

Sascha Wolfer / Alexander Kopenig / Marc Kupietz / Carolin Müller-Spitzer

WENIGER IST MEHR? EINE ANALYSE ZUR „NEIGUNG ZUM HINZUFÜGEN“ IM DEUTSCHEN ANHAND DES NEUEN HÄUFIGKEITSDATENSATZES DeReKoGram

Stellen Sie sich einmal vor, Sie würden gebeten, die Situation an Ihrer Arbeitsstelle zu verbessern. Welche Maßnahmen würden Ihnen einfallen? Oder Sie überarbeiten einen Text und möchten die Argumentation noch überzeugender gestalten: Wie würden Sie vorgehen? Oder etwas anders gefragt: Würden Sie eher etwas zur Arbeitssituation hinzufügen (zum Beispiel Pausenräume oder Kaffeemaschinen) oder etwas entfernen (zum Beispiel Meetings)? Oder wenn es um die Optimierung des Textes geht: Würden Sie eher Textmaterial hinzufügen oder entfernen?

Adams et al. (2021) zeigen in ihrem Nature-Artikel mit dem sprechenden Titel „People systematically overlook subtractive changes“, dass Menschen dazu tendieren, „additive Transformationen“ von Situationen, Objekten oder Ideen vorzunehmen, wenn sie diese verbessern sollen. In einem der berichteten Experimente gaben sie den Teilnehmenden die Aufgabe, eine Minigolfbahn zu verbessern. Dabei wurde ihnen eine Skizze der Bahn vorgelegt. Die Forschenden haben darauf geachtet, dass sowohl das Hinzufügen als auch das Entfernen von Elementen zu Verbesserungen der Minigolfanlage führen würden. So konnten sie zum Beispiel – um die Bahn etwas einfacher zu machen – einen Sandbunker (eine „Falle“, aus der man den Ball nur schlecht wieder herausspielen kann) entfernen.

MENSCHEN ÜBERSEHEN SUBTRAHIERENDE VERÄNDERUNGEN

Die Teilnehmenden sollten alle ihre Ideen auflisten, wie man die Bahn verbessern könnte. Dabei wurden sie auch darum gebeten, auf die Kosten ihrer vorgeschlagenen Änderungen zu achten. Trotz dieses Hinweises generierten nur 28% aller Teilnehmenden zumindest eine einzige Idee, die das Entfernen eines Elements betraf. Und selbst von jenen Teilnehmenden, die nochmals ausdrücklich darauf hingewiesen wurden, dass sie etwas zum aktuellen Aufbau der Bahn hinzufügen, aber auch entfernen können, generierten nur 43% zumindest einen Vorschlag, bei dem ein Element

entfernt würde. Adams et al. (2021) replizieren diesen *addition bias* (Dt. etwa: Neigung zum Hinzufügen) in sieben weiteren Experimenten. Dabei wurde der Typ der Aufgabe variiert, manche Aufgaben waren abstrakter oder mussten mit tatsächlich anfassbaren Objekten ausgeführt werden. Wenn die Teilnehmenden unter einem erhöhten kognitiven Druck standen (hierbei mussten sie eine zweite Aufgabe parallel zur eigentlichen erledigen), verstärkte sich dieser Effekt noch. Insgesamt schließen die Forschenden daraus, dass Menschen systematisch „subtrahierende Veränderungen“ übersehen, woraus die oben erwähnte generelle Neigung zu hinzufügenden Veränderungen entsteht. Oder wie Winter et al. (2023) es ausdrücken: „More is Better“. All dies hat – so die weitergehende Interpretation – gravierende Auswirkungen auf viele Lebensbereiche. Dies betrifft zum Beispiel übervolle Terminkalender, *institutional red tape* (ein feststehender Ausdruck im Englischen, der wohl am passendsten als „Amtschimmel“ ins Deutsche übersetzt werden könnte) oder negative Effekte auf unseren Planeten.

DIE NEIGUNG ZUM HINZUFÜGEN MANIFESTIERT SICH IN DER SPRACHE

Doch was hat das nun alles mit Sprache zu tun? Winter et al. (2023) können anhand von fünf verschiedenen Analysen zum Englischen zeigen, dass der *addition bias* seinen Niederschlag in der Sprachverwendung findet: Additive Wörter (*add, plus, increase*) sind häufiger als subtrahierende (*subtract, minus, decrease*), ihnen wird eine höhere emotionale Valenz¹ zugewiesen, sie sind den Wörtern *change* und *improve* ähnlicher im Sinne der distributionalen Semantik², und selbst das auf Unmengen an Sprachdaten trainierte Sprachmodell GPT-3 setzt den Satz „I suggest we change/improve this by ...“ eher mit „adding“ fort als mit „removing“. Ein weiterer interessanter Effekt zeigt sich bei Sequenzbildungen: So zeigen Winter et al. (2023), dass bei Paaren von Wörtern, die mit *and* oder *or* verbunden sind, eher das additive Wort zuerst genannt wird. Die Dreiwortfolgen (sogenannte Trigramme) *add or subtract* und *increase and decrease* sind also deutlich häufiger in englischen Korpora zu finden als *subtract or add*

Sascha Wolfer, Alexander Kopenig und Carolin Müller-Spitzer sind wissenschaftliche Mitarbeitende in der Abteilung Lexik und Marc Kupietz ist wissenschaftlicher Mitarbeiter in der Abteilung Digitale Sprachwissenschaft des Leibniz-Instituts für Deutsche Sprache, Mannheim.

und *decrease and increase*. All das interpretieren sie als Ausdruck des *addition bias*, der seinen Niederschlag eben auch in der Sprache findet. Oder anders ausgedrückt: Die Neigung zum Hinzufügen manifestiert sich auch sprachlich. In diesem Beitrag folgen wir dieser Interpretation, die sich primär auf den *addition bias* als Erklärungsfaktor stützt, also nur eine kognitive Dimension beachtet. Alternative, zusätzliche oder interagierende Erklärungsmuster, die sich auf rein phonologische oder morphologische Kriterien stützen könnten, sind somit nicht Gegenstand dieses Beitrags. Das bedeutet jedoch nicht, dass wir diesen ihre potenzielle Relevanz absprechen.

In diesem Beitrag möchten wir den oben beschriebenen Effekten der Einzelwörhäufigkeiten und den Unterschieden hinsichtlich der Dreiwortfolgen (im Folgenden werden wir diese als Trigramme bezeichnen) für die deutsche Sprache nachgehen. Hierzu werden wir eine neue Datenbasis verwenden, die am IDS in Kooperation der Programmbereiche Korpuslinguistik sowie Lexik empirisch und digital entstanden ist: eine große Frequenzliste mit Häufigkeiten aller Einzelwörter (Unigramme), Zweiwortverbindungen (Bigramme) und Trigramme aus dem Deutschen Referenzkorpus (DeReKo), die wir DeReKoGram genannt haben. Wir werden zunächst diese frei zum Download verfügbare Datenbasis vorstellen, bevor wir uns wieder dem *addition bias* im Deutschen zuwenden.

DeReKoGram: Ein neuer Datensatz für die linguistische Forschung

DeReKoGram (Koplenig/Kupietz/Wolfer 2022; Wolfer et al. 2023) ist ein Datensatz zu Worthäufigkeiten für Uni-, Bi- und Trigramme,³ der angereichert ist mit Informationen zu Lemma und Wortart (*part-of-speech*, POS). Datengrundlage ist ein großer Teil des Deutschen Referenzkorpus DeReKo mit ca. 43 Milliarden laufenden Wörtern. Im Gegensatz zu bspw. den Google-Books-Korpora bzw. dem daraus abgeleiteten Häufigkeitsdatensatz (Michel et al. 2010), der ähnlich aufgebaut ist wie DeReKoGram, haben wir keinen Häufigkeitsschwellenwert angesetzt, der übertroffen werden müsste, damit ein *n*-Gramm in den Datensatz aufgenommen wird. Das heißt, dass tatsächlich alle Wortformen aus dem zugrundeliegenden Korpus enthalten sind. Informationen zu Verfügbarkeit sowie Beispielcode für Python, R und Stata sind verfügbar unter <www.owid.de/plus/derekogram/> (Stand 25.5.2023).

Dieser Beispielcode ist in Form von Markup-Dokumenten so dokumentiert, dass er den Einstieg in die Arbeit mit DeReKoGram erleichtert.

DeReKoGram basiert auf 16 etwa gleich großen Korpusteilen (auch „Folds“ genannt), denen die Dokumente im Korpus zufällig zugeordnet wurden. Das macht nicht nur die Datenmenge einfacher handhabbar (zum Beispiel beim Download), sondern ermöglicht es auch, Analysen nur auf einer Teilmenge der Folds durchzuführen.⁴

DeReKoGram besteht aus Tabellen mit sehr vielen Zeilen (eine Zeile pro Uni-, Bi- oder Trigramm). Für Bigramme hat DeReKoGram sieben Spalten:

- die erste Wortform des Bigramms (zahlenkodiert)
- das vom TreeTagger (Schmid 1994) zugeordnete Lemma der ersten Wortform (zahlenkodiert)
- die vom TreeTagger zugeordnete Wortart der ersten Wortform
- die zweite Wortform des Bigramms (zahlenkodiert)
- das Lemma der zweiten Wortform (zahlenkodiert)
- die Wortart der zweiten Wortform
- die Häufigkeit des Bigramms in diesem Fold

Für Unigramme fallen dementsprechend drei Spalten weg, für Trigramme kommen drei Spalten hinzu. Die Wortformen und Lemmata sind zahlenkodiert, da dies die Dateigröße drastisch reduziert und die Verarbeitung der Daten in vielen Fällen erheblich beschleunigt. Für die Rückübersetzung der Zahlencodes in Wörter stellen wir „Wörterbücher“ bereit, die zwei Spalten enthalten: die Wortform/das Lemma und den zugehörigen Code. Auch diese Wörterbücher stellen wir für jede Einzeldatei bereit, was ebenfalls den Umgang mit den Daten erleichtern kann.

Eine Beispiel-Zeile in den Dateien sieht folgendermaßen aus:

```
für für APPR den die ART 1761184
```

Diese Zeile entstammt den Bigramm-Daten für Fold 12, und die Wortformen und Lemmata sind bereits aus den Zahlencodes zurückübersetzt. Wir können an dieser Zeile ablesen, dass das Bigramm *für den*, wobei dem ersten Wort das Lemma *für* (Wortart-Code APPR für Präposition) und dem zweiten Wort das Lemma *die* (ART für Artikel) zugeordnet ist, ca. 1,8 Millionen Mal in diesem Fold vorkommt.⁵

Der Additionsbias in DeReKoGram

DeReKoGram können wir nun dazu benutzen, den *addition bias* für die deutsche Sprache (zumindest so wie sie in DeReKo repräsentiert ist) zu überprüfen.⁶ Wir wollen dabei etwas weiter gehen als Winter et al. (2023) und überprüfen nicht nur Addition und Subtraktion. Wir erweitern die Analyse auf andere Gegenstandsbereiche, die einer gewissen Polarität unterworfen sind. Neben dem Gegenstandsbereich „Mathematik“ (*Addition* vs. *Subtraktion*) nehmen wir Wörter für die Bereiche „Größe/Gewicht“ (*groß* vs. *klein*), „Höhe“ (*oben* vs. *unten*), „Progression“ (*vor* vs. *zurück*), „Reichtum/Macht“ (*reich* vs. *arm*), „Wertigkeit“ (*gut* vs. *schlecht*) und „Wichtigkeit“ (*wichtig* vs. *unwichtig*) mit in die Untersuchungsmenge auf. Die Bezeichnungen für die Bereiche sollten dabei nicht überinterpretiert, sondern lediglich als „Etikett“ verstanden werden. Ausschlaggebend sind ohnehin nur die in Tabelle 1 verzeichneten Wörter bzw. Wortpaare, die durch die Gegenstandsbereiche grob gruppiert werden.

Für jedes dieser Wortpaare haben wir zunächst sowohl Typen als auch Lemmafrequenzen aus dem Unigramm-Datensatz von DeReKoGram extrahiert.

SIND POSITIV POLARISIERTE WÖRTER HÄUFIGER ALS NEGATIV POLARISIERTE WÖRTER?

Das Ergebnis dieser Abfragen ist in Abbildung 1 und Abbildung 2 ablesbar. Auf der Lemma-Ebene (Abb. 1) sind nur bei vier von 28 Paaren⁷ (14,3%) die negativ polarisierten Wörter häufiger. Die Paare mit den Nummern 1, 2, 18 und 19 fallen aus diesem Muster heraus. Die potenziellen Gründe hierfür sind interessant. Bei den Paaren 1 und 2 (*dazuzählen* vs. *abziehen*, *hinzufügen* vs. *entfernen*) könnte der gegenteilige Effekt daher stammen, dass die negativ polarisierten Wörter *abziehen* und *entfernen* in mehr Kontexten eingesetzt werden können als ihre positiv polarisierten Partnerwörter *dazuzählen* und *hinzufügen*. Sowohl *entfernen* als auch – vermutlich in höherem Maße – *abziehen* sind mehrdeutig. So kann *entfernen* im Gegensatz zu *hinzufügen* auch reflexiv eingesetzt werden. Zudem muss wohl konstatiert werden, dass die Einordnung von *abziehen* in den Gegenstandsbereich Mathematik aufgrund der vielen unterschiedlichen Bedeutungen, die dieses Lexem annehmen kann, zumindest fraglich ist. Auf Duden Online sind allein 18 Bedeutungen

Gegenstandsbereich	positiv polarisiert	negativ polarisiert	Paar-Nummer
Mathematik	<i>dazuzählen</i>	<i>abziehen</i>	1
	<i>hinzufügen</i>	<i>entfernen</i>	2
	<i>addieren</i>	<i>subtrahieren</i>	3
Größe	<i>groß</i>	<i>klein</i>	4
	<i>größer</i>	<i>kleiner</i>	5
	<i>riesig</i>	<i>winzig</i>	6
	<i>lang</i>	<i>kurz</i>	7
	<i>schwer</i>	<i>leicht</i>	8
Höhe	<i>hoch</i>	<i>tief</i>	9
	<i>hinauf</i>	<i>hinab</i>	10
	<i>aufwärts</i>	<i>abwärts</i>	11
	<i>aufsteigend</i>	<i>absteigend</i>	12
	<i>steigend</i>	<i>fallend</i>	13
Progression	<i>vorwärts</i>	<i>rückwärts</i>	14
	<i>vor</i>	<i>zurück</i>	15
Reichtum	<i>reich</i>	<i>arm</i>	16
	<i>mächtig</i>	<i>ohnmächtig</i>	17
	<i>wohlhabend</i>	<i>bedürftig</i>	18
	<i>finanzstark</i>	<i>finanzschwach</i>	19
Wertigkeit	<i>positiv</i>	<i>negativ</i>	20
	<i>wertvoll</i>	<i>wertlos</i>	21
	<i>gut</i>	<i>schlecht</i>	22
	<i>nützlich</i>	<i>nutzlos</i>	23
	<i>hilfreich</i>	<i>schädlich</i>	24
Wichtigkeit	<i>wichtig</i>	<i>unwichtig</i>	25
	<i>wesentlich</i>	<i>unwesentlich</i>	26
	<i>notwendig</i>	<i>unnötig</i>	27
	<i>bedeutend</i>	<i>unbedeutend</i>	28
	<i>relevant</i>	<i>irrelevant</i>	29

Tab. 1: Überprüfte Wörter sowie die zugeordneten Gegenstandsbereiche

von *abziehen* vermerkt, von denen nur eine eindeutig dem Gegenstandsbereich der Mathematik zuzuordnen ist. *abziehen* ist somit nur teilweise ein passendes Antonym zu *dazuzählen*. Das Verhältnis⁸ zwischen der Frequenz des positiv polarisierten Wortes und dem negativ polarisierten beträgt für Paar 1: $7.254 / 468.401 = 0,015$, für Paar 2: $173.526 / 1.897.760 = 0,091$.

Für die beiden Paare aus dem Gegenstandsbereich „Reichtum“ liegt keine unmittelbare Erklärung auf der Hand (wobei es Ausdruck einer Präferenz in Zeitungstexten sein könnte, eher finanzschwache als finanzstarke Gesellschaftsschichten zu fokussieren). Hier sind die Unterschiede allerdings auch deutlich kleiner, was sich in den Verhältnissen nieder-

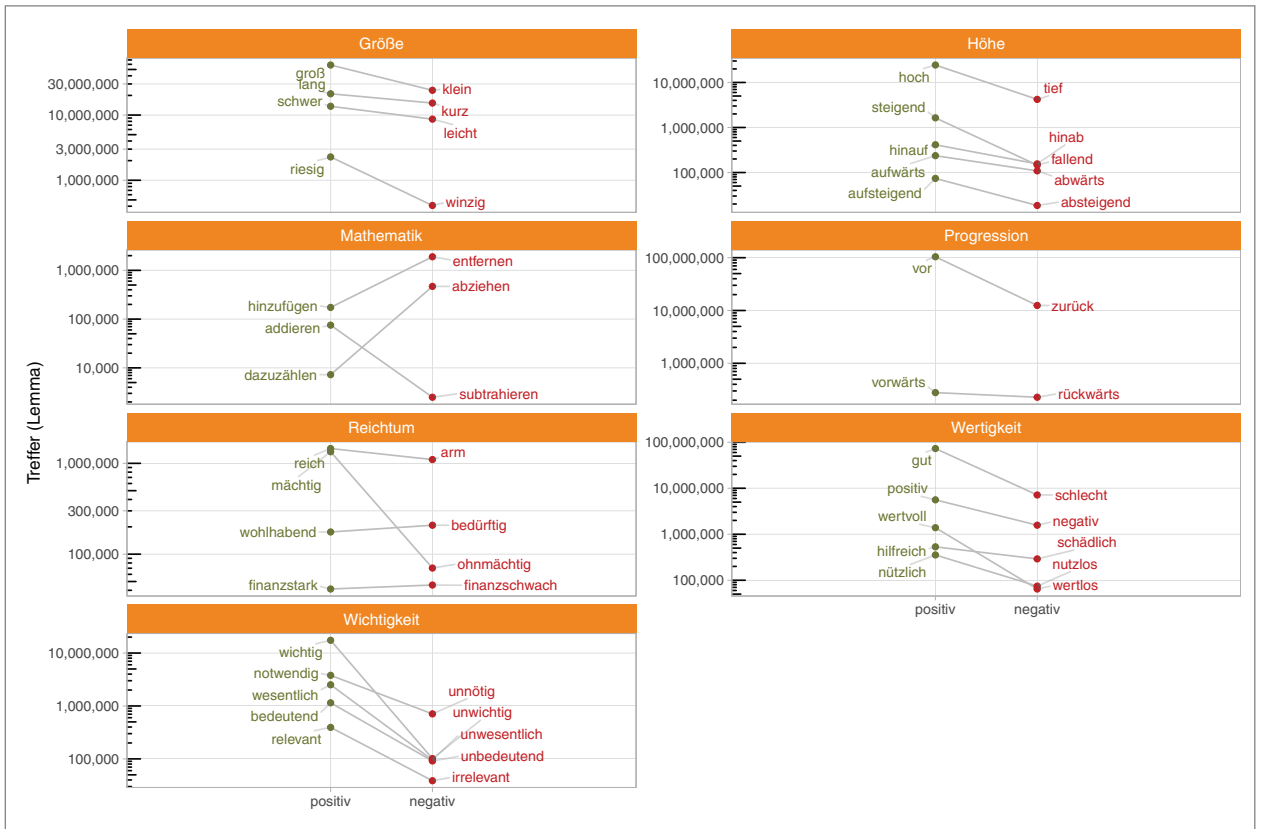


Abb. 1: Einzelwortfrequenzen für alle Suchwörter (Lemma-Ebene) aufgeteilt nach Gegenstandsbereichen. Links ist stets das positiv polarisierte Wort abgetragen, rechts das negativ polarisierte. Die y-Achse ist zur Basis 10 logarithmiert und variiert frei zwischen den Gegenstandsbereichen.

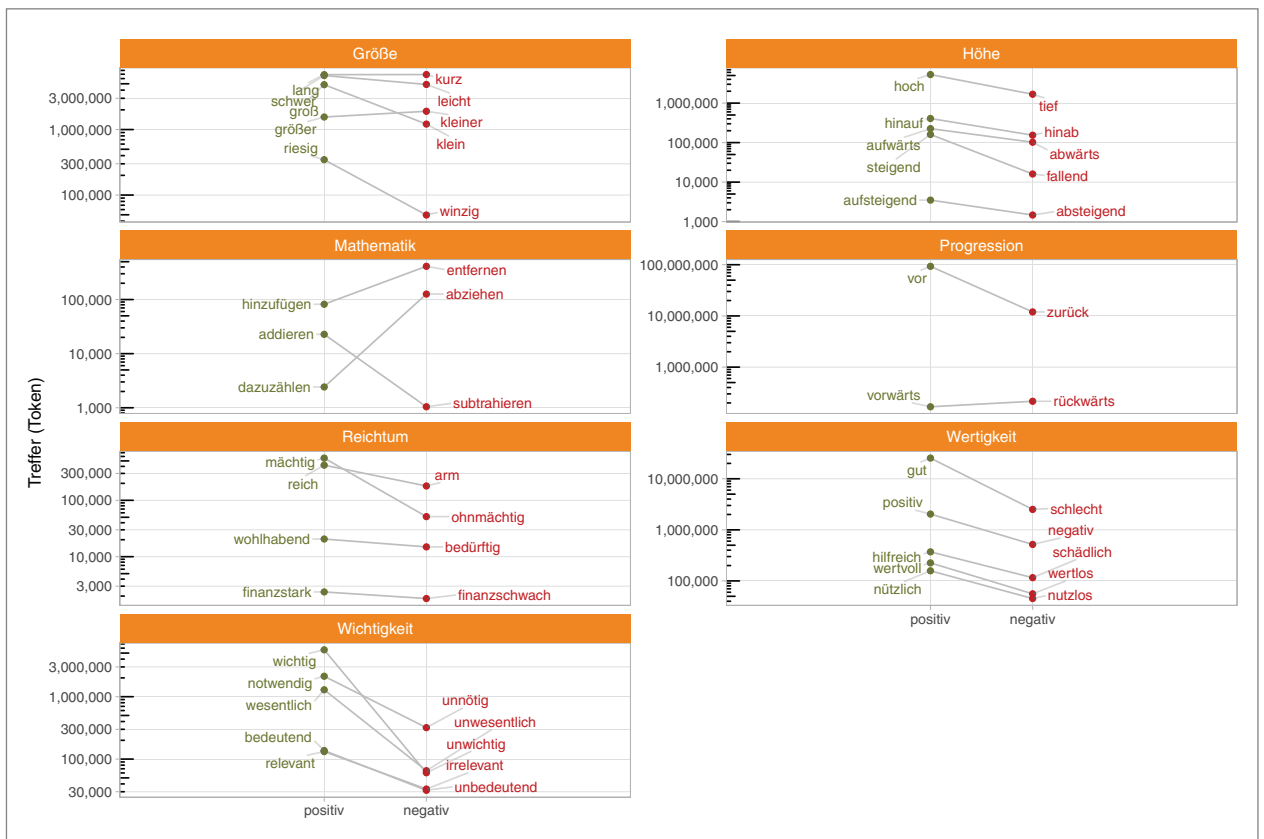


Abb. 2: Einzelwortfrequenzen für alle Suchwörter (Token-Ebene) aufgeteilt nach Gegenstandsbereichen. Die Achsen entsprechen Abbildung 1.

schlägt: Paar 18: $175.541 / 208.225 = 0,843$; Paar 19: $41.362 / 45.730 = 0,904$. Die Verhältnisse sind hier deutlich näher bei 1. Das zeigt, dass der Unterschied deutlich geringer ausgeprägt ist als für die Paare 1 und 2.

POSITIV POLARISIERTE WÖRTER SIND HÄUFIGER ALS NEGATIV POLARISIERTE

Auch auf Token-Ebene (Abb. 2), also den Vorkommen in genau dieser Form ohne jegliche Flexionsformen, sind jene Paare in der Mehrzahl, bei denen das positiv polarisierte Wort häufiger ist. Dies gilt nur für fünf von 29 Paaren (17,2%) nicht. Das sind die bereits bekannten Paare 1 (Verhältnis 0,019) und 2 (0,199) sowie Paar 5 (*größer* vs. *kleiner*, 0,817), Paar 7 (*lang* vs. *kurz*, 0,996) und 14 (*vorwärts* vs. *rückwärts*; Verhältnis 0,781). An den Verhältnissen kann abgelesen werden, dass

für die Paare 5, 7 und 14 die Unterschiede bei weitem nicht so ausgeprägt sind wie für die Paare 1 und 2. Insbesondere *lang* kommt mit 6.886.857 Treffern fast genauso häufig vor wie *kurz* mit einer Häufigkeit von 6.914.020.

Über alle Paare und Korpus-Folds hinweg zeigen Permutationstests,⁹ dass die positiv polarisierten Wörter höchst signifikant häufiger sind als die negativ polarisierten – und zwar in allen Gegenstandsbereichen. Das gilt für die Lemma- und die Token-Ebene und wird durch Wilcoxon-Rangsummentests bestätigt (höchster *p*-Wert = 0,0148 für „Größe“). Wir können daraus schließen, dass in der deutschen Sprache – zumindest in den untersuchten Gegenstandsbereichen und für unsere Wortauswahl – in der Tat eine Neigung besteht, positiv polarisierte Wörter häufiger zu verwenden als negativ polarisierte.

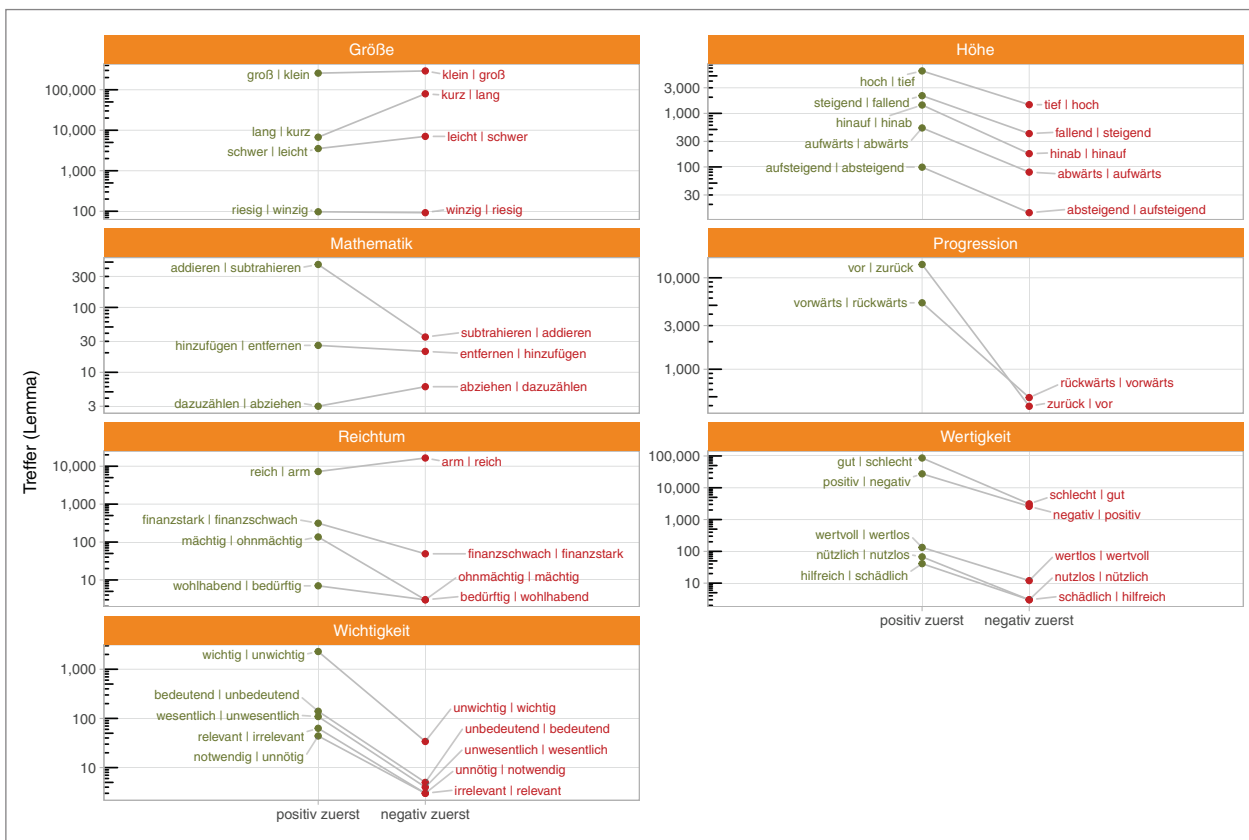


Abb. 3: Frequenzen für Paarbildungen (Lemma-Ebene), aggregiert nach dem zweiten Element des Trigramms und aufgeteilt nach Gegenstandsbereichen. Links ist die Häufigkeit für Trigramme mit dem positiv polarisierten Element an erster Stelle, rechts die Häufigkeit mit dem negativ polarisierten Element an erster Stelle abgetragen. Die y-Achse ist zur Basis 10 logarithmiert und variiert frei zwischen den Gegenstandsbereichen.



Abb. 4: Frequenzen für Paarbildungen (Lemma-Ebene), aggregiert nach dem zweiten Element des Trigramms und aufgeteilt nach Gegenstandsbereichen. Ist keine Verbindungslinie enthalten, sind keine Fälle für die jeweils andere Polarisierung vorhanden (Häufigkeit = 0). Die Achsen entsprechen Abbildung 3.

In einem weiteren Schritt haben wir den Trigramm-Datensatz von DeReKoGram benutzt, um alle Paarbildungen mit *und*, *oder* sowie *&* zu finden, und zwar sowohl mit dem positiv polarisierten Wort als auch mit dem negativ polarisierten Wort an erster Stelle. Für Paar Nummer 21 wurden bspw. alle Vorkommen der folgenden Trigramme extrahiert:

- wertvoll und wertlos
- wertvoll oder wertlos
- wertvoll & wertlos
- wertlos und wertvoll
- wertlos oder wertvoll
- wertlos & wertvoll

WERDEN POSITIVE WÖRTER HÄUFIGER ALS ERSTES GENANNT?

Wird auf Lemma-Ebene gesucht, werden alle Formen der beteiligten Wörter gefunden, zum Beispiel auch *wertvolle* und *wertlose*. Wir haben die Werte über das zweite Element des Trigramms (die Konjunktionen *und*, *oder*, *&*) aufsummiert (aggregiert). Damit bleiben für jedes Paar nur noch zwei Häufigkeitswerte übrig, nämlich jener für Trigramme mit dem positiv polarisierten Element an erster Stelle und jener mit dem negativ polarisierten Element an erster Stelle.

Für die oben genannte Aufzählung sind das die Häufigkeiten für *wertvoll|wertlos* und *wertlos|wertvoll*, wobei der senkrechte Strich | das zweite Element des Trigramms symbolisiert, über das aggregiert wurde. Die Ergebnisse sind in Abbildung 3 und Abbildung 4 abzulesen. Paare, bei denen keinerlei Treffer, egal in welcher Abfolge, gefunden wurden, sind nicht abgetragen (bspw. *dazuzählen* vs. *abziehen* auf Token-Ebene). Es deutet sich visuell bereits an, dass die positiv polarisierten Wörter in Paarbildungen wie bspw. *vor und/oder zurück* eher als erstes genannt werden als die negativ polarisierten (*zurück und/oder vor*).

Auf der Lemma-Ebene ist das nur für 5 der 28 Paare (17,9 %) nicht der Fall, nämlich *dazuzählen* vs. *abziehen* (Verhältnis: $3 / 6 = 0,5$), *reich* vs. *arm* ($7.248 / 16.477 = 0,440$), *groß* vs. *klein* ($257.301 / 291.488 = 0,883$), *lang* vs. *kurz* ($6.752 / 79.405 = 0,085$) sowie *schwer* vs. *leicht* ($3.527 / 7.053 = 0,500$). Hier sticht das Paar *lang* vs. *kurz* heraus, weil eine sehr deutliche Verschiebung in Richtung der Erstnennung des negativ polarisierten Wortes zu beobachten ist. Das ist auch leicht erklärbar, denn wir haben es hier mit dem feststehenden Ausdruck *über kurz oder lang* zu tun. Von allen 86.157 Vorkommen (auf der Lemma-Ebene) des Pairs *kurz* vs. *lang* in all seinen Abfolgen entfallen 76.017 Vorkommen (88,2 %) auf das Trigramm *kurz*

oder lang. Auf der Token-Ebene werden tatsächlich nur die Formen *kurz* und *lang* gefunden und keine anderen Formen wie bspw. *kürzer* oder *länger*. Hierbei ist der Anteil der 70.615 Treffer von *kurz* oder *lang* an allen 72.670 Treffern in allen Abfolgen mit 97,2% noch einmal deutlich höher.

Ähnlich, doch nicht ganz so extrem, verhält es sich mit dem Paar *arm* vs. *reich*. Hier ist es das Trigramm *arm und reich*, das auf der Lemma-Ebene von den 23.725 Vorkommen 12.097 auf sich vereint (51,0%). Auf Token-Ebene sind es 6.108 Treffer von 13.910 insgesamt (46,3%). Wohlbemerkt gilt das nur für die kleingeschriebene Variante. Das Trigramm *Arm und Reich* ist in dieser Analyse nicht beachtet.

WAS MEINEN GROßE STATISTISCHE SPRACHMODELLE DAZU?

Die Unterschiedlichkeit der verschiedenen Gegenstandsbe-
reiche, die sich in den Abbildungen bereits andeutet, schlägt
sich auch in den Permutationstests nieder. Diese zeigen für
alle Gegenstandsbereiche außer „Größe“ einen signifikan-
ten Unterschied der Differenzen der Mediane („positiv zu-
erst“ minus „negativ zuerst“) an. Dies wird durch die Wilco-
xon-Rangsummentests bestätigt ($p = 0,0541$ für „Größe“, alle
anderen p -Werte $< 0,05$).

Daraus lässt sich schließen, dass in der (über DeReKoGram
erfassten) deutschen Sprache in der Tat positiv polarisierte
Wörter in Paarnennungen mit *und* und *oder* eher als erstes
genannt werden. Dies gilt aber nicht unbedingt für Paar-
nennungen des Bereichs „Größe“. Hier ist es insbesondere
das Paar *lang* vs. *kurz*, das aufgrund der festen Fügung *über
kurz* oder *lang* aus dem Rahmen fällt.

Oben haben wir bereits erwähnt, dass Winter et al. (2023)
auch das statistische Sprachmodell GPT-3 nutzten, um den
addition bias im Englischen zu untersuchen. Ihr genaues
Vorgehen replizieren wir hier nicht für das Deutsche, aber
wir können ChatGPT (mit dem zugrundeliegenden Sprach-
modell GPT-3.5) einmal direkt fragen.

Prompt:

Mich interessiert Ihre ganz generelle Einschätzung dazu,
welches der beiden Wörter „hinzufügen“ oder „entfer-
nen“ positiver besetzt ist.

Antwort:

[...]

Im Allgemeinen wird das Wort „hinzufügen“ häufig als
positiver besetzt wahrgenommen als das Wort „entfer-
nen“. Das liegt daran, dass das Hinzufügen von etwas oft
mit Bereicherung, Erweiterung und positiven Verände-
rungen in Verbindung gebracht wird. Es impliziert, dass
etwas Neues oder Zusätzliches eingeführt wird, was als
positiv empfunden werden kann.

Auf der anderen Seite wird das Wort „entfernen“ oft mit
Verlust, Einschränkung oder dem Wegnehmen von et-
was assoziiert. Es kann daher eher als negativ oder zu-
mindest als weniger positiv besetzt empfunden werden.
Es bedeutet, dass etwas weggenommen oder beseitigt
wird, was potenziell als Veränderung oder Einschrän-
kung wahrgenommen werden kann.

Es ist jedoch wichtig zu beachten, dass diese Einschät-
zung nicht in allen Kontexten gilt. Es gibt Situationen, in
denen das Entfernen von etwas positiv sein kann, zum
Beispiel bei der Beseitigung von Hindernissen oder nega-
tiven Einflüssen. Ebenso kann das Hinzufügen in be-
stimmten Fällen als negativ empfunden werden, wenn es
zu Überlastung, Komplexität oder Einschränkungen führt.
[...]

Das ist eine interessante Antwort, da in den ersten beiden
Absätzen der *addition bias* tatsächlich explizit angesprochen
wird (mit einigen einschränkenden Zusätzen wie *im Allge-
meinen*, *oft* und *eher*). Im dritten Absatz wird diese Einschät-
zung dann mit Verweis auf die Kontextabhängigkeit relati-
viert.

Dem ist nicht mehr viel hinzuzufügen, und vermutlich wür-
den die meisten Menschen dieser aufgrund tiefer statisti-
scher Analyse sprachlicher Massendaten generierten Ein-

schätzung zustimmen oder ganz ähnliche Antworten auf die Frage geben, die wir ChatGPT gestellt haben. Auch wenn diese Antwort unsere empirischen Analysen bestätigen, zeigt unsere Untersuchung doch, dass das explizite Beschreiben solcherlei Kontraste nicht zusammenfallen muss mit dem tatsächlichen Sprachgebrauch, in dem solche *biases* nur implizit sind und sich oft nur mit der Analyse großer Datenbestände freilegen lassen. Mit DeReKoGram stellt das IDS nun einen Datensatz zur Verfügung, mit dem sich solche und vielerlei Fragen mehr bearbeiten lassen. ■

Anmerkungen

- ¹ Emotionale Valenz gibt an, welchen emotional-affektiven „Wert“ Menschen einem Konzept zuweisen, das von einem Wort beschrieben wird. Um solche Werte zu erheben, werden meist viele Personen befragt, die Wörter einschätzen müssen. So wird bspw. *Liebe* konsistent ein sehr hoher Valenzwert, *Krieg* ein sehr niedriger Valenzwert zugewiesen. Aus diesen menschlichen Bewertungen werden dann einzelne Valenzwerte für bestimmte Wörter berechnet. Wolfer und Hein (2022) setzen solche Valenzwerte dazu ein, die emotionale „Wirkung“ der Suffigierung mit *-los* zu untersuchen. Die Häufigkeitswerte, die dort berichtet werden, sind ebenfalls anhand von DeReKoGram ermittelt.
- ² Die distributionelle Semantik basiert auf dem Credo von Firth (1957, S. 179): „You shall know a word by the company it keeps“. Dabei wird die Ähnlichkeit zwischen zwei Wörtern quantifiziert über die Ähnlichkeit der Kontexte, in denen die beiden Wörter auftreten. Berechnet werden dabei Wortvektoren, die meist über die Kosinus-Ähnlichkeit miteinander verglichen werden. Das IDS stellt ebenfalls eine Datenquelle für Wortvektoren, die anhand von DeReKo erstellt sind, mit dem Namen DeReKoVecs zur Verfügung (Fankhauser & Kupietz 2022).
- ³ Will man über die Anzahl der Wörter in der Sequenz verallgemeinern, spricht man auch von „*n*-Grammen“.
- ⁴ Da die Korpusdokumente zufällig auf die Folds verteilt wurden, können viele Analysen von einem auf alle Folds verallgemeinert werden. Das ermöglicht es beispielsweise auch, DeReKoGram auf einfachen Arbeitsplatzrechnern einzusetzen und nicht nur auf großen Datenverarbeitungsservern.
- ⁵ Um die Zahl der möglichen Fragestellungen, die anhand von DeReKoGram bearbeitet werden können, noch zu erweitern, stellen wir zwei Varianten des Datensatzes zur Verfügung, nämlich mit und ohne Satzzeichen. Für Unigramme hat das lediglich die Auswirkung, dass bspw. Informationen dazu, wie viele Kommata im Korpus vorkommen, nicht in den Häufigkeitsdaten enthalten sind. Für Bi- und Trigramme jedoch hat das Auslassen von Satzzeichen vor der Berechnung der *n*-Gramm-Häufigkeiten tiefgreifendere Konsequenzen. Schauen wir uns als Beispiel den folgenden Satz an: „Sie glaubt, dass man Glasritzen sehr wohl mit Zeichen vergleichen könne.“ (Glasritzen ist Erfahrungssache, Andreas König, St. Galler Tagblatt vom 7.10.1999). Das zweite Trigramm dieses Satzes, das im Datensatz mit Satzzeichen vertreten ist, lautet „glaubt, dass“ (POS-Tags VVFIN, finites Verb; \$, Komma; KOUS, unterordnende Konjunktion mit Satz). Für den Datensatz ohne Satzzeichen lautet das zweite Trigramm „glaubt dass man“ (VVFIN; KOUS; PIS, substituierendes Indefinitpronomen). Was auf den ersten Blick wie ein vernachlässigbarer Unterschied aussieht, kann sich aber auf die Zahl möglicher Analysen auswirken: Möchte man zum Beispiel alle möglichen Fortsetzungen nach satzeinbettenden Prädikaten wie „glauben“ in mit „dass“ eingeleiteten Nebensätzen extrahieren, wird uns der Datensatz mit Satzzeichen nicht weiterhelfen, weil das Komma dafür sorgt, dass die Fortsetzung nicht im Trigramm enthalten ist. Der Datensatz ohne Satzzeichen stellt diese Information bereit.
- ⁶ Die hier gezeigten Analysen basieren auf einer Studie (Wolfer 2023), die in der Fachzeitschrift *Cognitive Science* als Antwort auf die Studie von Winter et al. (2023) erschienen ist. Dort werden noch weitere Analysen präsentiert (insbesondere im Vergleich zum Englischen), allerdings mit einer kleineren Anzahl an Wortpaaren.
- ⁷ Paar 5 fällt auf Lemma-Ebene weg, da den Token „größer“ und „kleiner“ stets die Lemmata „groß“ und „klein“ zugeordnet werden.
- ⁸ Ist dieser Wert kleiner 1, ist das negativ polarisierte Wort häufiger als das positiv polarisierte. Ist der Wert größer 1, ist das positiv polarisierte Wort häufiger, was der Annahme des *addition bias* entspricht.
- ⁹ Im Permutationstest wird die Information, ob ein Wort positiv oder negativ polarisiert ist, 50.000-mal zufällig den Wörtern zugeordnet. In keiner dieser 50.000 zufälligen Anordnungen ist der Median des Häufigkeitsunterschieds größer als in den tatsächlich beobachteten Daten.

Literatur

- Adams, Gabrielle S./Converse, Benjamin A./Hales, Andrew H./Klotz, Leidy E. (2021): People systematically overlook subtractive changes. In: *Nature* 592, 7853, S. 258-261. <<https://doi.org/10.1038/s41586-021-03380-y>>.
- Fankhauser, Peter/Kupietz, Marc (2022): Count-based and predictive language models for exploring DeReKo. In: Banski, Piotr/Barbareasi, Adrien/Clematide, Simon/Kupietz, Marc/Lüngen, Harald (Hg.): *Proceedings of the workshop on challenges in the management of large corpora (CMLC-10)*. Marseille: European Language Resources Association, S. 27-31. <<https://aclanthology.org/2022.cmlc-1.5>>.
- Firth, John Rupert (1957): *Papers in linguistics, 1934-1951*. London: Oxford University Press.
- Koplenig, Alexander/Kupietz, Marc/Wolfer, Sascha (2022): Testing the relationship between word length, frequency, and predictability based on the German reference corpus. In: *Cognitive Science* 46, 6. <<https://doi.org/10.1111/cogs.13090>>.
- Michel, Jean-Baptiste/Shen, Yuan Kui/Aiden, Aviva Presser et al. (2010): Quantitative analysis of culture using millions of digitized books. (Supporting Online Material II). In: *Science* 331, 6014, S. 176-182. <<https://doi.org/10.1126/science.1199644>>.
- Schmid, Helmut (1994): Probabilistic part-of-speech tagging using decision trees. In: *Proceedings of the International Conference on New Methods in Language Processing*. Manchester, UK. [Revidierte Fassung.] <www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf>.
- Winter, Bodo/Fischer, Martin H./Scheepers, Christoph/Myachykov, Andriy (2023): More is better: English language statistics are biased toward addition. In: *Cognitive Science* 47, 4. <<https://doi.org/10.1111/cogs.13254>>.
- Wolfer, Sascha (2023): Is More Always Better? Testing the Addition Bias for German Language Statistics. In: *Cognitive Science* 47, 9. <<https://doi.org/10.1111/cogs.13339>>.
- Wolfer, Sascha/Hein, Katrin (2022): Konsequenzen der *los*-Suffigierung im Deutschen: Korpushäufigkeit, emotional-affektive Effekte und konstruktionsgrammatische Perspektiven. In: *Zeitschrift für Wortbildung/Journal of Word Formation* 6, 2, S. 71-99. <<https://doi.org/10.3726/zwjw.2022.02.03>>.
- Wolfer, Sascha/Koplenig, Alexander/Kupietz, Marc/Müller-Spitzer, Carolin (2023): Introducing DeReKoGram: A Novel Frequency Dataset with Lemma and Part-of-Speech Information for German. In: *Data*, 8, 11. <<https://doi.org/10.3390/data8110170>>. ■