

Sandra Hansen & Thilo Weber

Wenn sie den richtigen Kandidat wählen

Zur Flexion der schwachen Maskulina im heutigen gesprochenen (und geschriebenen) Deutsch

Zusammenfassung: Die traditionell als „schwach“ bezeichneten Maskulina des Deutschen neigen bisweilen dazu, sich dem dominanten, „starken“ Flexionsmuster anzuschließen, insbesondere dazu, im Dativ/Akkusativ Singular endungslos aufzutreten (*dem/den Kandidat* statt *dem/den Kandidaten*). Formen dieser Art bilden einen klassischen Zweifelsfall. In unserem Beitrag untersuchen wir dieses Phänomen im spontanen, mündlichen Sprachgebrauch im Vergleich mit der Schriftsprache. Dabei gehen wir auch auf die Frage ein, wie sich mündliche und schriftliche Korpusdaten sinnvoll vergleichen lassen. Unsere Auswertung ergibt, dass endungslose Formen in den mündlichen Daten zwar erwartungsgemäß häufiger sind als in den schriftlichen, dass aber auch mündlich die endungshaltigen Varianten noch deutlich überwiegen. Mit Hilfe logistischer Regressionsanalyse untersuchen wir den Einfluss der „Medialität“ (Mündlichkeit vs. Schriftlichkeit) im Zusammenspiel mit bereits bekannten innersprachlichen (u.a. semantischen und phonotaktischen) Steuerungsfaktoren. Der Befund, dass endungslose Formen mündlich weiter verbreitet sind als schriftlich, lässt sich als synchroner Reflex einer allgemeineren diachronen Tendenz des Deutschen zum Abbau substantivischer Kasusflexion auffassen.

Schlüsselwörter: Substantivflexion, schwache Maskulina, logistische Regressionsanalyse, Standardvariation, grammatische Variation, Korpuslinguistik

1 Einleitung

In vielen Regionen des deutschen Sprachraums erleben standardnahe Sprachformen seit Jahrzehnten auch in der Mündlichkeit eine fortschreitende Domänenausweitung. Man kann darin das Spiegelbild des Rückgangs stärker regional gebundener Sprachformen sehen, insbesondere der Basisdialekte (vgl. etwa Adler et al. zum Niederdeutschen in Norddeutschland). Mit Weiß (2005) und Pröll (2021) befindet sich das Standarddeutsche gegenwärtig in einem Prozess, den Pröll als „Nativierung“ bezeichnet: Nach Jahrhunderten ohne Muttersprachler:innen wird das ursprünglich rein schriftsprachliche Standarddeutsche im 20. und 21. Jahr-

hundert immer öfter natürlich erworben und damit zur gesprochenen Erstsprache („Muttersprache“) eines wachsenden Teils der deutschsprachigen Bevölkerung – und das mit Auswirkungen auf seine grammatische Struktur. Aus diesen Entwicklungen folgt u.a., dass die Untersuchung der Grammatik des Standard- bzw. standardnahen Deutschen nicht ausschließlich auf schriftsprachlicher Grundlage erfolgen sollte, sondern dass zunehmend auch mündliche Daten berücksichtigt werden sollten.

Vor diesem Hintergrund befassen wir uns im vorliegenden Beitrag mit einem bekannten Variationsphänomen aus dem Bereich des Nominalsystems: der Deklination der sog. schwachen Maskulina. Traditionell werden Substantive dieser Flexionsklasse außer im Nominativ Singular mit dem Suffix *-(e)n* markiert (1a). Bekanntermaßen treten sie im Akkusativ und Dativ Singular bisweilen aber auch endungslos auf (1b); häufig wird darin eine Angleichung an das dominante, sog. starke Flexionsmuster gesehen („Flexionsklassenwechsel“, Dudengrammatik 2016: 213). Die Beispiele entstammen dem Forschungs- und Lehrkorpus *Gesprochenes Deutsch* (FOLK) des IDS (Deppermann & Hartung 2012; Schmidt 2016; Kaiser 2018), das die zentrale Datengrundlage des vorliegenden Beitrags bildet (vgl. Abschnitt 4.1).

- (1) a. wenn man en en toten **Menschen** sieht
(FOLK_E_00339_SE_01_T_02)
- b. und Geruch ist für n **Mensch** (0.22) instinktiv extrem wichtig
(FOLK_E_00426_SE01_T_01)

Bei der Variation zwischen Fällen wie (1a) und (1b) handelt es sich um einen klassischen Zweifelsfall: Dem *Zweifelsfälle-Wörterbuch* (2021: 907) zufolge ist das Weglassen der Endung „auch im geschriebenen Standarddeutsch so weit verbreitet, dass es nicht einfach als inkorrekt bezeichnet werden kann.“ Anders sieht dies jedoch z.B. der bekannte Sprachkritiker Bastian Sick (2005: 65): Ihm zufolge ist „die Unterlassung der Flexion [...] umgangssprachlich weit verbreitet, standardsprachlich jedoch gilt sie als falsch.“ Die schwachen Maskulina waren schon mehrfach Gegenstand z.T. umfangreicher korpuslinguistischer Untersuchungen (siehe etwa Kusová 2013, 2014; Paulfranz 2013; Schäfer 2019; Weber & Hansen 2024). Allerdings beschränken diese sich auf die Schriftsprache. Wie die oben angeführten Zitate aus dem *Zweifelsfälle-Wörterbuch* und Sick (2005) zeigen, wird in endungslosen Akkusativen und Dativen aber insbesondere ein Merkmal des gesprochenen Deutsch gesehen. Im vorliegenden Beitrag untersuchen wir auf empirischer, korpuslinguistischer Grundlage, wie verbreitet diese Formen dort tatsächlich sind und wie sie entlang inner- und außersprachlicher Variationsdimensionen verteilt sind. Wir schlagen vor, dass es gewinnbringend ist, die Flexion

der schwachen Maskulina in der Mündlichkeit im Kontext der gegenwärtigen Nativierung des Standarddeutschen zu sehen. Zum einen lassen sich aus diesem Ansatz konkrete Hypothesen ableiten. Zum anderen trägt er zu einer Erklärung für die zu erwartenden Unterschiede zwischen Mündlichkeit und Schriftlichkeit bei.

Der Beitrag gliedert sich wie folgt: Im nachfolgenden Abschnitt 2 geben wir die nötigen Hintergrundinformationen zur Flexion der schwachen und starken Maskulina. In Abschnitt 3 stellen wir den Nativierungsansatz gemäß Weiß (2005) und Pröll (2021) vor. In den Abschnitten 4 und 5 berichten wir die Durchführung und Ergebnisse unserer Korpusuntersuchungen. Wir konzentrieren uns zunächst auf das mündliche Korpus FOLK (Abschnitt 4), bevor wir anschließend schriftliche Korpora zum Vergleich heranziehen (Abschnitt 5). Dabei möchten wir die schwachen Maskulina auch als Anlass dazu nehmen, ein paar allgemeinere methodische Überlegungen dazu anzustellen, wie sich mündliche und schriftliche Korpusdaten in Bezug auf grammatische Variationsphänomene sinnvoll miteinander vergleichen lassen. In Abschnitt 6 fassen wir unsere Befunde zusammen und diskutieren ihre Konsequenzen für die traditionelle Unterscheidung zwischen einer Klasse „starker“ und einer Klasse „schwacher“ Maskulina.

2 Starke vs. schwache Maskulina

Die meisten Maskulina des Deutschen folgen der sog. starken Flexion (Dudengrammatik 2016: 194–220), vgl. Tab. 1. Bei diesem Muster tritt im Singular im Genitiv im prototypischen Fall das Suffix *-(e)s*¹ auf; Akkusativ und Dativ bleiben unflektiert, abgesehen vom mehr oder weniger relikthaft erhaltenen Dativ-*e* (siehe z.B. Konopka 2012). Als Pluralmarker begegnen neben *-e* ohne Umlaut (wie beim Beispieldsubstantiv *Tag* in Tab. 1) auch *-e* mit Umlaut (*Gäste*), endungslose Formen mit und ohne Umlaut (*Nägel*, *Brunnen*), *-s* (*Zoos*) und *-er* mit Umlaut (sofern der Stammvokal denn umlautfähig ist) (*Wälder*). Endet der Plural auf *-e*, *-el* oder *-er*, tritt im Dativ *-n* dazu. Der schwachen Flexion (vgl. Tab. 2), bei der außer im Nominativ Singular die Endung *-(e)n* auftritt, folgt hingegen nur eine kleine Minderheit der maskulinen Substantive.

¹ Siehe Bubenhofer, Hansen-Morath & Konopka (2014) zur Gesamtheit der Varianten der starken Genitivmarkierung und zu deren Verteilung.

Tab. 1: Starke Maskulinflexion.

| | Sg. | Pl. |
|-------------|----------------|--------------|
| Nom. | <i>Tag</i> | <i>Tage</i> |
| Akk. | <i>Tag</i> | <i>Tage</i> |
| Dat. | <i>Tag(e)</i> | <i>Tagen</i> |
| Gen. | <i>Tag(e)s</i> | <i>Tage</i> |

Tab. 2: Schwache Maskulinflexion.

| | Sg. | Pl. |
|-------------|-----------------|-----------------|
| Nom. | <i>Mensch</i> | <i>Menschen</i> |
| Akk. | <i>Menschen</i> | <i>Menschen</i> |
| Dat. | <i>Menschen</i> | <i>Menschen</i> |
| Gen. | <i>Menschen</i> | <i>Menschen</i> |

Wie Köpcke (1995) zeigt, haben viele schwache Maskulina prototypische phono-taktische und semantische Merkmale. So enden z.B. viele von ihnen auf Schwa und bezeichnen Menschen oder höhere Lebewesen (z.B. *Matrose*, *Kollege*, *Hase*). Schäfer (2019) zeigt in einer Korpusstudie anhand des DECOLW-Webkorpus (Schäfer & Bildhauer 2012), dass diese Eigenschaften einen Einfluss darauf haben, wie die Substantive im Sprachgebrauch tatsächlich flektiert werden. So zeigt er, dass insbesondere Substantive mit wenig prototypischen Eigenschaften (z.B. Einsilber wie *Mensch*, vgl. (1b), oder *Bär*) dazu neigen, dem dominanten, starken Muster angeglichen zu werden. Betroffen sind dabei v.a. der Akkusativ (z.B. *den Mensch*, vgl. (1b), oder *den Bär*) und Dativ (*dem Bär*), wo die Angleichung zu Endungslosigkeit führt. Die Ersetzung von *-(e)n* durch *-(e)s* im Genitiv (z.B. *des Bärs*) ist weniger verbreitet (vgl. Abschnitt 4.2). Um einen einheitlichen Begriff für alle Formen zu haben, die mit dem starken Paradigma (vgl. Tab. 1) vereinbar sind, bezeichnen wir nicht nur Genitivformen auf *-(e)s*, sondern auch endungslose Akkusativ- und Dativformen als „starke“ Formen.

Weber & Hansen (2024) bestätigen die Befunde Schäfers (2019) und zeigen darüber hinaus u.a., dass auch Registerunterschiede innerhalb der Schriftlichkeit eine Rolle spielen. So neigen Texte aus dem Webforen-Teil des DECOLW-Webkorpus, die häufiger spontan und unter geringerem Normdruck verfasst werden, in stärkerem Maße zu starken Formen als Texte aus dem pressesprachlichen DeReKo (Kupietz et al. 2018). Insgesamt spielen starke Formen jedoch in

beiden schriftlichen Korpora in quantitativer Hinsicht nur eine geringe Rolle. Schäfer (2019: 401) kommt für das DECOL-Webkorpus insgesamt auf einen Anteil von 2,2% starker Formen. Weber & Hansen (2024) ermitteln für DeReKo einen Anteil von ca. 1,5% und für den Webforen-Teil des DECOL-Webkorpus einen Anteil von gut 5%.

Im vorliegenden Beitrag untersuchen wir die Verwendung der schwachen Maskulina im gesprochenen Deutsch. Dabei interessiert uns, wie weit starke Formen dort verbreitet sind und wie sie entlang inner- und außersprachlicher Variationsdimensionen verteilt sind. Die Untersuchung erfolgt im Vergleich mit der Schriftsprache (Schäfer 2019; Weber & Hansen 2024).

3 Zur Nativierung des standarddeutschen Kasussystems

Unsere Untersuchung erfolgt unter der Annahme, dass das Standarddeutsche gegenwärtig einen Prozess der Nativierung durchläuft (Weiß 2005; Pröll 2021).

Weiß (2005) gliedert die Verankerung des Standarddeutschen in vier „Lebensalter“. Die Herausbildung der Standardvarietät durch gegenseitige Anpassung verschiedener frühneuzeitlicher Varietäten vollzieht sich zunächst ausschließlich als Schriftsprache (Stufe 1). In einem nächsten Schritt, beobachtbar in der Literatursprache ab etwa 1760, wird diese Schriftsprache an die gesprochene Sprache angenähert, dadurch wird sie **sprechbar** (Stufe 2). Als nächstes, einsetzend wohl im 19. Jahrhundert, wird die Standardsprache tatsächlich gesprochen, dabei jedoch weiterhin lediglich sekundär erlernt (Stufe 3). Seit der Mitte des 20. Jahrhunderts schließlich wird die Standardvarietät immer häufiger muttersprachlich erworben (Stufe 4) – mit starken regionalen Unterschieden: So haben in Norddeutschland standardnähere Sprachformen die niederdeutschen Basisdialekte inzwischen bereits weitgehend ersetzt, während in Süddeutschland und Österreich ein Dialekt-Standard-Kontinuum vorliegt. In der Schweiz schließlich herrscht weiterhin mediale Diglossie (keine Übernahme der Standardvarietät in die mündliche Kommunikation). Pröll (2021) liefert silbenphonologische Evidenz für die gegenwärtige Nativierung des Standarddeutschen. So weist er am Beispiel von Bayerisch-Schwaben eine Expansion des standarddeutschen Systems schwererer (zweimoriger) betonter Silben nach. Suprasegmentalia seien in Bezug auf die Nativierungsfrage insofern besonders aufschlussreich, als sie „einen unverfälschten Blick auf das Erstsprachsystem“ böten, da sie „notorisch schwer willentlich“

beeinflussbar und zugleich für Sprecher:innen in der Eigenwahrnehmung wenig salient“ seien (Pröll 2021: 324).

Die Nativierung des Standarddeutschen bleibt nicht ohne Folgen für seine Struktur. So kommt es zum einen zu einem Abbau von „artifiziellen“ Eigenschaften (Weiß 2005: 301), die wohl mit dem ursprünglich fehlenden muttersprachlichen Erwerb der Standardvarietät in Verbindung zu bringen seien. Als Beispiel für ein solches Merkmal aus dem Bereich der Syntax führt Weiß (2005: 296) affine Konstruktionen an (Breitbarth 2005). Umgekehrt werden „diejenigen Strukturen notwendig bzw. naheliegend [...], die ein L1-taugliches Kommunikationssystem typischerweise auszeichnen“ (Pröll 2021: 309). Insbesondere kann sich die Standardvarietät im Zuge ihrer Nativierung nun flexionsmorphologischen Entwicklungstendenzen anschließen, wie sie auch für andere natürlich erworbene Varietäten des Deutschen (bzw. Germanischen) wie etwa die Dialekte typisch sind (vgl. Pröll 2025). Dies im Gegensatz zu den vorangegangenen Jahrhunderten, in denen entsprechender Wandel im Standarddeutschen als einer Varietät ohne Muttersprachler:innen aber mit streng kodifizierter Schriftform ausgesetzt bzw. zumindest verlangsamt war.

Pröll (2025) zeigt dies anhand einer langfristigen Entwicklungstendenz des (hoch-)deutschen² Kasussystems, nämlich des Abbaus des formalen Kontrasts zwischen Nominativ und Akkusativ (vgl. Seiler 2022: 42–43). Anders als noch im Althochdeutschen ist dieser Kontrast im Neuhochdeutschen (jenseits von Pronomen) nur noch im Maskulinum erhalten. In alemannischen Dialekten ist er (wieder: jenseits von Pronomen) konsequent zu Ende geführt (vgl. Seiler 2022: 56–58). Pröll (2025) zeigt nun anhand des Indefinitartikels (*ein-*), dass auch das gesprochene Standarddeutsche weiterreichenden Nominativ-Akkusativ-Synkretismus aufweist als die Schriftsprache. So weist er auf Grundlage eines Korpus bundesdeutscher Fernsehnachrichten nach, dass die endungslose Form *ein*, im Maskulinum ursprünglich auf den Nominativ beschränkt, regelmäßig auch im Akkusativ auftritt.

In Bezug auf die schwachen Maskulina sind zwei langfristige Entwicklungstendenzen des deutschen Kasussystems relevant: Zum einen die generelle Tendenz zum Rückgang von Kasusmarkierung am Substantiv selbst: „Während die Substantive in den früheren Sprachstufen ihre Kasusmarkierung prototypisch am Wortkörper selber trugen, ist seit dem Althochdeutschen ein Abbauprozess zu beobachten“ (Dal & Eroms 2014: 6). Zu den Hauptträgern der Kasusflexion sind stattdessen die Substantivbegleiter (Adjektive und insbesondere Determinierer)

² Charakteristisch **Niederdeutsch** ist hingegen der Zusammenfall von Akkusativ und Dativ zu einem Kasus, der oft als „Obliquus“ bezeichnet wird (vgl. z.B. Berg 2013).

geworden. Z.T. folgt der Abbau substantivischer Kasusflexion aus dem Rückgang des Genitivs (s.u.). Denn mit diesem geht genau derjenige Kasus zurück, in dem zumindest die Maskulina und Neutra noch recht konsequent markiert werden. Der Wandel umfasst aber auch den bloßen Abbau von Kasussuffixen: Dal & Eroms (2014: 6) selbst führen als jüngeres Beispiel für diese Entwicklung den Schwund des starken Dativ-*e* an. In der Tilgung des Suffixes *-en* im Akkusativ und Dativ der schwachen Maskulina (1b) lässt sich ein weiterer Ausdruck dieser langfristigen Entwicklungstendenz sehen. In Dialekten wie z.B. dem Pfälzischen ist Post 1992: 121) zufolge das *-en*-Suffix im Akkusativ und Dativ vollständig abgebaut. Die Annahme, dass endungslose Formen im gesprochenen Deutsch weiter verbreitet sind als in der Schriftsprache (vgl. Abschnitt 1), lässt sich somit aus dem Nativierungsansatz ableiten.

Zum anderen ist der viel beachtete Rückgang des Genitivs relevant (vgl. z.B. Fleischer & Schallert 2011: 83–101; Scott 2014; Seiler 2022: 43–46). Während dieser in früheren Sprachstufen u.a. auch systematisch als Objektkasus auftrat, ist er heute weitgehend auf die Attributfunktion beschränkt. Die meisten Dialekte haben ihn vollständig aufgegeben. Scott (2014) geht davon aus, dass entsprechende Deflexionstendenzen im Mittelhochdeutschen durch die Standardisierung zum Teil aufgehalten wurden. Seiler (2022: 58) erklärt den Erhalt des Genitivs in der Standardsprache damit, dass er als stilistischer Marker für Hoch- und Schriftsprachlichkeit refunktionalisiert worden ist (hierzu Szczepaniak 2014; Pickl 2020). Wie sich zeigen wird, spielt mit dem Genitiv genau derjenige Kasus gesprochensprachlich nur eine geringe Rolle, in dem die morphologische Unterscheidung zwischen schwachen und starken Maskulina noch am stabilsten ist (vgl. Abschnitte 4.2 und 5).

4 Die schwachen Maskulina im gesprochenen Deutsch

4.1 Daten und Methoden

Wir untersuchen die Verwendung der schwachen Maskulina in der Mündlichkeit anhand des Forschungs- und Lehrkorpus *Gesprochenes Deutsch* (FOLK), welches am Leibniz-Institut für Deutsche Sprache angesiedelt ist (Deppermann & Hartung 2012; Schmidt 2016; Kaiser 2018). Dabei handelt es sich um „ein kontinuierlich wachsendes Korpus [...], das Gesprächsdaten aus unterschiedlichsten Bereichen des gesellschaftlichen Lebens (Arbeit, Freizeit, Bildung, öffentliches Leben, Dienst-

leistungen usw.) im deutschen Sprachraum beinhaltet.³ Die in FOLK enthaltenen Interaktionen sind u.a. hinsichtlich des Parameters der „Interaktionsdomäne“ stratifiziert. Unterschieden wird dabei zwischen „Privat“, „Institutionell“, „Öffentlich“ und „Sonstiges“.

Methodisch bauen wir auf Vorarbeiten aus Weber & Hansen (2024) auf. Dort wird anhand von DeReKo eine Liste von 1.156 Substantiven erstellt, die sich mindestens einmal schwach flektiert belegen lassen. Dabei wurden auch Wortbildungprodukte (z.B. *Doktorand*) einschließlich Präkonfixbildungen (z.B. *Ex-Polizist*) berücksichtigt; ausgeschlossen wurden aber prototypische Komposita wie z.B. *Militärpolizist*. Wir konzentrieren uns hier auf diejenigen Maskulina dieser Liste, für die die schwache Flexion das kodifizierte Muster bzw. zumindest eine kodifizierte Variante darstellt. Grundlage dafür bildet der entsprechende Eintrag bei Duden online. Ausgeschlossen werden Substantive, bei denen die *n*-Form (wieder: gemäß Duden online) auch im Nominativ Singular vorkommt (z.B. *der Friede(n)*, *der Funke(n)*) (hierzu u.a. Augustin 2012: 77–81, Paulfranz 2013: 90–91). Denn wenn man davon ausgeht, dass der Nominativ Singular immer dem Stamm entspricht, dann sind Akkusativ- und Dativformen auf *-n* bei diesen Substantiven schwierig zu analysieren: Sie sind ambig zwischen einer endungshaltigen Form zu einem *-n*-losen Stamm (*einen/einem/eines Funke-n*) und einer endungslosen Form zu einem *-n*-haltigen Stamm (*einen/einem/eines Funken-Ø*).

Sämtliche Belege für diese Substantive wurden aus FOLK extrahiert. Anschließend wurden manuell diejenigen Belege herausgesucht, in denen das Substantiv auf eine Akkusativ-, Dativ- oder Genitivform des bestimmten oder unbestimmten Artikels folgt (einschließlich Verschmelzungen von Präposition und Definitartikel), optional mit einem Adjektiv dazwischen. (In Fällen, in denen dem Substantiv kein flektierendes Element vorausgeht, gilt die endungslose Form als Regelfall, vgl. z.B. *DAAD-Preis für Student aus China*, Dudengrammatik 2016: 213, unsere Hervorhebung.) Nach diesem Schritt blieben 631 Belegkandidaten. All diese Kandidaten wurden auditiv überprüft und auf ihre Validität hin bewertet. Dabei ergaben sich einige Problemfälle, die mehrheitlich auf die Besonderheiten gesprochener Sprachdaten zurückzuführen sind.

Erstens: FOLK enthält neben überregionalen auch regionale Sprachformen. Dies zeigt sich z.B. in dem Beleg in (2), der von einer Sprecherin aus dem rheinfränkischen Sprachraum stammt.

³ <https://agd.ids-mannheim.de/folk.shtml>, abgerufen am 25.10.2022.

- (2) °h [dass ma do] am betreude Wohne als erschtes mol nochfrägt wenn ma den **Kunde** net ans Telefon kriegt
 (FOLK_E_00340_SE_01_T_03)

In dem Interaktionsausschnitt verwendet die Sprecherin die Form (*den*) *Kunde* – auf den ersten Blick eine *-n*-lose, unflektierte („starke“) Akkusativform. Allerdings zeigt bereits der kurze Transkriptausschnitt, dass die Sprecherin auch in ganz anderen morphologischen Kontexten standardsprachliches *-en* als Schwa realisiert (vgl. am *betreude Wohne*). Diese „n-Apokope“ führt dazu, dass die stark/schwach-Unterscheidung bei Substantiven, die auf Schwa enden, grundsätzlich neutralisiert wird. Wir haben uns entschlossen, bei Sprecher:innen, bei denen uns *n*-Apokopierung aufgefallen ist, Belege für Substantive auf Schwa auszuschließen.

Zweitens: In manchen Fällen sind Belege nicht eindeutig bestimmbar. Ein Beispiel dafür ist (3). Die Sprecherin verwendet zunächst die unflektierte Akkusativform *Mensch*, korrigiert sich dann aber selbst hin zur Form *Menschen*. Dabei erscheint uns aber unklar, ob es sich um eine Korrektur hin zu einer flektierten Singularform oder hin zu einer Pluralform handelt. Wir haben uns entschlossen, unklare Belege wie diesen auszuschließen.

- (3) glaubst du man kann ähm (0.44) ein **Mensch Menschen** mit einem Duft
 glücklich machen
 (FOLK_E_00426_SE_01_T_01)

Ein drittes Problem ergibt sich für das Lemma *Herr* – mit 218 Belegen das mit Abstand häufigste Lemma unter den Belegkandidaten. *Herr* wird in den meisten Fällen von einem Nachnamen gefolgt, der in den nicht-öffentlichen Interaktionen aus forschungsethischen Gründen maskiert wird. Oft erstreckt sich diese Maskierung auch auf das Wort *Herr*, sodass der Beleg auditiv nicht überprüfbar ist. Da damit zu rechnen ist, dass *n*-lose Formen (z.B. *mit Herr X*) gerade in den nicht-öffentlichen Interaktionen auftreten, würde ein Ausschluss ausschließlich der maskierten Belege vermutlich zu einer Verzerrung der Datenlage führen. Wir haben daher entschieden, das Substantiv *Herr* vollständig auszuklammern.

Nach Ausschluss der als nicht valide gewerteten sowie der *Herr*-Belege bleiben 397 Belege. Diese bilden die Grundlage für den Hauptteil der empirischen Untersuchung in den Abschnitten 4 und 5.

4.2 Ergebnisse

In diesem Abschnitt präsentieren wir die Verbreitung starker Formen schwacher Maskulina in FOLK und ihre Verteilung entlang grammatischer und außersprachlicher Variationsdimensionen. Angesichts der geringen Anzahl von nur 397 Belegen müssen wir uns dabei auf deskriptive Statistiken beschränken.

Insgesamt treten die schwachen Maskulina in 15% der Belege in starken Formen wie (1b) auf. Das ist deutlich mehr, als sich anhand von Vergleichsstichproben aus DeReKo und dem Webforen-Teil von DECOW ermitteln lässt (hierzu mehr in Abschnitt 5). Allerdings ist festzuhalten, dass auch im mündlichen Sprachgebrauch, wie er in FOLK dokumentiert ist, schwache Formen wie (1a) insgesamt deutlich überwiegen.

Abb. 1 zeigt die Verteilung der Formen über die verschiedenen Interaktionsdomänen, nach denen FOLK stratifiziert ist (Kaiser 2018: 521–522). In diesem und den folgenden Diagrammen werden die relativen Häufigkeiten in Form von farbigen Balken dargestellt. Die absoluten Häufigkeiten erscheinen zusätzlich als Zahlen in den Balken. „Öffentliche Interaktionen“ sind definiert als „Gespräche, die im Rahmen öffentlich zugänglicher und/oder massenmedial vermittelter Anlässe stattfinden.“ Sie umfassen u.a. die Schlichtungsgespräche zu Stuttgart 21 sowie TV-Debatten. „Private Interaktionen“ sind definiert als „informelle Gespräche mit Familie und/ oder Freunden“. „Institutionelle“ Interaktionen sind definiert als „Gespräche, die im Rahmen institutioneller Räumlichkeiten bzw. Handlungen mit Personen in der Rolle institutioneller bzw. professioneller Vertreter und mit den entsprechenden konstitutiven Aktivitäten stattfinden“; Beispiele sind Interaktionen am Arbeitsplatz oder in Ausbildungsstätten. Die Kategorie „Sonstiges“ umfasst Interviews und experimentelle Interaktionen. Wie Abb. 1 zeigt, treten die starken Formen häufiger in den privaten als in den institutionellen und öffentlichen Interaktionen auf. Wertet man insbesondere die öffentlichen Interaktionen als Indikator dafür, was den „mündlichen Gebrauchsstandard“ (im Sinne von Deppermann, Kleiner & Knöbl 2013) ausmacht, dann deutet sich an, dass das Weglassen der Endung insbesondere in dieser Sprachform eher untypisch ist.

Als nächstes betrachten wir die sprachinterne Verteilung der starken Formen in FOLK. Dabei konzentrieren wir uns auf die Faktoren Kasus, Phonotaktik und Semantik, die auch von Schäfer (2019) in den Blick genommen werden.

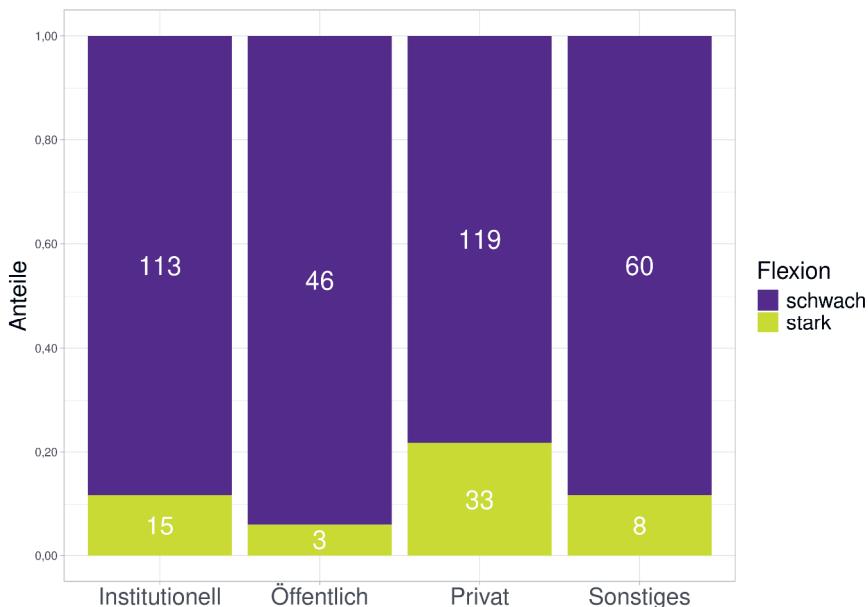


Abb. 1: Starke und schwache Formen in FOLK nach Interaktionsdomäne.

4.2.1 Kasus

In der Literatur herrscht Einigkeit darüber, dass starke Formen primär in Form von endungslosen Akkusativ- und Dativformen auftreten (*dem/den Bär*), und nur in geringerem Maße in Form von Genitivformen auf -(e)s (*des Bärs*) (Thieroff 2003: 109, 114; Schäfer 2019: 405; Dudenwörterbuch 2021: 907). Thieroff erklärt dies strukturalistisch, nämlich damit, dass bei der Angleichung der schwachen (vgl. Tab. 2) an die starke (vgl. Tab. 1) Flexion nicht nur die konkreten Formen innerhalb des starken Paradigmas eine Rolle spielen, sondern dass auch die abstrakte Struktur des starken Paradigmas relevant ist. In diesem Paradigma steht eine markierte Genitivform einer unmarkierten Nicht-Genitivform gegenüber (*des Tag(e)s* vs. *der/dem/den Tag*). Diese Distinktionsstruktur ist bei einem ursprünglich schwachen Maskulinum bereits dann erreicht, wenn lediglich der Akkusativ und der Dativ angeglichen werden (also endungslos verwendet werden), der Genitiv aber weiterhin mit -(e)n (statt mit -(e)s) versehen wird (vgl. *des Bären* vs. *der/dem/den Bär*).

Zusätzlich zu dieser strukturalistischen dürfte aber auch eine soziolinguistische Erklärung eine Rolle spielen (vgl. Weber & Hansen 2024). So ist ja der Genitiv an sich, d.h. der Genitiv als Kasuskategorie, ein stilistischer Marker geworden, der Hoch- oder Schriftsprachlichkeit signalisiert (vgl. Abschnitt 3). Vor diesem Hintergrund ist zwar davon auszugehen, dass er mündlich seltener vorkommt als schriftlich. Gleichzeitig ist aber auch zu erwarten, dass er, **wenn** er verwendet wird, morphologisch in einer Weise realisiert wird, die der schriftsprachlichen Norm entspricht: bei den schwachen Maskulina also schwach.

Wie Abb. 2 zeigt, lassen sich starke Formen in den ausgewerteten FOLK-Daten ausschließlich im Akkusativ (1b) und Dativ (4) belegen. Im Genitiv tritt in unseren Daten immer das schwache Suffix *-en* auf (5). Deutlich wird außerdem die geringe Beleganzahl für den Genitiv insgesamt (mehr dazu in Abschnitt 5). Dies steht im Einklang mit dem oben erwähnten Status des Genitivs als einem primär schriftsprachlichen Kasus.

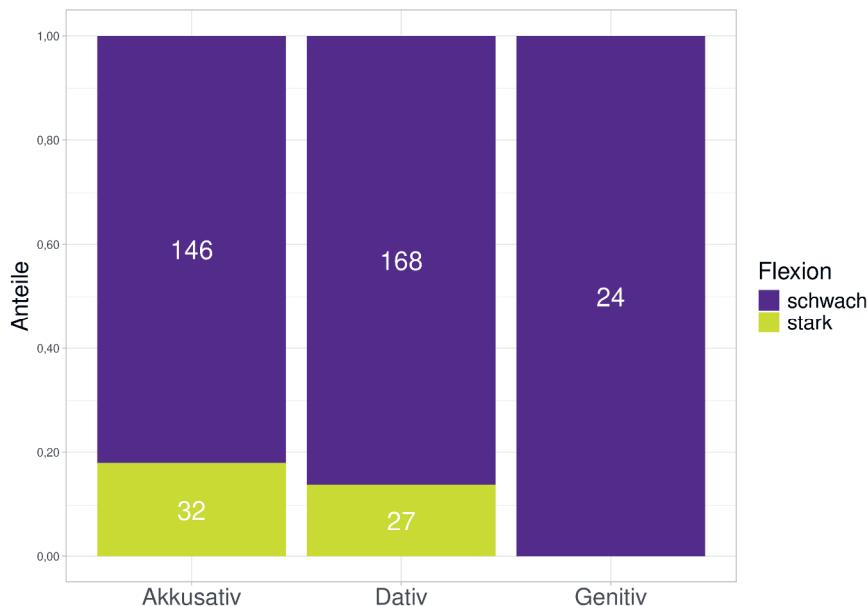


Abb. 2: Starke und schwache Formen in FOLK nach Kasus.

- (4) gibt es eine Verknüpfung zwischen Lustgarten hier(.) vom **Architekt** zum Lustgarten Berlin
(FOLK_E_00418_SE_01_T_01)
- (5) für die Unterbringung des **Soldaten**
(FOLK_E_00315_SE_01_T_01)

4.2.2 Phonotaktik

Aufbauend auf Köpcke (1995), teilt Schäfer (2019: 403) die schwachen Maskulina in vier phonotaktische Typen ein, die unterschiedlich typisch für diese Substantivklasse sind. Eher typisch sind Mehrsilber auf Schwa („polyschwa“, z.B. *Matrose*, *Kollege*) und Mehrsilber mit Endbetonung („polyult“, z.B. *Kandidat*, *Praktikant*). Eher untypisch sind Einsilber („mono“, z.B. *Bär*, *Mensch*) und Mehrsilber mit Nicht-Endbetonung („polynult“, z.B. *Bauer*). Schäfer (2019: 405) zeigt, dass starke Formen insbesondere bei den untypischen phonotaktischen Typen auftreten.

Wie Abb. 3 zeigt, ist auch in den FOLK-Daten der Anteil starker Formen am höchsten bei den untypischen Vertretern: den Mehrsilbern mit Nicht-Endbetonung, wie in Bsp. (6), und den Einsilbern, wie in Bsp. (1b).⁴ Bei den Mehrsilbern mit Endbetonung sind sie weniger häufig, ein Beispiel ist der Beleg in (4); bei den Mehrsilbern auf Schwa treten sie kaum auf; eines der wenigen Beispiele ist der Beleg in (7). Wie die Abbildung auch zeigt, kommen Mehrsilber mit Nicht-Endbetonung insgesamt nur selten vor. Dies steht im Einklang mit der Annahme, dass es sich bei ihnen um untypische Vertreter der schwachen Maskulina handelt.

- (6) meine Großeltern(.) die sind hier beim **Nachbar**
(FOLK_E_00342_SE_01_T_01)
- (7) °hhh und wenn ich beispielsweise den **Kollege** Wegner sehe
(FOLK_E_00390_SE_01_T_06)

⁴ Ein weiterer phonotaktischer Typ liegt bei Substantiven vor, bei denen die Akzentverteilung morphologisch konditioniert ist. Hierzu gehören die Substantive auf *-or*, (z.B. *Autor*, *Motor*), die gemäß Kodex im Singular stark flektieren, sich bisweilen aber der schwachen Flexion anschließen. Unter den Substantiven, für die (auch) die schwache Flexion kodifiziert ist, ist dieses Muster selten. In der Stichprobe findet sich nur ein Beleg dafür (beim Substantiv *Matador*); dieser wurde hier nicht berücksichtigt.

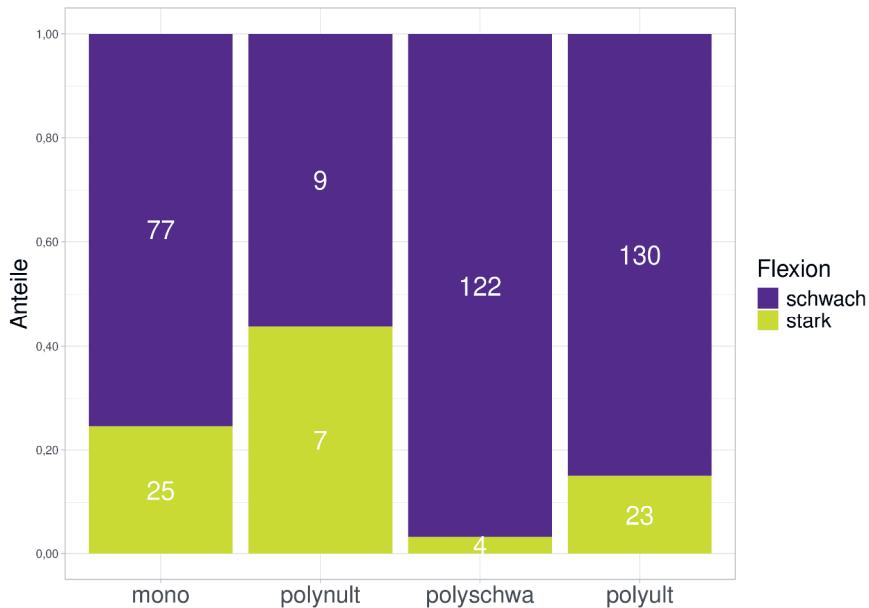


Abb. 3: Starke und schwache Formen in FOLK nach Phonotaktik.

4.2.3 Semantik

Charakteristisch für die schwachen Maskulina ist ferner, dass viele von ihnen Menschen oder höhere Lebewesen bezeichnen. Schäfer (2019: 402–403) unterscheidet zwischen der Referenz auf Menschen (*hum*), auf andere Lebewesen (*ani*) und auf Unbelebte (*ina*), und er zeigt, dass starke Formen bei Substantiven mit nicht-menschlichem Referenten eher auftreten als bei Substantiven mit menschlichem Referenten. Wie Abb. 4 zeigt, ergibt sich für die FOLK-Daten ein ähnliches Bild. Der Anteil starker Formen ist am höchsten bei den Substantiven bzw. Substantivverwendungen mit unbelebtem Referenten wie in Beispiel (8).⁵ Wie die Abbildung auch zeigt, kommen diese insgesamt nur selten vor. Dies steht im Einklang mit der Annahme, dass sie untypisch sind.

5 Einer der Belege ist semantisch unklar und bleibt in Abb. 4 unberücksichtigt.

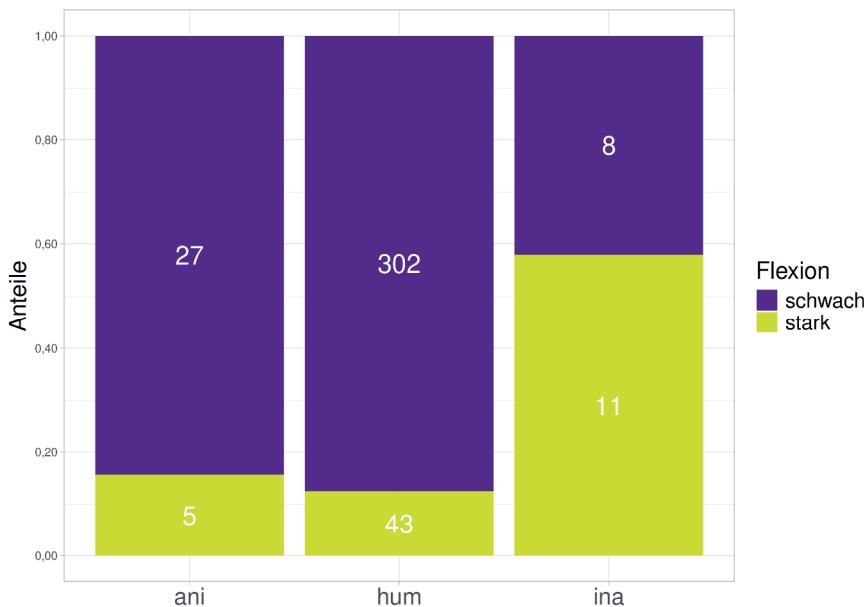


Abb. 4: Starke und schwache Formen in FOLK nach Semantik.

- (8) weil im **Paragraf** steht ja au die Lösung
 (FOLK_E_00166_SE_01_T_01)

4.2.4 Zwischenfazit

Starke Formen stellen auch im mündlichen Sprachgebrauch, wie er in FOLK dokumentiert ist, die deutlich seltener gewählte Variante dar, und zwar insbesondere, aber nicht nur, in den öffentlichen Interaktionen. In Bezug auf die sprachinternen Faktoren deutet sich eine ähnliche Verteilung an, wie sie zuvor auch anhand schriftlicher Daten ermittelt wurde. Bei „starken“ Formen handelt es sich in den ausgewerteten Daten immer um unflektierte Akkusativ- und Dativformen, und diese finden sich v.a. bei Substantiven bzw. Substantivverwendungen, die untypisch für die Klasse der schwachen Maskulina sind: nämlich bei Mehrsilbern mit Nicht-Endbetonung und bei Einsilbern; seltener bei Mehrsilbern mit Endbetonung und kaum bei Mehrsilbern auf Schwa; und sie finden sich eher bei Substantiven bzw. Substantivverwendungen mit unbelebtem Referenten als bei solchen mit belebtem Referenten.

5 Gesprochene vs. geschriebene Sprache

In diesem Abschnitt vergleichen wir die Verwendung der schwachen Maskulina in der Mündlichkeit (FOLK) mit ihrer Verwendung in der Schriftsprache. Herangezogen wird dazu zum einen das presesprachliche DeReKo, zum anderen der – informellere – Webforen-Teil des DECOW-Webkorpus – dieselben Korpora, die auch in Weber & Hansen (2024) untersucht wurden. Sucht man auch in diesen beiden Korpora nur nach denjenigen Substantiven unserer Lemmaliste (vgl. Abschnitt 4.1), die gemäß Duden online schwach flektieren und bei denen (wieder: gemäß Duden online) *-n* nicht auch im Nominativ auftritt bzw. auftreten kann, und klammert man auch dort das Substantiv *Herr* aus, dann treten starke Formen in einer DeReKo-Stichprobe ($n=3.684$) in 2% der Fälle auf, in einer Webforen-Stichprobe ($n=923$) in 8% der Fälle. Wie zu Beginn von Abschnitt 4.2 beschrieben, konnten wir für FOLK einen höheren Anteil von 15% starker Formen ermitteln. Die Ergebnisse für FOLK, DeReKo und DECOW werden vergleichend in Abb. 5 dargestellt. Wie sich zeigt, tritt in der DECOW-Stichprobe marginal auch das Genitivsuffix *-(e)ns* auf („misch“) (*des Pilotens, des Soldatens*), das als Kombination aus schwachem und starkem Suffix angesehen werden kann. Siehe hierzu Krischke (2012), der von „kasusaugmentierenden“ Genitiven spricht.

Die Zahlen stehen im Einklang mit der Annahme, dass starke Formen im mündlichen Sprachgebrauch weiter verbreitet sind als im schriftlichen, insbesondere im presesprachlichen Gebrauch. Im Folgenden möchten wir jedoch diskutieren, wie die in Abb. 5 dargestellten Unterschiede genau zu interpretieren sind: In welchem Maße sind sie darauf zurückzuführen, dass Sprecher:innen bzw. Schreiber:innen in den drei Korpora **selbst unter vergleichbaren grammatischen Bedingungen** unterschiedlich stark zu starken Formen neigen? Und in welchem Maße sind sie eher darauf zurückzuführen, dass der Faktor Korpus (FOLK vs. DECOW vs. DeReKo) mit einigen der grammatischen Faktoren konfundiert ist, die die Wahl zwischen starken und schwachen Formen beeinflussen? Mit anderen Worten: In welchem Maße röhrt der vergleichsweise hohe Anteil starker Formen in FOLK möglicherweise daher, dass dort einige der grammatischen Bedingungen vergleichsweise häufig vorkommen, die die starke Flexion grundsätzlich (d.h. auch in der Schriftlichkeit) begünstigen? Dieses zweite Szenario möchten wir im Folgenden anhand der Kategorien Kasus und Phonotaktik veranschaulichen.

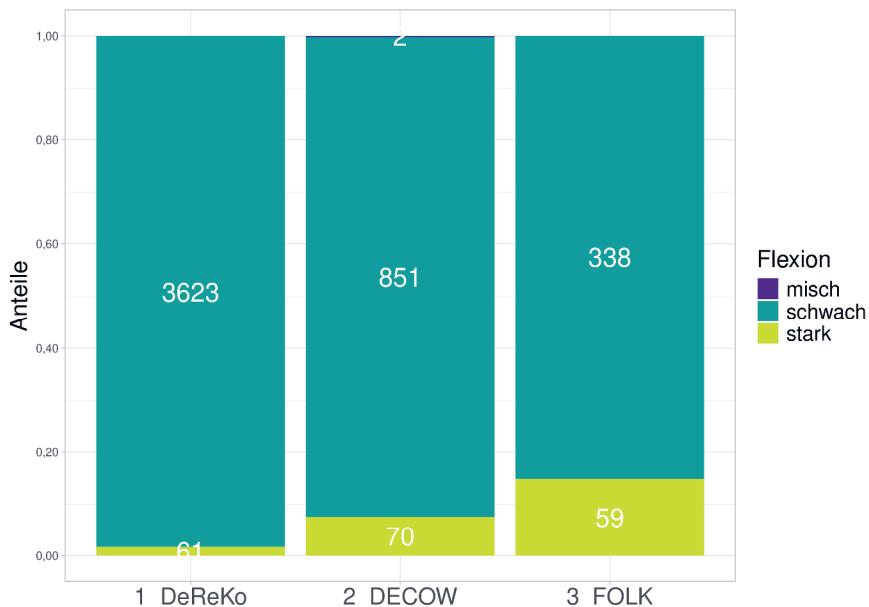


Abb. 5: Starke und schwache Formen in FOLK im Vergleich mit DeReKo und Webforen.

Abb. 6 zeigt die Verteilung der Belege aus den drei Korpora auf die drei nicht-nominativischen Kasus. Wie sich zeigt, gibt es deutliche Unterschiede bei den Genitivanteilen. So entfällt in DeReKo in etwa ein Drittel der Belege auf den Genitiv; in FOLK ist der Anteil deutlich geringer; DECOLW nimmt eine Mittelstellung ein. Wie bereits in Abschnitt 4.2 erwähnt, passt dieser Befund gut zu der Annahme, dass der Genitiv ein Marker für Schriftsprachlichkeit ist (vgl. Szczepaniak 2014; Pickl 2020). Als Alternative zum adnominalen Genitiv (*der Hut des Lehrers*) steht im Deutschen die *von*-Peripherase zur Verfügung (*der Hut von dem Lehrer*, Kopf & Bildhauer 2024), gesprochensprachlich auch der possessive Dativ (*dem Lehrer sein Hut*). Darüber hinaus neigen bekanntermaßen bestimmte (ursprünglich) Genitiv-regierende Präpositionen insbesondere in der Mündlichkeit zur Dativ-rektion, vgl. Beispiel (9).

- (9) wege dem Idiot
 (FOLK_E_00322_SE_01_T_01)



Abb. 6: Belege für schwache Maskulina in DeReKo, DECOW und FOLK nach Kasus.

Wie sich gezeigt hat, ist der Genitiv derjenige Kasus, in dem Sprecher:innen am stärksten am schwachen Muster festhalten, und zwar sowohl in der Mündlichkeit (vgl. Abb. 2) als auch in der Schriftlichkeit (z.B. Schäfer 2019: 405). Das bedeutet: Zu einem bestimmten Anteil dürften die unterschiedlich hohen Anteile starker Formen in den drei Korpora (vgl. Abb. 5) bereits durch die unterschiedlich hohen Genitivanteile zu erklären sein. Der „eigentliche“ grammatische Unterschied zwischen den drei Korpora beträfe dann eher die unterschiedliche Prominenz des Genitivs selbst, u.a. also den Umstand, dass z.B. in Umgebungen wie (9) je nach Korpus in unterschiedlichem Maße Dativ statt Genitiv auftritt. Dies führt dann lediglich indirekt zu unterschiedlich hohen Anteilen starker Formen schwacher Maskulina. Denn eine starke **Dativ**form wie *dem Idiot* in (9) ist eher möglich als eine starke **Genitiv**form wie *des Idiots*.

Abb. 7 zeigt die Verteilung der Belege aus den drei Korpora auf die hier unterschiedenen phonotaktischen Typen. Dabei fällt auf, dass der Typ polyult (z.B. *Präsident*, *Konkurrent*, *Finalist*) in DeReKo einen größeren Anteil ausmacht als in DECOW und FOLK. Der Typ mono (z.B. *Mensch*, Typ (im Sinne von ‚Mann‘)) hingegen hat einen geringeren Anteil. Wie sich gezeigt hat, neigt der Typ polyult in stärkerem Maße zur schwachen Flexion als der Typ mono, und zwar sowohl in

der Mündlichkeit (vgl. Abb. 3) als auch in der Schriftlichkeit (z.B. Schäfer 2019: 405). Das bedeutet: Zu einem bestimmten Anteil dürften die unterschiedlich hohen Anteile starker Formen in den drei Korpora (vgl. Abb. 5) bereits dadurch zu erklären sein, dass es in den DeReKo-Daten häufiger um *Präsidenten*, *Konkurrenten* oder *Finalisten* geht, in FOLK hingegen häufiger um *Menschen* oder *Typen*.

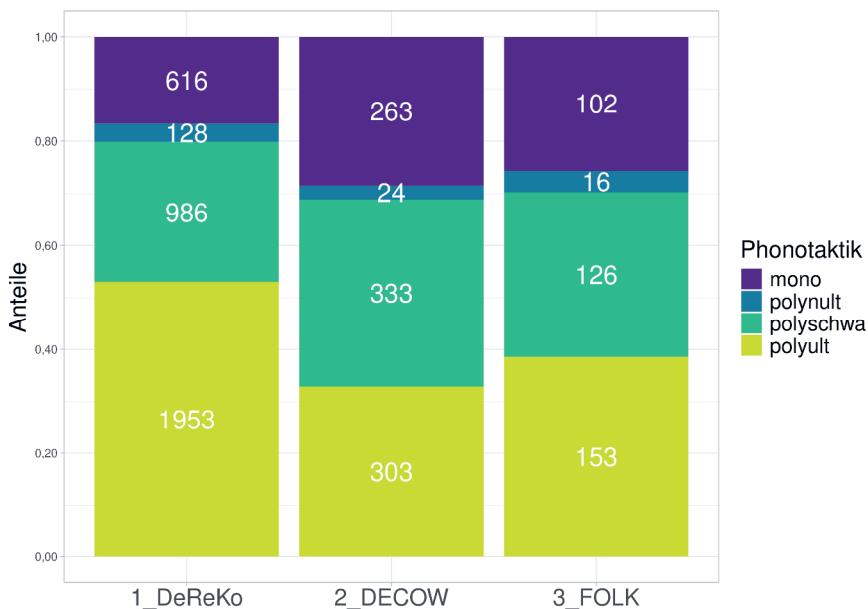


Abb. 7: Belege für schwache Maskulina in DeReKo, DECOLW und FOLK nach phonotaktischem Typ.

Um den Nachweis zu erbringen, dass die drei Korpora auch unter vergleichbaren grammatischen Bedingungen unterschiedlich stark zur starken Flexion neigen, müssen wir für ebendiese grammatischen Faktoren kontrollieren. Wir versuchen, dies umzusetzen, indem wir die Variation zwischen starker und schwacher Flexion in einem multivariaten statistischen Modell untersuchen, in dem wir den Faktor Korpus gemeinsam mit den grammatischen Steuerungsfaktoren berücksichtigen. Wir analysieren die Daten zu diesem Zweck mit einem gemischten logistischen Regressionsmodell (vgl. Bates et al. 2015). Als Datengrundlage verwenden wir die Gesamtheit der validen FOLK-Belege ($n=397$). Aus DeReKo und den DECOLW-Webforen ziehen wir hingegen neue Stichproben ($n=1.308$, $n=1.355$), die wir auf diejenigen Lemmata beschränken, die auch in FOLK mindestens ein-

mal vorkommen. Als Prädiktorvariablen (= unabhängige Variablen) nehmen wir eine Reihe derjenigen internen Faktoren auf, die (potentiell) einen Einfluss auf die Variation zwischen schwacher und starker Flexion haben. Das sind zunächst die hier bereits vorgestellten Faktoren Kasus, Phonotaktik und Semantik. Darüber hinaus ist die logarithmierte Frequenz des Substantivlemmas ein weiterer Prädiktor. So zeigt Schäfer (2019), dass frequente Substantive in geringerem Maße dazu tendieren, an das starke Muster angeglichen zu werden. Er erklärt dies mit dem höheren Grad an *Entrenchment* der schwachen Formen bei höher frequenten Substantiven. Außerdem berücksichtigen wir zwei weitere Prädiktoren, die auch in Weber & Hansen (2024) untersucht wurden: Erstens das (Nicht-)Auftreten einer Apposition: Damit wird eine Beobachtung aus dem Zweifelsfälle-Wörterbuch (2021: 907–908) aufgegriffen, wonach bei schwachen Maskulina, die eine Rangbezeichnung o.ä. bezeichnen, die Endung im Akkusativ und Dativ insbesondere dann weggelassen werde, wenn ein Eigenname folgt (z.B. *für den Präsident(en) Müller*). Zweitens das (Nicht-)Auftreten eines attributiven Adjektivs vor dem Substantiv. Damit greifen wir eine Beobachtung Rohdenburgs (1988) auf, wonach in nordniederdeutschen Dialekten ein vorangehendes Adjektiv auf *-n* die *n*-Form auch bei einem nachfolgenden schwachen Maskulinum begünstigt. Das Korpus, aus dem der jeweilige Beleg stammt, nehmen wir als weitere Prädiktorvariable auf. Das konkrete Substantiv-Lemma berücksichtigen wir als Zufallsfaktor. Dadurch versuchen wir den Umstand zu berücksichtigen, dass die verschiedenen Lemmata quantitativ ungleich über die drei Korpora verteilt sind. In Abhängigkeit der verschiedenen Prädiktoren wurde die Wahrscheinlichkeit modelliert, dass in einem gegebenen Fall starke oder schwache Flexion auftritt (Kriteriumsvariable = abhängige Variable).

Abb. 8 zeigt die Koeffizientenschätzungen und 95% Konfidenzintervalle für das gemischte logistische Regressionsmodell. Weitere Informationen zum Modell befinden sich im Anhang. Als Referenzausprägung (intercept) setzen wir für jeden kategorialen Prädiktor jeweils diejenige Ausprägung an, die gemäß Hypothese am stärksten die schwache Form begünstigt. Beim Prädiktor Kasus ist das die Ausprägung Genitiv, beim Prädiktor Phonotaktik die Ausprägung polyschwa, beim Prädiktor Semantik die Ausprägung hum, beim Prädiktor Apposition die Ausprägung „nein“, beim Prädiktor Adjektiv die Ausprägung „ja“ und beim Prädiktor Korpus die Ausprägung DeReKo. Im Diagramm aufgetragen sind für jeden Prädiktor die jeweils anderen Ausprägungen. Das Diagramm zeigt, ob und ggf. wie stark die Ausprägungen im Vergleich mit der jeweiligen Referenzausprägung die starke Flexion begünstigen (oder ggf. entgegen der Hypothese hemmen).

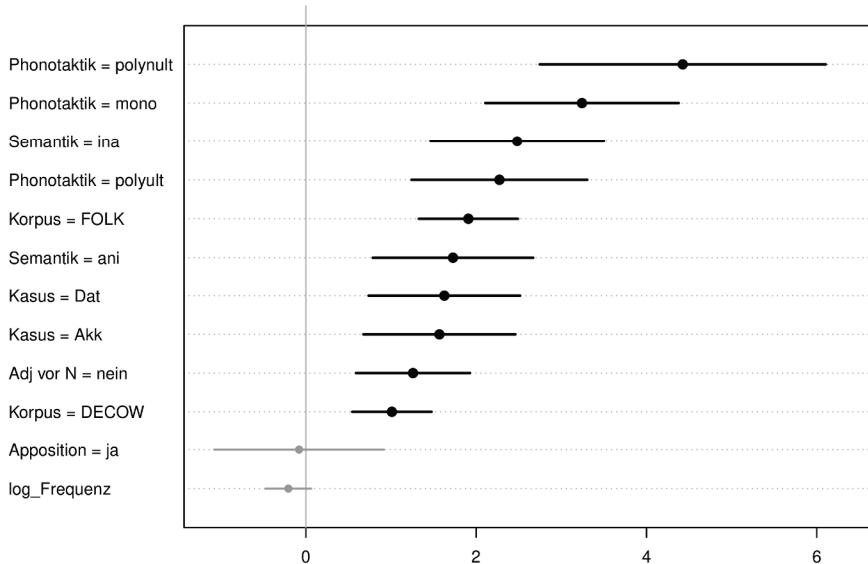


Abb. 8: Grafische Darstellung der Koeffizientenschätzungen. Die horizontalen Linien zeigen 95%-Konfidenzintervalle. Positive Koeffizienten erhöhen die Wahrscheinlichkeit für die starke Flexion. Marginales R^2 (nur feste Effekte) = 0,47, konditionales R^2 (gesamtes Modell) = 0,56.

Das Modell ermittelt signifikante Einflüsse der Prädiktoren Phonotaktik, Semantik, Kasus, Adjektiv und Korpus.

Zunächst zur Phonotaktik: Im Vergleich mit einem Substantiv auf Schwa (polyschwa, Referenzausprägung) neigen mehrsilbige Substantive mit Nicht-Endbetonung (polynult) am stärksten zur starken Flexion, gefolgt von Einsilbern (mono) und Mehrsilbern mit Endbetonung. Für den Prädiktor Semantik ergibt sich, dass Substantivverwendungen mit unbelebtem (ina) und belebtem nicht-menschlichem Referenten (ani) in stärkerem Maße zur starken Flexion tendieren als Substantivverwendungen mit menschlichem Referenten (Referenzausprägung, hum). Für den Prädiktor Kasus kann gezeigt werden, dass Akkusativ und Dativ in stärkerem Maße zur starken Flexion neigen als der Genitiv. Und für den Prädiktor Adjektiv ergibt sich, dass Substantivverwendungen ohne vorangehendes attributives Adjektiv stärker zur starken Flexion neigen als Verwendungen mit Adjektiv. Für das (Nicht-)Auftreten einer Apposition und die logarithmierte Frequenz des Substantivlemmas lässt sich kein Effekt feststellen.

In weiten Teilen decken diese Ergebnisse sich mit denen aus Schäfer (2019) und Weber & Hansen (2024). Abweichungen ergeben sich jedoch für die Prä-

diktoren Frequenz, Adjektiv und Apposition. Anders als Schäfer (2019: 406) finden wir hier keinen Einfluss der Frequenz des Substantivlemmas, was daran liegen könnte, dass die Frequenz unmittelbar mit dem Lemma verbunden ist, und wir – anders als Schäfer – ein gemischtes Modell rechnen, in dem wir das Lemma als Zufallsfaktor berücksichtigen. Weber & Hansen (2024) finden in einer größer angelegten Korpusstudie für den Vergleich zwischen Daten aus dem DeReKo und Daten aus den DECOLW-Webforen in einem gemischten Modell mit Lemma als Zufallsfaktor ebenfalls keinen Effekt für Frequenz, wohl aber in einem zum Vergleich gerechneten einfachen Modell ohne Lemma als Zufallsfaktor: Dort bestätigt sich Schäfers Befund: Je frequenter ein Substantiv, desto stärker hält es am schwachen Muster fest. Für den Prädiktor Adjektiv ermitteln Weber & Hansen (2024) in dem gemischten Modell ebenfalls keinen signifikanten Effekt – wohl aber in dem Modell ohne Lemma als Zufallsfaktor. Dort zeigt sich ein schwacher Effekt in die erwartete Richtung: So treten starke Formen eher dort auf, wo dem Substantiv kein attributives Adjektiv vorangeht. Weber & Hansen stellen die Vermutung auf, dass der Adjektiv-Effekt Lemma-abhängig sein könnte. Dass anders als in Weber & Hansen (2024) im hier ausgewerteten Datensatz selbst das gemischte Modell einen Effekt für den Prädiktor Adjektiv ermittelt, könnte möglicherweise daran liegen, dass hier eine andere Lemma-Auswahl berücksichtigt wurde: nämlich nur diejenigen, die mindestens einmal in FOLK belegt sind. Insgesamt muss die Rolle eines vorangehenden Adjektivs für die Wahl der Flexion aber weiterhin als ungeklärt gelten. Für den Prädiktor Apposition ermitteln Weber & Hansen (2024) einen schwachen Effekt in die erwartete Richtung: So begünstigt das Auftreten einer Apposition unmittelbar nach dem Substantiv die starke Flexion, der Hypothese gemäß dadurch, dass insbesondere in diesem Fall im Akkusativ und Dativ die Endung weggelassen wird. Im hier ausgewerteten Datensatz finden wir keinen entsprechenden Effekt. Möglicherweise hängt auch dies mit der anderen Lemma-Auswahl zusammen.

Für den Prädiktor Korpus, um den es hier schwerpunktmäßig geht, bestätigt sich zunächst der Befund aus Weber & Hansen (2024), dass die DECOLW-Webforen in stärkerem Maße als DeReKo zur starken Flexion neigen. Darüber hinaus zeigt sich in dieser Studie, dass auch FOLK stärker zur starken Flexion neigt als DeReKo. Das bedeutet: Auch unter ansonsten gleichen (grammatischen) Bedingungen ist die Wahrscheinlichkeit für die starke Flexion in FOLK und den Webforen höher als in DeReKo. Eine deskriptive Veranschaulichung kann z.B. anhand des Lexems *Mensch* vorgenommen werden: So enthalten die FOLK-Daten 37 Belege für dieses Lexem im Akkusativ oder Dativ; 10 davon werden endungslos realisiert (27,0%). In den Webforen beträgt der entsprechende Anteil lediglich 6,7% (9 von 134), in DeReKo gar nur 0,4% (1 von 283).

6 Fazit und Ausblick

Ziel dieses Beitrags war eine Untersuchung der Verbreitung und Verteilung starker Formen traditionell schwach flektierender Maskulina im heutigen gesprochenen Deutsch. Die Untersuchung erfolgte vor dem Hintergrund der Annahme, dass das Standarddeutsche sich gegenwärtig in einem Prozess der Nativierung befindet.

Wie sich gezeigt hat, treten starke Formen im mündlichen Korpus FOLK häufiger auf als in den schriftsprachlichen Vergleichskorpora; schwache Formen überwiegen aber auch in den mündlichen Daten noch recht deutlich – und zwar insbesondere, aber nicht nur, in den öffentlichen Interaktionen. Die sprachinterne Verteilung starker und schwacher Formen scheint in FOLK ähnlich auszufallen wie in schriftlichen Korpora: So handelt es sich bei „starken“ Formen ganz überwiegend um endungslose Akkusative und Dative; Genitive auf -(e)s ließen sich in den FOLK-Daten nicht belegen. Am konsequentesten wird noch bei Substantiven auf Schwa an der schwachen Form festgehalten; Formen ohne -n kommen dort kaum vor. Im Rahmen unseres Korpusvergleichs haben wir dafür plädiert, die unterschiedlichen Anteile starker und schwacher Formen im Zusammenspiel mit den grammatischen Steuerungsfaktoren zu betrachten. Mit Hilfe gemischter logistischer Regression haben wir gezeigt, dass FOLK aber auch dann noch stärker zu starken Formen neigt, wenn man für diese grammatischen Steuerungsfaktoren kontrolliert.

Wie bereits in Abschnitt 1 erwähnt, wird die Variation bei den schwachen Maskulina als Ausdruck von „Flexionsklassenwechsel“ aufgefasst, d.h. als Ausdruck einer Angleichung an die starke Flexion (vgl. Dudengrammatik 2016: 213). Möglicherweise ist diese Auffassung gerade für das gesprochene Deutsch ein wenig zu modifizieren. Wie sich gezeigt hat, scheint die Angleichung weitgehend auf Akkusativ und Dativ beschränkt zu sein – und damit auf diejenigen zwei Kasus, in denen sie zu Endungslosigkeit führt. Der einzige Kasus, in dem es zur Übernahme eines overten starken Suffixes käme, nämlich der Genitiv, zeigt kaum Variation. Dort ist der Gegensatz zwischen den traditionell schwach flektierenden und den traditionell stark flektierenden Substantiven also recht stabil. Gleichzeitig ist für den Genitiv jedoch zu beobachten, dass er insgesamt, d.h. als Kasuskategorie, gesprochensprachlich nur eine vergleichsweise geringe Rolle spielt (vgl. Abb. 6). In Bezug auf die Opposition zwischen einer Klasse schwacher Maskulina und einer Klasse starker Maskulina hat diese Tendenz zur Nicht-Verwendung des Genitivs jedoch einen ähnlichen Effekt wie die Angleichung der schwachen an die starke Genitivmarkierung: Sie ermöglicht im Singular ein einheitliches „Paradigma“ für die Substantive beider Klassen. Jedoch besteht dieses lediglich aus einer einzigen – der unflektierten – Form, die in allen drei verbleibenden Kasus (Nom-

nativ, Akkusativ, Dativ) verwendet werden kann (*der/den/dem Mensch/Tag*) (vgl. Weber & Hansen 2024). Da in diesem „Paradigma“ nicht nur das schwache Suffix, sondern auch das starke Genitivsuffix keine Rolle spielt, wäre zu überlegen, ob man anstatt von „Flexionsklassenwechsel“ nicht eher von einer Tendenz zur **Neutralisierung** der Flexionsklassenunterscheidung (im Singular) zu Gunsten der unflektierten Form sprechen sollte.

Das gesprochene Deutsch zeichnet sich im Vergleich mit der Schriftsprache somit durch einen Abbau von Kasusflexion am Substantiv aus. Damit tritt eine langfristige Entwicklungstendenz des deutschen Kasussystems in der Mündlichkeit stärker zu Tage als in der Schriftlichkeit. Das gesprochene Deutsch, wie es im FOLK dokumentiert ist, liegt in dieser Hinsicht im Vergleich mit der Schriftsprache ein wenig näher an (anderen) natürlich erworbenen Varietäten wie z.B. dem Pfälzischen (nach Post 1992: 121), in denen die Kasusflexion am Substantiv vollständig aufgegeben worden ist.

Statistischer Anhang

Die Parameter der statistischen Modelle wurden in R (R Core Team 2023) mit dem lme4-Paket (Bates et al. 2015) geschätzt. Für das gemischte Regressionsmodell ergibt sich folgende Modellspezifikation:

$$P(y_i = \text{stark}) = \text{logit}^{-1}(a_{j[i]} + \beta_0 + \beta_1 \text{Semantik}_i + \beta_2 \text{Korpus}_i + \beta_3 \text{Kasus}_i + \beta_4 \text{Phonotaktik}_i + \beta_5 \text{log.Frequenz}_i + \beta_6 \text{Adj_vor_N}_i + \beta_7 \text{Apposition}_i)$$

with

$$a_j \sim N(0, \sigma^2_a)$$

Die folgende Tabelle zeigt die Koeffizientenschätzungen für das Modell. Über den Einfluss eines Prädiktors auf die Wahrscheinlichkeit lassen sich genaue Aussagen nur machen, wenn die Ausprägungen aller anderen Prädiktoren berücksichtigt werden. Das Vorzeichen und der Betrag der Koeffizientenschätzer können interpretiert werden: Bei Koeffizienten > 0 steigt die vorhergesagte Wahrscheinlichkeit dafür, dass das Substantiv stark flektiert wird, bei negativen Koeffizienten nimmt sie ab.

Ein Indikator für die Modellgüte ist der Anteil der vom Modell erklärten Varianz in den Daten. Wir berichten im Folgenden das im MuMin-Paket (Bartoń 2019) implementierte Pseudo- R^2 von Nakagawa et al. (2017). Dieses ist auch für gemischte Modelle definiert und gibt zum einen die Varianz an, die durch die festen Effekte erklärt wird (marginales $R^2 = R^2_m$), und zum anderen die Varianz,

die durch das gesamte Modell einschließlich der zufälligen Effekte erklärt wird (konditionales $R^2 = R^2_c$).

Ergebnisse Regressionsmodell

Tab. 3: Parameterschätzungen für das gemischte Regressionsmodell.

| Variable | (Ausprägung) | $\hat{\beta}$ | Std. Error | z value | p |
|--------------|--------------|---------------|------------|---------|-------|
| INTERCEPT | | -F,G | H,I | -J,KH | <GGGH |
| KORPUS | DECOW | H,GH | G,NC | Q,NP | <GGGH |
| KORPUS | FOLK | H,KH | G,J | P,Q | <GGGH |
| SEMANTIK | ani | H,FJ | G,CI | J,P | <GGGH |
| SEMANTIK | ina | N,CI | GRN | Q,FI | <GGGH |
| KASUS | Akkusativ | H,RF | G,OP | J,OR | <GGGH |
| KASUS | Dativ | H,PJ | G,OR | J,RK | <GGGH |
| PHONOTAKTIK | mono | J,NC | G,RI | R,P | <GGGH |
| PHONOTAKTIK | polynult | Q,Q | G,IP | R,HF | <GGGH |
| PHONOTAKTIK | polyult | N,NF | GRJ | Q,JN | <GGGH |
| LOG.FREQUENZ | | -GNH | G,HC | -H,R | GHJ |
| APPOSITION | ja | -G,CI | G,RH | -G,HP | GI,F |
| ADJ_VOR_N | nein | H,NP | G,J,C | J,PI | <GGGH |

Marginales R^2 (nur feste Effekte) = G,OF; konditionales R^2 (gesamtes Modell) = GRK.

Literatur

- Adler, Astrid, Christine Ehlers, Reinhard Goltz, Andrea Kleene & Albrecht Plewnia (2016): *Status und Gebrauch des Niederdeutschen 2016. Erste Ergebnisse einer repräsentativen Erhebung*. Mannheim: Institut für Deutsche Sprache.
- Augustin, Hagen (2012): *Autor, Doktor, Friede, Funke*. Problemfälle der Flexion. In Marek Konopka & Roman Schneider (Hrsg.), *Grammatische Stolpersteine digital – Festschrift für Bruno Strecker zum 65. Geburtstag*, 71–82. Mannheim: Institut für Deutsche Sprache.
- Bartoń, Kamil (2019): *MuMIn: Multi-Model Inference*. R package version 1.43.6.
- Bates, Douglas, Martin Mächler, Ben Bolker & Stev Walker (2015): Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67 (1), 1–48.
- Berg, Kristian (2013): *Morphosyntax nominaler Einheiten im Niederdeutschen*. Heidelberg: Winter.
- Breitbarth, Anne (2005): *Live fast, die young – the short life of Early Modern German auxiliary ellipsis*. PhD diss., Tilburg University.

- Bubenhofer, Noah, Marek Konopka & Roman Schneider (2014): *Präliminarien einer Korpusgrammatik*. Tübingen: Narr.
- Dal, Ingerid & Hans-Werner Eroms (2014): *Kurze deutsche Syntax auf historischer Grundlage*. Berlin & Boston: De Gruyter.
- Deppermann, Arnulf & Martin Hartung (2012): Was gehört in ein nationales Gesprächskorpus? Kriterien, Probleme und Prioritäten der Stratifikation des „Forschungs- und Lehrkorpus Gesprochenes Deutsch“ (FOLK) am Institut für Deutsche Sprache (Mannheim). In Ekkehard Felder, Marcus Müller & Friedemann Vogel (Hrsg.), *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen*, 415–450. Berlin: De Gruyter.
- Deppermann, Arnulf, Stefan Kleiner & Ralf Knöbl (2013): „Standard usage“: towards a realistic conception of spoken standard German. In Javier Caro Reina, Peter Auer & Göz Kaufmann (Hrsg.): *Language variation. European perspectives IV. Selected papers from the Sixth International Conference on Language Variation in Europe (ICLaVE 6), Freiburg, June 2011*, 83–116. Amsterdam: Benjamins.
- Dudengrammatik 2016 = Wöllstein, Angelika & Dudenredaktion (Hrsg.) (2016): *Dudenband 4 – Die Grammatik: unentbehrlich für richtiges Deutsch*. 9., vollständig überarbeitete und aktualisierte Auflage. Berlin: Dudenverlag.
- Hennig, Mathilde (2000): *Tempus und Temporalität in geschriebenen und gesprochenen Texten*. Tübingen: Niemeyer.
- Kaiser, Julia (2018): Zur Stratifikation des FOLK-Korpus: Konzeption und Strategien. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 19, 515–552.
- Konopka, Marek (2012): Dem Manne kann geholfen werden — Wann kommt das Dativ-e zum Einsatz? In Marek Konopka & Roman Schneider (Hrsg.), *Grammatische Stolpersteine digital – Festschrift für Bruno Strecker zum 65. Geburtstag*, 115–124. Mannheim: Institut für deutsche Sprache.
- Köpcke, Klaus-Michael (1995): Die Klassifikation der schwachen Maskulina in der deutschen Gegenwartssprache – ein Beispiel für die Leistungsfähigkeit der Prototypentheorie. *Zeitschrift für Sprachwissenschaft* 14 (2), 159–180.
- Kopf, Kristin & Felix Bildhauer (2024): The genitive alternation in German. *Corpus Linguistics and Linguistic Theory*. Published online November 13, 2024.
<https://www.degruyter.com/document/doi/10.1515/cllt-2024-0017/html> (abgerufen am 15.11.2024)
- Krischke, Wolfgang (2012): Des Menschen Genitive. Normabweichende Genitiv-Varianten bei schwachen Maskulina. *Linguistik Online* 53 (3), 55–84.
- Kupietz, Marc, Harald Lüngen, Paweł Kamocki & Andreas Witt (2018): The German Reference Corpus DeReKo: New Developments – New Opportunities. In Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis & Takenobu Tokunaga (Hrsg.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 4353–4360. Miyazaki: European Language Resources Association (ELRA).
<https://www.aclweb.org/anthology/L18-1689.pdf> (abgerufen am 01.10.2020)
- Nakagawa, Shinichi, Paul C.D. Johnson & Holger Schielzeth (2017): The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface* 134 (14), 1–11.
- Paulfranz, Alexandra (2013): *Kasusmarkierungen der Gegenwartssprache in deutschen Lokal- und Regionaltageszeitungen*. Bamberg: University of Bamberg Press.
- Pickl, Simon (2020): Polarization and the emergence of a written marker. A diachronic corpus study of the adnominal genitive in German. *Journal of Germanic Linguistics* 32, 145–182.

- Post, Rudolf (1992): *Pfälzisch. Einführung in eine Sprachlandschaft*. 2. aktualisierte und erweiterte Auflage. Landau: Pfälzische Verlagsanstalt.
- Pröll, Simon (2021): Die Nativierung des Standarddeutschen. Mikrotypologische Evidenz für supra-segmentalen Wandel. *Zeitschrift für Angewandte Linguistik* 75, 305–329.
- Pröll, Simon (2025): Szenarien der Schriftlich-mündlich-Schere im Standard. In Nadine Proske, Thilo Weber, Monika Dannerer & Arnulf Deppermann (Hrsg.), *Gesprochenes Deutsch. Struktur, Variation, Interaktion*, 1–23. Berlin & Boston: De Gruyter.
- R Core Team (2023): *R: A Language and Environment for Statistical Computing*. Wien: R Foundation for Statistical Computing. <https://www.R-project.org/> (abgerufen am 31.10.2023)
- Rohdenburg, Günter (1988): Flexionsangleichung von Substantiven an attributive Adjektive und verwandte Erscheinungen im Nordniedersächsischen. In Heinrich Weber & Ryszard Zuber (Hrsg.), *Linguistik Parissette. Akten des 22. linguistischen Kolloquiums, Paris 1987*, 277–288. Tübingen: Niemeyer.
- Schäfer, Roland (2019): Prototype-driven alternations: The case of German weak nouns. *Corpus Linguistics and Linguistic Theory* 15 (2), 383–417.
- Schäfer, Roland & Felix Bildhauer (2012): Building large corpora from the web using a new efficient tool chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (Hrsg.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 486–493. Istanbul: ELRA.
- Seiler, Guido (2022): Wie viele Kasus hat das Deutsche? In Sarah Brommer & Kersten Sven Roth (Hrsg.), *Brückenschläge. Linguistik an den Schnittstellen*, 39–64. Tübingen: Narr.
- Schmidt, Thomas (2016): Good practices in the compilation of FOLK, the research and teaching corpus of spoken German. *International Journal of Corpus Linguistics* 21 (3), 396–418.
- Scott, Alan K. (2014): *The Genitive Case in Dutch and German. A Study of Morphosyntactic Change in Coded Languages*. Leiden, Boston: Brill.
- Sick, Bastian (2005): *Der Dativ ist dem Genitiv sein Tod. Folge 2. Neues aus dem Irrgarten der deutschen Sprache*. 4. Auflage. Köln: Kiepenheuer & Witsch.
- Szczepaniak, Renata (2014): Sprachwandel und sprachliche Unsicherheit: Der formale und funktionale Wandel des Genitivs seit dem Frühneuhochdeutschen. In Albrecht Plewnia (Hrsg.), *Sprachverfall? Dynamik – Wandel – Variation*, 33–49. Berlin & Boston: De Gruyter.
- Thieroff, Rolf (2003): Die Bedienung des Automatens durch den Mensch. Deklination der schwachen Maskulina als Zweifelsfall. *Linguistik Online* 16 (4), 105–117.
- Weber, Thilo & Sandra Hansen (2024): Schwankungen zwischen schwacher und starker Substantivflexion. In: Marek Konopka, Angelika Wöllstein & Ekkehard Felder (Hrsg.), *Substantivflexion, Attributsätze, Präfix- und Partikelverben*. (Bausteine einer Korpusgrammatik 3), 13–58. Heidelberg: Heidelberg University Publishing.
- Weiβ, Helmut (2005): Von den vier Lebensaltern einer Standardsprache. Zur Rolle von Spracherwerb und Medialität. *Deutsche Sprache* 33 (4), 289–307.
- Zweifelfälle-Wörterbuch = Hennig, Mathilde (Hrsg.) (2021): *Dudenband 9 – Sprachliche Zweifelfälle. Das Wörterbuch für richtiges und gutes Deutsch*. 9. Auflage. Berlin: Dudenverlag.

Korpora und Quellennachweise

DECOW: s. Schäfer & Bildhauer (2012)

Leibniz-Institut für Deutsche Sprache: *Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK) / Archiv für Gesprochenes Deutsch (AGD)*. <https://agd.ids-mannheim.de/folk.shtml> (abgerufen am 31.10.2023)

Leibniz-Institut für Deutsche Sprache (2022): *Deutsches Referenzkorpus / Archiv der Korpora geschriebener Gegenwartssprache 2022-I* (Release vom 08.03.2022). Mannheim: Leibniz-Institut für Deutsche Sprache. <http://www.ids-mannheim.de/DeReKo> (abgerufen am 18.09.2022)