zu finden.

ist die Prästudie äquivalent zur Hauptstudie. Nur der aktuelle Wohnort wurde in der Prästudie nicht abgefragt. Die einzige Abweichung in den Ausschlusskriterien ist, dass in der Prästudie keine Proband_innen ausgeschlossen wurden, die aus bestimmten Bundesländern kommen. Die Metadaten der Proband_innen und die Ergebnisse der Prästudie sind im digitalen Anhang aufgeschlüsselt.

¹⁸Ich danke Alexander Werth, Heike Zinsmeister, Tanja Stevanovic, Lena Schnee, Oliver Czulo, Susanne Flach, Alan Lena van Beek, Patrick Grommes, Konstanze Marx, Ilka Lemke und Katharina Böhnert für das Weiterleiten des Studienaufrufs.

4.4 Self-paced-reading-Studie zu Form-Schematizitätseffekten

Die Proband_innen sind durchschnittlich 24 Jahre alt mit einer Standardabweichung von fünf Jahren. 15 identifizieren sich als männlich, 91 als weiblich. 29 Proband_innen kommen aus Niedersachsen, 26 aus Nordrhein-Westfalen, 14 aus Schleswig-Holstein. Die restlichen Bundesländer (bis auf die ausgeschlossenen Bundesländer) entfallen auf maximal sechs Personen, viele Proband_innen haben in mehreren Bundesländern gelebt. Derzeit wohnen 26 der Proband_innen in Nordrhein-Westfalen, 21 in Niedersachsen, 17 in Hamburg und 13 in Mecklenburg-Vorpommern und Schleswig-Holstein. Andere Bundesländer sind maximal mit sechs Personen vertreten, eine Probandin lebt im Ausland. 95 der Proband_innen gaben an, täglich zu lesen, elf wöchentlich und eine Person monatlich. 104 Proband_innen sprechen in ihrem persönlichen Umfeld Standard, zwei Dialekt und eine Person eine Mischform. 74 Proband_innen haben Abitur, 18 einen Bachelor- und acht einen Masterabschluss. 94 nutzen die rechte Hand, elf die linke, eine Person beide Hände.

Die Ergebnisse der Studie zu Prototypizitätseffekten werden in Abschnitt 5.3 erläutert. Im nächsten Abschnitt wird die self-paced-reading-Studie zu Form-Schematizitätseffekten vorgestellt.

4.4 Self-paced-reading-Studie zu Form-Schematizitätseffekten

Der Einfluss von Form-Schematizität auf Variation wird anhand des Form-Schemas I¹⁹ schwacher Maskulina untersucht. Dafür wird neben anderen Studien eine self-paced-reading-Studie durchgeführt. In der self-paced-reading-Studie lesen Proband_innen Pseudosubstantive, die dem Form-Schema schwacher Maskulina zu unterschiedlichen Graden entsprechen. Diese werden mit schwachen und starken Flexionsformen präsentiert. Die Studie ist aus einer vorherigen Studie (Schmitt 2019a) entstanden, in der konfligierende Ergebnisse zum Einfluss des Form-Schemas schwacher Maskulina auf Lesezeiten festgestellt wurden (siehe Abschnitt 3.2.3). Daher wird der fragliche Einfluss des Form-Schemas auf Lesezeiten in der vorliegenden Studie mit einem ähnlichem Design überprüft. Um die generelle Aufmerksamkeit der Proband_innen gezielt auf Wortendungen zu lenken, wurde gefragt, welche Wortform gelesen wurde (Haben Sie *des Schettosen* oder *des Schettoses* gelesen?). Auf diese Weise sollte die Wahrscheinlichkeit verringert werden, dass Proband_innen Unterschiede in der Flexion überlesen.

¹⁹Im Folgenden wird verkürzend *Form-Schema* genutzt, da das Form-Schema II für die empirische Untersuchung nicht berücksichtigt wurde.

In diesem Abschnitt werden zunächst die der Studie zugrunde liegenden Fragestellungen und Hypothesen erläutert. Anschließend wird das methodische Vorgehen (Materialien, Versuchsdesign und -ablauf, Metadaten, Probandenakquise und Proband_innen) für die Studie beschrieben.

Der Einfluss von Form-Schemata wird nicht nur durch *self-paced reading* evaluiert: Schwache Maskulina mit unterschiedlichem Form-Schematizitätsgrad werden auch als Filler in den Studien zu Frequenz- und Prototypizitätseffekten (siehe Abschnitt 4.2 und 4.3) genutzt, sodass bezüglich des Einflusses von Form-Schematizität auf Variation die Ergebnisse mehrere Studien verglichen werden können.

4.4.1 Fragestellung und Hypothesen

Die Studie geht der Frage nach, welchen Einfluss das Form-Schema schwacher Maskulina auf die Prozessierung und Produktion schwacher und starker Formen von Maskulina nimmt. Dabei wird mit Pseudosubstantiven gearbeitet, die dem Form-Schema schwacher Maskulina zu unterschiedlichen Graden entsprechen (siehe Abschnitt 3.2.3 für ausführliche Erläuterungen zu Form-Schemata bei Maskulina): Es werden Pseudosubstantive getestet, die dem Prototyp des Form-Schemas entsprechen, in der Peripherie des Form-Schemas sind, und Substantive, deren phonologische Eigenschaften schwache Deklination ausschließen, da sie zum Form-Schema starker Maskulina gehören (siehe Abschnitt 4.4.2 für genauere Erläuterungen zu den verwendeten Materialien). Im Anschluss an die *self-paced reading task* erhalten die Teilnehmer_innen einen Lückentext, in dem sie die vorgegebenen Testwörter deklinieren müssen. Folgende Hypothesen werden überprüft:

4.4.1.1 Lesezeit

1. Schwache Formen von Substantiven, die prototypische Vertreter des Form-Schemas schwacher Maskulina sind, lösen geringere Lesezeiten aus als starke Formen. Im Gegensatz dazu lösen schwache Formen von Substantiven, die dem Form-Schema starker Maskulina angehören, höhere Lesezeiten aus als starke Formen.
2. Bei Testsubstantiven, die in der Peripherie des Form-Schemas schwacher Maskulina sind, sind vergleichbare Lesezeiten für beide Formen messbar.

Bestätigen sich die Hypothesen, können die nicht existenten Unterschiede in der Lesezeit bei Substantiven in der Peripherie des Form-Schemas als ein Hinweis auf Varianten interpretiert werden: Da nicht nur eine, sondern zwei Formen

4.4 Self-paced-reading-Studie zu Form-Schematizitätseffekten

mental gefestigt sind, stellt keine der beiden Formen eine Herausforderung für die Prozessierung dar.

4.4.1.2 Produktion

1. Substantive, deren Eigenschaften prototypisch zum Form-Schema schwacher Maskulina passen, werden vornehmlich schwach dekliniert. Im Gegensatz dazu werden Substantive, die der starken Flexion angehören, vorrangig stark dekliniert.
2. Bei Substantiven, die in der Peripherie des Form-Schemas sind, werden schwache sowie starke Formen gewählt.

4.4.2 Materialien

Folgende Pseudowörter werden für die self-paced-reading-Studie genutzt:

1. Prototyp des Form-Schemas schwacher Maskulina: *der Schettose*; *der Trun-take*
2. Peripherie des Form-Schemas schwacher Maskulina: *der Knatt*
3. Form-Schema starker Maskulina: *der Grettel*

Alle Pseudowörter werden über menschliche Referenten eingeführt, um den Einfluss der Semantik auf das Form-Schema und somit auf die Flexion konstant zu halten. Die genutzten Testitems sind Köpcke (2000b) entnommen, der in seinem Produktionsexperiment mit diesen Pseudowörtern den Einfluss von Form-Schemata auf das Flexionsverhalten nachweisen konnte. Um zu verhindern, dass Assoziationen mit tatsächlich existierenden Lexemen Einfluss auf das Leseverhalten nehmen, wurden 90 Student_innen²⁰ nach den ersten drei Assoziationen mit dem jeweiligen Pseudosubstantiv gefragt. Dabei wurden neben den Testwörtern Alternativwörter (*Trilch* und *Fletter*) abgefragt. Einen ähnlichen Assoziationstest führt auch Köpcke (2000b: 160) durch. Die ausgewählten Testitems lösten dabei keine problematischen Assoziationen aus, während *Fletter* mit *Vetter* und *Trilch* mit *Knilch* assoziiert wurde. Sowohl *Vetter* als gemischt flektierendes Substantiv als auch *Knilch* als starkes Substantiv könnten Einfluss auf die Deklination nehmen, weswegen sie nicht ausgewählt wurden. Siehe Anhang H für weitere Erläuterungen zum Assoziationstest und einen Überblick über die Assoziationen.

²⁰Es handelt sich um Student_innen der Heinrich-Heine-Universität Düsseldorf sowie der Johannes Gutenberg-Universität Mainz. Ich bedanke mich bei Johanna Flick und Jessica Nowak für die Hilfe bei der Datenerhebung.

4.4.3 Versuchsdesign

Die Versuchitems werden in Texten präsentiert, die wie Lexikonartikel aufgebaut sind, siehe Beispiel 7.²¹ Die Proband_innen haben die Aufgabe, die Texte zu lesen. Nach jedem Satz²² werden sie nach einem Wort gefragt, das im Satz vorkam (Haben Sie *Amtsschreiber* oder *Amtschreiber* gelesen?). Da nach jedem Satz nach dem Aufbau eines darin enthaltenen Worts gefragt wird, wird die Aufmerksamkeit der Proband_innen auf die Endungen der Testitems gelenkt, ohne den Untersuchungsgegenstand zu verraten. Zudem werden nicht nur Wortendungen abgefragt, sondern auch Fugenelemente (*Amtsschreiber* vs. *Amtschreiber*) und Umlaute in den Distraktortexten, die ebenfalls Pseudosubstantive enthalten (*Brachte* vs. *Brächte*).

(7) Schettose, der

Betonung: Schettóse

Amtsbezeichnung am ungarischen Hof, Amtsschreiber/Diener

Der Schettose war ähnlich gestellt wie ein Amtsschreiber, zusätzlich erfüllte er jedoch niedrigere Dienste. Der Schettose wurde vom König direkt ernannt. Somit war das Amt **des Schettosen** offensichtlich ein verantwortungsvoller Posten. Es ist davon auszugehen, dass die Übertragung dieses Amtes als Ehre galt. Interessanterweise ließ sich außer dem Amtsschreiber kein vergleichbares Amt an anderen Höfen beobachten. Dementsprechend war das Amt **des Schettoses** offensichtlich ein Spezifikum des ungarischen Königshofs.

Die Texte sind als *self-paced reading task* mit *moving window* gestaltet. Um spill-over-Effekte zu minimieren, wird auf eine Wort-für-Wort-Präsentation verzichtet, stattdessen werden den Proband_innen Wortgruppen präsentiert (McDonough & Trofimovich 2012: 121–122). Die Wortgruppen entsprechen meist Präpositional- und Nominalphrasen. Verben werden nicht einzeln gelesen, sondern gemeinsam mit einer weiteren Konstituente. Adverbien werden zu anderen Phrasen hinzu genommen oder einzeln präsentiert. Die Texte sind inklusive der Wortgruppeneinteilung und der Fragen zu den Testwörtern und Distraktoren (z. B. Haben Sie *des Schettoses* oder *des Schettosen* gelesen?) in Anhang G.1 zu finden.

²¹Im Beispiel werden zunächst schwache und dann starke Genitivformen von *Schettose* genutzt. Diese Reihenfolge war jedoch nicht fest, stattdessen wurde pseudorandomisiert zwischen dieser und der gegensätzlichen Reihenfolge gewechselt, wie unten erläutert wird.

²²Vereinzelt werden zwei Sätze gelesen, bevor nach dem Aufbau eines Worts in den Sätzen gefragt wird. Dies ist vor allem bei sehr kurzen Sätzen der Fall.

4.4 Self-paced-reading-Studie zu Form-Schematizitätseffekten

Die Testitems werden in einer vergleichbaren syntaktischen Umgebung präsentiert. So sind sie stets Teil der dritten Konstituente im Satz und stellen ein Genitivattribut dar. Genitivformen wurden gewählt, da hier die starken und schwachen Formen gleich lang sind. Die einzige Ausnahme davon bildet das Testsubstantiv *Truntake*, das im Dativ präsentiert wird. Auf diese Weise ist ein Vergleich von Genitiv- und Dativformen bei zwei prototypischen Vertretern des Form-Schemas schwacher Maskulina (*Schettose*, *Truntake*) möglich. Die Beispiele 8 bis 11 geben einen Überblick über die Sätze, in die die Testsubstantive eingeflochten sind:

(8) *Schettose*

- a. Somit war **das Amt des Schettosen/*Schettoses** offensichtlich ein verantwortungsvoller Posten.
- b. Dementsprechend war **das Amt des Schettosen/*Schettoses** offensichtlich ein Spezifikum des ungarischen Königshofs.

(9) *Knatt*

- a. Mit der Zeit setzte sich **der Beruf des Knatts/Knatten** anscheinend auch in anderen Regionen durch.
- b. Auch in der Stadt setzte sich **der Beruf des Knatts/Knatten** anscheinend nie durch.

(10) *Grettel*

- a. Deswegen ist **das Amt des Grettels/*Gretteln** in vielen Vereinen nur für langjährige Mitglieder vorgesehen.
- b. Dementsprechend ist **das Amt des Grettels/*Gretteln** in vielen Vereinen nicht besetzt.

(11) *Truntake*

- a. Jedoch kam **dem Truntaken/*TruntakeØ** üblicherweise keine Urteilskraft zu.
- b. Jedoch kam **dem Truntaken/*TruntakeØ** üblicherweise auch eine unangenehme Aufgabe zu: der Schuldspruch.

Die Testitems werden jeweils als Bigram (bspw. *des Gretteln* bzw. *des Grettels*) präsentiert. Da spill-over-Effekte möglich sind, wird das Element nach dem Versuchitem jeweils konstant gehalten (*offensichtlich, anscheinend, in vielen Vereinen, üblicherweise*), sodass die Lesezeiten der Elemente verglichen werden können. Nach den Sätzen, die Testitems enthalten, wird jeweils nach den Endungen der Testsubstantive gefragt (bspw. Haben Sie *des Knatten* oder *des Knatts* gelesen?).

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

Die Reihenfolge der Genitivformen ist pseudo-randomisiert: Ein Teil der Proband_innen liest zunächst die zu erwartenden Formen (*des Schettosen, des Knatts, des Grettels, dem Truntaken*) und dann nicht-erwartbare Formen (*des Schettoses, des Knatten, des Gretteln, den Truntake*). *Des Knatten* wurde dabei als weniger erwartbar als *des Knatts* eingestuft, weil *-n* die typeninfrequente Flexion darstellt. Der andere Teil der Proband_innen liest die Testitems in umgekehrter Reihenfolge. Die Reihenfolgenrandomisierung dient dazu, Reihenfolge-Effekte zu vermeiden: Proband_innen neigen bei *self-paced reading tasks* dazu, schneller zu werden (Ackermann & Zimmer 2017: 166–168), somit könnte die Lesezeit der Items durch die Reihenfolge beeinflusst werden. Auch die Reihenfolge der Formen in den Fragen ist auf die Gruppen verteilt: In der ersten Gruppe wird in der Frage nach dem Aufbau der Testitems stets die erwartbare Form zuerst genannt (Haben Sie *des Schettosen* oder *des Schettoses* gelesen?), in der zweiten Gruppe die nicht-erwartbare Form (Haben Sie *des Schettoses* oder *des Schettosen* gelesen?).

Im Gegensatz zu den Genitivformen der Testitems sind die Texte selbst nicht randomisiert, sondern in konstanter Reihenfolge gehalten. Dabei sind die Testitems nach potentieller Salienz sortiert: *Knatt* wird als erstes Testitem genutzt, da hier aufgrund der Peripherie des Form-Schemas keine Flexion präferiert werden sollte. *Schettose* wird als zweites Testitem präsentiert, darauf folgt *Grettel* und schließlich *Truntake*. Da schwache Maskulina weniger typenfrequent sind als starke, wird davon ausgegangen, dass die starke Deklination von schwachen Maskulina (*des Schettoses*) weniger salient ist als die schwache Deklination von starken Maskulina (*des Gretteln*) (siehe Abschnitt 3.2.1 zu Typen- und Tokenfrequenz von Maskulina). *Truntake* wurde nicht aufgrund seiner Salienz als letztes Item genutzt, sondern um zu verhindern, dass die dreisilbigen Testitems dicht aufeinander folgen.

Neben den Testtexten werden Distraktortexte genutzt. Hierfür wird auf Pseudosubstantive aus Köpcke (2000b) zurückgegriffen. Die Distraktoren (*die Taff, die Schirr, die Bracht*) werden mit variierenden Pluralformen (*die Täffe/Taffe/Taffs*) eingeflochten. Die Fragen zum Wortaufbau lenken die Aufmerksamkeit der Proband_innen auf die Flexionsformen der Distraktoren (*Brachte/Brächte/Brächten*) (siehe Anhang G.2 für einen Überblick über die Distraktortexte). Distraktor- und Testsätze wechseln sich im Experiment regelmäßig ab, nur der Text zum Testwort *Truntake* folgt direkt auf den Text zu *Grettel* (siehe Anhang G). Vor den Test- und Distraktortexten wird ein Beispielsatz präsentiert, um die Proband_innen an die Aufgabe zu gewöhnen. Auch im Beispiel wird nach dem Wortaufbau (*Haben Sie Pizzas oder Pizzen gelesen?*) gefragt, hierbei wird die Aufmerksamkeit bereits allgemein auf Wortendungen gelenkt und die Proband_innen mit der Aufgabe vertraut gemacht.

4.4 Self-paced-reading-Studie zu Form-Schematizitätseffekten

Nach der *self-paced reading task* werden den Proband_innen Lückensätze vorgelegt, in die Genitivformen der Pseudosubstantive ergänzt werden sollen. Die Sätze entsprechen dabei jeweils den Sätzen aus der *self-paced reading task* (siehe Anhang G.3 für eine Auflistung der Lückensätze). Auch die Distraktoren sind in das Produktionsexperiment eingeflochten. Im Anschluss an das Produktionsexperiment werden die Proband_innen gefragt, ob ihnen Varianten in den zu lesenden Texten aufgefallen sind, um überprüfen zu können, ob die Varianten in der *self-paced reading task* die Antworten im Produktionsexperiment beeinflusst haben.

4.4.4 Versuchsablauf

Anders als die lexical-decision- und sentence-maze-Studie wurde die self-paced-reading-Studie nicht online durchgeführt. Daher wurden die Proband_innen von mir sowie von studentischen Hilfskräften betreut, die vorab in den Versuchsablauf eingewiesen wurden. Die Betreuungspersonen hielten sich bei der Durchführung an einen schriftlichen Ablaufplan, der neben den Instruktionen an die Hilfskräfte²³ auch generelle Anweisungen zur Studie enthielt, die den Proband_innen mündlich gegeben wurden. Diese wurden abgelesen, um die Studie für alle Proband_innen möglichst konstant zu halten (Höhle 2010: 215). Die exakten Anweisungen zur *self-paced reading task* erhielten die Proband_innen per Bildschirm.

Vor den Distraktor- und Testtexten werden ein Begrüßungstext und ein kurzes Beispiel gezeigt (siehe Anhang G für Begrüßung und Beispiel), um die Proband_innen an das *self-paced reading* zu gewöhnen. Im Begrüßungstext werden die Proband_innen gebeten, die Texte in einer komfortablen Geschwindigkeit zu lesen und auf die Schreibung der Wörter zu achten. Nach Begrüßung und Beispiel werden die Proband_innen per Bildschirm aufgefordert, Verständnisfragen zu stellen, falls etwas unklar geblieben sein sollte. Die Metadaten wurden nach dem Lesen der Texte mündlich abgefragt und direkt in eine Excel-Tabelle eingetragen. Abschließend wurde den Proband_innen das Produktionsexperiment vorgelegt, das sie handschriftlich ausfüllten.

Die Studie wurde mithilfe eines Laptops (Windows 7) und eines Response Pads (RB-540) durchgeführt, dabei wurde das Programm DMDX genutzt. Dieses Programm erlaubt die Bestimmung von Lesezeiten mit einer Genauigkeit von Millisekunden (Forster & Forster 2003). Das Response Pad wurde hierfür im Keyboardmodus genutzt, sodass die Messgenauigkeit bei zehn bis zwölf ms liegt (Herstellerangabe unter https://cedrus.com/rb_series/). Das Response Pad

²³Ich bedanke mich bei Jonas Fehn und Patricia Pfeffer für die Unterstützung bei der Datenerhebung.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

hat den Vorteil, dass die Proband_innen nur aus vier Tasten auswählen müssen und nicht aus über 100 bei einer normalen Tastatur.

Die Messungen für alle Studien fanden in einem ruhigen Raum statt. Um Störgeräusche von außen zu minimieren, erhielten die Proband_innen zusätzlich einen Gehörschutz. Den Proband_innen wurde am Ende der Studie ein Aufklärungsformular vorgelegt und das Einverständnis eingeholt, die Daten anonymisiert zu verwenden.

4.4.5 Metadaten

Für die Studie wurden dieselben Metadaten abgefragt wie für die lexical-decision-Studie zum Einfluss der Frequenz auf Variation (siehe Abschnitt 4.2). Nur der aktuelle Wohnort wurde nicht abgefragt, da die Proband_innen Bamberger Student_innen darstellen und somit ein Wohnort in Bamberg und Umgebung angenommen werden kann. Zudem wurden die Fragen nach der präferierten sprachlichen Varietät und nach Dialektkenntnissen als eine Frage zusammengefasst und anders als in den anderen Studien nach dem vermuteten Untersuchungsgegenstand sowie nach Auffälligkeiten im Experiment gefragt, da die Metadaten am Ende des Experiments erhoben wurden.

Wie in den anderen Studien dienen die Metadaten vornehmlich dazu, sicherzustellen, dass die untersuchte Gruppe möglichst homogen ist. Die Frage nach der Leseintensität ist für *self-paced reading* relevant, da davon auszugehen ist, dass ungeübte Leser_innen längere Lesezeiten erzielen als geübte (Jegerski 2014: 43). Generell ist jedoch anzunehmen, dass es sich bei den Proband_innen um geübte Leser_innen handelt, da Student_innen getestet werden. Aufgrund des within-subject-Designs sollten allerdings auch ungeübte Leser_innen das Ergebnis kaum verzerren. Zudem sind Proband_innen als zufällige Effekte in das statistische Modell eingebunden.

Der vermutete Untersuchungsgegenstand und Strukturen, die Proband_innen im Experiment aufgefallen sind, werden abgefragt, um Proband_innen, die das Untersuchungsinteresse durchschaut haben, ausschließen zu können. Zusätzlich zu den Metadaten werden generelle Kommentare festgehalten, um Proband_innen (z. B. aufgrund von missverstandenen Anweisungen) ggf. ausschließen zu können.

4.4.6 Probandenakquise und Proband_innen

Als Proband_innen wurden Studierende der Germanistik an der Otto-Friedrich-Universität Bamberg akquiriert. Zusätzlich nahmen Proband_innen aus meinem persönlichen Umfeld teil. Um weitgehend auszuschließen, dass Fachwissen die

4.4 Self-paced-reading-Studie zu Form-Schematizitätseffekten

Studienergebnisse beeinflusst, wurden vornehmlich Studierende am Anfang des Bachelorstudiums für die Studie geworben, allerdings nahmen auch einige Masterstudent_innen an der Studie teil. Die Akquise lief über Seminare ab, die im Sommersemester 2018 an der Otto-Friedrich-Universität Bamberg angeboten wurden. Alle Proband_innen wurden darauf hingewiesen, dass die Teilnahme an den Studien freiwillig ist, und ihr Einverständnis, dass die Daten anonymisiert gespeichert und ausgewertet werden, wurde schriftlich eingeholt. Das Einverständnis konnte bis zum Ende der Datenerhebung zurückgezogen werden.

Es werden nur Teilnehmer_innen berücksichtigt, die Deutsch als L1 sprechen, monolingual sind und keine Legasthenie haben. Um sicherzustellen, dass die Teilnehmer_innen die Texte aufmerksam gelesen haben, wird das Antwortverhalten bei den Füllsätzen evaluiert. In den Füllsätzen sollten insgesamt sieben Fragen beantwortet werden. Proband_innen, die drei oder mehr dieser Sätze falsch beantwortet haben, werden ausgeschlossen. Außerdem werden Proband_innen ausgeschlossen, die bei den korrespondierenden Testitems (bspw. *Gretteln* und *Grettels*) beide Fragen falsch beantwortet haben, d. h. jeweils angaben, dass sie die gegensätzliche Flexion gelesen haben. Wie in den anderen Studien werden Proband_innen ausgeschlossen, deren Reaktionszeiten mehr als drei Standardabweichungen über oder unter dem Mittelwert eines Testitems liegen. Da in der self-paced-reading-Studie der direkte Vergleich zwischen den Flexionsformen relevant ist, werden hierbei starke und schwache Formen separat betrachtet.

Die Daten für die Studie wurden vom 04.6.2018 bis 29.06.2018 an der Otto-Friedrich-Universität Bamberg erhoben. Insgesamt wurden 61 Proband_innen akquiriert, von denen 54 Proband_innen in die Analyse aufgenommen wurden.

Alle Proband_innen machten zu allen Metadaten Angaben. Eine tabellarische Aufschlüsselung der Metadaten sind im digitalen Anhang. Das Durchschnittsalter beträgt 25 Jahre mit einer Standardabweichung von acht Jahren. 45 der Proband_innen identifizieren sich als weiblich, acht als männlich, eine Person machte keine Angabe. Die Proband_innen stammen vornehmlich aus Bayern (28). Die meisten Proband_innen sprechen in ihrem familiären Umfeld vornehmlich Standard (21), weitere 17 Proband_innen geben an, Standard mit dialektalem Einschlag zu sprechen. Die restlichen Proband_innen nannten einen Dialekt als bevorzugte sprechsprachliche Varietät. Die meisten Proband_innen nannten Abitur als höchsten Bildungsabschluss (40), eine Person Fachabitur oder mittlere Reife und die restlichen 12 höhere Abschlüsse. Die meisten Proband_innen lesen täglich (49), nur fünf geben an, gelegentlich zu lesen. 48 Proband_innen sind Rechtshänder_innen, davon wurde eine Person auf rechts umgeschult. Fünf Proband_innen sind Linkshänder_innen, eine Person nutzt beide Hände gleichwertig.

Die Ergebnisse der self-paced-reading-Studie zu Form-Schematizitätseffekten werden in Abschnitt 5.4 vorgestellt. Im folgenden Abschnitt wird das methodische Vorgehen in der lexical-decision- und der sentence-maze-Studie zu Form-Schematizitätseffekten erläutert.

4.5 Lexical-decision- und sentence-maze-Studie zu Form-Schematizitätseffekten

In diesem Abschnitt werden die lexical-decision- sowie die sentence-maze-Studie zum Einfluss von Form-Schematizität auf Variation beschrieben. Da die beiden Studien viele Ähnlichkeiten aufweisen, werden sie gemeinsam vorgestellt. Beide Studien arbeiten mit starken und schwachen Formen real existierender Substantive, die dem Form-Schema schwacher Maskulina zu unterschiedlichem Grad entsprechen. In der lexical-decision-Studie bewerten die Proband_innen die starken und schwachen Formen nach ihrer Bekanntheit, in der sentence-maze-Studie wählen die Proband_innen zwischen starken und schwachen Formen. Im Folgenden werden die Fragestellungen und Hypothesen sowie das methodische Vorgehen (Materialien, Versuchsdesign und -ablauf, Metadaten, Probandenakquise und Proband_innen) vorgestellt.

4.5.1 Fragestellung und Hypothesen

Die Studien gehen der Frage nach, welchen Einfluss das Form-Schema schwacher Maskulina auf Reaktionszeiten und die Wahl von Flexionsformen nimmt. Anders als in der self-paced-reading-Studie werden real existierende Substantive getestet. Die Testsubstantive entsprechen entweder dem Prototyp des Form-Schemas, sind in der Peripherie des Form-Schemas oder stellen starke Substantive dar. Dabei werden keine starken Maskulina auf *-el* genutzt, sondern monosyllabische wie *der Vogt* (siehe Abschnitt 4.5.2 für ausführliche Erläuterungen zu den Materialien). Folgende Hypothesen werden überprüft:

4.5.1.1 Reaktionszeiten

1. In der *lexical decision task* ist eine Interaktion zwischen Deklinationsklasse und Flexion zu beobachten: Schwache Formen lösen bei schwachen Maskulina unabhängig von deren Passgenauigkeit zum Form-Schema niedrige Reaktionszeiten aus, bei starken Maskulina hingegen hohe. Das Gegenteil wird für starke Formen angenommen. Es wird also nur ein Effekt der Deklinationsklasse, aber kein Form-Schematizitätseffekt erwartet.

4.5 Lexical-decision- und sentence-maze-Studie zu Form-Schematizität

2. In der *sentence maze task* evozieren Substantive, die dem Form-Schema peripher angehören, geringere Reaktionszeiten als Substantive, die prototypisch zum Form-Schema gehören oder stark sind.

Diese Hypothesen basieren auf den Ergebnissen der Prästudien. Die Prästudie ist im digitalen Anhang dokumentiert. Ursprünglich wurde für beide Studien erwartet, dass die Reaktionszeiten bei Substantiven niedrig sind, die eindeutig zum Form-Schema passen oder der starken Flexion angehören, aber erhöht bei Substantiven, die nur peripher dem Form-Schema angehören, da hier zwei Formen möglich sind und die Entscheidung daher verlangsamt sein könnte. Auch wenn in den Studien auf Tokenfrequenz kontrolliert wird (siehe Abschnitt 4.5.2), könnten sich die niedrigen Reaktionszeiten für die peripheren Substantive in der *sentence-maze-Studie* durch unterschiedliche Tokenfrequenzen erklären lassen, dies wird in der explorativen Analyse der Daten getestet. Ist dies der Fall, könnte es sein, dass Tokenfrequenz die zu erwartenden Form-Schematizitätseffekte überdeckt. Sind hingegen höhere Reaktionszeiten für die Substantive in der Peripherie des Form-Schemas messbar, kann dies als ein Hinweis auf Varianten interpretiert werden: Es existiert nicht nur eine einzige, stark gefestigte Form, stattdessen sind es zwei Formen. Hierdurch verlangsamt sich die Beurteilung der Formen.

4.5.1.2 Antwortverhalten

1. Passen die phonologischen Eigenschaften prototypisch zum Form-Schema schwacher Maskulina, wird die schwache Form *-(e)n* in der *lexical-decision-Studie* als bekannt bewertet bzw. in der *sentence-maze-Studie* gewählt. Bei Substantiven der starken Deklinationsklasse wird *-(e)n* hingegen als unbekannt bewertet bzw. nicht gewählt. Das Gegenteil wird für die Bekanntheit/Wahl der starken Form *-(e)s* angenommen.
2. In der *lexical-decision-Studie* weisen schwache Formen von Substantiven aus der Peripherie des Form-Schemas ähnliche Bekanntheitswerte auf wie schwache Formen von prototypischen Vertretern des Form-Schemas. Die starken Formen werden aber bei Substantiven in der Peripherie ebenfalls als bekannt bewertet. Dabei werden für die schwachen Formen von Substantiven in der Peripherie des Form-Schemas hohe Bekanntheitswerte erwartet (über 90 %), für starke hingegen nur eine Bekanntheit von ca. 50 %.
3. In der *lexical-decision-Studie* antworten Teilnehmer_innen bei Substantiven, die zum Form-Schema gehören, und Substantiven, die der starken De-

klination angehören, konsistent (Bekanntheit einer Form und Unbekanntheit der anderen). In der Peripherie des Form-Schemas sind hingegen auch inkonsistente Antworten möglich (Bekanntheit/Unbekanntheit beider Formen).

4. In der sentence-maze-Studie evozieren die Testsubstantive in der Peripherie des Form-Schemas ein ähnliches Antwortverhalten wie die Testsubstantive, die dem Form-Schema prototypisch entsprechen.

Die Hypothesen zu den Substantiven in der Peripherie des Form-Schemas basieren auf den Ergebnissen der Prästudien. Ursprünglich wurde erwartet, dass sowohl hinsichtlich schwacher als auch hinsichtlich starker Formen Variation in den Antworten messbar ist. Die auf Basis der sentence-maze-Prästudie erwartete ausbleibende Variation bei Substantiven in der Peripherie des Form-Schemas lässt sich durch das Studiendesign der sentence-maze-Studie erklären: Da die mental gefestigte Form *-(e)n* neben der mental noch nicht so stark gefestigten Alternative *-(e)s* präsentiert wird, kann die gefestigte Form die weniger gefestigte leichter dominieren.

Die hier vorgestellten Hypothesen sowie das methodische Vorgehen inklusive der geplanten statistischen Auswertung wurden vor der Datenerhebung registriert. Die Registrierungen lassen sich unter <https://osf.io/yx3tv> (*lexical decision task*) und <https://osf.io/bhxpz> (*sentence maze task*) aufrufen. Auf die statistische Auswertung wird zusammen mit den Ergebnissen in Abschnitt 5.4 eingegangen.

4.5.2 Materialien

Für die Studien werden Testitems genutzt, die jeweils im Genitiv präsentiert werden, sodass die starke und die schwache Form gleich lang sind. Die Zahlen in Klammern geben die relative Frequenz der Genitivformen der Substantive im Jahr 2018 an. Die Frequenz der Genitivformen wurde anhand der Wortverlaufskurven im DWDS (Berlin-Brandenburgische Akademie der Wissenschaften 2019) bestimmt.

1. Prototyp des Form-Schemas:
des Kollegen (0,58), *des Neffen* (0,05), *des Schützen* (0,41), *des Franzosen* (1,66), *des Gesellen* (0)
2. Peripherie:
des Grafen (0,36), *des Helden* (0,53), *des Zaren* (0,48), *des Fürsten* (0,29), *des Nachbarn* (0,86)

4.5 Lexical-decision- und sentence-maze-Studie zu Form-Schematizität

3. stark:

des Diebes (0,04), *des Freundes* (0,31), *des Vogts* (0), *des Kerls* (0), *des Feindes* (0,46)

In der sentence-maze-Studie werden jeweils nur die ersten drei Testwörter pro Ausprägung genutzt.

Für die Bestimmung der Frequenz der Substantive wurde in den Wortverlaufskurven des DWDS nach dem Genitiv der Wörter gesucht. Dabei war jeweils die frequentere Genitivform *-(e)n*, *-es* oder *-s* ausschlaggebend.²⁴ Bei der Suche wurde ein Determinierer vorangestellt (bspw. *@des @Kollegen*), um Fehltreffer aufgrund von Homonymien zu anderen Kasus- und Pluralformen zu vermeiden. Die Abfrage im DWDS wurde am 28.10.2019 durchgeführt.

Wie in der self-paced-reading-Studie (siehe Abschnitt 4.4) wird die Belebtheit konstant gehalten, indem nur Substantive genutzt werden, die auf Menschen referieren. Anders als in der self-paced-reading-Studie schließen die starken Substantive die schwache Flexion jedoch durch ihre phonologischen Eigenschaften nicht aus, denn sie ähneln Substantiven in der Peripherie des Form-Schemas.

Alle Substantive haben eine Frequenz von bis zu zwei Belegen pro Million Token im Jahr 2018. Auf diese Weise können Tokenfrequenzeffekte minimiert werden. Zwischen den Ausprägungen existieren jedoch leichte Frequenzunterschiede: Drei Substantive, die dem Prototyp des Form-Schemas entsprechen (*Schütze*, *Kollege*, *Franzose*), weisen eine relative Frequenz von über 0,4 auf. Die anderen zwei Substantive dieser Ausprägung (*Neffe*, *Geselle*) bewegen sich im Frequenzbereich von bis zu 0,05. Die Substantive in der Peripherie des Form-Schemas zeigen eine Frequenz von 0,3 bis 2,5. Ähnlich frequent sind zwei der starken Substantive (*Freund*, *Feind*). Die restlichen drei Substantive dieser Ausprägung (*Vogt*, *Kerl*, *Dieb*) sind mit einer Frequenz von bis zu 0,04 deutlich weniger frequent als die Substantive der anderen Ausprägungen. *Geselle* und *Neffe* sind jedoch ähnlich frequent wie die starken Substantive.

4.5.3 Versuchsdesign

Die Studien nutzen ein within-subject-Design mit zwei Faktoren: Form-Schematizität (mit drei Ausprägungen: Prototyp des Form-Schemas, Peripherie des Form-Schemas, stark) und Flexion (mit zwei Ausprägungen: stark, schwach). Pro Ausprägung werden fünf bzw. drei Testwörter genutzt. Die Proband_innen müssen

²⁴Die infrequenteren Formen lagen für *Nachbar* bei einer relativen Frequenz von 0,02 (*des Nachbars*), für *Freund*, *Dieb* und *Feind* bei 0,01 (*des Freunds*, *Diebs*, *Feinds*). Bei den anderen Testsubstantiven haben die alternativen Genitivformen eine relative Frequenz von 0.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

in der lexical-decision-Studie die Wortformen danach bewerten, ob sie sie kennen. In der sentence-maze-Studie haben die Proband_innen die Wahl zwischen schwacher und starker Flexion (siehe Abschnitt 4.2.3 und 4.3.3 zum generellen Versuchsdesign der Studien).

Die Testitems sind jeweils im Genitiv Singular. In der lexical-decision-Studie stehen sie jeweils mit Definitartikel. Als starke Endung wird i. d. R. die kurze Endung (-s) angezeigt. Wenn die lange Endung (-es) frequenter ist als die kurze, wie bei *Dieb*, *Freund* und *Feind*, wird diese genutzt. In der sentence-maze-Studie sind die Testitems in folgende Sätze eingeflochten:

(12) Prototyp des Form-Schemas

- a. Ich nehme an, dass die Arbeit **meines Kollegen/meines *Kolleges** wertgeschätzt wird.
- b. Ich habe gesehen, wie der Pfeil **des Schützen/des *Schützes** ins Schwarze traf.
- c. Mein Freund war nicht da, weil er zum Geburtstag **seines Neffen/seines *Neffes** gegangen ist.

(13) Peripherie des Form-Schemas

- a. Mein Neffe mag Superman nicht, weil ihm der Anzug **des Helden/des Helds** nicht gefällt.
- b. Sie konnte nicht fassen, dass das Anwesen **des Grafen/des Grafs** so groß ist wie ein Schloss.
- c. Ihm war nicht bewusst, dass die Macht **des Zaren/des Zars** einmal derart groß gewesen war.

(14) stark

- a. Er fährt nochmal zurück, weil er die Jacke **seines Freundes/seines *Freunden** aus Versehen eingesteckt hat und dieser sie auf der Arbeit benötigt.
- b. Der Krimi-Fan ahnt schon früh, dass der Plan **des Diebes/des *Dieben** zu simpel ist und nicht klappt.
- c. Er macht einen Ausflug, um die Burg **des Vogts/des *Vogten** zu betrachten.

Um die Testsätze ähnlich zu gestalten wie die Sätze zur Auxiliarselektion von *haben* und *sein*, werden Satzgefüge genutzt. Die Testitems sind jeweils als Genitivattribut in die Testsätze eingeflochten (... *weil er zum Geburtstag seines Neffen gegangen ist*). Sie stehen allesamt im Singular und werden mit einem

4.5 Lexical-decision- und sentence-maze-Studie zu Form-Schematizität

bestimmten Artikel bzw. mit einem Possessivpronomen (*ihres Neffen/*s*) determiniert. Die Proband_innen haben jeweils die Wahl zwischen starker und schwacher Endung (*des Helden/des Helds*). Der vorab durchgeführte Plausibilitätstest mit 58 Proband_innen war für alle Testsätze zu schwachen Maskulina unauffällig.²⁵

In der lexical-decision-Studie sind pro Block sieben bis acht Testsubstantive eingeflochten, dementsprechend finden sich pro Ausprägung zwei bis drei Testsubstantive in einem Block. Starke und schwache Formen eines Testsubstantivs kommen nicht in einem Block vor. In der sentence-maze-Studie sind pro Experimentblock zwei Testsätze enthalten. Diese entstammten verschiedenen Form-Schematizitätsausprägungen. Für nähere Erläuterungen zum Aufbau der Experimentblöcke in den Studien siehe Abschnitt 4.2.3 und 4.3.3. Die Tabellen D.1 und E.1 in Anhang D.1 und E.1 schlüsseln den Aufbau der Blöcke in den Studien auf, Tabelle E.1 bietet zudem einen Überblick über die Einteilung der Testsätze in die zu lesenden Abschnitte.

4.5.4 Versuchsablauf, Metadaten, Probandenakquise, Proband_innen

Der Versuchsablauf, die erhobenen Metadaten und die Probandenakquise entsprechen den Erläuterungen zur lexical-decision-Studie zu Frequenzeffekten (Abschnitt 4.2) bzw. zur sentence-maze-Studie zu Prototypizitätseffekten (Abschnitt 4.3). Die Stichprobengrößen von 50 (*lexical decision*) und 125 (*sentence maze*) wurde wiederum anhand einer Datensimulation berechnet, die auf Prästudien²⁶ basiert.

In der Prästudie zur lexical-decision-Studie zeigte sich eine Interaktion zwischen Deklinationsklasse und Flexionsform: Schwache Formen von schwachen Substantiven werden schneller beurteilt als starke Formen. Das Gegenteil ist bei starken Substantiven der Fall. Unterschiede zwischen prototypischen und peripheren Vertretern des Form-Schemas schwacher Maskulina waren nicht zu erkennen. Die angestrebte Stichprobengröße sollte also groß genug sein, um die

²⁵Alle Sätze wurden mit mindestens 3 auf einer Skala von 1 (nicht plausibel) bis 5 (plausibel) bewertet. Eine kleine Änderung wurde nach dem Test vorgenommen: *Gewertschätzt* wurde zu *wertgeschätzt*. Weitere Erläuterungen zum Plausibilitätstest sowie die Bewertungen der Proband_innen sind in Anhang F.

²⁶Die Stichprobengröße der Prästudien betrug 20 Proband_innen in der lexical-decision-Studie und 27 Proband_innen in der sentence-maze-studie nach Ausschluss von Proband_innen, die die Kriterien nicht erfüllten. Die Prästudien haben denselben Versuchsablauf wie die Hauptstudien. Nur der aktuelle Wohnort wurde in den Prästudien nicht abgefragt. Die Proband_innen wurden anhand derselben Kriterien wie in den Hauptstudien aus der Stichprobe ausgeschlossen.

4 Psycholinguistische Betrachtung der Einflussfaktoren: Methode

Interaktion zwischen Deklinationsklasse und Flexionsform messen zu können. Unter der Annahme, dass die Schätzung aus der Prästudie robust ist, sollten 50 Proband_innen bei einem α -Level von 0,01 für eine Teststärke von mindestens 0,8 ausreichend sein (für den Schätzwert der Ausprägung FORM-SCHEMA mit schwachen Formen im Vergleich zu den Schätzwerten der Ausprägung FORM-SCHEMA und PERIPHERIE mit starken Formen sowie der Ausprägung STARK mit schwachen und starken Formen). Insgesamt wurden 100 Proband_innen akquiriert, da in der Prästudie mit 12 von 32 Proband_innen 38 % der Teilnehmer_innen ausgeschlossen werden mussten.

In der Prästudie zur sentence-maze-Studie war erkennbar, dass die Substantive aus der Peripherie des Form-Schemas niedrigere Reaktionszeiten evozierten als Substantive der anderen Ausprägungen. Um hierfür eine Teststärke von 0,8 (für den Schätzwert der Ausprägung FORM-SCHEMA im Vergleich zum Schätzwert der Ausprägung PERIPHERIE) bei einem α -Level von 0,01 zu erreichen, sollte die Stichprobengröße 125 Proband_innen umfassen. Insgesamt werden 170 Proband_innen akquiriert, da in der Prästudie mit fünf von 32 Proband_innen 16 % ausgeschlossen wurden.

Insgesamt umfasst die Stichprobe der lexical-decision-Studie nach Ausschluss von Proband_innen 57 Teilnehmer_innen. Die Proband_innen der lexical-decision-Studie unterscheiden sich dabei nur minimal von denen der Frequenzstudie. Der Unterschied zur Frequenzstudie ergibt sich dadurch, dass Proband_innen jeweils neben den anderen Ausschlusskriterien auch anhand ihrer Reaktionszeit (mehr als drei Standardabweichungen über bzw. unter dem Mittelwert der jeweiligen Testitems) ausgeschlossen wurden. In der Stichprobe für die lexical-decision-Studie zu Form-Schematizitätseffekten sind daher mit 57 Proband_innen vier Proband_innen mehr als in der Frequenzstudie. Aufgrund der hohen Deckungsgleichheit mit der Frequenzstudie werden die Proband_innen an dieser Stelle nicht gesondert vorgestellt, siehe Abschnitt 4.2.6 für einen Überblick über die Proband_innen.

Die Stichprobe der sentence-maze-Studie umfasst nach Bereinigung der Daten 132 Proband_innen. Hinsichtlich der sentence-maze-Studie zu Form-Schematizität ergibt sich eine Abweichung von der Prototypizitätsstudie, da bei der Prototypizitätsstudie Proband_innen aus Bayern, Baden-Württemberg, Rheinland-Pfalz und dem Saarland ausgeschlossen wurden (siehe Abschnitt 4.3.5). Dies war für die Form-Schematizitätstudie nicht nötig.²⁷ Aufgrund dessen sind in der Stich-

²⁷Zwar diskutiert Ronneberger-Sibold 2020 einen Einfluss von mittel- und nordbayerischen Dialekten auf die Wahrnehmung des Genitivs bei schwachen Maskulina, sie untersucht jedoch Studierende mit stark dialektalem Sprachverhalten. Daher wurde für diese Studie davon abgesehen, pauschal Proband_innen aus Bayern auszuschließen.

4.5 *Lexical-decision- und sentence-maze-Studie zu Form-Schematizität*

probe mit 132 Proband_innen 25 Proband_innen mehr als in der Prototypizitätsstudie. Daher wird die Stichprobe hier noch einmal gesondert vorgestellt. Die Angaben zu Metadaten waren freiwillig, daher fehlen für einige Metadaten Angaben. Eine ausführliche Aufschlüsselung der Metadaten ist im digitalen Anhang zu finden.

Die Proband_innen der sentence-maze-Studie zu Form-Schematizität sind im Schnitt 25 Jahre alt mit einer Standardabweichung von 7,6 Jahren. 110 identifizieren sich als weiblich, 21 als männlich. 30 Proband_innen kommen aus Niedersachsen, 26 aus Nordrhein-Westfalen, 16 aus Schleswig Holstein. Die restlichen Bundesländer sind mit bis zu acht Proband_innen vertreten, einige Proband_innen gaben auch an, in mehreren Bundesländern aufgewachsen zu sein. 27 Proband_innen leben derzeit in Nordrhein-Westfalen bzw. Hamburg, 20 in Niedersachsen, 15 in Schleswig-Holstein und 14 in Mecklenburg-Vorpommern. Die anderen Bundesländer entfallen auf maximal acht Proband_innen, drei Proband_innen leben im Ausland. 117 geben an, täglich zu lesen, zwölf lesen wöchentlich und zwei monatlich. 123 Proband_innen sprechen in familiären Kontexten Standard, acht Dialekt und eine Person eine Mischform. 84 Proband_innen haben Abitur, 25 einen Bachelor- und zwölf einen Masterabschluss. 117 der Proband_innen nutzen vornehmlich die rechte Hand, 14 die linke, eine Person gab an, beide Hände gleichermaßen zu nutzen. Im nächsten Kapitel werden die Ergebnisse der Studien vorgestellt und diskutiert.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

In diesem Kapitel werden die Ergebnisse der einzelnen Studien vorgestellt. Für die Datenaufbereitung wurden Python (van Rossum & Drake 2009)¹, Version 3.8, sowie R (R Core Team 2019), Version 3.6.0, und RStudio (RStudio Team 2016), Version 1.1.453, genutzt. R und RStudio wurden auch für die Datenauswertung inklusive der Visualisierung der Daten verwendet.²

Die in diesem Kapitel vorgestellten Ergebnisse beziehen sich auf die Daten aus den Hauptstudien. Eine analoge Auswertung der Daten aus den Prästudien³ befindet sich im digitalen Anhang. Erläuterungen zur Zusammensetzung der Stichprobe und zur Bereinigung der Daten sind in Kapitel 4 zu finden, in dem auch das methodische Vorgehen vorgestellt wurde. Bevor die Ergebnisse diskutiert werden, werden die in der Analyse genutzten statistischen Modelle, (generalisierte) gemischte lineare Modelle und *random forests*, vorgestellt. Im Anschluss daran werden zunächst die Ergebnisse der Studie zu Frequenzeffekten analysiert, darauf folgen die Ergebnisse der Studie zu Prototypizitätseffekten und schließlich die der Studien zu Form-Schematizitätseffekten. Das Kapitel schließt mit einer Diskussion, in der die Ergebnisse der einzelnen Studien kontrastiert werden und der Einfluss von Frequenz, Prototypizität und Form-Schematizität auf Variation diskutiert und evaluiert wird.

Das Hauptinteresse der Studien liegt in den Reaktionszeiten, das Antwortverhalten wird als zusätzliche Evidenz für Variation analysiert. Dabei wird für alle Studien zuerst das Antwortverhalten der Proband_innen betrachtet und erst im

¹Folgende Pakete wurden genutzt: glob (van Rossum & Drake 2009); os (van Rossum & Drake 2009); pandas (McKinney 2010).

²Folgende Pakete wurden für die Datenaufbereitung, -auswertung und Visualisierung der Daten in R genutzt: afex (Singmann u. a. 2019); broom.mixed (Bolker & Robinson 2019); data.table (Dowle & Srinivasan 2019); ggbeeswarm (Clarke & Sherrill-Mix 2017); ggeffects (Lüdtke 2018); Hmisc (Harrell & Dupont 2019); lme4 (Bates u. a. 2015); MuMIn (Barton 2019); optimx (Nash & Varadhan 2011); party (Hothorn u. a. 2006, Strobl u. a. 2008); patchwork (Pedersen 2020); readr (Wickham u. a. 2018); readxl (Wickham & Bryan 2019); tidyverse (Wickham 2017); waffle (Rudis & Gandy 2019); xtable (Dahl u. a. 2019).

³Für die self-paced-reading-Studie wurde keine Prästudie durchgeführt.

Anschluss daran werden die Reaktionszeiten diskutiert, da das Antwortverhalten potentiell Einfluss auf die Reaktionszeiten nehmen kann. Daher ist es sinnvoll, die Muster im Antwortverhalten für die Analyse der Reaktionszeiten zu kennen, um sie ggf. berücksichtigen zu können.

Das Antwortverhalten und die Reaktionszeiten werden zunächst deskriptiv ausgewertet und visualisiert. Anschließend folgt die konfirmatorische statistische Auswertung mithilfe (generalisierter) gemischter linearer Modelle. Zusätzlich werden *random forests* hinzugezogen, wenn sich Konvergenzprobleme⁴ bei den gemischten linearen Modellen ergeben. Die konfirmatorische Auswertung überprüft die Hypothesen, die mit Ausnahme der self-paced-reading-Studie samt statistischer Auswertung bereits vorab registriert wurden. Auf die konfirmatorische Analyse folgt eine explorative statistische Auswertung, wenn sich interessante Muster in den Daten zeigen. Die explorative Auswertung wurde teilweise ebenfalls vorab registriert, wenn die Muster bereits in der Prästudie zu erkennen waren. Die explorative Auswertung hat zum Ziel, zu überprüfen, ob sich Muster, die sich in den Daten erkennen lassen, auch in den statistischen Modellen zeigen. Aufgrund ihres explorativen Charakters sind *p*-Werte für diese Analysen nicht aussagekräftig, da die Analysen sich nicht auf vorab registrierte Hypothesen beziehen. Daher werden die *p*-Werte für die explorativen Analysen nicht berichtet. Im folgenden Abschnitt wird auf die Modelle, die den statistischen Analysen zugrundeliegen, näher eingegangen.

5.1 (Generalisierte) gemischte lineare Modelle und *random forests*

In diesem Abschnitt werden zunächst (generalisierte) gemischte lineare Modelle vorgestellt und im Anschluss daran *random forests*.

5.1.1 (Generalisierte) gemischte lineare Modelle

Für die statistische Analyse der Reaktionszeiten und des Antwortverhaltens werden vornehmlich (generalisierte) gemischte lineare Modelle genutzt. Um die Logik (generalisierter) gemischter linearer Modelle zu verstehen, ist es sinnvoll,

⁴Bei Konvergenzproblemen kann ein Modell keine stabilen Schätzwerte für die unabhängigen Variablen berechnen. Dieses Problem besteht häufig, wenn komplexe Modelle auf einen zu kleinen Datensatz angewendet werden. Konvergenzprobleme sind nicht zwingend mit einem unzuverlässigen Schätzwert gleichzusetzen, jedoch können sie darauf hindeuten (Winter 2020: 265–266).

5.1 (Generalisierte) gemischte lineare Modelle und random forests

zunächst lineare sowie gemischte lineare Modelle zu betrachten. Diese Modelle werden für die Analyse der Reaktionszeiten genutzt. Im Anschluss daran werden generalisierte gemischte lineare Modelle erläutert, da diese auf gemischten linearen Modellen basieren. Die generalisierten Modelle werden für die Analyse des Antwortverhaltens gerechnet.

Lineare Modelle basieren auf Regression: Hierbei wird eine kontinuierliche Variable (zum Beispiel Reaktionszeit) als Funktion einer anderen Variable modelliert (Winter 2020: 69–72):

$$\text{Reaktionszeit} = b_0 + b_1 * \text{Frequenz} + e$$

b_0 ist der y-Achsenabschnitt (*intercept*), also der Punkt, an dem die Regressionslinie die y-Achse schneidet. b_1 ist die Steigung (*slope*) der Regressionslinie, die in diesem Beispiel durch die Frequenzausprägung (z. B. FREQUENT/INFREQUENT) beeinflusst wird. e gibt den Fehler an, also in der Beispielgleichung alles, was nicht von der Variable Frequenz abgedeckt wird (Winter 2020: 73–74). Der Fehler entspricht in diesem Fall den Residuen: Residuen geben an, wie weit der tatsächliche gemessene Wert von dem Wert abweicht, den das Modell vorhersagt. Sie werden berechnet, indem die tatsächlich gemessenen Werte von den vom Modell vorhergesagten Werten abgezogen werden (Winter 2020: 73–74). Die Gleichung lässt sich beliebig erweitern:

$$\text{Reaktionszeit} = b_0 + b_1 * \text{Frequenz} + b_2 * \text{Flexion} + e$$

In dieser Gleichung nimmt zusätzlich zu Frequenz Flexion (STARK/SCHWACH) Einfluss auf die Reaktionszeiten. In der Gleichung wird der Einfluss der Faktoren auf die Reaktionszeiten separat berechnet. Genauso kann eine Interaktion zwischen den Einflussfaktoren überprüft werden:

$$\text{Reaktionszeit} = b_0 + b_1 * \text{Frequenz} + b_2 * \text{Flexion} + b_3 * (\text{Frequenz} * \text{Flexion}) + e$$

Bei einer Interaktion liegt die Erwartung zugrunde, dass die Einflussfaktoren einzeln nur bedingt Einfluss auf Reaktionszeiten nehmen, aber beide Faktoren gemeinsam die Reaktionszeit beeinflussen (Winter 2020: 133–134). Ein gutes Beispiel für Interaktion ist der Einfluss von Wasser und Sonne auf Pflanzenwachstum: Viel Wasser und viel Sonne allein fördern das Pflanzenwachstum nicht, erst beides in Kombination.

Während lineare Modelle nur feste Effekte (z. B. Frequenz und Flexion) modellieren, werden in einem gemischten linearen Modell neben den festen Effekten auch zufällige Effekte berücksichtigt (Winter 2020: 236). Teilnehmer_innen und

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Testitems sind zufällige Effekte. Sie bringen Varianz in die Daten, die sich nicht durch die festen Effekte erklären lässt (Winter 2020: 236). Zufällige Effekte haben einen unvorhersagbaren, idiosynkratischen Einfluss auf die Daten (Winter 2013: 39), während feste Effekte über viele Experimente hinweg konstant sind (Winter 2020: 236). Man sollte also in vielen Experimenten bspw. einen Einfluss von Frequenz messen können, auch wenn die Teilnehmer_innen variieren.

In den Beispielen oben wird angenommen, dass die gemessenen Daten unabhängig voneinander sind. Wenn aber Proband_innen mehrere Versuchitems pro Kondition (bspw. pro Frequenzausprägung) beurteilen, ist dies nicht der Fall: Es kann sein, dass einige Proband_innen generell schneller sind oder auf bestimmte Versuchitems langsamer oder schneller reagieren als andere Proband_innen. Daher werden Proband_innen als zufälliger Effekt in linearen Modellen berücksichtigt, wenn mehrere Messpunkte pro Proband_in vorliegen (Winter 2020: 236). Dabei können *random intercepts* (zufällige y-Achsenabschnitte) und zusätzlich dazu *random slopes* (zufällige Steigungen) genutzt werden.⁵ *Random intercepts* ermöglichen es, unterschiedliche Reaktionszeiten für schnelle und langsame Proband_innen zu berechnen, mit *random slopes* können unterschiedliche Entwicklungen in den Reaktionszeiten modelliert werden: Zum Beispiel könnten Proband_innen mit niedrigen Reaktionszeiten mit der Zeit langsamer werden, aber Proband_innen mit hohen Reaktionszeiten konstant bleiben (Winter 2020: 237–240; 267–268). Wie Proband_innen werden auch Versuchitems als zufällige Effekte berücksichtigt, wenn mehrere Items pro Kondition (bspw. pro Frequenzausprägung) genutzt werden.

Linearen Modellen liegen Vorannahmen zugrunde. Der Annahme der Unabhängigkeit von Datenpunkten kann durch ein gemischtes lineares Modell begegnet werden. Weiterhin wird davon ausgegangen, dass eine Normalverteilung und Homoskedastizität (konstante Varianz) der Residuen vorliegt (Winter 2020: 74–75). Ist dies nicht der Fall, sind die Vorhersagen des Modells unzuverlässig. Für die gemischten linearen Modelle wurden die Vorannahmen der Normalverteilung und Homoskedastizität der Residuen jeweils wie von Winter (2020: 109) empfohlen visuell überprüft. Die entsprechenden Grafiken sind im digitalen Anhang.

⁵Es ist nicht immer leicht zu entscheiden, ob sowohl *random intercepts* als auch *random slopes* für ein Modell genutzt werden sollten (siehe Winter 2020: 241–244 für eine Diskussion). In den vorliegenden Studien ist es sinnvoll, für Proband_innen sowohl *random intercepts* als auch *random slopes* zu berücksichtigen. Für Versuchitems sind jedoch in fast allen Studien nur *random intercepts* möglich, da die Versuchitems über Konditionen hinweg variieren (Winter 2020: 243). Nur bei der Prototypizitätsstudie ist das nicht der Fall, weswegen hier auch *random slopes* für Versuchitems vorgesehen waren. In der Anwendung mussten *random slopes* jedoch aufgrund von *singular fits*, die anzeigen, dass die Effektstruktur für die vorliegenden Daten zu komplex ist, sowie Konvergenzproblemen weggelassen werden.

5.1 (Generalisierte) gemischte lineare Modelle und random forests

Die Vorannahmen werden in den Modellen, die Reaktionszeiten in Millisekunden berücksichtigen, verletzt. Daher werden die Reaktionszeiten logarithmiert. Mit logarithmierten Reaktionszeiten werden die Annahmen erfüllt.

Für die Analyse des Antwortverhaltens werden generalisierte gemischte lineare Modelle genutzt, um logistische Regressionen rechnen zu können. Das Antwortverhalten ist im Gegensatz zu Reaktionszeiten nicht kontinuierlich, sondern binär: Die Proband_innen können die Frage zur Bekanntheit der Formen in der lexical-decision-Studie nur bejahen oder negieren und in der sentence-maze-Studie nur zwischen zwei Formen wählen. Auch bei dem Produktionsexperiment im Rahmen der self-paced-reading-Studie lassen sich die Antworten auf schwache und andere Formen (stark auf *-(e)s*, ohne Endung, Doppelformen) unterbrechen.

In einem gemischten generalisierten Modell wird nicht von einer Normalverteilung der Daten ausgegangen, sondern von einer anderen, im Fall der logistischen Regression von einer binomialen (Winter 2020: 200). Anders als bei kontinuierlichen Variablen versucht das Modell nicht, Werte der kontinuierlichen Variable vorherzusagen, sondern die Wahrscheinlichkeit für eine Antwort zu bestimmen, etwa für eine affirmierende Antwort unter einem bestimmten Einflussfaktor (z. B. Frequenz) (Winter 2020: 200–201).

Für binomiale Verteilungen gilt:

$$y \sim \text{binomial}(N = 1, p)$$

N beschreibt die Anzahl an Vorkommnissen (*trials*) und kann auf 1 beschränkt werden: Hierdurch wird geschätzt, wie wahrscheinlich es für eine einzelne Antwort ist, affirmierend oder negierend zu sein (Winter 2020: 200). Binominale Verteilungen mit $N = 1$ sind Bernoulli-Verteilungen. p beschreibt die Wahrscheinlichkeit für eine affirmierende oder negierende Antwort bzw. für die Wahl einer schwachen oder starken Flexionsendung. Hier wird deutlich, warum die Daten binär sein müssen: Die Wahrscheinlichkeit (p) für eine Antwort ist auf diese Weise $1 - p$ für die andere Antwort. Wie bei den linearen Modellen kann auch bei den generalisierten linearen Modellen der Einfluss mehrerer Variablen auf die abhängige Variable berechnet werden. In der Gleichung ist dies am Beispiel der Wahrscheinlichkeit für eine affirmierende Antwort dargestellt:

$$\begin{aligned} p(\text{Affirmation}) = & b_0 + b_1 * \text{Einflussfaktor}_1 + b_2 * \text{Einflussfaktor}_2 \\ & (+b_3 * \text{Einflussfaktor}_3) \\ & (+\dots) \end{aligned}$$

Problematisch ist, dass die Gleichung alle möglichen Werte annehmen kann. Wahrscheinlichkeiten bewegen sich aber zwischen 0 und 1 (Winter 2020: 201). Daher muss das Ergebnis der Gleichung transformiert werden: Durch die logistische Transformation werden die Daten auf Werte zwischen 0 und 1 beschränkt (Winter 2020: 201–202). Bei einem Wert nahe 0 ist die Wahrscheinlichkeit für affirmierende Antworten bspw. gering, bei einem Wert nahe 1 hoch. Das Gegenteil ist dann für negierende Antworten der Fall. Mithilfe der logistischen Transformation kann die Wahrscheinlichkeit für affirmierende und negierende Antworten daher in Abhängigkeit von bestimmten Einflussfaktoren (z. B. Frequenz) modelliert werden.

Für die Berechnung der (generalisierten) gemischten linearen Modelle wird das Paket *lme4* (Bates u. a. 2015), Version 1.1-23, genutzt. Zusätzlich wird auf die Pakete *MuMIn* (Barton 2019), Version 1.43.17, und *afex* (Singmann u. a. 2019), Version 0.27-2, zurückgegriffen.

5.1.2 *Random forests*

Zusätzlich zu den generalisierten gemischten linearen Modellen werden *random forests* genutzt, um das Antwortverhalten zu analysieren. Sie werden wie von Levshina (2015: 275) empfohlen dann verwendet, wenn die zufälligen Effekte in den generalisierten gemischten linearen Modellen aufgrund von Konvergenzproblemen nur teilweise oder nicht berücksichtigt werden können und die Aussagekraft der Modelle daher eingeschränkt ist. *Random forests* bestehen aus vielen *conditional inference trees*. *Conditional inference trees* werden mithilfe eines Permutationsverfahrens berechnet. Durch die Permutation werden zufällige Muster erstellt, deren Basis die real existierenden Daten sind. Die permutierten Daten werden mit den tatsächlichen verglichen, sodass geprüft werden kann, ob das zufällige Muster die Daten besser erklären kann als das tatsächliche Muster (Tagliamonte & Baayen 2012: 159, Levshina 2015: 291–292). Zudem wird getestet, welcher Einflussfaktor die Verteilung der Daten am besten vorhersagen kann. Die Daten werden anhand dieses Faktors geteilt. Dann wird innerhalb der geteilten Daten getestet, welcher Einflussfaktor die Verteilung am besten vorhersagt. Dieser Vorgang wird so lange wiederholt, bis die tatsächlich gemessenen Daten keine signifikant besseren Voraussagen treffen als die permutierten Daten (Tagliamonte & Baayen 2012: 159, Levshina 2015: 291–292). *Random forests* entstehen, indem viele *conditional inference trees* berechnet werden. Diese basieren jeweils auf zufälligen Stichproben, die aus den Daten mit Zurücklegen gezogen wurden. Auf diese Weise können die Vorhersagen jedes *conditional inference trees* mit den

Daten, die nicht Teil der Stichprobe waren, verglichen und evaluiert werden (Tagliamonte & Baayen 2012: 179–180). Für die Berechnung der *random forests* wird das Paket *party* (Hothorn u. a. 2006, Strobl u. a. 2008), Version 1.3-5, genutzt. Für die Berechnung der Passgenauigkeit der *random forests* wird zusätzlich das Paket *Hmisc* (Harrell & Dupont 2019), Version 4.4-0, verwendet. Im folgenden Abschnitt werden die Ergebnisse der lexical-decision-Studie zu Frequenzeffekten vorgestellt.

5.2 Frequenz

In diesem Abschnitt werden das Antwortverhalten und die Reaktionszeiten in der lexical-decision-Studie diskutiert. Die Studie hat zum Ziel, anhand von starken Verben mit unterschiedlicher Tokenfrequenz den Einfluss von Frequenz auf Variation zu untersuchen (siehe Abschnitt 3.1 für ausführliche Erläuterungen zur Variation in der Konjugation). Alle in diesem Abschnitt diskutierten Daten beziehen sich auf die in Abschnitt 4.2.6 vorgestellten Proband_innen. Die Stichprobengröße beträgt 53 Personen. In der Studie wurden jeweils starke Testverben genutzt, die entweder frequent (relative Tokenfrequenz ab 13 Token pro Million Wortformen) oder infrequent sind (relative Tokenfrequenz bis zu zwei Token pro Million Wortformen). Die infrequenten Verben teilen sich zusätzlich in Verben ohne Schwankung in Korpora (Ratio von starken zu schwachen Formen liegt über 100 starken Formen zu einer schwachen Form) und Verben mit Schwankung in Korpora (Ratio von starken zu schwachen Formen liegt bei unter oder gleich 100 starken Formen zu einer schwachen Form). Pro Frequenzausprägung (FREQUENT, INFREQUENT OHNE SCHWANKUNG und INFREQUENT MIT SCHWANKUNG) wurden acht Testverben⁶ genutzt.

Innerhalb einer *lexical decision task* haben Proband_innen die starken und schwachen Partizip-II-Formen jedes Testverbs nach ihrer Bekanntheit bewertet (siehe Abschnitt 4.2 für ausführliche Erläuterungen). Es wird erwartet, dass starke Formen der Testverben unabhängig von ihrer Frequenz bekannt sind. Hinsichtlich der Bekanntheit schwacher Formen wird ein Frequenzeffekt angenommen: Für die Ausprägung FREQUENT wird erwartet, dass schwache Formen unbekannt sind (z. B. **gefahrt*). Für die infrequenten Verben wird antizipiert, dass schwache Formen bekannter sind, dabei wird davon ausgegangen, dass

⁶FREQUENT: *sprechen, fahren, tragen, ziehen, halten, schreiben, sinken, fliegen*; INFREQUENT OHNE SCHWANKUNG: *flechten, spinnen, schmelzen, anschwellen, kneifen, fechten, melken, dreschen*; INFREQUENT MIT SCHWANKUNG: *salzen, hauen, quellen, einsaugen, weben, glimmen, gären, sinnen* (siehe Abschnitt 4.2.2 für Erläuterungen zu den Testverben).

die schwachen Verbformen der Ausprägung INFREQUENT OHNE SCHWANKUNG (z. B. *gefechtet*) weniger bekannt sind als schwache Verbformen der Ausprägung INFREQUENT MIT SCHWANKUNG (z. B. *geglimmt*). Hinsichtlich der Reaktionszeiten wird unabhängig von der beurteilten Flexionsform (stark/schwach) für Verben der Ausprägung FREQUENT ein schnelleres Antwortverhalten erwartet als für die anderen Ausprägungen. Zwischen den Ausprägungen INFREQUENT OHNE SCHWANKUNG und INFREQUENT MIT SCHWANKUNG wird kein Unterschied in den Reaktionszeiten antizipiert (siehe Abschnitt 4.2.1 für genauere Erläuterungen).

5.2.1 Antwortverhalten

Für jede Frequenzausprägung (FREQUENT, INFREQUENT OHNE SCHWANKUNG, INFREQUENT MIT SCHWANKUNG) wurden starke und schwache Partizip-II-Formen von jeweils acht Testverben beurteilt (bspw. *gezieht* und *gezogen*). Insgesamt wurden dabei pro Frequenzausprägung 424 starke und schwache Formen auf Ihre Bekanntheit hin bewertet.⁷ Abbildung 5.1 zeigt einen Waffleplot mit den Antworten. In dem Waffleplot wird die Verteilung der Antworten zu der Frage *Kennen Sie diese Wortform?* für die einzelnen Frequenzausprägungen und Flexionsformen dargestellt. Jedes Rechteck entspricht einer Antwort. Dunkelgrau eingefärbt sind Rechtecke, die eine affirmierende Antwort abbilden, hellgrau Rechtecke, die eine verneinende Antwort darstellen.

Es zeigt sich ein klar verteiltes Antwortverhalten: Starke Formen werden unabhängig von der Frequenzausprägung als bekannt bewertet. Lediglich bei der Ausprägung INFREQUENT MIT SCHWANKUNG werden starke Formen vereinzelt (7,1 %) als unbekannt gekennzeichnet. Im Gegensatz zu den starken Formen hängt bei den schwachen Formen die Bekanntheit deutlich von der Frequenzausprägung ab: Während die schwachen Formen bei Testverben der Ausprägung FREQUENT (z. B. **gefahrt*) als unbekannt bewertet werden,⁸ zeigen sich bei den Verben der Ausprägung INFREQUENT OHNE SCHWANKUNG (z. B. *gefechtet*) Tendenzen zur Bekanntheit (24,7 %). Bei den Testverben der Ausprägung INFREQUENT MIT SCHWANKUNG (z. B. *geglimmt*) werden die schwachen Formen zu 51,4 % als bekannt bewertet.

⁷Die 424 Antworten pro Frequenzausprägung und Flexion ergeben sich aus den acht Testverben und den 53 Proband_innen.

⁸Die klare Antwortverteilung in der Ausprägung FREQUENT zwischen starken und schwachen Formen weist darauf hin, dass die Proband_innen die Stimuli einheitlich beantworten. Daher wären auch bei einer mehrmaligen Präsentation derselben Stimuli keine abweichenden Antworten zu erwarten (siehe Abschnitt 4.1.1 für weitere Erläuterungen zu diesem Problem bei *lexical decision tasks*), auch das Antwortverhalten in der lexical-decision-Studie zu Form-Schematizität weist auf ein konsistentes Antwortverhalten für die Stimuli der Ausprägung FORM-SCHEMA und STARK hin (siehe Abschnitt 5.4.1.2),

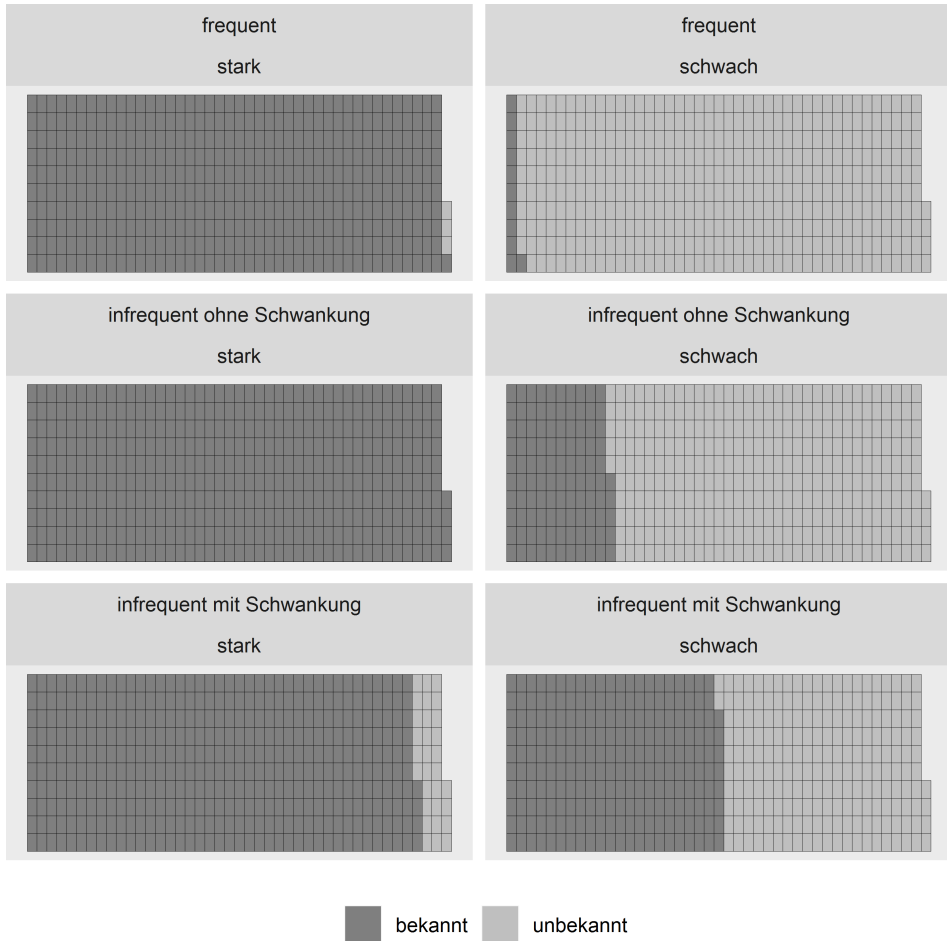


Abbildung 5.1: Antwortverhalten in der Frequenzstudie

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Ein näherer Blick auf die Ausprägungen INFREQUENT MIT SCHWANKUNG und INFREQUENT OHNE SCHWANKUNG⁹ zeigt, dass die Testverben innerhalb dieser Ausprägungen unterschiedliche Antwortverteilungen evozieren. Abbildung 5.2 gibt einen Überblick über das Antwortverhalten bei den Testverben der Ausprägung INFREQUENT OHNE SCHWANKUNG.¹⁰ Die Testverben sind absteigend nach ihrer Ratio von starken zu schwachen Formen geordnet. Oben ist also mit *flechten* das Verb mit der höchsten Ratio von starken zu schwachen Formen (807/1) und unten mit *dreschen* das Verb mit der geringsten Ratio von starken zu schwachen Formen (123/1) (siehe Abschnitt 4.2.2 für nähere Ausführungen zur Berechnung der Ratio zwischen starken und schwachen Formen). Insgesamt liegen 53 Antworten pro starker und schwacher Form jedes Testverbs vor, da 53 Proband_innen die Testverben beurteilten. Ein Rechteck steht für die Antwort einer Person.

Wie bereits aus Abbildung 5.1 deutlich wurde, werden die starken Formen durchweg als bekannt bewertet. Die schwachen Formen werden dagegen nur selten als bekannt gekennzeichnet. Der Anteil der Bekanntheit macht dabei maximal 35,8 % (*geflechtet*) aus. *Geschmolzt* und *angeschwellt* kommen nur auf einen Anteil von 9 % bzw. 11 % Bekanntheit, bei den restlichen Verben ist das Minimum an Bekanntheit schwacher Formen mit 24,5 % deutlich höher.

Abbildung 5.3 zeigt das Antwortverhalten für die Testverben der Ausprägung INFREQUENT MIT SCHWANKUNG. Auch hier wurden insgesamt 53 Antworten pro starker und schwacher Form der Testverben abgegeben. Die Testverben sind wieder nach ihrer Ratio geordnet: Oben ist mit *salzen* das Verb mit der größten Ratio von starken zu schwachen Formen (60/1) unten mit *sinnen* das mit der kleinsten (0,59/1) (zur Berechnung der Ratios siehe Abschnitt 4.2). Auch hier steht ein Rechteck für die Antwort einer Person.

Die starken Formen werden bei infrequenten Verben mit Schwankung nicht durchweg als bekannt angegeben. Dies ist nur bei *hauen* und *einsaugen* zu beobachten. Der Anteil der Bekanntheit starker Formen ist bei allen Verben bis auf *glimmen* jedoch hoch: Nur in einer (1,8 %) bis fünf Antworten (9,4 %) wird die Form als unbekannt bewertet. Im Gegensatz dazu fällt *geglommen* auf, dass nur zu 60,9 % als bekannt bewertet wurde.

⁹Ein Überblick über das Antwortverhalten bei den Testverben der Ausprägung FREQUENT befindet sich in Abbildung I.1.

¹⁰Das Testverb *spinnen* könnte eine Verzerrung der Daten bei den Verben der Ausprägung INFREQUENT OHNE SCHWANKUNG auslösen: Die schwache Form *gespinnt* könnte auch auf das homographe Verb *spinnen* im Sinne von ‚drehen‘ zurückgehen, das aus dem Englischen (*to spin*) entlehnt wurde und daher mit /s/ statt mit /ʃ/ ausgesprochen wird. Diese Bedeutung sollte jedoch im deutschen Sprachgebrauch weit weniger frequent sein als die Bedeutung ‚Wolle zu einem Faden drehen‘ bzw. ‚verrückt sein‘. Eine Beeinflussung der Ergebnisse durch das homographe Verb auf das Antwortverhalten kann jedoch nicht ausgeschlossen werden.

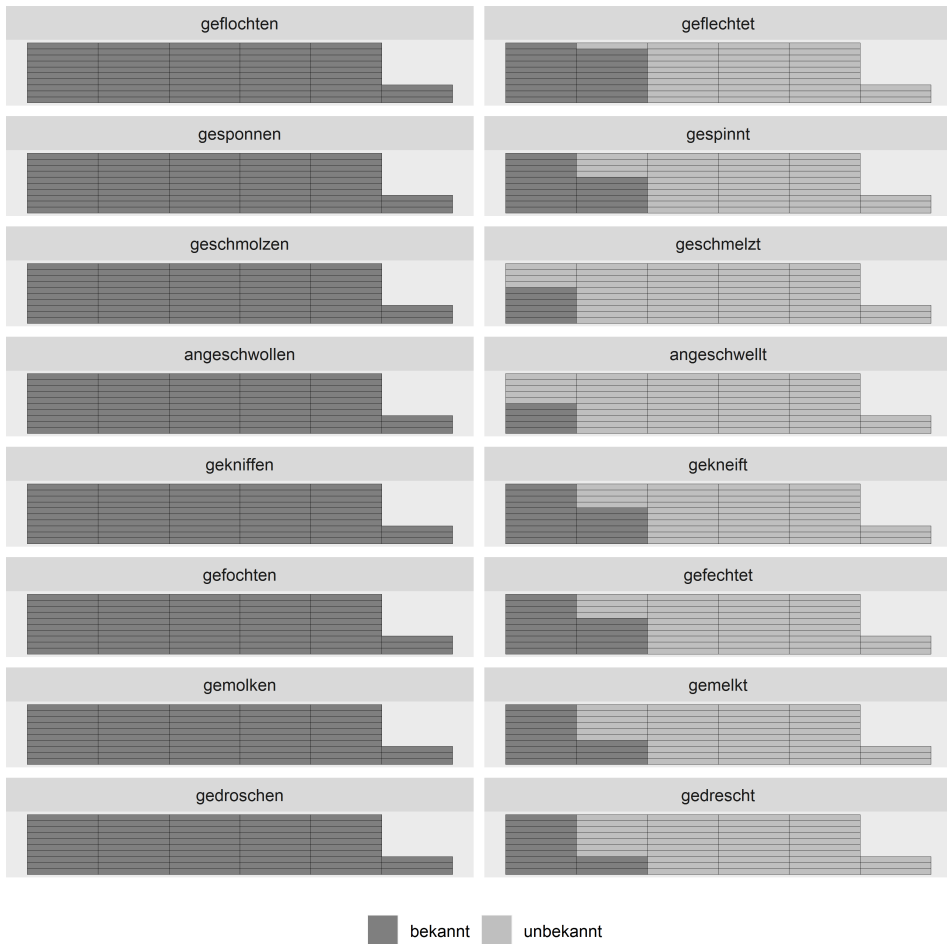


Abbildung 5.2: Antwortverhalten bei den Testverben der Ausprägung INFREQUENT OHNE SCHWANKUNG

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

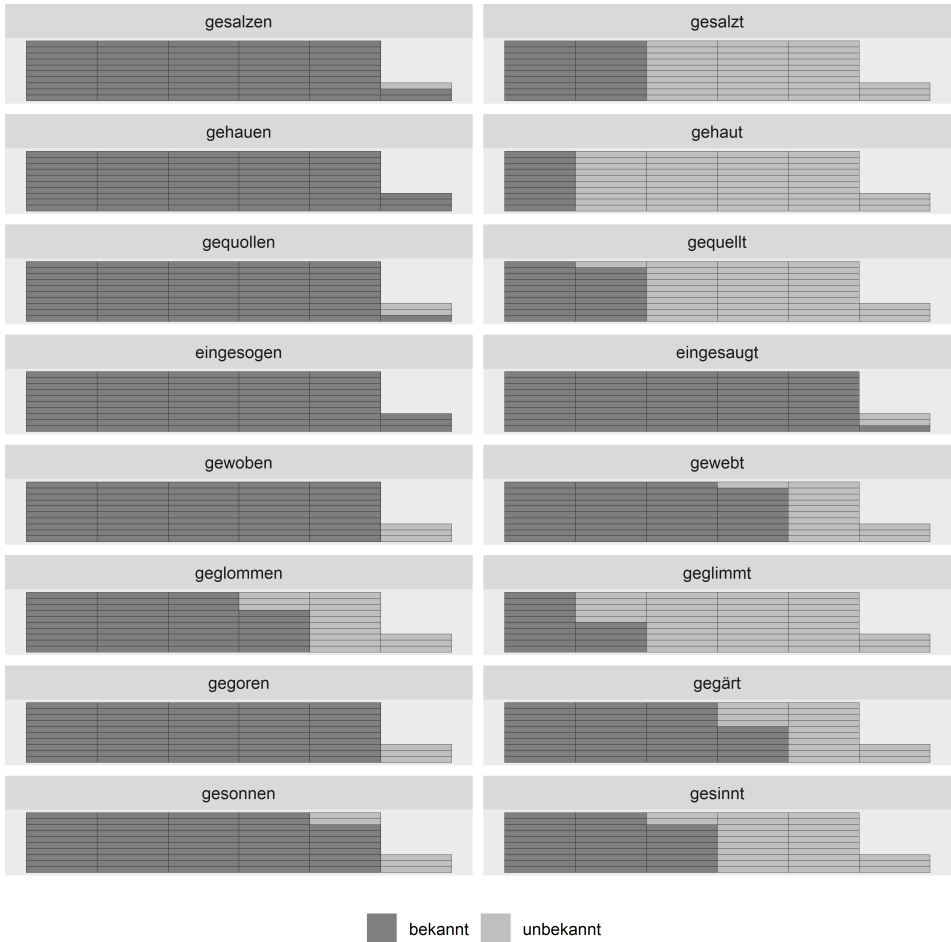


Abbildung 5.3: Antwortverhalten bei den Testverben der Ausprägung INFREQUENT MIT SCHWANKUNG

Die schwachen Formen werden sehr heterogen bewertet: *Gehaut* wird nur zu 18,8 % als bekannt gekennzeichnet. Dies könnte an der arealen Verteilung der schwachen Formen liegen: Diese sind laut der Dürscheid u. a. (2018) vorrangig in Österreich zu finden und machen dort in einigen Regionen 78 % der Verwendungen aus, aber in Deutschland ist *gehaut* sehr selten (2 % in Bayern). Dieser Umstand wurde bei der Konzeption der Studie nicht berücksichtigt. Die Teilnehmer_innen stammen allesamt aus Deutschland und nur 18 der 53 Proband_innen aus Bayern, was die geringe Bekanntheit erklärt. Deutlich höher als bei *gehaut* liegt der Anteil der Bekanntheit bei *geglimmt* (28,3 %), *gequellt* (35,8 %) und *gesalzt* (37,7 %). Dennoch sind die schwachen Formen bei diesen Verben vergleichsweise wenig bekannt: Die schwachen Formen anderer Testverben dieser Ausprägung werden zu über 50 % als bekannt bewertet, dabei fällt besonders *ingesaugt* mit 96,2 % Bekanntheit auf.

Bei *hauen*, *salzen* und *quellen* scheinen die schwachen Formen noch vergleichsweise wenig mental gefestigt zu sein, da die Bekanntheitswerte der schwachen Formen den höchsten Bekanntheitswerten der schwachen Formen von Verben der Ausprägung INFREQUENT OHNE SCHWANKUNG ähneln. Auch bei *glimmen* ist die Bekanntheit für schwache Formen gering. Jedoch weisen bei *glimmen* auch die starken Formen geringe Bekanntheitswerte auf. Sie scheinen somit nicht mehr stark gefestigt zu sein, aber die schwachen Formen sind noch nicht gefestigt, sodass beide Formen niedrige Bekanntheitswerte zeigen. Dies lässt darauf schließen, dass die Partizip-II-Formen des Verbs generell so infrequent sind, dass sich keine der Formen festigen kann. Diese Schlussfolgerung wird gestützt durch Cappellaro (2013) und Thornton (2019), die vergleichbare Effekte bei infrequenten Strukturen beobachten.

Während bei *glimmen* sowohl die starke als auch die schwache Form vergleichsweise hohe Anteile an Unbekanntheit aufweisen, ist bei *einsaugen* das Gegenteil der Fall: Hier liegt die Bekanntheit für die schwache Form bei 96,2 % und für die starke bei 100 %. Der Schluss liegt nahe, dass die starke und die schwache Form mit unterschiedlichen Funktionen assoziiert sind. Die schwache Form von *saugen* ist mit dem Einziehen von Staub oder Flüssigkeit in technische Geräte assoziiert, sodass bspw. *staubsaugen* nur schwache Formen aufweist (Nowak 2013, Duden 2020: 316–322). Um in der Studie die Assoziation der schwachen Form mit Technik zu stören, wurde das Präfixverb *einsaugen* genutzt (siehe Abschnitt 4.2.2), allerdings ist eine technische Bedeutung auch für *einsaugen* nicht ausgeschlossen. Die schwachen Formen von *saugen* könnten so stark mit technischen Geräten assoziiert sein, dass auch *ingesaugt* und *ingesogen* als Verbformen mit unterschiedlicher Konnotation wahrgenommen werden (*Der Saugbagger hat das Öl eingesaugt* vs. *Er hat den Zigarettenrauch eingesogen*). Diese Hy-

pothese müsste in einer weiterführenden Studie überprüft werden: Sie ließe sich zunächst in Korpora und darauf aufbauend bspw. mithilfe einer *sentence maze task* untersuchen, in der starke und schwache Formen von *saugen* und *einsaugen* in unterschiedlichen Kontexten genutzt werden. In der vorliegenden Studie zeigt sich hinsichtlich der logarithmierten Reaktionszeiten kein Unterschied zwischen starken (Mittelwert = 0,24; Standardabweichung = 0,41) und schwachen Formen (Mittelwert = 0,23; Standardabweichung = 0,36) von *einsaugen* (ausführlich zur Analyse der Reaktionszeiten siehe Abschnitt 5.2.2).

Für die konfirmatorische statistische Analyse der Daten wird ein generalisiertes gemischtes lineares Modell¹¹ genutzt, das Frequenzausprägung sowie Flexion als feste Effekte enthält: Antwortverhalten ~ Frequenz * Flexion + (1|Proband_in) + (1|Item). Die beiden Effekte können interagieren: Der Einfluss von Frequenz und Flexion wird also nicht separat betrachtet, indem der Einfluss der Frequenzausprägungen (bspw. FREQUENT vs. INFREQUENT OHNE SCHWANKUNG) unabhängig vom Einfluss der Flexion getestet wird und dann separat der Einfluss der Flexion (schwach vs. stark) unabhängig von der Frequenzausprägung. Stattdessen wird getestet, wie Frequenzausprägung und Flexion zusammen das Antwortverhalten ändern (bspw. die Ausprägung FREQUENT und schwache Flexion vs. FREQUENT und starke Flexion). Als zufälliger Effekt werden *random intercepts* für Proband_innen und Versuchitems (Lemmata) angesetzt. *Random slopes* für Proband_innen waren in der Präregistrierung vorgesehen, werden aber aufgrund von *singular fits* weggelassen, die anzeigen, dass die zufällige Effektstruktur für die vorliegenden Daten zu komplex ist. Da die Versuchitems über die Frequenzausprägungen nicht konstant sind, sind *random slopes* hierfür nicht sinnvoll (Winter 2020: 243).

Das Modell gibt Standardabweichungen um den y-Achsenabschnitt (*intercept*) für die zufälligen Effekte aus: Die Standardabweichung der Testitems ist mit 0,89 geringfügig größer als die der Teilnehmer_innen (0,85). Tabelle 5.1 zeigt die geschätzten Werte für die festen Effekte. Die Werte sind auf zwei Nachkommastellen gerundet. Der Wert für den y-Achsenabschnitt bezieht sich auf das Referenzlevel, für das die Ausprägung INFREQUENT OHNE SCHWANKUNG und schwache Formen gewählt wurden. Die drei nächsten Zeilen benennen den Effekt, wenn nur die Form (stark) oder nur die Frequenzausprägung (FREQUENT, INFREQUENT MIT SCHWANKUNG) geändert werden. Die letzten beiden Zeilen geben den Effekt für eine Änderung in der Frequenzausprägung und in der Form an. Aufgrund der Interaktion ist eine direkte Interpretation der Werte nur eingeschränkt möglich.

¹¹Das Modell hat mit verschiedenen Optimierungsmethoden nicht konvergiert. Da die verschiedenen Konvergierungsmethoden jedoch vergleichbare Werte ergaben, scheinen die geschätzten Werte der Modelle dennoch verlässlich zu sein.

Tabelle 5.1: Werte des Modells für die Unbekanntheit starker und schwacher Formen in der Frequenzstudie

	Wert	SE	<i>z</i>	<i>p</i>
(Intercept = infrequent ohne Schwankung, schwach)	1,36	0,00	423,67	<0,01
Flexion: stark	-20,90	0,00	-6522,19	<0,01
Frequenz: frequent	2,85	0,00	890,86	<0,01
Frequenz: infrequent mit Schwankung	-1,50	0,00	-468,60	<0,01
Flexion: stark & Frequenz: frequent	11,25	0,00	3513,81	<0,01
Flexion: stark & Frequenz: infrequent mit Schwankung	17,75	0,00	5526,05	<0,01

Festhalten lässt sich an dieser Stelle bereits, dass der Standardfehler (SE) für alle Werte minimal ist, sodass nur von einer sehr geringen Schwankung um die geschätzten Werte auszugehen ist. Um die Werte interpretieren zu können, bietet es sich an, diese in einer Kreuztabelle darzustellen (Winter 2020: 142–143).

Tabelle 5.2: Kreuztabelle der Werte des Modells für die Unbekanntheit starker und schwacher Formen in der Frequenzstudie

	schwach	stark
INFREQUENT OHNE SCHWANKUNG	1,36	$1,36 - 20,9 = -19,54$
FREQUENT	$1,36 + 2,85 = 4,21$	$1,36 - 20,9 + 2,85 + 11,25 = -5,44$
INFREQUENT MIT SCHWANKUNG	$1,36 - 1,5 = -0,14$	$1,36 - 20,9 - 1,5 + 17,75 = -3,29$

Der y-Achsenabschnitt von 1,36 wird für alle Frequenzausprägungen und Flexionsformen als Basis genutzt. Für die Berechnung der starken Formen der Ausprägung INFREQUENT OHNE SCHWANKUNG wird zusätzlich die Steigung von -20,9 berücksichtigt. Die anderen Zeilen folgen derselben Logik: Für die Frequenzausprägungen FREQUENT und INFREQUENT MIT SCHWANKUNG wird 2,85 bzw. -1,5 addiert und für die starken Formen dieser Ausprägung zusätzlich 11,25 und 17,75. Die Werte 11,25 und 17,75 ergeben sich aus der Interaktion aus Frequenzausprägung und Flexionsform.

Die so berechneten Werte lassen sich interpretieren. Die Rohwerte sowie die berechneten Werte geben *log odds* an. Aus den *log odds* lässt sich auf Wahrscheinlichkeiten schließen: *log odds* von 0 entsprechen einer Wahrscheinlichkeit von

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

0,5; *log odds* von unter bzw. über 0 korrespondieren dementsprechend mit Wahrscheinlichkeiten von unter bzw. über 0,5 (Winter 2020: 203–204). Nur für schwache Formen der Ausprägung INFREQUENT MIT SCHWANKUNG liegen die *log odds* mit $-0,14$ nahe 0. Die anderen *log odds* sind klar ausgeprägt. Für schwache Formen sind die *log odds* positiv, für starke negativ. Dies weist auf ein unterschiedliches Antwortverhalten für schwache und starke Formen hin.

Um die Wahrscheinlichkeit für negierende und affirmierende Antworten zu erhalten, müssen die Daten logistisch transformiert werden (Winter 2020: 206). Die Wahrscheinlichkeit für die Unbekanntheit von Flexionsformen in Abhängigkeit von Frequenzausprägungen ist in Abbildung 5.4 dargestellt. Schwache Formen werden durch Kreise dargestellt, starke durch Dreiecke.

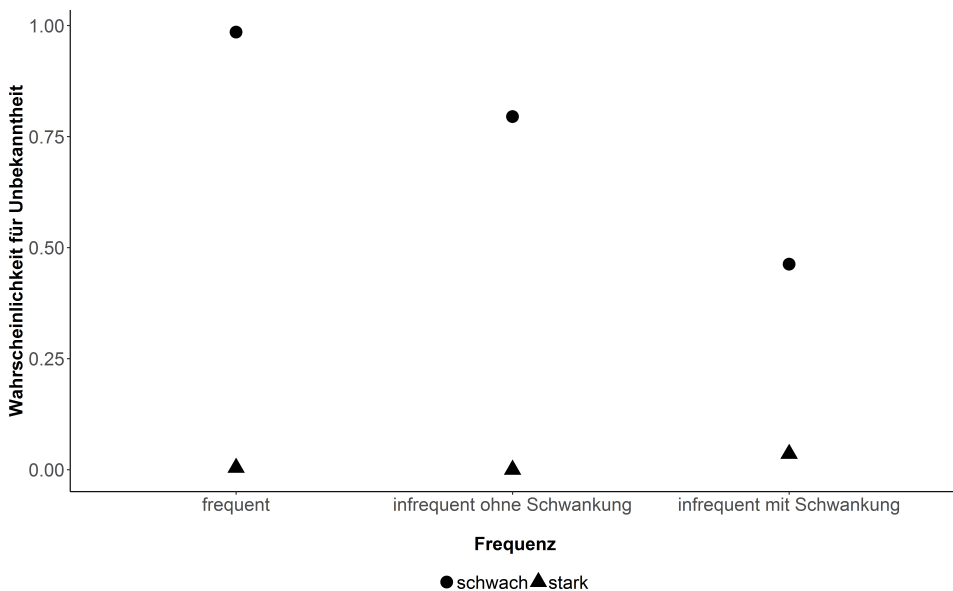


Abbildung 5.4: Wahrscheinlichkeit für die Unbekanntheit starker und schwacher Formen in der Frequenzstudie

Hinsichtlich der Flexion sagt das Modell eine klare Verteilung vorher: Für starke Flexion sind frequenzunabhängig Werte um 0 zu erwarten, die Wahrscheinlichkeit für Unbekanntheit liegt also bei 0. Dementsprechend ist Bekanntheit sehr wahrscheinlich. Es zeigt sich aber, dass bei der Frequenzausprägung INFREQUENT MIT SCHWANKUNG die Unbekanntheit etwas wahrscheinlicher ist als für die anderen Ausprägungen. Hinsichtlich der schwachen Flexion hängt die Wahrscheinlichkeit für Unbekanntheit von der Frequenzausprägung ab: Bei der Ausprägung FREQUENT liegt der Wert bei 1, Unbekanntheit ist also sehr wahrschein-

lich. Bei der Ausprägung `INFREQUENT OHNE SCHWANKUNG` ist Unbekanntheit mit ungefähr 0,8 etwas weniger wahrscheinlich. Für die Ausprägung `INFREQUENT MIT SCHWANKUNG` liegt die Wahrscheinlichkeit für Unbekanntheit bei ungefähr 0,5, sodass Unbekanntheit und Bekanntheit gleich wahrscheinlich sind.

Die 95 %-Konfidenzintervalle für die Wahrscheinlichkeitswerte sind in der Darstellung nicht zu sehen, da sie kaum um die Werte schwanken. Das 95 %-Konfidenzintervall gibt den Bereich an, in dem 95 % der Stichproben liegen, die den wahren Mittelwert einer Population enthalten (Field u. a. 2012: 45). Zieht man 100 Stichproben aus der gleichen Population, liegen 95 der Stichproben innerhalb des 95 %-Konfidenzintervalls. Für alle Frequenzausprägungen ist von systematischen Effekten auszugehen, da die 95 %-Konfidenzintervalle sich nicht überlappen.

Dies bestätigt auch ein Blick auf die p -Werte aus der Zusammenfassung des Modells: Alle Ausprägungen liegen unter dem α -Level von 0,01. Es ist somit davon auszugehen, dass die vorliegenden Daten unter der Nullhypothese (die Flexion und die Frequenzausprägung beeinflussen das Antwortverhalten nicht) sehr unwahrscheinlich sind. Die Effektstärke des Modells lässt sich mithilfe des R^2 -Werts¹² einschätzen. Dieser Wert gibt an, wie viel Varianz in den Daten durch das Modell abgedeckt wird (Winter 2020: 89–90). Für das Modell ergibt sich ein Wert von 0,92 (theoretisch) bzw. 0,91 (delta). Fast die gesamte Varianz in den Daten kann also mithilfe der festen Effekte erklärt werden. Bezieht man die zufälligen Effekte mit ein, erhöht sich der Wert nur leicht auf 0,95 bzw. 0,93.

Aufgrund der Konvergierungsprobleme bei dem generalisierten linearen Modell wird zusätzlich ein *random forest* mit Frequenzausprägung und Flexion als Einflussfaktoren gerechnet: Antwortverhalten ~ Frequenz + Flexion. Für den *random forest* wurden 1.000 *conditional inference trees* gerechnet, für die jeweils beide Einflussfaktoren genutzt wurden (Levshina 2015: 297). Abbildung 5.5 zeigt den *conditional inference tree* für die Einflussfaktoren.

Der wichtigste Einflussfaktor ist die Flexion: Für starke Formen liegt die Wahrscheinlichkeit für Unbekanntheit unabhängig von der Frequenzausprägung maximal bei 0,05. Diesen maximalen Wert weist die Ausprägung `INFREQUENT MIT SCHWANKUNG` auf. Bei schwachen Formen hängt die Bekanntheit stark von der

¹²Die Berechnung eines R^2 -Werts stellt für generalisierte gemischte lineare Modelle im Gegensatz zu gemischten linearen Modellen eine Herausforderung dar (Nakagawa u. a. 2017: 1–2). Für generalisierte Modelle wird sowohl die theoretische als auch die observierte Varianz berechnet. Die theoretische Varianz bezieht sich auf die mögliche Varianz in der Verteilung, die observierte auf Varianz in den beobachteten Werten (Nakagawa u. a. 2017: 4–6). Zur Berechnung der observierten Varianz wird die delta-Methode angewandt, weshalb der Wert delta genannt wird.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

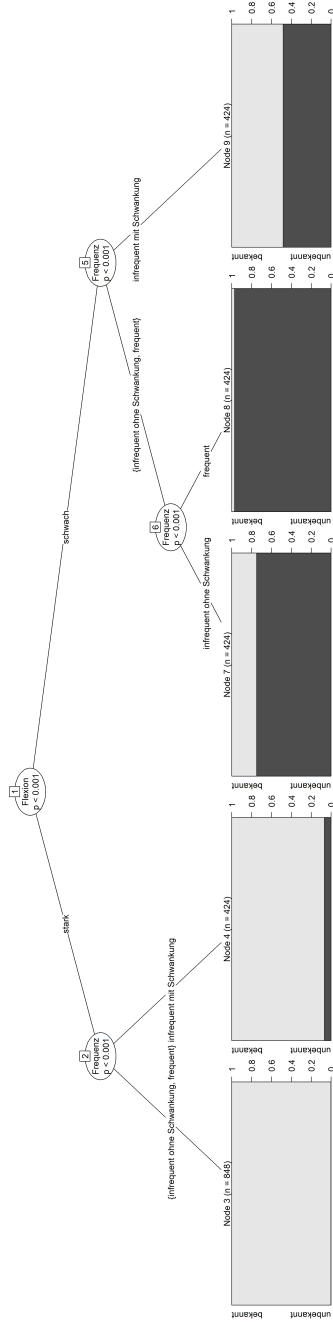


Abbildung 5.5: *Conditional inference tree* für die Bekanntheit bzw. Unbekanntheit starker und schwacher Formen in der Frequenzstudie

Frequenzausprägung ab: Bei FREQUENT ist die Wahrscheinlichkeit für Unbekanntheit mit 0,95 sehr hoch, bei INFREQUENT OHNE SCHWANKUNG ist Unbekanntheit mit 0,75 ebenfalls recht wahrscheinlich. Bei INFREQUENT MIT SCHWANKUNG ist hinsichtlich schwacher Formen Bekanntheit und Unbekanntheit ungefähr gleich wahrscheinlich. Die Ergebnisse des generalisierten gemischten linearen Modells werden also bestätigt. Für den *random forest* kann die *conditional variance importance* berechnet werden, die angibt, wie sehr die einzelnen Faktoren Einfluss auf das Modell nehmen (Levshina 2015: 298–299). Der Wert liegt für die Flexion bei 0,3 und für die Frequenzausprägung bei 0,06. Flexion (stark vs. schwach) beeinflusst das Antwortverhalten also deutlich stärker als die Frequenzausprägung (FREQUENT, INFREQUENT OHNE SCHWANKUNG, INFREQUENT MIT SCHWANKUNG). Da es sich bei den Testverben um starke Verben handelt, ist dieser Befund zu erwarten. Für den *random forest* wurde geprüft, wie gut er auf die Daten passt, hierfür wurde der C-Wert berechnet (Levshina 2015: 299). Der Wert ist mit 0,94 hoch, Tagliamonte & Baayen (2012: 156) schlagen 0,8 als Minimum für eine gute Passung vor.

Die in Abschnitt 4.2.1 aufgestellten Hypothesen haben sich bestätigt: Starke Formen der Testverben sind unabhängig von der Frequenzausprägung bekannt. Nur bei den infrequenten Verben mit Schwankung (z. B. *geglommen*) zeigen sich erste Tendenzen zur Unbekanntheit. Die Bekanntheit der schwachen Formen hängt hingegen von der Frequenzausprägung ab: Während schwache Formen bei frequenten Verben (z. B. *gezieht*) unbekannt sind, zeigen sich erste Tendenzen zur Bekanntheit bei infrequenten Verben ohne Schwankung (z. B. *geflechtet*). Bei infrequenten Verben mit Schwankung (z. B. *gesinnt*) halten sich Unbekanntheit und Bekanntheit für schwache Formen die Waage.

5.2.2 Reaktionszeiten

Der Blick auf die Reaktionszeiten zeigt, dass Testverben der Frequenzausprägung FREQUENT kürzere Reaktionszeiten hervorrufen als Testverben der anderen Frequenzausprägungen. Abbildung 5.6 zeigt die logarithmierten Reaktionszeiten¹³ in einem Beeswarmplot. Jeder Punkt im Plot steht für die Reaktionszeit einer Person für einen einzelnen Stimulus. Die Reaktionszeiten für starke Formen sind als hellgraue Dreiecke dargestellt, die Reaktionszeiten für schwache Formen als dunkelgraue Kreise. Auf den Beeswarmplots ist jeweils das arithmetische Mittel (M) abgebildet, die Fehlerbalken geben die Standardabweichung an. Zusätzlich

¹³Da im gemischten linearen Modell mit logarithmierten Reaktionszeiten gearbeitet wird, werden auch in der visuellen Betrachtung logarithmierte Reaktionszeiten berücksichtigt. Dies gilt für alle Studien.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

zum arithmetischen Mittel sowie der Standardabweichung (*standard deviation* SD) wird der Standardfehler (*standard error* SE) berichtet. Dieser fällt sehr klein aus, weshalb er nicht graphisch dargestellt wird. Die Standardabweichung gibt an, wie gut der Mittelwert die Daten repräsentiert (Field u. a. 2012: 39–40). Der Standardfehler misst, wie gut der Mittelwert der Stichprobe dem Mittelwert der Population entspricht (Field u. a. 2012: 42–43). Für beide Maße ist die Entsprechung höher, je kleiner der Wert ist.

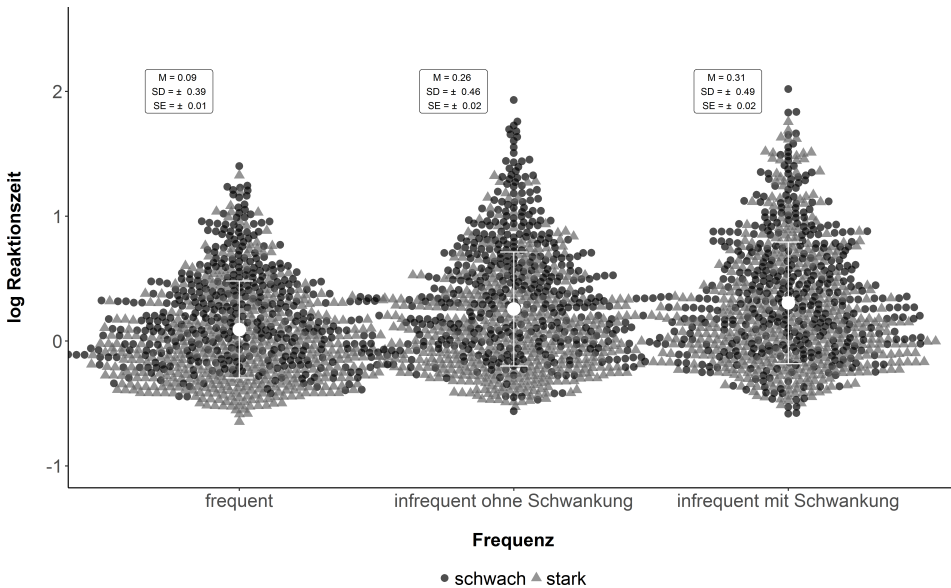


Abbildung 5.6: Reaktionszeiten in der Frequenzstudie

Alle Frequenzausprägungen zeigen eine ähnliche Verteilung der Reaktionszeiten. Allerdings ist deutlich zu erkennen, dass die Ausprägung **FREQUENT** insgesamt geringere Reaktionszeiten hervorruft als die anderen Frequenzausprägungen. Dies zeigt sich daran, dass die Reaktionszeiten nicht so stark streuen (von $-0,65$ bis $1,4$) wie die Reaktionszeiten der anderen Ausprägungen (von $-0,56$ bis $1,93$ bei **INFREQUENT OHNE SCHWANKUNG** bzw. von $-0,58$ bis $2,02$ bei **INFREQUENT MIT SCHWANKUNG**). Dies wird durch den Mittelwert gespiegelt, der mit $0,09$ und einer Standardabweichung von $0,39$ deutlich unter den Mittelwerten der Ausprägungen **INFREQUENT OHNE SCHWANKUNG** ($0,31$, Standardabweichung $0,49$) und **INFREQUENT MIT SCHWANKUNG** ($0,25$, Standardabweichung $0,49$) liegt.

Die Verben der Ausprägungen **INFREQUENT OHNE SCHWANKUNG** und **INFREQUENT MIT SCHWANKUNG** scheinen vergleichbare Reaktionszeiten hervorgerufen

zu haben: Die Reaktionszeiten sind ähnlich verteilt und auch die Differenz zwischen den Mittelwerten deutet mit 0,06 allenfalls auf einen vernachlässigbaren Unterschied hin.

Vergleicht man die Reaktionszeiten für starke und schwache Formen, so fällt auf, dass bei der Ausprägung *FREQUENT* die starken Formen schneller beurteilt wurden als die schwachen Formen. Dies ist zu erwarten, da die schwachen Formen im Gegensatz zu den starken nicht mental gefestigt sind. Eine schnelle Ablehnung kann daher nur durch den Rückgriff auf starke Formen stattfinden, die die schwachen Formen aufgrund der hohen Tokenfrequenz statistisch ausstechen (zum statistischen Vorkaufsrecht siehe Goldberg 2019: 74–94 sowie Abschnitt 2.1.3). Für die Ausprägung *INFREQUENT OHNE SCHWANKUNG* lässt sich ebenfalls eine Tendenz zu niedrigeren Reaktionszeiten bei starken Formen erkennen, bei *INFREQUENT MIT SCHWANKUNG* scheint die Flexion keinen Einfluss auf die Reaktionszeiten zu haben.

Der Blick auf die Mittelwerte der Reaktionszeiten für starke und schwache Formen bestätigt diesen Eindruck, wie Abbildung 5.7 zeigt. Die Reaktionszeiten für schwache Formen sind als Kreise dargestellt, die Reaktionszeiten für starke Formen als Dreiecke. Die Fehlerbalken zeigen die Standardabweichung.

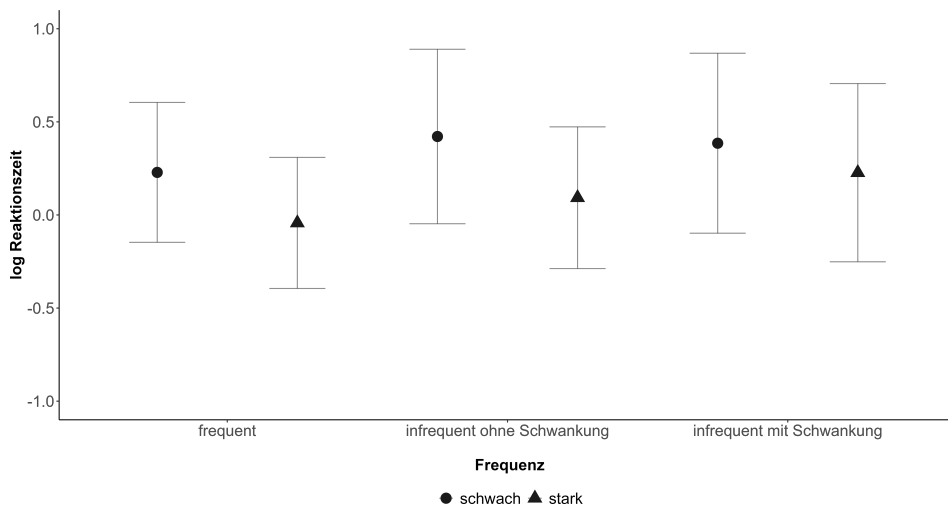


Abbildung 5.7: Reaktionszeiten in der Frequenzstudie getrennt nach starken und schwachen Formen

Die Reaktionszeiten für schwache Formen liegen für alle Frequenzausprägungen über den Reaktionszeiten für starke Formen. Dies ist deutlich bei den Ausprägungen *FREQUENT* und *INFREQUENT OHNE SCHWANKUNG* zu erkennen. Bei

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

INFREQUENT MIT SCHWANKUNG ist nur ein kleiner Vorteil für starke Formen zu sehen: Da schwache Formen bereits in Korpora attestiert sind, ist davon auszugehen, dass sie durch den Gebrauch mental gefestigt sind, sodass sie schneller verarbeitet werden können.

Auch in dieser Darstellung zeigt sich, dass die starken und schwachen Formen der Ausprägung FREQUENT jeweils geringere Reaktionszeiten aufweisen als die starken und schwachen Formen der anderen Ausprägungen. Die Reaktionszeiten für INFREQUENT OHNE SCHWANKUNG und INFREQUENT MIT SCHWANKUNG sind hingegen vergleichbar.

Neben dem oben diskutierten Einfluss des *entrenchments* könnte ein weiterer Faktor die Unterschiede in den Reaktionszeiten zwischen starken und schwachen Formen hervorrufen: Die Ablehnung der Frage nach der Bekanntheit könnte längere Reaktionszeiten hervorrufen als die Zustimmung zur Frage. Die schwachen Formen würden in diesem Fall aufgrund der Ablehnung höhere Reaktionszeiten hervorrufen.

Für schwache Formen von Verben der Ausprägung INFREQUENT MIT SCHWANKUNG, die ungefähr zur Hälfte als bekannt und unbekannt bewertet wurden, ist im arithmetischen Mittel ein kleiner Unterschied in den Reaktionszeiten zwischen der Zustimmung zur Frage nach der Bekanntheit (0,34; Standardabweichung 0,48) und der Ablehnung (0,43; Standardabweichung 0,48) zu beobachten, siehe Abbildung 5.8, für die nur schwache Formen der Ausprägung INFREQUENT MIT SCHWANKUNG berücksichtigt wurden.

Als bekannt bewertete Formen scheinen etwas schneller beurteilt zu werden als Formen, die als unbekannt bewertet wurden: Die Reaktionszeiten bei als bekannt bewerteten Formen clustern stärker um den Mittelwert und streuen weniger. Ein Einfluss der Bekanntheit von Formen auf Reaktionszeiten kann daher nicht ausgeschlossen werden, scheint aber nur gering zu sein. Auch in anderen Studien weisen affirmierende Antworten niedrigere Reaktionszeiten auf als negierende (Wegge 2001: 282–283).

Der Unterschied in den Reaktionszeiten zwischen starken und schwachen Formen lässt sich anhand der Scatterplots in Abbildung 5.9 näher betrachten. Die Abbildung vergleicht die Reaktionszeiten für starke und schwache Formen pro Frequenzausprägung. Ein Punkt gibt die Reaktionszeit einer Person für ein Testitem an. Die grauen Linien in den Plots dienen als Orientierung: Sie zeigen den Bereich an, in dem starke und schwache Formen dieselbe Reaktionszeit aufweisen würden. Wenn starke und schwache Formen keinen Effekt auf die Reaktionszeiten haben, ist also zu erwarten, dass diese sich gleichmäßig um die Linie verteilen. Ist hingegen eine Verzerrung nach oben oder unten zu erkennen, deutet dies auf Unterschiede in den Reaktionszeiten hin.

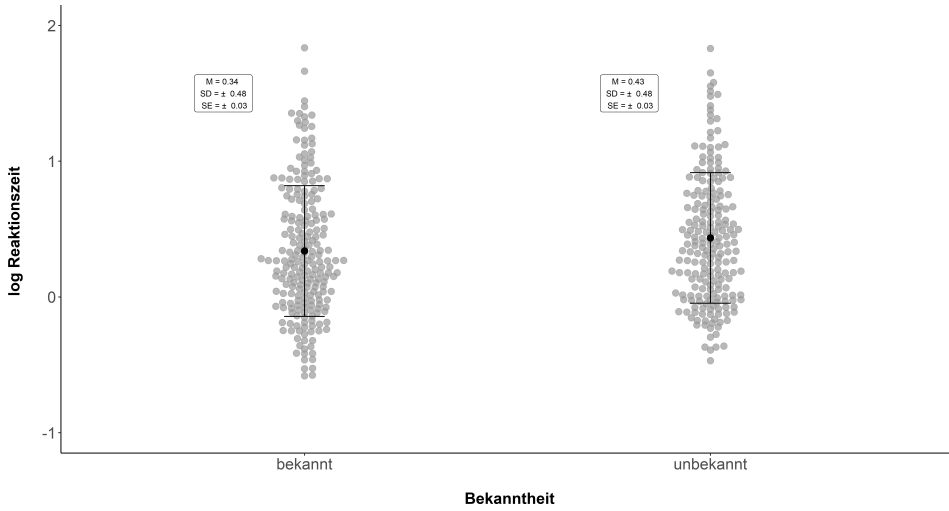


Abbildung 5.8: Reaktionszeiten für schwache Formen von Verben der Ausprägung INFREQUENT MIT SCHWANKUNG nach Bekanntheit

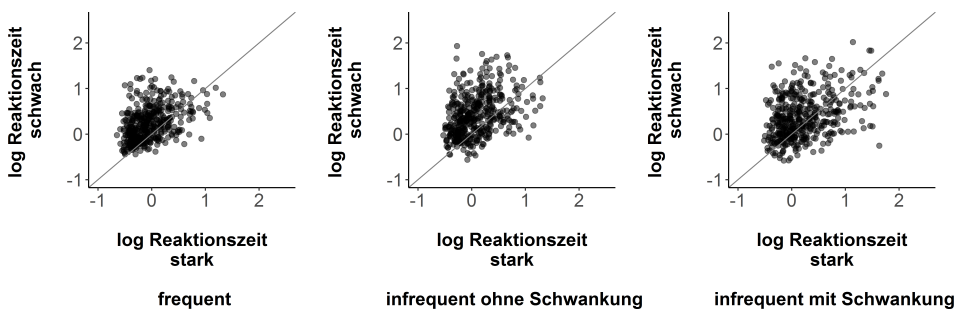


Abbildung 5.9: Scatterplots der Reaktionszeiten für starke und schwache Formen der Testverben

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Aus der Abbildung lässt sich deutlich erkennen, dass die Verben der Ausprägung FREQUENT und INFREQUENT OHNE SCHWANKUNG höhere Reaktionszeiten bei schwachen Formen hervorrufen als bei starken: Der Großteil der Punkte liegt oberhalb der grauen Linie. Für die Ausprägung INFREQUENT MIT SCHWANKUNG ist dies nicht mehr so deutlich zu erkennen: Zwar scheinen einige schwache Formen langsamer beurteilt worden zu sein als starke Formen, jedoch finden sich auch viele Punkte unterhalb der grauen Linie, die anzeigen, dass die starke Form langsamer bewertet wurde als die schwache. Zudem finden sich viele Punkte nahe der grauen Linie, sodass hier von vergleichbaren Reaktionszeiten für die Formen ausgegangen werden kann. Dies ist bspw. bei *glimmen* der Fall: Starke und schwache Formen weisen bei diesem Verb mit Mittelwerten von 0,23 für die schwache Form und 0,24 für die starke mit Standardabweichungen von 0,36 und 0,4 im Schnitt vergleichbare Reaktionszeiten auf. Auch in dieser Darstellung zeigt sich der Effekt der Frequenzausprägungen: Während die Verben der Ausprägung FREQUENT nur bis 1 schwanken, sind bei den anderen beiden Ausprägungen Schwankungen bis 2 zu erkennen.

In Bezug auf die Reaktionszeitunterschiede zwischen starken und schwachen Formen ist ein Blick auf die Differenz in den Reaktionszeiten zwischen starken und schwachen Formen pro Testverb interessant. Tabelle 5.3 zeigt die Differenz in den Reaktionszeiten zwischen starken und schwachen Formen: Bei negativen Zahlen wurde die starke Form schneller bewertet, bei positiven die schwache. Die Tabelle ist nach aufsteigender Differenz sortiert. Somit stehen die Verben oben, bei denen die starke Form schneller bewertet wurde als die schwache, und die Verben unten, bei denen die schwache Form schneller bewertet wurde als die starke.

Es zeigt sich zunächst erneut, dass die starken Formen generell schneller bewertet wurden als die schwachen. Nur bei *einsaugen* (0,01) und *glimmen* (-0,04) ist der Unterschied praktisch inexistent. Interessanterweise sind dies die Verben, die beim Antwortverhalten andere Bewertungen evozierten als die restlichen Testverben: Für *einsaugen* wurden beide Formen als bekannt angesehen, für *glimmen* beide Formen eher als unbekannt bewertet (siehe hierzu ausführlich Abschnitt 5.2.1).

Weiterhin fällt auf, dass die Verben, bei denen die Differenz zwischen starken und schwachen Formen $-0,25$ oder größer war, bis auf eine Ausnahme (*salzen*) den Ausprägungen FREQUENT und INFREQUENT OHNE SCHWANKUNG entstammen. Die Verben der Ausprägung INFREQUENT MIT SCHWANKUNG weisen dagegen (bis auf *salzen*) maximal einen Unterschied von $-0,19$ zwischen den Formen auf. Diese Differenz ist auch für einige Verben der Ausprägung FREQUENT zu finden. Die kleinste Differenz für frequente Verben liegt bei $-0,14$ (*tragen*). Für

Tabelle 5.3: Differenz zwischen den Reaktionszeiten für starke und schwache Formen pro Testverb

Verb	Frequenzausprägung	Differenz
<i>flechten</i>	infrequent ohne Schwankung	-0,53
<i>salzen</i>	infrequent mit Schwankung	-0,51
<i>fahren</i>	frequent	-0,43
<i>melken</i>	infrequent ohne Schwankung	-0,40
<i>halten</i>	frequent	-0,39
<i>kneifen</i>	infrequent ohne Schwankung	-0,36
<i>ziehen</i>	frequent	-0,31
<i>sinken</i>	frequent	-0,31
<i>spinnen</i>	infrequent ohne Schwankung	-0,29
<i>fechten</i>	infrequent ohne Schwankung	-0,27
<i>schmelzen</i>	infrequent ohne Schwankung	-0,27
<i>dreschen</i>	infrequent ohne Schwankung	-0,25
<i>anschwellen</i>	infrequent ohne Schwankung	-0,25
<i>fliegen</i>	frequent	-0,20
<i>schreiben</i>	frequent	-0,19
<i>sprechen</i>	frequent	-0,19
<i>hauen</i>	infrequent mit Schwankung	-0,19
<i>sinnen</i>	infrequent mit Schwankung	-0,16
<i>tragen</i>	frequent	-0,14
<i>gären</i>	infrequent mit Schwankung	-0,14
<i>quellen</i>	infrequent mit Schwankung	-0,12
<i>weben</i>	infrequent mit Schwankung	-0,12
<i>glimmen</i>	infrequent mit Schwankung	-0,04
<i>einsaugen</i>	infrequent mit Schwankung	0,01

Verben der Ausprägung INFREQUENT OHNE SCHWANKUNG ist die kleinste Differenz mit $-0,25$ hingegen deutlich höher. Der geringe Unterschied zwischen starken und schwachen Formen bei *tragen* könnte auf die Identität im Stammvokal zwischen Infinitiv und Partizip II zurückgehen, allerdings weisen auch andere frequente Verben (*halten*, *fahren*) Vokalidentität auf, bei denen eine höhere Differenz zwischen den Reaktionszeiten der starken und schwachen Form gemessen wurde. Der Blick auf die Differenzen zeigt insgesamt, dass Verben der Ausprägungen FREQUENT und INFREQUENT OHNE SCHWANKUNG vergleichbare Unterschiede in den Reaktionszeiten zwischen starken und schwachen Formen hervorrufen, während Verben der Ausprägung INFREQUENT MIT SCHWANKUNG deutlich geringere Reaktionszeitunterschiede als die anderen Ausprägungen aufweisen. Die einzige Ausnahme hiervon ist *salzen* ($-0,51$), das jedoch im Antwortverhalten eher Verben der Ausprägung INFREQUENT OHNE SCHWANKUNG gleicht als Verben der Ausprägung INFREQUENT MIT SCHWANKUNG (siehe Abschnitt 5.2.1).

Neben *salzen* sind in der Analyse des Antwortverhaltens *hauen* und *quellen* aufgefallen (siehe Abschnitt 5.2.1): Diese wurden aufgrund ihrer Ratios von starken zu schwachen Formen als INFREQUENT MIT SCHWANKUNG in die Studie integriert, allerdings war das Antwortverhalten für diese Verben vergleichbar mit dem Antwortverhalten für die Verben der Ausprägung INFREQUENT OHNE SCHWANKUNG. Um auszuschließen, dass das Antwortverhalten bei diesen Verben die Reaktionszeiten für die Ausprägung INFREQUENT MIT SCHWANKUNG verzerrt hat, wurden die Verben *salzen*, *hauen* und *quellen* zu Verben der Ausprägung INFREQUENT OHNE SCHWANKUNG umkodiert.¹⁴ Abbildung 5.10 zeigt die Beeswarmplots mit der Umkodierung. Zusätzlich werden das arithmerische Mittel (M), die Standardabweichung (SD) und der Standardfehler berichtet. Aufgrund der Umkodierung ist der Beeswarmplot der Ausprägung INFREQUENT MIT SCHWANKUNG schmaler und der Beeswarmplot der Ausprägung INFREQUENT OHNE SCHWANKUNG breiter als in Abbildung 5.6.

Es zeigt sich nach wie vor dieselbe Verteilung. Das abweichende Antwortverhalten bei *salzen*, *hauen* und *quellen* scheint die Reaktionszeiten also nicht verzerrt zu haben. Zudem wurde ein möglicher Einfluss der Stimulusreihenfolge und der Wortlänge, die bedingt durch die starken und schwachen Formen leicht schwankt, visuell überprüft. Hierbei wurde jedoch kein Hinweis auf einen Effekt gefunden, siehe Anhang I.1 für die entsprechenden Grafiken.

Der Einfluss von Frequenz und Flexion auf Reaktionszeiten wird mithilfe eines gemischten linearen Modells überprüft. Als feste Effekte werden die Frequenzausprägung sowie die Flexion genutzt, als zufällige Effekte Teilnehmer_in und

¹⁴*Glimmen* wurde nicht umkodiert, da für dieses Verb nicht nur die schwachen, sondern auch die starken Formen häufig als unbekannt bewertet wurden.

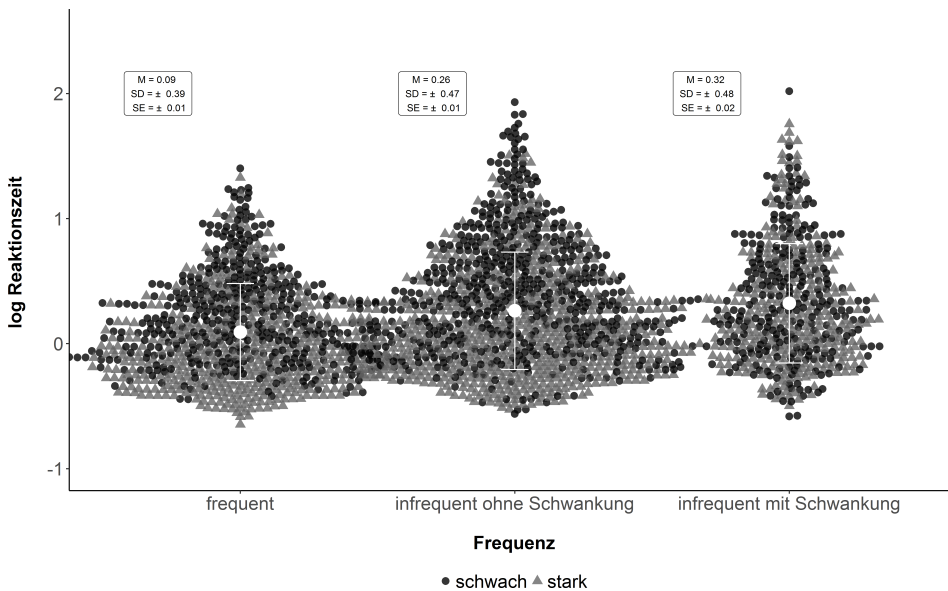


Abbildung 5.10: Reaktionszeiten für die umkodierte Frequenzausprägungen

Versuchitem (Lemma): Reaktionszeit \sim Frequenz + Flexion + (1|Proband_in) + (1|Item). Für beide zufälligen Effekte werden nur *random intercepts* berechnet. *Random slopes* werden für Teilnehmer_innen aufgrund von *singular fits* nicht genutzt, bei Versuchitems sind *random slopes* nicht sinnvoll, weil die Versuchitems über die Frequenzausprägungen hinweg nicht konstant sind.

Die Standardabweichung der Versuchitems beträgt nur 0,06. Bezüglich der Teilnehmer_innen ist mit 0,25 eine deutlich größere Standardabweichung um den y-Achsenabschnitt zu beobachten. Das ist zu erwarten, da Reaktionszeiten individuell stark variieren. Trotz der zufälligen Effekte bleibt mit 0,34 eine relativ hohe Residuenvarianz. Tabelle 5.4 zeigt die geschätzten Werte für die festen Effekte.¹⁵

Das Referenzlevel ist die Ausprägung INFREQUENT OHNE SCHWANKUNG mit schwachen Flexionsformen. Hierfür wird ein y-Achsenabschnitt von 0,38 mit einem Standardfehler von 0,04 geschätzt. Für die Ausprägung FREQUENT ändert sich der Wert um $-0,16$ mit einem Standardfehler von 0,03 auf 0,22. Die Reaktionszeiten werden für die Ausprägung FREQUENT also niedriger geschätzt als für INFREQUENT OHNE SCHWANKUNG. Für die Ausprägung INFREQUENT MIT SCHWAN-

¹⁵Die Abkürzung *df* in der Tabelle steht für Freiheitsgrade (*degrees of freedom*). Freiheitsgrade geben die Anzahl an Werten in der Berechnung an, die frei variieren können.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Tabelle 5.4: Werte des Modells für die Reaktionszeiten in der Frequenzstudie

	Wert	SE	<i>t</i>	df	<i>p</i>
(Intercept = frequent ohne Schwankung, schwach)	0,38	0,04	8,91	75,88	<0,01
Frequenz: frequent	-0,16	0,03	-4,75	23,26	<0,01
Frequenz: frequent mit Schwankung	0,05	0,03	1,42	23,26	0,17
Flexion: stark	-0,25	0,01	-18,77	2467,93	<0,01

KUNG ändert sich der Wert mit 0,05 und einem Standardfehler von 0,03 im Vergleich zur Ausprägung INFREQUENT OHNE SCHWANKUNG hingegen kaum.

Für starke Formen ändert sich der Wert im Vergleich zum Referenzlevel deutlich um -0,25 mit einem Standardfehler von 0,01 auf 0,23. Starke Formen werden also deutlich schneller bewertet als schwache. Für die Ausprägung FREQUENT und für starke Flexion liegen die *p*-Werte im Vergleich zum Referenzlevel (INFREQUENT OHNE SCHWANKUNG, schwache Formen) unter dem α -Level von 0,01. Die Daten sind also unter der Nullhypothese, nach der die Frequenzausprägungen sowie die Flexion keinen Einfluss auf Reaktionszeiten haben, unwahrscheinlich. Für die Ausprägung INFREQUENT MIT SCHWANKUNG liegt der *p*-Wert mit 0,17 hingegen deutlich über dem α -Level.

Wenn nur die festen Effekte berücksichtigt werden, ist die Effektstärke mit $R^2m = 0,117$ recht gering. Inklusive der zufälligen Effekte ist sie mit $R^2c = 0,443$ deutlich höher.¹⁶ Dennoch bleibt ein Anteil von 0,557 der Varianz nicht erklärt. Da Reaktionszeiten hochvariabel sind und durch viele Faktoren beeinflusst werden, ist ein hoher Anteil nicht erklärter Varianz jedoch erwartbar.

Abbildung 5.11 zeigt die vorhergesagten Reaktionszeiten. Schwache Formen sind als Kreise dargestellt, starke als Dreiecke.

¹⁶ R^2m und R^2c geben an, wie viel Varianz durch das Modell erklärt wird. Hierfür wird die Residuenquadratsumme berechnet, indem die Residuen aufsummiert und quadriert werden. Die Residuenquadratsumme des Modells wird durch die Residuenquadratsumme des Nullmodells geteilt, das lediglich den Mittelwert aller Reaktionszeiten als Vorhersage für die Datenpunkte nutzt. Je kleiner die Residuenquadratsumme des Modells im Vergleich zum Nullmodell ist, desto geringer ist der berechnete Wert. Dieser Wert wird von 1 abgezogen, um R zu erhalten: Ist R hoch, ist die erklärte Varianz hoch, ist R hingegen niedrig, ist auch die erklärte Varianz niedrig (Winter 2020: 75–77). Beim R^2m werden nur die festen Effekte berücksichtigt, beim R^2c zusätzlich zu den festen Effekten auch die zufälligen (Winter 2020: 264).

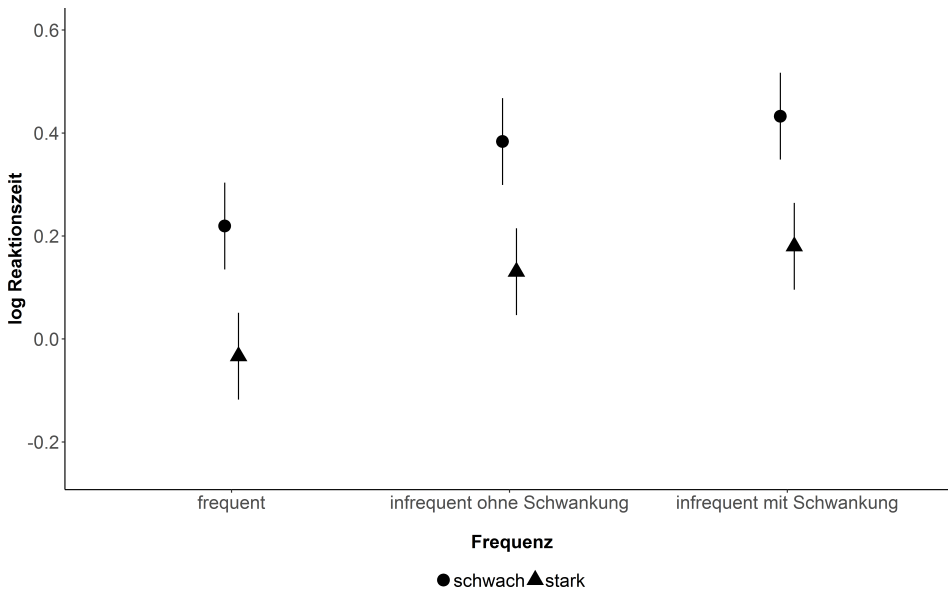


Abbildung 5.11: Vorhergesagte Reaktionszeiten in der Frequenzstudie

Die vorhergesagten Reaktionszeiten unterscheiden sich deutlich pro Frequenz-
ausprägung: Die starken und schwachen Formen neigen für die Ausprägung FRE-
QUENT mit $-0,03$ (stark) und $0,22$ (schwach) jeweils zu niedrigeren Reaktionszei-
ten als die starken und schwachen Formen der anderen Frequenzausprägungen.
Die vorhergesagten Reaktionszeiten für die Ausprägungen INFREQUENT OHNE
SCHWANKUNG und INFREQUENT MIT SCHWANKUNG unterscheiden sich mit $0,13$
bzw. $0,38$ und $0,18$ bzw. $0,43$ zwar leicht, die 95 %-Konfidenzintervalle überlap-
pen sich jedoch deutlich. Aus Abbildung 5.11 ist außerdem zu erkennen, dass
für starke und schwache Formen unterschiedliche Reaktionszeiten vorhergesagt
werden. Dies ist besonders deutlich bei den schwachen Formen der Ausprägung
FREQUENT zu erkennen, für die ähnlich hohe Reaktionszeiten wie für die starken
Formen der anderen Frequenzausprägungen vorhergesagt werden.

Die visuelle Betrachtung der Daten deutet darauf hin, dass Frequenz und Fle-
xion nicht unabhängig voneinander Einfluss auf die Reaktionszeiten nehmen,
sondern interagieren. So scheinen starke Formen von Verben der Ausprägung
FREQUENT schneller bewertet worden zu sein als schwache Formen von Verben
derselben Ausprägung. Allerdings deuten die Daten nicht auf einen Unterschied
in den Reaktionszeiten zwischen starken und schwachen Verbformen der Aus-
prägung INFREQUENT MIT SCHWANKUNG hin. Das obige Modell berücksichtigt
dies nicht, da es keine Interaktion vorsieht und daher den Einfluss der Frequenz

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

und der Flexion unabhängig voneinander betrachtet. Daher wird als explorative Analyse ein zweites gemischtes lineares Modell gerechnet, das eine Interaktion zwischen Form und Frequenzausprägung vorsieht: Reaktionszeit ~ Frequenz * Flexion + (1|Proband_in) + (1|Item). Dieses Modell wurde nicht als konfirmatorische Analyse gerechnet, da für die vorliegende Untersuchung in erster Linie Effekte der Frequenzausprägung und weniger Interaktionen mit starken und schwachen Formen von Interesse sind. Da es sich um eine explorative Analyse handelt, werden keine *p*-Werte berichtet. Die zufälligen Effekte sind identisch mit dem Modell ohne Interaktion. Tabelle 5.5 zeigt die vom Modell berechneten Werte für das Modell, Abbildung 5.12 die vorhergesagten Reaktionszeiten inklusive der 95 %-Konfidenzintervalle. Die Kreuztabelle der Werte (Tabelle I.1) befindet sich in Anhang I.1. Die Standardfehler sind für alle Werte mit bis zu 0,04 recht gering.

Tabelle 5.5: Werte des Modells für die Reaktionszeiten mit Interaktion zwischen Frequenz und Flexion

	Wert	SE	<i>t</i>	df
(Intercept = infrequent ohne Schwankung, schwach)	0,42	0,04	9,57	83,41
Frequenz: frequent	-0,19	0,04	-5,04	34,91
Frequenz: infrequent mit Schwankung	-0,04	0,04	-0,94	34,91
Flexion: stark	-0,33	0,02	-14,17	2467,94
Frequenz: frequent & Flexion: stark	0,06	0,03	1,74	2467,94
Frequenz: infrequent mit Schwankung & Flexion: stark	0,17	0,03	5,20	2467,94

Wie bei dem Modell ohne Interaktion lässt sich ein klarer Effekt der Frequenzausprägungen erkennen: Die starken Verbformen der Ausprägung *FREQUENT* werden durchweg schneller beurteilt als die starken Verbformen der anderen Frequenzausprägungen. Dasselbe ist für die schwachen Verbformen festzuhalten. Durch die Interaktion zeigt sich, dass die Reaktionszeiten bei Verben der Ausprägung *INFREQUENT MIT SCHWANKUNG* weniger stark durch die Formen beeinflusst werden als die Reaktionszeiten bei Verben der anderen Frequenzausprägungen: Bei *INFREQUENT MIT SCHWANKUNG* ist der Wert für schwache Formen bei 0,38 und für starke bei 0,23, während bei den Verben der Ausprägung *INFREQUENT OHNE SCHWANKUNG* schwache Formen bei 0,42 und starke nur bei 0,09 liegen. Bei Verben der Ausprägung *FREQUENT* ist der Unterschied mit 0,23 zu -0,04 ähnlich

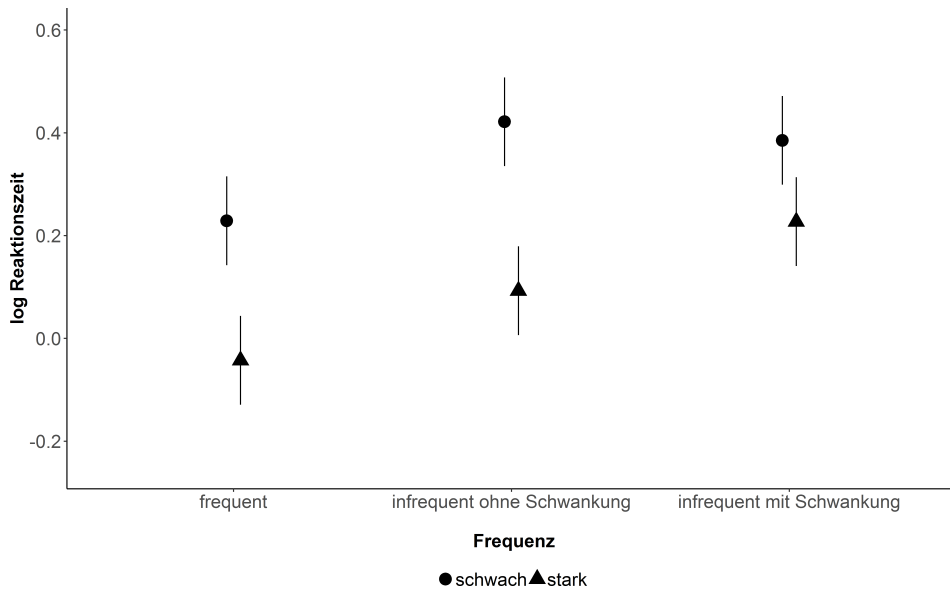


Abbildung 5.12: Vorhergesagte Reaktionszeiten mit Interaktion zwischen Frequenz und Flexion

deutlich wie bei *INFREQUENT OHNE SCHWANKUNG*. Wieder zeigt sich eine Überlappung der 95 %-Konfidenzintervalle für die Ausprägungen *INFREQUENT OHNE SCHWANKUNG* und *INFREQUENT MIT SCHWANKUNG*. Zudem überlappen sich bei der Ausprägung *INFREQUENT MIT SCHWANKUNG* die 95 %-Konfidenzintervalle für die starken und schwachen Formen, sodass ein systematischer Einfluss der Flexion auf die Reaktionszeiten für diese Ausprägung fraglich ist. Dies bestätigt die visuelle Betrachtung der Daten, in der kein Einfluss starker und schwacher Formen zu erkennen war. Die Effektstärke des Modells hat sich durch die Interaktion mit einem R^2m von 0,123 und einem R^2c von 0,444 kaum geändert (vorherige Werte: $R^2m = 0,117$; $R^2c = 0,443$). Die festen Effekte erklären mit der Interaktion jedoch etwas mehr Varianz.

Insgesamt zeigt die Analyse der Reaktionszeiten, dass die Frequenzausprägungen Einfluss auf Reaktionszeiten nehmen: Frequente Testverben werden schneller bewertet als infrequente. Zwischen infrequenten Testverben mit und ohne Schwankung lässt sich hingegen kein Unterschied in den Reaktionszeiten messen. Zudem zeigt sich, dass starke Formen schneller bewertet werden als schwache. Die Analyse bestätigt damit insgesamt die Hypothesen aus Abschnitt 4.2.1. Anders als vorab angenommen scheint die Bewertung als bekannt oder unbekannt die Reaktionszeiten zu beeinflussen, zudem zeigen sich Reaktionszeitunter-

schiede zwischen starken und schwachen Formen. Diese überdecken den Vorteil für frequente Verben jedoch nicht: Starke und schwache Formen der frequenten Verben wurden jeweils schneller beurteilt als starke und schwache Verbformen der anderen Ausprägungen. Die visuelle sowie die explorative Analyse legen nahe, dass starke Formen nur bei frequenten und infrequenten Testverben ohne Schwankung schneller bewertet werden als schwache Formen und sich die Reaktionszeiten für starke und schwache Formen bei infrequenten Verben mit Schwankung einander annähern. Im folgenden Abschnitt werden die Ergebnisse der sentence-maze-Studie zu Prototypizitätseffekten vorgestellt.

5.3 Prototypizität

In diesem Abschnitt werden das Antwortverhalten und die Reaktionszeiten in der sentence-maze-Studie ausgewertet. Die Studie hat zum Ziel, anhand der Auxiliarselektion von *haben* und *sein* in Sätzen mit verschiedenen Transitivitätsgraden den Einfluss der Prototypizität auf Variation zu untersuchen. Die in diesem Abschnitt diskutierten Ergebnisse beziehen sich auf die in Abschnitt 4.3.5 vorgestellten Proband_innen. Die Stichprobengröße beträgt 107 Personen. In der Studie wählten die Proband_innen zwischen *hat* und *ist* in Sätzen mit unterschiedlichen Transitivitätsgraden. Prototypisch transitive Testsätze enthalten ein belebtes Objekt und ein Direktionaladverbial (*Er sah, wie die Mutter die Kinder zur Schule gefahren hat/*ist*). Prototypisch intransitive Sätze enthalten kein Objekt (*Ich ging davon aus, dass die Oma am Sonntag zur Kur gefahren ist/*hat*). Zudem werden zwei ambige Testsätze genutzt: AMBIG I enthält ein Objekt, das ambig zwischen Objekt und Instrument ist (*Er erzählte mir, dass seine Tante eine Zeit lang ein Cabrio gefahren hat/ist*), AMBIG II enthält ein Akkusativobjekt mit Referenz auf eine unbelebte Entität, das als Adverbial aufgefasst werden kann (*In den Nachrichten sagten sie, dass der neue Fahrer die Strecke in Rekordzeit gefahren ist/hat*). Die vier Transitivitätsausprägungen (TRANSITIV, AMBIG I, AMBIG II und INTRANSITIV) werden jeweils mit drei Testverben präsentiert (*fahren, fliegen, reiten*) (siehe Abschnitt 4.3 für weitere Erläuterungen).

Für transitive Sätze werden *haben*-Antworten erwartet, für intransitive *sein*-Antworten. Für die ambigen Sätze wird basierend auf der Prästudie eine Tendenz zu *sein*-Antworten antizipiert, allerdings keine so klare wie bei den intransitiven Sätzen (siehe Abschnitt 3.3 für weitere Erläuterungen zur Selektion von *haben* und *sein*). Hinsichtlich der Reaktionszeiten wurden ursprünglich kürzere Reaktionszeiten für die Prototypen (TRANSITIV und INTRANSITIV) und längere für die Peripherie (AMBIG I und AMBIG II) erwartet. Basierend auf der Prästudie werden

allerdings kürzere Reaktionszeiten für die intransitiven Sätze erwartet als für die anderen Ausprägungen. Für die Ausprägungen TRANSITIV und AMBIG II werden vergleichbare Reaktionszeiten angenommen und für AMBIG I werden die höchsten Reaktionszeiten erwartet. Die erhöhten Reaktionszeiten der transitiven und ambigen Sätze im Vergleich zu den intransitiven könnte durch die Komplexität der Sätze bedingt sein. Sollten sich geringere Reaktionszeiten für die Ausprägung AMBIG II im Vergleich zur Ausprägung AMBIG I zeigen, könnte dies darauf hindeuten, dass AMBIG II von den Proband_innen nicht als ambig aufgefasst wird (siehe Abschnitt 4.3.1 für genauere Erläuterungen zu den Hypothesen).

5.3.1 Antwortverhalten

Die 107 Proband_innen wählten für jedes Versuchitem (*fahren, fliegen, reiten*) in jeder Transitivitätsausprägung zwischen *hat* und *ist*, sodass 107 Antworten pro Testverb und Transitivitätsausprägung vorliegen. Abbildung 5.13 zeigt das Antwortverhalten in einem Waffleplot. Ein Rechteck bildet jeweils eine Antwort ab. Rechtecke in dunkelgrau geben die Antworten wieder, bei denen sich die Proband_innen für *hat* entschieden, hellgraue Rechtecke dementsprechend Antworten für *ist*.

Es zeigt sich ein klarer Einfluss der Transitivität: Für die Ausprägung TRANSITIV wird *hat* gewählt, für INTRANSITIV *ist*. Während für TRANSITIV auch vereinzelt *sein*-Antworten gegeben werden, sind *haben*-Antworten für INTRANSITIV bis auf eine Ausnahme nicht zu finden. Zudem fällt das Testverb *reiten* auf, bei dem auch für TRANSITIV 44 % der Proband_innen *ist* wählen.

Der Blick auf AMBIG I und AMBIG II zeigt, dass hier vorwiegend *ist* gewählt wird: Bei AMBIG I liegt der Anteil von *hat* für die Testverben *fliegen* und *fahren* bei ca. 14 %, bei AMBIG II nur bei 0,03 %. Die Antwortverteilung für AMBIG I weicht von den Ergebnissen von Gillmann (2016: 286–291) ab. In den Sätzen der Ausprägung AMBIG I wird ein Konkretum genutzt, das als Fortbewegungsmittel interpretiert werden kann (*Er erzählte mir, dass seine Tante eine Zeit lang ein Cabrio gefahren hat/ist*). In ihrem Korpus stellt Gillmann (2016: 286–291) für *fahren* in solchen Sätzen eine klare Tendenz zur *haben*-Selektion fest und für *fliegen* Variation zwischen beiden Auxiliaren. Hinsichtlich der Ausprägung AMBIG II bestätigen sich die Ergebnisse von Gillmann (2016: 299–301) dagegen. Die Sätze der Ausprägung AMBIG II enthalten als Akkusativergänzung Abstrakta, die Pfad-Akkusative darstellen (*In den Nachrichten sagten sie, dass der neue Fahrer die Strecke in Rekordzeit gefahren ist/hat*). Für solche Sätze stellt Gillmann (2016: 299–301) ebenfalls eine hohe *sein*-Selektion fest.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

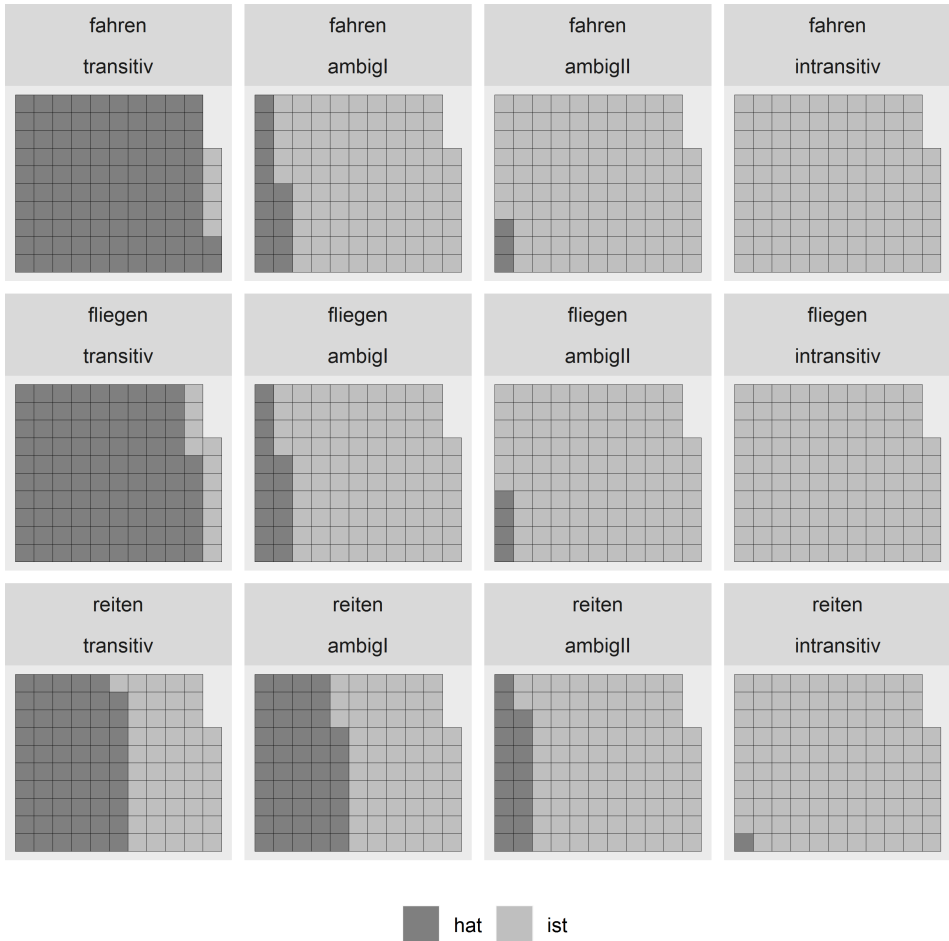


Abbildung 5.13: Antwortverhalten in der Prototypizitätsstudie

Auch in Bezug auf die ambigen Sätze verhält sich *reiten* anders als *fliegen* und *fahren*: Mit 43,9 % wählten weit mehr Proband_innen für AMBIG I *hat* als für *fliegen* und *fahren*. Auch für AMBIG II ist der Anteil an *haben*-Antworten bei *reiten* mit 16,8 % im Vergleich zu den anderen Testverben hoch.¹⁷ In den Daten von Gillmann (2016) tendiert *reiten* ebenfalls etwas stärker zu *haben* als andere Fortbewegungsverben. Als Erklärung hierfür können die geringere Tokenfrequenz sowie die Semantik von *reiten* angeführt werden, da es nicht nur auf eine Fortbewegung, sondern auch auf eine Aktivität referiert (Gillmann 2016: 273–280).

Um das Antwortverhalten für die einzelnen Testverben näher zu untersuchen, werden die Kombinationsmöglichkeiten der Auxiliare betrachtet, z. B. könnte eine Person folgende Kombination für ein Testverb gewählt haben: TRANSITIV: *hat*, AMBIG I: *sein*, AMBIG II: *sein*, INTRANSITIV: *sein*. Für Tabelle 5.6 wurde gezählt, wie häufig welche der möglichen Kombinationen pro Testverb gewählt wurde. Die Tabelle ist nach der Vorkommenshäufigkeit der Kombinationen für das Testverb *fahren* sortiert. Die Kombinationen beziehen sich auf die Abfolge TRANSITIV, AMBIG I, AMBIG II, INTRANSITIV.

Bei *fahren* und *fliegen* haben mit 80,4 % und 72,9 % die meisten Proband_innen nur für die Ausprägung TRANSITIV *hat* gewählt und für die anderen Ausprägungen *ist*. Mit 12,2 % bzw. 13,1 % folgt weit abgeschlagen eine Kombination, bei der *hat* für die Ausprägungen TRANSITIV sowie AMBIG I und *ist* für die Ausprägungen AMBIG II sowie INTRANSITIV gewählt wurde. Nur wenige Proband_innen wählen für alle Sätze *ist* (3,7 % bzw. 8,4 %). Andere Kombinationen liegen unter 4 %.¹⁸

¹⁷Die 20 Proband_innen, die ausgeschlossen wurden, da sie in Bayern, Baden-Württemberg, Rheinland-Pfalz oder im Saarland aufgewachsen sind bzw. leben (siehe Abschnitt 4.3.5), weisen ein ähnliches Antwortverhalten auf: Für intransitive Sätze wird bei *fahren* und *fliegen* ausnahmslos *ist* gewählt, für transitive *hat*. Auch AMBIG II ruft ausnahmslos *sein*-Antworten hervor. Bei AMBIG I finden sich nur vereinzelte *haben*-Antworten (vier für *fahren*, eine für *fliegen*). Die Tendenz zu *sein* in den ambigen Sätzen könnte bei diesen Proband_innen daher noch deutlicher sein als bei den Proband_innen in der Stichprobe. *Reiten* evoziert ein von den anderen Versuchsverben abweichendes Antwortverhalten: Die intransitiven Sätze rufen ausnahmslos *sein* hervor, aber *hat* ist in den anderen Ausprägungen häufiger als bei *fahren* und *fliegen*: Bei AMBIG II wird es in vier Antworten gewählt, bei AMBIG I in 13 Antworten. Der transitive Satz fällt mit relativ wenig *haben*-Antworten (11) auf. Insgesamt scheinen die Antworten denen der Proband_innen in der Stichprobe zu ähneln, allerdings könnte das Antwortverhalten aufgrund der wenigen Proband_innen verzerrt sein. Für Reaktionszeiten wird aufgrund der kleinen Stichprobengröße kein Vergleich zwischen den Proband_innen der Hauptstudie und den aus der Studie ausgeschlossenen Proband_innen angestellt.

¹⁸Ein Überblick über die Auxiliarkombinationen bei den Proband_innen, die aufgrund des Bundeslands aus der Stichprobe ausgeschlossen wurden, befindet sich in Tabelle I.2 in Anhang I.2. Wie bei den Proband_innen in der Hauptstudie tendieren *fahren* und *fliegen* hier eindeutig zu *hat* für transitive Sätze und *ist* für alle anderen Ausprägungen, während bei *reiten* mehr Variation in den Auxiliarkombinationen vorherrscht.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Tabelle 5.6: Häufigkeit der Auxiliarkombinationen für die Abfolge TRANSITIV, AMBIG I, AMBIG II, INTRANSITIV

Kombinationsmöglichkeit	<i>fahren</i>	<i>fliegen</i>	<i>reiten</i>
hat_ist_ist_ist	86 (80,4 %)	78 (72,9 %)	23 (21,5 %)
hat_hat_ist_ist	13 (12,2 %)	14 (13,1 %)	21 (19,6 %)
ist_ist_ist_ist	4 (3,7 %)	9 (8,4 %)	31 (29,0 %)
hat_ist_hat_ist	2 (1,9 %)	4 (3,7 %)	4 (3,7 %)
hat_hat_hat_ist	1 (0,9 %)	0 (0,0 %)	11 (10,3 %)
ist_hat_ist_ist	1 (0,9 %)	2 (1,9 %)	13 (12,2 %)
ist_hat_hat_ist	0 (0,0 %)	0 (0,0 %)	2 (1,9 %)
ist_ist_hat_ist	0 (0,0 %)	0 (0,0 %)	1 (0,9 %)
ist_ist_ist_hat	0 (0,0 %)	0 (0,0 %)	1 (0,9 %)
gesamt	107 (100 %)	107 (100 %)	107 (100 %)

Bei *reiten* sind die Auxiliarkombinationen weniger eindeutig verteilt: Mit 29 % wählen die meisten Proband_innen für alle Ausprägungen *ist*. Ein etwas geringerer Anteil (21,5 %) wählt nur für die Ausprägung TRANSITIV *hat* und für die restlichen Sätze *ist*. Mit 19,6 % ebenfalls recht häufig sind *haben*-Antworten für die Ausprägungen TRANSITIV sowie AMBIG I und *sein*-Antworten für die Ausprägungen AMBIG II sowie INTRANSITIV. Deutlich seltener ist die Kombination *hat* für AMBIG I und *ist* für alle anderen Ausprägungen (12,2 %) sowie die Kombination *hat* für alle Ausprägungen bis auf INTRANSITIV (10,3 %). Die anderen in der Tabelle aufgeführten Kombinationen liegen unter 4 %.

Während die Auxiliarkombinationen bei *fahren* und *fliegen* auf eine klare Präferenz für eine Trennung zwischen TRANSITIV (*haben*) und den anderen Ausprägungen (*sein*) hindeuten und einige Proband_innen die Auxiliargrenze zwischen TRANSITIV sowie AMBIG I (*haben*) und AMBIG II sowie INTRANSITIV (*sein*) ziehen, zeigt sich bei *reiten* zusätzlich zu diesen Antwortkombinationen eine Tendenz dazu, unabhängig von der Transitivität *sein* zu nutzen. Diese Antwortkombination ist bei *fahren* und *fliegen* nur vereinzelt zu finden. Zudem ist für *reiten* bei weitem keine Präferenz zu einem Wechsel des Auxiliars zwischen TRANSITIV und anderen Ausprägungen zu erkennen: Stattdessen sind drei Kombinationen (durchgängige *sein*-Antworten; *sein*-Antworten für alle Ausprägungen bis auf TRANSITIV; *hat* für TRANSITIV und AMBIG I, aber *ist* für AMBIG II und INTRANSITIV) ähnlich häufig. Dass *reiten* generell mehr Variation in der Auxiliarwahl zulässt als die anderen Testverben, verdeutlichen außerdem Kombinationen, die bei *reiten* zu ca.

10 % gewählt wurden, aber für *fahren* und *fliegen* maximal zweimal (*hat* für alle Ausprägungen bis auf INTRANSITIV, *ist* für alle Ausprägungen bis auf AMBIG I).

Die häufige Wahl von *sein* in allen Transitivitätsausprägungen bei *reiten* lässt den Schluss zu, dass Transitivität für einige Proband_innen keinen Einfluss auf die Auxiliarwahl bei *reiten* hat. In Bezug auf die Auxiliarwahl bei transitiv gebrauchtem *reiten* ist ein Blick auf die Korpusstudie von Gillmann (2016) interessant: In ihrem Korpus sind nur 19 *reiten*-Belege, die ein Objekt enthalten, das auf eine belebte Entität verweist (*Ich bin/habe ein Pferd geritten*).¹⁹ In zwei der 19 Belege ist das Auxiliar *sein* zu finden (Gillmann 2016: 282–283). *sein*-Selektion ist daher im Korpus nicht ausgeschlossen, aber weitaus weniger prominent als in der vorliegenden Studie. Sätze mit Direktionaladverbial wie der Satz der Ausprägung TRANSITIV (*Ich gehe davon aus, dass das Mädchen das Pferd zur Koppel geritten hat/ist*) kommen im Korpus nicht vor, sodass diesbezüglich kein Vergleich angestellt werden kann.

Die durchgängigen *sein*-Antworten bei *reiten* könnten auch durch die Transitivität des Testsatzes bedingt sein: Der transitive Satz könnte nur eingeschränkt als transitiv wahrgenommen worden sein. Pferde sind zwar belebt, aber nicht menschlich, sodass sie nicht prototypisch in der Patiensrolle vorkommen. Zudem lässt die Studie von Gillmann (2016) darauf schließen, dass *reiten* kaum mit Direktionaladverbial genutzt wird, sodass der transitive Satz eventuell als ungewöhnlich wahrgenommen wurde. Dies könnte auch eine Erklärung für die auffällige Auxiliarkombination bei *reiten* sein, bei der nur für AMBIG I *hat* und ansonsten *ist* gewählt wurde.

Neben dem Einfluss der Prototypizität könnte auch Tokenfrequenz eine Rolle spielen: *reiten* ist weniger tokenfrequent als *fliegen* und *fahren*. Es ist daher möglich, dass die Auxiliare weniger stark gefestigt sind als bei den frequenteren Verben.

Der Einfluss der Transitivität auf das Antwortverhalten wird für die konfirmatorische statistische Analyse mithilfe eines generalisierten gemischten linearen Modells überprüft.²⁰ Als fester Effekt, der Einfluss auf die Wahl zwischen *haben* und *sein* nimmt, wird die Transitivitätsausprägung der Sätze eingeflochten: Antwortverhalten ~ Transitivität + (1|Proband_in). Für Teilnehmer_innen werden

¹⁹Diese entsprechen dem Satz der Ausprägung AMBIG I in der Studie (*Mir wurde erzählt, dass mein Onkel früher ein schwarzes Pferd geritten hat/ist*).

²⁰*Reiten* evozierte bereits in der Prästudie ein anderes Antwortverhalten als die anderen Testverben. Aufgrund dessen wurde in der Präregistrierung ein Modell angegeben, das *reiten* nicht in der statistischen Analyse berücksichtigt. Das Modell ohne *reiten* konvergierte jedoch auch mit verschiedenen Optimierungsmethoden nicht, weshalb alle Testverben in das Modell einbezogen wurden.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

random intercepts genutzt, *random slopes* werden aufgrund von Konvergenzproblemen und *singular fits* nicht berücksichtigt. Aus demselben Grund werden die Versuchstems (*reiten*, *fahren*, *fliegen*) nicht als zufälliger Effekt einbezogen. Das Modell ist daher in seiner Aussagekraft eingeschränkt. Die Teilnehmer_innen variieren mit einer Standardabweichung von 0,8 um den y-Achsenabschnitt. Tabelle 5.7 zeigt die *log odds* für *sein*-Antworten, Abbildung 5.14 die auf den *log odds* basierenden Wahrscheinlichkeiten für *sein*-Antworten in Abhängigkeit von der Transitivitätsausprägung.

Tabelle 5.7: Werte des Modells für *sein*-Antworten in der Prototypizitätsstudie

	Wert	SE	<i>z</i>	<i>p</i>
(Intercept = ambig I)	1,29	0,17	7,83	<0,01
Transitivität: ambig II	1,44	0,26	5,62	<0,01
Transitivität: intransitiv	4,78	1,01	4,72	<0,01
Transitivität: transitiv	-2,87	0,23	-12,57	<0,01

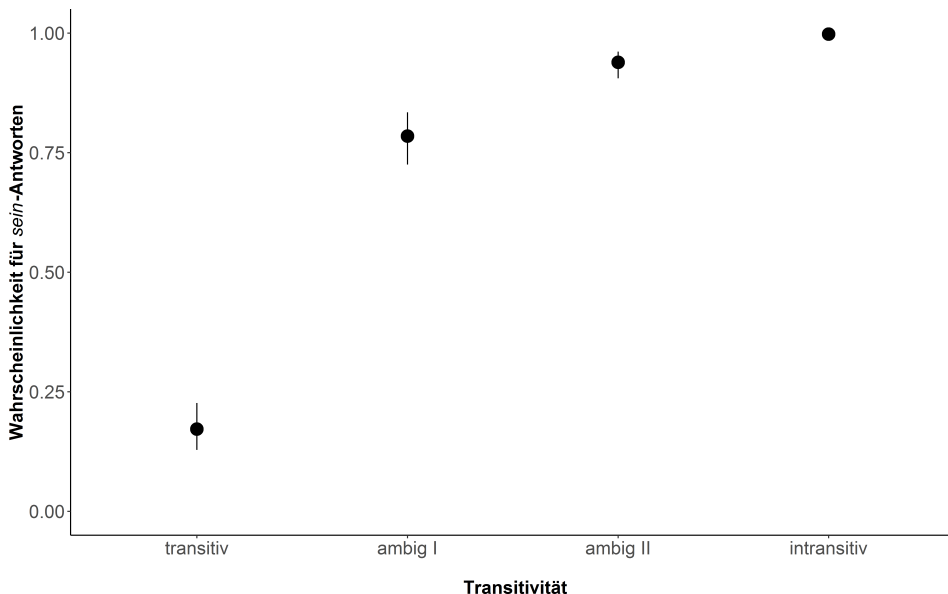


Abbildung 5.14: Wahrscheinlichkeit für *sein*-Antworten in der Prototypizitätsstudie

Für die Ausprägung TRANSITIV wird ein Wert von 0,17 vorhergesagt, d. h. die Wahrscheinlichkeit für *sein*-Antworten ist gering. Dementsprechend ist die Wahrscheinlichkeit für *haben*-Antworten hoch. Das Gegenteil ist bei INTRANSITIV und AMBIG II der Fall: Hier werden Werte von über 0,9 vorhergesagt. Auch AMBIG I zeigt mit 0,78 eine klare Tendenz dazu auf, *sein* zu evozieren. Die 95 %-Konfidenzintervalle für die einzelnen Werte sind jeweils sehr klein und keine Überlappungen zu erkennen. Dies deutet auf ein systematisches Antwortverhalten je nach Transitivitätsausprägung hin. Für alle Ausprägungen liegt der p -Wert unter dem α -Level von 0,01, die Daten sind also unter der Nullhypothese (die Transitivitätsausprägungen beeinflussen das Antwortverhalten nicht) unwahrscheinlich.

Die Effektstärke ist mit $R^2m = 0,66$ (theoretisch) bzw. 0,58 (delta) recht hoch, deckt aber bei Weitem nicht so viel Varianz ab wie das Modell zum Antwortverhalten in der Frequenzstudie (siehe Abschnitt 5.2.1). Da *reiten* sich anders verhält als die anderen Testverben und die Testverben nicht als zufälliger Effekt einbezogen werden konnten, ist dies nicht verwunderlich. Der R^2c , der zusätzlich zufällige Effekte einbezieht, ist mit 0,71 (theoretisch) und 0,62 (delta) etwas höher.

Wie in der Präregistrierung vorgesehen, wird aufgrund der Konvergenzprobleme zusätzlich ein *random forest* für das Antwortverhalten gerechnet (Antwortverhalten \sim Transitivität). Für den *random forest* wurden 1.000 *conditional inference trees* gerechnet, die den Einfluss der Transitivität überprüfen. Um die Ergebnisse des linearen Modells und des *random forests* besser vergleichen zu können, wurden auch hier, anders als in der Präregistrierung angegeben, alle Testverben einbezogen.

Der *conditional inference tree* deutet auf einen klaren Einfluss der Transitivitätsausprägungen hin, wie Abbildung 5.15 zeigt.

Während die Wahrscheinlichkeit für *sein* für die Ausprägung TRANSITIV bei 0,2 liegt, ist sie für alle anderen Ausprägungen deutlich höher: Für AMBIG I liegt sie bei 0,8, für AMBIG II bei 0,9 und für INTRANSITIV bei 1. Die Ergebnisse decken sich also mit denen des generalisierten gemischten linearen Modells. Der C-Wert des *random forest* ist mit 0,89 hoch, es liegt somit eine gute Passung vor.

Als explorative statistische Auswertung der Daten wird ein generalisiertes gemischtes lineares Modell gerechnet, in dem Transitivitätsausprägung und Versuchsverb als feste Effekte vorgesehen sind: Antwortverhalten \sim Transitivität + Versuchsverb + (1|Proband_in). Auf diese Weise kann der Einfluss der einzelnen Versuchsverben auf das Antwortverhalten gemessen werden. Dies ist von Interesse, da *reiten* ein anderes Verhalten zu evozieren scheint als die anderen Verben. Das abweichende Antwortverhalten scheint dabei kein Zufall zu sein, sondern systematisch, da *reiten* bereits in der Prästudie andere Antworten evozierte als

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

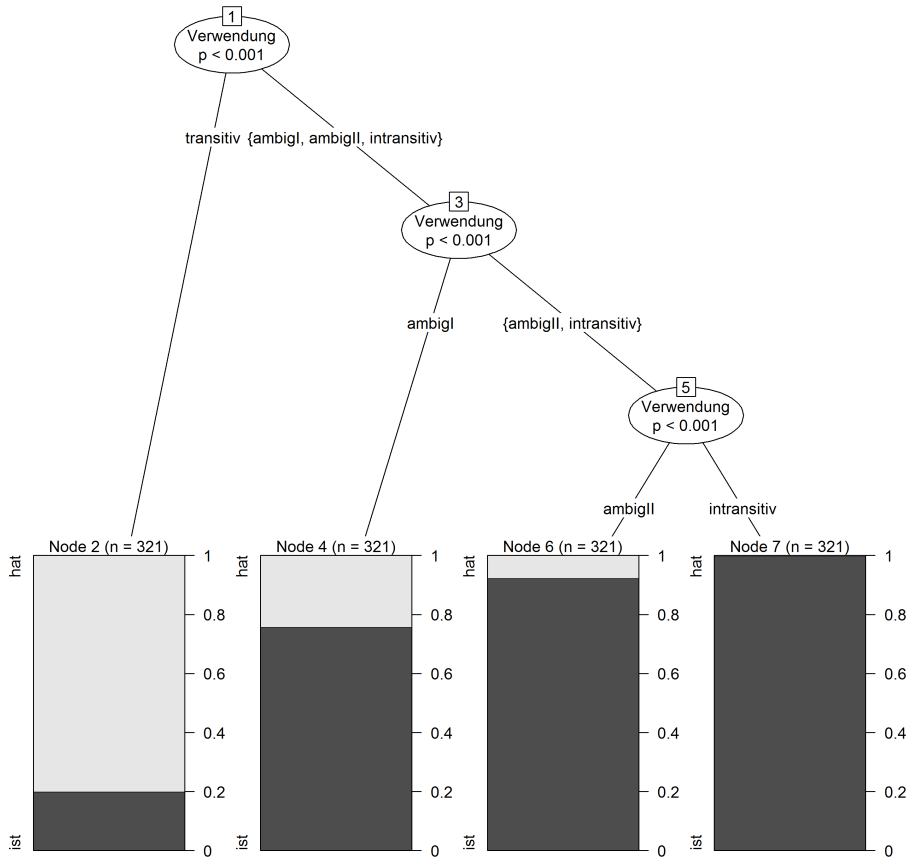


Abbildung 5.15: *Conditional inference tree* für die Wahl zwischen *haben* und *sein* in der Prototypizitätsstudie

fahren und *fliegen*. Proband_innen werden als zufällige Effekte mit *random intercepts* einbezogen, *random slopes* werden aufgrund von *singular fits* weggelassen. Die Standardabweichung für die Proband_innen um den y-Achsenabschnitt beträgt wie im Modell ohne Versuchsverb als festen Effekt 0,8. Die *log odds* für *sein*-Antworten finden sich in Tabelle 5.8. Abbildung 5.16 zeigt die auf den *log odds* basierenden Wahrscheinlichkeiten für *sein*-Antworten. Dabei stehen Kreise für *fahren*, Dreiecke für *fliegen* und Vierecke für *reiten*.

Tabelle 5.8: Werte des Modells für *sein*-Antworten in Abhängigkeit von Transitivität und Testverb

	Wert	SE	z
(Intercept, ambig I & fahren)	1,30	0,21	6,23
Transitivität: ambig II	1,44	0,26	5,62
Transitivität: intransitiv	4,79	1,01	4,72
Transitivität: transitiv	-2,87	0,23	-12,57
Item: fliegen	0,10	0,22	0,44
Item: reiten	-0,12	0,22	-0,55

Die Wahrscheinlichkeiten für *sein*- bzw. *haben*-Antworten haben sich hinsichtlich der Transitivitätsausprägungen nicht verändert. Auch hinsichtlich der Testverben unterscheiden sich die Wahrscheinlichkeiten kaum: Für alle Testverben ist eine klare Überlappung der 95 %-Konfidenzintervalle für die jeweiligen Ausprägungen zu sehen. Die Effektstärke des Modells hat sich durch die Berücksichtigung des Testverbs als festen Effekt nicht verändert.

Da *reiten* in Abhängigkeit von der Transitivitätsausprägung Unterschiede im Antwortverhalten hervorzurufen scheint, ist ein Blick auf die Interaktion zwischen Testverb und Transitivitätsausprägung interessant. Um die Interaktion näher zu betrachten, wird ein *random forest* mit Transitivitätsausprägung und Versuchsverb als Prädiktoren gerechnet (Antwortverhalten ~ Transitivität + Versuchsverb).²¹ Wieder wurden für den *random forest* 1.000 *conditional inference trees* gerechnet, die den Einfluss von Transitivität und Testitem überprüfen. Der *conditional inference tree* zeigt wie in der konfirmatorischen Analyse einen klaren Einfluss der Transitivitätsausprägungen, wie Abbildung 5.17 zeigt.

²¹Ein entsprechendes generalisiertes gemischtes lineares Modell mit Interaktion zwischen Transitivitätsausprägung und Versuchsverb und *random intercepts* für Teilnehmer_innen konvergierte auch mit verschiedenen Optimierungsmethoden nicht. Da die Optimierungsmethoden unterschiedliche Werte schätzen, wird das Modell nicht berücksichtigt. Der digitale Anhang enthält eine Dokumentation des Skripts zur Berechnung des Modells.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

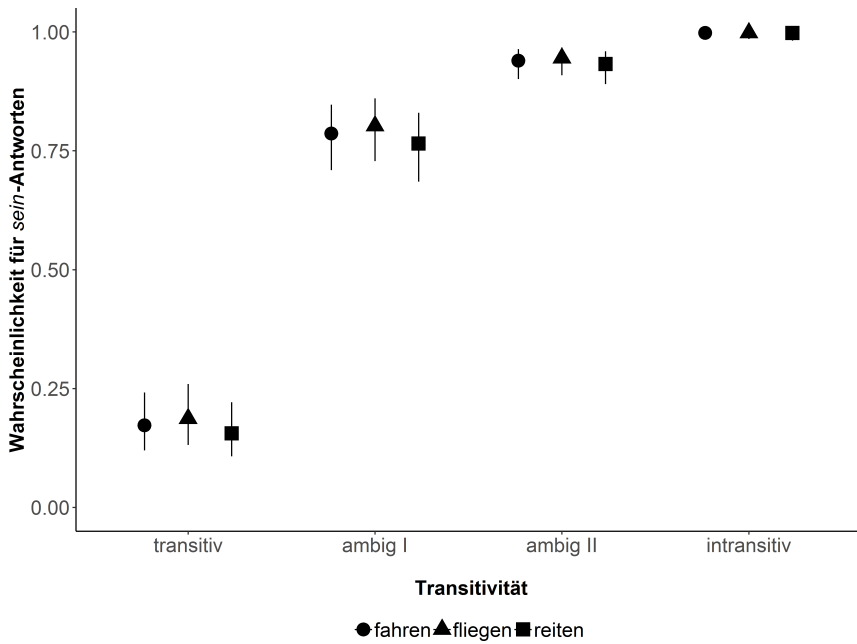


Abbildung 5.16: Wahrscheinlichkeit für *sein*-Antworten in Abhängigkeit von Transitivität und Testverb

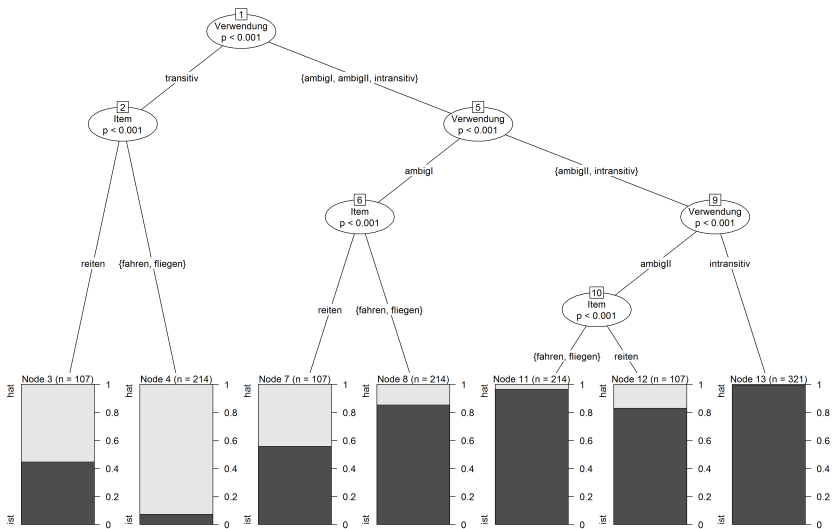


Abbildung 5.17: *Conditional inference tree* für die Wahl zwischen *haben* und *sein* in Abhängigkeit von Transitivität und Testverb

Zusätzlich zeigt sich der Einfluss der Testverben: *Reiten* verhält sich anders als *fahren* und *fliegen*. So liegt die Wahrscheinlichkeit für *sein*-Antworten bei der Ausprägung TRANSITIV für das Testverb *reiten* bei 0,45, während sie bei den anderen Versuchsverben nur bei 0,1 liegt. Auch bei AMBIG I und AMBIG II ist dies zu erkennen: Hier liegt die Wahrscheinlichkeit für *sein*-Antworten in Bezug auf *reiten* bei 0,55 bzw. 0,85, während sie in Bezug auf *fahren* und *fliegen* mit 0,85 bzw. 0,95 jeweils höher liegt. Nur bei der Ausprägung INTRANSITIV hat das Testverb keinen Einfluss. Für die Transitivitätsausprägung ergibt sich eine *conditional variance importance* von 0,26, für das Testverb von 0,003. Die Transitivitätsausprägung nimmt also mehr Einfluss auf das Modell als das Testverb, was sich auch darin spiegelt, dass *reiten* sich nur bei bestimmten Ausprägungen anders verhält als die anderen Testverben. Der C-Wert für das Modell liegt mit 0,93 höher als beim Modell ohne Testitem als gesonderten Einflussfaktor.

Insgesamt zeigt sich ein deutlicher Einfluss der Transitivität auf die Auxiliarauswahl: Transitive Sätze evozieren erwartbarerweise *haben*, intransitive *sein*. Auch bei den ambigen Sätzen herrschen *sein*-Antworten vor. Es ist jedoch ein Unterschied in den Antworten zwischen den ambigen Sätzen zu erkennen: AMBIG I evoziert etwas mehr *haben*-Antworten als AMBIG II. Der Einfluss der Prototypizität ist daher bei den Prototypen (transitive und intransitive Sätze) eindeutig festzustellen und in der Peripherie (ambige Sätze) zumindest zu erahnen. Der Einfluss der Transitivitätsausprägungen zeigt sich auch in der statistischen Analyse. Die Hypothesen aus Abschnitt 4.3.1 bestätigen sich somit insgesamt. Zudem fällt *reiten* durch ein von den anderen Testverben abweichendes Antwortverhalten auf: Dies wird vor allem in den häufigen *sein*-Antworten im transitiven Satz und den vergleichsweise häufigen *haben*-Antworten bei AMBIG I und II deutlich.

5.3.2 Reaktionszeiten

Die Reaktionszeiten beziehen sich auf die Wahl zwischen *hat* und *ist* in den Testätzen. Es zeigt sich ein deutlicher Unterschied zwischen intransitiven Sätzen und den anderen Ausprägungen. Diese evozieren höhere Reaktionszeiten als die intransitiven Sätze. Abbildung 5.18 zeigt einen Beeswarmplot der logarithmierten Reaktionszeiten für die einzelnen Transitivitätsausprägungen inklusive arithmetischem Mittel (M) und Standardabweichung (SD) als Fehlerbalken. Zudem wird der Standardfehler (SE) berichtet. Die Reaktionszeiten für *fahren* sind als Kreise dargestellt, die Reaktionszeiten für *fliegen* als Dreiecke und die für *reiten* als Vierecke.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

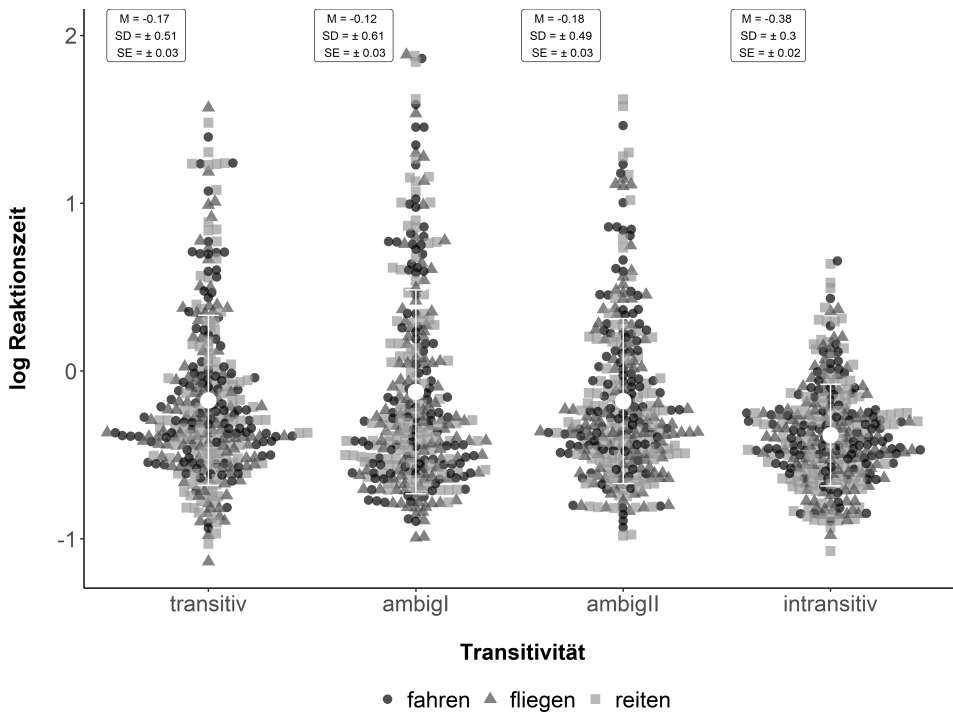


Abbildung 5.18: Reaktionszeiten in der Prototypizitätsstudie

Die Reaktionszeiten für die Ausprägung INTRANSITIV streuen weniger zu hohen Werten als die der anderen Transitivitätsausprägungen: Die intransitiven Sätze weisen maximal Reaktionszeiten von 0,66 auf, die anderen Ausprägungen von bis zu über 1,5. Dies zeigt sich auch im Mittelwert der Reaktionszeiten für die intransitive Sätze, der bei $-0,38$ mit einer Standardabweichung von 0,3 liegt. Die anderen Ausprägungen zeigen höhere Mittelwerte und Standardabweichungen: Bei AMBIG I und AMBIG II liegen die Mittelwerte bei $-0,12$ bzw. $-0,18$ und die Standardabweichungen bei 0,6 bzw. 0,4. Der Mittelwert für die Ausprägung TRANSITIV ist mit $-0,17$ und einer Standardabweichung von 0,5 vergleichbar mit dem Mittelwert für die Ausprägung AMBIG II. Insgesamt deuten die Beeswarmplots somit auf eine schnellere Verarbeitung der intransitiven Sätze im Vergleich zu den anderen Transitivitätsausprägungen hin.

Hinsichtlich der Testverben zeigt sich keine auffällige Verteilung in den Daten. Da *reiten* jedoch in der Prästudie andere Reaktionszeiten hervorrief als die anderen Testverben und *reiten* zudem ein anderes Antwortverhalten evoziert, wird in Abbildung 5.19 die Verteilung der Reaktionszeiten ohne das Testverb *reiten* betrachtet.

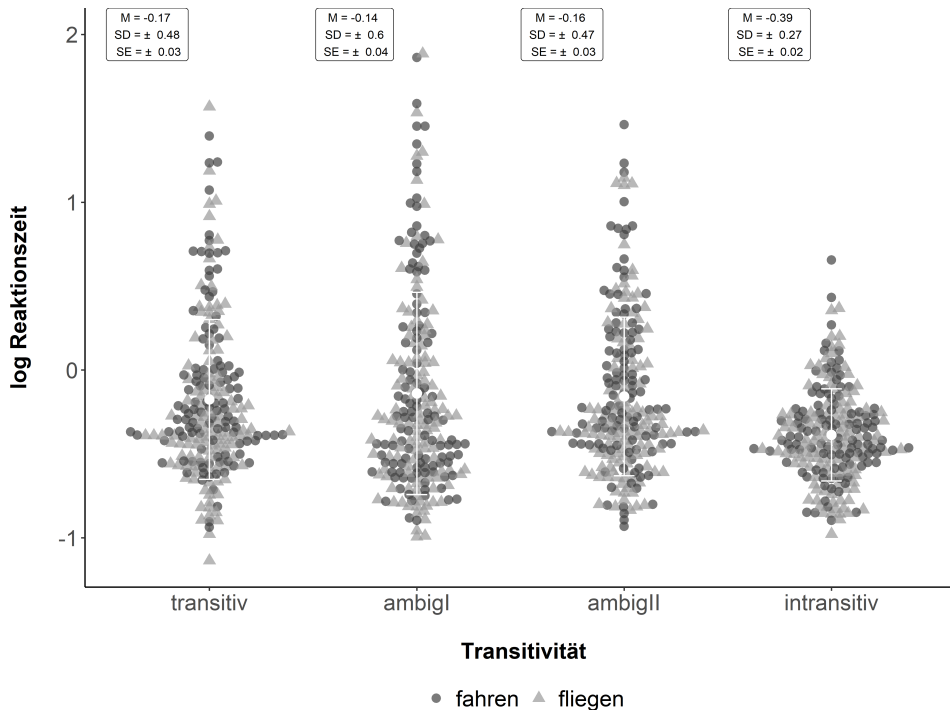


Abbildung 5.19: Reaktionszeiten in der Prototypizitätsstudie ohne *reiten*

Es sind vergleichbare Reaktionszeiten zu beobachten wie in den Beeswarmplots mit *reiten*, sodass ausgeschlossen werden kann, dass *reiten* einen gesonderten Einfluss auf die Reaktionszeiten nimmt. Ein Reihenfolge-Effekt konnte ebenfalls ausgeschlossen werden (siehe Abbildung I.4 in Anhang I.2). Da die Länge der kritischen Segmente bei *hat* und *ist* nicht variiert und beide Stimuli gleichzeitig zu sehen sind, wird der Einfluss der Länge nicht überprüft.

Um den Unterschied in den Reaktionszeiten zwischen den Transitivitätsausprägungen genauer in den Blick zu nehmen, werden Scatterplots betrachtet. Abbildung 5.20 vergleicht die Reaktionszeiten für die Ausprägung INTRANSITIV mit denen der anderen Ausprägungen. Ein Punkt stellt die Reaktionszeit einer Person für ein Testitem dar. Die Reaktionszeiten sind bei *fahren* als Kreise, bei *fliegen* als Dreiecke und bei *reiten* als Vierecke dargestellt. Die graue Linie ist als Orientierung gedacht, sie zeigt den Bereich an, in dem die jeweils miteinander verglichenen Ausprägungen die gleiche Reaktionszeit hervorrufen würden.

Die Ausprägungen TRANSITIV, AMBIG I und AMBIG II rufen höhere Reaktionszeiten hervor als die Ausprägung INTRANSITIV, was sich deutlich an den Reaktionszeiten oberhalb der grauen Linie zeigt. Nur bei den niedrigen Reaktionszeiten

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

(bis 0) wurden vergleichbare Zeiten gemessen. Ein Einfluss der Testverben auf die Reaktionszeiten ist wie in den Bienenwärmplots nicht zu erkennen.

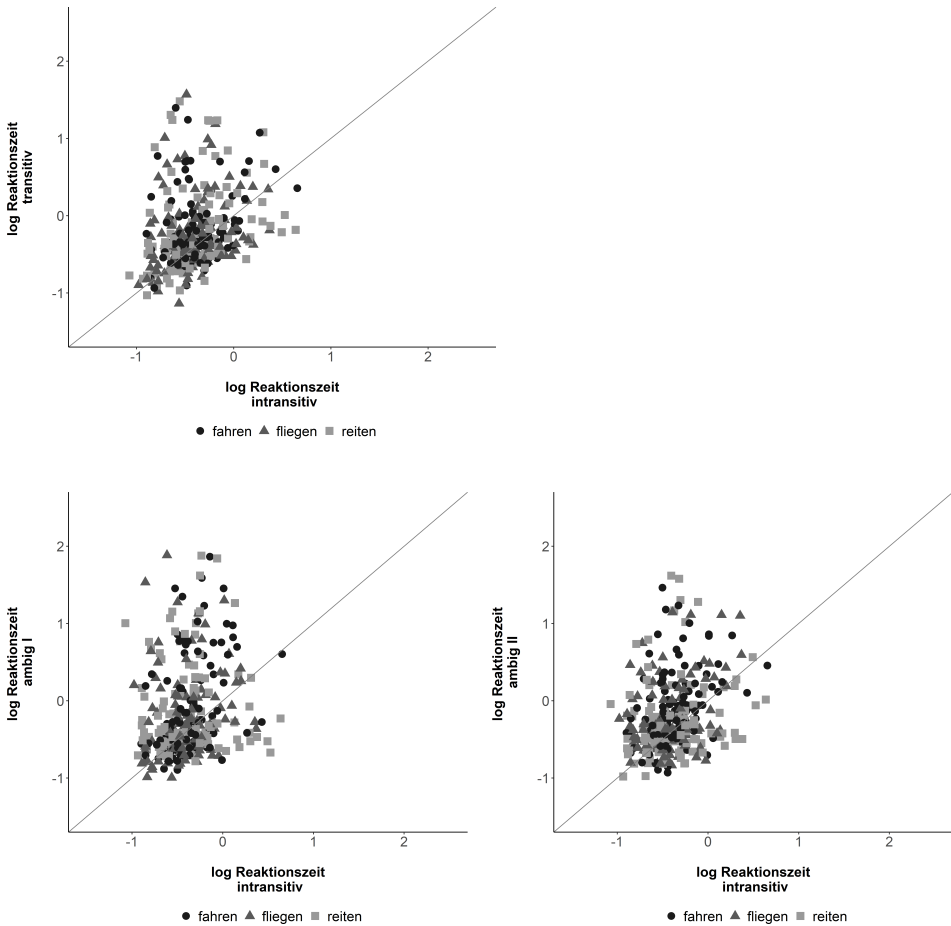


Abbildung 5.20: Scatterplots für INTRANSITIV in Kontrast zu den anderen Transitivitätsausprägungen

Abbildung 5.21 kontrastiert die Reaktionszeiten für die Ausprägung TRANSITIV mit denen der Ausprägungen AMBIG I und AMBIG II. Zudem werden die Reaktionszeiten der Ausprägungen AMBIG I und AMBIG II miteinander verglichen.

Im Vergleich zu der Ausprägung TRANSITIV streuen AMBIG I und AMBIG II stärker zu hohen Reaktionszeiten, was sich an den Punkten oberhalb der grauen Linie erkennen lässt. Anders als die Ausprägung INTRANSITIV weist TRANSITIV aber teilweise hohe Reaktionszeiten auf, wenn AMBIG I und AMBIG II geringe Reaktionszeiten zeigen, wie an den Punkten unterhalb der grauen Linie erkennbar

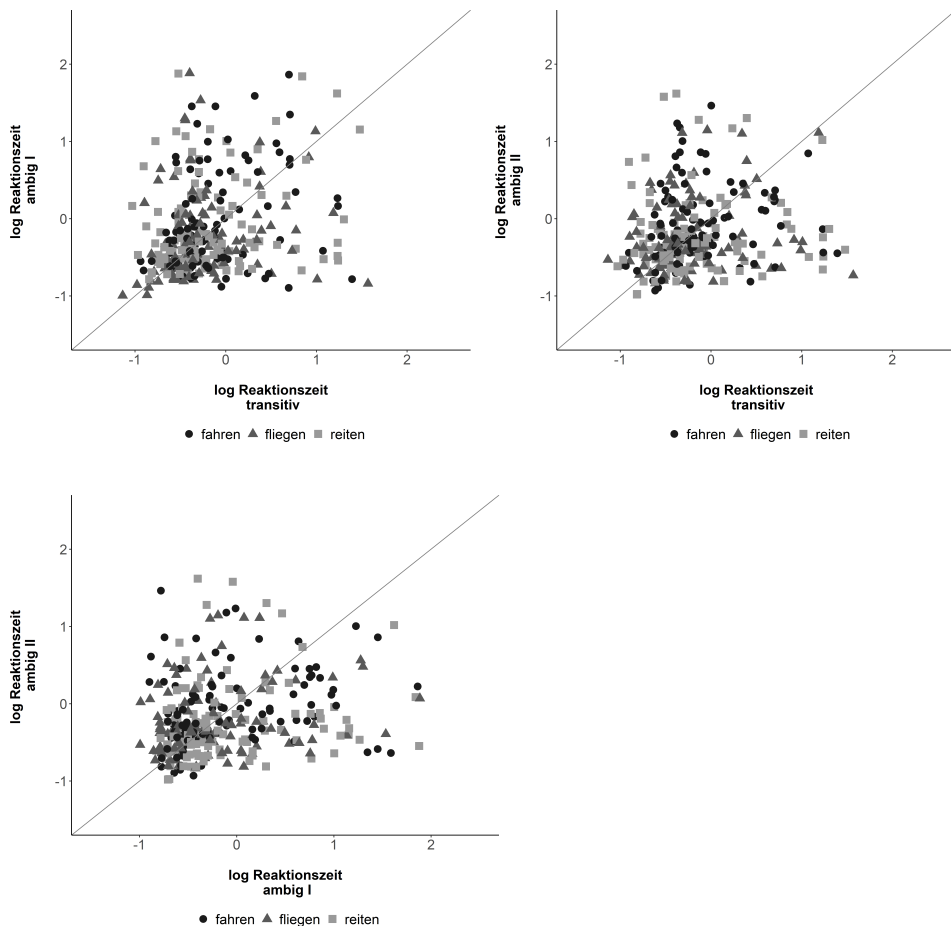


Abbildung 5.21: Scatterplots für TRANSITIV, AMBIG I und AMBIG II

ist. AMBIG I und AMBIG II evozieren vergleichbare Reaktionszeiten: Die Reaktionszeiten schwanken um die graue Linie. Es sind zudem sowohl hohe Reaktionszeiten für AMBIG I bei geringen Reaktionszeiten für AMBIG II als auch geringe Reaktionszeiten für AMBIG I bei hohen Reaktionszeiten für AMBIG II zu sehen. Hinsichtlich der Testverben ist wiederum kein Einfluss auf die Reaktionszeiten zu erkennen.

Die Reaktionszeiten deuten somit auf einen klaren Unterschied zwischen INTRANSITIV und den anderen Ausprägungen hin. Zudem scheint einen kleiner Unterschied zwischen der Ausprägung TRANSITIV und den Ausprägungen AMBIG I und AMBIG II zu bestehen, aber kein Unterschied zwischen AMBIG I und AMBIG II.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Für die konfirmatorische statistische Analyse wird ein gemischtes lineares Modell mit der Transitivitätsausprägung der Testsätze als festem Effekt genutzt: Reaktionszeit ~ Transitivität + (1|Proband_in) + (1|Item).²² Für Proband_in und Testverb werden jeweils *random intercepts* berücksichtigt, auf *random slopes* muss aufgrund von *singular fits* und Konvergierungsproblemen verzichtet werden.

Die Standardabweichung für die Versuchitems ist mit 0,04 sehr gering, für die Proband_innen ist sie erwartbarerweise mit 0,22 deutlich höher, da Reaktionszeiten individuell stark schwanken. Es bleibt eine Residuenvarianz von 0,44. Die vom Modell geschätzten Werte für die Reaktionszeiten sind in Tabelle 5.9 aufgeführt. Abbildung 5.22 zeigt die Reaktionszeiten, die das Modell vorhersagt.

Tabelle 5.9: Werte des Modells für die Reaktionszeiten in der Prototypizitätsstudie

	Wert	SE	<i>t</i>	df	<i>p</i>
(Intercept = ambig I)	-0,12	0,04	3,19	16,54	<0,01
Transitivität: ambig II	-0,06	0,03	-1,60	1174,70	0,11
Transitivität: intransitiv	-0,26	0,03	-7,50	1174,70	<0,01
Transitivität: transitiv	-0,05	0,03	-1,49	1174,70	0,14

Für TRANSITIV, AMBIG I und AMBIG II werden ähnliche Werte (-0,12 bis 0,17) vorhergesagt, deren 95 %-Konfidenzintervalle sich überlappen. Die vorhergesagten Reaktionszeiten für INTRANSITIV (-0,38) sind deutlich geringer und das 95 %-Konfidenzintervall überlappt nicht mit den 95 %-Konfidenzintervallen der Werte für die anderen Ausprägungen.

Nur für INTRANSITIV wird im Vergleich zum Referenzlevel AMBIG II das α -Level von 0,01 unterschritten. Die Ausprägung INTRANSITIV scheint daher systematisch geringere Reaktionszeiten hervorzurufen als die anderen Ausprägungen. Dies ist für die Ausprägung TRANSITIV nicht der Fall. Die Effektstärke des Modells ist mit einem R^2m von 0,039 sehr gering, mit den zufälligen Effekten ist sie mit einem R^2c von 0,24 deutlich höher.²³

²²Die Präregistrierung sieht vor, dass nur *fahren* und *fliegen* im Modell berücksichtigt werden. Da *reiten* aber anders als in der Prästudie mit den anderen Versuchsverben vergleichbare Reaktionszeiten hervorruft, wird auch *reiten* einbezogen.

²³Als explorative Analyse der Daten war in der Präregistrierung ein Modell geplant, das Transitivität und Versuchsverb als feste Effekte enthält, die interagieren. Dies war vorgesehen, weil *reiten* in der Prästudie andere Reaktionszeiten evozierte als die anderen Testverben. Da die visuelle Betrachtung der Daten keinen Hinweis auf eine Interaktion gibt, wird darauf verzichtet.

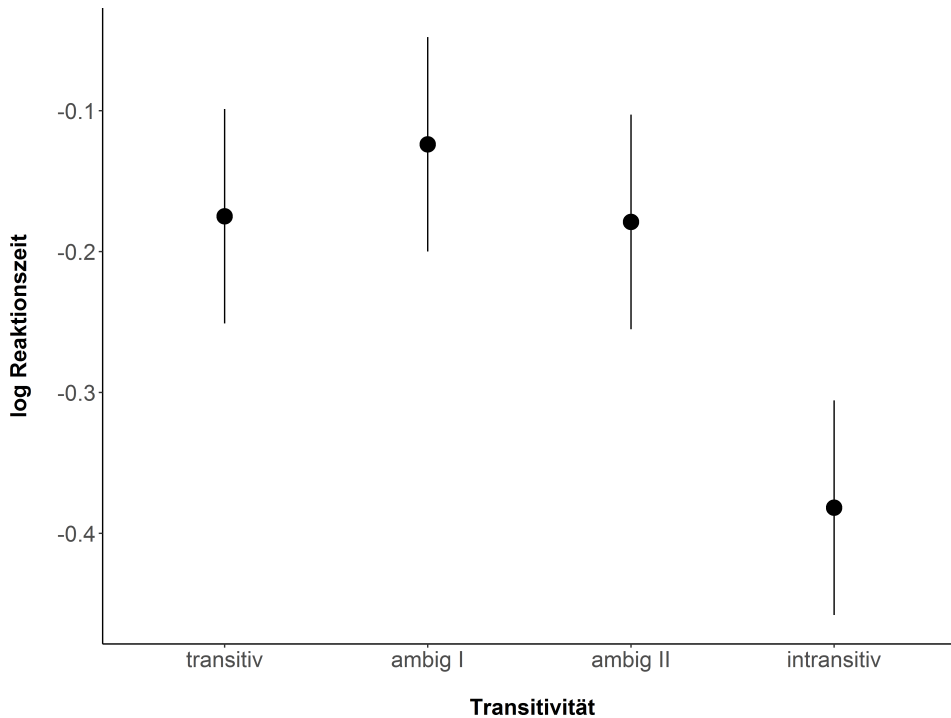


Abbildung 5.22: Vorhergesagte Reaktionszeiten in der Prototypizitätsstudie

Der geschätzte Wert für die Reaktionszeiten im transitiven Satz war in der Prästudie mit $-0,2$ deutlich niedriger als in der Hauptstudie ($-0,05$; siehe Tabelle 5.9). Da die Stichprobengrößenberechnung für eine Teststärke von $0,87$ auf diesem deutlich niedrigeren Wert beruhte, ist es möglich, dass die Teststärke zu gering war, um einen Unterschied zu entdecken. Allerdings ist der Schätzwert von $-0,05$ sehr klein, sodass ein Effekt zwar möglich, aber vernachlässigbar zu sein scheint. Ähnliches ist für AMBIG II festzuhalten: Auch hier liegt der Schätzwert aus der Prästudie mit $-0,2$ deutlich niedriger als in der Hauptstudie ($-0,06$ siehe Tabelle 5.9).

In den Hypothesen in Abschnitt 4.3.1 wurden geringere Reaktionszeiten für die Ausprägung INTRANSITIV erwartet als für die anderen Ausprägungen. Jedoch wurde auch erwartet, dass die Ausprägung TRANSITIV geringere Reaktionszeiten hervorruft als die ambigen Sätze, da für transitive Sätze nur ein Auxiliar (*haben*) mental stark gefestigt ist, während bei den ambigen Sätzen theoretisch beide Auxiliare möglich sind und in Korpora vorkommen. Aufgrund der Ergebnisse

der Prästudie wurde zudem die Möglichkeit in Betracht gezogen, dass die Ausprägung AMBIG II vergleichbare Reaktionszeiten hervorruft wie die Ausprägung TRANSITIV und beide geringere Reaktionszeiten evozieren als AMBIG I. Die Daten der Hauptstudie deuten hingegen darauf hin, dass die transitiven Sätze genauso schnell verarbeitet werden wie die ambigen Sätze und dass zwischen den ambigen Sätzen kein Unterschied in den Reaktionszeiten besteht.

Es stellt sich die Frage, warum die Ausprägung TRANSITIV höhere Reaktionszeiten evoziert als die Ausprägung INTRANSITIV. Eine mögliche Erklärung für die erhöhten Reaktionszeiten ist, dass die transitiven Sätze aufgrund des Objekts komplexer sind als die intransitiven Sätze. Diese Erklärung greift auch für die ambigen Sätze, die ebenfalls ein Objekt bzw. ein Adverbial aufweisen. Um zu überprüfen, ob die Reaktionszeitunterschiede allein durch die unterschiedliche Komplexität von intransitiven und transitiven Sätzen zu erklären sind, werden die Reaktionszeiten für die Auxiliarwahl in zwei transitiven Füllersätzen²⁴ mit den bereits analysierten Reaktionszeiten der Ausprägungen INTRANSITIV und TRANSITIV verglichen. Abbildung 5.23 zeigt BeeswarmpLOTS für die intransitiven Sätze, die transitiven Füllersätze und die transitiven Testsätze.

Die Reaktionszeiten für die transitiven Füllersätze streuen weniger zu hohen Reaktionszeiten als die Reaktionszeiten für die transitiven Testsätze, was sich auch in der geringeren Standardabweichung zeigt. Zudem ähneln die Reaktionszeiten für die Füllersätze in ihrer Verteilung den Reaktionszeiten für die intransitiven Sätze, auch wenn die Verteilung etwas höhere Reaktionszeiten zeigt. Dies wird auch durch die Mittelwerte deutlich: Die Reaktionszeiten der transitiven Füllersätze haben einen Mittelwert von $-0,28$ mit einer Standardabweichung von $0,29$. Der Mittelwert liegt damit zwischen dem Mittelwert der intransitiven Sätze ($-0,38$, Standardabweichung = $0,3$) und dem Mittelwert der transitiven Testsätze ($-0,18$, Standardabweichung = $0,51$). Das Komplexitätsargument scheint somit nur eingeschränkt zu greifen: Zwar sind die Reaktionszeiten für die transitiven Füllersätze höher als für die intransitiven Sätze, jedoch müssten sie ähnlich hoch sein wie die Reaktionszeiten für die transitiven Testsätze, wenn die Unterschiede in den Reaktionszeiten allein auf die Komplexität der Sätze zurückzuführen wären.

Die Daten legen daher eine andere Erklärung nahe: Wie Gillmann (2016: 267–268) in ihrer Korpusstudie zeigt, werden Bewegungsverben vornehmlich in intransitiven Sätzen genutzt. Daher ist die *sein*-Selektion bei Bewegungsverben

²⁴Die Füllersätze lauten *Ich weiß, dass der Vater die Kinder zur Schule gebracht hat* und *Er fährt nochmal zurück, weil er die Jacke seines Freundes aus Versehen eingesteckt hat und dieser sie auf der Arbeit benötigt*. Beide wurden vorab von Proband_innen als semantisch plausibel bewertet (4,4 und 4 auf einer Skala von 1 (nicht plausibel) bis 5 (plausibel), siehe Tabelle F.1.

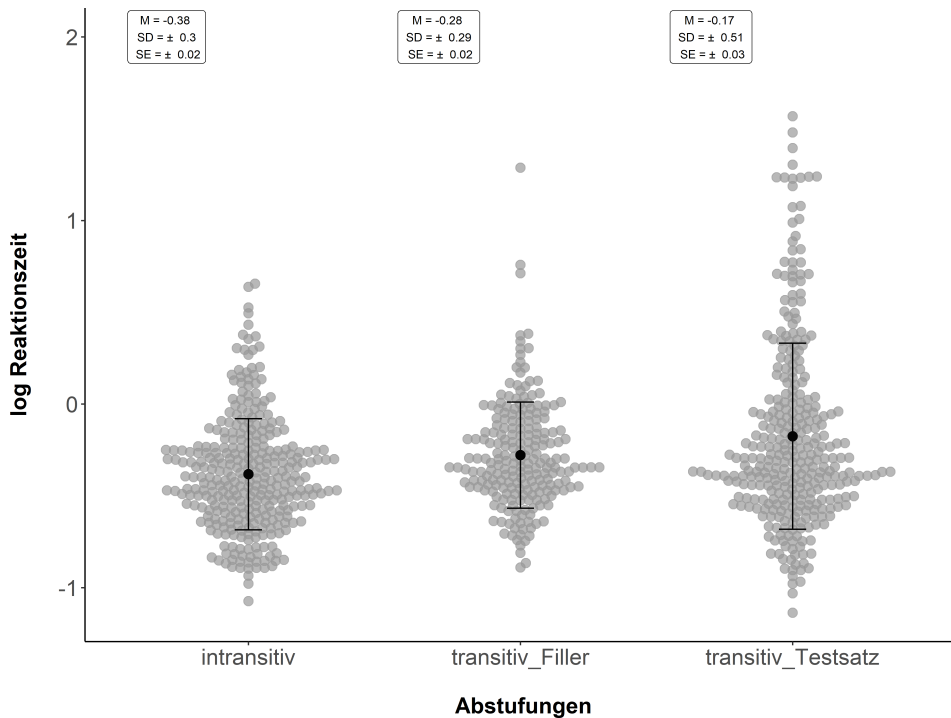


Abbildung 5.23: Reaktionszeiten für intransitive Sätze sowie transitive Filler- und Testsätze

viel häufiger als die *haben*-Selektion, weshalb Bewegungsverben mit *sein* assoziiert sind. In transitiven Sätzen muss entgegen dieser Assoziation *haben* gewählt werden. Der Wechsel von *sein* zu *haben* ist daher mit kognitivem Aufwand verbunden und könnte zu erhöhten Reaktionszeiten in den transitiven Testsätzen geführt haben. Diese Erklärung greift nicht für die ambigen Sätze, da hier vorrangig *sein* gewählt wurde. Aufgrund der Ambiguität sind hier beide Auxiliare ähnlich wahrscheinlich. Dies spiegelt sich auch ansatzweise im Antwortverhalten wider, sodass sich die erhöhten Reaktionszeiten hier mit der Wahlmöglichkeit erklären lassen.

Insgesamt zeigt sich, dass die intransitiven Sätze niedrigere Reaktionszeiten hervorrufen als die anderen Transitivitätsausprägungen. Für diese lassen sich keine Unterschiede in den Reaktionszeiten feststellen. Abweichend von diesem Befund waren für transitive Sätze basierend auf der Prästudie leicht erhöhte Reaktionszeiten im Vergleich zu intransitiven Sätzen erwartet worden, die vergleichbar mit den Reaktionszeiten der Ausprägung AMBIG II sind, aber unter den

Reaktionszeiten der Ausprägung AMBIG I liegen (siehe Abschnitt 4.3.1 für nähere Erläuterungen). Ein visueller Vergleich der Reaktionszeiten für transitive Filler- und Testsätze deutet darauf hin, dass die erhöhten Reaktionszeiten bei den transitiven Sätzen nicht allein auf die erhöhte Komplexität der transitiven Sätze zurückzuführen sind. Dies müsste in einer weiteren Studie statistisch geprüft werden. Der Schluss liegt nahe, dass Bewegungsverben generell mit *sein* assoziiert sind und die Selektion von *haben* in transitiven Sätzen daher einen kognitiven Aufwand darstellt, der sich in höheren Reaktionszeiten spiegelt. Somit lässt sich zumindest für einen der Prototypen (intransitive Sätze) ein schnelleres Antwortverhalten beobachten, für die Peripherie (ambige Sätze) hingegen ein langsameres. Die erhöhten Reaktionszeiten beim zweiten Prototyp (transitive Sätze) lassen sich durch das Wirken des anderen Prototyps erklären. Im folgenden Abschnitt werden die Ergebnisse der Studien zu Form-Schematizitätseffekten vorgestellt.

5.4 Form-Schematizität

In diesem Abschnitt werden das Antwortverhalten und die Reaktionszeiten in den Studien zu Form-Schematizitätseffekten vorgestellt. Diese überprüfen anhand des Form-Schemas schwacher Maskulina den Einfluss von Form-Schematizität auf Variation. Alle Daten beziehen sich auf die in Abschnitt 4.4.6 und 4.5.4 vorgestellten Proband_innen. Die Stichprobengröße beträgt 54 Personen für die *self-paced-reading*-Studie, 56 für die *lexical-decision*-Studie und 132 für die *sentence-maze*-Studie.

5.4.1 Antwortverhalten

Das Antwortverhalten wurde innerhalb der Studien auf verschiedene Weisen evoziert: Im Rahmen der *self-paced-reading*-Studie wurde ein Produktionsexperiment durchgeführt, in dem die Proband_innen Pseudosubstantive (*Schettose*, *Knatt*, *Grettel*, *Truntake*) in den Genitiv bzw. bei *Truntake* in den Dativ setzten (siehe Abschnitt 4.4 für weitere Erläuterungen).²⁵ In der *lexical-decision*-Studie wurden jeweils starke und schwache Formen von real existierenden Substantiven nach ihrer Bekanntheit bewertet und in der *sentence-maze*-Studie wur-

²⁵Die Antworten der Proband_innen auf die Frage, ob ihnen innerhalb der *self-paced reading task* Varianten aufgefallen sind, lassen nicht darauf schließen, dass die vorab gelesenen Varianten das Antwortverhalten im Produktionsexperiment beeinflusst haben. Die Antworten sind in den Rohdaten enthalten und im digitalen Anhang dokumentiert. Dennoch ist ein Einfluss der vorab durchgeführten *self-paced-reading task* auf das Antwortverhalten nicht komplett auszuschließen.

de zwischen starken und schwachen Formen von real existierenden Substantiven gewählt (siehe Abschnitt 4.5 für weitere Erläuterungen). Die Testsubstantive entsprechen jeweils dem Prototyp des Form-Schemas schwacher Maskulina (z. B. *Kollege, Nefte, Schettose*), der Peripherie des Form-Schemas (z. B. *Graf, Held, Knatt*) oder zählen zur starken Flexion (z. B. *Vogt, Freund, Grettel*) (siehe Abschnitt 3.2 für Erläuterungen zur Variation in der Deklination von Maskulina). Als starke Genitivform wurde in der lexical-decision- und der sentence-maze-Studie meist die kurze Endung genutzt (*des Grafs*), nur bei den Substantiven *Dieb, Freund* und *Feind* wurde die lange Endung präsentiert (*des Diebes*), da diese in den Wortverlaufskurven des DWDS (Berlin-Brandenburgische Akademie der Wissenschaften 2019) häufiger ist als die kurze Endung.

Für Substantive, die dem Prototyp des Form-Schemas schwacher Maskulina komplett entsprechen, wird erwartet, dass schwache Formen im Produktionsexperiment genutzt werden, in der lexical-decision-Studie als bekannt bewertet werden und in der sentence-maze-Studie gegenüber den starken Formen bevorzugt werden. Starke Formen sollten hingegen nur selten produziert, als bekannt bewertet und gegenüber schwachen Formen bevorzugt werden. Für starke Maskulina wird das gegenteilige Antwortverhalten und damit eine Tendenz dazu erwartet, starke Formen zu nutzen, als bekannt zu bewerten und gegenüber schwachen Formen zu bevorzugen. Hinsichtlich der Testitems in der Peripherie des Form-Schemas wird ein schwankendes Antwortverhalten antizipiert.

5.4.1.1 Produktionsexperiment in der self-paced-reading-Studie

In diesem Abschnitt werden die Ergebnisse des Produktionsexperiments vorgestellt, das Teil der self-paced-reading-Studie war. Abbildung 5.24 zeigt für die Ausprägungen FORM-SCHEMA (*Schettose/Truntake*), PERIPHERIE (*Knatt*) und STARK (*Grettel*) den Anteil schwacher Formen *-(e)n* im Vergleich zu anderen gewählten Formen *-(e)s*, Nullendungen, doppelte Formen). Bis auf *Truntake* sollten alle Substantive in den Genitiv gesetzt werden, bei *Truntake* wurde der Dativ elizitiert. Ein Rechteck entspricht jeweils einer Antwort. Dunkelgraue Rechtecke stehen für schwache Formen, hellgraue für andere Flexionsformen. Insgesamt wurden 54 Antworten gegeben.

Für *Schettose* und *Truntake* überwiegen erwartungsgemäß die schwachen Formen mit 79,6 % bzw. 90,7 % deutlich. Bei *Grettel* ist das Gegenteil der Fall: Hier entfallen nur 11,1 % der Antworten auf die schwache Form. *Knatt* zeigt Variation in den Antworten: Mit 55,6 % entfallen nur gut die Hälfte der Antworten auf die schwache Form. Die Ergebnisse decken sich mit den Ergebnissen der Produktionsexperimente von Köpcke (2000b) und Schmitt (2019a) (siehe Abschnitt 3.2.3) und bestätigen daher das Wirken des Form-Schemas in der Sprachproduktion.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

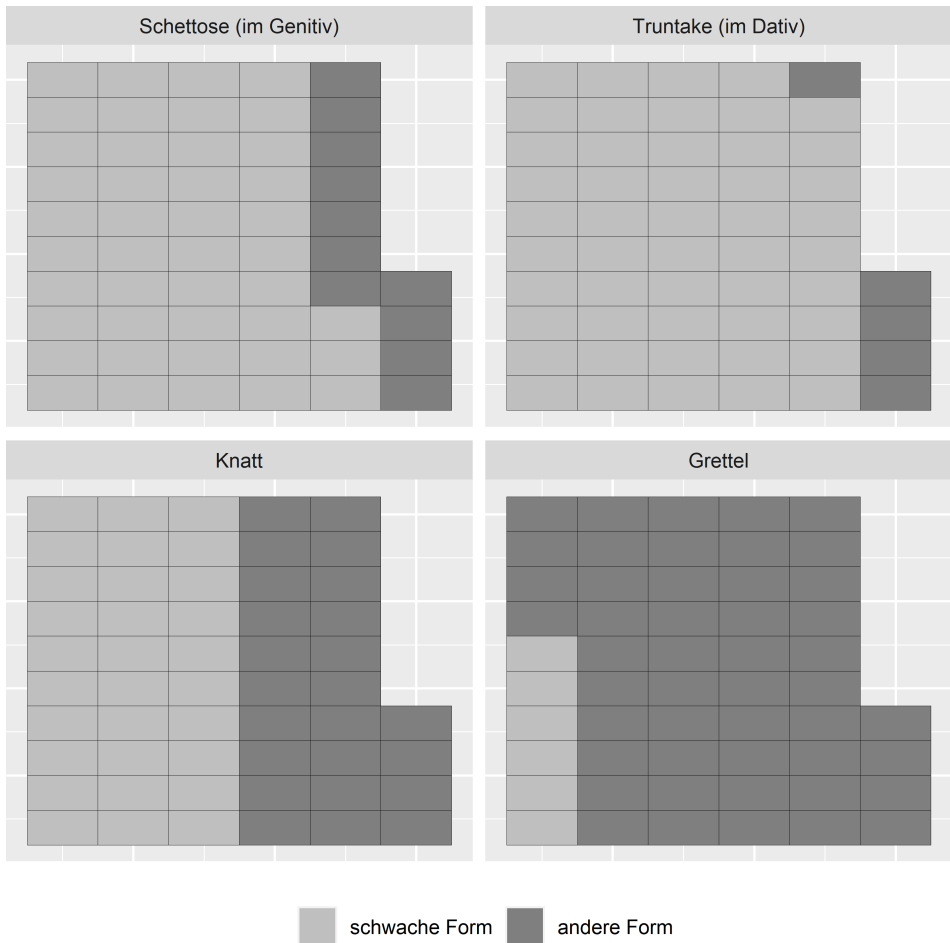


Abbildung 5.24: Antwortverhalten im Produktionsexperiment

Um die Formen, die in Abbildung 5.24 unter ANDERE zusammengefasst wurden, näher in den Blick zu nehmen, wird die Verteilung aller Formen für die einzelnen Testitems betrachtet. Tabelle 5.10 zeigt das Antwortverhalten für *Schettose* und *Truntake* gemeinsam, da beide Substantive zum Prototyp des Form-Schemas schwacher Maskulina gehören. *Schettose* wurde im Genitiv abgefragt, *Truntake* im Dativ. Schrägstriche werden in der Tabelle genutzt, wenn eine Form für ein Substantiv aufgrund des Kasusunterschieds nicht möglich ist. In der Tabelle ist die Form auf *-n* hervorgehoben, da diese die zu erwartenden Deklinationsform darstellt.

Tabelle 5.10: Antwortverhalten bei den Testsubstantiven *Schettose* und *Truntake*

Endung	<i>Schettose</i> (im Genitiv)	<i>Truntake</i> (im Dativ)
-n	43 (79,6 %)	49 (90,7 %)
-s	5 (9,3 %)	/
ohne	4 (7,4 %)	4 (7,4 %)
-ns	1 (1,9 %)	/
-n/-s	1 (1,9 %)	/
-n/ohne	0 (0 %)	1 (1,9 %)
gesamt	54 (100 %)	54 (100 %)

Die wenigen Antworten, die nicht auf *-n* entfallen, machen bei *Schettose* Genitivformen auf *-s* (9,3 %) und Nullendungen (7,4 %) aus. Zudem wählte eine Person bei *Schettose* *-ns* und eine weitere sowohl *-n* als auch *-s*. Obwohl Nullendungen mit 7,4 % bei *Schettose* nur sehr selten gewählt wurden, ist ein näherer Blick auf sie interessant: Nullendungen sind auch in den Produktionsexperimenten von Köpcke (2000b) und Schmitt (2019a) zu beobachten, die mit denselben Pseudosubstantiven arbeiten. Köpcke (2000b: 161) berichtet bei einer Stichprobengröße von 31 Proband_innen nur für 3 % der Testsubstantive, die dem Prototyp des Form-Schemas entsprechen, Nullendungen. Schmitt (2019a: 167) nennt mit 5,5 % einen ähnlich hohen Anteil an Nullendungen für *Schettose* wie in der vorliegenden Studie. Auch die Stichprobengrößen der Studien sind mit 55 und 54 Proband_innen ähnlich groß.

Nullendungen sind zu erwarten, da die Testsubstantive Pseudowörter darstellen und daher den Proband_innen unbekannt sind (Schmitt 2019a: 167). Deshalb liegt es nahe, dass das Prinzip der Morphemkonstanz greift: Seltene (und damit

auch unbekannte) Wörter sollten aufgrund ihrer geringen Tokenfrequenz möglichst wenig verändert werden, um sicherzustellen, dass sie wiedererkannt werden (Ackermann & Zimmer 2017: 150). Dies kann am besten erreicht werden, indem die Form der Wörter überhaupt nicht verändert wird. Das Morphemkonstanzprinzip lässt sich bei Wörtern in der Peripherie der deutschen Substantive beobachten, bspw. bei Eigennamen, Kurzwörtern und tokeninfrequenten nicht-nativen Wörtern (Ackermann & Zimmer 2017: 151–154). Da die Pseudowörter in Texten eingeflochten waren, die Lexikoneinträgen ähneln, ist es möglich, dass die Proband_innen die Pseudowörter als real existierende Wörter angesehen und sie aufgrund der geringen Bekanntheit wie periphere Substantive wahrgenommen haben.

Bei *Truntake* entfallen die wenigen von *-n* abweichenden Antworten auf die starke und somit endungslose Dativform (7,4 %) und eine Doppelnennung der starken endungslosen und der schwachen Form auf *-n*.

Auch für *Knatt* ist ein Blick auf die genauere Aufschlüsselung der gewählten Formen interessant, die Tabelle 5.11 zeigt. Für *Knatt* sind sowohl *-n* als auch *-s* erwartbar, weswegen beide Formen graphisch hervorgehoben sind.

Tabelle 5.11: Antwortverhalten bei dem Testsubstantiv *Knatt*

Endung	Häufigkeit
-en	30 (55,6 %)
-(e)s	16 (29,6 %)
-en/-(e)s	5 (9,3 %)
-en/ens	1 (1,9 %)
-ers	1 (1,9 %)
ohne	1 (1,9 %)
gesamt	54 (100 %)

Wie zu erwarten ist, entfallen bei *Knatt* die meisten Formen, die nicht auf *-n* enden, auf *-(e)s* (29,6 %).²⁶ Zudem sind bei *Knatt* mit 9,3 % Doppelnennungen von *-en* und *-(e)s* vergleichsweise häufig: Bei *Schettose* und *Truntake* wurden jeweils nur einmal Formen doppelt genannt. Zudem wird bei *Knatt* einmal *-en* und *-ens* doppelt genannt. Die Nullendung kommt bei *Knatt* einmal vor, genauso wie die Endung auf *-ers*. Auch diese Endung hat Schmitt (2019a: 167–168) bereits beobachtet. Sie kann als eine Anpassung von *Knatt* an das Form-Schema starker Maskulina gewertet werden: Formen auf *-er* werden immer stark dekliniert. Indem

²⁶Von den 16 Formen auf *-(e)s* sind fünf lang (*-es*) und elf kurz (*-s*).

-er an *Knatt* angehängt wird, wird das Variationspotential von *Knatt* aufgelöst (Schmitt 2019a: 168).

Grettel weist erwartungsgemäß den größten Anteil starker Formen auf. Diese entfallen zu 58,8 % auf -s, wie Tabelle 5.12 zeigt. -s ist in der Tabelle hervorgehoben, da -s die zu erwartende Form darstellt.

Tabelle 5.12: Antwortverhalten bei dem Testsubstantiv *Grettel*

Endung	Häufigkeit
-s	32 (59,3 %)
ohne	9 (16,7 %)
-s/ohne	7 (13 %)
-n	6 (11,1 %)
gesamt	54 (100 %)

Die Form auf -n ist bei *Grettel* auch in dieser detaillierten Betrachtung mit 11,1 % am seltensten. Der Anteil an Nullendungen ist dagegen mit 16,7 % im Vergleich zu den anderen Testsubstantiven am höchsten. Dies ist auch in der Studie von Schmitt (2019a: 167) der Fall, bei der die Nullendungen 32,7 % der gewählten Formen für *Grettel* ausmachen. Ähnlich häufig wie Nullendungen sind in der vorliegenden Studie Doppelnennungen von -s und Nullendung (13 %). In der Studie von Schmitt (2019a: 167) kommen für *Grettel* hingegen gar keine Doppelnennungen dieser Art vor. Der hohe Anteil an Nullendungen bei *Grettel* im Vergleich zu *Schettose* und *Knatt* könnte durch das Form-Schema schwacher Maskulina bedingt sein: Bei *Schettose* kann der naheliegende Anschluss an das Form-Schema die Möglichkeit der Nullendung ausstechen, bei *Knatt* könnte dies trotz der Peripherie des Form-Schemas ebenfalls der Fall sein. Dies ist für *Grettel* ausgeschlossen, daher zeigt *Grettel* eine klare Tendenz zu starken Formen (-s oder Nullendung).

Für die konfirmatorische statistische Analyse des Antwortverhaltens wird ein generalisiertes lineares Modell mit Form-Schematizitätsausprägungen als festem Effekt genutzt: Antwortverhalten ~ Form-Schematizität. Dabei wird das Antwortverhalten binär betrachtet: Den schwachen Formen werden nicht-schwache Formen (-e)s, -ers, Nullendung, doppelt genannte Formen) gegenübergestellt. Die Form-Schematizitätsausprägungen werden durch die Testitems (*Schettose*, *Truntake*, *Knatt* und *Grettel*) repräsentiert. Prototypische Vertreter des Form-Schemas sind mit *Schettose* und *Truntake* zweimal vertreten, jedoch einmal im Genitiv und

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

einmal im Dativ, weshalb sie als zwei Ausprägungen betrachtet werden. Aufgrund von *singular fits* können Proband_innen nicht als zufälliger Effekt berücksichtigt werden. Ursprünglich waren *random intercepts* für Proband_innen vorgesehen, *random slopes* sind nicht möglich, da pro Form-Schematizitätsausprägung nur ein Testitem genutzt wird. Die *singular fits* sind vermutlich der singulären Itemstruktur geschuldet: Hierdurch liegen zu wenig Beobachtungen vor, um die zufällige Effektstruktur zu modellieren. Daher ist die Annahme der unabhängigen Datenpunkte verletzt, weshalb zusätzlich zu diesem Modell ein *random forest* gerechnet wird. Die *log odds* für die Produktion schwacher Formen sind in Tabelle 5.13 gelistet. Abbildung 5.25 zeigt die auf den *log odds* basierende Wahrscheinlichkeit für die Wahl schwacher Formen im Produktionsexperiment.

Tabelle 5.13: Werte des Modells für die Wahl schwacher Formen im Produktionsexperiment

	Wert	SE	<i>z</i>	<i>p</i>
(Intercept = Schettose)	1,36	0,34	4,03	<0,01
Item: Grettel	-3,44	0,55	-6,27	<0,01
Item: Knatt	-1,14	0,43	-2,62	<0,01
Item: Truntake	0,92	0,58	1,59	0,11

Für *Schettose* und *Truntake* werden Werte von 0,8 bzw. 0,9 angesetzt, d. h. schwache Antworten sind sehr wahrscheinlich. Das Gegenteil ist für *Grettel* der Fall, das einen Wert von 0,1 aufweist. Für *Knatt* sind mit 0,56 beide Formen gleich wahrscheinlich. Die 95 %-Konfidenzintervalle sind jeweils relativ lang, aber überlappen sich bis auf die Intervalle von *Schettose* und *Truntake* nicht. Die Überlappung bei *Schettose* und *Truntake* ist erwartbar, da beide Substantive dem Prototyp des Form-Schemas schwacher Maskulina entsprechen. Abseits von *Schettose* und *Truntake* ist daher von systematischen Effekten im Antwortverhalten auszugehen, die durch die Form-Schematizitätsausprägungen ausgelöst werden. Dies bestätigt auch ein Blick auf die *p*-Werte, die für alle Ausprägungen bis auf *Truntake* im Vergleich zum Referenzlevel *Schettose* unter dem α -Level von 0,01 liegen. Die Effektstärke weist mit einem R^2m von 0,45 (theoretisch) und 0,39 (delta) auf einen mittleren Effekt hin. Sie ist jedoch im Vergleich zu den Modellen zum Antwortverhalten in den Studien zu Frequenz und Prototypizität relativ gering. Dies könnte daran liegen, dass das Antwortverhalten durch die Aufteilung in schwache und andere Formen stark vereinfacht wurde.

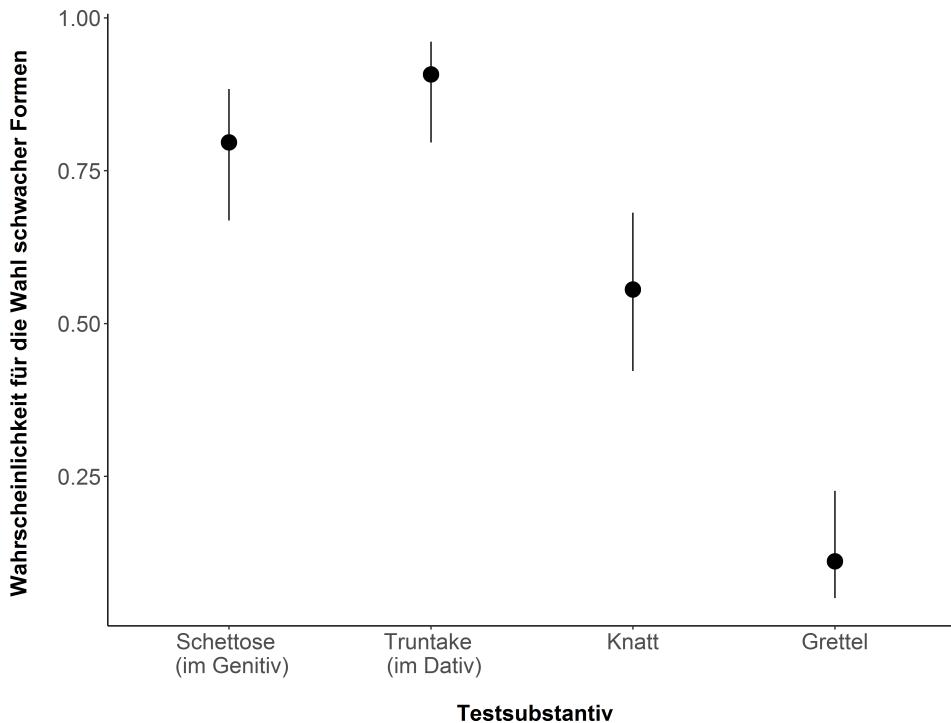


Abbildung 5.25: Wahrscheinlichkeit für die Wahl schwacher Formen im Produktionsexperiment

Zusätzlich zum generalisierten linearen Modell wird ein *random forest* mit Form-Schematizität als Einflussfaktor gerechnet: Antwortverhalten \sim Form-Schematizität. Für den *random forest* wurden 1.000 *conditional inference trees* gerechnet, die den Einfluss der Form-Schematizität überprüfen. Abbildung 5.26 zeigt den *conditional inference tree* für die Wahl schwacher und anderer Formen.

Der *conditional inference tree* bestätigt die Ergebnisse des generalisierten linearen Modells: Für *Truntake* und *Schettose* liegt die Wahrscheinlichkeit für schwache Formen bei 0,85, für *Grettel* hingegen lediglich bei 0,1. Für *Knatt* sind die schwache Form und andere Formen etwa gleich wahrscheinlich. Die Passung des *random forests* auf die Daten ist mit einem C-Wert von 0,84 gut.

Insgesamt zeigt sich das erwartete Antwortverhalten (siehe Abschnitt 4.4.1): *Truntake* und *Schettose* weisen vornehmlich schwache Formen auf, *Grettel* vornehmlich starke. Das Testsubstantiv *Knatt* schwankt zwischen stark und schwach und weist auch die meisten Doppelnennungen zwischen starker und schwacher Form auf. Die Studie repliziert damit die Ergebnisse von Köpcke

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

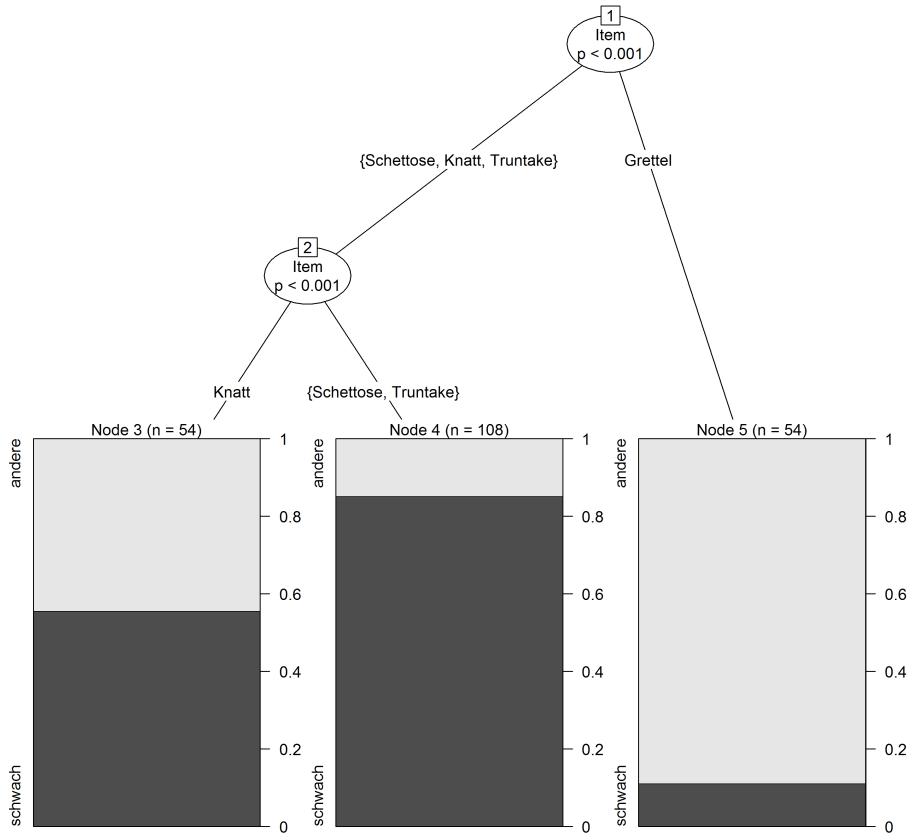


Abbildung 5.26: *Conditional inference tree* für die Wahl schwacher und anderer Formen im Produktionsexperiment

(2000b) und Schmitt (2019a). Interessant sind die Nullendungen, die sich vor allem bei *Grettel* zeigen. Diese geben einen Hinweis auf das Wirken des Morphemkonstanzprinzips bei unbekanntem Substantiven. Bei *Knatt* ist zudem einmal zu beobachten, dass das Testsubstantiv durch die Endung *-er* dem Form-Schema starker Maskulina (Substantive auf *-er* und *-el*) angepasst wird (Schmitt 2019a: 167–168).

5.4.1.2 Lexical-decision-Studie

In der lexical-decision-Studie bewerteten Proband_innen schwache und starke Flexionsformen der Testsubstantive nach ihrer Bekanntheit. Die Testsubstantive weisen unterschiedliche Form-Schematizitätsausprägungen auf. Dabei gehören die Substantive der Ausprägung FORM-SCHEMA (Testsubstantive: *Kollege, Nefte, Schütze, Franzose, Geselle*) und der Ausprägung PERIPHERIE (Testsubstantive: *Graf, Held, Zar, Fürst, Nachbar*) der schwachen Flexion an, Substantive der Ausprägung STARK (Testsubstantive: *Dieb, Freund, Vogt, Kerl, Feind*) gehören offensichtlich der starken Flexion an. Um die Ausprägung STARK und die starken Flexionsformen graphisch voneinander zu unterscheiden, wird mit Kapitälchen gearbeitet.

Das Antwortverhalten ergibt sich aus 280 Antworten, die pro Form-Schematizitätsausprägung und Flexionsform gegeben wurden.²⁷ Abbildung 5.27 zeigt einen Waffleplot mit den Antworten für die einzelnen Form-Schematizitätsausprägungen und Flexionsformen. Dunkelgraue Rechtecke zeigen die Antworten, bei denen die Proband_innen angaben, die Form zu kennen; hellgraue Rechtecke die Antworten, bei denen die Form als unbekannt bewertet wurde. Ein Rechteck steht jeweils für eine Antwort.

Für die Ausprägung FORM-SCHEMA zeigt sich ein deutlicher Unterschied zwischen starken und schwachen Formen: Die schwachen Formen sind zu 95 % bekannt, während die starken zu 90 % unbekannt sind. Das Gegenteil lässt sich für die Ausprägung STARK festhalten: Hier sind schwache Formen zu 6,8 % und starke Formen zu 90,7 % bekannt. Da die Ausprägungen FORM-SCHEMA und STARK verschiedenen Deklinationsklassen angehören, ist dieses Ergebnis zu erwarten. Spannend ist ein Blick auf die Ausprägung PERIPHERIE. Die schwachen Formen sind mit 96,4 % ähnlich bekannt wie die schwachen Formen der Ausprägung FORM-SCHEMA. Die starken Formen der Ausprägung PERIPHERIE sind mit 53,9 % jedoch deutlich bekannter als die starken Formen der Ausprägung FORM-SCHEMA. Die Variation zeigt sich somit nicht in den schwachen Formen der Ausprägung

²⁷Die 280 Antworten ergeben sich daraus, dass 56 Proband_innen jeweils fünf Testsubstantive pro Form-Schematizitätsausprägung und Flexionsform bewertet haben.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

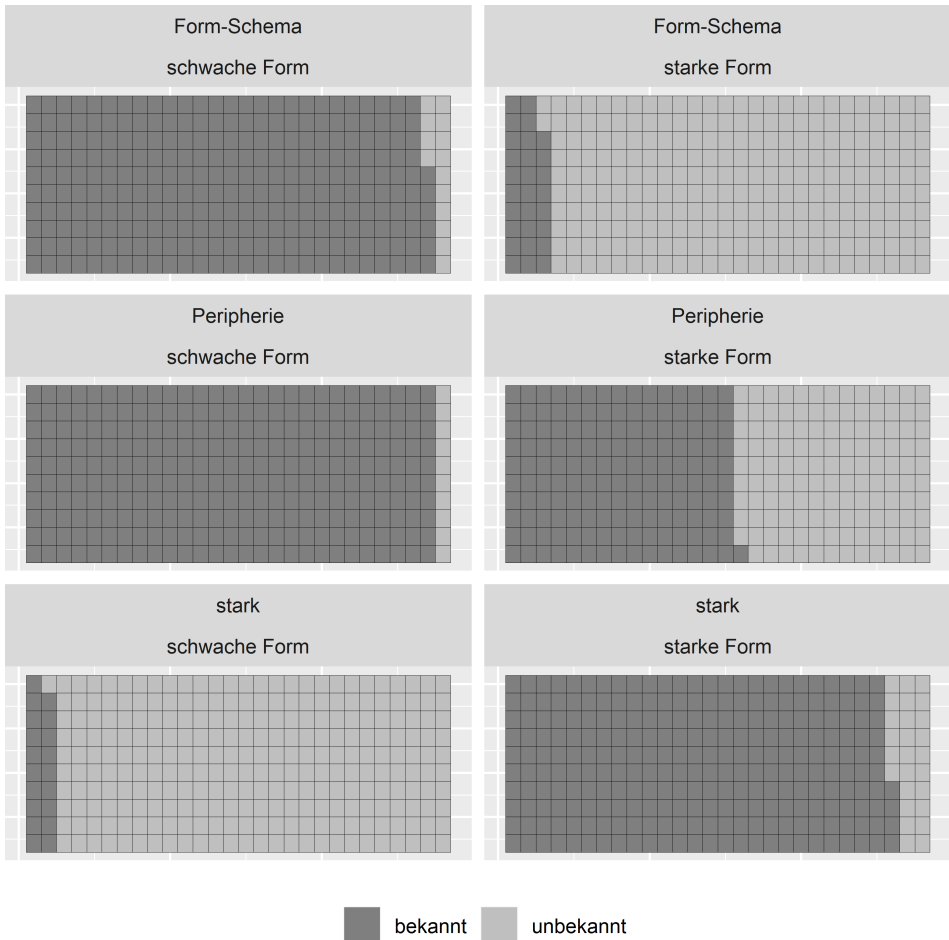


Abbildung 5.27: Antwortverhalten in der lexical-decision-Studie zu Form-Schematizität

PERIPHERIE, sondern in den starken, die bereits mental gefestigt zu sein scheinen. Ein ähnliches Antwortverhalten wurde in der Studie zu Frequenzeffekten bei den infrequenten Verben mit Schwankung in Korpora beobachtet (siehe Abschnitt 5.2.1).

Um herauszufinden, ob die Bekanntheit starker und schwacher Formen von den Versuchssubstantiven abhängt, werden diese näher in den Blick genommen. Pro Substantiv und Flexionsform wurden 56 Antworten gegeben, weil 56 Proband_innen an der Studie teilnahmen. Abbildung 5.28 zeigt die Bewertung für die Substantive der Ausprägung FORM-SCHEMA (*Neffe, Kollege*).

Die schwachen Formen befinden sich für alle Substantive auf einem ähnlichen Bekanntheitslevel von mindestens 92 %. Bei den starken Formen zeigen sich leichte Unterschiede: Während *des Franzoses* als völlig unbekannt eingestuft wird und *des Neffes* mit 94,6 % ebenfalls hohe Unbekanntheitswerte aufweist, liegen die Werte für *des Geselles* und *des Kolleges* niedriger (89,3 % Unbekanntheit). Mit 76,8 % weist *des Schützes* einen deutlich geringeren Anteil an Unbekanntheit auf als die anderen Testsubstantive. Dies könnte eventuell daran liegen, dass *Schütze* auch als Familienname verbreitet ist und ein Sternbild benennt. Zudem existiert das Wort *Schütz*, das auf einen elektromagnetischen Schalter referiert und den Genitiv auf *-es* bildet (Duden 2020). Allerdings ergibt die Suche nach *Schützes* im Archiv W der geschriebenen Sprache nur 210 Belege (DeReKo, Leibniz-Institut für Deutsche Sprache 2019b). Diese setzen sich vornehmlich aus dem Familiennamen *Schütze* im Genitiv oder im Plural zusammen. Trotz der kleinen Unterschiede zwischen den Testsubstantiven der Ausprägung FORM-SCHEMA deutet die visuelle Analyse auf eine systematische Bekanntheit der schwachen und eine systematische Unbekanntheit der starken Formen hin.

Ein entgegengesetztes Bild mit ähnlicher Eindeutigkeit ergibt sich für die Substantive der Ausprägung STARK (*Freund, Dieb*), wie Abbildung 5.29 zeigt.

Starke Formen sind weitgehend bekannt, die schwachen unbekannt. Allerdings zeigen sich auch hier kleine Unterschiede zwischen den Testsubstantiven. Hierbei fällt vor allem *des Vogts* mit vergleichsweise geringen Bekanntheitswerten auf (71,4 %). Auch hinsichtlich der schwachen Form fällt *Vogt* auf, diesmal mit etwas höheren Bekanntheitswerten als die restlichen Substantive (14,3 %). *Vogt* zählt zu den Testsubstantiven mit der geringsten Tokenfrequenz, sodass die Proband_innen das Substantiv eventuell nicht kannten und daher in der Bewertung unsicher waren.

Auch die Antwortverteilung bei den Substantiven der Ausprägung PERIPHERIE (*Vogt, Fürst*) scheint nicht stark von einzelnen Testsubstantiven beeinflusst worden zu sein, wie Abbildung 5.30 zeigt.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

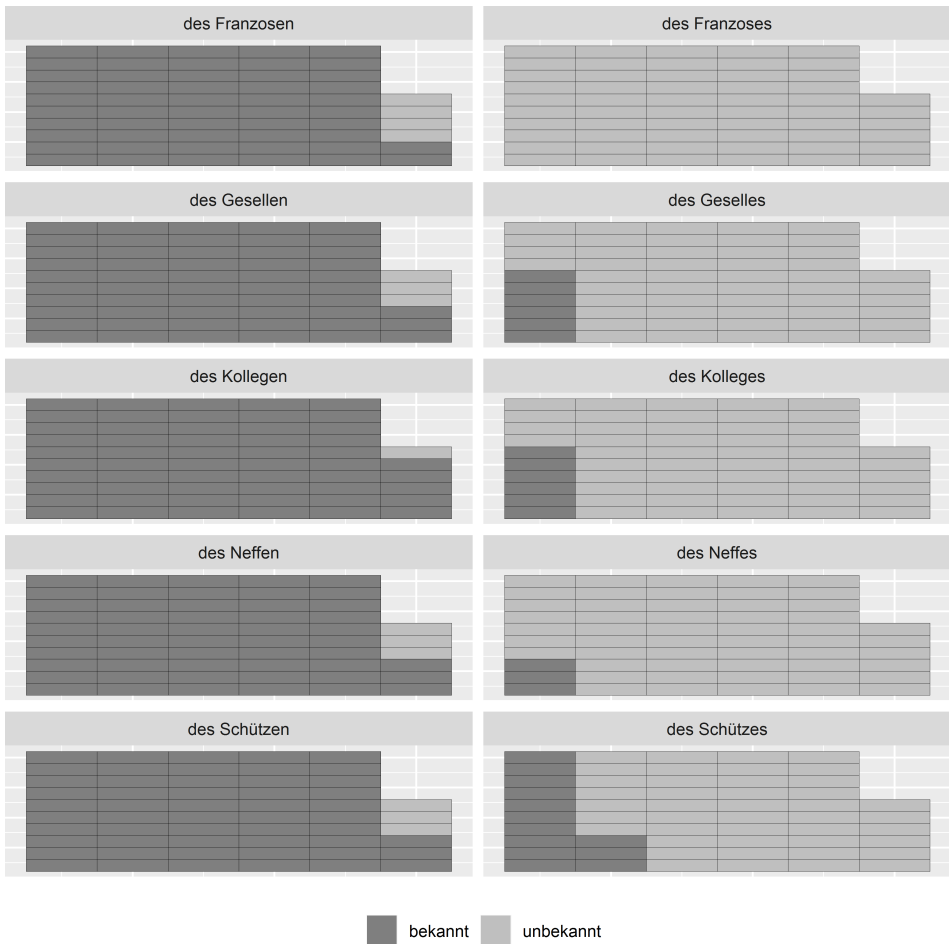


Abbildung 5.28: Antwortverhalten bei den Testsubstantiven der Ausprägung FORM-SCHEMA in der lexical-decision-Studie zu Form-Schematizität

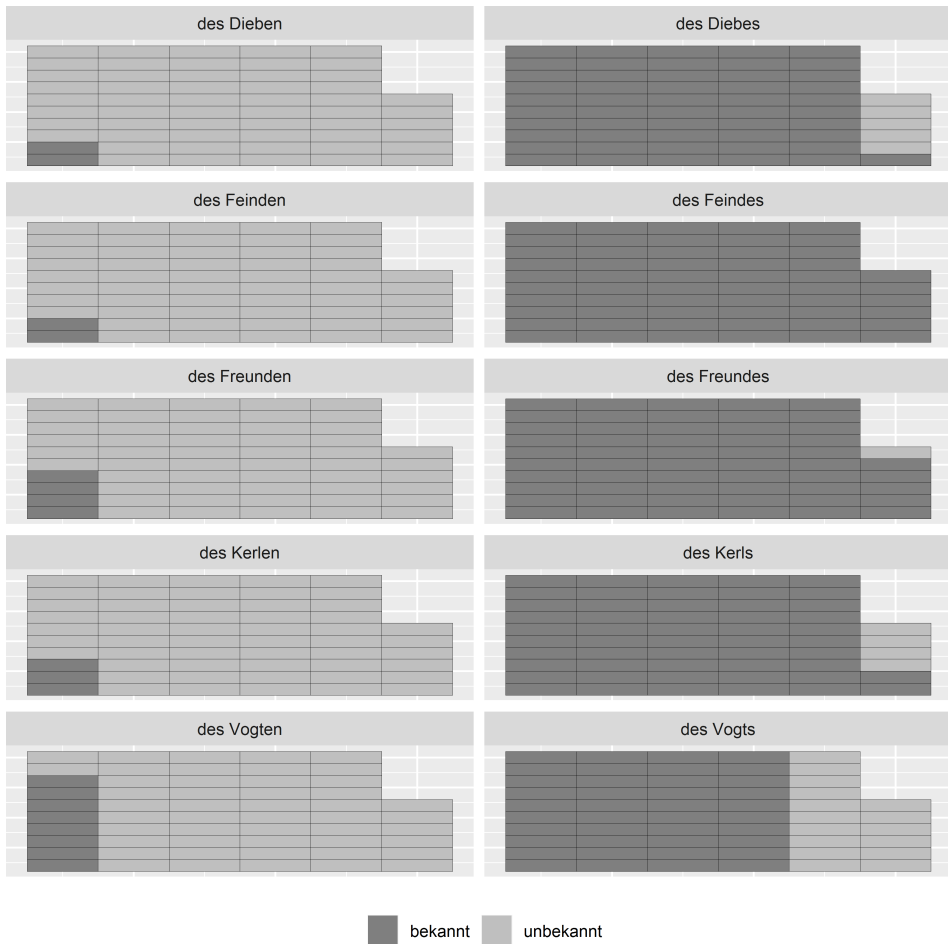


Abbildung 5.29: Antwortverhalten bei den Testsubstantiven der Ausprägung STARK in der lexical-decision-Studie zu Form-Schematizität

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

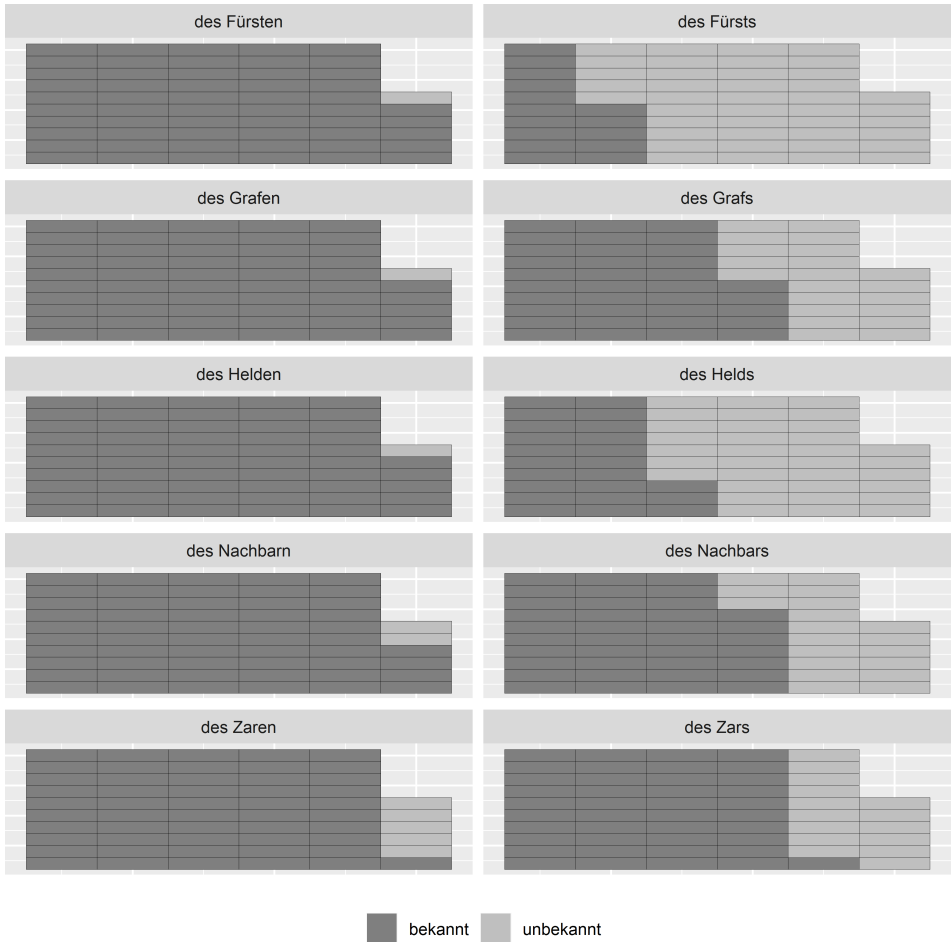


Abbildung 5.30: Antwortverhalten bei den Testsubstantiven der Ausprägung PERIPHERIE in der lexical-decision-Studie zu Form-Schematizität

Die schwachen Formen sind für alle Substantive bekannt, nur bei *des Zaren* sind die Bekanntheitswerte mit 91,1 % etwas geringer als bei den anderen Substantiven. Dafür weist *des Zars* mit 73,2 % vergleichsweise hohe Bekanntheitswerte auf. *Des Nachbars* und *des Grafs* sind mit 66,1 % und 62,5 % ähnlich bekannt. Deutlich weniger bekannt sind *des Helds* (41,1 %) und *des Fürsts* (26,8 %). Die Bekanntheit von *des Fürsts* ist damit ähnlich hoch wie die von *des Schützes*, das zum Form-Schema schwacher Maskulina zählt.

Die Unterschiede im Antwortverhalten zwischen den Testsubstantiven der Abstufung PERIPHERIE könnten durch Unterschiede in der Sonorität²⁸ der Stammauslaute bedingt sein: Die Formen *Helds* und *Fürsts* weisen aufgrund des Plosivs als Stammauslaut ein extrasilbisches *s* auf (Szczepaniak 2010: 114). Genitivformen auf *-s* sind bei Substantiven, deren Stamm auf Plosiv endet, daher generell selten. Substantive auf Plosiv tendieren stattdessen zu langen Genitivformen (*-es*): Lange Genitivformen machen bei Simplizia auf Plosiv (*des Kruges*) 79 % der Genitivformen aus und bei Simplizia auf *-st* (*des Dienstes*) sogar 89 % (Szczepaniak 2010: 113). Bei Frikativen im Auslaut ist die relative Frequenz der langen Genitivform hingegen deutlich geringer (66 %).

Die vergleichsweise geringe Frequenz von langen Genitivformen bei Frikativen im Stammauslaut könnte den höheren Bekanntheitswert von *des Grafs* im Vergleich zu *des Helds* und *des Fürsts* erklären: Die Beschränkungen in der Phonotaktik blockieren die kurze Endung hier nicht so stark wie bei den Plosiven. Die lange Form *-es* scheint für die ursprünglichen schwachen Maskulina dabei keine Option zu sein (**des Heldes*, **des Fürstes*). Bei *Zar* und *Nachbar* stellt die kurze Form aufgrund des vokalischen Auslauts keine phonotaktischen Probleme dar, was sich ebenfalls in der hohen Bekanntheit spiegelt. Insgesamt lässt sich daher festhalten, dass sich die globale Verteilung der Antworten innerhalb der einzelnen Testsubstantive widerspiegelt. Die wenigen Unterschiede im Antwortverhalten zwischen den Substantiven der Ausprägung PERIPHERIE lassen sich durch den Einfluss von Phonotaktik erklären.

Für die konfirmatorische Analyse des Antwortverhaltens wird ein generalisiertes gemischtes lineares Modell genutzt, das Form-Schematizität und Flexion als feste Effekte enthält, die interagieren: Antwortverhalten ~ Form-Schematizität * Flexion + (1|Proband_in) + (1|Item). Als zufällige Effekte werden *random intercepts* für Proband_in und Versuchitem (Lemma) berücksichtigt. *Random slopes* werden aufgrund von Konvergierungsproblemen und *singular fits* für Proband_innen nicht genutzt, für Versuchitems sind *random slopes* nicht sinnvoll,

²⁸Das Prinzip der Sonorität unterscheidet Laute nach ihrer Lautstärke (Szczepaniak 2010: 107). Dabei stellen Vokale die sonorsten Laute dar, es folgen Liquide, Nasale und schließlich Frikative und Plosive als unsonore Laute (Vennemann 1987: 167).

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

da für die Form-Schematizitätsausprägungen verschiedene Items genutzt werden. Die Proband_innen schwanken mit einer Standardabweichung von 0,71 um den y-Achsenabschnitt, die Versuchitems variieren geringer mit einer Standardabweichung von 0,51. In Tabelle 5.14 sind die *log odds* des Modells für die Unbekanntheit starker und schwacher Formen aufgelistet, die entsprechende Kreuztabelle (Tabelle I.3) der *log odds* befindet sich in Anhang I.3. Abbildung 5.31 zeigt die auf den *log odds* basierende Wahrscheinlichkeit für die Unbekanntheit starker und schwacher Formen pro Form-Schematizitätsausprägung. Kreise stehen für schwache Formen, Dreiecke für starke.

Tabelle 5.14: Werte des Modells für die Unbekanntheit starker und schwacher Formen in der lexical-decision-Studie zu Form-Schematizität

	Wert	SE	z	p
(Intercept = Form-Schema & schwach)	-3,28	0,39	-8,47	<0,01
Form-Schematizität: Peripherie	-0,38	0,54	-0,69	0,49
Form-Schematizität: stark	6,19	0,52	11,94	<0,01
Flexion: stark	5,77	0,39	14,76	<0,01
Form-Schematizität: Peripherie & Flexion: stark	-2,29	0,51	-4,46	<0,01
Form-Schematizität: stark & Flexion: stark	-11,23	0,56	-19,94	<0,01

Bei Substantiven der Ausprägung FORM-SCHEMA liegt die Wahrscheinlichkeit für Unbekanntheit für starke Formen bei 0,92 und für schwache Formen bei 0,04. Starke Formen sind also mit großer Wahrscheinlichkeit unbekannt und schwache bekannt. Das Gegenteil ist für die Ausprägung STARK der Fall: Starke Formen haben eine Wahrscheinlichkeit für Unbekanntheit von 0,07, schwache von 0,95. Die 95 %-Konfidenzintervalle sind recht klein und überlappen sich nicht. Bei der Ausprägung PERIPHERIE zeigt sich eine Wahrscheinlichkeit von 0,03 für schwache Formen, die Wahrscheinlichkeit für Unbekanntheit ist also ungefähr so gering wie für schwache Formen der Ausprägung FORM-SCHEMA. Bei starken Formen der Ausprägung PERIPHERIE liegt die Wahrscheinlichkeit bei 0,46, Bekanntheit und Unbekanntheit sind für starke Formen also in etwa gleich wahrscheinlich. Die starken Formen der Ausprägung PERIPHERIE haben auch ein langes 95 %-Konfidenzintervall, das jedoch nicht mit anderen 95 %-Konfidenzintervallen überlappt. Bis auf den *p*-Wert für schwache Formen der Ausprägung PERIPHERIE sind alle *p*-Werte für alle Formen und Ausprägungen im Vergleich zum Referenzlevel FORM-SCHEMA mit schwachen Formen jeweils unter dem α -Level von 0,01. Es ist daher von systematischen Effekten auszugehen. Der ausbleibende Unterschied in der Bewertung von schwachen Formen der Ausprägungen FORM-SCHEMA und

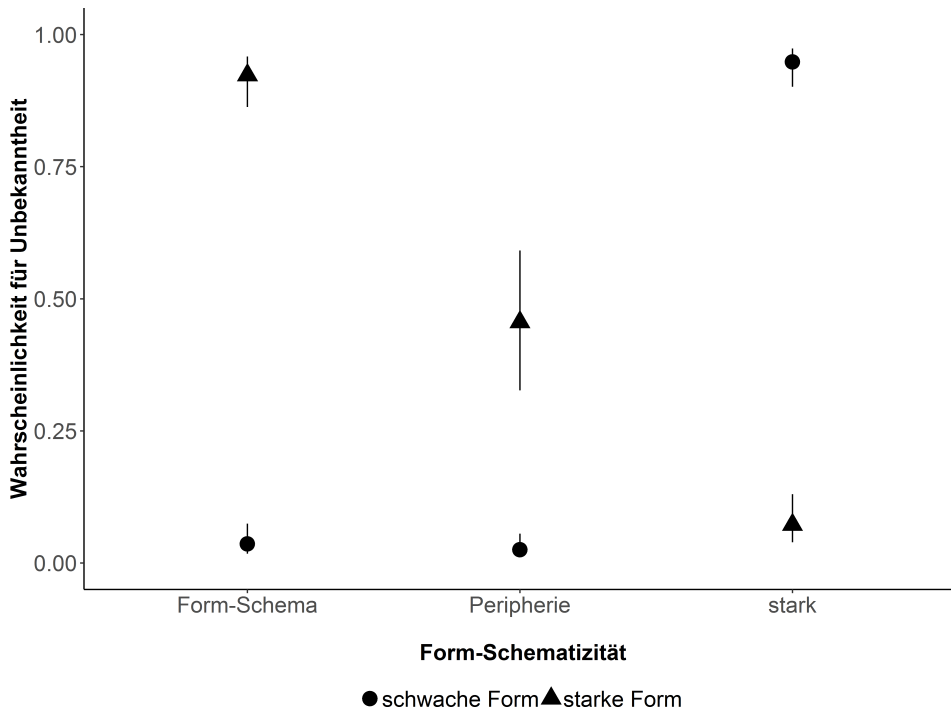


Abbildung 5.31: Wahrscheinlichkeit für die Unbekanntheit starker und schwacher Formen in der lexical-decision-Studie zu Form-Schematizität

PERIPHERIE war bereits in der visuellen Analyse zu erkennen. Die Effektstärke ist mit einem R^2m von 0,63 (theoretisch) bzw. 0,59 (delta) recht hoch. Unter Einbezug der zufälligen Effekte erhöht sie sich mit einem R^2c von 0,7 (theoretisch) und 0,65 (delta) leicht.

Die Analyse des Antwortverhaltens in der lexical-decision-Studie zeigt, dass starke und schwache Formen für starke und schwache Maskulina gegenteilige Bekanntheitsgrade aufweisen. Das ist aufgrund der Flexionklassenzugehörigkeit auch zu erwarten. Schwache Substantive in der Peripherie des Form-Schemas weisen hohe Bekanntheitsgrade für schwache Formen auf, bei starken Formen halten sich Bekanntheit und Unbekanntheit die Waage. Diese Verteilung lässt sich für alle Testsubstantive erkennen. Die Hypothesen aus Abschnitt 4.5.1 werden somit bestätigt.

5.4.1.3 Sentence-maze-Studie

In der sentence-maze-Studie wählten die Proband_innen zwischen starken *-(e)s* und schwachen *-(e)n* Genitivformen der Testsubstantive. Dabei wurden Testsubstantive der Ausprägungen FORM-SCHEMA (Testsubstantive: *Kollege, Neffe, Schütze*), PERIPHERIE (Testsubstantive: *Graf, Held, Zar*) und STARK (Testsubstantive: *Dieb, Freund, Vogt*) präsentiert. Pro Ausprägung liegen 396 Antworten vor.²⁹ Abbildung 5.32 zeigt das Antwortverhalten pro Form-Schematizitätsausprägung in einem Waffleplot. Rechtecke in dunkelgrau stehen für die Wahl der schwachen Form, Rechtecke in hellgrau für die Wahl der starken. Ein Rechteck steht für zwei Antworten.

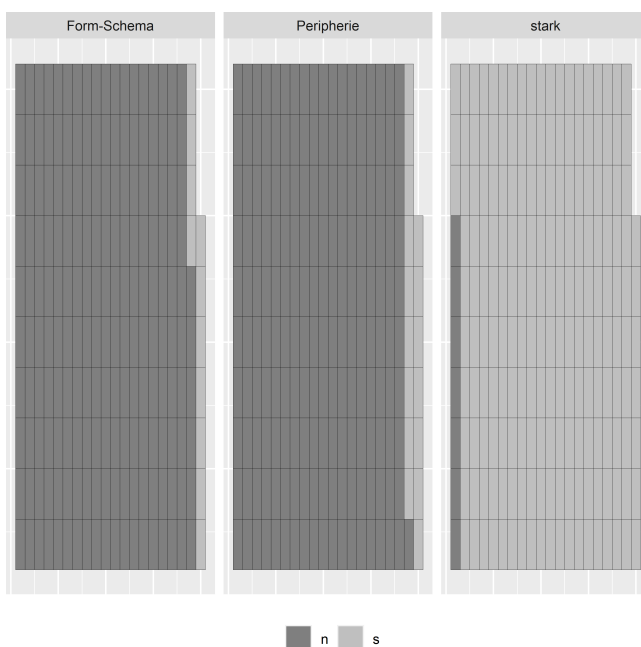


Abbildung 5.32: Antwortverhalten in der sentence-maze-Studie zu Form-Schematizität

Für Substantive der Ausprägung FORM-SCHEMA wählten die Proband_innen mit 94,2 % vorrangig schwache Formen, für Substantive der Ausprägung STARK mit 96,2 % vorrangig starke. Interessant ist ein Blick auf die Peripherie: Auch hier

²⁹Diese ergeben sich daraus, dass 132 Proband_innen Flexionsformen für jeweils drei Testsubstantive wählten.

herrschen mit 91,7 % schwache Formen vor. Es zeigt sich somit ein klar verteiltes Antwortverhalten. Die generelle Tendenz zu schwachen Formen in der Peripherie mag zunächst überraschen, da die starken Formen in der lexical-decision-Studie immerhin in ungefähr der Hälfte der Antworten als bekannt angesehen wurden. Diese Diskrepanz ist durch das Versuchsdesign bedingt: Anders als in der lexical-decision-Studie werden die Formen in der sentence-maze-Studie nicht einzeln bewertet, sondern beide stehen gleichzeitig zur Auswahl. Die mental gefestigte schwache Form kann die weniger gefestigte starke Form so leicht statistisch ausstechen (Goldberg 2019: 74–94, zum statistischen Vorkaufsrecht siehe Abschnitt 2.1.3).

Die globale Verteilung der Antworten zeigt sich auch im Antwortverhalten für die einzelnen Testsubstantive: Die starke Endung wurde bei den Substantiven der Ausprägung FORM-SCHEMA maximal zu 6,8 % gewählt und die schwache Endung für Substantive der Ausprägung STARK maximal zu 6,1 %, siehe die Abbildungen I.7 und I.8 in Anhang I.3. Nur bei der Ausprägung PERIPHERIE zeigen sich leichte Unterschiede zwischen den Testsubstantiven, wie Abbildung 5.33 zeigt. Pro Substantiv wurden 132 Antworten gegeben. Ein Rechteck steht für zwei Antworten.

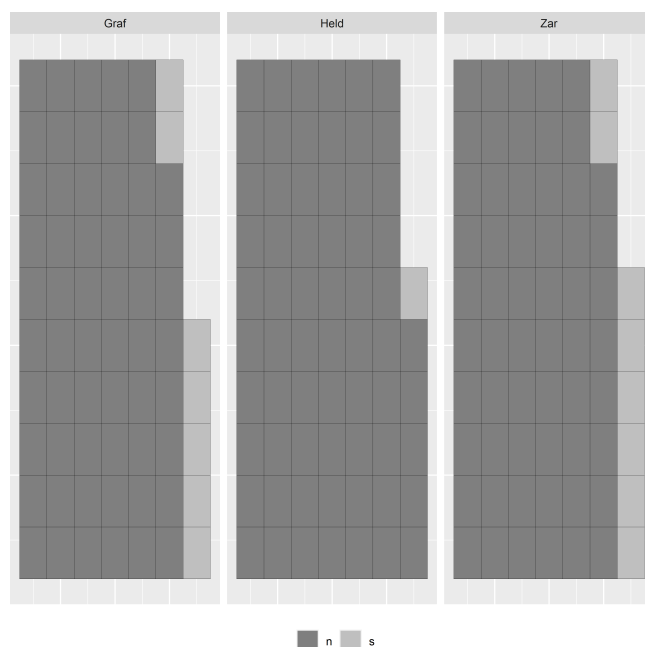


Abbildung 5.33: Antwortverhalten bei den Testsubstantiven der Ausprägung PERIPHERIE in der sentence-maze-Studie zu Form-Schematizität

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Das Tests substantiv *Held* weist mit 1,5 % deutlich weniger Antworten für starke Formen auf als die anderen Tests substantive, bei denen die Antworten für starke Formen jeweils um die 10 % ausmachen. In der lexical-decision-Studie wies *Held* im Vergleich zu *Graf* und *Zar* ebenfalls weniger Zustimmung für starke Formen auf (41,1 % Zustimmung für *Held*, 62,5 % für *Graf* und 73,2 % für *Zar*). Dies könnte wiederum an den Unterschieden in der Sonorität des Stammaslauts liegen: Da *Held* auf einen Plosiv auslautet, ist die kurze Genitivform -s phonotaktisch ungünstig, da sie zu einem extrasilbischen Element führt (Szczepaniak 2010: 114). Die Kombination aus -s und einem Frikativ (*Graf*s) oder vokalischem Auslaut (*Zar*s) ist hingegen phonotaktisch gesehen weniger bzw. überhaupt nicht problematisch (Szczepaniak 2010, siehe hierzu genauer Abschnitt 5.4.1.2). Unabhängig von den kleinen Unterschieden zwischen den Tests substantiven der Ausprägung PERIPHERIE ist festzuhalten, dass die schwachen Formen in der direkten Wahl zwischen stark und schwach bevorzugt werden.

Für die konfirmatorische statistische Analyse wird ein generalisiertes gemischtes lineares Modell mit Form-Schematizitätsausprägung als festem Effekt gerechnet: Antwortverhalten \sim Form-Schematizität + (1|Proband_in) + (1|Item). Als zufällige Effekte sind *random intercepts* für Proband_in und Versuchitem (Lemma) vorgesehen. *Random slopes* werden für Proband_innen aufgrund von *singular fits* nicht genutzt, für Versuchitems sind sie in diesem Modell nicht sinnvoll, da verschiedene Items für die Form-Schematizitätsausprägungen genutzt werden.

Die Proband_innen variieren mit einer Standardabweichung von 0,48 um den y-Achsenabschnitt, die Versuchitems etwas weniger mit einer Standardabweichung von 0,37. Die *log odds* für die Wahl starker Formen sind in Tabelle 5.15 aufgelistet. Abbildung 5.34 zeigt die auf den *log odds* basierende Wahrscheinlichkeit für die Wahl starker Formen.

Tabelle 5.15: Werte des Modells für die Wahl starker Formen in der sentence-maze-Studie zu Form-Schematizität

	Wert	SE	z	p
(Intercept = Form-Schema)	-2,92	0,34	-8,67	<0,01
Form-Schematizität: Peripherie	0,34	0,42	0,82	0,42
Form-Schematizität: stark	6,31	0,54	11,75	<0,01

Für Substantive der Ausprägungen FORM-SCHEMA und PERIPHERIE ist die Wahl starker Formen mit Werten von 0,05 und 0,07 sehr unwahrscheinlich, das Gegenteil ist für Substantive der Ausprägung STARK der Fall, bei denen die Wahrscheinlichkeit bei 0,97 liegt. Die 95 %-Konfidenzintervalle sind relativ klein und

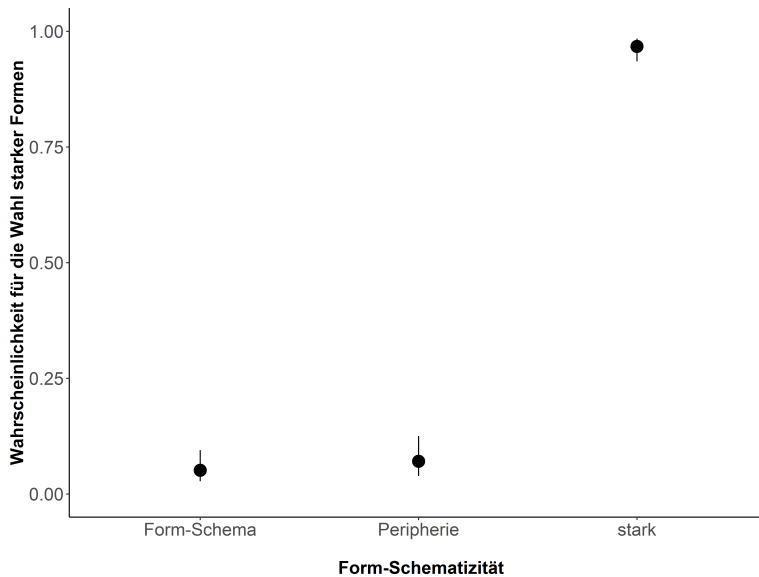


Abbildung 5.34: Wahrscheinlichkeit für die Wahl starker Formen in der sentence-maze-Studie zu Form-Schematizität

zeigen nur für die Ausprägungen FORM-SCHEMA und PERIPHERIE Überlappungen. Die p -Werte unterschreiten nur für die Ausprägung STARK im Vergleich zum Referenzlevel FORM-SCHEMA das α -Level von 0,01. Für die Ausprägungen FORM-SCHEMA und PERIPHERIE ist dagegen nicht von einem systematischen Unterschied im Antwortverhalten auszugehen. Die Effektstärke des Modells ist mit einem R^2m von 0,7 (theoretisch) und 0,66 (delta) hoch. Mit den zufälligen Effekten erhöht sie sich mit einem R^2c von 0,73 (theoretisch) und 0,69 (delta) nur leicht. Die Effektstärke des Modells ist vergleichbar mit der Effektstärke des Modells für die lexical-decision-Studie.

Insgesamt zeigt sich im Antwortverhalten eine klare Präferenz für schwache Formen bei den schwachen Maskulina unabhängig davon, ob sie zum Prototyp oder zur Peripherie des Form-Schemas gehören. Für starke Substantive zeigt sich erwartbarerweise eine klare Präferenz für starke Formen. Die Verteilung bestätigt somit die Hypothesen in Abschnitt 4.5.1. Die Bevorzugung für die schwachen Formen in der Peripherie des Form-Schemas lässt sich durch das Versuchsdesign erklären, in dem schwache und starke Formen nebeneinander präsentiert werden. Daher kann die mental gefestigte Form (schwach) die weniger gefestigte (stark) leichter statistisch ausstechen.

5.4.1.4 Zusammenfassung

In allen Studien ist der Einfluss der Form-Schematizität auf die Wahl der Flexionsendung zu erkennen. Die Unterschiede im Antwortverhalten zwischen den Studien lassen sich jeweils durch das Versuchsdesign erklären: Im Produktionsexperiment mussten die Proband_innen die Formen selbst bilden, in der lexical-decision-Studie wurden starke und schwache Formen separat beurteilt und in der sentence-maze-Studie jeweils zwischen der starken und schwachen Form eines Testsubstantivs gewählt.

Der Einfluss der Form-Schematizität zeigt sich am deutlichsten im Produktionsexperiment zu der self-paced-reading-Studie und in der lexical-decision-Studie. Wenn ein Substantiv prototypischer Vertreter des Form-Schemas schwacher Maskulina ist, werden schwache Formen im Produktionsexperiment gebildet, während starke Formen gemieden werden. Dasselbe zeigt sich in der lexical-decision-Studie: Schwache Formen prototypischer Vertreter des Form-Schemas werden als bekannt und starke Formen als unbekannt bewertet. Bei starken Maskulina ist das Gegenteil der Fall: Hier werden starke Formen im Produktionsexperiment gebildet und schwache gemieden. Genauso werden in der lexical-decision-Studie starke Formen von starken Maskulina als bekannt bewertet und schwache als unbekannt. Die Substantive in der Peripherie des Form-Schemas weisen im Produktionsexperiment und in der lexical-decision-Studie Schwankungen auf: Im Produktionsexperiment wird die schwache Form *-en* zu 55 % gewählt und die starke Form *-(e)s* zu 30 %. Zudem werden Doppelformen genannt. Auch in der lexical-decision-Studie ist in Hinblick auf starke Formen Variation zu erkennen: Nur 50 % der starken Formen von Substantiven in der Peripherie des Form-Schemas werden als bekannt bewertet, bei den schwachen liegt der Bekanntheitsgrad hingegen bei ca. 90 %. Der Einfluss des Form-Schemas schwacher Maskulina lässt sich somit sowohl in der Produktion als auch in der Bewertung von Formen nach Bekanntheit beobachten und wirkt bei Pseudosubstantiven genauso wie bei real existierenden. Die Einblicke in die Sprachproduktion zeigen zudem, dass bei Pseudosubstantiven weitere Formen (Nullendungen, Formen auf *-ers*) möglich sind.

Die Ergebnisse der sentence-maze-Studie weichen etwas von den Ergebnissen der anderen Studien ab. Zwar werden in der sentence-maze-Studie ebenfalls schwache und starke Formen je nach Flexionsklasse gewählt, jedoch ist in der Peripherie des Form-Schemas keine Variation zu beobachten. Der Einfluss des Form-Schemas ist daher in der sentence-maze-Studie nur insofern zu erkennen, als dass die Mitglieder des Form-Schemas schwacher Maskulina als solche erkannt werden, unabhängig davon, ob sie prototypische oder periphere Vertreter

darstellen. Das Ausbleiben der Schwankungen in der Peripherie geht auf das Studiendesign zurück: Die schwache und die starke Form wurden direkt nebeneinander als Antwortmöglichkeiten präsentiert, sodass die stärker gefestigte schwache Form die weniger gefestigte starke leichter statistisch ausstechen konnte. Diese These stützen die anderen Studien, in denen Schwankungen im Antwortverhalten für die Testsubstantive in der Peripherie des Form-Schemas zu beobachten sind. Die *sentence maze task* scheint sich daher nur bedingt zu eignen, um Variation zu untersuchen.

Insgesamt bestätigt die Analyse des Antwortverhaltens den Einfluss der Form-Schematizität auf die Deklination von Maskulina. Unabhängig von dem Untersuchungsdesign werden starke und schwache Formen je nach Flexionsklassenzugehörigkeit bevorzugt. Im Produktionsexperiment zu der *self-paced-reading*-Studie und in der *lexical-decision*-Studie lässt sich zudem Variation in der Peripherie des Form-Schemas schwacher Maskulina beobachten. Generell ist dabei festzuhalten, dass die schwachen Formen auch in der Peripherie des Form-Schemas schwacher Maskulina stark gefestigt sind, da sie hohe Bekanntheitswerte in der *lexical-decision*-Studie aufweisen und die starken Formen in der *sentence-maze*-Studie ausstechen. Umgekehrt scheinen die starken Formen von Substantiven in der Peripherie des Form-Schemas nicht so stark gefestigt zu sein, da sie in der *lexical-decision*-Studie nur zur Hälfte als bekannt bewertet werden. Das Produktionsexperiment zeigt jedoch, dass Proband_innen in der Flexion von Pseudosubstantiven aus der Peripherie des Form-Schemas zwischen starken und schwachen Formen schwanken. Dies verdeutlicht, dass die starken Formen von Substantiven in der Peripherie des Form-Schemas durchaus die schwachen in der Sprachproduktion statistisch ausstechen können. Es ist jedoch möglich, dass dies bei real existierenden Substantiven in der Peripherie des Form-Schemas aufgrund der Tokenfrequenzunterschiede zwischen schwachen und starken Formen unwahrscheinlicher ist.

5.4.2 Reaktionszeiten

In diesem Abschnitt werden die Reaktionszeiten der Form-Schematizitätsstudien analysiert und die Ergebnisse anschließend miteinander verglichen. Dabei ist zu bedenken, dass die Reaktionszeiten in den Studien nicht direkt vergleichbar sind, da sie Lesezeiten (*self-paced reading*) sowie kontextlose (*lexical decision*) und kontextbedingte (*sentence maze*) Reaktionszeiten widerspiegeln.

Für die einzelnen Studien werden verschiedene Reaktionszeitmuster erwartet: Für die *self-paced-reading*-Studie wird angenommen, dass schwache Formen bei

Testsubstantiven, die dem Prototyp des Form-Schemas entsprechen, schnell gelesen werden, starke hingegen langsam. Das Gegenteil wird für starke Testsubstantive erwartet. Für die Formen in der Peripherie des Form-Schemas wird kein Unterschied in den Lesezeiten antizipiert (siehe hierzu genauer Abschnitt 4.4). Für die lexical-decision- und die sentence-maze-Studie wurden ursprünglich geringe Reaktionszeiten für den Prototyp des Form-Schemas und starke Maskulina erwartet und höhere für die Peripherie. In der Peripherie des Form-Schemas sind schwache und starke Formen ähnlich wahrscheinlich, weshalb die Beurteilung der Formen verlangsamt sein könnte. Bei den anderen Ausprägungen kann eine Form die andere hingegen klar statistisch ausstechen. Dies wurde jedoch in der Prästudie nicht bestätigt, sodass die Hypothesen angepasst wurden (siehe hierzu ausführlich Abschnitt 4.5). Für die lexical-decision-Studie wird auf Basis der Prästudie eine Interaktion zwischen Form-Schematizitätsausprägung und Flexionsform erwartet: Schwache Formen werden für schwache Substantive (sowohl bei prototypischen als auch bei peripheren Vertretern des Form-Schemas) schnell bewertet und starke langsam, das Gegenteil ist für starke Substantive der Fall. In der sentence-maze-Studie wird auf Basis der Prästudie angenommen, dass die Wahl zwischen starken und schwachen Formen bei Substantiven in der Peripherie des Form-Schemas schneller abläuft als bei den Substantiven der anderen Form-Schematizitätsausprägungen (siehe hierzu Abschnitt 4.5).

5.4.2.1 Self-paced-reading-Studie

In diesem Abschnitt werden die Lesezeiten innerhalb der self-paced-reading-Studie analysiert. Abbildung 5.35 zeigt die logarithmierten Lesezeiten in Bienenwabenplots für die einzelnen Testsubstantive samt arithmetischem Mittel und Standardabweichung der starken und schwachen Formen. Das arithmetische Mittel (M), die Standardabweichung (SD) und der Standardfehler (SE) sind zudem gesondert aufgeführt. Dunkelgraue Kreise zeigen Lesezeiten für schwache Formen, hellgraue Dreiecke Lesezeiten für starke Formen.

Die Testsubstantive weisen vergleichbare Lesezeiten auf, wobei *Knatt* und *Grettel* vermutlich aufgrund ihrer Kürze etwas geringere Lesezeiten evozieren. Starke und schwache Formen sind pro Testsubstantiv recht gleichmäßig auf die Lesezeiten verteilt. Bei *Grettel* fällt ein Datenpunkt mit deutlich geringerer Lesezeit als die anderen Datenpunkte ins Auge. Dieser weicht jedoch nicht mehr als drei Standardabweichungen vom Mittelwert ab und wird daher nicht als extremer Wert betrachtet (siehe Abschnitt 4.4.6). Abseits davon sind die Lesezeiten für starke und schwache Formen ähnlich. Nur bei *Schettose* weist die starke Form

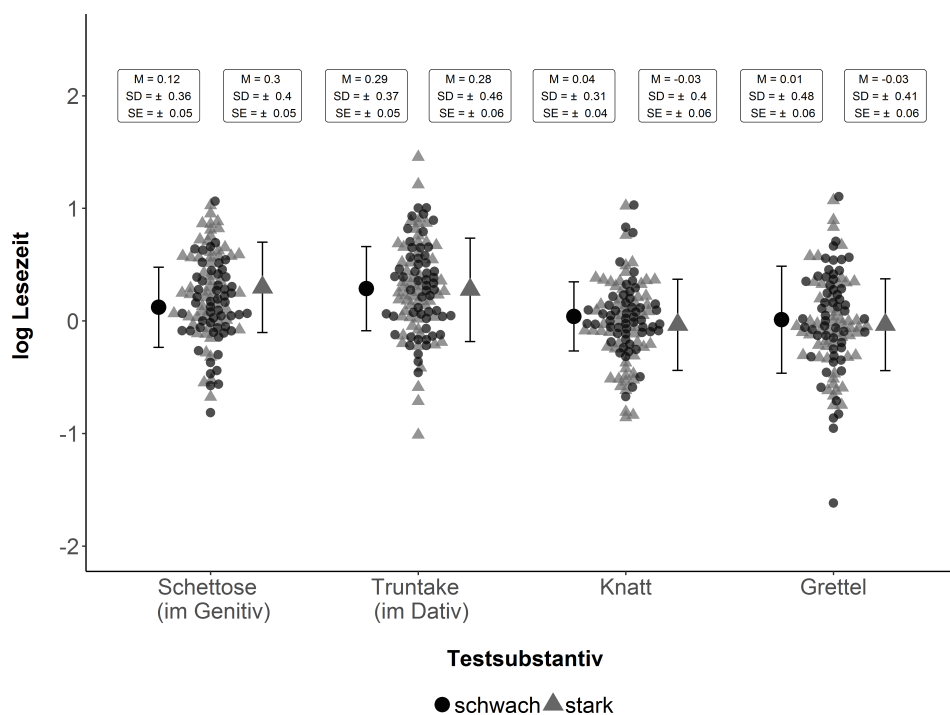


Abbildung 5.35: Lesezeiten in der self-paced-reading-Studie

höhere Lesezeiten auf als die schwache. Dies zeigt sich in der Verteilung der Lesezeiten und in den Mittelwerten: Die Lesezeiten für die starke Form von *Schettose* haben einen Mittelwert von 0,3 mit einer Standardabweichung von 0,4, während der Mittelwert für die schwache Form mit 0,12 und einer Standardabweichung von 0,36 deutlich niedriger liegt. Die anderen Testsubstantive weisen maximale Differenzen von 0,07 zwischen den Mittelwerten auf.

Die visuelle Analyse der direkt nach den Testitems gelesenen Elemente lässt keine spill-over-Effekte erkennen,³⁰ auch hinsichtlich der Fragen zu den Testitems (Haben Sie *des Schettoses* oder *des Schettosen* gelesen?) ergab sich kein Unterschied in den Reaktionszeiten (siehe Abbildungen I.10 und I.11 in Anhang I.3). Zudem weisen die Daten nicht auf Reihenfolgeeffekte hin (siehe Abbildung I.12 in Anhang I.3). Längeneffekte können für *des Knatten* im Vergleich zu dem um

³⁰Die Elemente nach den Testitems wurden jeweils konstant gehalten, auf *des Schettoses* bzw. *des Schettosen* folgte also dasselbe Wort bzw. dieselbe Wortgruppe (siehe Abschnitt 4.4.3). Unterschiede in den Lesezeiten zwischen den Elementen nach den Testitems würden damit auf spill-over-Effekte hindeuten.

einen Buchstaben kürzeren starken Genitiv (*des Knatts*) ausgeschlossen werden (siehe Abbildung I.13 in Anhang I.3), für *Truntake* ist es hingegen möglich, dass die im Dativ kürzere starke Form (*dem Truntake*) vereinzelt kürzere Lesezeiten hervorruft als die schwache (siehe Abbildung I.14 in Anhang I.3). Bei den anderen Testsubstantiven sind starke und schwache Formen gleich lang.

Um den Unterschied in den logarithmierten Lesezeiten zwischen starken und schwachen Formen näher bestimmen zu können, werden Scatterplots der Testitems betrachtet, siehe Abbildung 5.36. Ein Punkt steht für die Lesezeit einer Person für die starke bzw. schwache Form. Die graue Linie markiert den Bereich, in dem die starke und die schwache Form dieselben Lesezeiten hervorrufen würden.

Bei *Schettose* scheinen die Lesezeiten für die starke Endung etwas erhöht zu sein, wie die Punkte oberhalb der grauen Linie zeigen. Bei *Truntake* ist hingegen kein Effekt der schwachen und starken Endung zu sehen, die Punkte verteilen sich regelmäßig um die graue Linie. Der ausbleibende Effekt bei *Truntake* könnte der unterschiedlichen Länge der schwachen (*dem Truntaken*) und starken (*dem Truntake*) Form geschuldet sein, wie oben bereits erwähnt wurde. Bei *Knatt* und *Grettel* scheinen schwache Endungen etwas erhöhte Lesezeiten zu evozieren als starke, wie die Punkte unterhalb der grauen Linie zeigen. Dieser Trend ist jedoch weniger ausgeprägt als der Trend zu höheren Lesezeiten für die starke Form bei *Schettose*.

Für die statistische Analyse der Reaktionszeiten wird ein gemischtes lineares Modell mit Form-Schematizitätsausprägung und Flexionsform als feste Effekte gerechnet, die interagieren: $\text{Lesezeit} \sim \text{Form-Schematizität} * \text{Flexion} + (1|\text{Proband_in})$. Die Form-Schematizitätsausprägungen werden jeweils durch die Testitems repräsentiert. Teilnehmer_innen werden als zufällige Effekte mit *random intercepts* berücksichtigt, *random slopes* sind nicht möglich, da pro Ausprägung nur ein Testitem genutzt wird. Die Standardabweichung der Proband_innen um den y-Achsenabschnitt beträgt 0,23 mit einer restlichen Residuenvarianz von 0,32. Die Struktur der zufälligen Effekte ist damit vergleichbar mit der der zufälligen Effekte in den Modellen zu Reaktionszeiten in der Frequenz- und Prototypizitätsstudie. Die geschätzten Werte des Modells sind in Tabelle 5.16 aufgeschlüsselt, eine entsprechende Kreuztabelle (Tabelle I.6) befindet sich in Anhang I.3. Abbildung 5.37 zeigt die vom Modell vorhergesagten Lesezeiten. Kreise stellen schwache Formen dar, Dreiecke starke.

Für *Schettose* ist ein Unterschied zwischen starker und schwacher Form zu erkennen, während für die anderen Testsubstantive starke und schwache Formen nah beieinander liegen: Bei *Truntake* lässt sich kein Unterschied zwischen

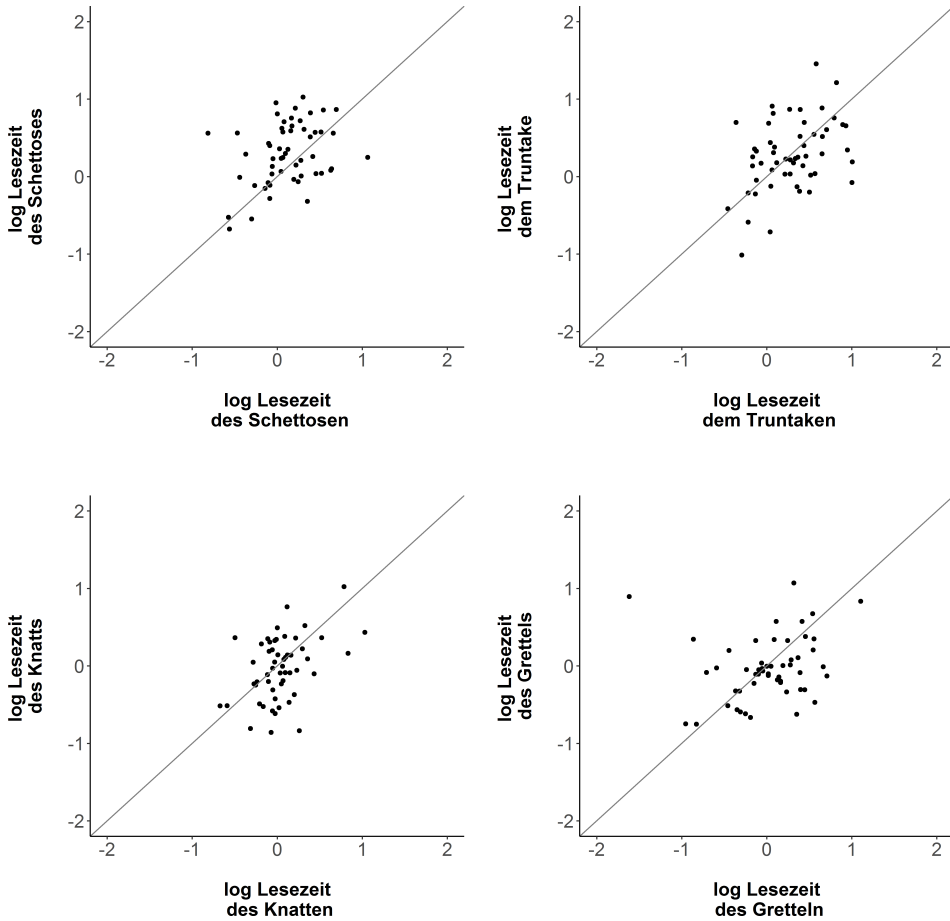


Abbildung 5.36: Scatterplots der Lesezeiten für starke und schwache Formen der Testsubstantive

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Tabelle 5.16: Werte des Modells für die Lesezeiten in der self-paced-reading-Studie

	Wert	SE	<i>t</i>	df	<i>p</i>
(Intercept = Schettose & schwach)	0,12	0,05	2,24	236,10	0,03
Endung: stark	0,18	0,06	2,85	378,00	<0,01
Item: Grettel	-0,11	0,06	-1,78	378,00	0,08
Item: Knatt	-0,08	0,06	-1,31	378,00	0,19
Item: Truntake	0,17	0,06	2,67	378,00	<0,01
Endung: stark & Item: Grettel	-0,22	0,09	-2,53	378,00	0,01
Endung: stark & Item: Knatt	-0,25	0,09	-2,86	378,00	<0,01
Endung: stark & Item: Truntake	-0,19	0,09	-2,13	378,00	0,03

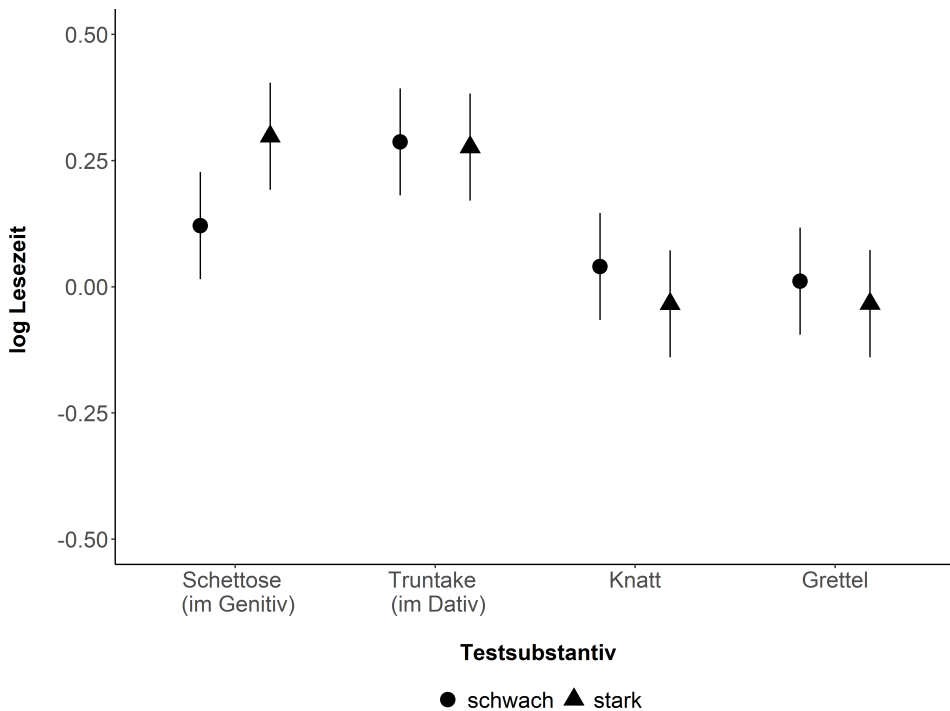


Abbildung 5.37: Vorhergesagte Lesezeiten in der self-paced-reading-Studie

der starken und schwachen Form feststellen. Bei *Knatt* und *Grettel* tendieren starke Formen dazu, etwas schneller gelesen zu werden als schwache, die 95 %-Konfidenzintervalle überlappen sich jedoch deutlich. Es ist somit nicht davon auszugehen, dass die Unterschiede systematisch sind. Auch bei *Schettose* ist eine Überlappung der 95 %-Konfidenzintervalle festzustellen, jedoch liegen die Mittelwerte jeweils außerhalb der Überlappung. Dennoch ist es möglich, dass die Unterschiede nicht systematisch sind. Interessant ist der Unterschied zwischen *Schettose* und *Truntake*: Während die schwache Form von *Schettose* bei 0,12 liegt und die starke bei 0,3, ist bei *Truntake* sowohl die schwache als auch die starke Form nah bei 0,3.

Die p -Werte sind für dieses Modell nur bedingt interessant, da sie immer im Vergleich zum Referenzlevel *Schettose* mit schwacher Form stehen und daher keine Aussage über die Unterschiede zwischen starken und schwachen Formen der anderen Ausprägungen machen. Die Lesezeiten für die starke Form von *Schettose* liegen im Vergleich zum Referenzlevel unter dem α -Level von 0,01, was auf einen systematischen Effekt für die starke und die schwache Form von *Schettose* hindeutet. Da sich die Konfidenzintervalle jedoch überlappen, ist ein Zufallsbefund dennoch möglich.

Auch *Truntake* mit schwacher Form und *Knatt* mit starker Form unterschreiten das α -Level. Bei *Grettel* mit starker Form wird das α -Level hingegen knapp verfehlt. Die Konfidenzintervalle für die schwache Form von *Schettose* und die starken Formen von *Knatt* und *Grettel* überlappen sich jeweils nicht. Zumindest für *Knatt* kann somit von systematischen Unterschieden ausgegangen werden, bei *Grettel* kann der Unterschied auch zufällig sein. Es ist also möglich, dass die schwache Form von *Schettose* systematisch langsamer gelesen wird als die starke Form von *Knatt*. Dies ließe sich als ein genereller Vorteil der starken Formen bei Substantiven deuten, die die starke Flexion nicht qua Form-Schema komplett ausschließen. Allerdings ist es in diesem Zusammenhang verwunderlich, dass die starke Form von *Grettel* keinen systematischen Unterschied zu der schwachen von *Schettose* aufweist.

Hinsichtlich möglicher Lesezeitunterschiede zwischen den starken und schwachen Formen der Testitems deutet das Modell lediglich für *Schettose* auf einen systematischen Effekt hin. Aufgrund der Überlappung der 95 %-Konfidenzintervalle ist ein Zufallsbefund aber nicht auszuschließen. Wenig überraschend ist die Effektstärke des Modells mit einem R^2m von 0,11 klein, erst unter Einbezug der zufälligen Effekte ist sie mit einem R^2c von 0,41 höher.

Die Lesezeiten scheinen kaum von den Form-Schematizitätsausprägungen beeinflusst worden zu sein. Dies steht diametral zu den Hypothesen (siehe Abschnitt 4.4.1) und dem Antwortverhalten im Produktionsexperiment, das eindeu-

tig von Form-Schematizität beeinflusst ist. Nur bei starken und schwachen Formen von *Schettose* lässt sich ein Effekt erahnen, der jedoch nicht groß ist und aufgrund der Überlappung der 95 %-Konfidenzintervalle auch nicht zwingend systematisch ist. Auch in der Studie von Schmitt (2019a: 169–171) waren Form-Schematizitätseffekte in den Lesezeiten kaum zu greifen: Hier wurde ebenfalls mit den Testsubstantiven *Schettose*, *Knatt* und *Grettel* gearbeitet. Nur für *Grettel* waren Unterschiede in der Lesezeit zwischen starker und schwacher Form festzustellen, für *Schettose* hingegen nicht. Um zuverlässige Aussagen über den Einfluss der Form-Schematizität auf Lesezeiten treffen zu können, müsste eine weitere Studie mit mehr Versuchitems und teststärkenbasierter Stichprobengröße durchgeführt werden. Sollten die Unterschiede zwischen starken und schwachen Formen bei *Schettose* und *Grettel* systematisch sein und der ausbleibende Unterschied bei *Grettel* in dieser Studie einen β -Fehler darstellen, würde dies auf einen Einfluss der Form-Schematizität hinweisen: Die jeweils unwahrscheinlichen Formen würden dann Probleme in der Prozessierung auslösen, die sich in erhöhten Lesezeiten niederschlagen würden. Nur für die Peripherie wäre dies nicht zu beobachten, da hier beide Formen ähnlich wahrscheinlich sind. In diesem Fall wäre ein genauere Blick auf den anscheinend ausbleibenden Effekt bei *Truntake* interessant: Da die schwache Form im Dativ um einen Buchstaben länger ist als die starke, könnte dies die Lesezeit der starken Form verkürzen und so einen möglichen Form-Schematizitätseffekt überdecken.

5.4.2.2 Lexical-decision-Studie

In diesem Abschnitt werden die Reaktionszeiten in der lexical-decision-Studie betrachtet. Diese sind aufgrund der Aufgabenstellung (Bewertung der Bekanntheit statt Lesen) und aufgrund der Nutzung von real existierenden Testsubstantiven anders gelagert als in der self-paced-reading-Studie. Abbildung 5.38 zeigt die logarithmierten Reaktionszeiten in Beeswarmplots für die einzelnen Form-Schematizitätsausprägungen samt arithmetischem Mittel (M) und Standardabweichung (SD), diese werden samt Standardfehler (SE) auch gesondert berichtet. Reaktionszeiten für starke Formen sind als hellgraue Dreiecke dargestellt, Reaktionszeiten für schwache Formen als dunkelgraue Kreise.

Eine deutliche Interaktion zwischen Form-Schematizitätsausprägung und Flexionsform ist zu erkennen: Während Substantive der Ausprägungen FORM-SHEMA und PERIPHERIE hohe Reaktionszeiten für starke Formen aufweisen und niedrige für schwache, ist das Gegenteil für Substantive der Ausprägung STARK der Fall. Die starken Formen sind bei schwachen Maskulina mental nicht gefestigt

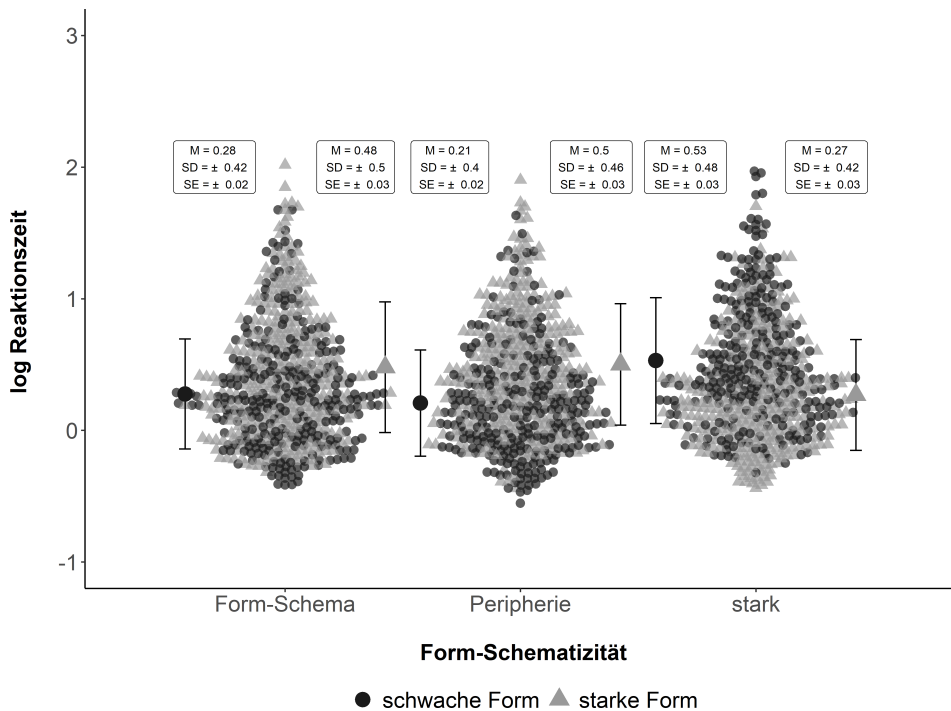


Abbildung 5.38: Reaktionszeiten in der lexical-decision-Studie zu Form-Schematizität

und lösen daher höhere Reaktionszeiten aus als die mental gefestigten schwachen Formen. Dies scheint unabhängig von der Passgenauigkeit der Substantive zum Form-Schema der Fall zu sein, da der Vorteil für schwache Formen sowohl bei prototypischen als auch bei peripheren Vertretern des Form-Schemas zu erkennen ist. Umgekehrt besteht für die starken Substantive ein Vorteil für starke Formen.

Die Reaktionszeiten spiegeln damit das Antwortverhalten. Nur bei Substantiven aus der Peripherie des Form-Schemas ist dies in Hinblick auf starke Formen nicht der Fall: Da die starken Formen bei Substantiven der Ausprägung PERIPHERIE in 50 % der Antworten als bekannt bewertet wurden, scheinen sie stärker gefestigt zu sein als starke Formen bei Substantiven der Ausprägung FORM-SCHEMA, die nur zu 10 % als bekannt bewertet wurden (siehe Abschnitt 5.4.1.2 zum Antwortverhalten in der lexical-decision-Studie). Daher wäre zu erwarten, dass die Reaktionszeiten für starke und schwache Formen bei Substantiven der Ausprägung PERIPHERIE näher beieinander liegen als bei Substantiven der Ausprägung FORM-SCHEMA. Dies ist aber in den Reaktionszeiten nicht zu erkennen.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Abseits der Interaktion ist kein Unterschied in den Reaktionszeiten zwischen den Ausprägungen zu erkennen: Die Verteilungen decken dieselbe Reaktionszeitspanne ab und auch die Mittelwerte ähneln sich stark: Sie liegen zwischen 0,35 und 0,4 mit Standardabweichungen von ca. 0,5. Die Daten wurden auf Längen- und Reihenfolge-Effekte überprüft (siehe Abbildungen I.5 und I.6 in Anhang I.3), die anhand der visuellen Analyse ausgeschlossen werden konnten.

Ein Einfluss der starken und schwachen Formen auf Reaktionszeiten war bereits in der lexical-decision-Studie zu Frequenzeffekten bei starken Verben zu erkennen (siehe Abschnitt 5.2.2). Wie in der Frequenzstudie könnte auch in dieser Studie die Ablehnung der Frage nach der Bekanntheit der Form höhere Reaktionszeiten hervorgerufen haben als die Zustimmung zur Frage. Die starken Formen von Substantiven in der Peripherie des Form-Schemas wurden etwa zur Hälfte als bekannt oder unbekannt angesehen, daher können sie einen Hinweis auf einen möglichen Einfluss der Bewertung geben. Abbildung 5.39 zeigt die Reaktionszeiten für starke Formen von Substantiven der Ausprägung PERIPHERIE nach ihrer Bekanntheit.

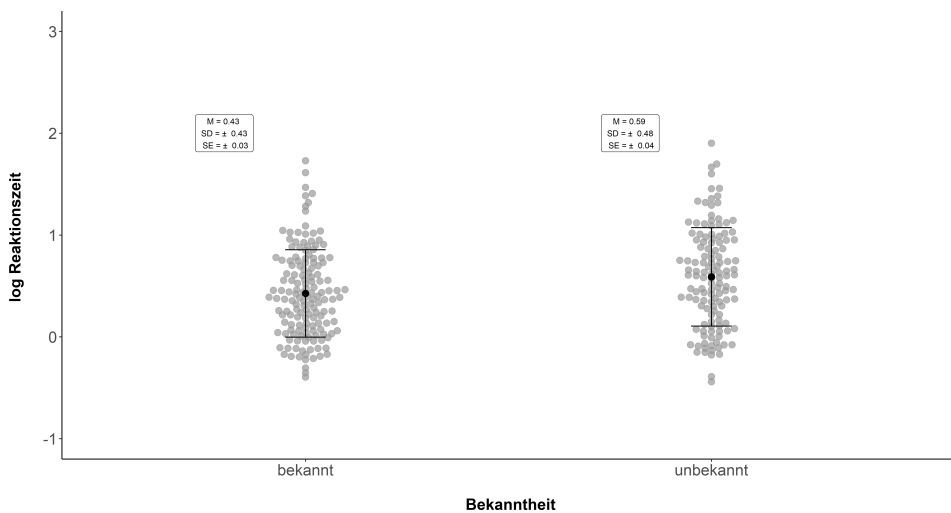


Abbildung 5.39: Reaktionszeiten für starke Formen von Substantiven der Ausprägung PERIPHERIE nach Bekanntheit

Es zeigt sich eine ähnliche Verteilung für beide Bewertungen, allerdings tendieren die Reaktionszeiten bei als bekannt bewerteten Substantiven eher zum Mittelwert und streuen weniger. Daher ergeben sich auch leichte Unterschiede im Mittelwert: Der Mittelwert für die Antworten, die der Bekanntheit der Form zustimmten, liegt mit 0,43 und einer Standardabweichung von 0,3 unter dem Mittelwert für ablehnende Antworten (0,59; Standardabweichung 0,48). Ein Einfluss

der Bewertung als bekannt oder unbekannt auf Reaktionszeiten ist daher wie in der Frequenzstudie nicht auszuschließen.

Um die Interaktion zwischen Form-Schematizitätsausprägung und Flexionsform näher betrachten zu können, werden Scatterplots in den Blick genommen. Abbildung 5.40 kontrastiert die logarithmierten Reaktionszeiten für starke und schwache Formen von Substantiven der einzelnen Form-Schematizitätsausprägungen. Ein Punkt steht für die Reaktionszeiten einer Person für ein Testitem. Die graue Linie dient als Orientierung: Sie zeigt den Bereich an, in dem starke und schwache Formen gleich schnell bewertet wurden.

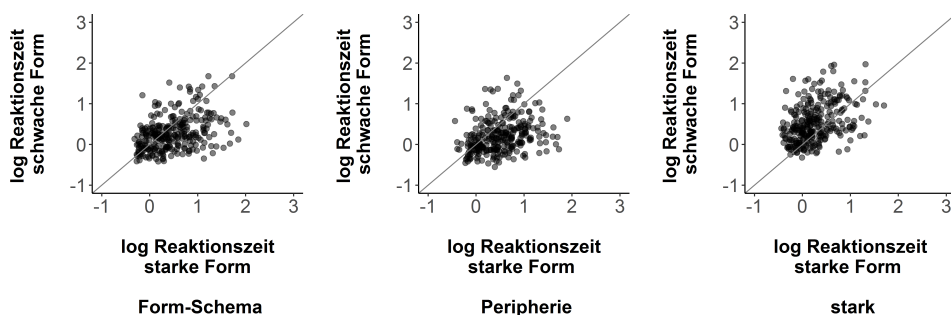


Abbildung 5.40: Scatterplots der Reaktionszeiten für starke und schwache Formen der Testsubstantive

Die Interaktion ist auch in dieser Darstellung deutlich zu erkennen: Während die Ausprägungen FORM-SCHEMA und PERIPHERIE vornehmlich bei starken Formen höhere Reaktionszeiten hervorrufen, was an den Punkten unterhalb der grauen Linie zu erkennen ist, evoziert die Ausprägung STARK höhere Reaktionszeiten für schwache Formen, was anhand der Punkte oberhalb der grauen Linie deutlich wird.

Im Versuchsdesign wurde weitgehend auf Tokenfrequenz kontrolliert (siehe Abschnitt 4.5.2). Da die Testsubstantive in ihrer Tokenfrequenz aber dennoch leicht variieren, wird überprüft, ob Tokenfrequenz einen Einfluss auf die Reaktionszeiten nimmt. Hierfür wurden die Testsubstantive anhand ihrer Tokenfrequenz in drei Gruppen geteilt: FREQUENT für Substantive mit einer relativen Tokenfrequenz von über 0,47 Belegen pro Million Token (*Franzose, Nachbar, Kollege, Held, Zar*), MITTELFREQUENT für Substantive mit einer Tokenfrequenz von 0,28 bis 0,47 (*Feind, Schütze, Graf, Freund, Fürst*) und INFREQUENT für Substantive mit einer Tokenfrequenz von unter 0,05 (*Neffe, Dieb, Geselle, Kerl, Vogt*).³¹ Die drei Frequenzausprägungen wurden genutzt, um die Vergleichbarkeit mit den

³¹Dieses Vorgehen wurde vorab registriert, siehe <https://osf.io/yx3tv>.

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Form-Schematizitätsausprägungen aufrecht zu erhalten, daher wurden auch pro Frequenzausprägung fünf Substantive genutzt. Die Frequenzunterschiede zwischen den Frequenzausprägungen ergeben sich daher allein durch die Anordnung der Testsubstantive nach Frequenz und durch die Einteilung in fünf Substantive pro Frequenzausprägung. Da die Testsubstantive nicht auf Unterschiede in ihrer Tokenfrequenz hin ausgewählt wurden, sind die Abstufungen zwischen den Frequenzausprägungen z. T. subtil. Dennoch fällt der Frequenzunterschied zwischen FREQUENT (ab 0,47) und INFREQUENT (unter 0,05) recht deutlich aus.

Abbildung 5.41 zeigt Beeswarmplots für die Reaktionszeiten nach Frequenzausprägung samt arithmetischem Mittel (M) und Standardabweichung (SD), zusätzlich wird der Standardfehler (SE) berichtet.

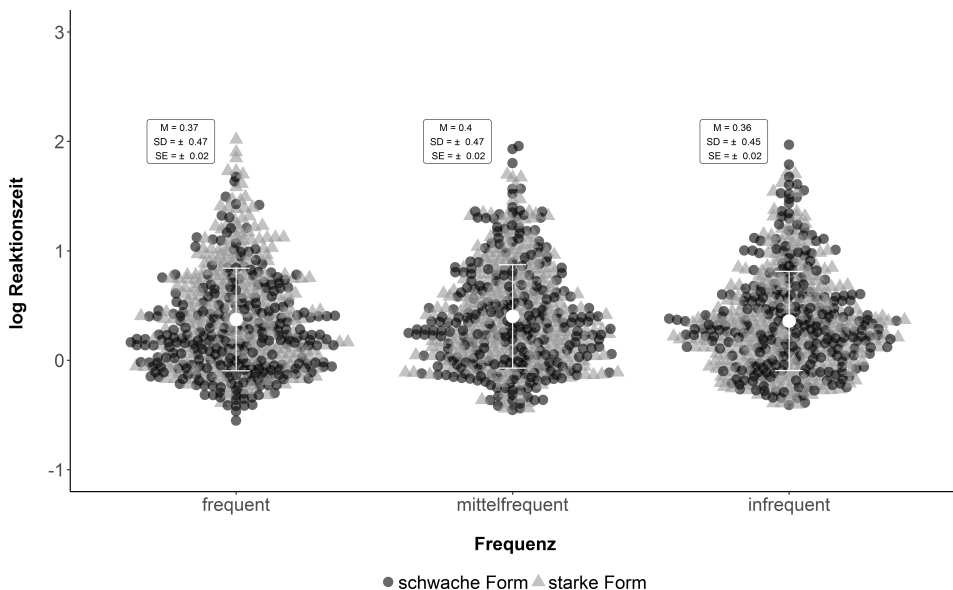


Abbildung 5.41: Reaktionszeiten nach Frequenzausprägung in der lexical-decision-Studie zu Form-Schematizität

Der Einfluss der starken und schwachen Formen deutet sich nur noch in der Ausprägung FREQUENT an: Die schwachen Formen haben eher niedrige Reaktionszeiten hervorgerufen, die starken eher hohe. Bei den anderen Ausprägungen lässt sich kein Unterschied zwischen starken und schwachen Formen erkennen. Die Ausprägung FREQUENT ist die einzige, die nur schwache Maskulina enthält, während die anderen beiden Ausprägungen durchmischte sind. Ein Frequenzefekt ist nicht zu erkennen: Die Verteilung der Reaktionszeiten hat sich nicht geändert, die Mittelwerte liegen bei ca. 0,4 mit Standardabweichungen von 0,5.

Für die konfirmatorische statistische Analyse der Daten wird ein gemischtes lineares Modell mit Form-Schematizitätsausprägung und Flexionsform als feste Effekte genutzt, die interagieren können: Reaktionszeit ~ Form-Schematizität * Flexion + (1|Proband_in) + (1|Item). Für Teilnehmer_innen und Versuchitems (Lemma) werden *random intercepts* berücksichtigt. *Random slopes* werden für Teilnehmer_innen wegen *singular fits* nicht genutzt. Bei den Versuchitems sind keine *random slopes* möglich, da sie über die Form-Schematizitätsausprägungen hinweg variieren. Die Standardabweichung um den y-Achsenabschnitt ist für Versuchitems mit 0,05 sehr gering, Proband_innen weisen mit 0,27 erwartbarerweise eine höhere Standardabweichung auf. Die restliche Residuenvarianz beträgt 0,35. Die geschätzten Werte des Modells sind in Tabelle 5.17, eine entsprechende Kreuztabelle (Tabelle I.4) befindet sich in Anhang I.3. Abbildung 5.42 zeigt die vom Modell vorhergesagten Reaktionszeiten. Kreise stehen für schwache Formen, Dreiecke für starke.

Tabelle 5.17: Werte des Modells für die Reaktionszeiten in der lexical-decision-Studie zu Form-Schematizität

	Wert	SE	<i>t</i>	df	<i>p</i>
(Intercept = Form-Schema & schwach)	0,28	0,05	5,87	71,43	<0,01
Form-Schematizität: Peripherie	-0,07	0,04	-1,59	24,64	0,12
Form-Schematizität: stark	0,25	0,04	5,84	24,64	<0,01
Flexion: stark	0,20	0,03	6,78	1609,90	<0,01
Form-Schematizität: Peripherie & Flexion: stark	0,09	0,04	2,14	1609,90	0,03
Form-Schematizität: stark & Flexion: stark	-0,46	0,04	-10,94	1609,90	<0,01

Auch bei den vorhergesagten Reaktionszeiten ist die Interaktion zwischen Form-Schematizitätsausprägung und starken und schwachen Flexionsformen klar zu erkennen. Die 95 %-Konfidenzintervalle der Ausprägungen überlappen sich jeweils hinsichtlich der schneller und langsamer bewerteten Formen, sodass abseits der Interaktion kein Einfluss der Form-Schematizitätsausprägung auf Reaktionszeiten zu erkennen ist. Es scheint somit vorrangig relevant zu sein, welche Formen beurteilt werden: Die mental gefestigten Formen werden schneller bewertet als die mental weniger bzw. nicht gefestigten. Starke Formen der Ausprägung FORM-SCHEMA liegen im Vergleich zum Referenzlevel FORM-SCHEMA mit schwachen Formen unter dem α -Level von 0,01. Dasselbe gilt für die Ausprägung STARK mit starken und schwachen Formen. Hier ist also jeweils von systematischen Effekten auszugehen. Der Unterschied zur Ausprägung PERIPHERIE ist hingegen über α -Level. Somit ist nicht von einem Unterschied in den Reaktionszeiten zwischen FORM-SCHEMA und PERIPHERIE auszugehen, was bereits in der

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

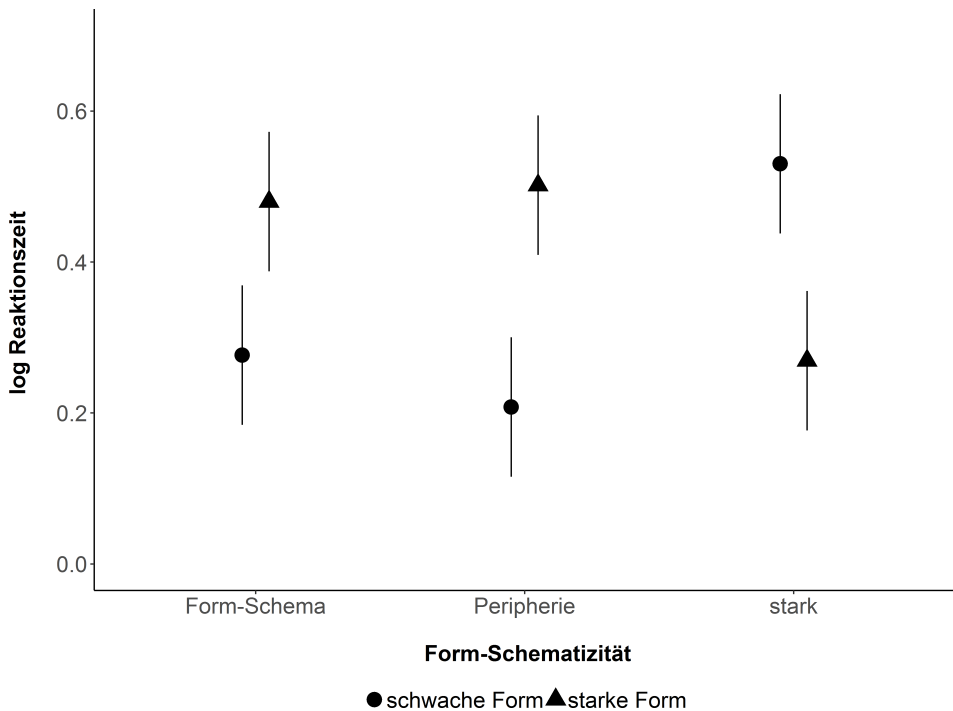


Abbildung 5.42: Vorhergesagte Reaktionszeiten in der lexical-decision-Studie zu Form-Schematizität

visuellen Analyse sichtbar war. Die Effektstärke des Modells ist mit einem R^2m von 0,08 sehr klein, mit den zufälligen Effekten ist sie mit einem R^2c von 0,4 deutlich höher.

Für die explorative Analyse der Daten werden die Frequenzausprägungen der Substantive in einem gemischten linearen Modell betrachtet. Auch hier ist zusätzlich die Flexionsform als fester Effekt vorgesehen, die beiden Effekte können interagieren: Reaktionszeit \sim Frequenz * Flexion + (1|Proband_in) + (1|Item). Für die zufälligen Effekte Proband_in und Testitem sind *random intercepts* vorgesehen, *random slopes* werden wegen *singular fits* nicht genutzt bzw. sind für die Testitems nicht möglich, da die Testitems über die Frequenzausprägungen hinweg variieren. Die zufälligen Effekte ähneln stark dem oben diskutierten Modell zum Einfluss der Form-Schematizitätsausprägungen: Items variieren mit 0,05 und Proband_innen mit 0,27 um den y-Achsenabschnitt. Die restliche Residuenvarianz beträgt 0,37. Die Ergebnisse des Modells sind in Tabelle 5.18, die entsprechende Kreuztabelle (Tabelle I.5) befindet sich in Anhang I.3. Abbildung 5.43 zeigt die vorhergesagten Reaktionszeiten.

Tabelle 5.18: Werte des Modells für die Reaktionszeiten in der lexical-decision-Studie zu Form-Schematizität in Abhängigkeit von Frequenz

	Wert	SE	<i>t</i>	df
(Intercept = frequent & schwach)	0,26	0,05	5,42	72,08
Frequenz: infrequent	0,12	0,04	2,79	25,66
Frequenz: mittelfrequent	0,12	0,04	2,77	25,66
Flexion: stark	0,23	0,03	7,40	1609,89
Frequenz: infrequent & Flexion: stark	-0,27	0,04	-6,16	1609,89
Frequenz: mittelfrequent & Flexion: stark	-0,19	0,04	-4,21	1609,89

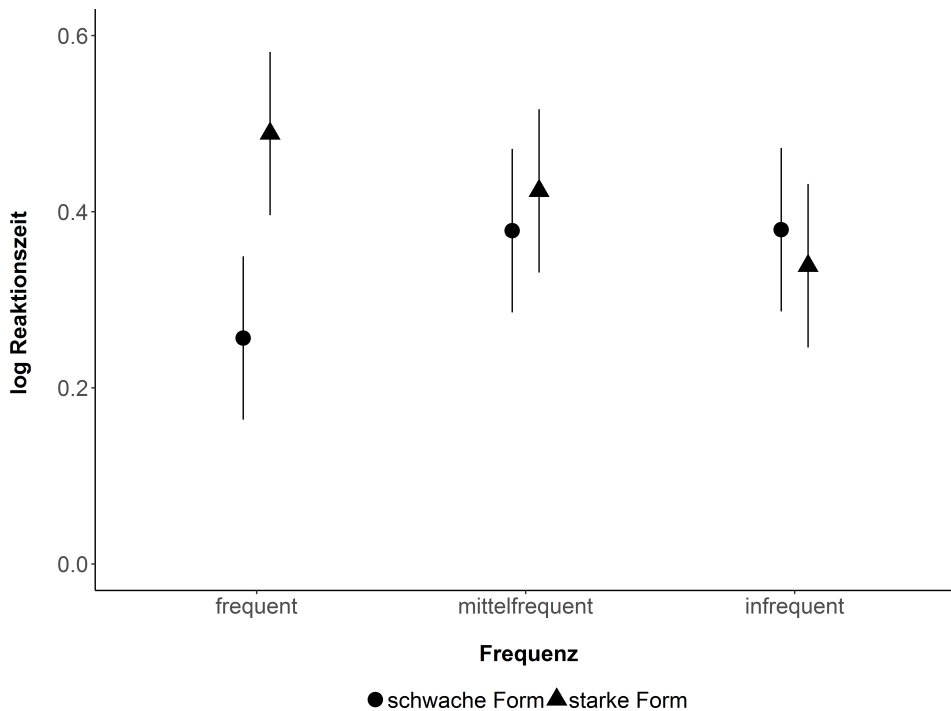


Abbildung 5.43: Vorhergesagte Reaktionszeiten in der lexical-decision-Studie zu Form-Schematizität in Abhängigkeit von Frequenz

Der Einfluss von starken und schwachen Formen ist wieder zu erkennen: Da in der Ausprägung *FREQUENT* nur schwache Maskulina enthalten sind, werden starke Formen langsamer bewertet als schwache. Bei *MITTELFREQUENT* und *INFREQUENT* ist dieser Effekt nicht zu sehen, da hier Substantive beider Klassen vorkommen. Ein Frequenzeffekt lässt sich ebenfalls nicht erkennen. Zwar weisen die Substantive der Ausprägung *FREQUENT* für schwache Formen niedrigere Werte auf als die anderen Ausprägungen, dies lässt sich aber dadurch erklären, dass die anderen Ausprägungen auch starke Maskulina enthalten und die schwachen Formen daher nicht so stark gefestigt sind. Diese Erklärung wird von den starken Formen der Ausprägung *FREQUENT* gestützt, die die höchsten Reaktionszeiten evozieren: Da in der Ausprägung *FREQUENT* nur schwache Maskulina enthalten sind, sind hier die starken Formen nicht gefestigt.

Die 95 %-Konfidenzintervalle sind für fast alle Werte recht lang, nur für die starken und schwachen Formen der Ausprägung *FREQUENT* zeigen sich keine Überlappungen. Ein systematischer Einfluss der Frequenzausprägungen scheint somit nicht gegeben zu sein. Die Effektstärke ist mit einem R^2m von 0,02 und einem R^2c von 0,36 deutlich unterhalb der Effektstärke des Modells zum Einfluss der Form-Schematizitätsausprägungen.

Um die Reaktionszeiten innerhalb der lexical-decision-Studie zu Form-Schematizität besser einordnen zu können, ist ein Vergleich der Reaktionszeiten mit den Reaktionszeiten in der Studie zu Frequenzeffekten bei starken Verben (siehe Abschnitt 5.2) hilfreich, da in beiden Fällen zwischen zwei Flexionsformen gewählt wurde. Dabei ist zu bedenken, dass starke und schwache Formen von Verben und Substantiven nicht direkt vergleichbar sind. Aber dennoch ist eine Gegenüberstellung interessant, da auf diese Weise Unterschiede in den Reaktionszeitmustern zwischen der Frequenz- und der Form-Schematizitätsstudie herausgearbeitet werden können. Für den Vergleich wurden die Frequenz- und Form-Schematizitätsausprägungen zusammengefasst: *FREQUENT/FORM-SCHEMA*, *INFREQUENT OHNE SCHWANKUNG/PERIPHERIE* und *INFREQUENT MIT SCHWANKUNG/STARK* sind die Kategorien, die miteinander verglichen werden. Abbildung 5.44 zeigt die Mittelwerte für die logarithmierten Reaktionszeiten samt Standardabweichung. Auf BeeswarmpLOTS wurde aufgrund der Komplexität der Daten verzichtet. Substantive sind als Kreise dargestellt, Verben als Dreiecke. Reaktionszeiten für schwache Formen sind dunkelgrau eingefärbt, Reaktionszeiten für starke Formen hellgrau.

Während die Mittelwerte der Reaktionszeiten bei den Substantiven auf einem Niveau bleiben und man lediglich die Interaktion zwischen Flexionsklasse und Flexionsform beobachten kann, sind die Mittelwerte der Reaktionszeiten bei den Verben zusätzlich zum Einfluss der Flexionsform frequenzabhängig gestaffelt:

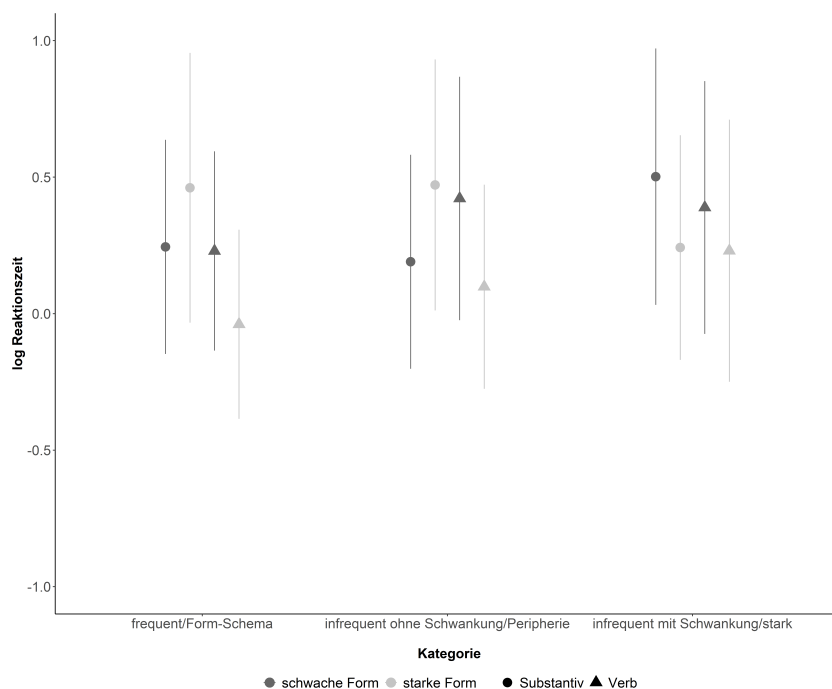


Abbildung 5.44: Vergleich der Reaktionszeiten für Substantive und Verben in den Studien zu Form-Schematizität und Frequenz

Die Reaktionszeiten für frequente Verben sind geringer als für infrequente Verben mit und ohne Schwankung.

Innerhalb der Kategorieausprägungen weisen die Reaktionszeiten für starke und schwache Flexionsformen interessante Unterschiede zwischen Verben und Substantiven auf: Der Mittelwert der Reaktionszeiten für starke Formen von Verben der Ausprägung *FREQUENT* ist deutlich niedriger als der Mittelwert der Reaktionszeiten für starke und schwache Formen von Substantiven der Ausprägung *FORM-SCHEMA*. Der Mittelwert für schwache Formen von frequenten Verben gleicht hingegen dem Mittelwert der schwachen Formen von Substantiven der Ausprägung *FORM-SCHEMA*. Die Unterschiede in den Mittelwerten verdeutlichen die Tokenfrequenzunterschiede zwischen den Testverben und -substantiven der Ausprägung *FREQUENT/FORM-SCHEMA*, die zu einem unterschiedlichen Grad an *entrenchment* führen: Die schwachen Substantivformen sind aufgrund der geringeren Tokenfrequenz weniger stark gefestigt als die starken Verbformen. Daher lösen die starken Verbformen geringere Reaktionszeiten aus als die

Substantivformen. Die schwachen Verbformen sind selbst nicht gefestigt und können daher nur schnell abgelehnt werden, da sie von den frequenten starken Verbformen statistisch ausgestochen werden. Deshalb sind die Reaktionszeiten der schwachen Verbformen im Vergleich zu den starken Verbformen leicht erhöht und auf dem Niveau der schwachen Substantivformen.

Bei den Verben der Ausprägung *INFREQUENT OHNE SCHWANKUNG* verschieben sich die Mittelwerte der Reaktionszeiten im Vergleich zu den frequenten Verben deutlich nach oben: Nun sind die starken Verbformen auf einem Niveau mit den schwachen Substantivformen der Ausprägung *PERIPHERIE* und die schwachen Verbformen auf einem Niveau mit den starken Substantivformen derselben Ausprägung. Ähnlich verhalten sich die Verben der Ausprägung *INFREQUENT MIT SCHWANKUNG*: Auch hier ist die jeweils schneller beurteilte Verbform auf einem Niveau mit der schneller beurteilten Substantivform. Bei den infrequenten Verben sind die Reaktionszeiten der starken Formen somit unabhängig von Schwankungen in Korpora jeweils vergleichbar mit den Reaktionszeiten der mental gefestigteren Form der Testsubstantive (schwache Formen bei der Ausprägung *PERIPHERIE*, starke Formen bei der Ausprägung *STARK*). Vergleichbare Reaktionszeiten für die jeweils mental gefestigten bzw. nicht gefestigten Formen sind erwartbar, da die infrequenten Testverben wie die Testsubstantive eine Tokenfrequenz von unter zwei Belegen pro Million Token aufweisen.

Um den Vergleich zwischen Frequenz- und Form-Schematizitätseffekten statistisch zu fassen, wird als explorative Analyse ein gemischtes lineares Modell gerechnet, das Wortart, Kategorie (*FREQUENT/FORM-SCHEMA*, *INFREQUENT OHNE SCHWANKUNG/PERIPHERIE* und *INFREQUENT MIT SCHWANKUNG/STARK*) und Flexionsform beinhaltet, die festen Effekte können interagieren: Reaktionszeit ~ Wortart * Kategorie * Flexion + (1|Proband_in) + (1|Item). Als zufällige Effekte werden *random intercepts* für Teilnehmer_innen und Versuchitems genutzt. Die zufälligen Effekte sind vergleichbar mit den vorherigen Modellen. Die Ergebnisse des Modells sind in Tabelle 5.19 aufgeführt, hierfür wurde *INFREQUENT OHNE SCHWANKUNG* durch *ohne* und *INFREQUENT MIT SCHWANKUNG* durch *mit* abgekürzt. Die darauf aufbauende Kreuztabelle (Tabelle I.7) befindet sich in Anhang I.4. Abbildung 5.45 (S. 318) zeigt die vorhergesagten Reaktionszeiten. Sie sind für schwache Formen dunkelgrau und für starke hellgrau eingefärbt. Kreise stehen für Substantive, Dreiecke für Verben.

Die vorhergesagten Reaktionszeiten unterscheiden sich sowohl für Substantive als auch für Verben deutlich je nach Flexionsform. Die schwachen Formen weisen für Substantive der Ausprägungen *FORM-SCHEMA* und *PERIPHERIE* geringere Reaktionszeiten auf als die starken, das Gegenteil ist für Substantive der Ausprägung *STARK* der Fall. Die Verben zeigen unabhängig von der Frequenzausprägung

Tabelle 5.19: Werte des Modells für die Reaktionszeiten bei Substantiven und Verben in den Studien zu Form-Schematizität und Frequenz

	Wert	SE	<i>t</i>	df
(Intercept = frequent/Form-Schema & Substantiv & schwach)	0,24	0,05	4,99	97,47
Kategorie: mit/stark	0,26	0,04	5,79	65,85
Kategorie: ohne/Peripherie	-0,05	0,04	-1,23	65,85
Wortart: Verb	-0,01	0,04	-0,37	65,85
Flexion: stark	0,22	0,03	7,04	3580,96
Kategorie: mit/stark & Wortart: Verb	-0,10	0,06	-1,73	65,85
Kategorie: ohne/Peripherie & Wortart: Verb	0,25	0,06	4,36	65,85
Kategorie: mit/stark & Flexion: stark	-0,48	0,04	-10,95	3580,96
Kategorie: ohne/Peripherie & Flexion: stark	0,06	0,04	1,49	3580,96
Wortart: Verb & Flexion: stark	-0,49	0,04	-12,38	3580,96
Kategorie: mit/stark & Wortart: Verb & Flexion: stark	0,59	0,06	10,58	3580,96
Kategorie: ohne/Peripherie & Wortart: Verb & Flexion: stark	-0,12	0,06	-2,16	3580,96

niedrigere Reaktionszeiten für starke Formen. Die 95 %-Konfidenzintervalle überlappen sich in dieser Hinsicht nur bei starken und schwachen Formen der infrequenten Verben mit Schwankung. Bei Verben ist zusätzlich zum Einfluss der Flexionsformen ein Frequenzeffekt zu sehen: Die Ausprägung *FREQUENT* weist geringere Reaktionszeiten auf als die anderen Frequenzausprägungen. Eine Überlappung des 95 %-Konfidenzintervalls der Ausprägung *FREQUENT* mit den 95 %-Konfidenzintervallen der anderen Ausprägung ist dabei nicht zu erkennen. Für die Form-Schematizitätsausprägungen lässt sich hingegen kein systematischer Unterschied in den Reaktionszeiten beobachten. Die Effektstärke ist mit einem R^2m von 0,13 und einem R^2c von 0,46 vergleichbar mit der Effektstärke des Modells der Frequenzstudie.

Insgesamt zeigt die Betrachtung der Reaktionszeiten in der lexical-decision-Studie vorrangig Effekte der schwachen und starken Formen: Je nach Flexionsklasse der Substantive werden schwache Formen schneller oder langsamer beurteilt als starke, da sie je nach Flexionsklasse stärker bzw. schwächer mental

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

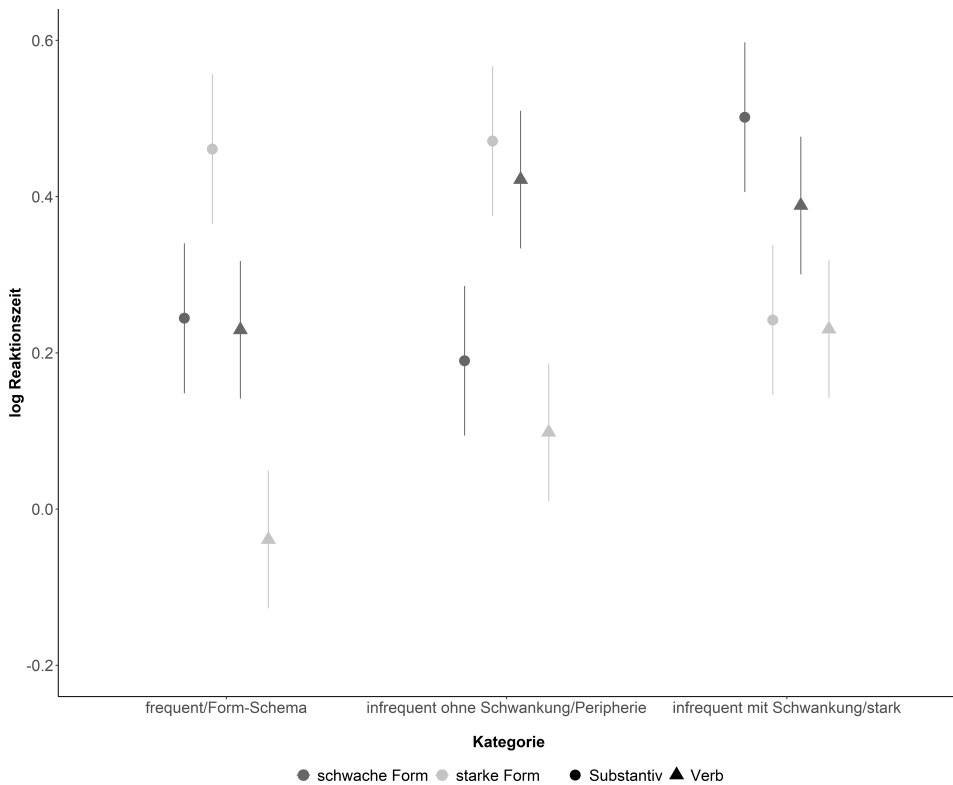


Abbildung 5.45: Vorhergesagte Reaktionszeiten für Substantive und Verben in den Studien zu Form-Schematizität und Frequenz

gefestigt sind. Zudem könnte die Beurteilung der Formen (Bekanntheit/Unbekanntheit) die Reaktionszeiten beeinflussen. Ein Effekt der Form-Schematizität lässt sich hingegen nicht erkennen. Die Substantive der einzelnen Form-Schematizitätsausprägungen evozieren ähnliche Reaktionszeiten, auch Tokenfrequenz scheint keinen Einfluss auf die Reaktionszeiten zu haben. Anders als in der Frequenzstudie, in der neben dem Einfluss der Flexionsformen auch Frequenzeffekte beobachtet wurden, scheint in der lexical-decision-Studie zu Form-Schematizität somit nur die Flexionsform die Reaktionszeiten zu beeinflussen. Die Ergebnisse bestätigen damit die in Abschnitt 4.5.1 vorgestellten Hypothesen, die anhand der Ergebnisse der Prästudie aufgestellt wurden.

Ursprünglich wurde erwartet, dass Substantive, die dem Form-Schema nur peripher angehören, für starke und schwache Formen höhere Reaktionszeiten evozieren, da die Wahl zwischen zwei mental gefestigten Formen besteht. Dabei weist auch das Antwortverhalten in der Studie darauf hin, dass in der Peripherie des Form-Schemas beide Formen mental gefestigt sind: Schwache Formen

von Substantiven der Ausprägung PERIPHERIE wurden zu 90 % und starke Formen zu 50 % als bekannt bewertet. In den Reaktionszeiten zeigt sich jedoch kein Unterschied zwischen Peripherie und den anderen Ausprägungen. Starke und schwache Formen von Substantiven der Ausprägung PERIPHERIE werden jeweils ähnlich schnell bewertet wie starke und schwache Formen von Substantiven der Ausprägung FORM-SCHEMA. Die Existenz einer alternativen Form scheint die Reaktionszeiten in der lexical-decision-Studie daher nicht zu beeinflussen.

Neben generell erhöhten Reaktionszeiten bei Substantiven in der Peripherie im Vergleich zu den anderen Form-Schematizitätsausprägungen wäre auch eine Annäherung der Reaktionszeiten von starken und schwachen Formen bei Substantiven der Peripherie denkbar. Beide Formen sind mental gefestigt und sollten daher ähnliche Reaktionszeiten hervorrufen. Genau das ist bei den infrequenten Verben mit Schwankung, deren neue schwache Formen in der Frequenzstudie ebenfalls zu 50 % als bekannt bewertet wurden, zu beobachten, aber nicht bei den Substantiven in der Peripherie des Form-Schemas in der lexical-decision-Studie: Bei Substantiven in der Peripherie des Form-Schemas weisen die starken Formen höhere Reaktionszeiten auf als schwache Formen. Daher ist davon auszugehen, dass die starken Formen in der Peripherie des Form-Schemas im Vergleich zu schwachen nur wenig gefestigt sind, sodass sie von den schwachen Formen dominiert werden.

5.4.2.3 Sentence-maze-Studie

In diesem Abschnitt werden die Reaktionszeiten in der sentence-maze-Studie analysiert. Diese sind anders strukturiert als die Reaktionszeiten in der lexical-decision-Studie, da hier zwischen zwei Formen gewählt wurde und nicht beide Formen einzeln auf ihre Bekanntheit hin bewertet wurden.

Abbildung 5.46 zeigt die logarithmierten Reaktionszeiten für die Form-Schematizitätsausprägungen in Beeswarmplots samt arithmetischem Mittel (M) und Standardabweichung (SD). Diese werden samt Standardfehler (SE) auch gesondert berichtet.

Die Verteilungen der Reaktionszeiten sehen pro Form-Schematizitätsausprägung unterschiedlich aus: Während bei der Ausprägung FORM-SCHEMA viele Datenpunkte um den Mittelwert herum clustern, finden sich bei den anderen Ausprägungen viele Datenpunkte unterhalb des Mittelwerts. Auch in den Mittelwerten sind Unterschiede zu erkennen: Die Ausprägung FORM-SCHEMA weist mit 0,58 bei einer Standardabweichung von 0,42 den höchsten Mittelwert auf, die Ausprägung STARK liegt knapp darunter mit einem Mittelwert von 0,53 und einer Standardabweichung von 0,4. Geringere Reaktionszeiten weisen die Testsubstantive der Ausprägung PERIPHERIE mit einem Mittelwert von 0,4 und ei-

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

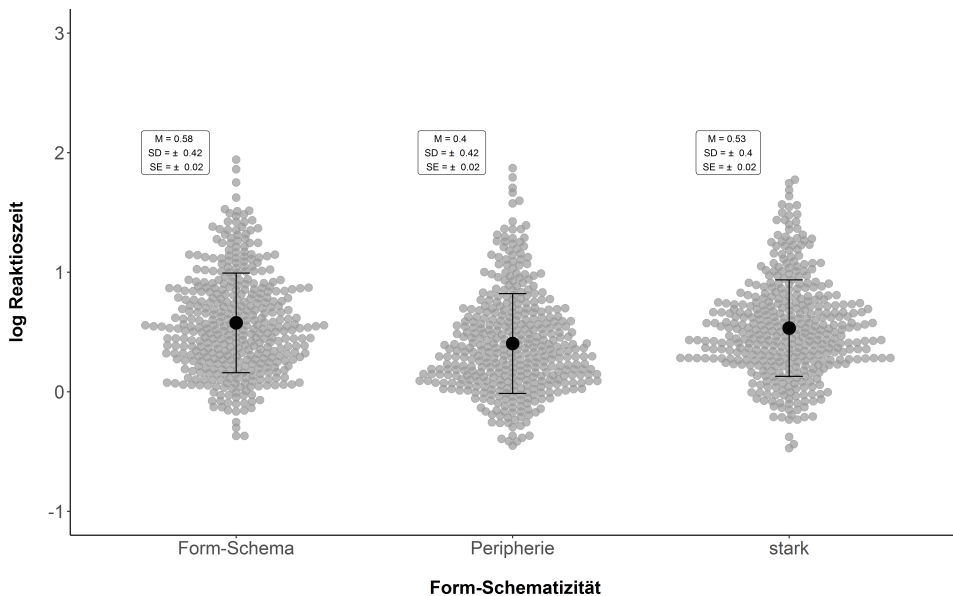


Abbildung 5.46: Reaktionszeiten in der sentence-maze-Studie zu Form-Schematizität

ner Standardabweichung von 0,42 auf. Ein Effekt durch Form-Schematizität ist somit nicht zu erkennen, da die Testsubstantive der Peripherie geringere statt höhere Reaktionszeiten aufweisen als die Substantive der anderen Form-Schematizitätsausprägungen. Höhere Reaktionszeiten wären bei einem Form-Schematizitätseffekt für die Peripherie zu erwarten, da in der Peripherie beide Formen, die zur Wahl stehen, mental gefestigt sind und nicht nur jeweils eine wie bei den anderen Form-Schematizitätsausprägungen. Die visuelle Überprüfung auf Reihenfolge-Effekte ergab keine Auffälligkeiten (siehe Abbildung I.9 in Anhang I.3). Auf Längeneffekte muss nicht geprüft werden, da beide Stimuli jeweils gleichzeitig sichtbar sind.

Um zu überprüfen, ob Tokenfrequenz Einfluss auf die Reaktionszeiten nimmt, wurden die Testsubstantive in verschiedene Frequenzklassen geteilt: Dabei wird wie in der lexical-decision-Studie zu Form-Schematizität zwischen FREQUENT (ab 0,48 Belegen pro Million Token; *Kollege, Held, Zar*), MITTELFREQUENT (von 0,3 bis 0,41; *Schütze, Graf, Freund*) und INFREQUENT (bis 0,05; *Neffe, Dieb, Vogt*) unterschieden.³² Wie in der lexical-decision-Studie wurden auch in dieser Studie drei Frequenzausprägungen gewählt, um die Vergleichbarkeit mit den Form-Schematizitätsausprägungen sicherzustellen. Abbildung 5.47 zeigt die Beeswarmplots mit den Frequenzausprägungen.

³²Die Aufteilung in Frequenzklassen wurde vorab registriert, siehe <https://osf.io/bhxpz>.

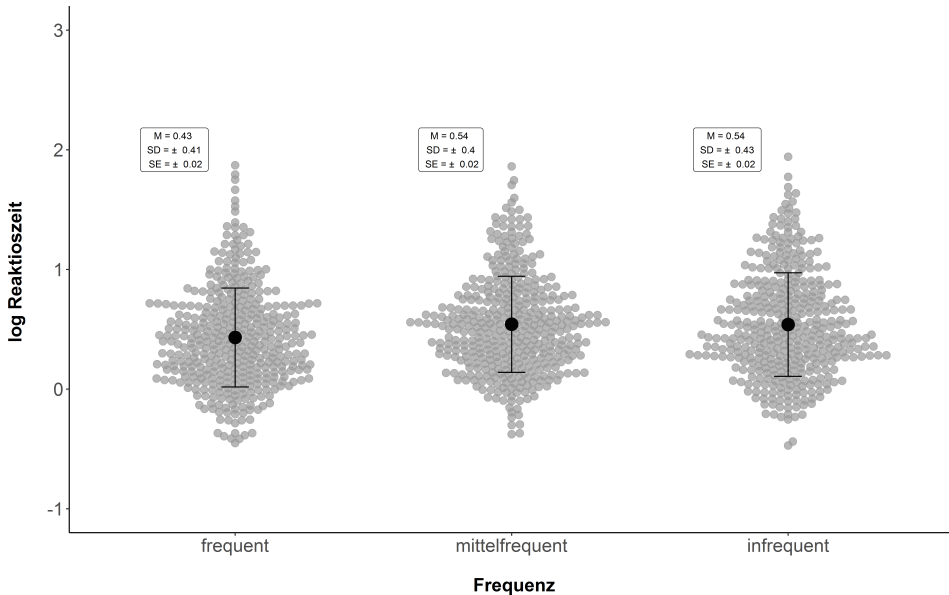


Abbildung 5.47: Reaktionszeiten nach Frequenzausprägung in der sentence-maze-Studie zu Form-Schematizität

Die Reaktionszeiten der Ausprägung `FREQUENT` scheinen um niedrigere Reaktionszeiten zu clustern als die der anderen Ausprägungen. Die Ausprägungen `MITTELFREQUENT` und `INFREQUENT` weisen dagegen eine ähnliche Verteilung der Reaktionszeiten auf. Die Spanne der Reaktionszeiten ist für alle Ausprägungen vergleichbar (`FREQUENT`: -0,55 bis 2,02; `MITTELFREQUENT`: -0,47 bis 1,96; `INFREQUENT`: -0,41 bis 1,97). In den Mittelwerten zeigen sich Unterschiede zwischen den Ausprägungen: Die Reaktionszeiten der Ausprägung `FREQUENT` haben einen Mittelwert von 0,4 mit einer Standardabweichung von 0,4. Die Mittelwerte der anderen Ausprägungen liegen mit 0,5 und Standardabweichungen von 0,4 etwas höher. Es ist somit ein kleiner Frequenzeffekt zu erkennen.

Für die konfirmatorische statistische Analyse wird ein gemischtes lineares Modell mit Form-Schematizität als festem Effekt gerechnet: $\text{Reaktionszeit} \sim \text{Form-Schematizität} + (1|\text{Proband_in}) + (1|\text{Item})$. Teilnehmer_innen und Versuchitems werden mit *random intercepts* berücksichtigt. Für Teilnehmer_innen werden wegen *singular fits* keine *random slopes* genutzt. Für Versuchitems sind *random slopes* nicht sinnvoll, da unterschiedliche Versuchitems pro Form-Schematizitätsausprägung genutzt werden. Die Versuchitems variieren lediglich mit einer Standardabweichung von 0,09 um den y-Achsenabschnitt, die Proband_innen deutlich stärker mit 0,26. Die restliche Residuenvarianz liegt bei 0,31. Die geschätzten

5 Psycholinguistische Betrachtung der Einflussfaktoren: Ergebnisse

Werte des Modells sind in Tabelle 5.20 aufgelistet, Abbildung 5.48 zeigt die vorhergesagten Reaktionszeiten.

Tabelle 5.20: Werte des Modells für die Reaktionszeiten in der sentence-maze-Studie zu Form-Schematizität

	Wert	SE	<i>t</i>	df	<i>p</i>
(Intercept = Form-Schema)	0,58	0,06	9,87	12,14	<0,01
Form-Schematizität: Peripherie	-0,17	0,08	-2,26	8,88	0,05
Form-Schematizität: stark	-0,04	0,08	-0,58	8,88	0,58

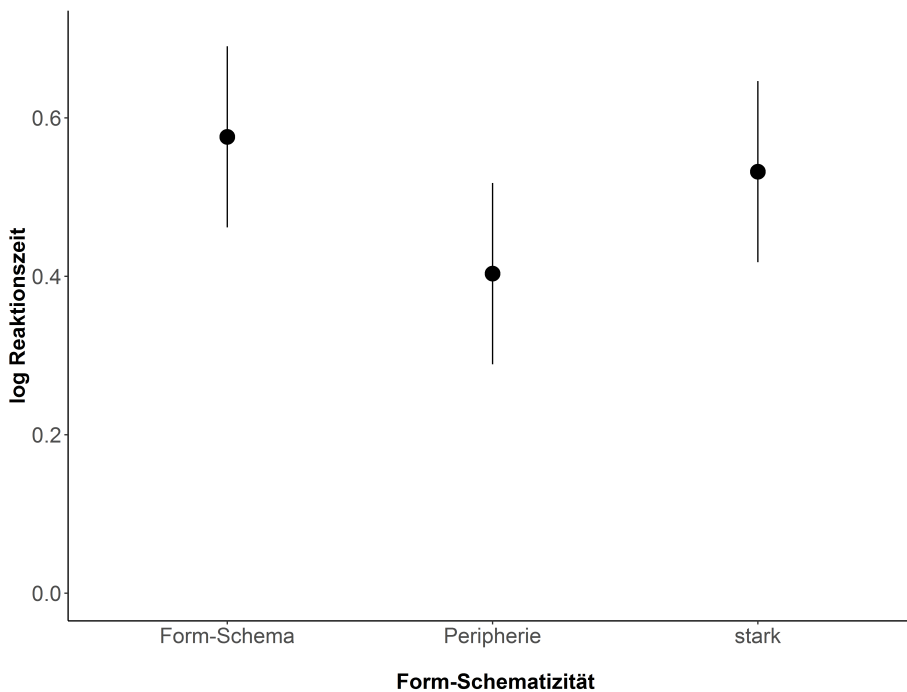


Abbildung 5.48: Vorhergesagte Reaktionszeiten in der sentence-maze-Studie zu Form-Schematizität

Für die Ausprägungen FORM-SCHEMA und STARK werden mit 0,58 und 0,53 vergleichbare Reaktionszeiten vorhergesagt. Die vorhergesagten Reaktionszeiten für die Ausprägung PERIPHERIE sind mit 0,4 geringer. Die 95 %-Konfidenzintervalle der Ausprägungen sind jedoch recht lang und überlappen sich deutlich. Die p -Werte für die Unterschiede zwischen den Ausprägungen sind jeweils über dem α -Level von 0,01, sodass die Daten unter der Nullhypothese (die Form-Schematizitätsausprägungen beeinflussen die Reaktionszeiten nicht) nicht überraschend wären. Es ist daher nicht von systematischen Effekten auszugehen. Die Effektstärke des Modells ist mit einem R^2m von 0,03 dementsprechend gering, bezieht man die zufälligen Effekte mit ein, ist sie mit einem R^2c von 0,45 deutlich höher.

Die Schätzwerte für die Reaktionszeiten aus der Prästudie, anhand derer die benötigte Stichprobengröße für die Hauptstudie berechnet wurde, sind mit den Werten im aktuellen Modell vergleichbar. Es ist daher davon auszugehen, dass die Stichprobengröße von 132 Proband_innen es ermöglicht, einen Effekt mit einer Wahrscheinlichkeit von 0,8 zu entdecken. Die Wahrscheinlichkeit, fälschlicherweise nicht von einem systematischen Effekt auszugehen, liegt daher nur bei 0,2. Es scheint somit kein Effekt vorzuliegen.

Um zu überprüfen, ob die Unterschiede in der Tokenfrequenz der Testsubstantive die Reaktionszeiten beeinflussen, wird als explorative Analyse dasselbe Modell mit Frequenzausprägungen statt Form-Schematizitätsausprägungen gerechnet: Reaktionszeit \sim Frequenz + (1| Proband_in) + (1|Item). Dafür werden dieselben zufälligen Effekte wie für das Modell mit Form-Schematizitätsausprägungen genutzt. Die zufälligen Effekte sind vergleichbar mit dem vorherigen Modell: Versuchitems variieren mit einer Standardabweichung von 0,1 um den y -Achsenabschnitt, Proband_innen mit 0,26. Die restliche Residuenvarianz beträgt 0,3. Die geschätzten Werte des Modells sind in Tabelle 5.21 aufgelistet. Abbildung 5.49 zeigt die vorhergesagten Reaktionszeiten.

Tabelle 5.21: Werte des Modells für die Reaktionszeiten in der sentence-maze-Studie zu Form-Schematizität in Abhängigkeit von Frequenz

	Wert	SE	t	df
(Intercept = frequent)	0,43	0,07	6,56	11,38
Frequenz: infrequent	0,11	0,09	1,24	8,92
Frequenz: mittelfrequent	0,11	0,09	1,26	8,92

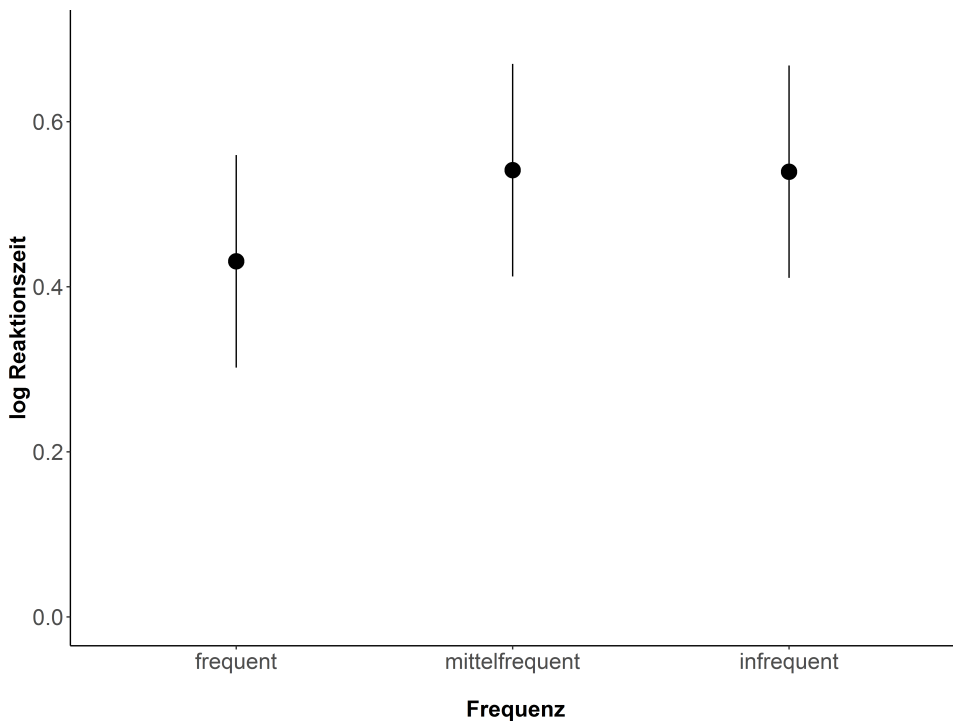


Abbildung 5.49: Vorhergesagte Reaktionszeiten in der sentence-maze-Studie zu Form-Schematizität in Abhängigkeit von Frequenz

Die Substantive der Ausprägung FREQUENT zeigen mit 0,43 leicht niedrigere Werte als die anderen Ausprägungen mit 0,54. Die 95 %-Konfidenzintervalle überlappen sich jedoch deutlich, sodass nicht von einem systematischen Effekt ausgegangen werden kann. Die Effektstärke ist mit einem R^2m von 0,015 noch niedriger als die Effektstärke für das Modell mit Form-Schematizitätsausprägungen, mit zufälligen Effekten ist sie mit einem R^2c von 0,45 höher.

In den Reaktionszeiten ist kein Einfluss der Form-Schematizität zu erkennen. Die Hypothesen aus Abschnitt 4.5.1 bestätigen sich daher nicht. Für Substantive aus der Peripherie des Form-Schemas waren Reaktionszeiten erwartet worden, die von den anderen Form-Schematizitätsausprägungen abweichen. Dabei wurden einerseits höhere Reaktionszeiten in Betracht gezogen, da bei Substantiven in der Peripherie des Form-Schemas im Gegensatz zu den Substantiven der anderen Ausprägungen zwischen zwei (allerdings unterschiedlich stark) mental gefestigten Formen gewählt wird. Andererseits wären auch niedrigere Reaktionszeiten erwartbar, weil die Testsubstantive aus der Peripherie des Form-

Schemas zu den tokenfrequenten Substantiven³³ zählen. In den Daten zeigen sich allenfalls leicht geringere Reaktionszeiten für Substantive in der Peripherie des Form-Schemas im Vergleich zu den anderen Ausprägungen, die jedoch nicht systematisch zu sein scheinen. Ein Frequenzeffekt ist in den Daten zu erahnen, jedoch deutet die explorative Analyse allenfalls auf einen kleinen Effekt hin. Ein möglicher Frequenzeffekt müsste daher in einer weiteren Studie überprüft werden. Auch das Antwortverhalten in der sentence-maze-Studie zeigt lediglich einen klaren und erwartbaren Unterschied zwischen den Substantivklassen, jedoch nicht zwischen prototypischen und peripheren Vertretern des Form-Schemas. Der Grund für das Nullergebnis hinsichtlich der Reaktionszeiten und des Antwortverhaltens könnte darin liegen, dass das Versuchsdesign eine direkte Wahl zwischen stark gefestigten und weniger bzw. nicht gefestigten Formen vorsieht, die sich nur wenig voneinander unterscheiden.

5.4.2.4 Zusammenfassung

Für die einzelnen Studien ließ sich kein Form-Schematizitätseffekt in den Reaktionszeiten erkennen. In der self-paced-reading-Studie werden starke und schwache Formen gleich schnell gelesen, nur bei *Schettose* sind leichte Unterschiede in der Lesezeit zu beobachten. Ob diese systematisch sind, bleibt aufgrund der Überlappung der 95 %-Konfidenzintervalle jedoch fraglich. Auch in der sentence-maze-Studie konnte kein Effekt der Form-Schematizität festgestellt werden. Die Versuchssubstantive aus der Peripherie des Form-Schemas evozieren leicht geringere Reaktionszeiten als die anderen Ausprägungen, dieser Unterschied scheint jedoch nicht systematisch zu sein. Falls der Unterschied systematisch sein sollte, muss näher untersucht werden, wie er mit Tokenfrequenz zusammenhängt, da die frequenten Testsubstantive ebenfalls geringere Reaktionszeiten aufweisen als mittel- und infrequente. Auch hier ist fraglich, ob es sich um einen systematischen Zusammenhang handelt.

In den Reaktionszeiten der lexical-decision-Studie zeigen sich weder Einflüsse der Form-Schematizität noch der Frequenz, allerdings eine Interaktion zwischen Flexionsformen und Deklinationsklasse: Schwache Formen werden bei schwachen Maskulina unabhängig von ihrer Passgenauigkeit zum Form-Schema schnell beurteilt und starke langsam. Das Gegenteil ist für starke Maskulina der Fall. Unabhängig davon, ob der Unterschied in den starken und schwachen Formen an sich oder in der Bewertung als bekannt oder unbekannt begründet ist, zeigt dies, dass Proband_innen für starke und schwache Formen bei Maskulina

³³Mit *Held* und *Zar* zählen zwei der drei tokenfrequenten Testsubstantive zur Peripherie des Form-Schemas.

sensibel sind. Zudem deuten die Reaktionszeiten darauf hin, dass schwache Flexionsformen bei Substantiven in der Peripherie des Form-Schemas stark gefestigt sind und daher starke Formen dominieren, obwohl die starken Formen bereits zu 50 % als bekannt bewertet werden. Im folgenden Abschnitt werden die Ergebnisse der Studien zu Frequenz-, Prototypizität- und Form-Schematizitätseffekten miteinander verglichen und diskutiert.

5.5 Diskussion der Ergebnisse

Die in diesem Abschnitt vorgestellten Untersuchungen hatten zum Ziel, den Einfluss von Frequenz, Prototypizität und Form-Schematizität auf Variation psycholinguistisch zu überprüfen, indem Reaktionszeiten von Stimuli verglichen wurden, die nach Frequenz, Prototypizität sowie Form-Schematizität gestaffelt sind. Zusätzlich zu den Reaktionszeiten wurde das Antwortverhalten der Proband_innen hinsichtlich der Einflussfaktoren untersucht, um Reaktionszeiten und Antwortverhalten in Kontrast setzen zu können.

In diesem Abschnitt werden zunächst die Ergebnisse zu Frequenz, dann zu Prototypizität und schließlich zu Form-Schematizität diskutiert und miteinander verglichen. Hierbei wird jeweils anhand der Reaktionszeiten und des Antwortverhaltens überprüft, inwiefern Frequenz, Prototypizität und Form-Schematizität Variation beeinflussen. Der Fokus liegt dabei auf den Reaktionszeiten, da sie das Kerninteresse der Studien spiegeln. Der Abschnitt schließt mit einer Zusammenfassung samt einer Evaluation der Eignung von *lexical decision*, *sentence maze* und *self-paced reading tasks* zur Elizitation von Reaktionszeiten.

5.5.1 Frequenz

Der Einfluss von Frequenz wurde in einer *lexical decision task* mithilfe von starken Verben untersucht, die unterschiedliche Tokenfrequenzen aufweisen. Hierbei wurde zwischen tokenfrequent (*ziehen*, *sprechen*) und tokeninfrequent (*flechten*, *glimmen*) unterschieden. Die tokeninfrequenten Verben wurden zusätzlich nach attestierter Schwankung in Korpora in zwei Gruppen geteilt: Es wurden infrequente Verben genutzt, die in Korpora noch nicht zwischen starken und schwachen Formen schwanken (*flechten*, *dreschen*), und infrequente Verben, die dies bereits tun (*glimmen*, *sinnen*). Die Proband_innen beurteilten die starke und die schwache Partizip-II-Form der Verben danach, ob sie die Form kennen (siehe Abschnitt 4.2 für ausführliche Erläuterungen zum Design der Studie). Mit diesem Versuchsdesign ließ sich ein deutlicher Einfluss von Frequenz auf Reaktionszeiten feststellen. Frequenz spiegelte sich dabei deutlich stärker als die Einflussfaktoren Prototypizität und Form-Schematizität in den Reaktionszeiten wider.

Starke und schwache Formen tokenfrequenter starker Verben evozieren jeweils geringere Reaktionszeiten als starke und schwache Formen tokeninfrequenter starker Verben, unabhängig davon, ob diese bereits in Korpora schwanken oder nicht. Dieser Unterschied verdeutlicht das *entrenchment* tokenfrequenter Elemente: Die starken Formen frequenter starker Verben sind mental stark gefestigt und daher schnell zu verarbeiten. Der Vorteil für schwache Formen frequenter Verben gegenüber den schwachen Formen infrequenter Verben kann jedoch nicht allein durch Frequenz erklärt werden, denn schwache Formen tokenfrequenter starker Verben werden selten genutzt und sind daher nicht mental gefestigt. Hier greift das statistische Vorkaufsrecht (Goldberg 2019: 74–94; siehe Abschnitt 2.1.3): Da eine tokenfrequente Alternative (starke Form) existiert, kann die selten genutzte Alternative (schwache Form) schnell als unbekannt bewertet werden. Aus demselben Grund werden die schwachen Formen der frequenten Verben langsamer beurteilt als die starken, denn erst der Rückgriff auf die starken Formen erlaubt die Beurteilung der schwachen. Die Reaktionszeiten für die schwachen Formen der frequenten Testverben sind dabei vergleichbar mit den Reaktionszeiten für die starken Formen der anderen Ausprägungen.

Die schnelle Bewertung von schwachen Formen tokenfrequenter starker Verben im Vergleich zu den schwachen Formen infrequenter Verben verdeutlicht die Relevanz des statistischen Vorkaufsrechts: Fokussiert man nur auf die Frequenz einzelner Formen, wäre zu erwarten, dass kein Vorteil für die schwachen Formen der tokenfrequenten starken Verben besteht, da diese so gut wie nie genutzt werden. Dies ist offensichtlich nicht der Fall. Stattdessen wurden die schwachen Formen ebenfalls schnell beurteilt, da sie aufgrund des statistischen Vorkaufsrechts schnell abgelehnt werden können.

Die Häufigkeit, mit der schwache Verbformen in Korpora vorkommen, nimmt keinen Einfluss auf die Schnelligkeit der Beurteilung in der Frequenzstudie. Allerdings weisen infrequente starke Verben ohne Schwankung Unterschiede in den Reaktionszeiten zwischen starken und schwachen Formen auf: Wie bei den tokenfrequenten starken Verben werden auch bei den infrequenten starken Verben ohne Schwankung starke Formen schneller beurteilt als schwache. Dies ist bei den infrequenten starken Verben mit Schwankung nur für wenige Testverben der Fall. Daher ergibt sich bei den infrequenten Verben mit Schwankung kein Unterschied zwischen starken und schwachen Formen.

Der Unterschied in den Reaktionszeiten für starke und schwache Formen bei den infrequenten starken Verben ohne Schwankung könnte im unterschiedlichen Grad an *entrenchment* der Formen begründet sein: Die starken Formen sind zwar aufgrund der geringen Tokenfrequenz bereits nicht mehr so stark gefestigt, sodass das statistische Vorkaufsrecht nur noch eingeschränkt greift, aber die schwachen Formen werden noch nicht verwendet, sodass sie mental nicht

gefestigt sind und daher höhere Reaktionszeiten hervorrufen als die starken Formen. Bei infrequenten Verben mit Schwankung sind die schwachen Formen hingegen bereits durch Gebrauch gefestigt. Daher ist der Vorteil für starke Formen nicht mehr zu erkennen. Neben dem *entrenchment* könnte die Bewertung der Formen die Reaktionszeiten beeinflussen: Die Bewertung als unbekannte Form scheint längere Reaktionszeiten hervorzurufen als die Bewertung als bekannte Form. Dies war sowohl bei schwachen Formen der infrequenten Verben mit Schwankung als auch bei den starken Formen der Substantive aus der Peripherie des Form-Schemas schwacher Maskulina in der Form-Schematizitätsstudie zu erkennen. Für diese Formen wurde die Frage nach der Bekanntheit zu ungefähr 50 % bejaht bzw. verneint, wobei die Bewertung als bekannt im Mittel jeweils geringere Reaktionszeiten aufwies als die Bewertung als unbekannt. Dass affirmierende Antworten geringere Reaktionszeiten auslösen als negierende, ist auch in anderen Studien zu beobachten (Wegge 2001: 282–283).

Neben den Reaktionszeiten verdeutlicht das Antwortverhalten in der Studie den Frequenzeffekt: Während bei frequenten Verben schwache Formen als unbekannt bewertet werden, werden sie bei den infrequenten Verben ohne Schwankung vereinzelt als bekannt angesehen und bei Verben mit Schwankung zur Hälfte. Die Beurteilung der starken Formen wird dagegen nicht von der Frequenzausprägung beeinflusst.

Wie die Studie zeigt, sind Variation und Frequenz eng verbunden: Starke Formen von tokenfrequenten Verben sind mental stark gefestigt, was sich einerseits in den niedrigen Reaktionszeiten und andererseits in den hohen Bekanntheitswerten der Formen zeigt. Zudem gibt die geringe Bekanntheit schwacher Formen einen indirekten Hinweis auf das *entrenchment* der starken Formen. Hier ist Variation sehr unwahrscheinlich. Die starken Formen tokeninfrequenter starker Verben sind hingegen weniger stark gefestigt, was sich in den höheren Reaktionszeiten zeigt. Zwar weisen die starken Formen auch bei den infrequenten starken Verben einen hohen Grad an Bekanntheit auf, jedoch sind die schwachen Formen deutlich häufiger bekannt als bei den frequenten starken Verben. Die Studie verdeutlicht somit, dass Variation für tokeninfrequente Verben generell möglich ist. Insgesamt lässt sich der Einfluss der Frequenz somit als ein Zwischenschritt modellieren: Im ersten Schritt ist relevant, wie frequent ein Verb ist. Die Tokenfrequenz definiert, ob überhaupt Variation möglich ist. Bei geringer Tokenfrequenz können starke Formen die schwachen nicht mehr so leicht dominieren. Dies spiegelt sich in erhöhten Reaktionszeiten für starke Formen bei tokeninfrequenten Verben. Variation ist somit wahrscheinlich. Im zweiten Schritt werden schwache Formen durch Gebrauch gefestigt. Erst hier wird Variation in Korpora sichtbar.

Zwischen den infrequenten Verben mit und ohne Schwankung in Korpora lassen sich Unterschiede im Antwortverhalten feststellen, aber keine Unterschiede in Hinblick auf die Reaktionszeiten. Die ausbleibenden Unterschiede in den Reaktionszeiten lassen darauf schließen, dass die mentale Repräsentation der starken Formen unabhängig vom Vorkommen in Korpora nicht mehr so stark gefestigt ist wie bei den frequenten starken Verben. Variation ist daher auch bei den Verben ohne attestierte Schwankung bereits angelegt, da das statistische Vorkaufsrecht nur noch eingeschränkt greifen kann. Dennoch scheinen die starken Formen bei den infrequenten Verben ohne Schwankung noch einen kleinen Vorteil gegenüber den schwachen zu haben, der bei den infrequenten Verben mit Schwankung nicht mehr besteht. Der Wegfall des Vorteils für starke Formen deutet darauf hin, dass sich die mentale Repräsentation der schwachen Formen bei den infrequenten Verben mit Schwankung festigt und sich der Repräsentation der starken Formen annähert. Die Variation in der Konjugation ist damit zweifach von Frequenz beeinflusst: Zunächst wirkt das statistische Vorkaufsrecht starker Formen. Wenn dieses durch den Verlust der Tokenfrequenz an Einfluss verliert, wird die Nutzung schwacher Formen ermöglicht. Im nächsten Schritt werden die schwachen Formen durch Gebrauch tokenfrequenter und mental gefestigt, sodass sich die mentalen Repräsentationen beider Formen anpassen. Im folgenden Abschnitt werden die Ergebnisse zum Einfluss der Prototypizität diskutiert.

5.5.2 Prototypizität

Um den Einfluss der Prototypizität auf Variation zu untersuchen, wurde die Auxiliarselektion von *haben* und *sein* mithilfe einer *sentence maze task* in Sätzen mit unterschiedlichen Transitivitätsgraden untersucht: Transitive Sätze mit einem Objekt, das auf eine belebte Entität referiert (*die Kinder nach Hause fahren*), und intransitive Sätze ohne Akkusativobjekt (*zur Kur fahren*) wurden kontrastiert. Zusätzlich wurde zwischen zwei ambigen Konstruktionen unterschieden. Bei der Ausprägung AMBIG I kann das Objekt als Patiens oder Instrument aufgefasst werden (*ein Cabrio fahren*), die Ausprägung AMBIG II enthält eine Akkusativergänzung, die ein Adverbial darstellt (*die Strecke fahren*) (siehe Abschnitt 4.3 für ausführliche Erläuterungen zum methodischen Vorgehen). Die Sätze enthielten jeweils Bewegungsverben (*fahren, fliegen, reiten*) und waren als Nebensätze konzipiert, sodass das zu wählende Auxiliar direkt auf das Partizip II der Bewegungsverben folgte (z. B. *Er sah, wie die Mutter die Kinder zur Schule gefahren hat/*ist*). Im Gegensatz zu den Ergebnissen der Frequenzstudie sind die Ergebnisse der Prototypizitätsstudie weniger eindeutig.

Die intransitiven Sätze evozieren deutlich geringere Reaktionszeiten als die anderen Transitivitätsausprägungen (TRANSITIV, AMBIG I, AMBIG II). Diese rufen jeweils miteinander vergleichbare Reaktionszeiten hervor. Der visuelle Vergleich zwischen den transitiven Testsätzen und transitiven Füllersätzen ohne Bewegungsverb (*die Kinder zur Schule bringen*) deutet darauf hin, dass die transitiven Testsätze höhere Reaktionszeiten evozieren als die Füllersätze. Dieser Zusammenhang müsste in einer weiteren Studie statistisch überprüft werden. Sollten transitive Sätze mit Bewegungsverben höhere Reaktionszeiten evozieren als transitive Sätze ohne Bewegungsverb, kann dies auf zwei Ursachen zurückgeführt werden: Bewegungsverben werden vorrangig in intransitiven Sätzen genutzt (Gillmann 2016: 267–268), sodass die Verwendung von Bewegungsverben in transitiven Sätzen einen höheren Prozessierungsaufwand darstellen könnte. Zudem lassen sich die erhöhten Reaktionszeiten als Einfluss des *sein*-Schemas interpretieren. Bewegungsverben sind aufgrund der häufigen Verwendung in intransitiven Sätzen mit *sein* assoziiert. Nur in transitiven Sätzen ist diese Assoziation gestört und *haben* wird selektiert. Der Wechsel von generell überwiegendem *sein* bei Bewegungsverben zu in transitiven Sätzen genutztem *haben* könnte daher als Erklärung für die erhöhten Reaktionszeiten dienen und auf die Wirksamkeit der beiden Schemata hindeuten. Während sich die Wirksamkeit des *haben*- und des *sein*-Schemas in den Reaktionszeiten nur ansatzweise zeigt, ist es im Antwortverhalten deutlich zu erkennen: Transitive Sätze mit Bewegungsverben evozieren *haben*, intransitive *sein*.

Die prototypisch organisierte Verbindung zwischen den Funktionen des *haben*- und *sein*-Schemas lässt sich im Antwortverhalten nur ansatzweise feststellen. Die ambigen Sätze evozieren vornehmlich *sein*-Antworten. Jedoch ist bei *fliegen* und *fahren* ein Unterschied zwischen AMBIG I (Objekt ambig zwischen Patiens und Instrument) mit ca. 14 % *haben*-Antworten und AMBIG II (Akkusativergänzung ist ein Adverbial) mit nur 0,03 % *haben*-Antworten zu erkennen. Damit tendiert die Ausprägung AMBIG II, die nah am intransitiven Pol ist, mehr zu *sein*-Antworten als die Ausprägung AMBIG I, die nah am transitiven Pol ist.

Zusammen mit dem Antwortverhalten lassen sich die im Vergleich zu den intransitiven Sätzen erhöhten Reaktionszeiten in den ambigen Sätzen interpretieren: Aufgrund der überwiegenden *sein*-Antworten können die Reaktionszeiten nicht durch die Umstellung von *sein* auf *haben* erklärt werden. Zudem kann die Komplexität der Sätze nur eingeschränkt als Auslöser der erhöhten Zeiten gesehen werden, da die transitiven Füllersätze, die eine vergleichbare Komplexität wie die ambigen Sätze aufweisen, geringere Reaktionszeiten evozieren als die ambigen Sätze. Die erhöhten Reaktionszeiten könnten daher auf die Ambiguität der Sätze zurückgehen. Somit lässt sich ein Prototypizitätseffekt zumindest

erahnen: Prototypische Vertreter werden schnell verarbeitet, periphere langsamer, da beide Auxiliare möglich sind. Der Prototypizitätseffekt eines Prototyps (Transitivität) wird jedoch durch den anderen Prototyp (Intransitivität und Bewegungsverb) und dessen starke Assoziation mit *sein* überlagert. Da der Unterschied zwischen Test- und Füllersätzen nur visuell beobachtet wurde, muss dies jedoch eine Hypothese bleiben, die in weiteren Studien überprüft werden muss.

In der Studie fällt das Testverb *reiten* auf, da es ein deutlich anderes Antwortverhalten aufweist als *fahren* und *fliegen*. Vor allem die durchgängigen *sein*-Antworten unabhängig vom Transitivitätsgrad der Sätze stechen dabei ins Auge. Die *sein*-Antworten im transitiven Satz könnten ebenfalls ein Prototypizitätseffekt sein, da Pferde keine Menschen sind und somit aufgrund der geringeren Belebtheit nicht prototypisch in der Patiensrolle vorkommen. Allerdings scheint sich *reiten* insgesamt anders zu verhalten als die anderen Bewegungsverb, wie die vergleichsweise vielen *haben*-Antworten in den ambigen Sätzen und die von den anderen Bewegungsverb abweichenden Auxiliarkombinationen zeigen: Anders als bei *fahren* und *fliegen*, bei denen die Kombination aus *haben*-Wahl für transitive Sätze und *sein*-Wahl für alle anderen Ausprägungen deutlich überwiegt, kommt diese Kombination bei *reiten* ähnlich häufig vor wie die durchgängige *sein*-Wahl für alle Ausprägungen und die Kombination aus *haben*-Wahl für TRANSITIV und AMBIG I und *sein*-Wahl für INTRANSITIV und AMBIG II.

Der Einfluss von Prototypizität auf Variation lässt sich in der sentence-maze-Studie trotz der oben genannten Einschränkungen erkennen. Die Auxiliare sind bei den prototypischen Funktionen der Schemata (transitiv; intransitiv und Bewegungsverb) mental gefestigt, was sich in den niedrigen Reaktionszeiten und den klar verteilten Antworten zeigt. Die erhöhten Reaktionszeiten für die transitiven Sätze lassen sich mit dem Wechsel von einem zum anderen Prototyp erklären. Im Übergangsbereich zwischen transitiv und intransitiv sind die Auxiliare nicht so stark gefestigt, was sich in höheren Reaktionszeiten und einer leicht erhöhten Variation in den Antworten zeigt. Die Variation in den Antworten ist allerdings weniger stark als angenommen, da in den ambigen Sätzen *sein* als Defaultauxiliar für Bewegungsverb zu greifen scheint. Die durchgängigen *sein*-Antworten bei *reiten* sowie die erhöhten Reaktionszeiten in den transitiven Sätzen könnten darauf hindeuten, dass sich die *sein*-Selektion bei Bewegungsverb auf transitive Kontexte ausweiten wird. Dies muss zum gegenwärtigen Zeitpunkt jedoch eine kaum überprüfbare Hypothese bleiben. Im folgenden Abschnitt werden die Ergebnisse zum Einfluss der Form-Schematizität diskutiert.

5.5.3 Form-Schematizität

Der Einfluss der Form-Schematizität auf Variation wurde innerhalb einer *self-paced reading task* mit anschließendem Produktionsexperiment, einer *lexical decision task* und einer *sentence maze task* mithilfe von schwachen und starken Maskulina untersucht. Dabei wurde in der *self-paced-reading-Studie* mit Pseudosubstantiven gearbeitet, in den anderen Studien wurden real existierende Substantive genutzt. Die Versuchitems gehörten entweder zum Prototyp des Form-Schemas schwacher Maskulina (*Kollege, Schettose*), waren in der Peripherie des Form-Schemas (*Graf, Knatt*) oder gehörten der starken Flexion an (*Vogt, Grettel*). In der *self-paced-reading-Studie* lasen die Proband_innen starke und schwache Formen der Testsubstantive, in der *lexical-decision-Studie* bewerteten sie starke und schwache Formen der Testsubstantive danach, ob ihnen die Form bekannt ist, und in der *sentence-maze-Studie* wählten sie zwischen der starken und der schwachen Form der Testsubstantive (siehe Abschnitt 4.4 und 4.5 für ausführliche Erläuterungen zum methodischen Vorgehen). Anders als in der Frequenz- und der Prototypizitätsstudie lässt sich der Einfluss der Form-Schematizität in den Reaktionszeiten allenfalls erahnen.

Lediglich in der *self-paced-reading-Studie* zeigt sich ein kleiner Form-Schematizitätseffekt für das Testitem *Schettose*: Hier sind die Lesezeiten für die schwache Form etwas niedriger als die Lesezeiten für die starke Form. Dies ist aufgrund der hohen Passgenauigkeit von *Schettose* zum Form-Schema schwacher Maskulina auch zu erwarten. Es ist aber fraglich, ob dieser Effekt systematisch ist. Wenn er systematisch ist, stellt sich die Frage, warum kein gegenteiliger Effekt bei *Grettel* zu erkennen ist, bei dem aufgrund des Form-Schemas starker Maskulina nur starke Formen zu erwarten sind. In den anderen Studien zeigte sich gar kein Einfluss der Form-Schematizität auf Reaktionszeiten. Das Antwortverhalten wurde dagegen in allen Studien deutlich durch Form-Schematizität beeinflusst: Für die prototypischen Vertreter des Form-Schemas schwacher Maskulina wurden jeweils schwache Formen bevorzugt, für starke Maskulina jeweils starke. Variation war für die Peripherie festzustellen. Nur für die *sentence-maze-Studie* war das aufgrund der Gegenüberstellung starker und schwacher Formen nicht der Fall: Die stärker gefestigte Form (schwach) kann durch die Gegenüberstellung die weniger gefestigte (stark) leichter statistisch ausstechen und wird daher bevorzugt.³⁴

Form-Schematizität nimmt somit Einfluss auf Variation: Die prototypischen Vertreter des Form-Schemas schwacher Maskulina sowie Substantive, die dem

³⁴Dieses Problem stellte sich für die Prototypizitätsstudie nicht, da *haben* und *sein* jeweils nach den Bewegungsverben präsentiert wurden und die Auxiliare daher erst mit den Bewegungsverben verknüpft werden mussten, sodass die stark gefestigte Form nicht der weniger bzw. nicht gefestigten gegenüberstand.

Form-Schema nicht angehören, evozieren jeweils ein klar verteiltes Antwortverhalten. Variation ist in der Peripherie des Form-Schemas zu beobachten. Das Antwortverhalten in den Form-Schematizitätsstudien verdeutlicht damit nicht nur den Einfluss von Form-Schemata, sondern auch von Prototypizität auf Variation. Es stellt sich daher die Frage, warum sich das Antwortverhalten nicht in den Reaktionszeiten spiegelt.

Eine Erklärung für die Diskrepanz zwischen Antwortverhalten und Reaktionszeiten, die auch Schmitt (2019a: 171–172) diskutiert, ist, dass die Proband_innen die Endungen der Substantive nicht wahrnehmen. Dies ist für die vorliegenden Studien jedoch auszuschließen, da in der self-paced-reading-Studie der Aufbau der zu lesenden Wörter abgefragt wurde (Haben Sie *des Schettosen* oder *des Schettoses* gelesen?) und in der sentence-maze-Studie zwischen beiden Formen gewählt wurde. In der lexical-decision-Studie zeigt sich die Sensibilität für starke und schwache Formen einerseits im Antwortverhalten und andererseits hinsichtlich der Reaktionszeiten. Die Reaktionszeiten interagieren in Hinblick auf Deklinationsklasse und die Flexionsform: Schwache Formen von schwachen Maskulina werden schnell beurteilt und starke langsam, das Gegenteil ist für starke Maskulina der Fall.³⁵ Dies deutet auf Sensibilität für Formen hin und zeigt, dass Form-Schematizität nicht nur für die Produktion von Formen relevant ist, sondern auch für die Prozessierung, zumindest in dem Umfang, in dem die Methode Einblick in Prozessierung gewährt.

Hinsichtlich der self-paced-reading-Studie ist es zudem möglich, dass die Pseudosubstantive das Ergebnis beeinflusst haben, da hier keine der Formen durch Gebrauch mental gefestigt ist. Ein Unterschied in den Reaktionszeiten kann nur durch Rückgriff auf die Form-Schemata schwacher und starker Maskulina zustande kommen und ist daher eventuell sehr subtil. Wie anhand der Antworten im Produktionsexperiment zu erkennen ist, wählten viele Proband_innen Nullendungen für das Substantiv *Grettel*, sodass es möglich ist, dass eine Nullendung in diesem Fall die am leichtesten zu prozessierende Option gewesen wäre. Diese Erklärung greift nicht für die anderen Studien, da hier mit real existierenden Substantiven gearbeitet wurde und die Formen der jeweiligen Deklinationsklasse daher durch Gebrauch gefestigt sind. Dies wird auch durch das Antwortverhalten in den Studien sowie durch die Interaktion in den Reaktionszeiten zwischen Flexionsform und Deklinationsklasse in der lexical-decision-Studie deutlich. Die Reaktionszeiten spiegeln das Antwortverhalten zumindest in dieser Hinsicht.

³⁵Auch wenn die Interaktion vorrangig an der Ablehnung oder Zustimmung hinsichtlich der Bekanntheit von Formen hängen sollte, würde sie implizit auf die Sensibilität für Formen hinweisen, da das Antwortverhalten von der Deklinationsklassenzugehörigkeit abhängt.

Die Diskrepanz zwischen Reaktionszeiten und Antwortverhalten lässt abseits der methodischen Überlegungen auch theoretische Schlüsse zu. So legt die lexical-decision-Studie nahe, dass die starken Formen von Substantiven in der Peripherie des Form-Schemas trotz der Variation im Antwortverhalten nicht stark gefestigt sind: Zwar liegt die Bekanntheit der starken Formen für Testsubstantive in der Peripherie des Form-Schemas bei 50 %. Dies spiegelt sich aber nicht in den Reaktionszeiten.³⁶ Das statistische Vorkaufsrecht für die schwachen Formen scheint bereits eingeschränkt zu sein, wie das Antwortverhalten zeigt, die neuen starken Formen scheinen jedoch noch nicht genug gefestigt zu sein, um Reaktionszeiten hervorzurufen, die mit den schwachen Formen vergleichbar sind. In der Frequenzstudie weisen die infrequenten Verben mit Schwankung in Korpora ein vergleichbares Antwortverhalten wie die Substantive aus der Peripherie des Form-Schemas in der lexical-decision-Studie zu Form-Schematizität auf. Allerdings ist für die infrequenten Verben mit Schwankung eine Angleichung der Reaktionszeiten zwischen starken und schwachen Formen zu erkennen, was für die Substantive in der Peripherie nicht der Fall ist. Auch dies deutet darauf hin, dass die starken Formen von Substantiven in der Peripherie des Form-Schemas vergleichsweise wenig gefestigt sind.

Da die Testsubstantive in der Peripherie anders als die infrequenten Verben mit Schwankung in Korpora nicht danach ausgewählt wurden, ob sie bereits Schwankungen in Korpora aufweisen, könnte es sein, dass die starken Formen von Substantiven in der Peripherie tatsächlich im Sprachgebrauch zu selten vorkommen, um ähnlich gefestigt zu sein wie die schwachen Formen von infrequenten starken Verben mit Schwankung in Korpora. Die dennoch zu beobachtende hohe Bekanntheit der starken Formen von Substantiven aus der Peripherie des Form-Schemas lässt sich vor diesem Hintergrund mit Rückgriff auf die Relevanz der Kategorien *TEMPUS* und *KASUS* erklären: *Kasus* ist weniger relevant als *Tempus*, da das *Tempus* die Bedeutung eines Verbs stark verändert, während der *Kasus* nur anzeigt, wie Entitäten zueinander stehen und daher die Semantik eines Substantivs nicht beeinflusst (Nübling & Dammal 2004: 179–180, Nübling 2016: 158–159). Aufgrund der geringeren Relevanz ist Variation für die Kategorie *KASUS*

³⁶Im Fließtext wird hinsichtlich der lexical-decision-Studie nur eine mögliche Angleichung der Reaktionszeiten für starke und schwache Formen bei Substantiven aus der Peripherie des Form-Schemas diskutiert. Neben dieser Möglichkeit wäre auch ein generell langsames Antwortverhalten für starke und schwache Formen der Substantive in der Peripherie im Vergleich zu den anderen Ausprägungen denkbar. Dies wurde auch ursprünglich erwartet. Im Gegensatz zu den anderen Ausprägungen existieren in der Peripherie zwei Formen, die (zwar unterschiedlich stark, aber dennoch beide) mental gefestigt sind und daher beide bekannt sein können. Die Schnelligkeit, mit der eine Form beurteilt wird, scheint jedoch nicht durch die theoretische Existenz einer anderen Form beeinflusst worden zu sein.

wahrscheinlicher, sodass die starken Formen von Substantiven in der Peripherie des Form-Schemas trotz des geringen *entrenchments* bereits als bekannt bewertet werden. Zudem ist die Genitivmarkierung mit *-(e)s* vs. *-(e)n* sehr viel subtiler als der Ablaut.

Um das Zusammenspiel aus *entrenchment* und Relevanz näher zu untersuchen, wäre es für eine weitere Studie interessant, analog zu den Verben mit und ohne Schwankung in Korpora die Testsubstantive in der Peripherie des Form-Schemas nach ihrer Ratio von schwachen zu starken Formen in Substantive mit und ohne Schwankung in Korpora zu teilen. Hierbei wäre aufgrund der geringen Relevanz der Kategorie KASUS unabhängig von der attestierten Schwankung in Korpora eine Akzeptanz der starken Formen von ca. 50 % zu erwarten und für Substantive mit Schwankung in Korpora aufgrund des höheren *entrenchments* starker Formen eine Angleichung der Reaktionszeiten für starke und schwache Formen.

In der sentence-maze-Studie könnte Tokenfrequenz eine Erklärung für die Diskrepanz zwischen Reaktionszeit und Antwortverhalten sein. Ursprünglich wurde angenommen, dass periphere Vertreter des Form-Schemas in der sentence-maze-Studie höhere Reaktionszeiten hervorrufen als prototypische Vertreter des Form-Schemas sowie starke Maskulina, weil in der Peripherie zwischen zwei mental gefestigten Formen gewählt werden muss und die Wahl daher verlangsamter sein könnte. Dies ist in den Reaktionszeiten jedoch nicht zu erkennen. Die Substantive in der Peripherie neigen in der sentence-maze-Studie sogar zu geringeren Reaktionszeiten als die Substantive der anderen Ausprägungen, wobei dieser Unterschied zufällig sein kann. Die unterschiedlichen Tokenfrequenzen der Testsubstantive könnten als Erklärung für die niedrigeren Reaktionszeiten der Testsubstantive aus der Peripherie des Form-Schemas dienen. Zwei der drei Testsubstantive aus der Peripherie des Form-Schemas zählen zu den tokenfrequenten Testsubstantiven in der Studie. Die tokenfrequenten Testsubstantive evozieren geringere Reaktionszeiten als die mittel- und infrequenten, jedoch kann auch dieser Unterschied zufällig sein. Sollte der Effekt systematisch sein, wäre es möglich, dass Tokenfrequenz den Form-Schematizitätseffekt überschreibt und Substantive in der Peripherie aufgrund der höheren Tokenfrequenz etwas schneller verarbeitet werden als andere Substantive. Ist dies der Fall, könnte man den Schluss ziehen, dass Form-Schematizität die Deklinationsklassenzugehörigkeit grundlegend steuert und periphere Vertreter durch Tokenfrequenz in der schwachen Klasse gehalten werden: Da tokenfrequente Formen immer wieder aktiviert und somit mental gefestigt werden, bleibt die Flexion trotz der Peripherie des Form-Schemas schwach. Dies muss jedoch eine Hypothese bleiben, da die vergleichsweise niedrigen Reaktionszeiten sowohl bei den frequenten Substantiven als auch bei den Substantiven aus der Peripherie zufällig sein könnten.

Insgesamt zeigt sich hinsichtlich der Form-Schematizität ein zweigeteiltes Ergebnis: Das Form-Schema beeinflusst das Antwortverhalten und in der Peripherie des Form-Schemas ist Variation im Antwortverhalten zu erkennen. Die Reaktionszeiten spiegeln den Einfluss des Form-Schemas hingegen nicht. Es kann sein, dass Tokenfrequenzeffekte die Form-Schematizität überlagern. Dies müsste jedoch in weiteren Studien überprüft werden. Zudem ist es möglich, dass die in den Studien genutzten Verfahren zur Elizitation von Reaktionszeitunterschieden sich nicht für die Messung von Form-Schematizitätseffekten eignen. Im nächsten Abschnitt wird daher neben einer Zusammenfassung der Ergebnisse eine Evaluation der Reaktionszeitmessverfahren in Bezug auf alle Studien vorgenommen.

5.5.4 Zusammenfassung und Evaluation der Reaktionszeitmessverfahren

Der Einfluss von Frequenz, Prototypizität und (Form-)Schematizität auf Variation lässt sich im Antwortverhalten klar beobachten. In den Reaktionszeiten zeigt sich der Einfluss hingegen nur teilweise: Der Einfluss der Frequenz ist sehr deutlich zu erkennen, ein Einfluss von Schemata ist für *haben* und *sein* ebenfalls festzustellen. Die erhöhten Reaktionszeiten in den ambigen Sätzen lassen sich zudem als Effekt der Prototypizität sehen, der jedoch nur teilweise im Antwortverhalten gespiegelt wird, da auch die ambigen Sätze stark zur Selektion von *sein* neigen. Form-Schematizität zeigt sich hingegen überhaupt nicht in den Reaktionszeiten.

Es ist möglich, dass die gemessenen Reaktionszeiten zu ungenau sind, um Effekte von Form-Schematizität in der Prozessierung beobachten zu können. Generell zeigen die Ergebnisse der Frequenz- und der Prototypizitätsstudie, dass *lexical decision* und *sentence maze tasks* geeignet sind, um Unterschiede in Reaktionszeiten herauszuarbeiten. Es stellt sich aber die Frage, ob *self-paced reading* zweckmäßig ist, um Prozessierungsunterschiede zwischen Deklinationsformen zu messen. In Hinblick auf salientere Variationsfälle scheint *self-paced reading* geeignet zu sein, wie Schmitt (2019b) anhand der Kasusselektion von *wegen* zeigt, jedoch könnten die Unterschiede zwischen den Deklinationsklassen zu subtil für die Methode sein. Dies könnte auch darin begründet sein, dass der Kasus ein Kongruenzphänomen ist: In den Form-Schematizitätsstudien war die Funktion GENITIV immer schon durch den Determinierer markiert, sodass die starken und schwachen Endungen der Substantive keine so große Rolle für die Prozessierung spielen könnten. Zwar sind die Proband_innen offensichtlich sensitiv für die Genitivendungen der Substantive, wie das Antwortverhalten zeigt, dennoch könnten die Endungen in der Prozessierung nur eine geringe Rolle spielen. Die durch

die Endungen hervorgerufenen Unterschiede in der Prozessierung könnten daher sehr subtil sein, da die Endungen nicht eigenständig die Funktion GENITIV anzeigen.

Zusätzlich dazu ist für die *sentence maze task* festzuhalten, dass diese für die Erforschung von Variation in der Flexion weniger gut geeignet zu sein scheint, da mental stark gefestigte und weniger stark gefestigte Formen gleichzeitig präsentiert werden. Für alle Messmethoden lassen die ausbleibenden Reaktionszeitunterschiede in Hinblick auf Form-Schematizität bei gleichzeitiger Beeinflussung des Antwortverhaltens darauf schließen, dass Unterschiede in der Prozessierung existieren, die gemessenen Reaktionszeiten jedoch zu wenig sensitiv für diese sind. Hier wäre es sinnvoll, mit feineren Messmethoden wie bspw. EEG zu arbeiten.

Die Ergebnisse der Frequenz- und der Prototypizitätsstudie verdeutlichen trotz der genannten Einschränkungen das Potential von Reaktionszeitmessungen für die Erforschung grammatischer Varianten. So zeigt die Prototypizitätsstudie, dass Unterschiede in der Verarbeitung von Prototypen existieren und die Frequenzstudie verdeutlicht, dass Reaktionszeiten Hinweise auf Variation liefern können, die noch nicht in der Sprachproduktion sichtbar ist. Im nächsten Kapitel wird anhand der Kernergebnisse der Dissertation ein Fazit gezogen und ein Ausblick gegeben.

6 Zusammenfassung und Ausblick

In diesem Kapitel soll ein abschließender Blick auf den Einfluss von Frequenz, Prototypizität und (Form-)Schematizität auf sprachliche Variation geworfen werden. Dabei werden die Einflussfaktoren zunächst theoretisch modelliert. Zudem wird ihr Potential, Variation und Stabilität zu beeinflussen, skizziert. Darüber hinaus wird beschrieben, wie die Einflussfaktoren die Wahrscheinlichkeit für Variation steigen bzw. sinken lassen. Die Vorhersagen des Modells werden im Anschluss mit den empirischen Befunden in dieser Arbeit verglichen und es wird diskutiert, inwiefern sich der Einfluss der Faktoren in Reaktionszeiten spiegelt. Für jeden Einflussfaktor wird ein Ausblick auf offene Forschungsfragen und weiteren Forschungsbedarf gegeben. Abschließend wird diskutiert, wie die Einflussfaktoren auf Variation gewichtet werden können.

6.1 Frequenz

Frequenz beeinflusst Sprache und Variation grundlegend, da eine hohe Frequenz sowohl *entrenchment* als auch Kategorisierung und Schematisierung ermöglicht (Langacker 2008: 220). Somit lassen sich Prototypizität als Kategorisierungsprinzip und (Form-)Schematizität als Ergebnis von Abstraktion und damit von Schematisierung auch als Konsequenzen von Frequenz sehen.

Hinsichtlich der Frequenz ist zwischen Typen- und Tokenfrequenz zu unterscheiden, wie in Abschnitt 2.1.1 herausgearbeitet wurde. Typenfrequenz hängt eng mit der Generalisierbarkeit eines Flexionsmusters und dessen Produktivität zusammen (Bybee & Thompson 1997: 384). Gemeinsam mit der Variabilität der bereits in einer Kategorie existierenden Exemplare und der Ähnlichkeit eines neuen potentiellen Mitglieds der Kategorie zu bereits bestehenden Exemplaren nimmt Typenfrequenz Einfluss auf die Abdeckung (*coverage*) einer Kategorie (Goldberg 2019: 62–65). Das Konzept der Abdeckung fragt danach, wie gut eine Kategorie attestiert ist, wenn zu den bereits bestehenden Mitgliedern ein neues hinzutritt. Eine Kategorie wird dabei als Cluster in einem hyperdimensionalen Raum betrachtet. Hat eine Kategorie viele Mitglieder, die variable Eigenschaften aufweisen, ist das Cluster sehr vielfältig. Ein potentielles neues Mitglied führt

daher zu einer Kategorie mit guter Abdeckung, denn es ist wahrscheinlich, dass Ähnlichkeiten zu bereits existierenden Mitgliedern bestehen. Anhand dieser exemplarbasierten Kategorisierung kann aus einem vielfältigen Cluster ein abstraktes Schema generalisiert werden, wie bspw. das Dentalsuffix *-t* und die Funktion [+Vergangenheit] (Goldberg 2019: 136). Dieses abstrakte Schema kann als Default für die Konjugation genutzt werden, da Neuzugänge aufgrund der Abstraktheit des Schemas immer zu einer gut abgedeckten Kategorie führen.

Ein typeninfrequentes Cluster mit Mitgliedern, die in ihren Eigenschaften wenig variabel sind, kann hingegen weniger gut abstrahiert werden: Ein potentielles neues Mitglied muss den bereits bestehenden Exemplaren sehr ähnlich sein, um eine gute Abdeckung zu ermöglichen (Goldberg 2019: 62–65). Ist dies nicht der Fall, entsteht durch den Neuzugang eine Kategorie mit schlechter Abdeckung, da der Neuzugang zu einem brüchigen Cluster führt. Dies kann leicht an einem Beispiel verdeutlicht werden: Verben der Ablautreihe 3a sind eine wenig variable Kategorie mit guter Abdeckung, da sie sich in ihren phonologischen Eigenschaften stark ähneln und so das klar konturierte phonologische Cluster [#_I + ŋ + (C)] bilden. Ein Verb wie *googeln* passt nicht zu diesem Cluster, da es /u:/ statt /ɪ/ als Infinitivstammvokal aufweist und einen Plosiv statt einem Nasal. *Googeln* hat in der ersten Silbe somit keine phonologische Verbindung zu den restlichen Mitgliedern der Ablautreihe 3a. Wenn *googeln* wie Verben der Ablautreihe 3a konjugiert würde, wäre die Abdeckung der Kategorie (Verben der Ablautreihe 3a und *googeln*) daher schlecht. Deswegen wird *googeln* auch nicht wie die Verben der Ablautreihe 3a konjugiert, sondern nach dem abstrakten Schema mit Dentalsuffix (*googelte/gegoogelt*).

Das Beispiel zeigt, dass die Möglichkeit zur Generalisierung bei typeninrequenten Clustern mit Mitgliedern, die wenig variabel sind, eingeschränkt ist (Goldberg 2019: 65). Die phonologische Form der Verben ist daher Teil der Generalisierung. In diesem Fall bildet eine Kategorie mit geringer Typenfrequenz somit ein Form-Schema aus, in dem die Funktion [+Vergangenheit] mit der Form [#_I + ŋ + (C)] assoziiert ist. An dieser Stelle zeigt sich bereits der grundlegende Einfluss, den Frequenz auf Form-Schematizität hat.

Die enge Verknüpfung von Typenfrequenz mit Generalisierbarkeit führt zu einem grundlegenden Einfluss der Typenfrequenz auf Variation. Sie gibt bei Variation zwischen Flexionsklassen die Schwankungsrichtung vor: Mitglieder der typeninrequenten Klasse können zu Formen der typenfrequenten Klasse schwanken (*geglommen* > *geglimmt*). Eine umgekehrte Schwankungsrichtung ist hingegen generell unwahrscheinlich, sie kann jedoch bspw. durch Form-Schemata ausgelöst werden (*gewinkt* > *gewunken*).

Wichtig ist hierbei, dass eine Funktion (z. B. [+Vergangenheit]) mit mehreren Formen assoziiert ist, von denen eine typenfrequent ist, während die anderen typeninfrequent sind. Die typenfrequenterere Form (Dentalsuffix) wäre theoretisch immer eine mögliche Option, weswegen sie die typeninfrequenten Formen (z. B. Ablaut auf /o:/, /i:/) ablösen kann. Tokenfrequenz beeinflusst, ob Variation tatsächlich stattfindet: Tokenfrequente Elemente sind mental stark gefestigt und somit leicht und schnell zu verarbeiten (Rubenstein u. a. 1970, Bybee & Thompson 1997: 380, Schneider 2014: 13). Es ist daher für die Prozessierung von Vorteil, wenn tokenfrequente Elemente einer typeninfrequenten Flexionsklasse angehören: Da sie häufig genutzt werden, müssen sie nicht durch ein generalisierbares Muster gestützt werden. Hierdurch sind kurze und gut unterscheidbare, da differente Flexionsformen möglich (Werner 1989: 42–43, Nübling 2000: 226–227). Zudem können die Formen der typeninfrequenten Klasse aufgrund ihrer hohen Tokenfrequenz die theoretisch immer möglichen regelmäßigen Formen statistisch ausstechen (Goldberg 2019: 83–84). Für tokeninfrequente Formen der typeninfrequenten Klasse besteht dieser Vorteil nicht mehr, da sie weniger stark gefestigt sind. Der Anschluss an ein generalisierbares Muster bietet sich daher an. Dies lässt sich anhand der starken Verben gut verdeutlichen: Die starke Konjugation ist typeninfrequent. Wenn ein starkes Verb tokeninfrequent wird, können die starken Formen die schwachen aufgrund des geringeren *entrenchments* nicht mehr so stark dominieren und statistisch ausstechen, wie dies bei tokenfrequenten und daher stark gefestigten Formen der Fall wäre. Daher werden bei Verben, die ursprünglich der starken Klasse angehören, schwache Formen möglich und durch Nutzung gefestigt. Zudem ist beim Wechsel von starken Verben zur schwachen Konjugation ein Zwischenschritt möglich: Tokeninfrequente starke Verben, die nach einem typeninfrequenten Ablautmuster flektieren, können sich dem typenfrequenten Muster *x-o-o* anschließen. Auf diese Weise kann es zu einem doppelten Anschluss an Typenfrequenz kommen: zunächst an das typenfrequente Flexionsverhalten innerhalb der typeninfrequenten Klasse (*x-o-o*) und bei zu geringer Tokenfrequenz schließlich an die typenfrequente schwache Flexionsklasse (*der Hund ball > boll > bellte*) (Nowak 2013: 178–183).

Das typen- und tokenfrequenzbasierte statistische Vorkaufsrecht kann in der Variation zwischen starken und schwachen Verben und in der Variation zwischen schwachen und starken Maskulina beobachtet werden (siehe Abschnitt 3.1.1 und 3.2.1). Bei der Variation zwischen *haben* und *sein* ist dies nicht der Fall, da es sich hierbei um zwei Schemata handelt, die dem Makroschema [X_{NP} X_{VAFIN} *ge-V-t/en*] und [+Perfekt] angehören: Die Form *sein* ist mit der Funktion [– Transitivity, + Telicity oder + Bewegungssemantik] assoziiert, die Form *haben* dagegen mit der Funktion [+Transitivity, – Telicity, – Bewegungs-

semantik] (Gillmann 2016: 316–319, siehe Abschnitt 3.3). Hier ist also nicht eine Funktion mit Formen verknüpft, die sich in der Typenfrequenz unterscheiden, sondern zwei separate Schemata liegen vor, die durch prototypisch organisierte Funktionen verbunden sind. Daher greift anders als bei der Variation in der Flexion die Kombination aus typen- und tokenfrequenzbasiertem statistischem Vorkaufsrecht nicht (Goldberg 2019: 74–94). Es ist allenfalls möglich, dass Konstruktionen, die sich im Übergangsbereich zwischen den Funktionen befinden, aufgrund der Typenfrequenz zu *haben* neigen. Das könnte bspw. bei *das Pferd geritten haben/sein* der Fall sein, denn hier kann das Objekt als Patiens oder Instrument interpretiert werden. Das war im Antwortverhalten der Prototypizitätsstudie jedoch nicht zu beobachten. Stattdessen scheinen Bewegungsverben so stark mit der *sein*-Selektion assoziiert zu sein, dass diese Assoziation auch in ambigen Fällen greift. Unabhängig von der Typenfrequenz spielt Tokenfrequenz aber eine Rolle in der Auxiliarselektion: So weisen bspw. tokenfrequente Bewegungsverben eine höhere Stabilität in der Auxiliarselektion auf als tokeninfrequente (Gillmann 2016: 272–274, siehe Abschnitt 3.3.1). Auch das Antwortverhalten in der Prototypizitätsstudie weist auf einen Tokenfrequenzeffekt hin, da das vergleichsweise wenig frequente Bewegungsverb *reiten* mit auffällig vielen *sein*-Antworten in transitiven Sätzen und einem hohen Anteil von *haben*-Antworten bei den ambigen Sätzen im Vergleich zu den anderen Testverben die höchste Variation im Antwortverhalten zeigt.

Tokenfrequenz wurde in Abschnitt 2.1.2 als ein guter Prädiktor für die Schnelligkeit und Genauigkeit vorgestellt, mit der Proband_innen auf Stimuli in psycholinguistischen Studien reagieren (Divjak & Caldwell-Harris 2015: 55). Allerdings stellt Tokenfrequenz als pure Wortwiederholung eine Vereinfachung dar, da bspw. der Kontext in dieser Operationalisierung nicht berücksichtigt wird. Zudem haben viele Faktoren, die mit Frequenz korrelieren, wie bspw. Abfolgewahrscheinlichkeiten und Kontextdiversität, ebenfalls Einfluss auf die Prozessierung (McDonald & Shillcock 2003, Baayen 2012). Es stellt sich daher die Frage, ob Frequenz selbst einen umfassenden Einfluss auf Prozessierung und das Verhalten von Proband_innen nimmt oder ob der Einfluss besser durch einen Faktor erklärt werden kann, der mit Frequenz korreliert. Die Berücksichtigung von Kontexten, die bspw. durch den Blick auf Kontextdiversität möglich ist, scheint die Varianz im Verhalten von Proband_innen dabei besser zu erklären als Frequenz im Sinne einer puren Wortwiederholung (Baayen 2012: 181–191). Tokenfrequenz bietet dennoch einen schnellen und einfachen Zugriff auf die Varianz im Verhalten von Proband_innen. Man sollte sich aber über die Korrelation mit anderen Faktoren bewusst sein, die möglicherweise die Varianz besser erklären als Tokenfrequenz.

Daher wurden Typen- und Tokenfrequenz in Abschnitt 2.1.3 als Teil des statistischen Lernens in ein bayesianisches Modell integriert.

Bayesianische Modelle arbeiten mit bedingten Wahrscheinlichkeiten und berechnen die Wahrscheinlichkeit für ein Ereignis (Norris 2006: 330), z. B.: Wenn ein bestimmter Input gegeben ist, welche mentale Repräsentation passt zum Input? Diese Frage stellen sich Sprecher_innen in der Sprachverarbeitung. Wenn Sprecher_innen ein Wort lesen oder hören, gleichen sie das Wort mit mentalen Repräsentationen ab und überprüfen, welche Repräsentation am besten zum Input passt. Ist der Input nicht gut zu verstehen, noch nicht ganz gesprochen bzw. gebärdet oder unleserlich geschrieben, sind mehrere mentale Repräsentationen möglich. Um herauszufinden, welche Repräsentation zum Input passt, greifen Sprecher_innen auf die priore Wahrscheinlichkeit zurück. Frequenz ist Teil der prioren Wahrscheinlichkeit: Hört man [ba:], ist die Wahrscheinlichkeit höher, dass [lt] folgt als [jezja:nɪf], weil *bald* frequenter ist als die Wortform *bayesianisch*. Im Kontext dieser Dissertation ist das Gegenteil der Fall: *Bayesianisch* wird häufiger genutzt als *bald*, somit ist beim Wortanfang <ba> die Wahrscheinlichkeit für <yesianisch> höher als für <ld>. Auch der Kontext kann somit Teil der prioren Wahrscheinlichkeit sein. Frequenz nimmt in diesem Modell daher zusammen mit anderen Faktoren (bspw. Kontext und Abfolgewahrscheinlichkeiten) Einfluss auf die priore Wahrscheinlichkeit (Norris 2006: 330–331). Genauso beeinflusst Frequenz das statistische Vorkaufsrecht (Goldberg 2019: 75–77): Existieren zwei Formen mit derselben Funktion, sind zunächst beide Formen möglich. Ist eine der Formen tokenfrequenter als die andere, hat diese aufgrund der höheren Frequenz eine größere priore Wahrscheinlichkeit genutzt zu werden und sticht die andere Form daher statistisch aus.

Im empirischen Teil der Arbeit ist der Einfluss von Tokenfrequenz deutlich zu erkennen (siehe Abschnitt 5.2). Die frequenten starken Verben evozieren deutlich geringere Reaktionszeiten als die infrequenten starken Verben. Dies war jeweils für die starken (*gezogen*) als auch für die schwachen Formen (*gezieht*) zu beobachten. Die Ergebnisse stützen das Prinzip des statistischen Vorkaufsrechts: Schwache Formen sind bei frequenten starken Verben kaum gefestigt. Da aber eine tokenfrequente Alternative existiert, können die Formen schnell als unbekannt bewertet werden. Daher werden die schwachen Formen von frequenten Verben schneller beurteilt als die schwachen Formen von infrequenten Verben. Die schwachen Formen werden bei den frequenten Verben erwartbarerweise langsamer beurteilt als die starken, da die starken Formen selbst mental gefestigt sind und Proband_innen anders als bei den schwachen Formen in der Bewertung nicht auf die mentale Repräsentation einer anderen stark gefestigten Form zurückgreifen müssen. Die schwachen Formen der frequenten Verben

6 Zusammenfassung und Ausblick

weisen dabei vergleichbare Reaktionszeiten wie die starken Formen der infrequenten Verben auf.

Innerhalb der infrequenten starken Verben zeigt sich, dass das Vorkommen von schwachen Formen im Sprachgebrauch keinen Einfluss auf die Verarbeitungsgeschwindigkeit hat. Sowohl die Verben ohne attestierte Schwankung in Korpora als auch die Verben mit attestierter Schwankung weisen vergleichbare Reaktionszeiten auf (zur Operationalisierung von attestierter und nicht-attestierter Schwankung siehe Abschnitt 4.2). Dies gibt einen Hinweis darauf, dass die starken Formen infrequenter Verben bereits nicht mehr so stark gefestigt sind wie bei den frequenten Verben. Daher können sie die schwachen Formen weniger stark dominieren. Die ausbleibenden Reaktionszeitunterschiede zwischen Verben mit und ohne Schwankung weisen damit auf eine mögliche Variation bei Verben ohne attestierte Schwankung hin, die im Sprachgebrauch noch nicht sichtbar ist.

Nur ein kleiner Unterschied in den Reaktionszeiten lässt sich zwischen den Verben mit und ohne Schwankung erkennen: Die starken Formen von Verben ohne Schwankung werden wie bei den frequenten Verben schneller bewertet als schwache Formen. Bei den Verben mit Schwankung besteht hingegen kein Vorteil für starke Formen. Die schwachen Formen scheinen durch den Gebrauch bereits gefestigt zu sein und können daher schneller bewertet werden.

Es zeigen sich zwei aufeinander aufbauende Schritte: Im ersten Schritt beeinflusst das statistische Vorkaufsrecht, wie sehr die starken Formen die schwachen dominieren. Wenn die starken Formen die schwachen stark dominieren, ist Variation unwahrscheinlich. Wenn die starken Formen aufgrund geringer Tokenfrequenz die schwachen nicht mehr so stark dominieren können, ermöglicht dies den Gebrauch schwacher Formen, sodass Variation wahrscheinlich wird. Die schwachen Formen werden im zweiten Schritt durch Gebrauch gefestigt und daher schneller verarbeitet. Der erste Schritt zu Variation lässt sich bereits in Reaktionszeiten fassen, in Korpora wird Variation erst im zweiten Schritt sichtbar.

Für anschließende Studien wäre es sinnvoll, mit weiteren starken Verben zu arbeiten und die Operationalisierung von in Korpora attestierter und nicht-attestierter Schwankung zu verfeinern, indem bspw. mit feingliedrigeren Abstufungen in der Ratio zwischen starken und schwachen Formen gearbeitet wird und eine mögliche regionale Schwankung der Verben vorab systematischer überprüft wird. Hierbei wäre es auch wichtig, Faktoren wie Kontextdiversität zu berücksichtigen, die mit Frequenz korrelieren. Zudem wäre es interessant, mit feineren Frequenzabstufungen zu arbeiten, um auszuloten, ab welcher Tokenfrequenz das statistische Vorkaufsrecht für starke Formen nur noch bedingt greift.

6.2 Prototypizität

Prototypizität erwächst aus exemplarbasiertem Lernen und ist daher eng mit Frequenz verknüpft: Prototypische Vertreter einer Kategorie sind tokenfrequent und außerdem teilen sie viele Eigenschaften mit anderen Vertretern der Kategorie, daher sind ihre Eigenschaften typenfrequent (Ellis u. a. 2014: 85, Taylor 2015: 564). Zudem weisen prototypische Vertreter nur wenige Eigenschaften auf, die Vertreter anderer Kategorien zeigen, sodass den Eigenschaften von prototypischen Vertretern eine hohe *cue validity* zukommt (siehe Abschnitt 2.2.1). Prototypizität ergibt sich daher aus Token- und Typenfrequenz. Aufgrund dieser engen Verknüpfung von Prototyp und Frequenz werden prototypisch organisierte Kategorien anders als in der Standardversion der Prototypensemantik, die in Abschnitt 2.2.1 vorgestellt wird, in Abschnitt 2.2.3 als Ergebnis exemplarbasierten Lernens modelliert: Elemente mit ähnlichen Eigenschaften werden als Cluster gefasst, dabei bilden die häufigen Elemente einen Ankerpunkt, um den sich die anderen Elemente gruppieren (Goldberg 2006: 89, Ellis 2016: 242, Goldberg 2019: 51–73). Die Kategorie wird über das Prinzip der Abdeckung erweitert. Das Resultat dieses exemplarbasierten Lernens ist eine Kategorie mit prototypischen und peripheren Mitgliedern (Ross & Makin 1999: 216–222).

Dass frequente prototypische Vertreter als Ankerpunkt von Kategorien fungieren, zeigt sich in Abschnitt 2.2.2 bspw. darin, dass prototypische Vertreter einer Konstruktion als Erstes erlernt werden und Kategorien mit frequenten prototypischen Vertretern schneller erlernt werden als Kategorien mit weniger frequenten prototypischen Vertretern (Goldberg 2006: 198–200, Ellis & Ferreira-Junior 2009, Ellis u. a. 2014: 58). Prototypische Vertreter einer Kategorie werden aufgrund ihrer Tokenfrequenz und der vielen Eigenschaften, die sie mit anderen Mitgliedern der Kategorie teilen, immer wieder gefestigt (Ellis u. a. 2014: 85). Sie sind daher in ihrer Verwendung stabil, sodass Variation hier unwahrscheinlich ist. Bei peripheren Vertretern ist Variation hingegen wahrscheinlich (Ágel 2008: 67–68). Dies liegt auch daran, dass die Peripherie den Übergangsbereich zwischen zwei Clustern darstellen kann. Das ist bspw. bei der Auxiliarselektion von *haben* und *sein* der Fall: Die beiden Schemata sind durch prototypisch organisierte Funktionen miteinander verbunden, dabei sind im Übergangsbereich zwischen den Funktionen beide Auxiliare möglich (siehe Abschnitt 3.3.2).

Bei Flexionsklassen lässt sich das Prinzip der Prototypizität nutzen, um das Flexionsverhalten prototypisch zu modellieren: Verben und Substantive lassen sich anhand ihres Verhaltens in prototypisch stark und prototypisch schwach teilen (siehe Abschnitt 3.1.2 und 3.2.2). Die gemischte Deklination der Maskulina erscheint in dieser Betrachtung als Übergangsbereich zwischen der schwachen und

starken Deklinationsklasse. Zudem lassen sich auch prototypische und periphere Eigenschaften der typeninrequenten Klasse (starke Verben/schwache Maskulina) ableiten. Die Eigenschaften, die nur die prototypischen Vertreter der typeninrequenten Klasse aufweisen (z. B. Imperativhebung, Genitivprofilierung), sind periphere Eigenschaften, prototypische Eigenschaften zeigen hingegen auch die peripheren Mitglieder der infrequenten Klasse (z. B. Ablaut, Plural auf *-n*) (Bittner 1985: 57, Nowak 2013: 157). Daraus ergibt sich, dass viele Mitglieder einer Flexionsklasse die prototypischen Merkmale der Klasse aufweisen und nur wenige die peripheren Eigenschaften. Den peripheren Flexionseigenschaften kommt außerdem i. d. R. eine geringe intraparadigmatische Frequenz und eine geringere Relevanz nach Bybee (1985) zu als den prototypischen Eigenschaften. Bei einem Klassenwechsel werden die peripheren Eigenschaften daher zuerst aufgegeben und die prototypischen Eigenschaften zuletzt. Die prototypische Modellierung der Flexionsklassen lässt sich somit nutzen, um Schwankungen von der typeninrequenten zur typenfrequenten Klasse zu prognostizieren. So ist eine Aufgabe der prototypischen Flexionseigenschaften unwahrscheinlich, so lange die peripheren Eigenschaften noch stabil im Sprachgebrauch zu beobachten sind.

Der Einfluss der Prototypizität auf Variation zeigt sich im empirischen Teil der Arbeit nur bedingt (siehe Abschnitt 5.3). Die prototypischen Vertreter des *haben*- und *sein*-Schemas evozieren jeweils das erwartbare Auxiliar, aber in der Peripherie ist kaum Schwankung zu beobachten, denn in beiden ambigen Konditionen¹ überwiegen *sein*-Antworten (siehe Abschnitt 4.3 für ausführliche Erläuterungen zum Untersuchungsdesign). Im Antwortverhalten der Studie zu Form-Schematizität zeigt sich jedoch Variation in der Peripherie des Form-Schemas. Hinsichtlich der Reaktionszeiten lässt sich der Einfluss von Prototypizität ansatzweise feststellen: In der Prototypizitätsstudie wurde mit den intransitiven Sätzen zumindest einer der Prototypen schneller verarbeitet als die anderen Ausprägungen. Dabei deuten die erhöhten Lesezeiten bei den transitiven Sätzen auch auf das Wirken der Prototypizität hin: Durch die häufige Verwendung in intransitiven Sätzen sind Bewegungsverben so stark mit *sein* assoziiert, dass der Wechsel zu *haben* in transitiven Sätzen einen Prozessierungsaufwand darstellt, der sich in höheren Reaktionszeiten niederschlagen könnte. Die generelle Tendenz von Bewegungsverben zu *sein* könnte auch die klare *sein*-Präferenz der Proband_innen in den Sätzen im Übergangsbereich zwischen den beiden Schemata erklären.

Wie die transitiven Sätze weisen die Sätze in der Peripherie, AMBIG I und II, erhöhte Reaktionszeiten auf. Die erhöhten Reaktionszeiten sind aufgrund der vor-

¹Die ambigen Sätze enthielten ein Objekt, das entweder Patiens oder Instrument sein kann (AMBIG I: *ein Cabrio fahren*), bzw. eine Akkusativergänzung, die als Adverbial interpretiert werden kann (AMBIG II: *ein Rennen fahren*).

wiegenden *sein*-Wahl nicht mit dem Wechsel zu *haben* zu erklären und könnten auf Variationspotential hindeuten, da in den ambigen Sätzen theoretisch beide Auxiliare möglich sind.

In zukünftigen Studien sollte das Wirken der Schemata dahingehend überprüft werden, ob transitive Sätze ohne Bewegungsverb tatsächlich geringere Reaktionszeiten hervorrufen als transitive Sätze mit Bewegungsverb. Die vorliegende Studie gibt einen Hinweis darauf, der jedoch systematisch überprüft werden müsste. Zudem wäre in Bezug auf einzelne Bewegungsverben eine genauere Analyse der Auxiliarselektion bei *reiten* lohnend, das sich im Antwortverhalten von den anderen Testverben abhebt. Neben der prototypisch organisierten Transitivität wäre es lohnenswert, die ebenfalls prototypisch organisierten Einflussfaktoren Telizität und Bewegungssemantik auf die Auxiliarselektion in den Blick zu nehmen. In Bezug auf Telizität wäre es sinnvoll, *degree achievements* näher zu untersuchen. Hierbei könnte man telische Verben (*die Hefe ist/*hat verschimmelt*) mit *degree achievements* (*die Hefe ist/hat geschimmelt*) vergleichen sowie *degree achievements* in Sätzen mit telischer Lesart (*Das Messer ist/*hat von oben bis unten gerostet*) und atelischer Lesart (*Das Messer hat/*ist jahrelang vor sich hin gerostet*) gegenüberstellen. In Bezug auf Bewegungssemantik wäre ein Blick auf kontextuell konstruierte Bewegungssemantik (*Ich bin/*habe durch den Saal getanzt* vs. *Ich habe/*bin immer gern getanzt*) lohnenswert.

Zudem sollte die Prototypizitätsskala von starken und schwachen Verben sowie Maskulina psycholinguistisch überprüft werden: Hierfür könnte man bspw. Verben mit Imperativhebung nutzen (z. B. *geben*) und diese mit und ohne Imperativhebung (*gib/geb*) sowie mit starken und schwachen Präteritalformen (*gab/*gebte*) präsentieren und überprüfen, ob die Schwankung im Imperativ geringere Reaktionszeitunterschiede hervorruft als die Variation in der Präteritalform.

6.3 (Form-)Schematizität

Wie Prototypizität ergibt sich (Form-)Schematizität aus Frequenz: Wenn eine Konstruktion typenfrequente und hochvariable Slots aufweist, können diese abstrahiert und damit schematisiert werden. Das Ergebnis ist ein Schema (Ellis 2002a: 168–170, Bybee 2010: 80). Die vorliegende Arbeit führt die Termini *Form-Schematizität* und *Form-Schema* ein, um die Variation in der Flexion modellieren zu können: In Hinblick auf Flexionsklassen liegt Form-Schematizität vor, da eine Funktion nicht nur mit einer, sondern mit mehreren Formen assoziiert ist (siehe Abschnitt 2.3.2). Dabei weist eine der Formen eine hohe Typenfrequenz

und eine hohe Variabilität zwischen den Mitgliedern auf, weswegen sie generalisierbar ist. Die anderen Formen haben dagegen eine geringere Typenfrequenz und geringe Variabilität zwischen den Mitgliedern und sind daher nur eingeschränkt generalisierbar (Bybee 2010: 69, Goldberg 2019: 64–65). In der Folge sind die typeninfrequenten Formen nicht-schematisch oder teilschematisch. In der vorliegenden Arbeit werden Assoziationen aus einer Funktion und einer teilschematischen Form als Form-Schemata bezeichnet, wenn Form-Schematizität vorliegt, also eine Funktion und mehrere Formen mit unterschiedlichen Schematizitätsgraden miteinander verknüpft sind. Form-Schemata nehmen Einfluss auf Variation, wie im weiteren Verlauf des Abschnitts erläutert wird.

Das Prinzip der Form-Schematizität ist in der Deklination und in der Konjugation zu beobachten und lässt sich anhand der Konjugation am besten veranschaulichen, da hier der typenfrequenten Form viele typeninfrequente Formen gegenüberstehen. Eine typeninfrequente Form kann nur für ein Verb greifen (z. B. *sein*) oder für mehrere Verben mit ähnlichen phonologischen Eigenschaften (z. B. Verben der Ablautreihe 3a). Aufgrund der Ähnlichkeit in der Phonologie bilden die Verben ein Cluster, sodass die ähnlichen phonologischen Eigenschaften abstrahiert werden können (z. B. [#_I + η + (C)]), da sie häufig vorkommen. Ein Form-Schema entsteht. Hierbei handelt es sich streng genommen nicht um ein einzelnes, sondern um mehrere ineinander verschachtelte Form-Schema: Die Form *trank* ist mit [+Vergangenheit] verbunden, aufgrund dieser Verbindung und der Verbindung bei anderen Form-Funktions-Paaren (z. B. *getrunken* und [+Partizip II]) ist das Verb *trinken* mit starker Flexion assoziiert. Auf einer abstrakteren Ebene ist die Form [#_I + η + (C)] mit starkem Flexionsverhalten assoziiert, da mehrere Verben mit ähnlichen Eigenschaften existieren, die [+Vergangenheit] durch den Ablaut auf /a/ ausdrücken.

Prototypizität ist Form-Schemata eingeschrieben, da die Assoziation zwischen teilschematischer Form und Funktion probabilistisch ist (siehe Abschnitt 2.3.1): Je mehr Eigenschaften der Form des Form-Schemas erfüllt werden, desto wahrscheinlicher ist eine Assoziation mit der Funktion (Rumelhart 1980: 39, Bybee & Moder 1983: 262–264). Es existieren daher prototypische und periphere Mitglieder eines Form-Schemas: Prototypische Vertreter des Form-Schemas [#_I + η + (C)] sind bspw. *trinken* und *senken*, periphere hingegen *sinnen* und *rinnen*, da der Nasal vorne und nicht hinten im Mundraum gebildet wird. Form-Schematizität ist somit sowohl mit Frequenz als auch mit Prototypizität eng verbunden.

Form-Schemata sind typisch für eine typeninfrequente Flexionsklasse, aber auch in der variablen typenfrequenten Klasse können teilschematische Subcluster entstehen, die dann ein Form-Schema bilden, wie bspw. Maskulina auf *-er* (*Angler*), die mit starker Deklination assoziiert sind (Köpcke 2000b: 160–161). Für

Variation sind vorrangig Form-Schemata in typeninfrequenten Klassen relevant: Durch Form-Schemata besteht eine Typenfrequenz innerhalb von typeninfrequenten Klassen, denn bestimmte Eigenschaften kommen häufig vor und lassen sich daher, wenn auch nur bedingt, abstrahieren. Als typenfrequentes Cluster führen Form-Schemata zu Stabilität (siehe Abschnitt 2.3.3). So variieren bspw. Verben, die einem Form-Schema angehören, weniger, auch wenn sie genauso tokeninfrequent sind wie Verben, die keinem Form-Schema angehören (Köpcke 1999: 57–58). Die Flexion wird gestützt, da die Verben einem teilschematischen Flexionsmuster folgen. Für Verben wie *trinken*, die einem Form-Schema angehören, ist Variation also unwahrscheinlich. Variation ist bei Verben möglich, die Form-Schemata nur teilweise entsprechen wie bspw. *sinnen*. Form-Schemata führen somit nicht nur zu Stabilität, sondern auch zu Variation. Sie können außerdem Variation hin zur tokeninfrequenten Klasse auslösen: Verben aus der schwachen Flexionsklasse, die in ihren phonologischen Eigenschaften einem Form-Schema ähneln, schwanken zur starken Flexion (*gewinkt* > *gewunken*) (Köpcke 1999: 56).

Form-Schemata können unterschiedliche Schematizitätsgrade aufweisen: Das Form-Schema der Verben der Ablautreihe 3a ist mit der Form [#_I + η + (C)] relativ konkret, das Form-Schema *x-o-o* dagegen weitaus abstrakter (siehe hierzu Abschnitt 3.1.3). Das Form-Schema *x-o-o* springt bei geringer Tokenfrequenz und geringer Passgenauigkeit zu konkreteren Form-Schemata ein und stützt so die starke Konjugation.

Wie bereits deutlich wurde, können innerhalb einer typeninfrequenten Flexionsklasse mehrere Form-Schemata als typenfrequente Cluster existieren, dies ist bspw. bei starken Verben der Fall. Form-Schemata können aber auch eine gesamte Flexionsklasse strukturieren, wie bei den schwachen Maskulina.² Das Form-Schema schwacher Maskulina ist weit komplexer als die Form-Schemata starker Verben, da es neben prosodisch-phonologischen Eigenschaften (Dreisilblichkeit, Pänultimabetonung, Schwa) auch eine semantische Eigenschaft (hoher Belebtheitsgrad) aufweist. Substantive wie *Kollege*, die dem prosodisch-phonologischen und dem semantischen Prototyp des Form-Schemas entsprechen, flektieren stabil schwach, für sie ist Variation unwahrscheinlich. Dasselbe gilt für Substantive wie *Stein*, die weder den prosodisch-phonologischen noch den semantischen Eigenschaften des Form-Schemas entsprechen: Aufgrund der fehlenden prosodisch-phonologischen und semantischen Eigenschaften ähneln sie den schwachen Substantiven so wenig, dass sie nicht als Teil der schwachen Flexionsklasse wahrge-

²Es existieren zwar zwei Form-Schemata, die die schwachen Maskulina prägen, jedoch sind sie in ihren Eigenschaften sehr ähnlich. Dies ist bei den starken Verben nicht der Fall.

nommen werden. Variation ist jeweils in der Kombination aus Prototyp und Peripherie wahrscheinlich: Monosyllabische Substantive mit hohem Belebtheitsgrad (formale Peripherie, semantischer Prototyp) und zwei- und dreisilbige Substantive mit niedrigem Belebtheitsgrad (formaler Prototyp, semantische Peripherie) schwanken zwischen schwachen und starken Formen oder wechseln das Genus (*des Grafen/Grafs, des Willen/Willens, der Krake/die Krake*) (siehe Abschnitt 3.2.3). Das Form-Schema nimmt dabei auch Einfluss auf die Art der Variation: Zwei- und dreisilbige Substantive auf Schwa mit geringem Belebtheitsgrad schließen sich der starken Flexion an, indem sie *-ns* im Genitiv annehmen (*des Willen > des Willens*) und schließlich *-n* im Nominativ (*der Wille > der Willen*). Zweisilbigen Substantiven steht zudem die Möglichkeit offen, das Genus zu wechseln (*der Krake > die Krake*) (Köpcke 2000a: 108–109). Einsilbige Substantive mit hohem Belebtheitsgrad weisen hingegen zunächst endungslose Dativ- und Akkusativformen auf (*den Grafen > den GrafØ*) und wechseln dann das Genitivsuffix zu *-s* (*des Grafen > des Grafs*) (Köpcke 2000a: 108–109).

Im empirischen Teil der Arbeit zeigt sich der Einfluss von Form-Schemata nur im Antwortverhalten (siehe Abschnitt 5.4). Für Substantive, die dem Prototyp des Form-Schemas entsprechen, werden schwache Formen klar bevorzugt: Im Produktionsexperiment zur self-paced-reading-Studie werden für diese Substantive vorrangig schwache Formen gebildet, in der lexical-decision-Studie werden schwache Formen als bekannt und starke als unbekannt bewertet und in der sentence-maze-Studie vorrangig schwache Formen gewählt. Dies gilt für real existierende, aber auch für Pseudosubstantive. Die schwachen Formen bei Pseudosubstantiven verdeutlichen, dass das Form-Schema produktiv ist. Für Substantive, die zur starken Flexion gehören, wurden erwartbarerweise starke Formen genutzt. Substantive in der Peripherie des Form-Schemas weisen mit Ausnahme der sentence-maze-Studie Variation auf: Im Produktionsexperiment werden sowohl schwache als auch starke Formen gewählt, in der lexical-decision-Studie werden starke Formen zu 50 % als bekannt bewertet. Die ausbleibende Variation in der sentence-maze-Studie kann mit dem Versuchsdesign erklärt werden, da die häufigere Form direkt neben der selteneren präsentiert wurde und die häufige die seltene Form somit leichter statistisch ausstechen kann. In den Reaktionszeiten ist hingegen kein systematischer Einfluss der Form-Schematizität zu erkennen. In der self-paced-reading-Studie liegen die Unterschiede zwischen den Reaktionszeiten für die schwache und die starke Form von *Schettose* zwar unter dem α -Level, die 95 %-Konfidenzintervalle überlappen jedoch. In der visuellen Analyse der Reaktionszeiten der sentence-maze-Studie sind Einflüsse von Tokenfrequenz nicht auszuschließen, obwohl die Tokenfrequenz der Testsubstantive in der Studie (siehe Abschnitt 4.5) möglichst konstant gehalten wurde. Die Unter-

schiede in der Reaktionszeit scheinen aber auch in Hinblick auf Tokenfrequenz nicht systematisch zu sein.

In weiteren Untersuchungen sollte der Diskrepanz zwischen Antwortverhalten und Reaktionszeiten nachgegangen werden. Dabei ist es möglich, dass die in den Studien verwendeten Verfahren zu wenig sensitiv für die Messung von Prozessierungsunterschieden sind und daher auf feinere Methoden wie EEG zurückgegriffen werden sollte. Da der Einfluss von Form-Schematizität im Antwortverhalten zu erkennen ist, ist es unwahrscheinlich, dass er sich überhaupt nicht auf die Prozessierung auswirkt. Neben der Verwendung von feineren Messmethoden wäre es interessant, die Interaktion von Tokenfrequenz und Form-Schemata näher zu betrachten und Form-Schemata bei starken Verben psycholinguistisch zu untersuchen. Viele Unterschiede in den Flexionseigenschaften von starken und schwachen Verben sind salienter als die Unterschiede in den Flexionseigenschaften von starken und schwachen Maskulina, sodass mögliche Prozessierungsunterschiede auch in Reaktionszeiten sichtbar sein könnten. Es wäre sinnvoll, mit Verben der Ablautreihe 3a zu arbeiten, da diese eine klar umrissene phonologische Form aufweisen. Hierbei bietet sich an, Reaktionszeiten von Verben der Ablautreihe 3a mit unterschiedlicher Tokenfrequenz (z. B. *singen* vs. *wringen*) zu messen und diese mit Reaktionszeiten von Verben zu vergleichen, die keinem Form-Schema angehören und jeweils eine ähnliche Tokenfrequenz aufweisen.

Ein weiterer Faktor, der in der Arbeit ausgeklammert wurde, ist der Einflussfaktor des Mediums. Alle Experimente wurden visuell durchgeführt. Es wäre interessant, die Ergebnisse mit auditiven Stimuli zu kontrastieren. Dies gilt für alle Einflussfaktoren und Variationsphänomene, aber insbesondere für die Form-Schemata. Da diese phonologisch konditioniert sind, ist es möglich, dass sie einen größeren Einfluss auf Variation haben, wenn sie auditiv präsentiert werden, als nur mittelbar durch Schrift.

6.4 Gewichtung der Einflussfaktoren

Da Typen- und Tokenfrequenz die Existenz von Prototypizität und (Form-)Schematizität bedingen, ist der Einfluss von Frequenz auf Variation am stärksten zu gewichten. Gleichzeitig korreliert Frequenz am stärksten mit anderen Einflussfaktoren, sodass der Blick auf Frequenz allein eine Vereinfachung darstellen kann. Um diesem Problem zu begegnen, wird der Einfluss von Frequenz als Teil der prioren Wahrscheinlichkeit für das Vorkommen einer Struktur in einem bayesianischen Modell gefasst. Diese Modellierung schließt den Einfluss weiterer Faktoren nicht aus und kann den Einfluss der Frequenz auf Prozessierung erklären.

Prototypizität ergibt sich aus dem Einfluss von Frequenz: Prototypische Vertreter einer Kategorie haben eine hohe Tokenfrequenz und teilen viele Eigenschaften miteinander. Hieraus ergibt sich eine Typenfrequenz. Beides stützt die prototypischen Vertreter. (Form-)Schematizität ergibt sich aus Frequenz, da Typenfrequenz die Abstraktion von Mustern ermöglicht. Prototypizität und Form-Schematizität sind dabei eng miteinander verwandt. Prototypizität fokussiert auf die Struktur einer Kategorie und betrachtet, wie die Kategorie über periphere Mitglieder der Kategorie mit anderen Kategorien verbunden ist. Kategorien ergeben sich aus wiederkehrenden Eigenschaften von Konstruktionen und lassen sich somit wie Konstruktionen als Assoziationen aus Form und Funktion fassen. Bei Form-Schematizität ist eine Funktion nicht nur mit einer, sondern mit mehreren Formen verbunden. Hier wird ebenfalls die prototypisch organisierte Verbindung zwischen Form und Funktion betrachtet, nur in Bezug auf die Assoziation einer Funktion mit mehreren Formen. Trotz dieser engen Verknüpfung ist die Unterscheidung zwischen separaten Schemata, deren Funktionen prototypisch miteinander verbunden sind, und der in dieser Arbeit ausgearbeiteten Form-Schematizität, bei der eine Funktion mit mehreren Formen verbunden ist, essentiell. Bei Form-Schematizität spielt Typenfrequenz eine viel größere Rolle als bei separaten Schemata, da das statistische Vorkaufsrecht nur dann greift, wenn eine Funktion mit mehreren Formen assoziiert ist. Separate Schemata können zudem Elemente in ihre Slots zwingen: Auch wenn *tanzen* normalerweise keine Bewegung benennt, sondern eine Aktivität, selektiert *tanzen* das Auxiliar *sein* statt *haben*, wenn der Kontext eine Bewegung impliziert (*Ich bin durch den Raum getanzt*). Dieses Prinzip ist bei Form-Schemata ausgeschlossen, da keine separaten Funktionen vorliegen.

Trotz dieser Gewichtung ist nicht davon auszugehen, dass Frequenz die anderen Einflussfaktoren immer überlagert. Stattdessen handelt es sich um ein Zusammenspiel. Dies zeigt sich in Hinblick auf Typenfrequenz, Form-Schemata und Tokenfrequenz bei der Variation in der Deklination: Zunächst bedingt die Typenfrequenz die Schwankungsrichtung hin zu starken Maskulina, das Form-Schema bedingt die schwachen Substantive, die zur starken Flexion schwanken, und innerhalb der Peripherie des Form-Schemas bedingt die Tokenfrequenz, ob tatsächlich Variation stattfindet.

Der Einfluss von Frequenz, Prototypizität und (Form-)Schematizität auf Variation wird in allen Variationsphänomenen (Variation in der Konjugation, Deklination und Auxiliarselektion) sichtbar. Daher wäre es lohnenswert, den Einfluss der Faktoren auf andere Variationsphänomene zu untersuchen, z. B. auf Variation in der Kasusreaktion von Präpositionen oder auf Variation zwischen langen und kurzen Genitivendungen.

Empirisch ließ sich der Einfluss der Faktoren jeweils klar im Antwortverhalten der Proband_innen fassen, aber nur bedingt in den Reaktionszeitmessungen. Hierdurch werden die Herausforderungen sichtbar, die Reaktionszeitmessungen mit sich bringen. Der Einfluss von Frequenz auf Reaktionszeiten ist messbar, der Einfluss von Prototypizität jedoch aufgrund der starken Assoziation von Bewegungssemantik und *sein*-Selektion nur bedingt. Die Form-Schematizität ließ sich gar nicht in Reaktionszeitmessungen fassen.

Gleichzeitig verdeutlichen die Studien aber auch das Potential von Reaktionszeitmessungen: Im Fall der Verben ohne attestierte Schwankung in Korpora können Reaktionszeiten Hinweise auf Variation geben, auch wenn diese noch nicht bzw. kaum in der Sprachproduktion zu beobachten ist. Hinsichtlich der Selektion von *haben* und *sein* zeigen die erhöhten Reaktionszeiten bei transitiven Sätzen die feste Assoziation von *sein* mit Bewegungssemantik. Der empirische Teil der Dissertation unterstreicht, wie fundamental der Blick auf Prozessierung für die Modellierung von Variation und Grammatik aus einer gebrauchsbasierten Sicht ist. Erst mit Blick auf die Prozessierung zeigt sich das Variationspotential, das noch nicht im Sprachgebrauch sichtbar ist. Zudem zeigt Prozessierung, dass Prototypeneffekte sich nicht unbedingt in Stabilität äußern müssen, sondern komplexer sind, wie im Fall der erhöhten Reaktionszeiten bei transitiven Sätzen mit Bewegungsverben. Diese Erkenntnisse lassen sich nicht mit dem alleinigen Blick auf den Sprachgebrauch gewinnen. Die Hypothesen zu Variation, die auf Basis von Korpusdaten gemacht werden, sollten daher anhand von psycholinguistischen Studien systematisch überprüft werden. Reaktionszeitstudien bieten Einblicke in Variation, die eine bloße Betrachtung der Sprachproduktion nicht leisten kann. Auf diese Weise ermöglichen es Reaktionszeiten, Variation im Allgemeinen und Variationsphänomene im Speziellen besser zu verstehen.

Anhang A: Digitaler Anhang

Der digitale Anhang der Dissertation kann über https://osf.io/6yxzt/?view_only=bdd5d42c44b8459e83c40af807c32e60 aufgerufen werden.

Anhang B: Ablautreihen des Deutschen

Tabelle B.1: Überblick über die Ablautreihen im Deutschen nach Bergmann u. a. (2016: 209)

1	<i>reiten</i>	<i>ritt</i>	<i>geritten</i>
/aɪ/ + C	/aɪ/	/ɪ/	/ɪ/
2	<i>bieten</i>	<i>bot</i>	<i>geboten</i>
/i:/ + C	/i:/	/o:/	/o:/
3a	<i>sinken</i>	<i>sank</i>	<i>gesunken</i>
/ɪ/ + Nasal + C	/ɪ/	/a/	/ʊ/
3b	<i>helfen</i>	<i>half</i>	<i>geholfen</i>
/ɛ/ + Liquid + C	/ɛ/	/a/	/ɔ/
4	<i>nehmen</i>	<i>nahm</i>	<i>genommen</i>
/e:/ + Nasal/Liquid	/e:/	/ɑ:/	/ɔ/
5	<i>geben</i>	<i>gab</i>	<i>gegeben</i>
/e:/ + C	/e:/	/ɑ:/	/e:/
6	<i>fahren</i>	<i>fuhr</i>	<i>gefahren</i>
/ɑ:/ + C	/ɑ:/	/u:/	/ɑ:/
7	<i>schlafen</i>	<i>schlief</i>	<i>geschlafen</i>
Vokal außer /e:/	X (im Bsp. /ɑ:/)	/i:/	X (im Bsp. /ɑ:/)

Tabelle B.1 orientiert sich an Bergmann u. a. (2016: 209). Die Ablautreihe 1 und 2 werden dort in zwei Subgruppen geteilt. Da die Subgruppen für die Dissertation nicht relevant sind, wurde in Tabelle B.1 auf die Trennung verzichtet. Die hier vorgestellten Ablautreihen lassen sich für das Neuhochdeutsche nicht mehr ansetzen, hier ist von ca. 40 Ablautalternanzen auszugehen (Nübling u. a. 2013: 285–286). In der Dissertation werden die Ablautreihen dennoch als Orientierung genutzt.

Anhang C: Frequenz der Präteritalformen in der Frequenzstudie

Tabelle C.1: Frequenz von schwachen und starken Präteritalformen der infrequenten Testverben

schwache Form	Belege	starke Form	Belege
<i>flocht</i>	1.327	<i>flechtete</i>	15
<i>sponn</i>	139	<i>spinnnte</i>	12
<i>schwoll an</i>	516	<i>schwellte an</i>	0
<i>schmolz</i>	8.281	<i>schmelzte</i>	20
<i>kniff</i>	1.704	<i>kneifte</i>	15
<i>focht</i>	4.729	<i>fechtete</i>	14
<i>molk</i>	161	<i>melkte</i>	110
<i>drosch</i>	7.980	<i>dreschte</i>	18
<i>hieb</i>	5.844	<i>haute</i>	5.526
<i>quoll</i>	3.605	<i>quellte</i>	8
<i>sog ein</i>	17	<i>saugte ein</i>	20
<i>wob</i>	1.080	<i>webte</i>	378
<i>glomm</i>	322	<i>glimmte</i>	291
<i>es gor</i>	2	<i>es gärte</i>	83
<i>sann</i>	1.677	<i>sinnte</i>	12
<i>sonn</i>	1.809		

Das Verb *salzen* ist nicht in Tabelle C.1 aufgeführt, da es das Präteritum ausschließlich schwach bildet. Für die Präteritalformen von *gären* wurde nach *es gor* gesucht, da die Suche nach *gor* zu viele Fehltreffer (z. B. *gor* als Variante zu *gar* wie in *gor ned*) enthält.

Anhang D: Lexical-decision-Studie

D.1 Aufbau der Blöcke in der lexical-decision-Studie

Die Abfolge der Blöcke A bis D ist in der Studie randomisiert. In Tabelle D.1 sind sie in alphabetischer Reihenfolge aufgeführt. *o.S.* steht für *ohne Schwankung*, *m.S.* für *mit Schwankung*. Die Tabelle gibt zudem Auskunft über die Position der Stimuli innerhalb der Blöcke. Die Spalte *Art* unterscheidet Testsubstantive und Testverben, zudem sind einige Items mit *Evaluation* und *Designtest* markiert. *Evaluation* benennt Items, die zur Evaluation der Konzentriertheit der Proband_innen genutzt wurden, sie stellen allesamt Filler dar. Auch die Testitems mit *Designtest* wurden hierfür genutzt, sie hätten aber zusätzlich als Items genutzt werden können, die testen, ob das Versuchdesign funktioniert. Dies war jedoch nicht nötig. Die Spalte *Ausprägung* benennt die Frequenz- bzw. Form-Schematizitätsausprägung, der die Testverben bzw. -substantive angehören.

Tabelle D.1: Blockdesign in der lexical-decision-Studie

Position im Block	Block	Stimulus	Art	Ausprägung
1	Beispiel	die Tische	Evaluation	
2	Beispiel	du gebst	Evaluation	
3	Beispiel	die Autoe	Evaluation	
4	Beispiel	du siehst	Evaluation	
1	Einführung	des Laptopes	Evaluation	
2	Einführung	des Transporteres	Evaluation	
3	Einführung	des Bodens	Evaluation	
4	Einführung	des Spiels	Evaluation	
5	Einführung	des Praktikumes	Evaluation	
1	A	des Kontoos	Evaluation	
2	A	getragen	Testverb	frequent
3	A	des Helds	Testsubstantiv	Peripherie
4	A	gefahren	Testverb	frequent
5	A	geflechtet	Testverb	infrequent o. S.

D Lexical-decision-Studie

Position im Block	Block	Stimulus	Art	Ausprägung
6	A	des Flügeles	Evaluation	
7	A	gezogen	Testverb	frequent
8	A	des Neffen	Testsubstantiv	Form-Schema
9	A	gesinnt	Testverb	infrequent m. S.
10	A	des Feinden	Testsubstantiv	stark
11	A	gegoren	Testverb	infrequent m. S.
12	A	des Zaren	Testsubstantiv	Peripherie
13	A	gebracht	Designtest	
14	A	gequollen	Testverb	infrequent m. S.
15	A	des Kollegen	Testsubstantiv	Form-Schema
16	A	gedroschen	Testverb	infrequent o. S.
17	A	des Franzoses	Testsubstantiv	Form-Schema
18	A	gemelkt	Testverb	infrequent o. S.
19	A	des Fürsts	Testsubstantiv	Peripherie
20	A	des Angleres	Evaluation	
21	A	angeschwollen	Testverb	infrequent o. S.
22	A	des Ultimatumes	Designtest	
23	A	geschrieben	Testverb	frequent
24	A	des Kerls	Testsubstantiv	stark
25	A	gewebt	Testverb	infrequent m. S.
1	B	des Frühstückes	Evaluation	
2	B	gebracht	Designtest	
3	B	des Helden	Testsubstantiv	Peripherie
4	B	gefährten	Testverb	frequent
5	B	des Manteles	Evaluation	
6	B	geflochten	Testverb	infrequent o. S.
7	B	des Feindes	Testsubstantiv	stark
8	B	gesonnen	Testverb	infrequent m. S.
9	B	des Kerlen	Testsubstantiv	stark
10	B	gequellt	Testverb	infrequent m. S.
11	B	des Zars	Testsubstantiv	Peripherie
12	B	gedrescht	Testverb	infrequent o. S.
13	B	des Internetes	Designtest	
14	B	gemolken	Testverb	infrequent o. S.
15	B	gezieht	Testverb	frequent
16	B	des Kolleges	Testsubstantiv	Form-Schema
17	B	gegärt	Testverb	infrequent m. S.
18	B	des Neffes	Testsubstantiv	Form-Schema
19	B	gewoben	Testverb	infrequent m. S.
20	B	des Franzosen	Testsubstantiv	Form-Schema
21	B	angeschwellt	Testverb	infrequent o. S.

D.1 Aufbau der Blöcke in der lexical-decision-Studie

Position im Block	Block	Stimulus	Art	Ausprägung
22	B	des Fürsten	Testsubstantiv	Peripherie
23	B	geschrieben	Testverb	frequent
24	B	getragt	Testverb	frequent
25	B	des Wetteres	Evaluation	
1	C	des Abendes	Evaluation	
2	C	gespinnt	Testverb	infrequent o. S.
3	C	des Schützen	Testsubstantiv	Form-Schema
4	C	gesprochen	Testverb	frequent
5	C	des Orchesteres	Evaluation	
6	C	geflogen	Testverb	frequent
7	C	gegangen	Designstest	
8	C	des Geselles	Testsubstantiv	Form-Schema
9	C	geschmolzen	Testverb	infrequent o. S.
10	C	des Grafs	Testsubstantiv	Peripherie
11	C	gefechtet	Testverb	infrequent o. S.
12	C	des Vogten	Testsubstantiv	stark
13	C	ingesogen	Testverb	infrequent m. S.
14	C	des Löffeles	Evaluation	
15	C	gesalzt	Testverb	infrequent m. S.
16	C	des Ultimatum	Designstest	
17	C	gesinkt	Testverb	frequent
18	C	des Dieben	Testsubstantiv	stark
19	C	gekniffen	Testverb	infrequent o. S.
20	C	des Nachbarn	Testsubstantiv	Peripherie
21	C	gehaut	Testverb	infrequent m. S.
22	C	des Regenes	Evaluation	
23	C	gehalten	Testverb	frequent
24	C	geglommen	Testverb	infrequent m. S.
25	C	des Freundes	Testsubstantiv	stark
1	D	des Hebeles	Evaluation	
2	D	gesponnen	Testverb	infrequent o. S.
3	D	des Grafen	Testsubstantiv	Peripherie
4	D	gegeht	Designstest	
5	D	des Internets	Designstest	
6	D	gesunken	Testverb	frequent
7	D	des Vogts	Testsubstantiv	stark
8	D	gespracht	Testverb	frequent
9	D	gesalzen	Testverb	infrequent m. S.
10	D	des Nebeles	Evaluation	
11	D	geschmelzt	Testverb	infrequent o. S.

D Lexical-decision-Studie

Position im Block	Block	Stimulus	Art	Ausprägung
12	D	des Messeres	Evaluation	
13	D	eingesaugt	Testverb	infrequent m. S.
14	D	des Diebes	Testsubstantiv	stark
15	D	des Wagens	Evaluation	
16	D	gekneift	Testverb	infrequent o. S.
17	D	des Schützes	Testsubstantiv	Form-Schema
18	D	gehalten	Testverb	frequent
19	D	des Gesellen	Testsubstantiv	Form-Schema
20	D	gefliert	Testverb	frequent
21	D	des Nachbars	Testsubstantiv	Peripherie
22	D	gehauen	Testverb	infrequent m. S.
23	D	des Freunden	Testsubstantiv	stark
24	D	geglimmt	Testverb	infrequent m. S.
25	D	gefochten	Testverb	infrequent o. S.

D.2 Überblick über die Filleritems

Folgende Filleritems wurden genutzt:

- | | | |
|---------------------|------------------------|-------------------|
| (1) des Abendes | (12) des Manteles | (23) des Wetteres |
| (2) des Angleres | (13) des Messeres | (24) die Autoe |
| (3) des Bodens | (14) des Nebeles | (25) die Tische |
| (4) des Flügeles | (15) des Orchesteres | (26) du gebst |
| (5) des Frühstückes | (16) des Praktikumes | (27) du siehst |
| (6) des Hebeles | (17) des Regenes | (28) gebracht |
| (7) des Internetes | (18) des Spiels | (29) gebringt |
| (8) des Internets | (19) des Transporteres | (30) gegangen |
| (9) des Kontoes | (20) des Ultimatumes | (31) gegeht |
| (10) des Laptopes | (21) des Ultimatum | |
| (11) des Löffeles | (22) des Wagens | |

Anhang E: Sentence-maze-Studie

E.1 Aufbau der Blöcke in der sentence-maze-Studie

Die Experimentblöcke A bis D waren in der Studie in randomisiert. In Tabelle E.1 (S. 366f.) werden sie in tabellarischer Reihenfolge aufgeführt. Zudem wird der Beispiel- und der Einführungsblock gezeigt (in der Tabelle werden sie mit *Bsp.* und *Einf.* abgekürzt). Die Tabelle enthält neben den Testsätzen die Position der Testsätze im Block und benennt die Testitems sowie die Transitivitäts- bzw. Form-Schematizitätsausprägung der Testitems. Die senkrechten Striche (|) markieren die Abschnitte, in die die Sätze geteilt sind. Der obere Satz in der Tabelle besteht jeweils aus den zu erwartenden Möglichkeiten, der untere aus den unerwarteten. Bei Varianten sind selbstverständlich beide Möglichkeiten erwartbar, hier ist die Einteilung nur zufällig.

E.2 Überblick über die Füllsätze

Folgende Sätze wurden genutzt:

- (1) Ich wusste nicht, dass meine Tante durch die USA reisen will.
- (2) Ich weiß, dass der Vater die Kinder zur Schule gebracht hat.
- (3) Er geht spazieren, weil die Sonne scheint.
- (4) Ich habe gehört, dass mein Neffe seit gestern Fahrradfahren kann.
- (5) Er erzählt, dass sein Bruder in den USA herum reisen will.
- (6) Er teilte mir mit, dass der Kurier die Ware gestern morgen liefern sollte.

Tabelle E.1: Blockdesign in der sentence-maze-Studie

Block	Satz	Testitem
1 Bsp.	Erst wähle ich den Satzanfang, dann wähle ich die weiteren Wörter Stück für Stück aus, die den Satz am besten fortsetzen	
1 Bsp.	Er wähle ich den Satzanfang, trotz wehle du das weiteren Wörter Apfel für Apfel vor, dass den Zeichen am Besten vortsetzen	
1 Einf.	Ich wusste nicht, dass meine Tante durch die USA reisen will	
1 Einf.	Ich wuste nicht, weil meine Onkel durch den USA schwimmen sollen	
2 Einf.	Ich weiß, dass der Vater die Kinder zur Schule gebracht hat	
2 Einf.	Ich weis, das der Fater die Junge zur Tennisplatz geklettert ist	
3 Einf.	Er geht spazieren, weil die Sonne scheint	
3 Einf.	Er get spazieren, dass die Wolke regnet	
1 A	Sie konnte nicht fassen, dass das Anwesen des Grafen so groß ist wie ein Schloss	Graf
1 A	Sie konnte nicht fassen, das das Burg des Grafs so glitzernd hat als ein Schloß	Peripherie
2 A	Er sah, wie die Mutter die Kinder zur Schule gefahren hat	fahren
2 A	Er sa, obwohl die Vater das Junge zur Schule geschwommen ist	transitiv
3 A	Ich nehme an, dass die Arbeit meines Kollegen wertgeschätzt wird	Kollege
3 A	Ich neme an, das der Leistung meines Kolleges wertgeschetzt muss	Form-Schema
4 A	In der Zeitung stand, dass der Pilot die Distanz in Rekordzeit geflogen ist	fliegen
4 A	In das Zeitung stand, weil mein Nachbarin der Rennen in Besetzt geschwommen hat	ambig II
5 A	Er fährt nochmal zurück, weil er die Jacke seines Freundes aus Versehen eingesteckt hat und dieser sie auf der Arbeit benötigt	Freund
5 A	Er fährt nochmal zurück, dass sie die Anorak seines Freunden aus Absicht eingesteckt ist oder diese ihn auf das Arbeit benötigt	stark
6 A	Meine Oma erzählte, dass sie in den Siebzigern regelmäßig durch den Schwarzwald geritten ist	reiten
6 A	Meine Ohma erzählte, trotz er in den 70igern Regel mäßig durch den Schwartswald geschwommen hat	intransitiv
1 B	Mein Freund war nicht da, weil er zum Geburtstag seines Neffen gegangen ist	Neffe
1 B	Meine Freund war nicht da, dass sie zum Geburtstag seines Neffes geschwommen hat	Form-Schema
2 B	Ich nehme an, dass der Pilot die Passagiere zum nächstgelegenen Flughafen geflogen hat	fliegen
2 B	Ich neme an, ob der Stewardess das Ware zum nächst gelegenen Bahnhof geschwommen ist	transitiv

Block	Satz	Testitem
3 B	Mein Neffe mag Superman nicht, weil ihm der Anzug des Helden nicht gefällt	Held
3 B	Mein Nefe mag Superman nicht, dass ihre Tante das Aufzug des Helden nicht gefällt	Peripherie
4 B	Mir wurde erzählt, dass mein Onkel früher ein schwarzes Pferd geritten hat	reiten
4 B	Mir wurde erzählt, trotzdem mein Tante früher ein schwarzes Stute geschwommen ist	ambig I
5 B	Ich habe gehört, dass mein Neffe seit gestern Fahrradfahren kann	
5 B	Ich habe gehöhrt, trotz meine Nefte seid gestern Fahrrad fahren müssen	
6 B	Ich ging davon aus, dass die Oma am Sonntag zur Kur gefahren ist	fahren
6 B	Ich gehe davon aus, das der Oma am Sonntag zum Kur geklettert hat	intransitiv
1 C	Ich habe gesehen, wie der Pfeil des Schützen ins Schwarze traf	Schütze
1 C	Ich habe geseen, trotzdem der Feil des Schützes ins Blaue trafte	Form-Schema
2 C	Ich wusste nicht, dass mein Bruder am Mittwoch in die USA geflogen ist	fliegen
2 C	Ich wusste nicht, weil mein Schwester an das Mittwoch in das USA geschwommen hat	intransitiv
3 C	Er erzählt, dass sein Bruder in den USA herum reisen will	
3 C	Er erzeht, trotz seine Bruder auf die USA herumreisen müssen	
4 C	Im Radio hieß es, dass der Reiter den Parcours fehlerlos geritten ist	reiten
4 C	Im Radio hies es, trotzdem die Jockey die Hindemis fehlerfrei geflogen hat	ambig II
5 C	Er macht einen Ausflug, um die Burg des Vogts zu betrachten	Vogt
5 C	Er macht eine Ausflug, dass die Haus des Vogten zu singen	stark
6 C	Er erzählte mir, dass seine Tante eine Zeit lang ein Cabrio gefahren ist	fahren
6 C	Er erzälte mir, das ihre Tante eine zeitlang eine Cabrio geflogen hat	ambig I
1 D	Er teilte mir mit, dass der Kurier die Ware gestern morgen liefern sollte	
1 D	Er teilte mir mit, obwohl die Lieferant die Wahre gestern Morgen essen durften	
2 D	Ich kann kaum glauben, dass meine Schwester neulich einen Helikopter geflogen hat	fliegen
2 D	Ich kan kaum glauben, das meine Bruder neulich einen Flugzeug gefahren ist	ambig I
3 D	Ihm war nicht bewusst, dass die Macht des Zars einmal derart groß gewesen war	Zar
3 D	Ihm war nicht bewusst, ob die Einfluss des Zars ein Mal der Art gross gestaltet wurde	Peripherie
4 D	In den Nachrichten sagten sie, dass der neue Fahrer die Strecke in Rekordzeit gefahren ist	fahren
4 D	In die Nachrichten sagten sie, das der neue Farer das Runde in Beszet geklettert hat	ambig II
5 D	Der Krimi-Fan ahnt schon früh, dass der Plan des Diebes zu simpel ist und nicht klappt	Dieb
5 D	Der Krimi-Fan ant schon früh, das der Tatik des Dieben zu blau hat oder nicht klappt	stark
6 D	Ich gehe davon aus, dass das Mädchen das Pferd zur Koppel geritten hat	reiten
6 D	Ich gehe dahvon aus, trotzdem die Junge das Hengst zur Koppel geflogen ist	transitiv

Anhang F: Plausibilitätstest zur sentence-maze-Studie

Für den Plausibilitätstest zur sentence-maze-Studie wurden alle Sätze aus der Prototypizitäts- und Form-Schematizitätsstudie inklusive der Füllsätze abgefragt. Zusätzlich dazu wurden Distraktorsätze genutzt. Diese dielten einerseits zur Ablenkung, andererseits als Vergleichspunkt: Da die Proband_innen die Sätze nach ihrer semantischen Plausibilität beurteilen sollten, war der Inhalt der Distraktorsätze unwahrscheinlich oder unmöglich (siehe Abschnitt F.2), um einen Kontrast zu den Testsätzen zu generieren. Die Daten wurden mithilfe von SoSciSurvey (Leiner 2019), Version 3.1.06, erhoben. Die Stichprobengröße für den Plausibilitätstest beträgt nach Bereinigung der Daten 58 Proband_innen. Weitere Informationen zum Plausibilitätstest sind im digitalen Anhang.

F.1 Ergebnisse des Plausibilitätstests

In diesem Abschnitt werden die Ergebnisse des Plausibilitätstests für die Testsätze der Studien zu Prototypizität und Form-Schematizität samt den Ergebnissen für die transitiven Fillersätze in der Prototypizitätsstudie präsentiert. Die Tabellen F.1 und F.2 zeigen die Ergebnisse für die Testsätze in der Prototypizitätsstudie und der Form-Schematizitätsstudie. Die Ergebnisse zu den restlichen Fillersätzen und den Distraktorsätzen sind im digitalen Anhang. Die einzelnen Testsätze werden in den Abschnitten 4.3.2 und 4.5.3 aufgeschlüsselt.

In den Tabellen F.1 und F.2 reichte die Skala bei *reiten_transitiv*, *fahren_ambig II*, *Zar* und *Dieb* nur von 1 bis 4. Aufgrund der geringen Akzeptanz wurde *reiten_ambig I* geändert. Ursprünglich lautete der Satz *Mir wurde erzählt, dass mein Onkel während seines Studiums ein schwarzes Pferd geritten hat/ist*. Er wurde zu *Mir wurde erzählt, dass mein Onkel früher ein schwarzes Pferd geritten hat/ist* geändert und so den anderen Sätzen mit *reiten* ähnlicher gemacht. Im Nachhinein wurde außerdem *Parkour* in *reiten_ambig II* zu *Parcours* geändert. Zudem war bei *reiten_ambig II* ursprünglich *Bestzeit* statt *Rekordzeit* vorgesehen. Die Änderung entstand durch einen Übertragungsfehler. In Tabelle F.2 wurde im Nachhinein nur *gewertschätzt* im Testsatz zum Substantiv *Kollege* zu *wertgeschätzt* geändert.

F Plausibilitätstest zur sentence-maze-Studie

Tabelle F.1: Ergebnisse des Plausibilitätstests für die Testsätze der Prototypizitätsstudie

Testsatz	1	2	3	4	5	gesamt	Durchschnitt
fliegen _{transitiv}	9	7	12	16	14	58	3,33
fliegen _{ambigI}	4	10	11	12	21	58	3,62
fliegen _{ambigII}	3	5	5	18	27	58	4,05
fliegen _{intransitiv}	7	10	9	10	22	58	3,52
reiten _{transitiv}	5	10	19	24	0	58	3,07
reiten _{ambigI}	18	11	9	9	11	58	2,72
reiten _{ambigII}	3	6	4	10	35	58	4,17
reiten _{intransitiv}	5	10	14	14	15	58	3,41
fahren _{transitiv}	1	2	0	10	45	58	4,66
fahren _{ambig_I}	2	5	8	16	27	58	4,05
fahren _{ambig_II}	2	2	17	37	0	58	3,53
fahren _{intransitiv}	6	6	7	17	22	58	3,74
Filler _{Jacke} ^a	2	7	9	10	30	58	4,02
Filler _{Vater} ^b	2	0	3	18	35	58	4,45

^aFiller_{Jacke}: Er fährt nochmal zurück, weil er die Jacke seines Freundes aus Versehen eingesteckt hat und dieser sie auf der Arbeit benötigt

^bFiller_{Vater}: Ich weiß, dass der Vater die Kinder zur Schule gebracht hat

Tabelle F.2: Ergebnisse des Plausibilitätstests für die Testsubstantive

Testsubstantiv	1	2	3	4	5	gesamt	Durchschnitt
Kollege	3	6	13	18	18	58	3,72
Neffe	1	4	6	10	37	58	4,53
Schütze	1	2	4	9	42	58	4,53
Graf	5	8	9	8	28	58	3,79
Held	1	1	3	19	34	58	4,45
Zar	3	8	16	31	0	58	3,29
Vogt	5	5	6	17	25	58	3,90
Freund	2	7	9	10	30	58	4,02
Dieb	1	8	19	30	0	58	3,34

F.2 Distraktorsätze

Folgende Sätze wurden genutzt:

- (1) Ich habe in der Zeitung gelesen, dass meine Nichte gestern im Schwimmbad war.
- (2) Ich habe gehört, dass mein Bruder den Marathon in einer halben Stunde gelaufen ist.
- (3) In den Nachrichten hieß es, dass Hamburg die sonnigste Stadt Deutschlands ist.
- (4) Im Radio hieß es, dass du vorhin falsch geparkt hast.
- (5) Der Polizist ahnt schon früh, dass der Plan des Täters wegklettert.
- (6) In der Zeitung steht, dass Deutschland keine Hauptstadt hat.
- (7) Ihm war nicht bewusst, dass der Hund des Grafen sprechen konnte.
- (8) Mir hat er gesagt, dass er früher gerne Äpfel geritten ist.
- (9) Mein Freund war nicht da, weil er bei der Queen zum Kaffee eingeladen war.
- (10) Mir hat er gesagt, dass er früher eine schwarze Kuh geritten ist.
- (11) Ich nehme an, dass der Pilot die Passagiere zum Mars geflogen hat.
- (12) Er fährt den weiten Weg nochmal zurück, weil er die Zahnpasta seines Freunds vergessen hat.
- (13) Der Polizist ahnt schon früh, dass der Plan des Täters ein fliegendes Taxi involviert.
- (14) Ich weiß, dass die Kinder ihren Vater zur Schule gebracht haben.
- (15) Meine Oma erzählte, dass sie früher im Sommer Skifahren war.

Anhang G: Self-paced-reading-Studie

G.1 Texte zu den Testitems

Die Abfolge der Endungen bei den Testitems ist pseudorandomisiert. Eine Gruppe liest zuerst die erwartbare Form (-n für *Schettose* und *Truntake*, -s für *Grettel* und *Knatt*), dann die unerwartete (-s für *Schettose*, endungslos bei *Truntake*, -(e)n für *Grettel* und *Knatt*). Bei *Knatt* sind beide Formen erwartbar, aber -s aufgrund der Typenfrequenz wahrscheinlicher. Auch bei den Fragen zu den Testitems wird in dieser Gruppe zuerst die erwartete und dann die unerwartete genannt (Haben Sie *des Schettosen* oder *des Schettoses* gelesen?). Diese Verteilung wird auch hier präsentiert. Die andere Gruppe liest die gegenteilige Verteilung. Die senkrechten Striche (|) markieren die Wortgruppen, in die die Sätze geteilt wurden.

G.1.1 Prototyp des Form-Schemas: *Schettose* und *Truntake*

Schettose, der

(Betonung: Schettóse)

Amtsbezeichnung am ungarischen Hof, Amtsschreiber/Diener

Der Schettose | war ähnlich gestellt | wie ein Amtsschreiber, | zusätzlich erfüllte er | jedoch niedrigere Dienste. | Frage₁

Der Schettose | wurde vom König | direkt ernannt. | Somit war | das Amt | des Schettosen | offensichtlich | ein verantwortungsvoller Posten. | Frage₂

Es ist | davon auszugehen, | dass die Übertragung dieses Amtes | als Ehre galt. | Frage₃

Interessanterweise | ließ sich | außer dem Amtsschreiber | kein vergleichbares Amt | an anderen Höfen | beobachten. | Frage₄

Dementsprechend war | das Amt | des Schettosen | offensichtlich | ein Spezifikum | des ungarischen Königshofs. | Frage₅

G Self-paced-reading-Studie

1. Haben Sie *Amtschreiber* oder *Amtsschreiber* gelesen?
2. Haben Sie *des Schettosen* oder *des Schettoses* gelesen?
3. Haben Sie *des Amtes* oder *des Amts* gelesen?
4. Haben Sie *Amtschreiber* oder *Amtsschreiber* gelesen?
5. Haben Sie *des Schettosen* oder *des Schettoses* gelesen?

Truntake, der

(Betonung: Truntáke)

Vorsitzender bei Gerichtshöfen

Der Truntake war | im 17. Jahrhundert | in englischen Gerichtshöfen | weit verbreitet. | Frage₁

Der Truntake | wurde dabei | formal wie ein Richter | des Hofes | behandelt. | Frage₂

Jedoch kam | dem Truntaken | üblicherweise | keine Urteilskraft zu. | Frage₃

Der Truntake | stand dem Gericht | jedoch vor | und achtete | auf die Einhaltung | der Gerichtsordnung. | Frage₄

Häufig wurden | angehende Richter | des Hofes | mit diesem Amt betraut. | Frage₅

Sie galten | als vielversprechende Aspiranten | auf verantwortungsvolle Posten. | Frage₆

Dementsprechend beehrt war | der Posten. | Jedoch kam | dem Truntake | üblicherweise | auch eine unangenehme Aufgabe zu: | der Schuldspruch. | Frage₇

1. Haben Sie *Gerichthöfen* oder *Gerichtshöfen* gelesen?
2. Haben Sie *des Hofes* oder *des Hofes* gelesen?
3. Haben Sie *dem Truntaken* oder *dem Truntake* gelesen?
4. Haben Sie *Gerichtsordnung* oder *Gerichtordnung* gelesen?
5. Haben Sie *des Hofes* oder *des Hofes* gelesen?
6. Haben Sie *verantwortungsvoll* oder *verantwortungsvoll* gelesen?
7. Haben Sie *dem Truntaken* oder *dem Truntake* gelesen?

G.1.2 Peripherie des Form-Schemas: *Knatt*

Knatt, der

Veraltete Berufsbezeichnung, Feuerhüter

Dieser Beruf | war über Jahrhunderte | zentraler Bestandteil | einer funktionierenden Dorfgemeinschaft. | Frage₁

Er war zunächst | vor allem | in Franken und Bayern üblich. | Mit der Zeit | setzte sich | der Beruf | des Knatts | anscheinend | auch in anderen Regionen durch. | Frage₂

Der Knatt | wachte über große Feuerstellen | und läutete | im Brandfall | die Feuerturmglocken. | Frage₃

Mit der Verbreitung der Feuerwehren | wurde der Knatt | für die Dorfgemeinschaft | zunehmend bedeutungslos. | Frage₄

Auch in der Stadt | setzte sich | der Beruf | des Knatten | anscheinend | nie durch. | Frage₅

1. Haben Sie *Dorfgemeinschaft* oder *Dorfsgemeinschaft* gelesen?
2. Haben Sie *des Knatts* oder *des Knatten* gelesen?
3. Haben Sie *Brandsfall* oder *Brandfall* gelesen?
4. Haben Sie *Dorfgemeinschaft* oder *Dorfsgemeinschaft* gelesen?
5. Haben Sie *des Knatts* oder *des Knatten* gelesen?

G.1.3 Form-Schema der starken Flexion: *Grettel*

Grettel, der

(Betonung: Gréttel)

Funktionsträger in Heimatvereinen, Trachtenwart

Der Grettel | ist für die Instandhaltung | der Trachten | von Heimatvereinen zuständig. | Frage₁

Die Trachten | sind oft | hochwertig und empfindlich. | Deswegen ist | das Amt | des Grettels | in vielen Vereinen | nur für langjährige Mitglieder | vorgesehen. | Frage₂

G Self-paced-reading-Studie

Außerdem übernimmt | der Grettel | repräsentative Aufgaben. | Die Ausführung dieses Ehrenamts | nimmt deshalb | viel Zeit | in Anspruch. | Frage₃

Dementsprechend ist | das Amt | des Gretteln | in vielen Vereinen | nicht besetzt. | Frage₄

1. Haben Sie *Instandhaltung* oder *Instandhaltung* gelesen?
2. Haben Sie *des Grettels* oder *des Gretteln* gelesen?
3. Haben Sie *des Ehrenamtes* oder *des Ehrenamts* gelesen?
4. Haben Sie *des Grettels* oder *des Gretteln* gelesen?

G.2 Texte zu den Distraktoren

Hinsichtlich der Distraktoren wurde die Reihenfolge der Pluralformen nicht pseudorandomisiert, sodass beide Gruppen identische Texte lasen. Die senkrechten Striche (|) markieren die Wortgruppen, in die die Sätze geteilt wurden.

Taff, die

Veraltete Berufsbezeichnung, eine Art Hellseherin

Der Beruf | war | innerhalb des osmanischen Reichs | weit verbreitet. | Frage₁

Nahezu | in jedem Dorf | arbeiteten Taffs. | Frage₂

Zu den Aufgaben der Täfte | zählte die Begleitung | von Geburten | durch Gebete. | Frage₃

Zudem sagten sie | das Schicksal | des neugeborenen Kindes | vorher. | Frage₄

In der Gesellschaft | waren Taffe | hoch angesehen. | Frage₅

1. Haben Sie *des Reichs* oder *des Reiches* gelesen?
2. Haben Sie *Taffe* oder *Taffs* gelesen?
3. Haben Sie *Täfte* oder *Taffe* gelesen?
4. Haben Sie *des Kinds* oder *des Kindes* gelesen?
5. Haben Sie *Täfte* oder *Taffe* gelesen?

Schirr, die

Adelstitel, ähnlich einer Gräfin

Der Adelstitel der Schirr | wurde | in Sachsen eingeführt. | Frage₁

Schirrs waren | Gräfinnen gleichgestellt, | hatten aber | in die gräfliche Familie |
eingeheiratet. | Frage₂

Die Schirre wurden | inoffiziell als Gräfin | angesprochen, | weswegen sich | der
Titel | nie durchsetzte. | Frage₃

1. Haben Sie *Adelstitel* oder *Adelstitel* gelesen?
2. Haben Sie *Schirrs* oder *Schirre* gelesen?
3. Haben Sie *Schirrs* oder *Schirre* gelesen?

Bracht, die

Veraltete Berufsbezeichnung, eine Art Schneiderin

Zu den Aufgaben von Brächten | zählte das Ausbessern | von Kleidungsstücken.
| Frage₁

Dabei waren Brachts | oft in Schneidereien | beschäftigt | und ähnlich gestellt |
wie Lehrlinge. | Frage₂

Zusätzlich zu den Ausbesserungsarbeiten | waren Brachten | damit betraut, |für
das leibliche Wohl | der Schneiderfamilie | zu sorgen. | Frage₃

1. Haben Sie *von Brächten* oder *von Brachten* gelesen?
2. Haben Sie *Brachte* oder *Brachts* gelesen?
3. Haben Sie *Ausbesserungsarbeiten* oder *Ausbesserungsarbeiten* gelesen?

G.3 Produktionsexperiment

Folgende Sätze wurden genutzt. Weitere Informationen zum Produktionsexperiment sind im digitalen Anhang.

- (1) Viele ___ (Taff) arbeiteten in Dörfern.
- (2) Der Beruf des ___ (Knatt) diente dem Feuerschutz.
- (3) Die meisten ___ (Schirr) wurden als Gräfin angesprochen.
- (4) Das Amt des ___ (Schettose) war am ungarischen Königshof üblich.
- (5) Die meisten ___ (Bracht) arbeiteten in Schneidereien.
- (6) Das Amt des ___ (Grettel) ist in Heimatvereinen üblich.
- (7) Zunächst kam ___ (Truntake) üblicherweise keine Urteilstkraft zu.

Anhang H: Assoziationstest zur self-paced-reading-Studie

Im Assoziationstest wurden neben den in der Studie genutzten Testsubstantiven *Schettose*, *Truntake*, *Knatt* und *Grettel* als Alternativen *Trilch* und *Fletter* abgefragt, die ebenfalls Köpcke (2000b) entnommen wurden. Die Pseudosubstantive wurden mit Artikel (z. B. *der Grettel*) präsentiert, sodass das Genus der Wörter vorgegeben war. Die Daten wurden mithilfe von ausgedruckten Fragebögen erhoben. Die Stichprobe umfasst 90 Proband_innen, Metadaten wurden nicht erhoben. Weitere Informationen sind im digitalen Anhang zu finden.

Für Tabelle H.1 wurden alle Assoziationen berücksichtigt, die mehr als zehnmal genannt wurden. Hierbei wurde auch gezählt, wie häufig keine Assoziation aufgeschrieben wurde. Den Prozentzahlen liegt als Gesamtanzahl 270 (90 Proband_innen * 3 Assoziationen) zugrunde, da jede_r Proband_in maximal drei Angaben machen konnte. Assoziierte Substantive wurden lexemweise einbezogen. Alle anderen Wortarten wurden unter der jeweiligen Wortart zusammengefasst, da diese die Deklinationsklasse nicht beeinflussen sollten. Die Anzahl der Assoziationen wurde ermittelt, indem die Nennungen einzelner Lexeme ausgezählt wurden. Die einzigen Ausnahmen hiervon bilden die Assoziationskategorien *Betrunkener* sowie *Gretel*: Für *Betrunkener* wurden verschiedene Bezeichnungen zusammengefasst (*Trunkenbold*, *jemand, der getrunken hat*, *Alkoholiker*), im Fall von *Gretel* wurde auch *Hänsel und Gretel* als Teil der Assoziationskategorie angesehen, jedoch nicht die Assoziationen wie *Grimm* und *Märchen*, da diese auf einer abstrakteren Ebene agieren. Bei *Vetter* wurden vier Schreibungen mit *F* (*Fetter*) als diesem Lexem zugehörig interpretiert, da diese Lesart wahrscheinlicher ist als eine Konversion des Adjektivs *fett*.

Es ist festzuhalten, dass bei allen Pseudosubstantiven *n.a.* die häufigste Antwortkategorie darstellt. Somit scheint keines der Pseudosubstantive starke intersubjektive Assoziationen auszulösen. *Schettose* wird nur mit schwach flektierenden Substantiven (*Matrose*, *Schotte*) assoziiert. Da *Schettose* dem Form-Schema I entspricht (siehe Abschnitt 3.2.3), sind die Assoziationen mit schwachen Maskulina nicht verwunderlich und für das Experiment eher zuträglich denn schädigend, da die Assoziationen die schwache Flexion unterstützen können. Die restlichen

H Assoziationstest zur self-paced-reading-Studie

Tabelle H.1: Überblick über die häufigsten Assoziationen mit den Pseudosubstantiven

Assoziation	<i>Schettose</i>	Assoziation	<i>Truntake</i>
n.a.	103 (38 %)	n.a.	105 (62 %)
<i>Steckdose</i>	18 (7 %)	ADJ	20 (12 %)
<i>Matrose</i>	13 (5 %)	VV	20 (12 %)
ADJ	12 (4 %)	<i>Truthahn</i>	16 (9 %)
<i>Hose</i>	12 (4 %)	<i>Tentakel</i>	12 (7 %)
<i>Schotte</i>	11 (4 %)	<i>Betrunkener</i>	12 (7 %)
Assoziation	<i>Trilch</i>	Assoziation	<i>Knatt</i>
n.a.	81 (30 %)	n.a.	95 (35 %)
<i>Milch</i>	33 (12 %)	ADJ	35 (13 %)
<i>Knilch</i>	24 (9 %)	VV	26 (10 %)
<i>Trichter</i>	20 (7 %)	<i>Knast</i>	23 (9 %)
ADJ	15 (6 %)		
VV	11 (4 %)		
Assoziation	<i>Fletter</i>	Assoziation	<i>Grettel</i>
n.a.	94 (35 %)	n.a.	130 (48 %)
VV	34 (13 %)	<i>Gretel</i>	45 (17 %)
<i>Vetter</i>	18 (7 %)	<i>Zettel</i>	15 (6 %)
ADJ	14 (5 %)	ADJ	10 (4 %)
<i>Fledermaus</i>	13 (5 %)		

Assoziationen entfallen auf Feminina (*Steckdose*, *Hose*), welche die Deklination nicht beeinflussen sollten, da das Testitem als Maskulinum eingeführt wird. Die häufigste Assoziation (*Steckdose*) rührt vermutlich daher, dass die Proband_innen das Substantiv initial betonten (*Schéttose*). Aufgrund dessen wurde in der self-paced-reading-Studie ein Hinweis auf die Betonung des Substantivs ergänzt. *Truntake* scheint auf den ersten Blick nicht problematisch zu sein, allein die Assoziation *Truthahn* könnte Probleme bereiten, da dies ein starkes Maskulinum ist. Zudem scheint *Truntake* negativ konnotiert zu sein, wie die Assoziationen *Betrunkener*, *Trunkenbold* zeigen. Aufgrund dieser negativen Assoziation wurde *Schettose* als Testitem für die Genitivformen gewählt. Da für Substantive, die dem Form-Schema I komplett entsprechen, zusätzlich zum Genitiv der Dativ getestet wurde, wurde hierfür *Truntake* als Testitem eingeflochten. Wie bei *Schotto-*

se wurde ein Hinweis zur Betonung ergänzt (*Truntáke*), um die Assoziation mit schwachen Flexionsverhalten zu erleichtern.

Bei dem Pseudowortpaar *Trilch* und *Knatt* lassen sich zwei problematische Assoziationen feststellen: *Trilch* wurde mit dem starken Maskulinum *Knilch* assoziiert, *Knatt* mit dem starken Maskulinum *Knast*. Beide Assoziationen wurden ähnlich oft genannt, sodass die Wahl zwischen *Trilch* und *Knatt* nach semantischen Faktoren getroffen wurde. Da das Pseudowort im Experiment als Mensch eingeführt wird, ist bei *Knatt* eine Assoziation mit *Knast* unwahrscheinlich, während bei *Trilch* eine Assoziation mit *Knilch* durchaus möglich ist, da *Knilch* auf eine menschliche Entität referiert. Aufgrund dessen wurde *Knatt* als Testitem gewählt. Bei *Fletter* ist die Assoziation mit *Vetter* problematisch, da dieses Substantiv gemischt flektiert (Duden 2020). Zwar werden im Experiment keine Pluralformen abgefragt, jedoch ist es möglich, dass sich bei *Vetter* ähnlich wie bei *Autor* aufgrund der Nähe zum Form-Schema schwache Formen im Singular finden lassen (siehe Abschnitt 3.2.3). Aufgrund dessen wurde *Grettel* als Versuchitem gewählt. Die häufige Assoziation mit *Gretel* ist unproblematisch, da es sich bei *Grettel* um ein Femininum handelt.

Anhang I: Ergebnisse der Studien

I.1 Ergebnisse der Frequenzstudie

In der Abbildung I.1 steht ein Rechteck für eine Antwort.

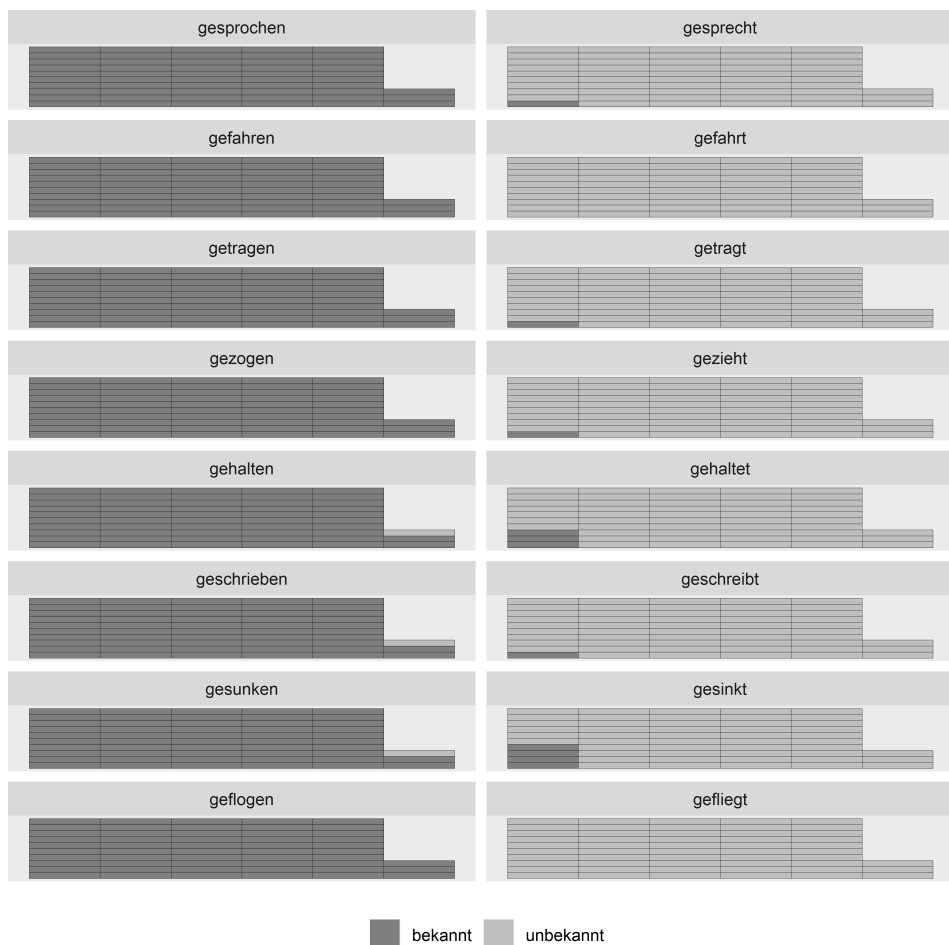


Abbildung I.1: Antwortverhalten bei den Testverben der Ausprägung FREQUENT

I Ergebnisse der Studien

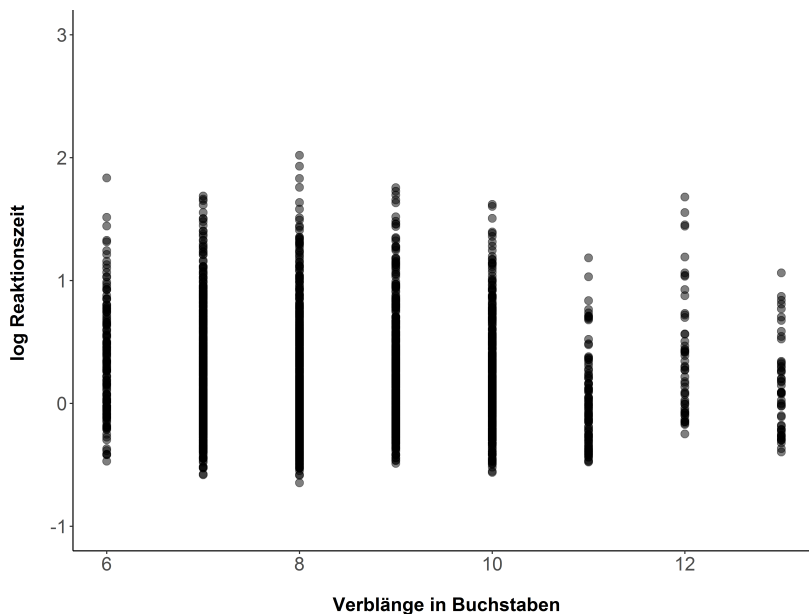


Abbildung I.2: Reaktionszeiten im Abhängigkeit von der Testverblänge in der Frequenzstudie

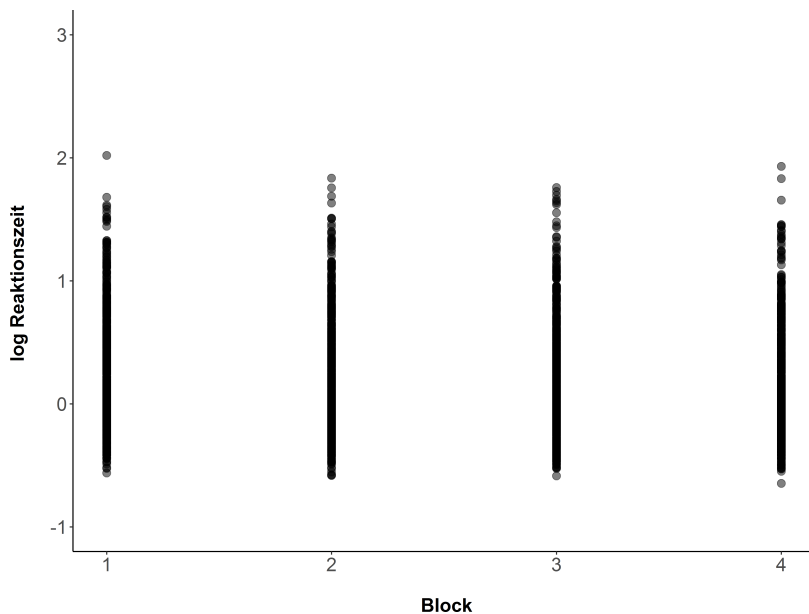


Abbildung I.3: Reaktionszeiten in Abhängigkeit vom Testblock in der Frequenzstudie

I.2 Ergebnisse der Prototypizitätsstudie

Tabelle I.1: Kreuztabelle der Werte des Modells für Reaktionszeiten mit Interaktion zwischen Frequenz und Flexion. o.S.: ohne Schwankung; m.S.: mit Schwankung.

	schwach	stark
infrequent o.S.	0,42	$0,42 - 0,33 = \mathbf{0,09}$
frequent	$0,42 - 0,19 = \mathbf{0,23}$	$0,42 - 0,33 - 0,19 + 0,06 = \mathbf{-0,04}$
infrequent m.S.	$0,42 - 0,04 = \mathbf{0,38}$	$0,42 - 0,33 - 0,04 + 0,17 = \mathbf{0,22}$

I.2 Ergebnisse der Prototypizitätsstudie

Tabelle I.2: Häufigkeit der Auxiliarkombinationen für die Abfolge TRANSITIV, AMBIG I, AMBIG II, INTRANSITIV bei den von der Stichprobe ausgeschlossenen Proband_innen

Kombinationsmöglichkeit	<i>fahren</i>	<i>fliegen</i>	<i>reiten</i>
hat_ist_ist_ist	16 (80 %)	18 (90 %)	3 (15 %)
hat_hat_ist_ist	4 (20 %)	1 (5 %)	7 (35 %)
ist_ist_ist_ist	0 (0 %)	1 (5 %)	3 (15 %)
ist_hat_ist_ist	0 (0 %)	0 (0 %)	3 (15 %)
ist_hat_hat_ist	0 (0 %)	0 (0 %)	2 (10 %)
hat_hat_hat_ist	0 (0 %)	0 (0 %)	1 (5 %)
ist_ist_hat_ist	0 (0 %)	0 (0 %)	1 (5 %)
gesamt	20 (100 %)	20 (100 %)	20 (100 %)

I Ergebnisse der Studien

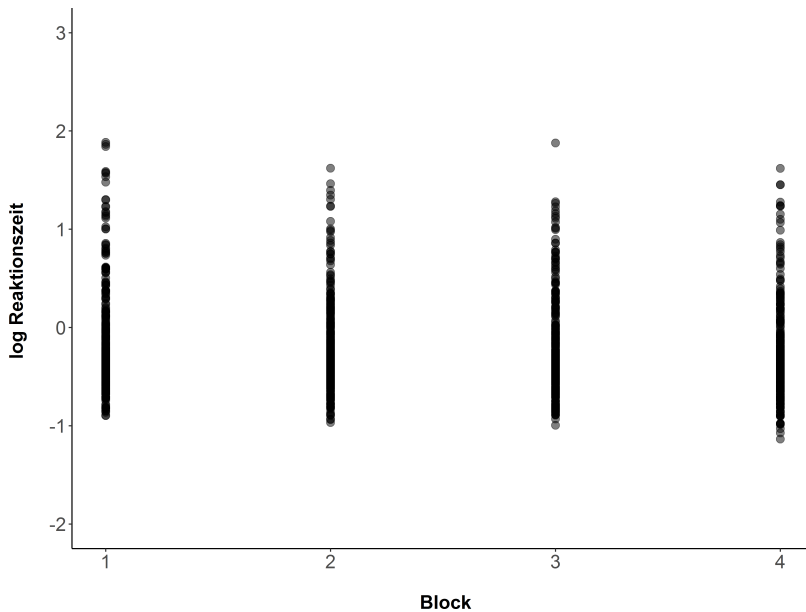


Abbildung I.4: Reaktionszeiten in Abhängigkeit vom Testblock in der Prototypizitätsstudie

I.3 Ergebnisse der Form-Schematizitätsstudien

I.3.1 Lexical-decision-Studie

Tabelle I.3: Kreuztabelle der Werte des Modells für die Unbekanntheit starker und schwacher Formen in der lexical-decision-Studie zu Form-Schematizität

	schwach	stark
Form-Schema	-3,28	$-3,28 + 5,77 = 2,49$
Peripherie	$-3,28 - 0,38 = -3,66$	$-3,28 + 5,77 - 0,38 - 2,29 = -0,18$
stark	$-3,28 + 6,19 = 2,91$	$-3,28 + 5,77 + 6,19 - 11,23 = -2,55$

I.3 Ergebnisse der Form-Schematizitätsstudien

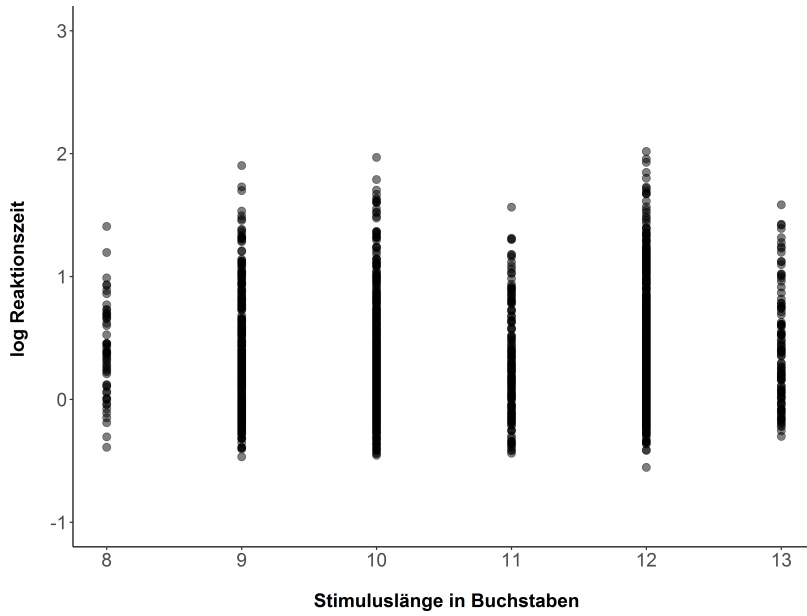


Abbildung I.5: Reaktionszeiten im Abhängigkeit von der Testitemlänge in der lexical-decision-Studie zu Form-Schematizität

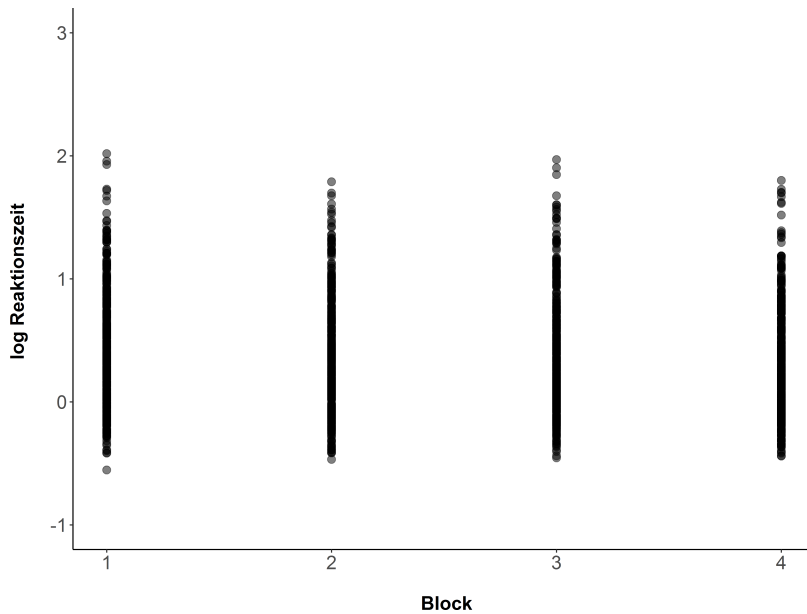


Abbildung I.6: Reaktionszeiten in Abhängigkeit vom Testblock in der lexical-decision-Studie zu Form-Schematizität

I Ergebnisse der Studien

Tabelle I.4: Kreuztabelle der Werte des Modells für die Reaktionszeiten der lexical-decision-Studie zu Form-Schematizität

	schwach	stark
Form-Schema	0,28	$0,28 + 0,2 = \mathbf{0,48}$
Peripherie	$0,28 - 0,07 = \mathbf{0,21}$	$0,28 + 0,2 - 0,07 + 0,09 = \mathbf{0,50}$
stark	$0,28 + 0,25 = \mathbf{0,53}$	$0,28 + 0,2 + 0,25 - 0,46 = \mathbf{0,27}$

Tabelle I.5: Kreuztabelle der Werte des Modells für die Reaktionszeiten in der lexical-decision-Studie zu Form-Schematizität in Abhängigkeit von Frequenz

	schwach	stark
frequent	0,26	$0,26 + 0,23 = \mathbf{0,49}$
mittelfrequent	$0,26 + 0,12 = \mathbf{0,38}$	$0,26 + 0,23 + 0,12 - 0,19 = \mathbf{0,42}$
infrequent	$0,26 + 0,12 = \mathbf{0,38}$	$0,26 + 0,23 + 0,12 - 0,27 = \mathbf{0,34}$

I.3.2 Sentence-maze-Studie

In den Abbildungen I.7 und I.8 steht ein Rechteck für zwei Antworten.

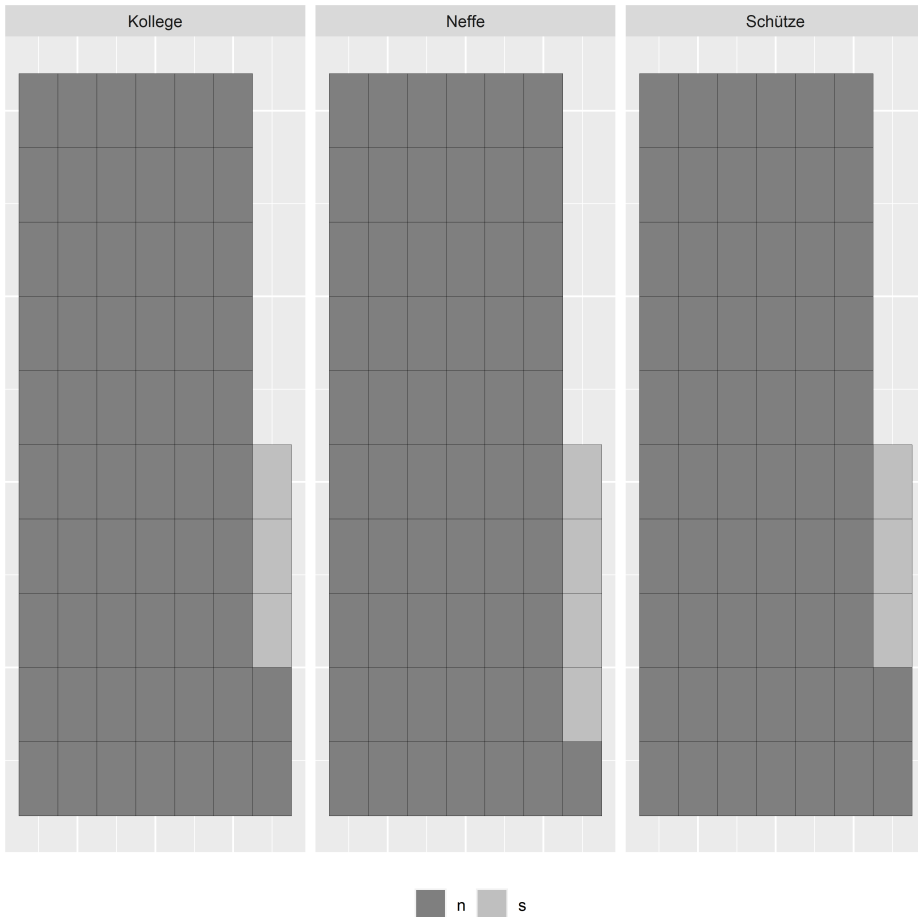


Abbildung I.7: Antwortverhalten bei den Testsubstantiven der Ausprägung FORM-SCHEMA in der sentence-maze-Studie zu Form-Schematizität

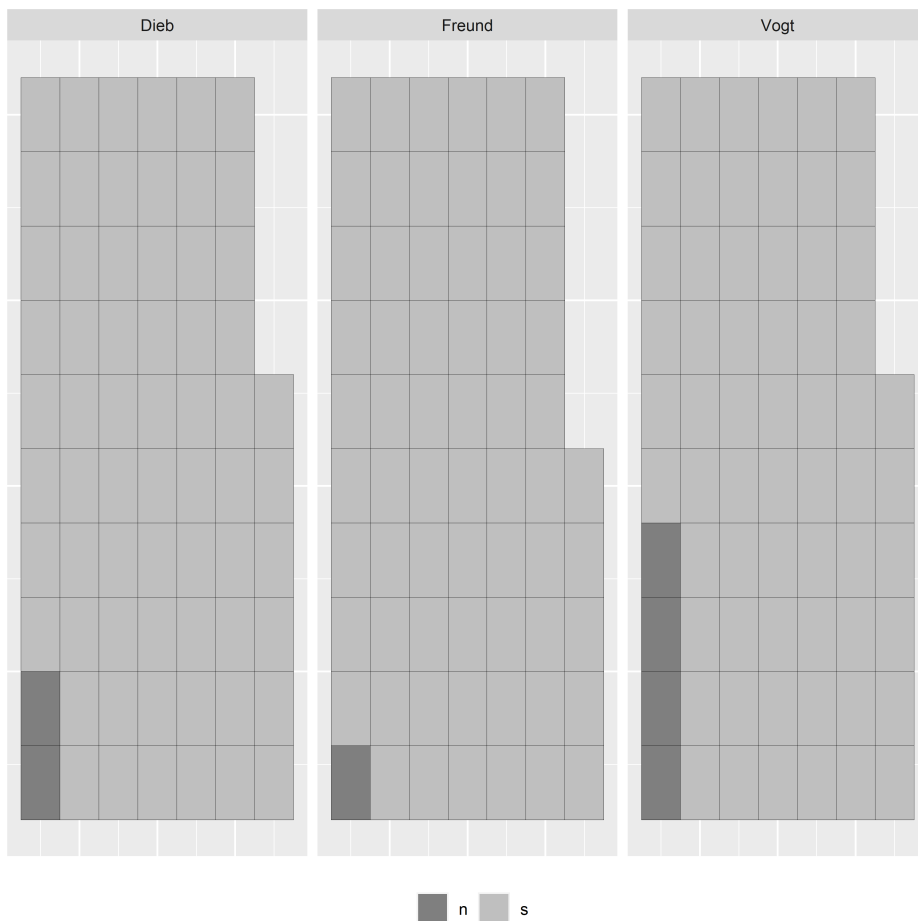


Abbildung I.8: Antwortverhalten bei den Testsubstantiven der Ausprägung STARK in der sentence-maze-Studie zu Form-Schematizität

I.3 Ergebnisse der Form-Schematizitätsstudien

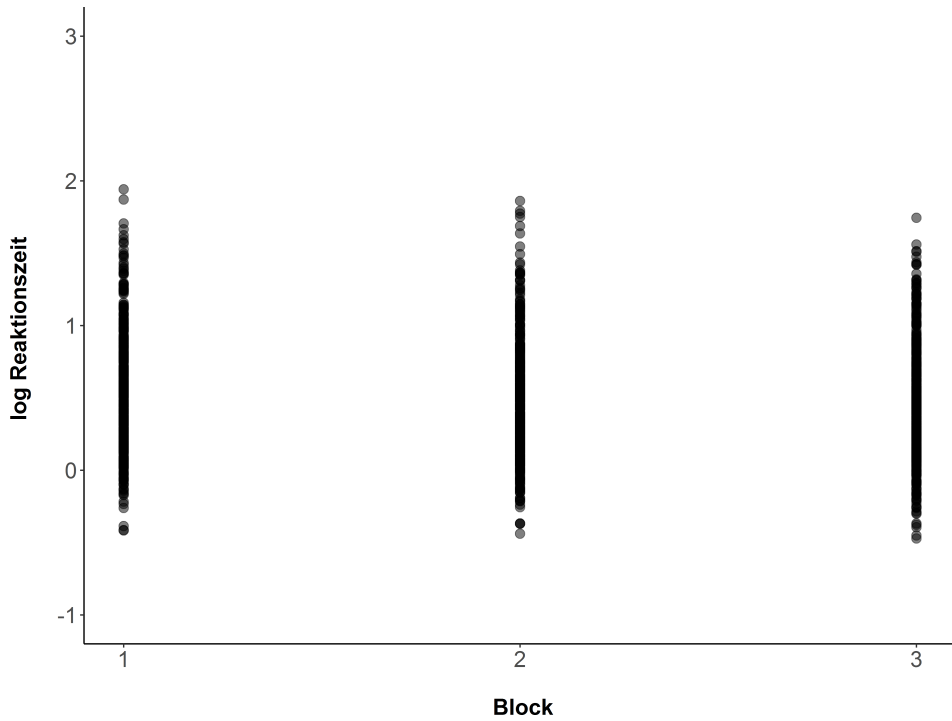


Abbildung I.9: Reaktionszeiten in Abhängigkeit vom Testblock in der sentence-maze-Studie zu Form-Schematizität

I.3.3 Self-paced-reading-Studie

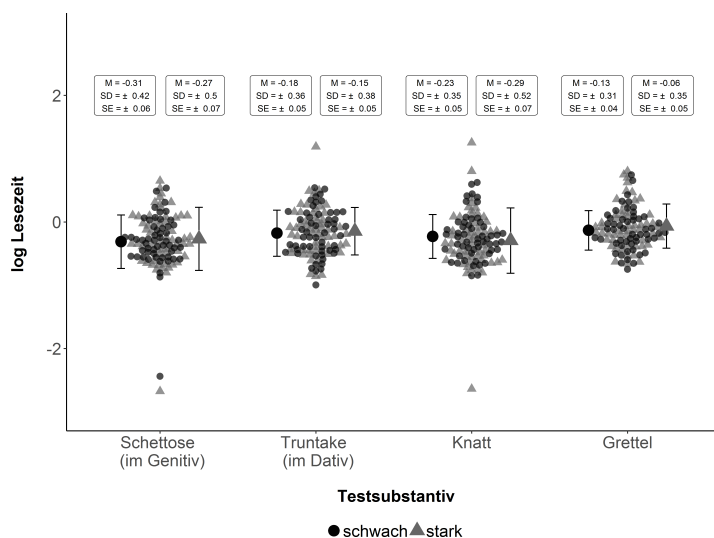


Abbildung I.10: Lesezeiten in Abhängigkeit von dem Element nach dem Testitem in der self-paced-reading-Studie (*stark* und *schwach* beziehen sich auf die Flexionsform des Testitems)

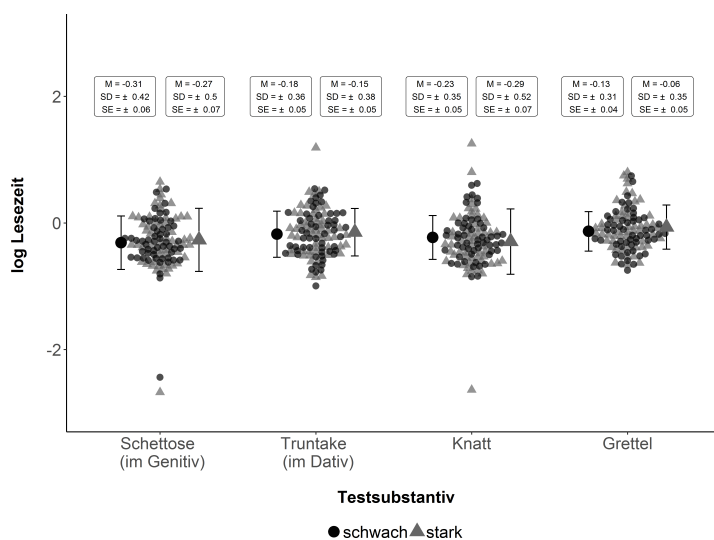


Abbildung I.11: Reaktionszeiten in Abhängigkeit von den Fragen zu den Testitems in der self-paced-reading-Studie

I.3 Ergebnisse der Form-Schematizitätsstudien

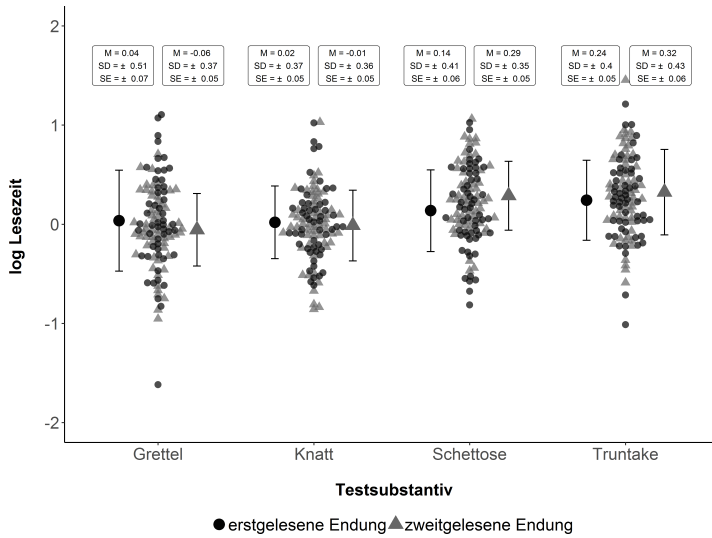


Abbildung I.12: Lesezeiten in Abhängigkeit von der Reihenfolge der Testitems in der self-paced-reading-Studie

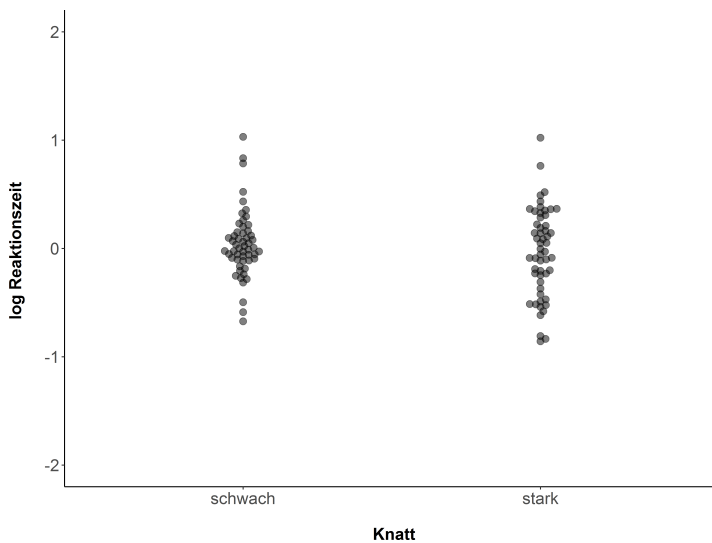


Abbildung I.13: Lesezeiten in Abhängigkeit von starken und schwachen Formen bei *Knatt*

I Ergebnisse der Studien

Wie Abbildung I.13 zeigt, evozieren beide Formen vergleichbare Lesezeiten, so dass ein Einfluss der starken und schwachen Form sowie der Länge des Testsubstantivs ausgeschlossen werden kann. Würde die Länge der Formen eine Rolle spielen, müsste die starke Form (*des Knatts*) zu niedrigeren Lesezeiten neigen, da sie um einen Buchstaben kürzer sind als die schwache (*des Knatten*).

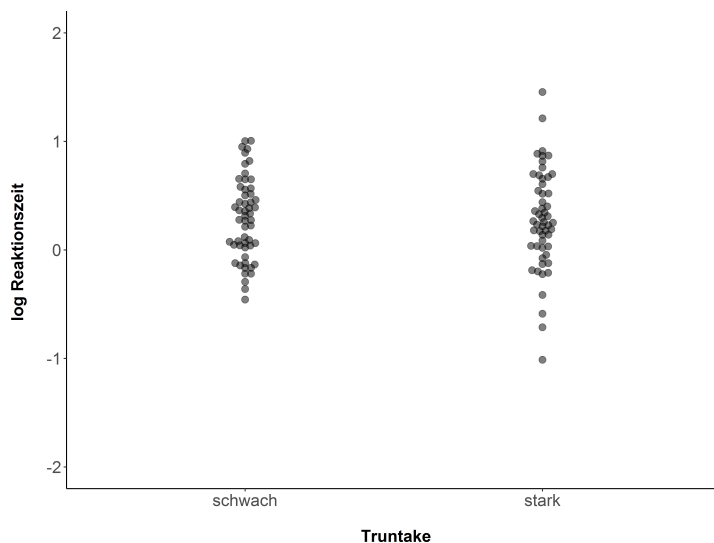


Abbildung I.14: Lesezeiten in Abhängigkeit von starken und schwachen Formen bei *Truntake*

Wie Abbildung I.14 zeigt, streut die starke Form stärker zwischen niedrigeren und höheren Lesezeiten als die schwache Form. Die kürzeren Lesezeiten bei der starken Form könnten einen Längeneffekt darstellen, da die starke Form (*dem Truntake*) um einen Buchstaben kürzer ist als die schwache (*dem Truntaken*). Umgekehrt könnten die erhöhten Lesezeiten bei der starken Form einen Effekt der Flexionsform darstellen, da die starke Form nicht erwartbar ist.

Tabelle I.6: Kreuztabelle der Werte des Modells für die Lesezeiten in der self-paced-reading-Studie

	schwach	stark
Schettose	0,12	$0,12 + 0,18 = \mathbf{0,40}$
Truntake	$0,12 + 0,17 = \mathbf{0,29}$	$0,12 + 0,18 + 0,17 - 0,19 = \mathbf{0,28}$
Knatt	$0,12 - 0,08 = \mathbf{0,04}$	$0,12 + 0,18 - 0,08 - 0,25 = \mathbf{-0,03}$
Grettel	$0,12 - 0,11 = \mathbf{0,01}$	$0,12 + 0,18 - 0,11 - 0,22 = \mathbf{-0,03}$

I.4 Vergleich zwischen den Reaktionszeiten der Verben und der Substantive in der lexical-decision-Studie

Tabelle I.7: Kreuztabelle der Werte des Modells für die Reaktionszeiten bei Substantiven und Verben in den Studien zu Form-Schematizität und Frequenz

	Substantiv		Verb	
	schwach	stark	schwach	stark
frequent/ Form-Schema	0,24	$0,24 + 0,22$ = 0,46	$0,24 - 0,01$ = 0,23	$0,24 - 0,01 + 0,22$ $- 0,49$ = -0,04
ohne/ Peripherie	$0,24 - 0,05$ = 0,19	$0,24 + 0,22$ $- 0,05 + 0,06$ = 0,47	$0,24 - 0,01$ $- 0,05 + 0,25$ = 0,43	$0,24 - 0,01 + 0,22$ $- 0,05 + 0,25 - 0,49$ $+ 0,06 - 0,12$ = 0,1
mit/stark	$0,24 + 0,26$ = 0,5	$0,24 + 0,22$ $+ 0,26 - 0,48$ = 0,24	$0,24 - 0,01$ $+ 0,26 - 0,10$ = 0,39	$0,24 - 0,01 + 0,22$ $+ 0,26 - 0,10 - 0,49$ $- 0,48 + 0,59$ = 0,23

Literaturverzeichnis

- Ackermann, Tanja & Christian Zimmer. 2017. Morphologische Schemakonstanz: Eine empirische Untersuchung zum funktionalen Vorteil nominalmorphologischer Wortschonung im Deutschen. In Nanna Fuhrhop, Renata Szczepaniak & Karsten Schmidt (Hrsg.), *Sichtbare und hörbare Morphologie* (Linguistische Arbeiten), 145–176. Berlin/Boston: de Gruyter.
- Ágel, Vilmos. 2008. Bastian Sick und die Grammatik: Ein ungleiches Duell. *Informationen Deutsch als Fremdsprache* 35(1). 64–84.
- Alegre, Maria & Peter Gordon. 1999. Frequency effects and the representational status of regular inflections. *Journal of Memory and Language* 40(1). 41–61. DOI: 10.1006/jmla.1998.2607.
- Alexiadou, Artemis, Elena Anagnostopoulou & Martin Everaert. 2004. Introduction. In Artemis Alexiadou, Elena Anagnostopoulou & Martin Everaert (Hrsg.), *The unaccusativity puzzle*, 1–22. Oxford: Oxford University Press.
- Audring, Jenny & Geert Booij. 2016. Cooperation and coercion. *Linguistics* 54(4). 617–637.
- Auer, Peter. 2014. Anmerkungen zum Salienzbeffir in der Soziolinguistik. *Linguistik online* 66(4). 7–20.
- Augst, Gerhard. 1975. *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache* (Forschungsberichte des Instituts für Deutsche Sprache 25). Tübingen: Narr.
- Baayen, R. Harald. 2012. Demythologizing the word frequency effect: A discriminative learning perspective. In Gary Libben, Gonia Jarema & Chris Westbury (Hrsg.), *Methodological and analytic frontiers in lexical research*, 171–195. Amsterdam/Philadelphia: Benjamins.
- Baayen, R. Harald. 2014. Experimental and psycholinguistic approaches to studying derivation. In Rochelle Lieber & Pavol Štekauer (Hrsg.), *The Oxford handbook of derivational morphology* (Oxford Handbooks in Linguistics), 95–118. Oxford: Oxford University Press.
- Baayen, R. Harald, Petar Milin, Dusica Filipović Đurđević, Peter Hendrix & Marco Marelli. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118(3). 438–481. DOI: 10.1037/a0023851.

- Bar, Moshe. 2004. Visual objects in context. *Nature Reviews. Neuroscience* 5(8). 617–629. DOI: 10.1038/nrn1476.
- Barca, Laura & Giovanni Pezzulo. 2012. Unfolding visual lexical decision in time. *PloS one* 7(4). 1–9. DOI: 10.1371/journal.pone.0035932.
- Bärenfänger, Olaf. 2002. Merkmals- und Prototypensemantik: Einige grundsätzliche Überlegungen. *Linguistik online* 12(3). 3–17.
- Barlow, Michael & Suzanne Kemmer. 1994. A schema-based approach to grammatical description. In Susan D. Lima, Roberta L. Corrigan & Gregory K. Iverson (Hrsg.), *The reality of linguistic rules* (Studies in Language Companion Series 26), 19–42. Amsterdam/Philadelphia: Benjamins.
- Bartke, Susanne, Frank Rösler, Judith Streb & Richard Wiese. 2005. An ERP-study of German ‘irregular’ morphology. *Journal of Neurolinguistics* 18(1). 29–55. DOI: 10.1016/j.jneuroling.2004.10.001.
- Barton, Kamil. 2019. *MuMIn: Multi-Model Inference*. Version 1.43.17. <https://CRAN.R-project.org/package=MumIn> (4 August, 2020).
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). Version 1.1-23, 1–48. <https://cran.r-project.org/web/packages/lme4/index.html> (4 August, 2020).
- Bates, Elizabeth, Inge Bretherton & Snyder Lynn. 1988. *From first words to grammar: Individual differences and dissociable mechanisms*. New York: Cambridge University Press.
- Battig, William F. & William E. Montague. 1969. Category norms of verbal items in 56 categories: A replication and extension of the connecticut category norms. *Journal of Experimental Psychology Monograph* 80(3). 1–46.
- Beretta, Alan, Carrie Campbell, Thomas H. Carr, Jie Huang, Lothar M. Schmitt, Kiel Christianson & Yue Cao. 2003. An ER-fMRI investigation of morphological inflection in German reveals that the brain makes a distinction between regular and irregular forms. *Brain and Language* 85(1). 67–92. DOI: 10.1016/S0093-934X(02)00560-6.
- Bergmann, Rolf, Claudine Moulin & Nikolaus Ruge. 2016. *Alt- und Mittelhochdeutsch: Arbeitsbuch zur Grammatik der älteren deutschen Sprachstufen und zur deutschen Sprachgeschichte*. 9. Aufl. Göttingen: Vadenhoeck & Ruprecht.
- Berlin-Brandenburgische Akademie der Wissenschaften. 2019. *DWDS: Digitales Wörterbuch der deutschen Sprache*. <https://dwds.de/> (4 August, 2020).
- Binanzer, Anja, Valentina Cristante & Andreas Bittner. 2020. Verarbeitung von deutschen W-Fragen mit starken und schwachen Maskulina durch ein- und mehrsprachige Kinder und Erwachsene: Evidenz für morphosyntaktische Schemata? In Anja Binanzer, Jana Gamper & Verena Wecker (Hrsg.), *Prototy-*

- pen – Schemata – Konstruktionen* (Reihe Germanistische Linguistik), 375–410. Boston: de Gruyter.
- Bittner, Andreas. 1985. Wie schwach sind die starken Verben? Überlegungen und Vorschläge zur Klassifizierung der neuhochdeutschen Verben. In Wolfgang Ullrich Wurzel (Hrsg.), *Studien zur Morphologie und Phonologie I* (Linguistische Studien. Reihe A. Arbeitsberichte 126), 51–74. Berlin: Akademie Verlag.
- Bittner, Andreas. 1996. *Starke „schwache“ Verben – schwache „starke“ Verben: Deutsche Verbflexion und Natürlichkeit* (Studien zur deutschen Grammatik 51). Tübingen: Stauffenburg.
- Bittner, Andreas. 2010. Aspekte diachronischer Fundierung: Historische Linguistik und mentale Repräsentation flexionsmorphologischen Wissens. In Arne Ziegler (Hrsg.), *Historische Textgrammatik und Historische Syntax des Deutschen*, 237–260. Berlin/Boston: de Gruyter.
- Bittner, Dagmar. 2003 [1991]. *Von starken Feminina und schwachen Maskulina: Die neuhochdeutsche Substantivflexion – Eine Systemanalyse im Rahmen der natürlichen Morphologie*. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung.
- Blumenthal-Dramé, Alice. 2012. *Entrenchment in usage-based theories: What corpus data do and do not reveal about the mind*. Berlin/Boston: de Gruyter.
- Blutner, Reinhard. 1995. Prototypen und Kognitive Semantik. In Gisela Harras (Hrsg.), *Die Ordnung der Wörter: Kognitive und lexikalische Strukturen* (Jahrbuch des Instituts für Deutsche Sprache 1993), 227–270. Berlin/Boston: de Gruyter. DOI: 10.7767/9783205127956-001.
- Boas, Hans Christian. 2013. Cognitive construction grammar. In Thomas Hoffmann & Graeme Trousdale (Hrsg.), *The Oxford handbook of construction grammar* (Oxford Handbooks in Linguistics), 233–254. Oxford: Oxford University Press.
- Bolker, Ben & David Robinson. 2019. *broom. Mixed: Tidying methods for mixed models*. Version: 0.2.6. <https://CRAN.R-project.org/package=broom.mixed> (4 August, 2020).
- Booij, Geert. 2010. *Construction morphology* (Oxford Linguistics). Oxford: Oxford University Press.
- Booij, Geert. 2012. Construction morphology, a brief introduction. *Morphology* 22. 343–346.
- Bornkessel-Schlesewsky, Ina & Matthias Schlewsky. 2009. *Processing syntax and morphology: A neurocognitive perspective*. Oxford: Oxford University Press.
- Brainerd, Charles J., Valerie F. Reyna & T. J. Forrest. 2002. Are young children susceptible to the false-memory illusion? *Child Development* 73(5). 1363–1377.

- Braune, Wilhelm & Frank Heidermanns. 2018. *Althochdeutsche Grammatik*. 16. Aufl. (Sammlung kurzer Grammatiken germanischer Dialekte. A: Hauptreihe 1: Laut- und Formenlehre). Berlin/Boston: de Gruyter.
- Bücker, Jörg. 2015. Schema – Muster – Konstruktion. In Christa Dürscheid & Jan Georg Schneider (Hrsg.), *Handbuch Satz, Äußerung, Schema*, 445–463. Berlin/Boston: de Gruyter.
- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form* (Typological Studies in Language 9). Amsterdam/Philadelphia: J. Benjamins.
- Bybee, Joan L. 1988. Morphology as lexical organization. In Michael Hammond & Michael Noonan (Hrsg.), *Theoretical morphology*, 119–141. London: Academic Press.
- Bybee, Joan L. 1991. Natural morphology: The organization of paradigms and language acquisition. In Thom Huebner & Charles A. Ferguson (Hrsg.), *Cross-currents in second language acquisition and linguistic theories*, 67–91. Amsterdam/Philadelphia: Benjamins.
- Bybee, Joan L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425–455. DOI: 10.1080/01690969508407111.
- Bybee, Joan L. 2000. The phonology of the lexicon: Evidence for lexical diffusion. In Michael Barlow & Suzanne Kemmer (Hrsg.), *Usage-based models of language*, 65–85. Stanford: Center for the Study of Language and Information.
- Bybee, Joan L. 2002a. Mechanisms of change in grammaticization: The role of frequency. In Brian D. Joseph & Richard D. Janda (Hrsg.), *The handbook of historical linguistics*, 602–623. Malden: Blackwell.
- Bybee, Joan L. 2002b. Sequentiality as the basis of constituent structure. In Talmy Givón & Bertram F. Malle (Hrsg.), *The evolution of language out of pre-language*, 109–134. Amsterdam/Philadelphia: Benjamins.
- Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711–733.
- Bybee, Joan L. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.
- Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Bybee, Joan L. 2013. Usage-based theory and exemplar representations of constructions. In Thomas Hoffmann & Graeme Trousdale (Hrsg.), *The Oxford handbook of construction grammar* (Oxford Handbooks in Linguistics), 49–69. Oxford: Oxford University Press.
- Bybee, Joan L. & David Eddington. 2006. A usage based approach to Spanish verbs of 'becoming'. *Language* 82(2). 323–355.

- Bybee, Joan L. & Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In Joan L. Bybee & Paul Hopper (Hrsg.), *Frequency and the emergence of linguistic structure* (Typological Studies in Language 45), 1–24. Amsterdam/Philadelphia: Benjamins. DOI: 10.1075/tsl.45.01byb.
- Bybee, Joan L. & Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59(2). 251–270.
- Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: Chicago University Press.
- Bybee, Joan L. & Dan I. Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58(2). 265–289.
- Bybee, Joan L. & Sandra A. Thompson. 1997. Three frequency effects in syntax. *Annual Meeting of the Berkeley Linguistics Society* 23(1). 378–388.
- Cappellaro, Chiara. 2013. Overabundance in diachrony: A case study. In Silvio Cruschina, Martin Maiden & John Charles Smith (Hrsg.), *The boundaries of pure morphology: Diachronic and synchronic perspectives*, 209–219. Oxford: Oxford University Press.
- Carroll, Ryan, Ragnar Svare & Joseph Salmons. 2012. Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics* 2(2). 153–172.
- Carstensen, Kai-Uwe, Christian Ebert, Cornelia Ebert, Susanne Jekat, Ralf Klambunde & Hagen Langer. 2010. *Computerlinguistik und Sprachtechnologie: Eine Einführung*. 3. Aufl. Heidelberg: Spektrum.
- Clahsen, Harald, Sonja Eisenbeiss, Meike Hadler & Ingrid Sonnenstuhl. 2001. The mental representation of inflected words: An experimental study of adjectives and verbs in German. *Language* 77(3). 510–543. DOI: 10.1353/lan.2001.0140.
- Clahsen, Harald, Sonja Eisenbeiss & Ingrid Sonnenstuhl-Henning. 1997. Morphological structure and the processing of inflected words. *Theoretical Linguistics* 23(3). 201–249.
- Clahsen, Harald, Meike Hadler & Helga Weyerts. 2004. Speeded production of inflected words in children and adults. *Journal of Child Language* 31(3). 683–712. DOI: 10.1017/S0305000904006506.
- Clahsen, Harald & Monika Rothweiler. 1993. Inflectional rules in children's grammars: Evidence from German participles. In Geert Booij & Jaap Marle (Hrsg.), *Yearbook of morphology 1992* (Yearbook of Morphology), 1–34. Dordrecht: Springer. DOI: 10.1007/978-94-017-3710-4_1.
- Clarke, Erik & Scott Sherrill-Mix. 2017. *ggbeeswarm: Categorical scatter (violin point) plots*. Version 0.6.0. <https://CRAN.R-project.org/package=ggbeeswarm> (4 August, 2020).

- Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. 2. Aufl. Chicago: Chicago University Press.
- Cumming, Geoff & Robert Calin-Jageman. 2017. *Introduction to the new statistics: Estimation, open science, and beyond*. New York/London: Routledge.
- Dąbrowska, Ewa. 2014. Words that go together: Measuring individual differences in native speakers' knowledge of collocations. *The Mental Lexicon* 9(3). 401–418. DOI: 10.1075/ml.9.3.02dab.
- Dahl, David B., David Scott, Charles Roosen, Arni Magnusson & Jonathan Swinton. 2019. *xtable: Export tables to LaTeX or HTML*. Version 1.8-4. <https://CRAN.R-project.org/package=xtable> (4 August, 2020).
- Dammel, Antje. 2008. Flexionsmorphologische Irregularität – ein graduelles Phänomen: Doch wie lassen sich Grade bestimmen? In Cornelia Stroh & Aina Urdrze (Hrsg.), *Morphologische Irregularität*, 1–28. Bochum: Brockmeyer.
- Dammel, Antje. 2011. *Konjugationsklassenwandel: Prinzipien des Ab-, Um-, Aus- und Aufbaus verbalflexivischer Allomorphie in germanischen Sprachen* (Studia Linguistica Germanica 103). Berlin/Boston: de Gruyter.
- Dammel, Antje & Jessica Nowak. 2011. „Schwarke“ Verben? Ein flexionsmorphologisches Spannungsfeld im Kontrast: Luxemburgisch – Deutsch. In Peter Gilles & Melanie Wagner (Hrsg.), *Linguistische und soziolinguistische Bausteine der Luxemburgistik* (Mikroglottika Minority Language Studies 4), 25–48. Frankfurt am Main: Peter Lang.
- Dammel, Antje & Jessica Nowak. 2014. Zur Reorganisation starker Verbklassen im West- und Ostfriesischen. *Philologia Frisica* 2014. 307–328.
- Dammel, Antje, Jessica Nowak & Mirjam Schmuck. 2010. Strong-verb paradigm leveling in four Germanic languages: A category frequency approach. *Journal of Germanic Linguistics* 22(04). 337–359. DOI: 10.1017/S1470542710000097.
- Dammel, Antje & Damaris Nübling. 2006. The superstable marker as an indicator of categorial weakness? *Folia Linguistica* XL(1-2). 97–113.
- Daugherty, Kim G. & Mark S. Seidenberg. 1994. Beyond rules and exceptions: A connectionist approach to inflectional morphology. In Susan D. Lima, Roberta L. Corrigan & Gregory K. Iverson (Hrsg.), *The reality of linguistic rules* (Studies in Language Companion Series 26), 353–388. Amsterdam/Philadelphia: Benjamins.
- Detges, Ulrich & Richard Waltereit. 2002. Grammaticalization vs. Reanalysis: A semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21(2). 151–195.
- Di Meola, Claudio. 2000. Deutsche Präpositionen im Überblick: Form, Stellung und Rektion. *Pandaemonium Germanicum: Revista de Estudos Germanísticos* 4. 321–368.

- Diependaele, Kevin, Marc Brysbaert & Peter Neri. 2012. How noisy is lexical decision? *Frontiers in Psychology* 3. 1–9. DOI: 10.3389/fpsyg.2012.00348.
- Diessel, Holger. 2017. Usage-based linguistics. In Mark Aronoff (Hrsg.), *Oxford research encyclopedia of linguistics*. New York: Oxford University Press.
- Divjak, Dagmar & Catherine L. Caldwell-Harris. 2015. Frequency and entrenchment. In Ewa Dąbrowska & Dagmar Divjak (Hrsg.), *Handbook of cognitive linguistics* (Handbücher zur Sprach- und Kommunikationswissenschaft 39), 53–75. Berlin/Boston: de Gruyter.
- Dowle, Matt & Arun Srinivasan. 2019. *data. Table: Extension of data. Frame*. Version 1.13.0. <https://CRAN.R-project.org/package=data.table> (4 August, 2020).
- Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ* (Studies in Linguistics and Philosophy 7). Dordrecht/Boston: D. Reidel.
- Duden. 2016. *Das Wörterbuch der sprachlichen Zweifelsfälle: Richtiges und gutes Deutsch*. 8. Aufl. (Der Duden in zwölf Bänden 9). Berlin: Dudenverlag.
- Duden. 2020. *Duden Online*. <https://www.duden.de> (22 März, 2015).
- Dürscheid, Christa, Stephan Elspaß, Arne Ziegler u. a. 2018. hauen. In *Variante(n)grammatik des Standarddeutschen*. Mannheim: Institut für Deutsche Sprache. <http://mediawiki.ids-mannheim.de/VarGra/index.php/Hauen>.
- Eisenberg, Peter. 1991. Syllabische Struktur und Wortakzent: Prinzipien der Prosodik deutscher Wörter. *Zeitschrift für Sprachwissenschaft* 10(1). 37–64.
- Eisenberg, Peter. 2013. *Das Wort*. 4. Aufl. (Grundriss der deutschen Grammatik 1). Stuttgart: Metzler.
- Eisenberg, Peter & Gerhard Voigt. 1990. Grammatikfehler? *Praxis Deutsch* 102. 10–15.
- Elio, Renée & John R. Anderson. 1984. The effects of information order and learning mode on schema abstraction. *Memory & Cognition* 12(1). 20–30.
- Ellis, Nick C. 2002a. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24(2). 143–188. DOI: 10.1017/S0272263102002024.
- Ellis, Nick C. 2002b. Reflections on frequency effects in language processing. *Studies in Second Language Acquisition* 24(2). 297–339. DOI: 10.1017/S0272263102002140.
- Ellis, Nick C. 2008. The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The Modern Language Journal* 92(2). 232–249.
- Ellis, Nick C. 2012. What can we count in language, and what counts in language acquisition, cognition, and use? In Stefan Thomas Gries & Dagmar Divjak

- (Hrsg.), *Frequency effects in language learning and processing* (Trends in Linguistics), 7–33. Berlin/Boston: de Gruyter.
- Ellis, Nick C. 2013. Second language acquisition. In Thomas Hoffmann & Graeme Trousdale (Hrsg.), *The Oxford handbook of construction grammar* (Oxford Handbooks in Linguistics), 365–378. Oxford: Oxford University Press.
- Ellis, Nick C. 2016. Frequency in language learning and language change: The contributions to this volume from a cognitive and psycholinguistic perspective. In Heike Behrens & Stefan Pfänder (Hrsg.), *Experience counts: Frequency effects in language* (Linguae & Litterae 54), 239–254. Berlin/Boston: de Gruyter.
- Ellis, Nick C. & Fernando Ferreira-Junior. 2009. Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics* 7. 188–221. DOI: 10.1075/arcl.7.08ell.
- Ellis, Nick C., Matthew B. O'Donnell & Ute Römer. 2014. Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality. *Linguistic Approaches to Bilingualism* 4(4). 405–431.
- Ellis, Nick C. & Richard Schmidt. 1998. Rules or associations in the acquisition of morphology? Frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes* 13(2/3). 307–336.
- Elman, Jeffrey L. 1992. Grammatical structure and distributed representations. In Steven Davis (Hrsg.), *Connectionism: Theory and practice*, 138–194. Oxford: Oxford University Press.
- Enger, Hans-Olav & Tore Nessel. 2011. Constraints on diachronic development: The animacy hierarchy and the relevance constraint. *Language Typology and Universals* 64(3). 193–212. DOI: 10.1002/9781118266892.ch1.
- Epstein, Richard. 1994. The development of the definite article in French. In William Pagliuca (Hrsg.), *Perspectives on grammaticalization* (Current Issues in Linguistic Theory 109), 63–79. Philadelphia/Amsterdam: Benjamins.
- Fabricius-Hansen, Catherine. 1977. Zur Klassifizierung der starken Verben im Neuhochdeutschen. *Deutsche Sprache* 5. 193–205.
- Fenk-Oczlon, Gertraud. 1990. Ökonomieprinzipien in Kognition und Kommunikation. In Norbert Boretzky, Werner Enninger & Thomas Stolz (Hrsg.), *Spielarten der Natürlichkeit – Spielarten der Ökonomie: Beiträge zum 5. Essener Kolloquium über Grammatikalisierung* (Bochum–Essener Beiträge zur Sprachwandelforschung 8), 37–51. Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Fenk-Oczlon, Gertraud. 1991. Frequenz und Kognition – Frequenz und Markiertheit. *Folia Linguistica* 25(3-4). 361–394.
- Field, Andy, Jeremy Miles & Zoë Field. 2012. *Discovering statistics using R*. Los Angeles: Sage.

- Fillmore, Charles. 1977. Scenes-and-frames-semantics. In Antonio Zampolli (Hrsg.), *Linguistic Structures Processing*, 55–81. Amsterdam/New York: North Holland Publishing.
- Fine, Alex & T. Florian Jaeger. 2013. Syntactic priming in language comprehension allows linguistic expectations to converge on the statistics of the input. *Proceedings of the Annual Meeting of the Cognitive Science Society* 35. 2279–2284.
- Fischer, Hanna. 2018. *Präteritumschwund im Deutschen: Dokumentation und Erklärung eines Verdrängungsprozesses*. Berlin/Boston: de Gruyter.
- Flexner, Arthur J. & Gordon H. Bower. 1975. Further evidence regarding instructional effects on frequency judgments. *Bulletin of the Psychonomic Society* 6(3). 321–324.
- Flick, Johanna. 2016. Der *am*-Progressiv und parallele *am V-en sein*-Konstruktionen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 138(2). 479. DOI: 10.1515/bgsl-2016-0017.
- Flick, Johanna. 2020. *Die Entwicklung des Definitartikels im Althochdeutschen: Eine kognitiv-linguistische Korpusuntersuchung* (Empirically Oriented Theoretical Morphology and Syntax 6). Berlin: Language Science Press. DOI: 10.5281/zenodo.3932780.
- Forster, Kenneth I. 2010. Methodological and analytic frontiers in lexical research (part I). *The Mental Lexicon* 5(3). 347–357. DOI: 10.1075/ml.5.3.05for.
- Forster, Kenneth I. & Jonathan C. Forster. 2003. DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35(1). 116–124. DOI: 10.3758/BF03195503.
- Forster, Kenneth I., Christine Guerrera & Lisa Elliot. 2009. The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods* 41(1). 163–171. DOI: 10.3758/BRM.41.1.163.
- Fowler, Carol A. & Jonathan Housum. 1987. Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26. 489–504.
- Frisch, Stefan A., Nathan R. Large, Bushra Zawaydeh & David B. Pisoni. 2001. Emergent phonotactic generalizations in English and Arabic. In Joan L. Bybee & Paul Hopper (Hrsg.), *Frequency and the emergence of linguistic structure* (Typological Studies in Language 45), 159–179. Amsterdam/Philadelphia: Benjamins.
- Fromkin, Victoria, Robert Rodman & Nina M. Hyams. 2014. *An introduction to language*. 10. Aufl. Boston: Wadsworth.

- Gardner, Michael K., Ernst Z. Rothkopf, Richard Lapan & Toby Lafferty. 1987. The word frequency effect in lexical decision: Finding a frequency-based component. *Memory & Cognition* 15(1). 24–28.
- Gernsbacher, Morton Ann. 1984. Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General* 113(2). 256–281.
- Gervain, Judit, Marina Nespou, Reiko Mazuka, Ryota Horie & Jacques Mehler. 2008. Bootstrapping word order in prelexical infants: A Japanese-Italian cross-linguistic study. *Cognitive Psychology* 57(1). 56–74. DOI: 10.1016/j.cogpsych.2007.12.001.
- Gervain, Judit & Janet F. Werker. 2013. Prosody cues word order in 7-month-old bilingual infants. *Nature Communications* 4. 1490. DOI: 10.1038/ncomms2430.
- Gillmann, Melitta. 2011. Die Grammatikalisierung des *sein*-Perfekts: Eine korpuslinguistische Untersuchung zur Hilfsverbselektion der Bewegungsverben im Deutschen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 133(2). 203–234. DOI: 10.1515/bgsl.2011.027.
- Gillmann, Melitta. 2016. *Perfektkonstruktionen mit haben und sein: Eine Korpusuntersuchung im Althochdeutschen, Altsächsischen und Neuhochdeutschen* (Studia Linguistica Germanica 128). Boston: de Gruyter.
- Gillmann, Melitta. 2018. Manner of motion and semantic transitivity: A usage-based perspective on change and continuity in the system of the German perfect auxiliaries *haben* and *sein*. In Antje Dammel, Matthias Eitelmann & Mirjam Schmuck (Hrsg.), *Reorganising grammatical variation: Diachronic studies in retention, redistribution and refunctionalisation of linguistic variants*. (Studies in Language Companion Series 203), 234–267. Amsterdam/Philadelphia: Benjamins.
- Goldberg, Adele E. 1998. Patterns of experience in patterns of language. In Michael Tomasello (Hrsg.), *The new psychology of language*, 203–219. Mahwah: Lawrence Erlbaum Associates.
- Goldberg, Adele E. 2006. *Constructions at work*. Oxford: Oxford University Press.
- Goldberg, Adele E. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton: Princeton University Press.
- Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical Access. *Psychological Review* 105(2). 251–279.
- Gouvea, Ana C., Colin Phillips, Nina Kazanina & David Poeppel. 2010. The linguistic processes underlying the P600. *Language and Cognitive Processes* 25(2). 149–188. DOI: 10.1080/01690960902965951.
- Grainger, Jonathan. 1990. Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language* 29. 228–244.

- Grainger, Jonathan, Kevin J. O'Regan, Arthur M. Jacobs & Juan Segui. 1989. On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psycholinguistics* 45(3). 189–195.
- Haberlandt, Karl. 1994. Methods in reading research. In Morton Ann Gernsbacher (Hrsg.), *Handbook of psycholinguistics*, 1–31. San Diego: Academic Press.
- Hahne, Anja & Angela D. Friederici. 1999. Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience* 11(2). 194–205.
- Haiman, John. 1994. Ritualization and the development of language. In William Pagliuca (Hrsg.), *Perspectives on grammaticalization* (Current Issues in Linguistic Theory 109), 3–28. Philadelphia/Amsterdam: Benjamins. DOI: 10.1075/cilt.109.07hai.
- Hanulikova, Adriana, Petra M. van Alphen, Merel M. van Goch & Andrea Weber. 2012. When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience* 24(4). 878–887. DOI: 10.1162/jocn_a_00103.
- Hare, Mary L., Michael Ford & William Marslen-Wilson. 2001. Ambiguity and frequency effects in regular verb inflection. In Joan L. Bybee & Paul Hopper (Hrsg.), *Frequency and the emergence of linguistic structure* (Typological Studies in Language 45), 181–200. Amsterdam/Philadelphia: Benjamins. DOI: 10.1075/tsl.45.10har.
- Harnisch, Rüdiger. 1988. Natürliche Morphologie und morphologische Ökonomie: Ein Vermittlungsversuch angesichts der Morphologien natürlicher Sprachen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 41(4). 426–437.
- Harrell, Frank E. Junior & Charles Dupont. 2019. *Hmisc: Harrell miscellaneous*. Version 4.4-0. <https://CRAN.R-project.org/package=Hmisc> (4 August, 2020).
- Hasher, Lynn & Rose T. Zacks. 1984. Automatic processing of fundamental information. *American Psychologist* 39(12). 1372–1388.
- Helfman, Gene S. 2009. *The diversity of fishes: Biology, evolution, and ecology*. 2. Aufl. Oxford: Wiley-Blackwell.
- Hinze, Christian & Klaus-Michael Köpcke. 2007. Was wissen Grundschüler über die Verwendung der Perfektauxiliare *haben* und *sein*? In Klaus-Michael Köpcke & Arne Ziegler (Hrsg.), *Grammatik in der Universität und für die Schule* (Reihe Germanistische Linguistik), 95–128. Tübingen: Max Niemeyer Verlag.
- Höhle, Barbara (Hrsg.). 2010. *Psycholinguistik* (Studienbuch Sprachwissenschaft). Berlin: Akademie-Verlag.
- Hopper, Paul & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.

- Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3, S. 651–674). Version 1.3-5. <http://www.biomedcentral.com/1471-2105/9/307> (4 August, 2020).
- Hotzenköcherle, Rudolf. 1962. Entwicklungsgeschichtliche Grundzüge des Neuhochdeutschen. *Wirkendes Wort* 12. 321–331.
- Hundt, Markus. 2000. Deutschlands meiste Kreditkarte: Probleme der Wortartenklassifikation. *Deutsche Sprache* 28. 1–24.
- Indefrey, Peter, Colin Brown, Peter Hagoort, Hans Herzog, M. Sach & Rüdiger J. Seitz. 1997. A PET study of cerebral activation patterns induced by verb inflection. *Neuroimage* 5. 548.
- Jackendoff, Ray. 2004. Categorization, fuzziness, and family resemblances. In Bas Aarts (Hrsg.), *Fuzzy grammar*, 109–129. Oxford/New York: Oxford University Press.
- Jaeger, Jeri J., Alan H. Lockwood, David L. Kemmerer, Robert D. van Valin, Brian W. Murphy & Hanif G. Khalak. 1996. A positron emission tomographic study of regular and irregular verb morphology in English. *Language* 72(3). 451–497. DOI: 10.2307/416276.
- Jegerski, Jill. 2014. Self-paced-reading. In Bill VanPatten & Jill Jegerski (Hrsg.), *Research methods in second language psycholinguistics* (Second Language Acquisition Research Series), 20–49. New York: Routledge.
- Jespersen, Ole. 1917. *Negation in English and other languages*. Kopenhagen: Høst.
- Joeres, Rolf. 1996. *Der Friede oder der Frieden*: Ein Normproblem der Substantivflexion. *Sprachwissenschaft* 21. 301–336.
- Johnson, Keith. 2005. Speaker normalization in speech perception. In David B. Pisoni & Robert Ellis Remez (Hrsg.), *The handbook of speech perception* (Blackwell Handbooks in Linguistics), 363–389. Malden: Blackwell. DOI: 10.1002/9780470757024.ch15.
- Johnson, Keith. 2007. Decisions and mechanisms in exemplar-based phonology. In Maria-Josep Sole, Patrice Speeter Beddor & Manjari Ohala (Hrsg.), *Experimental approaches to phonology*, 25–40. Oxford: Oxford University Press.
- Johnson, Matt A. & Adele E. Goldberg. 2013. Evidence for automatic accessing of constructional meaning: Jabberwocky sentences prime associated verbs. *Language and Cognitive Processes* 28(10). 1439–1452. DOI: 10.1080/01690965.2012.717632.
- Jones, Lara L. & Zachary Estes. 2012. Lexical priming: Associative, semantic, and thematic influences on word recognition. In James S. Adelman (Hrsg.), *Visual word recognition*, 44–72. Hove: Psychology Press.

- Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20. 137–194.
- Kaan, Edith, Anthony Harris, Edward Gibson & Phillip Holcomb. 2000. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15(2). 159–201. DOI: 10.1080/016909600386084.
- Kaan, Edith & Tamara Yvonne Swaab. 2003. Repair, revision and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience* 15(1). 98–110.
- Keller, Frank, Subahshini Gunasekharan & Neil Mayo. 2009. Timing accuracy of web experiments: A case study using the WebExp software package. *Behavior Research Methods* 41(1). 1–12.
- Kelly, Michael H. & J. Kathryn Bock. 1986. Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language* 25. 59–74.
- Kemmer, Suzanne & Michael Barlow. 2000. Introduction: A usage-based conception of language. In Michael Barlow & Suzanne Kemmer (Hrsg.), *Usage-based models of language*, vii–xxviii. Stanford: Center for the Study of Language and Information.
- Keuleers, Emmanuel, Kevin Diependaele & Marc Brysbaert. 2010. Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology* 1. 1–15. DOI: 10.3389/fpsyg.2010.00174.
- Kirchner, Robert, Roger K. Moore & Tsung-Ying Chen. 2010. Computing phonological generalization over real speech exemplars. *Journal of Phonetics* 38(4). 540–547. DOI: 10.1016/j.wocn.2010.07.005.
- Kleiber, Georges. 1993. *Prototypensemantik: Eine Einführung* (Narr-Studienbücher). Tübingen: Narr.
- Köpcke, Klaus-Michael. 1993. *Schemata bei der Pluralbildung im Deutschen: Versuch einer kognitiven Morphologie* (Studien zur deutschen Grammatik 47). Tübingen: Narr.
- Köpcke, Klaus-Michael. 1995. Die Klassifikation der schwachen Maskulina in der deutschen Gegenwartssprache: Ein Beispiel für die Leistungsfähigkeit der Prototypentheorie. *Zeitschrift für Sprachwissenschaft* 14. 159–180.
- Köpcke, Klaus-Michael. 1999. Prototypisch starke und schwache Verben der deutschen Gegenwartssprache. In Matthias Butt & Nanna Fuhrhop (Hrsg.), *Variation und Stabilität in der Wortstruktur* (Germanistische Linguistik 141-142), 45–60. Hildesheim/Zürich/New York: Georg Olms.
- Köpcke, Klaus-Michael. 2000a. Chaos und Ordnung: Zur semantischen Remotivierung einer Deklinationsklasse im Übergang vom Mhd. zum Nhd. In Andreas

- Bittner, Dagmar Bittner & Klaus-Michael Köpcke (Hrsg.), *Angemessene Strukturen*, 107–122. Hildesheim/Zürich/New York: Georg Olms.
- Köpcke, Klaus-Michael. 2000b. Starkes, Schwaches und Gemischtes in der Substantivflexion des Deutschen: Was weiß der Sprecher über die Deklinationsparadigmen? In Rolf Thieroff, Matthias Tamrat, Nanna Fuhrhop & Oliver Teuber (Hrsg.), *Deutsche Grammatik in Theorie und Praxis*. 155–170. Tübingen: Niemeyer.
- Köpcke, Klaus-Michael. 2005. „Die Prinzessin küsst den Prinz“: Fehler oder gelebter Sprachwandel? *Didaktik Deutsch* 18. 67–83.
- Köpcke, Klaus-Michael, Sarah Schimke & Verena Wecker. 2021. Processing of German noun plurals: Evidence for first- and second-order schemata. *Word Structure* 14(1). 1–24. DOI: 10.3366/word.2021.0173.
- Köpcke, Klaus-Michael & Verena Wecker. 2017. Source- and product-oriented strategies in L2 acquisition of plural marking in German. *Morphology* 27(1). 77–103. DOI: 10.1007/s11525-016-9292-5.
- Köpcke, Klaus-Michael & David A. Zubin. 2018. Motive für die Genusvariation von *der/die Krake* und *der/die Python*. In Kerstin Kazzazi, Karin Luttermann, Sabine Wahl & Thomas A. Fritz (Hrsg.), *Worte über Wörter: Festschrift zu Ehren von Elke Ronneberger-Sibold* (Stauffenburg Festschriften), 215–232. Tübingen: Stauffenburg.
- Krischke, Wolfgang. 2012. Des Menschens Genitive: Normabweichende Genitiv-Varianten bei schwachen Maskulina. *Linguistik online* 53(3). 55–84.
- Krott, Andrea & Riadh Lebib. 2013. Electrophysiological evidence for a neural substrate of morphological rule application in correct wordforms. *Brain Research* 1496. 70–83. DOI: 10.1016/j.brainres.2012.12.012.
- Kuperman, Victor, Denis Drieghe, Emmanuel Keuleers & Marc Brysbaert. 2013. How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology* 66(3). 563–580.
- Kusová, Jana. 2014. *Morphologische Varianz der peripheren Substantivparadigmen im geschriebenen Gegenwart Deutsch: Schwache Maskulina, starke Feminina und gemischte Substantive*. Wien: Praesens Verlag.
- Kutas, Marta & Steven A. Hillyard. 1983. Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition* 11(5). 539–550. DOI: 10.3758/BF03196991.
- Labov, William. 2004 [1973]. The boundaries of words and their meanings. In Bas Aarts (Hrsg.), *Fuzzy grammar*, 67–89. Oxford/New York: Oxford University Press.

- Lakoff, George. 1973. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic* 2. 458–508.
- Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago/London: University of Chicago Press.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar: Volume I: Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1991. *Concept, image, and symbol: The cognitive basis of grammar* (Cognitive Linguistics Research 1). Berlin/Boston: de Gruyter.
- Langacker, Ronald W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Langacker, Ronald W. 2014. *Essentials of cognitive grammar*. Oxford: Oxford University Press.
- Leibniz-Institut für Deutsche Sprache. 2019a. *COSMAS I/II (corpus search, management and analysis system)*. <http://www.ids-mannheim.de/cosmas2> (4 August, 2020).
- Leibniz-Institut für Deutsche Sprache. 2019b. *Das Deutsche Referenzkorpus DeReKo*. <http://www.ids-mannheim.de/kl/projekte/korpora/> (4 August, 2020).
- Leiner, Daniel J. 2019. *SoSci Survey*. Version: 3.1.06. <https://www.soscisurvey.de> (4 August, 2020).
- Leiss, Elisabeth. 1992. *Die Verbalkategorien des Deutschen: Ein Beitrag zur Theorie der sprachlichen Kategorisierung*. Berlin/Boston: de Gruyter.
- Lerche, Anne. 2019. *Frequency effects on entrenchment: Converging evidence from ongoing language change*. Freiburg: Albert-Ludwigs-Universität. (Diss.). DOI: 10.6094/UNIFR/149608.
- Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam/Philadelphia: Benjamins.
- Lewis, Michael B., Simon Gerhand & Hadyn D. Ellis. 2001. Re-evaluating age-of-acquisition effects: Are they simply cumulative-frequency effects? *Cognition* 78. 189–205.
- Lexer, Matthias. 1872–1878. *Mittelhochdeutsches Handwörterbuch*. <http://woerterbuchnetz.de> (4 August, 2020).
- Lüdecke, Daniel. 2018. ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software* 3(26). Version: 0.15.1, 772. <https://www.rdocumentation.org/packages/ggeffects/versions/0.14.3> (4 August, 2020).
- Mangasser-Wahl, Martina. 2000. Roschs Prototypentheorie: Eine Entwicklung in drei Phasen. In Martina Mangasser-Wahl (Hrsg.), *Prototypentheorie in der Linguistik* (Stauffenburg Linguistik 10), 15–31. Tübingen: Stauffenburg.

- Mater, Erich. 1967. *Deutsche Verben 3: Gesamtverzeichnis der Grundwörter/Stellung der Kompositionsglieder*. Leipzig: VEB Bibliographisches Institut.
- McDonald, Scott A. & Richard C. Shillcock. 2001. Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech* 44(3). 295–323.
- McDonald, Scott A. & Richard C. Shillcock. 2003. Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research* 43(16). 1735–1751. DOI: 10.1016/S0042-6989(03)00237-2.
- McDonough, Kim & Pavel Trofimovich. 2012. How to use psycholinguistic methodologies for comprehension and production. In Alison Mackey & Susan M. Gass (Hrsg.), *Research methods in second language acquisition. A practical guide* (Guides to Research Methods in Language and Linguistics), 117–138. Chichester/Malden: Wiley-Blackwell.
- McKinney, Wes. 2010. Data structures for statistical computing in Python. In Stéfan van der Walt & Jarrod Millman (Hrsg.), *Proceedings of the 9th Python in Science conference (SCIPY 2010)*, 56–61. DOI: 10.25080/Majora-92bf1922-00a.
- Mitchel, Don C. 2013. On-line methods in language processing: Introduction and historical review. In Manuel Carreiras & Clifton Charles (Hrsg.), *The on-line study of sentence comprehension*, 15–32. New York/Hove: Psychology Press.
- Mitchel, Tom M. 1997. *Machine learning*. Singapur: McGraw-Hill.
- Morrison, Catriona M. & Andrew W. Ellis. 1995. Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21(1). 116–133.
- Morrison, Catriona M. & Andrew W. Ellis. 2000. Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology* 91. 167–180.
- Nakagawa, Shinichi, Paul C. D. Johnson & Holger Schielzeth. 2017. The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface* 14(134). 1–11. DOI: 10.1098/rsif.2017.0213.
- Nash, John C. & Ravi Varadhan. 2011. Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software* 43(9). Version: 2020-4.2, 1–14. <http://www.jstatsoft.org/v43/i09/> (4 August, 2020).
- Nesset, Tore. 2008. *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface*. Berlin: de Gruyter.
- Nesset, Tore & Laura A. Janda. 2010. Paradigm structure: Evidence from Russian suffix shift. *Cognitive Linguistics* 21(4). 699–725.
- NLTK-Team. 2020. *NLTK: Natural language toolkit*. Version: 3.5. <http://www.nltk.org> (4 August, 2020).

- Norris, Dennis. 2006. The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review* 113(2). 327–357. DOI: 10.1037/0033-295X.113.2.327.
- Nowak, Jessica. 2010a. Im Spannungsfeld starker und schwacher Verben: Zur Entstehung einer „8. Ablautreihe“ im Deutschen, Niederländischen und Luxemburgischen. In Antje Dammel, Sebastian Kürschner & Damaris Nübling (Hrsg.), *Kontrastive Germanistische Linguistik* (Germanistische Linguistik 206–209), 429–472. Hildesheim/Zürich/New York: Georg Olms.
- Nowak, Jessica. 2010b. On the emergence of an eighth ablaut class in German and Dutch. *Journal of Germanic Linguistics* 22(4). 361–380.
- Nowak, Jessica. 2011. Zur Herausbildung semantischer Differenzierungen bei Konjugationsdubletten. In Jörg Riecke (Hrsg.), *Historische Semantik* (Jahrbuch für germanistische Sprachgeschichte), 312–325. Berlin/Boston: de Gruyter.
- Nowak, Jessica. 2013. *Spinnen – spinn? – gesponnen*: Die Alternanz *x-o-o* als Alternative zum „Schwachwerden“. In Petra Maria Vogel (Hrsg.), *Sprachwandel im Neuhochdeutschen* (Jahrbuch für germanistische Sprachgeschichte), 170–185. Berlin/Boston: de Gruyter.
- Nowak, Jessica. 2015. *Zur Legitimation einer 8. Ablautreihe: Eine kontrastive Analyse zu ihrer Entstehung im Deutschen, Niederländischen und Luxemburgischen*. Hildesheim/Zürich/New York: Olms.
- Nowak, Jessica. 2016. Die Ablautalternanz *x-o-o* als partielle Regularisierungsstrategie starker Verben im Deutschen. In Andreas Bittner & Klaus-Michael Köpcke (Hrsg.), *Regularität und Irregularität in Phonologie und Morphologie* (Lingua Historica Germanica: Studien und Quellen zur Geschichte der deutschen Sprache und Literatur 13), 127–152. Berlin/Boston: de Gruyter. DOI: 10.1515/9783110486094-008.
- Nowak, Jessica. 2018. Ablaut reorganisation: The case of German *x-o-o*. In Antje Dammel, Matthias Eitelmann & Mirjam Schmuck (Hrsg.), *Reorganising grammatical variation: Diachronic studies in retention, redistribution and re-functionalisation of linguistic variants*. (Panion Series 203), 151–176. Amsterdam/Philadelphia: Benjamins.
- Nübling, Damaris. 1998. Wie die Alten sangen: Zur Rolle von Frequenz und Allomorphie beim präteritalen Numerusausgleich im Frühneuhochdeutschen. *Zeitschrift für Sprachwissenschaft* 17(2). 185–203.
- Nübling, Damaris. 2000. *Prinzipien der Irregularisierung: Eine kontrastive Analyse von zehn Verben in zehn germanischen Sprachen* (Linguistische Arbeiten 415). Berlin/Boston: de Gruyter. DOI: 10.1515/9783110915082.

- Nübling, Damaris. 2001. Von *oh mein Jesus!* zu *oje!* Der Interjektionalisierungspfad von der sekundären zur primären Interjektion. *Deutsche Sprache* 29. 20–45.
- Nübling, Damaris. 2008. Was tun mit den Flexionsklassen? Deklinationsklassen und ihr Wandel im Deutschen und seinen Dialekten. *Zeitschrift für Dialektologie und Linguistik* 75(3). 282–330.
- Nübling, Damaris. 2016. Flexionsklassen als morphologischer Ballast? In Andreas Bittner & Klaus-Michael Köpcke (Hrsg.), *Regularität und Irregularität in Phonetik und Morphologie* (Lingua Historica Germanica 13), 153–176. Berlin/Boston: de Gruyter.
- Nübling, Damaris & Antje Dammel. 2004. Relevanzgesteuerter morphologischer Umbau im Frühneuhochdeutschen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 126(2). 177–207.
- Nübling, Damaris, Antje Dammel, Janet Duke & Renata Szczepaniak. 2013. *Historische Sprachwissenschaft des Deutschen: Eine Einführung in die Prinzipien des Sprachwandels*. 5. Aufl. (Narr Studienbücher). Tübingen: Narr.
- Osterhout, Lee, Philip Holcomb & David A. Swinney. 1994. Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20(4). 786–803.
- Paul, Hermann. 1917. *Deutsche Grammatik: Band II. Teil III: Die Flexionslehre*. Halle: Max Niemeyer.
- Pavlov, Vladimir. 1995. *Die Deklination der Substantive im Deutschen: Synchronie und Diachronie*. Frankfurt am Main: Lang.
- Pedersen, Thomas Lin. 2020. *Patchwork: The composer of plots*. Version 1.0.1. <https://github.com/thomasp85/patchwork> (4 August, 2020).
- Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman & Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51(51, S. 195–203). Version: 3.1.5, 195–203. DOI: 10.3758/s13428-018-01193-y. <https://www.psychopy.org/> (4 August, 2020).
- Penke, Martina, Helga Weyerts, Matthias Gross, Elke Zander, Thomas F. Münte & Harald Clahsen. 1997. How the brain processes complex words: An event-related potential study of German verb inflections. *Cognitive Brain Research* 6. 37–52.
- Perek, Florent & Adele E. Goldberg. 2017. Linguistic generalization on the basis of function and constraints on the basis of statistical preemption. *Cognition* 168. 276–293.

- Pfänder, Stefan & Heike Behrens. 2016. Experience counts: An introduction to frequency effects in language. In Heike Behrens & Stefan Pfänder (Hrsg.), *Experience counts* (Linguae & Litterae 54), 1–20. Berlin/Boston: de Gruyter.
- Pfeifer, Wolfgang. 2000. *Etymologisches Wörterbuch des Deutschen*. 5. Aufl. München: Deutscher Taschenbuch Verlag.
- Pinker, Steven & Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28. 73–193.
- Pinker, Steven & Alan Prince. 1991. Regular and irregular morphology and the psychological status of rules of grammar. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 17. 230–251. DOI: 10.3765/bls.v17i0.1624.
- Plunkett, Kim & Virginia Marchman. 1991. U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* 38. 43–102.
- Poitou, Jacques. 2004a. Prototypentheorie und Flexionsmorphologie. *Linguistik online* 19(2). 71–93.
- Poitou, Jacques. 2004b. Prototypentheorie und lexikalische Semantik. *LYLIA (Lyon-Linguistique-Allemande)* 3. 1–24.
- Pollack, Irwin, Herbert Rubenstein & Louis Decker. 1959. Intelligibility of known and unknown message sets. *Journal of the Acoustical Society of America* 31. 273–279.
- Primus, Beatrice. 2012. *Semantische Rollen*. Heidelberg: Winter.
- PsychoPy Team. 2019. *Pavlovia*. <https://pavlovia.org> (4 August, 2020).
- R Core Team. 2019. *R: A language and environment for statistical computing*. Version: 3.6.0. <https://www.R-project.org> (4 August, 2020).
- Randall, Janet, Angeliek van Hout, Jürgen Weissenborn & Harald Baayen. 2004. Acquiring unaccusativity: A cross-linguistic look. In Artemis Alexiadou, Elena Anagnostopoulou & Martin Everaert (Hrsg.), *The unaccusativity puzzle*, 332–353. Oxford: Oxford University Press.
- Rayner, Keith & Susan A. Duffy. 1986. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition* 14(3). 191–201. DOI: 10.3758/BF03197692.
- Rayner, Keith, Gretchen Kambe & Susan A. Duffy. 2000. The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology* 53(4). 1061–1080. DOI: 10.1080/713755934.
- Roehm, Dietmar, Ina D. Bornkessel, Hubert Haider & Matthias Schlesewsky. 2005. When case meets agreement: Event-related potential effects for morphology-based conflict resolution in human language comprehension. *Neuroreport* 16(8). 875–878.

- Roehm, Dietmar, Antonella Sorace & Ina Bornkessel-Schlesewsky. 2013. Processing flexible form-to-meaning mappings: Evidence for enriched composition as opposed to indeterminacy. *Language and Cognitive Processes* 28(8). 1244–1274. DOI: 10.1080/01690965.2012.712143.
- Rogers, Timothy T. & James L. McClelland. 2003. *Semantic cognition: A parallel distributed processing approach*. Cambridge: MIT Press.
- Ronneberger-Sibold, Elke. 2020. Wie dekliniert man *der Truntáke* in Bayern? Schemata und Prototypen in der Deklination von Kunstwörtern in verschiedenen deutschen Varietäten. In Anja Binanzer, Jana Gamper & Verena Wecker (Hrsg.), *Prototypen – Schemata – Konstruktionen* (Reihe Germanistische Linguistik), 411–438. Boston: de Gruyter.
- Rosch, Eleanor. 1973. Natural categories. *Cognitive Psychology* 4. 328–350.
- Rosch, Eleanor. 1975a. Cognitive reference points. *Cognitive Psychology* 7. 532–547.
- Rosch, Eleanor. 1975b. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* 104(3). 192–233.
- Rosch, Eleanor. 2004 [1978]. Principles of categorization. In Bas Aarts (Hrsg.), *Fuzzy grammar*, 91–108. Oxford/New York: Oxford University Press.
- Rosch, Eleanor & Carolyn B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7. 573–605.
- Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson & Penny Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8(3). 382–439. DOI: 10.1016/0010-0285(76)90013-X.
- Rosch, Eleanor, Carol Simpson & R. Scott Miller. 1976. Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance* 2(4). 491–502. DOI: 10.1037//0096-1523.2.4.491.
- Ross, Brian H. & Valerie S. Makin. 1999. Prototype versus exemplar models in cognition. In Robert J. Sternberg (Hrsg.), *The nature of cognition*, 205–241. Cambridge: MIT Press.
- RStudio Team. 2016. *RStudio: Integrated development for R*. Version: 1.1.453. <http://www.rstudio.com/> (4 August, 2020).
- Rubenstein, Herbert, Lonnie Garfield & Jane A. Millikan. 1970. Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior* 9(5). 487–494.
- Rubin, David C. 1974. The subjective estimation of relative syllable frequency. *Perception & Psycholinguistics* 16(1). 193–196.
- Rudis, Bob & Dave Gandy. 2019. *Waffle: Create waffle chart visualizations*. Version 1.0.1. <https://gitlab.com/hrbrmstr/waffle> (4 August, 2020).

- Rumelhart, David E. 1980. Schemata: The building blocks of cognition. In Rand J. Spiro (Hrsg.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education*, 33–58. Hillsdale: Erlbaum.
- Rumelhart, David E. & James L. McClelland. 1986. On learning the past tenses of English verbs. In James L. McClelland & David E. Rumelhart (Hrsg.), *Parallel distributed processing*, 216–271. Cambridge: MIT Press.
- Ruoff, Arno. 1981. *Häufigkeitswörterbuch gesprochener Sprache: Gesondert nach Wortarten alphabetisch, rückläufig-alphabetisch und nach Häufigkeit geordnet*. Berlin/Boston: de Gruyter.
- Saffran, Jenny R. 2003. Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science* 12(4). 110–114.
- Saffran, Jenny R., Richard N. Aslin & Elissa L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274(5294). 1926–1928.
- Sandberg, Bengt. 1998. Zur Distribution *Funken – Funke* nach semanto-syntaktischen Merkmalen. *Sprachwissenschaft* 23. 381–412.
- Schäfer, Lea. 2023. *Syntax and morphology of Yiddish dialects. Findings from the language and culture archive of Ashkenazic Jewry* (Deutsche Dialektgeographie 132). Hildesheim/Zürich/New York: Georg Olms Verlag.
- Schäfer, Roland. 2019. Prototype-driven alternations: The case of German weak nouns. *Corpus Linguistics and Linguistic Theory* 15(2). 383–417. DOI: 10.1515/cllt-2015-0051.
- Scherr, Elisabeth. 2018. Bildung des Perfekts. In Christa Dürscheid, Stephan Elspaß & Arne Ziegler (Hrsg.), *Variantengrammatik des Standarddeutschen*. Mannheim: Institut für Deutsche Sprache. http://mediawiki.ids-mannheim.de/VarGra/index.php/Bildung_des_Perfekts.
- Schmid, Hans-Jörg. 2000. Methodik der Prototypentheorie. In Martina Mangasser-Wahl (Hrsg.), *Prototypentheorie in der Linguistik* (Stauffenburg Linguistik 10), 33–53. Tübingen: Stauffenburg.
- Schmitt, Eleonore. 2019a. Conflicting evidence for mental schemas in language production and processing. In Constanze Juchem-Grundmann, Michael Pleyer & Monika Pleyer (Hrsg.), *Jahrbuch der Deutschen Gesellschaft für Kognitive Linguistik*, 157–176. Berlin/Boston: de Gruyter.
- Schmitt, Eleonore. 2019b. How do cases of doubt cause doubts? The mismatch between acceptance and processing as a cause for cases of doubt. In Eleonore Schmitt, Renata Szczepaniak & Annika Vieregge (Hrsg.), *Sprachliche Zweifelsfälle: Definition, Erforschung, Implementierung* (Germanistische Linguistik), 97–136. Hildesheim/Zürich/New York: Georg Olms.

- Schmuck, Mirjam. 2010. Relevanzgesteuerter verbalmorphologischer Wandel im Deutschen und Schwedischen. In Antje Dammel, Sebastian Kürschner & Damaris Nübling (Hrsg.), *Kontrastive Germanistische Linguistik* (Germanistische Linguistik 206-209), 523–551. Hildesheim/Zürich/New York: Georg Olms Verlag.
- Schneider, Ulrike. 2014. *Frequency, chunks and hesitations: A usage-based analysis of chunking in English*. Freiburg: New Ideas in Human Interaction Studies.
- Seidenberg, Mark S. 1992. Connectionism without tears. In Steven Davis (Hrsg.), *Connectionism: Theory and practice*, 85–137. Oxford: Oxford University Press.
- Shannon, Thomas F. 1992. Split intransitivity in German and Dutch: Semantic and pragmatic parameters. In Rosina Lippi-Green (Hrsg.), *Recent developments in Germanic linguistics* (Current Issues in Linguistic Theory 93), 97–113. Amsterdam/Philadelphia: Benjamins.
- Shannon, Thomas F. 1995. Toward a cognitive explanation of perfect auxiliary variation: Some modal and aspectual effects in the history of Germanic. *American Journal of Germanic Linguistics & Literatures* 7(2). 129–163.
- Shapiro, Bernard J. 1969. The subjective estimation of relative word frequency. *Journal of Verbal Learning and Verbal Behavior* 8. 248–251.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (Hrsg.), *Grammatical categories in Australian languages*, 112–171. New Jersey: Humanity Press.
- Singmann, Henrik, Ben Bolker, Jake Westfall & Frederik Aust. 2019. *Afex: Analysis of factorial experiments*. Version 0.27-2. <https://CRAN.R-project.org/package=afex> (4 August, 2020).
- Smolka, Eva, Patrick H. Khader, Richard Wiese, Pienie Zwitserlood & Frank Röslér. 2013. Electrophysiological evidence for the continuous processing of linguistic categories of regular and irregular verb inflection in German. *Journal of Cognitive Neuroscience* 25(8). 1284–1304. DOI: 10.1162/jocn_a_00384.
- Solms, Hans-Joachim. 1984. *Die morphologischen Veränderungen der Stammvokale der starken Verben im Frühneuhochdeutschen: Untersucht an Texten des 14.–18. Jahrhunderts*. Bonn: Rheinische Friedrich-Wilhelms-Universität. (Diss.).
- Sonderegger, Stefan. 1979. *Grundzüge deutscher Sprachgeschichte: Diachronie des Sprachsystems*, Bd. 1: Einführung – Genealogie – Konstanten. Berlin/New York: de Gruyter.
- Sorace, Antonella. 2000. Gradients in auxiliary selection with intransitive verbs. *Language* 76(4). 859–890.
- Stefanowitsch, Anatol. 2009. Bedeutung und Gebrauch in der Konstruktionsgrammatik: Wie kompositionell sind modale Infinitive im Deutschen? *Zeitschrift für Germanistische Linguistik* 37(3). 211. DOI: 10.1515/ZGL.2009.036.

- Stefanowitsch, Anatol. 2011. Keine Grammatik ohne Konstruktionen: Ein logisch-ökonomisches Argument für die Konstruktionsgrammatik. In Stefan Engelberg, Anke Holler & Kristel Proost (Hrsg.), *Sprachliches Wissen zwischen Lexikon und Grammatik* (Jahrbuch des Instituts für Deutsche Sprache 2010), 181–210. Berlin: de Gruyter.
- Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin & Achim Zeileis. 2008. Conditional variable importance for random forests. *BMC Bioinformatics* 9(307). Version 1.3-5. <http://www.biomedcentral.com/1471-2105/9/307> (4 August, 2020).
- Szczepaniak, Renata. 2010. *Während des Flug(e)s/Ausflug(e)s?* German short and long genitiv endings between norm and variation. In Albrecht Plewnia & Alexandra N. Lenz (Hrsg.), *Grammar between norm and variation* (VarioLingua 28), 103–126. Frankfurt am Main/New York: Peter Lang.
- Szczepaniak, Renata. 2011a. Gemeinsame Entwicklungspfade in Spracherwerb und Sprachwandel? Kognitive Grundlagen der onto-historiogenetischen Entwicklung der satzinternen Großschreibung. In Klaus-Michael Köpcke & Arne Ziegler (Hrsg.), *Grammatik – Lehren, Lernen, Verstehen: Zugänge zur Grammatik des Gegenwartsdeutschen* (Reihe Germanistische Linguistik 293), 341–359. Berlin/Boston: de Gruyter.
- Szczepaniak, Renata. 2011b. *Grammatikalisierung im Deutschen: Eine Einführung*. 2. Aufl. (Narr Studienbücher). Tübingen: Narr.
- Tabor, Whitney & Bruno Galantucci. 2000. Ungrammatical influence: Evidence for dynamical language processing. In Lila R. Gleitman & Aravind K. Joshi (Hrsg.), *Proceedings of the 22nd annual meeting of the Cognitive Science Society*, 505–510. Mahwah: Lawrence Erlbaum Associates.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178. DOI: 10.1017/S0954394512000129.
- Taraban, Roman & Bret Roark. 1996. Competition in learning language-based categories. *Applied Psycholinguistics* 17(02). 125–148. DOI: 10.1017/S014271640000761X.
- Taylor, John R. 1990. Schemas, prototypes, and models: In search of the unity of the sign. In Savas L. Tsohatzidis (Hrsg.), *Meanings and prototypes*, 521–535. London: Routledge.
- Taylor, John R. 1995. *Linguistic categorization: Prototypes in linguistic theory*. 2. Aufl. Oxford: Clarendon Press.
- Taylor, John R. 1998. Syntactic constructions as prototype categories. In Michael Tomasello (Hrsg.), *The new psychology of language*, 177–202. Mahwah: Lawrence Erlbaum Associates.

- Taylor, John R. 2015. Prototype effects in grammar. In Ewa Dąbrowska & Dagmar Divjak (Hrsg.), *Handbook of cognitive linguistics* (Handbücher zur Sprach- und Kommunikationswissenschaft 39), 462–579. Berlin/Boston: de Gruyter.
- Teuber, Oliver. 2005. *Analytische Verbformen im Deutschen: Syntax, Semantik, Grammatikalisierung* (Germanistische Linguistik. Monographien 18). Hildesheim/Zürich/New York: Olms.
- Thieroff, Rolf. 2003. Die Bedienung des Automaten durch den Mensch: Deklination der schwachen Maskulina als Zweifelsfall. *Linguistik online* 16(4). 105–117.
- Thornton, Anna M. 2019. Overabundance in morphology. In Mark Aronoff (Hrsg.), *Oxford research encyclopedia of linguistics*. New York: Oxford University Press.
- Tomasello, Michael. 1992. *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Tomasello, Michael. 1998. Introduction: A cognitive-functional perspective on language structure. In Michael Tomasello (Hrsg.), *The new psychology of language*, vii–xxiii. Mahwah: Lawrence Erlbaum Associates.
- Tomasello, Michael. 2009. The usage-based theory of language acquisition. In Edith Laura Bavin (Hrsg.), *The Cambridge handbook of child language* (Cambridge Handbooks in Linguistics), 69–88. Cambridge: Cambridge University Press.
- van Rossum, Guido & Fred L. Drake. 2009. *Python 3: Reference manual*. Version: 3.8. Scotts Valley: CreateSpace. <https://www.python.org/> (4 August, 2020).
- Vendler, Zeno. 1974. *Linguistics in philosophy*. 3. Aufl. Ithaca: Cornell University Press.
- Vennemann, Theo. 1987. Muta cum Liquida: Worttrennung und Syllabierung im Gotischen (Mit einem Anhang zur Worttrennung in der Pariser Handschrift der althochdeutschen Isidor-Übersetzung.) *Zeitschrift für deutsches Altertum und deutsche Literatur* 116(3). 165–204.
- Vieregge, Annika. 2019. Speakers' doubts about prepositional case government in German. In Eleonore Schmitt, Renata Szczepaniak & Annika Vieregge (Hrsg.), *Sprachliche Zweifelsfälle. Definition, Erforschung, Implementierung* (Germanistische Linguistik), 69–96. Hildesheim/Zürich/New York: Georg Olms Verlag.
- Wegener, Heide. 1985. „Er bekommt widersprochen“: Argumente für die Existenz eines Dativpassivs im Deutschen. *Linguistische Berichte* 96. 127–139.
- Wegera, Klaus-Peter & Hans-Joachim Solms. 2000. Morphologie des Frühneuhochdeutschen. In Werner Besch, Anne Betten, Oskar Reichmann & Stefan Sonderegger (Hrsg.), *Sprachgeschichte. Ein Handbuch zur Geschichte der deut-*

- schen Sprache und ihrer Erforschung. Band 2.* 2. Aufl. (Handbücher zur Sprach- und Kommunikationswissenschaft), 1542–1554. Berlin/New York: de Gruyter.
- Wegge, Jürgen. 2001. Motivation, information processing and performance: Effect of goal setting on basic cognitive processes. In Anastasia Efklides, Julius Kuhl & Richard M. Sorrentino (Hrsg.), *Trends and prospects in motivation research*, 269–296. Dordrecht: Kluwer.
- Werner, Otmar. 1989. Sprachökonomie und Natürlichkeit im Bereich der Morphologie. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42. 34–47.
- Weyerts, Helga, Martina Penke, Ulrike Dohrn, Harald Clahsen & Thomas F. Münte. 1997. Brain potentials indicate differences between regular and irregular German plurals. *Neuroreport* 8(4). 957–962. DOI: 10.1097/00001756-199703030-00028.
- Whaley, C. P. 1978. Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior* 17. 143–154.
- Whittlesea, Bruce W. A. 1987. Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13(1). 3–17.
- Wickham, Hadley. 2017. *Tidyverse: Easily install and load the tidyverse*. Version 1.2.1. <https://CRAN.R-project.org/package=tidyverse> (4 August, 2020).
- Wickham, Hadley & Jennifer Bryan. 2019. *Readxl: Read excel files*. Version 1.3.1. <https://CRAN.R-project.org/package=readxl> (4 August, 2020).
- Wickham, Hadley, Jim Hester & Romain Francois. 2018. *Readr: Read rectangular text data*. Version 1.3.1. <https://CRAN.R-project.org/package=readr> (4 August, 2020).
- Winter, Bodo. 2013. *Linear models and linear mixed effects models in R with linguistic applications*. <https://arxiv.org/abs/1308.5499> (4 August, 2020).
- Winter, Bodo. 2020. *Statistics for linguists: An introduction using R*. New York/London: Routledge.
- Wittgenstein, Ludwig. 2004 [1953/1968]. Family resemblances. In Bas Aarts, David Denison, Evelien Keizer & Gergana Popova (Hrsg.), *Fuzzy grammar*, 41–44. Oxford/New York: Oxford University Press.
- Witzel, Naoko, Jeffrey Witzel & Kenneth Forster. 2012. Comparisons of online reading paradigms: Eye tracking, moving-window, and maze. *Journal of Psycholinguistic Research* 41(2). 105–128. DOI: 10.1007/s10936-011-9179-x.
- Wurzel, Wolfgang Ullrich. 1984. *Flexionsmorphologie und Natürlichkeit: Ein Beitrag zur morphologischen Theoriebildung* (Studia grammatica). Berlin: Akademie Verlag.

Literaturverzeichnis

- Zevin, Jason D. & Mark S. Seidenberg. 2002. Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language* 47(1). 1–29.
- Zipf, George Kingsley. 1968 [1935]. *The psycho-biology of language: An introduction to dynamic philology*. Cambridge: MIT.
- Zipf, George Kingsley. 1972 [1949]. *Human behavior and the principle of least effort: An introduction to human ecology*. New York: Hafner.

Personenregister

- Ackermann, Tanja, 32, 131, 212, 280
Ágel, Vilmos, 57, 60, 69, 133, 345
Alegre, Maria, 32
Alexiadou, Artemis, 168
Anderson, John R., 69
Audring, Jenny, 58
Auer, Peter, 22
Augst, Gerhard, 4, 100, 101, 119
- Baayen, R. Harald, 27, 28, 36, 43,
177–180, 230, 231, 243, 342
Bar, Moshe, 72
Barca, Laura, 178
Bärenfänger, Olaf, 50
Barlow, Michael, 2, 9, 17, 29, 30,
71–73, 76
Bartke, Susanne, 32
Barton, Kamil, 225, 230
Bates, Douglas, 225, 230
Bates, Elizabeth, 11
Battig, William F., 56
Behrens, Heike, 26, 27
Beretta, Alan, 32, 34
Bergmann, Rolf, 99, 357
Binanzer, Anja, 137
Bittner, Andreas, 5, 6, 48, 97–99, 101,
108–110, 112, 113, 119, 121,
346
Bittner, Dagmar, 125, 135, 140, 148,
149
Blumenthal-Dramé, Alice, 21, 22, 75,
76
- Blutner, Reinhard, 53
Boas, Hans Christian, 74
Bock, J. Kathryn, 55, 56
Bolker, Ben, 225
Booij, Geert, 58, 73, 75, 81, 93
Bornkessel-Schlesewsky, Ina, 32, 33
Bower, Gordon H., 24
Brainerd, Charles J., 73
Braune, Wilhelm, 111, 117
Bryan, Jennifer, 225
Bücker, Jörg, 70, 71, 77, 78
Bybee, Joan L., 2–4, 6, 9, 12–14,
17–23, 31, 34, 43–45, 47, 49,
51, 61–66, 74–78, 81, 83–87,
90, 91, 93, 102, 103, 107, 109,
115, 118, 126, 128, 129, 163,
339, 341, 346–348
- Caldwell-Harris, Catherine L., 12,
24, 28, 29, 43, 91, 342
Calin-Jageman, Robert, 3
Cappellaro, Chiara, 237
Carroll, Ryan, 101
Carstensen, Kai-Uwe, 34
Clahsen, Harald, 32, 34, 106, 107, 110,
187
Clarke, Erik, 225
Comrie, Bernard, 135
Cumming, Geoff, 3
- Dąbrowska, Ewa, 22
Dahl, David B., 225

Personenregister

- Dammell, Antje, 98, 100, 105, 108,
109, 111, 112, 119, 129–131,
334
- Daugherty, Kim G., 26, 31
- Detges, Ulrich, 20, 22, 23
- Di Meola, Claudio, 57
- Diependaele, Kevin, 178, 179
- Diessel, Holger, 9, 22, 64, 66, 67, 75,
76
- Divjak, Dagmar, 12, 24, 28, 29, 43,
91, 342
- Dowle, Matt, 225
- Dowty, David R., 162
- Drake, Fred L., 225
- Duffy, Susan A., 24
- Dupont, Charles, 225, 231
- Dürscheid, Christa, 237
- Eddington, David, 22, 63–65
- Eisenberg, Peter, 57, 123, 124, 126,
131, 132
- Elio, Renée, 69
- Ellis, Andrew W., 27
- Ellis, Nick C., 9–11, 13, 17–19, 21, 25,
26, 33, 34, 45, 49, 51, 53, 57,
65, 68, 74, 89–92, 106, 345,
347
- Elman, Jeffrey L., 30
- Enger, Hans-Olav, 135
- Epstein, Richard, 136
- Estes, Zachary, 178
- Fabricius-Hansen, Catherine, 114
- Fenk-Oczlon, Gertraud, 20, 21, 50,
64
- Ferreira-Junior, Fernando, 57, 345
- Field, Andy, 36, 241, 244
- Fillmore, Charles, 71
- Fine, Alex, 39, 40
- Fischer, Hanna, 190
- Flexser, Arthur J., 24
- Flick, Johanna, 75, 135, 136
- Forster, Jonathan C., 213
- Forster, Kenneth I., 181, 182, 184–186,
203, 213
- Fowler, Carol A., 21, 27
- Friederici, Angela D., 38, 40
- Frisch, Stefan A., 24, 25
- Fromkin, Victoria, 197
- Galantucci, Bruno, 39, 180
- Gandy, Dave, 225
- Gardner, Michael K., 29
- Gernsbacher, Morton Ann, 27
- Gervain, Judit, 11
- Gillmann, Melitta, 1, 5, 6, 69, 81,
154–159, 161–165, 167–170,
172–174, 200, 201, 257, 259,
261, 274, 330, 342
- Goldberg, Adele E., 4, 9, 11, 14–19,
25, 29, 40–43, 49, 51, 57–68,
71, 73–79, 84, 90–93, 101,
108, 119, 125, 147, 158, 188,
245, 295, 327, 339–343, 345,
348
- Goldinger, Stephen D., 25, 61
- Gordon, Peter, 32
- Gouvea, Ana C., 38, 58
- Grainger, Jonathan, 26
- Haberlandt, Karl, 178
- Hahne, Anja, 38, 40
- Haiman, John, 23
- Hanulikova, Adriana, 40
- Hare, Mary L., 30, 32, 33
- Harnisch, Rüdiger, 101
- Harrell, Frank E. Junior, 225, 231
- Hasher, Lynn, 24

- Heidermanns, Frank, 111, 117
Helfman, Gene S., 52
Hillyard, Steven A., 38
Hinze, Christian, 170, 171
Höhle, Barbara, 197, 213
Hopper, Paul, 2, 3, 21, 135, 136, 159,
160
Hothorn, Torsten, 225, 231
Hotzenköcherle, Rudolf, 128
Housum, Jonathan, 21, 27
Hundt, Markus, 47

Indefrey, Peter, 32

Jackendoff, Ray, 50
Jaeger, Jeri J., 32
Jaeger, T. Florian, 39, 40
Janda, Laura A., 109
Jegerski, Jill, 181, 182, 214
Jespersen, Ole, 23
Joeres, Rolf, 142
Johnson, Keith, 15
Johnson, Matt A., 74
Jones, Lara L., 178
Jurafsky, Daniel, 36, 38

Kaan, Edith, 38
Keller, Frank, 181
Kelly, Michael H., 55, 56
Kemmer, Suzanne, 2, 9, 17, 29, 30,
71–73, 76
Keuleers, Emmanuel, 178, 179
Kirchner, Robert, 15
Kleiber, Georges, 46, 48–53, 55, 58,
64, 92
Köpcke, Klaus-Michael, 1, 7, 32, 69,
70, 80, 83, 84, 86–88, 93, 94,
114–116, 118, 121, 124, 125,
131–136, 138–142, 144–149,
152, 170, 171, 209, 212, 277,
279, 283, 348–350, 379
Krischke, Wolfgang, 122, 126, 127,
132, 144, 145, 148, 149
Krott, Andrea, 32
Kuperman, Victor, 179
Kusová, Jana, 139
Kutas, Marta, 38

Labov, William, 45, 49, 50, 59, 60
Lakoff, George, 45, 47, 50, 52, 53, 55,
58, 74
Langacker, Ronald W., 9, 22, 46, 76,
136, 339
Lebib, Riadh, 32
Leiner, Daniel J., 369
Leiss, Elisabeth, 163
Lerche, Anne, 13, 109, 111
Levshina, Natalia, 230, 241, 243
Lewis, Michael B., 27
Lexer, Matthias, 145
Lüdecke, Daniel, 225

Makin, Valerie S., 61, 63, 64, 66, 67,
92, 345
Mangasser-Wahl, Martina, 46
Marchman, Virginia, 31
Mater, Erich, 115
McClelland, James L., 4, 25, 30, 31,
72, 100
McDonald, Scott A., 27, 28, 37, 342
McDonough, Kim, 180, 181, 210
McKinney, Wes, 225
Mervis, Carolyn B., 45, 52, 56, 92
Mitchel, Don C., 181–183
Mitchel, Tom M., 36
Moder, Carol Lynn, 45, 77, 78, 81,
84–87, 93, 115, 118, 348
Montague, William E., 56

Personenregister

- Morrison, Catriona M., 27
- Nakagawa, Shinichi, 241
- Nash, John C., 225
- Nesset, Tore, 73, 80, 109, 135
- Norris, Dennis, 34–38, 343
- Nowak, Jessica, 1, 5, 17, 33, 44, 48,
97, 100–108, 110, 112–114,
116–121, 192, 193, 237, 341,
346
- Nübling, Damaris, 17, 22, 44, 98,
100–104, 107, 109, 110, 112,
122, 123, 126, 129–133, 135,
138, 142, 144, 151, 334, 341,
357
- Osterhout, Lee, 38
- Paul, Hermann, 145
- Pavlov, Vladimir, 124, 125
- Pedersen, Thomas Lin, 225
- Peirce, Jonathan, 196, 205
- Penke, Martina, 32
- Perek, Florent, 42, 43
- Pezzulo, Giovanni, 178
- Pfänder, Stefan, 26, 27
- Pfeifer, Wolfgang, 141
- Pinker, Steven, 31
- Plunkett, Kim, 31
- Poitou, Jacques, 56, 58–60, 83, 132,
149, 150
- Pollack, Irwin, 24
- Primus, Beatrice, 136
- Prince, Alan, 31
- Randall, Janet, 167
- Rayner, Keith, 24, 203
- Roark, Bret, 53
- Robinson, David, 225
- Roehm, Dietmar, 1, 38, 165, 166, 173,
174
- Rogers, Timothy T., 25
- Ronneberger-Sibold, Elke, 146, 222
- Rosch, Eleanor, 45, 50, 52, 53, 55, 56,
58, 59, 92
- Ross, Brian H., 61, 63, 64, 66, 67, 92,
345
- Rothweiler, Monika, 34
- Rubenstein, Herbert, 24, 178, 341
- Rubin, David C., 24
- Rudis, Bob, 225
- Rumelhart, David E., 4, 30, 31, 71, 72,
78, 93, 100, 348
- Ruoff, Arno, 101
- Saffran, Jenny R., 11, 28
- Sandberg, Bengt, 142
- Schäfer, Lea, 105
- Schäfer, Roland, 1, 124, 125, 127, 132,
139, 146, 151
- Scherr, Elisabeth, 206
- Schlesewsky, Matthias, 32, 33
- Schmid, Hans-Jörg, 46, 48, 50, 51, 56,
59, 64
- Schmidt, Richard, 33, 34, 106
- Schmitt, Eleonore, 57, 79, 80, 124,
147, 148, 153, 207, 277,
279–281, 285, 306, 333, 336
- Schmuck, Mirjam, 100
- Schneider, Ulrike, 17, 21, 22, 44, 90,
341
- Seidenberg, Mark S., 26, 27, 31, 34
- Shannon, Thomas F., 154, 158, 160,
161
- Shapiro, Bernard J., 24
- Sherrill-Mix, Scott, 225
- Shillcock, Richard C., 27, 28, 37, 342
- Silverstein, Michael, 135

- Simpson, Carol, 45
Singmann, Henrik, 225, 230
Slobin, Dan I., 77
Smolka, Eva, 33
Solms, Hans-Joachim, 105, 109, 128
Sonderegger, Stefan, 128
Sorace, Antonella, 161, 165
Srinivasan, Arun, 225
Stefanowitsch, Anatol, 74
Strobl, Carolin, 225, 231
Swaab, Tamara Yvonne, 38
Szczepaniak, Renata, 97, 128, 135,
136, 194, 291, 296
- Tabor, Whitney, 39, 180
Tagliamonte, Sali A., 230, 231, 243
Taraban, Roman, 53
Taylor, John R., 45–51, 53, 55, 58, 64,
75, 78, 92, 345
Teuber, Oliver, 161–163
Thieroff, Rolf, 124, 126, 129, 149, 153
Thompson, Sandra A., 4, 9, 12–14,
17, 19–22, 43, 44, 90, 91, 107,
135, 136, 159, 160, 339, 341
Thornton, Anna M., 237
Tomasello, Michael, 9, 58
Trofimovich, Pavel, 180, 181, 210
- van Rossum, Guido, 225
Varadhan, Ravi, 225
Vendler, Zeno, 162
Vennemann, Theo, 291
Vieregge, Annika, 57
Voigt, Gerhard, 57
- Waltereit, Richard, 20, 22, 23
Wecker, Verena, 32, 80, 83
Wegener, Heide, 14
Wegera, Klaus-Peter, 128
- Wegge, Jürgen, 246, 328
Werker, Janet F., 11
Werner, Otmar, 17, 107, 341
Weyerts, Helga, 32
Whaley, C. P., 24, 178
Whittlesea, Bruce W. A., 61
Wickham, Hadley, 225
Winter, Bodo, 226–230, 238–241,
252
Wittgenstein, Ludwig, 49
Witzel, Naoko, 182, 183, 185, 186
Wurzel, Wolfgang Ullrich, 83, 126,
129, 130, 133
- Zacks, Rose T., 24
Zevin, Jason D., 27
Zimmer, Christian, 32, 131, 212, 280
Zipf, George Kingsley, 10, 21, 89
Zubin, David A., 141

Frequenz. Prototyp. Schema.

Die Arbeit entwickelt ein gebrauchsbasiertes Modell zur Entstehung grammatischer Varianten. Dieses wird auf drei Variationsphänomene angewandt: Variation in der Konjugation (*geglimmt/geglossen*), Variation in der Deklination (*des Bären/Bärs*) und Variation in der Selektion zwischen *haben* und *sein* im Perfekt (*ich bin/habe Auto gefahren*). Zudem wird das Modell psycholinguistisch überprüft. Das Modell greift auf den gebrauchsbasierten Ansatz der Kognitionslinguistik zurück und erarbeitet Frequenz, Prototyp und Schema als grundlegende Einflussfaktoren darauf, wie wahrscheinlich Variation und Stabilität in einem Sprachsystem sind: Bei allen Variationsphänomenen sind neben der Variation auch stabile Verwendungen zu beobachten (*geflogen/*geflogen, des Matrosen/*des Matroses, ich bin gegangen/*ich habe gegangen*).

Im theoretischen Teil der Arbeit werden Frequenz, Prototyp und Schema als kognitive Einflussfaktoren auf Variation und Stabilität modelliert und anschließend ihr Einfluss auf die drei Variationsphänomene theoretisch beleuchtet. Im empirischen Teil der Arbeit wird der Einfluss der Faktoren Frequenz, Prototyp und Schema anhand von Reaktionszeitmessungen überprüft.

Das in der Arbeit entwickelte Modell fasst Variation und Stabilität von Sprache probabilistisch und prognostiziert auf diese Weise Variation. Der Rückgriff auf Reaktionszeiten erlaubt es, in der Sprachverarbeitung Variationspotential zu erkennen, das noch nicht im Sprachgebrauch sichtbar ist. Die Arbeit verdeutlicht damit den zentralen Stellenwert, den Variation in der Sprache einnimmt, erweitert mit der Verbindung aus Kognitions- und Psycholinguistik bestehende Forschung und ermöglicht einen systematischen, empirisch überprüfbaren Zugang zu Variation.