

Franziska Buchmann

# Die Bindestrichschreibung im Wort

Eine qualitative und quantitative Analyse von ausgewählten Schreibungen in Zeitungstexten des Deutschen

Die bisherige Forschung zum Bindestrich hat u.a. gezeigt, dass dieser in Komposita (z.B. *Panorama-Fenster*, *Uni-Gebäude*) bzw. in Wortbildungen, die komposita-ähnliche Strukturen aufweisen (z.B. *Ex-Freund*, *Maxi-Rock*), genutzt wird, wenn das Erstglied fremdwortartig ist. Dieser Aufsatz fokussiert nun Schreibungen mit Bindestrich im DeReKo.<sup>1</sup> Ausgewählte Bindestrichschreibungen mit fremdwortartigen Erstgliedern werden im DeReKo von ihrem ersten Auftreten an bis heute (gestaffelt nach Jahrzehnten) in einer quantitativ-orientierten Querschnittsuntersuchung analysiert. Dabei stehen drei Fragenkomplexe im Fokus.

Wann kommen die Bindestrichschreibungen in diesen Komposita bzw. Wortbildungen erstmals vor? Werden die Wortbildungen mit ihrem ersten Vorkommen im DeReKo mit Bindestrich geschrieben oder tritt die Bindestrichschreibung nachträglich in etablierte Wortbildungsprodukte?

Wie ist ihr quantitatives Verhältnis zur entsprechenden bindestrichlosen Äquivalentschreibung? Wird also eine bindestrichhaltige oder eine bindestrichlose Schreibung über die Jahrzehnte etabliert?

Wenn die Setzung eines Bindestrichs zeitgleich mit der Wortbildung selbst auftaucht, lassen sich dann Rückschlüsse auf die Integration des fremden Erstglieds ziehen? Es wird also immer auch die Vorkommenshäufigkeit des Erstglieds bzw. des Fremdworts alleine untersucht.

Die fremdwortartigen Erstglieder sind Fremdwörter, Kurzwörter, deren Langform ein Fremdwort ist, sowie morphologische Einheiten, die einen ungeklärten morphologischen Status aufweisen und typischerweise als Konfixe oder fremde Präfixe beschrieben werden.

---

<sup>1</sup> Daten einsehbar unter <https://hdl.handle.net/10932/00-06D3-E454-61FF-FF01-6>.

# 1 Warum ist das eine interessante Fragestellung?

## 1.1 Forschungsstand zur Bindestrichschreibung

Der Bindestrich kann mit Bredel (2008; 2020) als ein besonderer Gebrauch vom Divis gesehen werden. Unter dem Divis werden verschiedene Vorkommen des kurzen, horizontalen Strichs zusammengefasst: 1. der Trennstrich am Zeilenende in *Aufsatz-<sup>[Zeilenende]</sup>manuskript*, 2. der Ergänzungsstrich in *be- und entladen*, 3. der Bindestrich in *Alumni-Tag* und 4. einen Durchkopplungsstrich in *dein ewiges Ist-mir-doch-egal* (Beispiele aus Bredel 2008: 111–112 sowie Buchmann 2015).

Gallmann (1985) untersucht die Funktionen des Bindestrichs und unterteilt ihn in den (Kopplungs-)Bindestrich, den Trennstrich sowie den Ergänzungsstrich.<sup>2</sup> Für den Bindestrich nimmt Gallmann eine Grenzsignalfunktion zwischen Morphemen an (1985: 66–67) (vgl. dazu auch Gallmann 1989; 1996 mit der Zusammenfassung von Apostroph und Bindestrich als Grenzsignale zum Begriff *Syngraphem*). Explizit nur mit dem Trennstrich am Zeilenende beschäftigen sich Geilfuß-Wolfgang (2007) und Neef (2008), nur mit dem Ergänzungsstrich beschäftigt sich Smith (2000). Bernabei (2003) systematisiert auf empirischer Basis Bindestrichschreibungen im Gegenwartsdeutschen unter einer orthographischen Perspektive, im Sinne einer Ratgeberliteratur. Buchmann (2015) analysiert ebenfalls auf empirischer Basis Schreibungen mit Bindestrich unter einer graphematischen Perspektive und mit Blick auf ein Gesamtsystem der Wortzeichen. Eine historische Analyse von (Bindestrich-)Schreibungen an der Kompositionsfuge findet sich in Kopf (2018). Die Schreibungen mit Divis im Deutschen sind trotz dieser Forschungsarbeiten immer noch mit Fragen behaftet.

In Buchmann (2015: 217–289) wurden Bindestrichschreibungen qualitativ analysiert, die zwischen Januar 1995 und Juli 2008 im *Mannheimer Morgen*, einer regionalen Tageszeitung, erschienen. Das Korpus selbst umfasste zum Zeitpunkt der Datenerhebung im Juli 2008 240.215.316 fortlaufende Wörter; nach einer Bereinigung der Daten per Hand umfasste die analysierte Datenmenge ca. 270.000 Bindestrichschreibungen. Die Liste zählt sowohl die verschiedenen Wörter mit Bindestrich als auch die auftauchenden Flexionsformen eines Worts mit Binde-

---

<sup>2</sup> Anders als bei Bredel ist *Divis* hier nicht der Oberbegriff für die verschiedenen Vorkommen des kurzen, geraden Strichs, sondern er ist ein Synonym für den Begriff *Bindestrich*. Ich schließe mich im Folgenden Bredel (2008) an und verwende den Begriff *Divis* als Oberbegriff, wenn ich auf alle Vorkommen Bezug nehme, und den Begriff *Bindestrich*, wenn ich auf das Vorkommen als Bindestrich Bezug nehme.

strich als je einen Type.<sup>3</sup> Weiterhin wurde erfasst, wie viele Treffer pro Eintrag im *Mannheimer Morgen* zu finden waren, also zugehörige Token. Die Daten wurden einerseits nach Wortarten und andererseits nach Anzahl der enthaltenen Bindestriche sortiert: Es kommen substantivische und adjektivische Komposita vor und einige wenige Syntagmen, die als ein syntaktisches Wort fungieren: *Bürger-Büro*, *kfz-freien*, *Hans-Dampf-in-allen-Gassen*. Die gefundenen Bildungen sind aber nicht gleich verteilt: Innerhalb der gesamten Datenmenge finden sich auf ca. 270.000 Schreibungen beispielsweise nur 7076 Adjektive mit einem oder mehr Bindestrichen (vgl. dazu Buchmann 2015: 254).

**Tab. 1:** Anzahl der Bindestriche

Schreibungen mit 1 Bindestrich	248.000
Schreibungen mit 2 Bindestrichen	26.000
Schreibungen mit 3 Bindestrichen	1000
Schreibungen mit 4 Bindestrichen	109
Schreibungen mit 5 oder mehr Bindestrichen	11

Im Gegenwartsdeutschen werden morphologische Wörter in einer ununterbrochenen Zeichenkette zwischen zwei Spatien geschrieben. Dafür werden Buchstaben bzw. Grapheme verwendet, nur der erste Buchstabe wird ggf. großgeschrieben: <Ununterbrochene Zeichenketten aus Buchstaben bilden den Kernbereich der Zusammenschreibung.> (vgl. dazu Fuhrhop 2008: 194). U.a. Wortzeichen sorgen dafür, dass diese Zusammenschreibung im Kernbereich unterbrochen wird; die entstandene Schreibung ist graphisch markiert bzw. eine Schreibung des Randbereichs:

- (1) Un-unterbrochene Zeichen-Ketten aus Buchstaben bilden den Kern-Bereich der Zusammen-Schreibung.
- (2) Ununterbroch'ne Zeich'n-Kett'n bild'n den Kernbereich der Zusamm'-Schreibung.
- (3) 6-jährige Kinder sind schulpflichtig.
- (4) ABM-Maßnahmen werden seit 2012 nicht mehr gefördert.

<sup>3</sup> Da die Untersuchung nicht statistisch-quantitativ ausgelegt war, bestand auch keine Notwendigkeit, die Daten so aufzubereiten. Es ging damals um die Frage, welche Bindestrichschreibungen tauchen überhaupt in einem Korpus der Standardsprache auf und lassen sich die gefundenen Schreibungen grammatisch interpretieren und systematisieren.

- (5) „Harry Potter“-Geschichten werden immer noch gerne gelesen.

Neben der Anwesenheit eines Wortzeichens kann auch die Verwendung von weiteren Zeichen, die keine Minuskeln sind, zu einer Abweichung vom Kernbereich der Zusammenschreibung führen. Hier findet sich eine Art doppelte Markiertheit: Zum einen enthalten die Beispiele den Bindestrich selbst (Markiertheit 1. Grades), zum anderen sind die Wortelemente selbst auch abweichend von der ununterbrochenen Kleinschreibung gestaltet (Markiertheit 2. Grades): Ziffern, Majuskeln, Anführungszeichen sind nur ein Teil der Zeichen, die hier zu finden sind.

Ich werde mich im Folgenden der Einfachheit halber hauptsächlich auf Schreibungen beziehen, die einen Bindestrich enthalten. Wer genaueres über die mehrbindestrichhaltigen Schreibungen erfahren will, sei auf die Ausführungen in Buchmann (2015) verwiesen. Die Bindestrichschreibungen lassen sich in zwei große Gruppen unterteilen: 1. die obligatorischen Bindestrichschreibungen, die durch rein graphematische Kriterien hervorgerufen werden, sowie 2. die fakultativen Bindestrichschreibungen, die durch morphologische Kriterien hervorgerufen werden.

In (6)–(8) stehen typische Bindestrichschreibungen, die durch graphische Kriterien hervorgerufen werden. Die Schreibungen weisen eine Markiertheit 2. Grades auf, da nicht nur der Bindestrich selbst sie markiert, sondern auch die Wortformen selbst eine besondere, graphische Struktur aufweisen:

- (6) BMX-Fahrrad, enjoy-Tarif, FlowTex-Betrugsfall
- (7) 35 000-Marke, 0190er-Nummern
- (8) Messdiener/innen-Treffen, „Big Brother“-Kandidat, Tel.-Nummer, Geburtstags-Rock’n’Roll

Die graphischen Auszeichnungen der Wortformen sind vielfältig (hier bezogen auf die Markiertheit 2. Grades). Es lassen sich vom Kernbereich abweichende Gebräuche der Majuskeln und Minuskeln finden (durchgängige Majuskelschreibung in *BMX*, durchgängige Minuskelschreibung in *enjoy*, Binnenmajuskeln in *FlowTex*); Ziffern als Wortglieder und sogar Kombinationen aus Ziffern und Buchstaben *35 000* vs. *0190er*; Wortglieder mit Sonderzeichen in *Messdiener/innen*, mit Anführungszeichen „*Big Brother*“ oder mit anderen Wortzeichen in *Tel.* und *Rock’n’Roll*. Eine ausführliche Diskussion der Schreibungen findet sich in Buchmann (2015: 219–259).

Keines dieser Wörter weist eine bindestrichlose Alternativschreibung auf und die bindestrichlose Alternativschreibung wird auch nicht akzeptabler, wenn das Zweitglied (nach dem nun weggelassenen Bindestrich) klein geschrieben wird:

- (9) \*Ajugend, \*Ajugend, \*BMXFahrrad, \*BMXfahrrad, \*enjoyTarif,  
\*enjoytarif, \*FlowTexBetrugsfall, \*FlowTexbetrugsfall
- (10) \*35 000Marke, \*35 000marke, \*0190erNummern, \*0190ernummern
- (11) \*Messdiener/innenTreffen, <sup>7</sup>Messdiener/innentreffen<sup>4</sup>,  
\*, „Big Brother“Kandidat, \*, „Big Brother“kandidat, \*Tel.Nummer,  
\*Tel.nummer, \*GeburtstagsRock’n’Roll, \*Geburtstagsrock’n’roll

Durch den Bindestrich können die morphologischen Bildungen<sup>5</sup> überhaupt erst in einem graphematischen Wort zusammengeschrieben werden.<sup>6</sup>

Kommen wir nun zu den fakultativen Bindestrichschreibungen, in denen morphologische Kriterien dazu führen, dass der Bindestrich gesetzt wird (eine ausführliche Diskussion dieser und weiterer Fälle finden sich in Buchmann 2015: 236–249):

- (12) Alumni-Tag, Panorama-Glasschiebedach, Telefon-Terror
- (13) Casablanca-Bühne, Olympia-Gastgeber, Werder-Angreifer
- (14) Uni-Sprecher, Lok-Führer
- (15) Ex-Freundin, Öko-Steuer, Bio-Produkte

Die Bindestrichschreibungen in (12)–(15) enthalten Wortformen, die selbst keine graphischen Besonderheiten aufweisen. Es handelt sich also um eine Markiertheit 1. Grades. Nur der Bindestrich selbst sorgt für eine Abweichung vom Kernbereich der Zusammenschreibung im Sinne einer ununterbrochenen Zeichenkette. Die Bindestrichschreibungen hier werden durch rein morphologische Kriterien her-

---

4 <Messdiener/innentreffen> ist zwar vorstellbar, kommt aber im MM-Korpus nicht vor. Der Schrägstrich als Sonderzeichen ist hier ein pragmatisches Zeichen, er zeigt hier zwei (biologische) Geschlechter an: *Messdiener* und *Messdienerinnen*. Eine vertiefte Debatte zur graphischen Auszeichnung der biologischen Geschlechter bzw. der sozialen Geschlechter, im Sinne einer gendersensiblen Sprache, ist ein eigenes Thema und bedarf einer umfassenderen Analyse als es eine Fußnote hier zulassen würde.

5 Die meisten hier gezeigten Schreibungen sind Komposita bis auf *0190er-Nummern*. Der Bildungsprozess ist unklar. Fun Fact: Seit dem 30.06.2010 existieren keine 0190er-Nummern mehr. Es handelte sich um Bezahltelefonnummern, die für Serviceleistungen aller Art und durch Sicherheitslücken am PC auch für sog. Dialer-Betrügereien bei der Einwahl ins Internet genutzt wurden (vgl. dazu <https://www.sueddeutsche.de/digital/0190-nummern-das-stoehnen-hat-einende-1.622365>, letzter Zugriff am 27.09.2022).

6 Dies gilt für die hier gezeigten morphologischen Bildungen: Natürlich finden sich Schreibungen wie <Schüler/innen> (mit 642 Treffern im MM-Korpus). Die Klammer lässt eine bindestrichlose Schreibung zu: *(Theater)Projekte werden finanziert*. Auch die Schreibung *(Theater-)Projekte* ist möglich. Hier kann man eine Variation beobachten, die noch zusätzlicher Klärung bedarf.

vorgerufen: Fremdwörter, Eigennamen, Kurzwörter und Einheiten mit einem ungeklärten morphologischen Status sorgen für eine Bindestrichschreibung. Dabei muss konstatiert werden, dass diese Kriterien natürlich gebündelt auftauchen können: *Casablanca* ist Fremdwort und Eigenname zugleich, *Uni* ist Fremdwort und Kurzwort zugleich. In *Bio-Produkte* findet sich neben der Einheit mit ungeklärtem Status (*Bio*) ein weiteres Fremdwort (*Produkte*).

## 1.2 Methodik der durchgeführten Analyse

Im Folgenden werden nur fakultative Bindestrichschreibungen in den Fokus der Analyse gestellt. Die qualitative Auswertung der Bindestrichschreibung aus Buchmann (2015) erfolgte anhand des *Mannheimer Morgen*-Korpus, das die Jahre 1995 bis 2008 abdeckt. Dieses Korpus erlaubt damit keine Zeitschnitte mit größeren Abständen und damit auch keine diachron orientierte Analyse. Daher wurden für die folgende quantitativ ausgerichtete Analyse Bindestrichschreibungen aus der ersten Datenerhebung im *DeReKo 2022-I* kontrolliert.<sup>7</sup> Das DeReKo konzentriert sich mit seinen Texten auf die zweite Hälfte des 20. Jahrhunderts (vgl. dazu <https://www.ids-mannheim.de/digspra/kl/projekte/korpora/archiv-1/>). Die gefundenen Daten lassen sich also nach Jahrzehnten filtern. Jedes Jahrzehnt bildet dann einen Zeitschnitt (vgl. dazu die gezeigten Beispiele in Tab. 4–11).

Es wurden zunächst Bindestrichschreibungen aus dem MM-Korpus ausgewählt, deren Wortbestandteile einem oder mehreren der oben genannten morphologischen Kriterien *Fremdwort*, *Kurzwort*, *Präfix/Konfix/ungeklärt*, *Eigenname* entsprechen. Pro morphologischem Kriterium wurden dann im MM-Korpus verschiedene passende Erstglieder ausgesucht. Hier wurde versucht, die Tokenzahlen zu berücksichtigen.<sup>8</sup> Es sollten Wortverbindungen gefunden werden, die neben ihrem Vorkommen im MM-Korpus auch tatsächlich mit einer gewissen Häufigkeit im DeReKo vorkommen. Die ausgewählten Wortverbindungen wurden dann im DeReKo gemeinsam mit ihren alternativen bindestrichlosen Schreibungen ge-

<sup>7</sup> Alle hier genannten Zahlen beziehen sich auf den Stand September/Oktober 2022.

<sup>8</sup> Dieses Kriterium ist leider sehr weich und damit angreifbar. Eine reine Auswahl über die Tokenzahl im MM-Korpus ist leider kaum möglich. Die Bindestrichschreibungen mit den höchsten Tokenzahlen entsprechen nicht den fakultativen Bindestrichschreibungen, sondern den obligatorischen. Und die interessieren ja hier in diesem Zusammenhang genau nicht. Bindestrichschreibungen mit sehr kleinen Tokenzahlen (unter 10) haben häufig auch sehr wenige alternative bindestrichlose Schreibungen. D.h. diese Wörter kommen einfach selten in Zeitungstexten vor, eine diachrone Überprüfung ist damit u.U. gar nicht leistbar. Letztendlich sind die Ausgangstokenzahlen im MM-Korpus sehr unterschiedlich.

sucht. In einem letzten Schritt wurden dann aus dem DeReKo auch andere, sinnvolle Komposita zu diesen ersten Wortverbindungen ergänzt.

Neben dem morphologischen Kriterium und der Tokenzahl sollten die untersuchten Wortverbindungen auch in der zweiten Hälfte des 20. Jahrhunderts neu im Deutschen auftauchen; dies trifft mit ziemlicher Sicherheit auf Wortverbindungen mit *Internet*, *Inliner* oder *Aids* zu, aber auch *Asbest* war ein Thema des Jahrtausendwechsels. Bei der Kategorie „fremdes Präfix/Konfix/ungeklärter Status“ sind es die Einheiten *bio-*, *öko-*, *ex-*.

Herausgekommen ist ein kleines Korpus mit mehr oder weniger zufällig ausgewählten Wortverbindungen: Es handelt sich um 91 Bindestrichschreibungen, im Sinne von Types. Alle 91 Bildungen wurden im Kontext ihrer (bindestrichlosen) Alternativschreibungen untersucht.<sup>9</sup> 51 Wortverbindungen werden hier detailliert gezeigt werden. Damit lassen sich natürlich keine Rückschlüsse auf alle Bindestrichschreibungen dieser Art ziehen.<sup>10</sup> Trotzdem zeigt sich, dass die analysierten Wortverbindungen im Großen und Ganzen ein sehr ähnliches Verhalten aufweisen.

Zum Umgang mit gefundenen Schreibungen: Die untersuchten Eigennamen werden hier im Aufsatz nicht präsentiert. Die Klassen der Eigennamen sind sehr divers. In den Oberkategorien *Personennamen*, *Tiernamen*, *Ortsnamen*, *Objektnamen*, *Ereignisnamen*, *Phänomennamen* finden sich unzählige Subklassen, denen hier kaum gerecht werden kann (vgl. dazu u.a. Nübling, Fahlbusch & Heuser 2015). Namen weisen ein eigenes grammatisches Verhalten auf, das sich auch in der Schreibung niederschlägt (vgl. zum syntaktischen Verhalten Eisenberg 2020: 158–160; zur Schreibung Fuhrhop 2020: 35–41). Sie können aus einem Wort bestehen oder aus mehreren Wörtern: *Airbus* als Unternehmensname vs. *Deutsche Bahn* als Unternehmensname. Darüber hinaus müssten Komposita, die einen Namen enthalten, von Verbindungen, die als Ganzes ein Name sind, unterschieden werden. Die Übergänge sind fließend: *Bahn-Chef* vs. *Johann-Gottfried-Herder*.

---

<sup>9</sup> Ich danke Thees Becker für die tatkräftige Unterstützung bei der Datensammlung. Die ausgewählten Komposita bzw. Wortverbindungen mussten aufwendig per Hand in COSMAS eingegeben und gefiltert werden. Siehe dazu auch Fußnote 10.

<sup>10</sup> Die Erzeugung von quantitativ-statistisch auswertbaren Daten im Sinne einer Korpuslinguistik weist aktuell immer noch die gleiche Problematik auf wie schon 2008 bei der Erstellung des Korpus für die qualitative Untersuchung der Bindestrichschreibungen in Buchmann (2015). Das DeReKo und die Suchanfragesyntax von COSMAS sind nicht entsprechend aufbereitet, dass alle bindestrichhaltigen Schreibungen eines Zeitschnitts automatisiert gesucht und extrahiert werden können. Dies liegt vor allem an der Annotation des Bindestrichs. D.h. alle Bindestrichschreibungen müssen per Hand in COSMAS eingegeben werden, man muss also die Wortverbindungen kennen und dann die Schreibung überprüfen.

*Gymnasium*, aber *Carl von Ossietzky Universität Oldenburg*. Die Anwesenheit von Durchkopplungsbindestrichen kann auch davon abhängen, ob diese im namensgebenden Akt als Schreibung festgelegt wurden.

Die Einschränkung auf die Fremdwörter, Kurzwörter und Einheiten mit ungeklärtem Status hat den positiven Effekt, dass die verbleibenden Kategorien alle lose dem Fremdwortbereich angegliedert werden können. Neben den Fremdwörtern sind auch die hier gezeigten Kurzwörter (*Lok*, *Aids*<sup>11</sup>, *Profi*) sowie die Einheiten mit ungeklärtem Status (*ex-*, *bio-*, *öko-*) dem Fremdwortbereich zuzuschlagen.

Eine weitere Einschränkung muss hier erwähnt werden: Eine durchgängige Majuskelschreibung des gesamten Kompositums wurde als Überschriftschreibung gewertet. Diese Variation kommt bei fast jedem der hier gezeigten Wörter vor und sie umfasst i.d.R. wenige Treffer. Kommt sie vor, ist das in den Tabellen ausgezeichnet und sie ist dann in die hier gezeigten Varianten jeweils inkludiert. Enthalten die hier untersuchten Wortverbindungen hingegen einen Trennstrich am Zeilenende, so ist dieses Vorkommen nicht in die hier gezeigten Varianten der Schreibungen inkludiert. Dies hat vor allem praktische Gründe.<sup>12</sup> Man muss annehmen, dass die Tokenzahlen der bindestrichlosen Schreibungen höher wären, wenn diese Schreibungen inkludiert wären.

---

11 *Aids* ist vielleicht das einzige der hier gezeigten Lexeme, dessen Eigennamenstatus präsenter ist als der Kurzwortstatus (vgl. dazu auch Abschnitt 2.3).

12 Jede Schreibung muss per Hand im DeReKo gesucht werden. COSMAS schlägt dem Nutzer dann zunächst Wortformen vor; an dieser Stelle können Treffer automatisch ausgeschlossen werden, die nicht relevant sind. Die Trennstriche tauchen als vorgeschlagene Wortform bei den bindestrichhaltigen Wortverbindungen auf und werden in einem ersten Schritt aus den für die Bindestrichschreibung relevanten, anzuzeigenden Treffern ausgeschlossen. Warum wird am folgenden Beispiel deutlich: *Bundesliga-Geschichte* weist 3572 Token seit 1960 auf. Vorgeschlagene Wortformen in COSMAS sind auch *Bundesliga-geschichte* (4 Token) und *bundesliga-geschichte* (1 Token). Mindestens der Bindestrich in *Bundesliga-geschichte* dürfte aufgrund der Konkatenation des Bindestrichs ein Trennstrich am Zeilenende sein. In einem zweiten Schritt wird die Zusammenschreibung gesucht. *Bundesligageschichte* weist 1406 Token seit 1980 auf. Hier müssten für jeden Zeitschnitt die Treffer für die Trennstriche am Zeilenende aufwendig kontrolliert und zur bindestrichlosen Zusammenschreibung addiert werden. Das mag für die 4 Token *Bundesliga-geschichte* machbar sein. Aber es ist nicht gesagt, dass der gefundene Trennstrich der einzige Trennstrich in der entsprechenden Wortform ist: *Bundesligageschichte* weist deutlich mehr Silbengrenzen auf, an denen ein Trennstrich stehen kann, als die Silbengrenze, an der der Bindestrich steht. Dies müsste ebenfalls per Hand kontrolliert und ggf. zu den bindestrichlosen Zusammenschreibungen addiert werden. Dieses Vorgehen würde zwar sehr exakte Token für die bindestrichlose Zusammenschreibung ergeben, es ist aber auch per Hand kaum durchzuführen. Die Treffer der bindestrichlosen Zusammenschreibungen beziehen sich also auf Zusammenschreibungen ohne Trennstriche am Zeilenende.



### 1.3 Hypothesen zur Interpretation der gefundenen Schreibungen

Wie sollen nun die gefundenen Schreibungen interpretiert werden? Tatsächlich stehen verschiedene Szenarien und damit Interpretationsansätze im Raum.

Die erste Hypothese ist: Mit zunehmendem Gebrauch eines Kompositums nimmt die Bindestrichschreibung ab. Morphologische Wörter werden zusammengeschrieben; graphematisch handelt es sich um eine ununterbrochene Zeichenkette (vgl. zur Getrennt- und Zusammenschreibung im Deutschen Fuhrhop 2007; Jacobs 2005). Die Bindestrichschreibung interveniert hier, sie führt zu einer graphischen Markiertheit des gesamten Kompositums. Lässt sich also innerhalb eines Kompositums mit zunehmendem oder fortlaufendem Gebrauch ein Wandel feststellen von einer durch Bindestrich unterbrochenen Zusammenschreibung zu einer Zusammenschreibung ohne Bindestrich, so handelt es sich um die Überführung einer Schreibung von einer graphischen Markiertheit 1. Grades, ausgelöst durch das Fremdwort im Erstglied, in eine graphische Unmarkiertheit.

Eine zweite Hypothese schließt daran an: Ist die Bindestrichschreibung ein Integrationsmarker für die Fremdwörter bzw. die fremden Wortbestandteile selbst? Dies ist eindeutig der Fall, wenn der bindestrichauslösende fremdwortartige Erstbestandteil erstmals mit dem hier untersuchten Kompositum im DeReKo auftritt: Der Übergang von einer bindestrichhaltigen Zusammenschreibung zu einer bindestrichlosen Zusammenschreibung stellt dann ein Zeichen für die Integration des Fremdworts selbst dar (vgl. zum Fremdwortbegriff sowie zum grammatischen Verhalten von Fremdwörtern vor allem Eisenberg 2018).

Eine dritte Hypothese lässt sich mit einem anderen Szenarium verbinden: Kommt das bindestrichauslösende Fremdwort selbst schon vor der hier untersuchten Wortbildung im DeReKo vor, so muss es sich nicht um Integration des Fremdworts handeln.<sup>13</sup> Denn dieses ist ja schon im Deutschen vorhanden und

---

<sup>13</sup> Die Interpretation einer solchen Abfolge der Wörter ist leider nicht einfach: Natürlich kann die Beteiligung eines Fremdworts an Komposition im Deutschen an sich als Integrationsmarker interpretiert werden. Dann müssten aber verschiedene Sachverhalte rund um das Fremdwort geprüft werden: 1. Ist das Kompositum, in dem das Fremdwort steht, im Deutschen gebildet oder eine Entlehnung? Dann kann man entscheiden, ob die Komposition im Deutschen stattgefunden hat und ob das Fremdwort also Anteil hat an dem Wortbildungsmuster des Deutschen. 2. Ist das Fremdwort als wortfähiger Stamm ins Deutsche gekommen oder als Teil einer Wortbildung? Hat sich also der fremde Stamm aus einer Wortbildung im Deutschen herausbewegt und ist hier wortfähig geworden oder war er das schon? Auch mit einer solchen Information ist eine gezieltere Aussage über den morphologischen Integrationsprozess des Fremdworts im Deutschen möglich. 3. Ist das hier untersuchte Kompositum das erste Kompositum mit dem entsprechenden

wird u.U. wortfähig gebraucht. Wenn die bindestrichhaltige Schreibung von der bindestrichlosen Schreibung abgelöst wird, kann es sich um eine Art Etablierung des Kompositums selbst handeln, also um die Aufnahme eines Wortbildungsprodukts in den Wortschatz des Deutschen. Diese These findet sich auch schon bei Satkauskaite (2008): Sie untersucht in den beiden Korpora Archiv W-öffentlich sowie Archiv N2-öffentlich (ebenfalls vom IDS über COSMAS zugänglich) die beiden Fremdwörter *Computer* und *Web* hinsichtlich ihrer Schreibung in ausgewählten Komposita<sup>14</sup>, und zwar zwischen 1983 und 2008 (Satkauskaite 2008: 254). Sie findet zunächst eine schwankende Schreibung der Komposita hinsichtlich der Verwendung des Bindestrichs, dann zeigen allen Schreibungen eine Tendenz zur Zusammenschreibung ohne Bindestrich (Satkauskaite 2008: 256–259). Sie interpretiert diesen Befund (ohne diesen näher auszuführen) als Konventionalisierung des Kompositums (Satkauskaite 2008: 259).

Mit Schmid (2008: 2–3) lässt sich die Etablierung neuer Wörter (*establishing*) unter drei Perspektiven und mit drei Stadien fassen:

**Tab. 2:** Drei Perspektiven und drei Stadien, die ein Wort bei seiner Etablierung durchläuft (vgl. Schmid 2008: 2)

		Perspectives		
		Structural perspective	Socio-pragmatic perspective	Cognitive perspective
Stages	creation	(product of) nonce-formation	(process of) nonce-formation	pseudo-concept
	consolidation	stabilization	spreading	(process of) hypostatization
	establishing	lexicalized lexeme	institutionalized lexeme	hypostatized concept

Fremdwort? Ist dies der Fall, kann es sich um eine Integration des Fremdworts handeln. Gibt es jedoch Komposita mit dem Fremdwort, die früher im Deutschen auftauchen als die hier analysierten, muss es sich nicht um morphologische Integration des Fremdworts handeln, kann es aber.

<sup>14</sup> Es handelt sich um *Computerprogramm*, *-spezialist*, *-kriminalität*, *-bildschirm*, *-simulation*, *-animation* sowie *Webseite*, *-browser*, *-adresse*, *-design*. Für alle Komposita wurden auch die zugehörigen Flexionsformen untersucht.

Die erste Perspektive ist eine strukturelle, die zweite ist eine sozio-pragmatische und die dritte eine kognitive. Unter struktureller Perspektive ist ein Wort dann etabliert, wenn sich seine Form und Bedeutung stabilisiert haben und es lexikalisiert ist. Lexikalisierung meint hier eine Idiomatisierung, die einhergehen kann mit dem Verlust von Motiviertheit oder Transparenz. Unter sozio-pragmatischer Perspektive lässt sich etablierten Wörtern eine größere Akzeptanz zusprechen, im Sinne eines Gebrauchs durch viele verschiedene Sprecher. Mögliche Ambiguitäten und Kontextabhängigkeiten sind abgebaut, sodass die Sprecher das Wort problemlos verstehen und seine Bedeutung identifizieren können (Schmid 2008: 4). Unter kognitiver Perspektive werden im Zuge der Etablierung neuer Wörter bzw. Konzepte zwei Begriffe genannt: zum einen der Begriff Hypostasierung (in der Sprachphilosophie oder Semantik), zum anderen der Begriff Entrenchment (in der Psycholinguistik). Hier steht einerseits die Bildung eines neuen semantischen Konzepts im Vordergrund, das mit dem neugebildeten Wort benannt wird, und andererseits der Eingang in das mentale Lexikon der Sprecher (Schmid 2008: 5; vgl. dazu auch Ziem & Fritsche 2018, die Entrenchment als Verfestigung eines sprachlichen Ausdrucks zu einer kognitiven bzw. sprachlichen Einheit definieren).

In dieser Untersuchung wird erstens untersucht, wann eine Wortbildung das erste Mal im DeReKo auftaucht, und zweitens, wie sich die bindestrichhaltigen und die bindestrichlosen Schreibungen dieser Wortbildung über die Zeitschnitte im DeReKo verteilen. Darüber hinaus wird vermerkt, in wie vielen Texten die entsprechenden Schreibungen auftauchen. Als viertes wird analysiert, wann und in welchem Umfang die bindestrichauslösenden fremdwortartigen Erstglieder als wortfähige Einheiten im DeReKo vorkommen, und zwar immer dann, wenn diese auch suchbar sind. Mit diesen Daten können vorsichtige Rückschlüsse auf die strukturelle und sozio-pragmatische Perspektive der Etablierung eines Wortes gezogen werden. Und es ist eine Hypothese, dass mit zunehmender Etablierung, im Sinne von Gebrauch und Streuung der Wortverbindungen über verschiedene Texte hinweg, die Bindestrichschreibung zugunsten einer bindestrichlosen Zusammenschreibung zurückgeht.

## 2 Quantitatives Vorkommen ausgewählter fakultativer Bindestrichschreibungen im DeReKo

**Tab. 3:** Absolute und prozentuale Verteilung der Schreibungen aller Wortbildungen mit fremdwortartigen Erstgliedern

Type	Token mit Bindestrich		Token ohne Bindestrich		Gesamt
Asbest_Sanierung	213	11,9%	<b>1578</b>	<b>88,1%</b>	1791
Asbest_Belastung	54	7,8%	<b>636</b>	<b>92,2%</b>	690
Asbest_Entsorgung	79	26,7%	<b>217</b>	<b>73,3%</b>	296
Asbest_Schäden	31	18,1%	<b>140</b>	<b>81,9%</b>	171
Asbest_Klage	13	31,7%	<b>28</b>	<b>68,3%</b>	41
Internet_Auktionshaus	<b>2.154</b>	<b>69,4%</b>	948	30,6%	3.102
Internet_Seite	11.910	8,4%	<b>130.556</b>	<b>91,6%</b>	142.466
Web_Seite	5.087	4%	<b>122.332</b>	<b>96%</b>	127.419
Inline_Skater	<b>6.191</b>	<b>62,6%</b>	3.698	37,4%	9.889
Inline_Skating	<b>3.349</b>	<b>67,8%</b>	1.593	32,2%	4.942
Inline_Skates	3.981	58,2%	2.857	41,8%	6.838
Shopping_Center	3.719	39,2%	<b>5.763</b>	<b>60,8%</b>	9.482
Gala_Abend	3.303	45,9%	3.887	54,1%	7.190
Pop_Song	700	14,6%	<b>4.098</b>	<b>85,4%</b>	4.798
Rock_Song	181	13,5%	<b>1.159</b>	<b>86,5%</b>	1.340
Tsunami_Opfer	<b>1.565</b>	<b>90,6%</b>	162	9,4%	1.727
Tsunami_Katastrophe	<b>3.471</b>	<b>90,2%</b>	377	9,8%	3.848
Zoo_Tiere	162	6,7%	<b>2.251</b>	<b>93,3%</b>	2.413
Zoo_Direktor	1.293	16,3%	<b>6.629</b>	<b>83,7%</b>	7.922
Zoo_Besuch	197	11,4%	<b>1.524</b>	<b>88,6%</b>	1.721
Zoo_Besucher	406	16,4%	<b>2.077</b>	<b>83,6%</b>	2.483
Lok_Führer	202	0,4%	<b>50.835</b>	<b>99,6%</b>	51.037
Lok_Schuppen	137	1,7%	<b>8.033</b>	<b>98,3%</b>	8.170
Aids_Kranke	1.822	47,1%	2.049	52,9%	3.871
Aids_Virus	<b>3.309</b>	<b>78,5%</b>	904	21,5%	4.213
Aids_Erreger	<b>1.486</b>	<b>77%</b>	443	23%	1.929
Profi_Fußball	4.709	19,3%	<b>19.650</b>	<b>80,7%</b>	24.359
Profi_Karriere	3.278	21,2%	<b>12.214</b>	<b>78,8%</b>	15.492

Type	Token mit Bindestrich		Token ohne Bindestrich		Gesamt
Profi_Spieler	490	14,9%	<b>2.798</b>	<b>85,1%</b>	3.288
Profi_Spielerin	126	11,8%	<b>940</b>	<b>88,2%</b>	1.066
Profi_Boxer	1.273	20,6%	<b>4.911</b>	<b>79,4%</b>	6.184
Profi_Boxerin	152	24,8%	<b>462</b>	<b>75,2%</b>	614
Ex_Mann	<b>11.130</b>	<b>83,9%</b>	2.142	16,1%	13.272
Ex_Manager	<b>7.252</b>	<b>93%</b>	546	7%	7.798
Ex_Bundeskanzler	<b>3.779</b>	<b>90,1%</b>	416	9,9%	4.195
Ex_Ministerpräsident	<b>5.823</b>	<b>92,9%</b>	443	7,1%	6.266
Ex_Bundesligist	<b>704</b>	<b>99,6%</b>	3	0,4%	707
Ex_Profi	<b>17.810</b>	<b>98%</b>	366	2%	18.176
Ex_Bundestrainer	<b>2.699</b>	<b>99%</b>	26	1%	2725
Bio_Produkte	3.953	37,6%	<b>6.556</b>	<b>62,4%</b>	10.509
Bio_Äpfel	240	50,8%	232	49,2%	472
Bio_Bananen	145	55,8%	115	44,2%	260
Bio_Tonne	707	8,1%	<b>7.974</b>	<b>91,9%</b>	8681
Bio_Abfall	241	7,1%	<b>3.130</b>	<b>92,9%</b>	3371
Bio_Ethik	107	3,6%	<b>2.869</b>	<b>96,4%</b>	2976
Bio_Bauer	1.362	20,8%	<b>5.194</b>	<b>79,2%</b>	6556
Öko_Strom	2.312	8,8%	<b>23.874</b>	<b>91,2%</b>	26.186
Öko_Steuern	2.947	8,1%	<b>33.301</b>	<b>91,9%</b>	36.248
Öko_Partei	4.634	42,6%	6.236	57,4%	10.870
Öko_Bauer	245	32,8%	<b>503</b>	<b>67,2%</b>	748
Öko_Anbieter	95	52,5%	86	47,5%	181

In Tab. 3 lassen sich die Gesamtzahlen (absolut und prozentual) der hier untersuchten Wortverbindungen nachvollziehen. Es zeigt sich schon auf den ersten Blick, dass über alle Zeitschnitte hinweg in vielen Fällen die bindestrichlose Schreibung die bevorzugte ist. Ein Blick in die Zeitschnitte wird dies nicht nur bestätigen, sondern sogar verstärken (s. Abschnitt 2.1).

Darüber hinaus muss noch auf einen weiteren Aspekt hingewiesen werden. In 1.2 wird sehr deutlich gezeigt, wie schwierig und teilweise zufällig die Auswahl der Wortverbindungen erfolgen musste. In Tab. 3 wird nun überdeutlich, dass sich die untersuchten Wortbildungen sehr stark voneinander unterscheiden und

jenseits der oben genannten Merkmale keine homogene Gruppe bilden.<sup>15</sup> Die Wortbildungen mit Fremdwort und Kurzwort sind eindeutig Komposita, die mit fremdem Präfix bzw. Konfix nur dann, wenn man akzeptiert, dass diese wortfähig im DeReKo vorkommen.<sup>16</sup> Es musste also mit Blick auf die vorhandenen Daten mit Augenmaß entschieden werden, welche sinnvollen Komposita bzw. Wortverbin-

---

15 Die Komposita mit Fremdwort im Erstglied sind in vielerlei Hinsicht voneinander verschieden: 1. neben typischen Determinativkomposita finden sich auch speziellere Rektionskomposita, z.B. *Asbest-Entsorgung*, *Zoo-Besuch*, 2. die Variation hinsichtlich *Fremdwort* befindet sich nicht nur im Erstglied, sondern teilweise auch im Zweitglied: *Asbest-Belastung* vs. *Asbest-Sanierung*, 3. die Fremdwörter kommen aus unterschiedlichen Sprachen, z.B. die Anglizismen *Popsong*, *Inline-Skater*, der Gräzismus *Zoo*, der Latinismus *Asbest*, der Hispanismus *Gala*.

Für die Schreibungen mit Kurzwort werden drei Kurzwörter in verschiedenen Wortverbindungen analysiert: *Lok* und *Profi* sind als Kurzwort deutlich länger im Deutschen vorhanden als *Aids*. Darüber hinaus ist *Aids* nicht nur ein Kurzwort, sondern auch der Name der Krankheit (vgl. dazu Abschnitt 2.3).

Die Wortverbindungen mit den Konfixen/fremden Präfixen sind ebenfalls nicht in sich konsistent: Alle eint das fremde Präfix/Konfix/die Einheit mit ungeklärtem grammatischen Status. *Ex-*, *bio-* und *öko-* können wortfähig verwendet werden, d.h. sie können im Satz als Wortform alleine stehen (vgl. Buchmann 2015: 246–249). Die Bestandteile nach dem Bindestrich können Fremdwörter (*Manager*, *Profi*, *Ministerpräsident*) sein oder Wörter des Kernbereichs (*Äpfel*, *Tonne*, *Bauer*, *Anbieter*). *Ministerpräsident* ist zudem ein Titel; *Bundesligist* steht im Zusammenhang mit *Bundesliga*, das als Objektname aufgefasst werden kann.

16 Im DeReKo lassen sich u.a. folgende Beispiele eines wortfähigen Gebrauchs finden. Man kann an den Einträgen deutlich sehen, dass dieser Gebrauch bis in die 80er Jahre zurückgeht.

1. *Sie ist Mitte vierzig und liebt einen jungen Burschen, kocht für ihn und für die gern reinschneien-de Nachbarin, auch mal für den Ex und die Tochter (glänzend: Elena Rublack), die einen Fiesling freit.* (Z89/DEZ.00234 Die Zeit, 08.12.1989, 68; Schiefelge Leben)

2. *Hellendorf ging es darum, das Vertrauen in Bio zu wecken und zu zeigen, welchen Kontrollen Biobetriebe unterworfen sind.* (K97/JUN.40794 Kleine Zeitung, 01.06.1997, Ressort: Völkermarkt; Vertrauen in Biokost wecken)

3. *[...] Schweine und dergleichen zum Anfassen und die netten Leute von der Ufa-Fabrik, die das hier alles wieder aufmöbeln und uns mit dem diskreten Charme der Alternativkultur bewirteten, bio bis zum letzten Reiskorn.* (Z82/JUN.00161 Die Zeit, 11.06.1982, 51; Punk-Wolf und sprechende Bäume)

4. *Nur wo Öko draufsteht, ist auch Öko drin* (U92/MAR.01912 Süddeutsche Zeitung, 10.03.1992, 0; Nur wo Öko draufsteht, ist auch Öko drin)

5. *Da steht also Wolfgang Niedecken, der alte Öko, auf der Bühne und muss sich wirklich blöde vorkommen, zumal da er sein Liedchen noch einmal singen muss – wegen mangelnder Euphorie des Publikums.* (U99/OKT.79502 Süddeutsche Zeitung, 05.10.1999, 21, Ressort: MEDIEN; Die Leinwand lebt)

6. *Denn man kann nicht sicher sein, daß wirklich jeder Stromanbieter, der sich öko nennt, auch wirklich hält, was er verspricht.* (L05/JUN.06223 Berliner Morgenpost, 29.06.2005, 6; Ökostrom - reine Vertrauenssache)

dungen im DeReKo vorkommen. Im Folgenden werde ich hypothesengeleitet die Ergebnisse der Detailanalyse darstellen.

## 2.1 Hypothese 1: Die Wortverbindungen etablieren eine bindestrichlose Zusammenschreibung

Die Beobachtungen aus Tab. 3 können mit Blick auf die Zeitschnitte verfeinert werden. Unabhängig von der Gesamttokenzahl und über alle Zeitschnitte hinweg streben die meisten Wörter oder Wortverbindungen der Zusammenschreibung ohne Bindestrich und damit einer graphematisch unmarkierten Schreibung entgegen. Die graphematisch markierte Schreibung mit Bindestrich ist also nicht die bevorzugte Schreibung. Die Wortverbindungen lassen sich dabei in 4 Gruppen einordnen:

- (1) In Gruppe 1 überwiegt in allen Zeitschnitten die bindestrichlose Schreibung (9 Types): z.B. *Asbestbelastung*, *Zoodirektor*, *Zoobesucher*, *Lokführer*, *Bioethik*, *Ökostrom*.
- (2) In Gruppe 2 findet ein Wechsel der Schreibvariante statt, und zwar von einer Variantenschreibung zu einer bindestrichlosen Schreibung (28 Types): z.B. *Asbestklage*, *Asbestentsorgung*, *Internetseite*, *Webseite*, *Inlineskater*, *Shoppingcenter*, *Zoobesuch*, *Lokschnitten*, *Aidskranke*, *Biobauer*, *Bioabfall*, *Ökopartei*.
- (3) In Gruppe 3 überwiegt in allen Zeitschnitten die bindestrichhaltige Schreibung (11 Types): Das sind alle Verbindungen mit *Tsunami* und *Ex* sowie die Wortverbindungen *Aids-Virus*, *Aids-Erreger*.
- (4) In einer letzten Gruppe 4 finden sich drei Wortverbindungen, die auch in den Zeitschnitten echte Variantenschreibungen sind (3 Types): *Bio-Äpfel/Bioäpfel*, *Bio-Bananen/Biobananen*, *Öko-Anbieter/Ökoanbieter*.

37 von 51 Types zeigen also in den Zeitschnitten eindeutig die Tendenz zur bindestrichlosen Zusammenschreibung. 3 weitere Types weisen eine echte Variantenschreibung auf. Lediglich 11 Types etablieren über die Zeitschnitte hinweg eine bindestrichhaltige Schreibung. Das sind nur 21,57% der hier untersuchten Komposita mit fremdwortartigen Erstgliedern.

Im Folgenden werden Beispiele aus der Gruppe 2 gezeigt, da diese hier besonders interessant sind. Aus einer Variantenschreibung heraus etabliert sich die bindestrichlose Zusammenschreibung. In einigen Fällen besteht die Variantenschreibung tatsächlich zu einem 50-50-Verhältnis (z.B. in *Asbest\_Entsorgung*, *Zoo\_Besuch*) und in anderen Fällen wird zunächst die bindestrichhaltige Schrei-

bung bevorzugt, bevor diese von der bindestrichlosen abgelöst wird (z.B. in *Aids\_Kranke*, *Web\_Seite*, *Shopping\_Center*).

Tab. 4: Internet\_Seite

	Internet-Seite (inkl. INTERNET-SEITE)		Internetseite (inkl. INTERNETSEITE)	
1990–1999	1.124	62,03%	688	37,96%
2000–2009	7.533	17,95%	34.435	82,05%
2010–2019	3.050	3,73%	78.965	96,27%
2020–	230	1,38%	16.468	98,62%
	11.910		130.556	

Tab. 5: Web\_Seite

	Web-Seite		Webseite (inkl. WEBSEITE)	
1990–1999	938	61,03%	599	38,97%
2000–2009	2.562	11,22%	20.275	88,78%
2010–2019	1.558	1,62%	94.398	98,38%
2020–	29	0,41%	7.060	99,59%
	5.087		122.332	

Tab. 6: Shopping\_Center

	Shopping-Center (inkl. SHOPPING-CENTER)		Shoppingcenter (inkl. SHOPPINGCENTER)	
1960–1969	9	81,82%	2	18,18%
1970–1979	20	95,24%	1	4,76%
1980–1989	31	83,78%	6	16,21%
1990–1999	609	76,70%	185	23,30%
2000–2009	1.522	46,83%	1.728	53,17%
2010–2019	1.397	29,22%	3.384	70,78%
2020–	131	22,28%	457	77,72%
	3.719		5.763	



Tab. 7: Inline\_Skater

	Inline-Skater (inkl. INLINE-SKATER)		Inlineskater (inkl. INLINESKATER)	
1990–1999	<b>2.503</b>	<b>79,89%</b>	630	20,11%
2000–2009	<b>2.955</b>	<b>63,81%</b>	1.676	36,19%
2010–2019	702	34,48%	<b>1.334</b>	<b>65,52%</b>
2020–	31	34,83%	<b>58</b>	<b>65,17%</b>
	<b>6.191</b>		3.698	

Tab. 8: Aids\_Kranke

	Aids-Kranke (inkl. AIDS-KRANKE)		Aidskranke (inkl. AIDSKRANKE)	
1980–1989	<b>282</b>	<b>84,18%</b>	53	15,82%
1990–1999	604	50,12%	601	49,88%
2000–2009	654	38,81%	<b>1.031</b>	<b>61,19%</b>
2010–2019	271	44,35%	340	55,65%
2020–	11	31,43%	<b>24</b>	<b>68,57%</b>
	1.822		<b>2.049</b>	

Tab. 9: Asbest\_Entsorgung

	Asbest-Entsorgung		Asbestentsorgung	
1980–1989	1	50%	1	50%
1990–1999	18	21,69%	<b>65</b>	<b>78,31%</b>
2000–2009	25	33,33%	<b>50</b>	<b>66,67%</b>
2010–2019	33	27,27%	<b>88</b>	<b>72,73%</b>
2020–	2	13,33%	<b>13</b>	<b>86,67%</b>
	79		<b>217</b>	

Tab. 10: Zoo\_Besuch

	Zoo-Besuch		Zoobesuch (inkl. ZOOBESUCH)	
1950–1959	1	50%	1	50%
1960–1969	2	50%	2	50%
1970–1979	1	25%	3	75%
1980–1989	1	20%	4	80%
1990–1999	24	15,29%	<b>133</b>	<b>84,71%</b>
2000–2009	77	12,32%	<b>548</b>	<b>87,68%</b>
2010–2019	80	10,60%	<b>675</b>	<b>89,40%</b>
2020–	11	6,51%	<b>158</b>	<b>93,49%</b>
	197		<b>1.524</b>	

Tab. 11: Bio\_Abfall

	Bio-Abfall		Bioabfall (inkl. BIOABFALL)	
1980–1989	2	66,67%	1	33,33%
1990–1999	104	11,98%	<b>764</b>	<b>88,02%</b>
2000–2009	69	8,87%	<b>709</b>	<b>91,13%</b>
2010–2019	56	3,98%	<b>1.350</b>	<b>96,02%</b>
2020–	10	3,16%	<b>306</b>	<b>96,84%</b>
	241		<b>3.130</b>	

Betrachtet man die zeitliche Abfolge aller Bindestrichschreibungen gegenüber der entsprechenden bindestrichlosen Schreibungen, lässt sich darüber hinaus Folgendes in den hier untersuchten Wortverbindungen mit fremdwortartigen Erstgliedern feststellen: Bei Erstnennung einer Wortverbindung im DeReKo taucht diese Wortverbindung sowohl in einer bindestrichhaltigen als auch in einer bindestrichlosen Schreibvariante auf. Jeder Zeitschnitt umfasst ein Jahrzehnt und vorsichtig interpretiert könnte man sagen, dass beide Varianten in etwa gleichzeitig – also im Verlauf des gleichen Jahrzehnts – im DeReKo auftauchen. Diese Beobachtung gilt für 30 der hier untersuchten Komposita. Bei weiteren 16 Komposita taucht die Bindestrichschreibung vor der Zusammenschreibung auf, also in einem früheren Zeitschnitt als die bindestrichlose Zusammenschreibung. Lediglich bei fünf Komposita zeigt sich erst eine bindestrichlose Schreibung und danach eine bindestrichhaltige Schreibung.

**Tab. 12:** Zeitliche Abfolge der Schreibvarianten mit Bindestrich (BIS) und ohne Bindestrich (ZUS) in allen Wortverbindungen

BIS-Schreibung und ZUS-Schreibung im gleichen Zeitschnitt	Asbest_Sanierung, Asbest_Entsorgung, Asbest_Schäden, Internet_Auktionenhaus, Internet_Seite, Web_Seite, Inline_Skater, Inline_Skates, Inline_Skating, Shopping_Center, Gala_Abend, Tsunami_Opfer, Zoo_Direktor, Zoo_Tiere, Zoo_Besuch, Lok_Führer, Lok_Schuppen, Aids_Kranke, Aids_Erreger, Aids_Virus, Profi_Boxer, Profi_Boxerin, Ex_Ministerpräsident, Bio_Produkte, Bio_Tonne, Bio_Abfall, Öko_Steuer, Öko_Strom, Öko_Anbieter
BIS-Schreibung vor ZUS-Schreibung	Asbest_Klage, Pop_Song, Rock_Song, Tsunami_Katastrophe, Profi_Fußball, Profi_Karriere, Profi_SpielerIn, Ex_Mann, Ex_Manager, Ex_Bundestrainer, Ex_Profi, Ex_Bundesligist, Bio_Bauer, Bio_Äpfel, Öko_Partei, Öko_Bauer
ZUS-Schreibung vor BIS-Schreibung	Asbest_Belastung, Zoo_Besucher, Profi_Spieler, Ex_Bundeskanzler, Bio_Ethik

Es lässt sich also vorsichtig folgern, dass in den allermeisten Fällen (hier 90 Prozent der untersuchten Komposita) die Bindestrichschreibung nicht in etablierte Zusammenschreibungen oder morphologische Bildungen hineintritt, sondern dass sie mit dem ersten Auftauchen der Bildung selbst im DeReKo als Schreibvariante genutzt wird und dann entweder abgebaut wird oder auf einem niedrigen Level stagniert. Ansonsten ergibt sich ein eher uneinheitliches Bild, denn weder die Erstglieder noch die Zweitglieder lassen Vorhersagen zu, in welcher zeitlichen Reihenfolge die Schreibvarianten im DeReKo auftauchen.

## 2.2 Hypothese 2: Die Ablösung der Bindestrichschreibung als Integrationsmarker für das fremdwortartige Erstglied

Der Abbau der Bindestrichschreibung kann dann als Integrationsmarker für das fremdwortartige Erstglied gewertet werden, wenn der entsprechende Erstbestandteil erstmals mit dem hier untersuchten Kompositum im DeReKo auftritt. Daher wurde zusätzlich zur Verteilung der Schreibvarianten in den Zeitschnitten der bindestrichauslösende Wortbestandteil selbst überprüft, und zwar immer dann, wenn dies im DeReKo auch suchbar war. Aufgrund der Annotation des Korpus und der Suchanfragesyntax in COSMAS lassen sich *ex-*, *bio-*, *öko-* als wortfähige Einheiten im DeReKo nicht systematisch suchen. Es ist vielleicht Zufall, dass unter den 16 Komposita, in denen die bindestrichhaltige Schreibung vor der

bindestrichlosen Schreibung auftaucht, Verbindungen mit *ex-*, *bio-* und *öko-* zu finden sind.

Im Ergebnis hat sich gezeigt, dass die bindestrichauslösenden Wortbestandteile als wortfähige Einheiten entweder vor den Wortverbindungen oder zumindest im gleichen Zeitschnitt im DeReKo nachweisbar sind. Auch übersteigt die Tokenzahl der Wortbestandteile alleine in den meisten Fällen die Tokenzahl der hier untersuchten Wortverbindung. Dies sei an ein paar Beispielen verdeutlicht:

**Tab. 13:** *Asbest* – Erstbelege und Gesamttoken

Type	Erstbeleg	Gesamttoken
Asbest	1940	14.384
Asbest_Belastung	1970	690
Asbest_Sanierung	1980	1.791
Asbest_Entsorgung	1980	296
Asbest_Schäden	1980	171
Asbest_Klage	1990	41

**Tab. 14:** *Aids* – Erstbelege und Gesamttoken

Type	Erstbeleg	Gesamttoken
Aids	1980	75.900
Aids_Kranke	1980	3.871
Aids_Erreger	1980	1.929
Aids_Virus	1980	4.213

Die Schreibung von <Aids> wirft sicherlich einige Fragen auf. *Aids* als Benennung der Krankheit taucht im DeReKo als Erstbeleg ab 1980 auf. Frühere Einträge aus den 1950er Jahren deuten auf den englischen Plural *aids* hin, also die Hilfen oder Gehilfen; das ist hier aber nicht gemeint. Ab 1980 taucht *Aids* in beiden Schreibvarianten auf: <AIDS> und <Aids>, wobei sich 7.308 Token auf die Schreibung <AIDS> vereinen und 75.900 auf die Schreibung <Aids>. Mit Beginn der Nennung ist die Normalschreibung <Aids> die präferierte. Das Gleiche gilt auch für <AIDS-Erreger> mit 45 Token in allen Zeitschnitten, <AIDS-Virus> mit 156 Token in allen Zeitschnitten und <AIDS-Kranke> mit 142 Token in allen Zeitschnitten.

**Tab. 15:** *Zoo* – Erstbelege und Gesamttoken

Type	Erstbeleg	Gesamttoken
Zoo	1940	146.939
Zoo_Direktor	1940	7.922
Zoo_Besucher	1940	2.483
Zoo_Tiere	1950	2.413
Zoo_Besuch	1950	1.721

Für die Erstbelege für *Zoo*- und *Profi*- sehen die Einträge auf den ersten Blick aus wie ein Widerspruch zu dem Versuch, dass die untersuchten Wortverbindungen in der 2. Hälfte des 20. Jahrhunderts im Deutschen auftauchen und vorkommen sollen (vgl. dazu Abschnitt 1.2 zur Methodik). Im DeReKo taucht *Zoo* alleine ab dem Zeitschnitt 1940–1949 auf (genauer: 25.01.1947 in einem *Spiegel*-Artikel). Es ist aber vermutlich schon länger im Deutschen. Eine Suche im DTA zeigt gar keine Treffer für *Zoo* oder *Zoologischer Garten*. Ab 1605 findet sich aber der Begriff *Thiergarten* ([www.deutschestextarchiv.de](http://www.deutschestextarchiv.de)). Im DWDS-Kernkorpus 1900–1999 ist *Zoo* bzw. *Zoologischer Garten* ab 1900 nachweisbar. Ab 1950 übersteigt *Zoo* im DeReKo die Tokenzahl von 100; *Zoo\_Direktor*, *Zoo\_Besucher*, *Zoo\_Tiere*, *Zoo\_Besuch* sind zwar ab 1940 bzw. 1950 im DeReKo nachweisbar, Tokenzahlen über 100 erreichen diese Wortverbindungen aber erst ab 1990.

**Tab. 16:** *Profi* – Erstbelege und Gesamttoken

Type	Erstbeleg	Gesamttoken
Profi	1940	112.897
Profi_Fußball	1940	24.359
Profi_Karriere	1940	15.492
Profi_Boxer	1950	6.184
Profi_Spieler	1960	3.288
Profi_Spielerin	1970	1.066
Profi_Boxerin	1990	614

*Profi* als Kurzwort kommt im DWDS-Kernkorpus 1900–1999 im Jahr 1930 erstmals vor, der zweite Beleg im Zusammenhang mit Boxsport im Jahr 1957. Das passt also durchaus zu den Daten im DeReKo. Im DeReKo übersteigt *Profi* die Tokenzahl von

100 ab 1960, *Profī\_Fußball*, *Profī\_Karriere*, *Profī\_Boxer* und *Profī\_Spielerin* ab 1990, *Profī\_Spielerin*, *Profī\_Boxerin* ab 2000.

Blendet man alle Nebenschauplätze aus, die sich für die jeweiligen Wortverbindungen und ihre Erstglieder ergeben, bleibt der Befund stehen: Die fremden Erstglieder weisen in allen Bedingungen (Fremdwort, fremdes Kurzwort, fremdes Präfix/Konfix) deutlich höhere Tokenzahlen auf als die Wortverbindungen selbst und sie sind entweder im gleichen Zeitschnitt wie die Wortverbindungen im DeReKo nachweisbar oder sogar deutlich davor. Damit lässt sich für die hier analysierten Wortverbindungen nicht nachweisen, dass der Übergang von der bindestrichhaltigen Schreibung zu einer bindestrichlosen Schreibung ein Integrationsmarker für die fremden Erstglieder ist.

Im Anschluss müssten nun die Erstbelege der Fremdwörter/Kurzwörter/fremden Präfixe/Konfixe vertiefend untersucht werden, um zu analysieren, ob diese in anderen Komposita erstmals im DeReKo auftauchen und erst danach eine Wortfähigkeit entwickeln. Oder ob sie, wie es sich hier andeutet, mit ihrem ersten Auftauchen im DeReKo wortfähig sind.

## 2.3 Hypothese 3: Die Zusammenschreibung als Zeichen für die Etablierung der Wortverbindung im Deutschen

Die dritte Hypothese nimmt genau den gezeigten Befund aus 2.2 auf: Kommt das bindestrichauslösende Fremdwort selbst schon vor der hier untersuchten Wortbildung im DeReKo vor, so kann der Wechsel der bindestrichhaltigen Schreibung zur bindestrichlosen Schreibung nicht als Integration des Fremdworts gewertet werden. Dieses Fremdwort ist schon im Deutschen vorhanden und wird u.U. wortfähig gebraucht. Eine andere Interpretation legt dann nahe, dass es sich bei dem Wechsel von der bindestrichhaltigen zur bindestrichlosen Zusammenschreibung um eine Art Etablierung des Kompositums selbst handelt, also um die Aufnahme eines Wortbildungsprodukts in den Wortschatz des Deutschen.

In Abschnitt 1.3 wurden die drei Dimensionen der Etablierung nach Schmid 2008 vorgestellt: die strukturelle, die sozio-pragmatische und die kognitive. Mit Blick auf die hier analysierten Wortverbindungen sollte die strukturelle Perspektive unstrittig sein: Morphologisch lassen sich alle hier untersuchten Wortverbindungen in vorhandene Wortbildungsmuster des Deutschen einordnen. Es handelt sich im Großen und Ganzen um Komposita oder sie lassen sich als solche reinterpretieren (*Asbest\_Belastung* vs. *Inline\_Skating*, *Bio\_Tonne*). Inwiefern die Bildungen idiomatisiert oder demotiviert sind und ihre morphologische Transparenz verlieren, müsste für jede Wortverbindung einzeln geprüft werden. Auf den ersten Blick sind alle Bildungen zumindest morphologisch transparent und keines

idiomatisiert, in dem Sinne, dass seine Gesamtbedeutung nicht aus der Summe der Einzelbedeutungen zusammengesetzt werden könnte (vgl. dazu Fleischer & Barz 2012: 44–47). Graphematisch werden morphologische Wörter zusammengeschieden und es bedarf spezieller Gründe, warum in die unmarkierte Zusammenschreibung ein Wortzeichen, hier der Bindestrich, interveniert. In Abschnitt 2.1 wurde gezeigt, dass lediglich 21,57% der hier analysierten Wortbildungen über alle Zeitschnitte hinweg eine Bindestrichschreibung ausprägen. 78,43% bilden also im Laufe ihres Gebrauchs eine unmarkierte bindestrichlose Schreibung aus und fügen sich so nicht nur morphologisch, sondern auch graphematisch in das grammatische System des Deutschen ein. Man könnte also schlussfolgern, dass die Wortverbindungen strukturell etabliert sind (vgl. Abschnitt 1.3, Tab. 2).

Was können Gründe für die dauerhafte Herausbildung einer Bindestrichschreibung in einem morphologischen Wort sein? Zu den 11 Types, die eine bindestrichhaltige Schreibung als Hauptvariante aufweisen, gehören: *Aids-Erreger*, *Aids-Virus*, *Tsunami-Katastrophe*, *Tsunami-Opfer* sowie alle sieben Wortverbindungen mit *Ex-*, also *Ex-Mann*, *Ex-Manager*, *Ex-Bundeskanzler*, *Ex-Ministerpräsident*, *Ex-Bundesligist*, *Ex-Profi*, *Ex-Bundestrainer*. Warum sich alle *ex*-Verbindungen anders verhalten als die anderen Wortverbindungen mit Konfix (hier *bio-* und *öko-*), bleibt spekulativ. *Aids* und *Tsunami* hingegen sind möglicherweise durch ihnen gemeinsame Beobachtung erklärbar: Beide können als Eigennamen interpretiert werden. *Tsunami* als Name eines bestimmten Wetterphänomens und *Aids* als Name einer Krankheit. Beide Einheiten können beiden Kriterien – Fremdwort und Eigenname – zugeschlagen werden. Und speziell zu *Aids*: Es ist nicht nur seinem phonologischen Verhalten (und seiner Herkunft) nach ein Fremdwort sowie seiner Bedeutung nach die Benennung einer Krankheit, es ist darüber hinaus morphologisch auch ein Kurzwort. Syntaktisch wird es i.d.R. ohne Artikel gebraucht. Im DeReKo gibt es nur sehr wenige Vorkommen mit Artikel bzw. Artikelpronomen: *das Aids*, *dieses Aids*, *mein/dein/sein/ihr Aids*, und zwar seit Mitte der 80er Jahre.<sup>17</sup> Hier wurde es bislang als fremdes Kurzwort ein-

17 1. *Es gibt das afrikanische Aids, das immer noch zwangsläufig tödlich ist, es gibt das westeuropäisch-amerikanische Aids, das seit Einführung von Haart im Jahre 1996 nurmehr eine chronische Krankheit ist. Es gibt das Aids der Frauen, deren Ansteckungsgefahr angeblich höher ist als die der Männer. Es gibt das Aids der Schwulen, das mittlerweile eine Art Eigenleben zu führen scheint, außerhalb global engagierter Diskurse.* (T05/DEZ.00315 die tageszeitung, 02.12.2005, 14; Das gibt zu denken)

2. *Man wird schon vorher ein Stück weit beerdigt, aber ich kämpfe noch für mein Leben. Ich will dieses Aids nicht akzeptieren, bis zur letzten Stunde nicht! Ich fühle mich wie ein verwundetes Tier und brülle immer lauter.*“ (Z93/MAR.00084 Die Zeit, 05.03.1993, 19; Man wird schon vorher ein Stückchen beerdigt)

geordnet, immer mit dem Hinweis darauf, dass es eben gleichzeitig auch ein Eigenname ist.

Die eigennamenhaltigen Wortverbindungen in dem für diese Untersuchung zusammengestellten Korpus (vgl. Abschnitt 1.2) werden in diesem Aufsatz nicht gezeigt, es hat sich bei diesen aber gezeigt, dass die bindestichhaltigen Schreibungen dort z.T. sehr viel präsenter sind als die bindestrichlosen Schreibungen.

**Tab. 17:** Ausgewählte Schreibungen mit Eigennamen im Erstglied

Type	Token mit Bindestrich	Token ohne Bindestrich	Gesamt
Telekom_Fahrer	1.190 (94,67%)	67 (5,33%)	1.257
Airbus_Chef	4.184 (99,78%)	9 (0,22%)	4.193
Bush_Regierung	10.240 (99,71%)	30 (0,29%)	10.270
Merkel_Regierung	824 (99,4%)	5 (0,6%)	829
Alzheimer_Krankheit	2.627 (70,98%)	1.074 (29,02%)	3.701
Alzheimer_Patient	132 (75,43%)	43 (24,57%)	175
Alzheimer_Diagnose	108 (91,53%)	10 (8,47%)	118

Hier sind Beispiele gezeigt, die verschiedene Typen von Eigennamen im Erstglied enthalten und in denen über alle Zeitschnitte hinweg die bindestrichhaltige Schreibung die bevorzugte Variante ist. Es wäre also eine (neue) Hypothese, dass Eigennamen deutlich zugänglicher für Bindestrichschreibungen sind als Fremdwörter und wenn eine Einheit sowohl ein Eigenname als auch ein Fremdwort ist, sich diese Einheit möglicherweise eher wie ein Eigenname als wie ein Fremdwort verhält. Dies passt zu den Analysen des grammatischen Verhaltens von Eigennamen im Deutschen von Nübling (2012). Sie untersucht die zunehmende Deflexion der Eigennamen und interpretiert diese als Stärkung der Autonomie des Wortkörpers (vgl. auch Nübling 2014, wo sie den zunehmenden Apostrophgebrauch im Deutschen ebenfalls als wortkörperschonend interpretiert). Eine Bindestrichschreibung in einem Kompositum, das einen Eigennamen enthält, könnte ebenfalls den Wortkörper des Eigennamens schonen.

Mit Schmid (2008: 3) gehört die Streuung einer Wortverbindung über verschiedene Texte und damit Sprecher zur sozio-pragmatischen Perspektive der

---

3. Frank, 37, *Programmgestalter bei NBC in Hollywood, unter dessen linkem Auge eine erbsengroße Geschwulst blüht, lebt noch immer im Exil. Sein Aids ist "Hautkrebs". Er will niemanden erschrecken.* (S85/OKT.00354 Der Spiegel, 28.10.1985, 200; "Vergnügt euch, aber seht euch vor")



Etablierung neuer Wortverbindungen. Keine der hier gezeigten Wortverbindungen ist auf einen Text oder Autor beschränkt. Ganz im Gegenteil: Mit ihrer ersten Verwendung sind die Wortverbindungen und ihre Schreibvarianten sofort auf unterschiedliche Texte im DeReKo verteilt. Auch hier ließe sich also schlussfolgern, dass die hier gezeigten Wortverbindungen sozio-strukturell etabliert sind (vgl. Abschnitt 1.3, Tab. 2).

**Tab. 18:** *Asbest\_Belastung* – Streuung in den Texten

<b>Asbest-Belastung</b>		<b>Asbestbelastung</b>	
1970–1979	---	1970–1979	1 in 1
1980–1989	---	1980–1989	4 in 4
1990–1999	12 in 12	1990–1999	97 in 87
2000–2009	28 in 26	2000–2009	213 in 189
2010–2019	13 in 13	2010–2019	281 in 226
2020–	1 in 1	2020–	40 in 36
<b>54 Treffer in 52 Texten</b>		<b>636 Treffer in 543 Texten</b>	

**Tab. 19:** *Aids\_Virus* – Streuung in den Texten

<b>Aids-Virus (inkl. AIDS-VIRUS)</b>		<b>Aidsvirus (inkl. AIDSVIRUS)</b>	
1980–1989	648 in 339	1980–1989	29 in 18
1990–1999	1.178 in 981	1990–1999	310 in 249
2000–2009	1.103 in 959	2000–2009	421 in 357
2010–2019	353 in 311	2010–2019	120 in 107
2020–2029	27 in 26	2020–2029	24 in 22
<b>3.309 Treffer in 2.616 Texten</b>		<b>904 Treffer in 753 Texten</b>	

Tab. 20: *Ex\_Mann* – Streuung in den Texten

Ex-Mann (inkl. EX-MANN)		Exmann (inkl. EXMANN)	
1960–1969	2 in 2	1960–1969	---
1970–1979	5 in 5	1970–1979	1 in 1
1980–1989	44 in 40	1980–1989	5 in 3
1990–1999	1.269 in 1.088	1990–1999	188 in 172
2000–2009	3.770 in 3.139	2000–2009	818 in 708
2010–2019	5.249 in 4.236	2010–2019	1.064 in 905
2020–2029	791 in 664	2020–2029	66 in 53
11.130 Treffer in 9.174 Texten		2.142 Treffer in 1.842 Texten	

Bleibt als letztes die kognitive Perspektive und diese ist die am schwierigsten greifbare: Grundsätzlich lässt sich die Frage stellen, wie morphologisch einfache und morphologisch komplexe Wörter im mentalen Lexikon gespeichert und beim Lesen oder Hören verarbeitet werden (vgl. dazu u.a. Aitchison 1987/1997 zitiert nach Hoffmann 2019: 634–647 und neuer ein Zeitschriftenband zu morphologisch komplexen Wörtern Leminen et al. (Hrsg.) 2016). Und es ist die Frage, ob sich mit den hier gezeigten Daten Rückschlüsse auf die Repräsentation von Komposita mit Fremdwortanteil im mentalen Lexikon nativer Sprecher des Deutschen ziehen lassen. Fest steht: Keine der hier gezeigten Wortbildungen wurde in einem psycholinguistischen Verarbeitungsexperiment untersucht.

Schmid (2008: 10–13) referiert den einschlägigen Forschungsstand (aus der experimentellen Psycholinguistik) und fasst Folgendes für neugebildete Komposita zusammen. Einfluss auf die Verarbeitung eines neugebildeten Kompositums haben folgende Faktoren:

1. die Anzahl der Kotext- und Kontextinformationen,
2. die Transparenz der morphologischen Struktur,
3. die Vorkommenshäufigkeit der Konstituenten des Kompositums sowie
4. die Vorkommenshäufigkeit der ersten Konstituente,
5. die Größe der (morphologischen) Wortfamilie der Konstituenten,
6. die semantische Relation zwischen den beiden Konstituenten.

Um das Entrenchment, also den Eingang eines neugebildeten Wortes ins mentale Lexikon zu erfassen, ist die Frequenz des Wortes selbst und seiner Konstituenten von Interesse (Schmid 2008: 19–21). Anschließend an Schmid (2008) führen Stefanowitsch & Flach (2017) aus, wie Entrenchment mithilfe der Vorkommenshäufigkeit korpusbasiert gemessen werden kann. Zunächst erläutern sie sehr umsich-

tig, warum echte Sprecherdaten überhaupt als empirische Basis für die Frage nach einem kognitiven Konzept wie Entrenchment genutzt werden dürfen. Sie zeigen dann anhand von (englischen) ADJ-Noun-Verbindungen und (englischen) Phrasen verschiedene Verfahren, wie die Vorkommenshäufigkeit der gesamten Bildung mit den Vorkommenshäufigkeiten der Konstituenten miteinander verrechnet werden können. Darüber hinaus zeigen sie, wie schwierig die Interpretation der so errechneten Daten ist. Es stellt sich nun die Frage, ab welcher Vorkommenshäufigkeit bzw. ab welchem errechneten Wert eine Wortverbindung als im mentalen Lexikon gespeichert bzw. in einer Sprache etabliert gilt. Diese Frage kann hier kaum beantwortet werden und diese Frage beantworten auch Stefanowitsch & Flach (2017) (und auch Schmid 2008) nicht explizit, denn die errechneten Maße bringen immer nur die untersuchten Wortverbindungen zueinander in Relation, aber nicht die Wortverbindungen zu allen Wörtern und Wortverbindungen im zugrundeliegenden Korpus (bei Stefanowitsch & Flach 2017 verschiedene Referenzkorpora für das Britische und Amerikanische Englisch, hier für diesen Aufsatz das DeReKo).

Was lässt sich dennoch aus den hier gezeigten Daten über die (kognitive) Etablierung einer Wortverbindung schließen? Zumindest die Vorkommenshäufigkeit der Wortverbindung selbst und der beiden Konstituenten ist offensichtlich ausschlaggebend für Entrenchment. Für die hier gezeigten Wortverbindungen wurde die Tokenzahl der Verbindung selbst und des Erstglieds erfasst. Das Zweitglied müsste nun ergänzt werden, um die Berechnungen nach Stefanowitsch & Flach (2017) durchzuführen. Im Moment lässt sich festhalten: Alle hier gezeigten Wortverbindungen zeigen sehr unterschiedliche Vorkommenshäufigkeiten im DeReKo; es war eine Voraussetzung bei der Datenauswahl, dass die Wortverbindungen in einem gewissen bzw. ausreichendem Maße im DeReKo vorkommen. Es sollten ausdrücklich keine Hapax Legomena analysiert werden (vgl. dazu Abschnitt 1.2). Ob nun auch aus psycholinguistischer Perspektive die Wortverbindungen als etabliert bzw. aufgenommen ins mentale Lexikon gelten können, soll offengelassen werden. Es reicht an dieser Stelle vielleicht zu zeigen, dass die anderen beiden Perspektiven (strukturell und sozio-pragmatisch) eine Etablierung der Wortverbindungen im Deutschen nahelegen.

### 3 Schlussbetrachtungen

In dieser Untersuchung stehen Wortverbindungen bzw. Komposita im Fokus, die einen Wortbestandteil enthalten, der eine Bindestrichschreibung auslösen kann. In Buchmann (2015: 236–253) wurden diese Schreibungen als fakultative Binde-

strichschreibungen analysiert und ausführlich beschrieben. Sie enthalten ein Fremdwort, ein Kurzwort, einen Eigennamen oder einen Wortbestandteil, der einen ungeklärten morphologischen Status hat. Sie stammen alle aus der Tageszeitung *Mannheimer Morgen*. Auf Grundlage dieses Korpus wurde hier ein zweites Korpus mit 91 Komposita bzw. morphologisch komplexen Wörtern gebildet. Nimmt man die eigennamenhaltigen Bindestrichschreibungen heraus, bleiben 51 übrig. Die verbliebenen Wörter enthalten Fremdwörter, Kurzwörter, die als Langform ebenfalls einen Fremdwortstatus aufweisen, sowie Einheiten mit einem ungeklärten morphologischen Status. Letztere werden in der Forschungsliteratur als fremde Präfixe oder Konfixe beschrieben, was ebenfalls dazu führt, dass sie in den Fremdwortbereich fallen. Letztendlich ergibt sich durch den Ausschluss der Eigennamen eine relativ homogene Gruppe von Wörtern.

Es wurde zum einen erfasst, wie viele Token im DeReKo (Version 2022-I) in den einzelnen Jahrzehnten mit Bindestrich und wie viele ohne Bindestrich auftauchen. Zum anderen wurden die Token des Erstglieds selbst erfasst. Die Untersuchung hat ein paar interessante und überraschend einheitliche Ergebnisse hervorgebracht.

Es wurden drei Hypothesen formuliert, die nacheinander bearbeitet wurden.

1. Mit zunehmendem Gebrauch einer bindestrichhaltigen Wortverbindung nimmt die Bindestrichschreibung ab. Damit verbunden war auch die Frage, wann der Bindestrich erstmals in einer Wortverbindung auftaucht.
2. Die Herausbildung einer bindestrichlosen Schreibung ist ein Zeichen für die Integration des Fremdworts im Erstglied der Wortverbindung.
3. Die Herausbildung einer bindestrichlosen Schreibung ist ein Zeichen für die Etablierung der Wortverbindung im Deutschen bzw. im mentalen Lexikon der Sprecher.

Die erste Hypothese konnte verifiziert werden, darüber hinaus hat sich gezeigt, dass der Bindestrich nicht in vorhandene morphologische Bildungen tritt. I.d.R. lässt sich mit den ersten Vorkommen der Wortverbindungen eine Variation hinsichtlich der Schreibung mit oder ohne Bindestrich beobachten, aus dieser Variation entwickelt sich dann i.d.R. die bindestrichlose Zusammenschreibung. Die Wortverbindungen mit Fremdwort als Erstglied gliedern sich also in das graphematische System der Zusammen- und Getrennschreibung ein und entwickeln eine unmarkierte Zusammenschreibung.

Die zweite Hypothese konnte falsifiziert werden. Zumindest für die hier untersuchten Wortverbindungen ließ sich keine Integration des Fremdworts nachweisen. Dieses ist i.d.R. schon vorher oder zumindest im gleichen Zeitschnitt im DeReKo wortfähig nachweisbar und es weist über alle Zeitschnitte hinweg deut-

lich größere Tokenzahlen auf als die Wortverbindungen selbst. Hier müsste nun geprüft werden, wann und ggf. in welchen Wortverbindungen die Fremdwörter ins Deutsche kommen bzw. im Deutschen gebildet werden, um die Integration der Fremdwörter nachzuvollziehen.

Für die dritte Hypothese konnten Indizien gesammelt werden: Zumindest für zwei der drei Dimensionen der Etablierung von morphologisch komplexen Wörtern konnte gezeigt werden, dass die Wortverbindungen etabliert sind – nämlich die strukturelle und die sozio-pragmatische Perspektive. Strukturell gliedern sich die Wortverbindungen in morphologische Wortbildungsprozesse des Deutschen ein, sie nehmen darüber hinaus die unmarkierte Zusammenschreibung des Deutschen an. Sie werden in vielen verschiedenen Texten genutzt und weisen damit eine breite Streuung über das Referenzkorpus des Deutschen auf.

Die hier untersuchten Komposita wurden relativ kontrolliert zusammengetragen bzw. ausgewählt und weisen dennoch eine gewisse Zufälligkeit auf. Es handelt sich nicht um einen repräsentativen Querschnitt und daher können keine Rückschlüsse auf alle Bindestrichschreibungen im Deutschen des 20. und 21. Jahrhunderts gezogen werden. Trotzdem haben sich deutlich erkennbare und im Grunde auch überraschend eindeutige Tendenzen gezeigt, die mit einem größeren Korpus überprüft werden sollten.

## Literaturverzeichnis und Korpora

- Bernabei, Dante (2003): *Der Bindestrich. Ein Vorschlag zur Systematisierung*. Frankfurt a.M. u.a.: Peter Lang.
- Bredel, Ursula (2008): *Das Interpunktionssystem des Deutschen. Ein kompositionelles System zur Online-Steuerung des Lesens*. Tübingen: Niemeyer.
- Bredel, Ursula (2020): *Interpunktion*. 2. Aufl. Heidelberg: Winter.
- Buchmann, Franziska (2015): *Die Wortzeichen im Deutschen*. Heidelberg: Winter.
- Eisenberg, Peter (2018): *Das Fremdwort im Deutschen*. 3., überarbeitete und erweiterte Aufl. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/les.2011.027>
- Eisenberg, Peter (2020): *Grundriss der deutschen Grammatik. Der Satz*. Unter Mitarbeit von Rolf Schöneich. Stuttgart, Weimar: Metzler.
- Fleischer, Wolfgang & Irmhild Barz (2012): *Wortbildung der deutschen Gegenwartssprache*. 4. Aufl. Völlig neu bearbeitet von Irmhild Barz unter Mitarbeit von Marianne Schröder. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110256659>
- Fuhrhop, Nanna (2007): *Zwischen Wort und Syntagma. Zur grammatischen Fundierung der Getrennt- und Zusammenschreibung*. Tübingen: Niemeyer. <https://doi.org/10.1515/zrs.2010.041>
- Fuhrhop, Nanna (2008): Das graphematische Wort (im Deutschen): Eine erste Annäherung. *Zeitschrift für Sprachwissenschaft* 27 (2), 189–228. <https://doi.org/10.1515/ZFSW.2008.010>
- Fuhrhop, Nanna (2020): *Orthografie*. 5. Aufl. Heidelberg: Winter.
- Gallmann, Peter (1985): *Graphische Elemente der geschriebenen Sprache. Grundlagen für eine Reform der Orthographie*. Tübingen: Niemeyer. <https://doi.org/10.1515/9783111630380>
- Gallmann, Peter (1989): Syngropheme an und in Wortformen. Bindestrich und Apostroph im Deutschen. In Peter Eisenberg & Hartmut Günther (Hrsg.), *Schriftsystem und Orthographie*. Tübingen: Niemeyer, 85–110. <https://doi.org/10.1515/9783111372266.85>
- Gallmann, Peter (1996): Interpunktion (Syngropheme). In Hartmut Günther & Otto Ludwig (Hrsg.), *Schrift und Schriftlichkeit. Writing and its use*, 1456–1467. Berlin, New York: De Gruyter. <https://doi.org/10.1515/9783110147445.2.9.1456>
- Geilfuß-Wolfgang, Jochen (2007): *Worttrennung am Zeilenende. Über die deutschen Worttrennungsregeln, ihr Erlernen in der Grundschule und das Lesen getrennter Wörter*. Tübingen: Niemeyer. <https://doi.org/10.1515/zrs.2010.043>
- Hoffmann, Ludger (Hrsg.) (2019): *Sprachwissenschaft. Ein Reader*. 4. aktualisierte und erweiterte Aufl. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110588972>
- Jacobs, Joachim (2005): *Spatien. Zum System der Getrennt- und Zusammenschreibung im heutigen Deutsch*. Berlin, New York: De Gruyter. <https://doi.org/10.1515/9783110919295>
- Kopf, Kristin (2018): Fugenelement und Bindestrich in der Compositions-Fuge. In Nanna Fuhrhop, Renata Szczepaniak & Karsten Schmidt (Hrsg.), *Sichtbare und hörbare Morphologie*, 177–204. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110528978-007>
- Leminen Alina, Harald Clahsen, Minna Lehtonen & Mirjana Bozic (Hrsg.) (2016): Morphologically complex Words in the Mind/Brain. *Frontiers in Human Neuroscience* 10. <https://doi.org/10.3389/978-2-88919-803-0>
- Neef, Martin (2008): Worttrennung am Zeilenende. Überlegungen zur Bewertung und Analyse orthographischer Daten. *Zeitschrift für Germanistische Linguistik* 35, 283–314. <https://doi.org/10.1515/zgl.2007.022>

- Nübling, Damaris (2012): Auf dem Wege zu Nicht-Flektierbaren: Die Deflexion der deutschen Eigennamen diachron und synchron. In Björn Rothstein (Hrsg.), *Nicht-flektierende Wortarten*, 224–246. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110276619.224>
- Nübling, Damaris (2014): Sprachverfall? Sprachliche Evolution am Beispiel des diachronen Funktionszuwachses des Apostrophs im Deutschen. In Albrecht Plewnia & Andreas Witt (Hrsg.), *Sprachverfall? Dynamik – Wandel – Variation*. Jahrbuch 2013 des Instituts für deutsche Sprache, 99–123. Berlin, Boston: De Gruyter. <https://doi.org/10.1515/9783110343007.99>
- Nübling, Damaris, Fabian Fahlbusch & Rita Heuser (2015): *Namen. Eine Einführung in die Onomastik*. 2., überarbeitete und erweiterte Aufl. Tübingen: Narr Francke Attempto.
- Satkauskaitė, Danguolė (2008): Die Rolle des Bindestrichs bei der Konventionalisierung von Zusammensetzungen. *Kalbotyra* 59 (3), 153–260. <https://doi.org/10.15388/Klbt.2008.7613>
- Schmid, Hans-Jörg (2008): New Words in the Mind: Concept-formation and Entrenchment of Neologisms. *Anglia. Journal of English Philology* 126 (1), 1–36. <https://doi.org/10.1515/angl.2008.002>
- Smith, George (2000): Word Remnants and Coordination. In Rolf Zhieroff, Matthias Tamrat, Nanna Fuhrhop & Oliver Teuber (Hrsg.), *Deutsche Grammatik in Theorie und Praxis*, 57–68. Tübingen: Niemeyer. <https://doi.org/10.1515/9783110933932.57>
- Stefanowitsch, Anatol & Susanne Flach (2017): The Corpus-Based Perspective on Entrenchment. In Hans-Jörg Schmid (Hrsg.), *Entrenchment and the Psychology of Language Learning. How We Reorganize and Adapt Linguistic Knowledge*, 101–127. Berlin: De Gruyter.
- Ziem, Alexander & Björn Fritsche (2018): Von der Sprache zur (Konstruktion von) Wirklichkeit. In Ekkehard Felder & Andreas Gardt (Hrsg.), *Wirklichkeit oder Konstruktion? Sprachtheoretische und interdisziplinäre Aspekte einer brisanten Alternative*, 243–276. Berlin u.a.: De Gruyter. <https://doi.org/10.1515/9783110563436-012>

## Verwendete Korpora

*DeReKo 2022-I* (2022): <https://cosmas2.ids-mannheim.de/cosmas2-web/>

*Mannheimer Morgen* (1995–2008): <https://cosmas2.ids-mannheim.de/cosmas2-web/> [www.dwds.de](http://www.dwds.de)

Kernkorpus des 20. Jahrhunderts: [www.deutschestextarchiv.de](http://www.deutschestextarchiv.de)

Ohne Autor: Informationen zum Aufbau des DeReKo am IDS Mannheim <https://www.ids-mannheim.de/digspra/kl/projekte/korpora/archiv-1/> (letzter Zugriff 30.09.2022).

