

Kollokationen im Lernerwörterbuch – Anspruch und Wirklichkeit

Jupp Möhring (Leipzig)

Abstract

This article outlines an approach which explores how frequent lexical collocations can be identified for a representative extract of the German lexicon. It takes into account corpus-linguistic and phraseological criteria. The lexical foundation for this study is a corpus-driven and thematically-classified frequency dictionary of German. Additionally, the extensive DWDS corpus and a manual linguistic selection were used to identify typical collocations. Finally, after identifying 56 high-frequency collocations, four of the foremost learner dictionaries in the field of German as a foreign language were analyzed quantitatively for the presence of these collocations.

1 Einführung

Die Relevanz lexikalischer Kollokationen für Fremdsprachenlerner, in diesem Artikel bezogen auf jene des Deutschen als Fremdsprache, wurde bereits mehrfach deutlich herausgestellt (cf. Hausmann 1984: 400ff; Herbst 1996: 389ff, Reder 2006: 59ff; Steyer 2008: 186; Wallner 2010: 197f.) und wird deshalb in diesem Beitrag ebenso wenig nochmals diskutiert wie der Terminus Kollokation selbst (cf. dazu u. a. Bahns 1996: 6ff; Bubenhofer 2009: 111ff; Hausmann 2004; Konecny 2010: 77ff; Steyer 2008: 187ff). Auch die sich aus der Lernerrelevanz folgerichtig ergebende lernerlexikographische Forderung, Kollokationen "nicht länger als zufälliges Accessoire eines Wörterbucheintrages [zu] betrachten" (Viehweger 1989: 892), "da sie einen Datentyp darstellen, der einem vorrangigen Benutzerbedürfnis entspricht" (Viehweger 1989: 892 nach Steyer 2000: 105), ist schon gestellt (cf. Bahns 1996: 37f.) und die Beschreibung der Darstellung von Kollokationen in verschiedenen Wörterbüchern konsequenterweise ein wichtiger Aspekt der Wörterbuchkritik (cf. u. a. Engelberg/Lemnitzer 2001: 189ff; Köster/Neubauer 2002; Lehr 1998; Markus/Korhonen 2005; Reder 2006: 167 ff; Steyer 2008: 191ff). Jedoch scheinen nicht nur die Wörterbuchautoren Probleme bei der Kollokationsauswahl und -verzeichnung zu haben, denn auch "[d]ie Untersuchung der Kollokationen in den Lernerwörterbüchern geschieht im allgemeinen mit relativ geringer Systematik" (Jehle 1990: 117). Was aber genau unter einer ausführlichen und systematischen Analyse zu verstehen ist, wird in der einschlägigen Literatur sehr unterschiedlich definiert (cf. Bahns 1996: 68). Dies führt dazu, dass die Auswahl des zu überprüfenden Lexikonausschnitts oft auf vagen Kriterien basiert (cf. Jehle 1990: 264).

Dieser Beitrag möchte einen Vorschlag unterbreiten, wie semantische, phraseologische und korpuslinguistische Kriterien kombiniert für eine systematische Wörterbuchanalyse genutzt werden können. So werden besonders gebräuchliche und somit lernerrelevante lexikalische Kollokationen für einen Ausschnitt des Lexikons identifiziert. Im Anschluss daran werden

vier¹ einsprachige Lernerwörterbücher (LWB) quantitativ² auf die Eintragung dieser Kollokationen hin untersucht.

2 Kollokationen

Als Kollokationen werden in diesem Beitrag binäre³ Wortverbindungen gewertet, welche aus einer nominalen Basis und einem verbalen Kollokator bestehen (cf. Hausmann 1984: 401; Reder 2001: 2). Andere Strukturtypen (cf. Bahns 1993: 33) bleiben unberücksichtigt. Kollokationen sind nicht satzwertig, semantisch kompositionell, strukturell variant und weisen häufig eine eingeschränkte Kombinationsfähigkeit⁴ auf (Neubert 1966: 109; Reder 2004: 532f.). Diese eingeschränkte Kombinierbarkeit ist von besonderer Relevanz für den Fremdsprachenlerner, da dadurch sprachvergleichend häufig unvorhersehbare Kontraste in der Kombinationsfähigkeit (Scherfer 2001: 9) und somit Interferenzfehler auftreten.

3 Auswahl der Kollokationen

3.1 Die Ausgangswörter

Um die Aussagefähigkeit der Wörterbuchanalyse zu gewährleisten, wurde auf thematische Vielfalt und hohe Gebräuchlichkeit der Kollokationen geachtet.

Thematische Vielfalt wurde gewährleistet, indem der onomasiologisch geordnete Grundwortschatz "Grund- und Aufbauwortschatz Deutsch als Fremdsprache nach Themen" (Tschirner 2008b) genutzt wurde. Dieser Wortschatz umfasst die 4000 häufigsten Wörter des Deutschen⁵, eingeordnet in 16 Sachgruppen. Diese sind an Profile Deutsch (Glabionat 2010) und somit am *Gemeinsamen europäischen Referenzrahmen für Sprachen* orientiert. Als Ausgangswörter für die Kollokationsidentifikation wurden aus dem "Grund- und Aufbauwortschatz Deutsch als Fremdsprache nach Themen" (Tschirner 2008b) die häufigsten Nomen der 16 Sachgruppen sowie die zehn häufigsten Verben aus der Sonderkategorie *die 50 häufigsten Verben* extrahiert. Diese Auswahl hat gegenüber Auswahlmethoden, welche z. B. eine breite alphabetische Streuung (cf. Bahns 1996: 63f.; Jehle 1990: 259ff; Markus/Korhonen 2005) oder die erfolgte Aufnahme einer Mehrwortverbindung in einem anderen Wörterbuch als Kriterium definieren, wesentliche Vorteile. So ist die alphabetische Verteilung des Wortschatzes sehr ungleichmäßig (cf. Haß-Zumkehr 2001: 383f.) und der Eintrag in einem Wörterbuch keine Garantie für die tatsächliche Gebräuchlichkeit. Häufigkeitsaspekte spielen hierbei eine ungleich relevantere Rolle (cf. Tschirner 2005: 134ff). Die Kollokationsidentifikation auf Grundlage der 16 Nomen wird im Folgenden baseninduzierte Analyse genannt.

Neben der baseninduzierten Analyse wurde eine kollokatoreninduzierte Analyse durchgeführt, wobei die zehn häufigsten Verben des Deutschen die Datengrundlage bildeten⁶.

¹ Der damit einhergehende quantitative Vergleich der Kollokationsaufnahme der relevantesten einsprachigen LWB ist ein Desiderat der Wörterbuchforschung, welche i. d. R. lediglich ein WB gründlich analysiert, gelegentlich noch Vergleiche zu einem anderen (L)WB zieht, jedoch bislang nicht auf einer empirisch verlässlichen Datenbasis auf alle vier LWB reflektiert.

² Zu weiteren Analyseebenen cf. Steyer 2008: 193.

³ Die Beschränkung auf binäre Einheiten wird vor allem durch die weitere Arbeit mit dem DWDS Korpus notwendig, da die dort angebotene Kollokationsanalyse nur für binäre Einheiten verfügbar ist.

⁴ Zu möglichen Ursachen dieser cf.: Karacic 2004: 293; Kromann 1989: 267; Nesselhauf 2005: 30f.; Reder 2004: 534 und Scherfer 2001: 18.

⁵ Die empirische Grundlage dafür bildet das 4,2 Millionen Token große, ausgewogene und repräsentative Herder/BYU-Korpus der deutschen Gegenwartssprache (Jones/Tschirner 2006; Tschirner 2008a).

⁶ Von der Analyse unbeachtet blieben darunter Hilfsverben, Modalverben und Funktionsverben, da eine Analyse von verbnominalen Verbindungen in einem Korpus bei Verben, welche im Deutschen zur Bildung von Zeitformen genutzt werden, kaum zielführende Aussagekraft besitzt, sollte das Korpus nicht in die

Tabelle 1 zeigt die als Analysebasis genutzten Nomen und Verben mit der Angabe des Ranks⁷ im thematisch geordneten Frequenzwörterbuch.

Name der Sachgruppe	Rank im WB	Ausgangswort
Nomen		
1. Personalien, Informationen zur Person	103	Frau
2. Wohnen	159	Haus
3. Umwelt	186	Stadt
4. Reisen und Verkehr	104	Mensch
5. Verpflegung	970	Essen ⁸
6. Einkaufen	251	Geld
7. Öffentliche und private Dienstleistungen	350	System
8. Körper und Gesundheit	179	Hand
9. Wahrnehmung und Bewegung	323	Sinn
10. Ausbildung	89	Beispiel
11. Arbeitswelt	167	Arbeit
12. Sprache	231	Grund ⁹
13. Freizeit und Unterhaltung	204	Ende
14. Persönliche Beziehungen und Kontakte	305	Entwicklung
15. Politik und Gesellschaft	160	Fall
16. Allgemeine Begriffe	51	Jahr
17. Strukturwörter		Kein Nomen

(Fortsetzung auf der nächsten Seite)

Funktionen des Verbs als Vollverb oder Hilfsverb unterscheiden können. Dies ist bei der Analyse mit Hilfe des DWDS Kernkorpus nicht möglich. Die Funktionsverben wurden von der Analyse ausgeschlossen, da viele ihrer signifikanten nominalen Partner mit ihnen tatsächlich Funktionsverbgefüge bilden. Die Analyse zielt jedoch darauf ab, eine möglichst große Anzahl hochfrequenter lexikalischer Kollokationen zu finden, um deren Darstellung in den großen einsprachigen LWB zu überprüfen. Als Funktionsverben gelten hier nach Heine (2004: 74): *bringen, finden, gehen, halten, kommen, nehmen, setzen, stehen, stellen, treten* und *ziehen*. Aus technischen Gründen musste außerdem das Verb *machen* von der Analyse ausgeschlossen werden, da hierfür keine Daten in der Kollokationsdatenbank des DWDS vorliegen.

⁷ Der Rank gibt die Platzierung in der Häufigkeitsliste an, wobei Rank 1 das häufigste Wort definiert, Rank 2 das zweithäufigste usw. Der Rank ergibt sich aus der absoluten Häufigkeit eines Lexems inkl. aller Nennformen innerhalb des Herder/BYU-Korpus. Ist der Wert der Häufigkeit zwischen mehreren Lexemen identisch, so wird alphabetisch geordnet.

⁸ Das Nomen *Essen* wird in "Deutsch als Fremdsprache nach Themen" nicht separat aufgeführt, sondern lediglich das Verb *essen* (Rank 657). Die Zuordnung erfolgte manuell.

⁹ Für eine ausführliche Beschreibung von Kollokationen mit der Basis *Grund* cf. Steyer 2008: 198ff.

18. 50 häufigste Verben		Verb
	46	sagen
	57	geben
	81	sehen
	82	lassen
	112	bleiben
	118	liegen
	123	heißen
	124	denken
	140	tun
	143	glauben

Tabelle 1: Die als Analysebasis genutzten Nomen und Verben

Insgesamt wurden somit 26 Ausgangswörter für die anschließende Kollokationsidentifikation genutzt, welche einen thematisch ausgewogenen Ausschnitt des Wortschatzes darstellen. Zudem gehören sie durchweg zu den 1000 häufigsten Wörtern des Deutschen und sind somit essenziell für jeden Fremdsprachenlerner (cf. Tschirner 2008a: 195).

3.2 Die Kollokationsidentifikation

Für die Identifikation der statistisch signifikantesten Kollokationspartner wurde das DWDS Kernkorpus¹⁰ genutzt, da es mit 100 Millionen laufenden Textwörtern der Mindestgröße zur Analyse von Mehrworteinheiten (Tschirner 2005: 137) entspricht und für die Analyse von Wörterbüchern geeignet ist (Geyken 2004: 74, 79). Es ist repräsentativ aufgebaut, stellt gut geeignete Recherchewerkzeuge bereit und ermöglicht, bis auf wenige, urheberrechtliche Einschränkungen, Zugang zu bibliografisch markierten Belegen¹¹.

Mit Hilfe der Kollokationsabfrage des DWDS wurden im Rahmen der baseninduzierten Analyse die gemäß dem Assoziationsmaß Log-Likelihood¹² 20 signifikantesten verbalen Kollokationspartner zu jedem Nomen identifiziert. Für die kollokatoreninduzierte Analyse wurden auf die gleiche Weise die 20 signifikantesten nominalen Kollokationspartner gefunden. Die Kollokationssuche bei DWDS erfolgte in einer Kollokationsspanne von jeweils fünf Wörtern links und rechts des untersuchten Ausgangswortes. Für die 26 Ausgangswörter lagen somit insgesamt 520 statistisch signifikante Kollokationen (cf. Tabellen im Anhang)¹³ vor. Diese Wortkombinationen beinhalteten jedoch nicht ausschließlich lexikalische Kollokationen, sondern auch freie Wortverbindungen (worunter auch alle semantisch und/oder grammatisch nicht verbundenen Kombinationen gezählt werden), Funktionsverbgefüge und Idiome, weshalb anhand der eingangs erwähnten Kriterien manuell

¹⁰ Verfügbar unter <http://www.dwds.de>, Stand 30. Mai 2010.

¹¹ Das DWDS ist eins von inzwischen mehreren großen verfügbaren Korpora der deutschen Sprache, die sich zur Kollokationssuche eignen. An dieser Stelle sei ergänzend auf das Deutsche Referenzkorpus *DeReKo* des IDS Mannheim und das *Wortschatzportal* der Universität Leipzig verwiesen (cf. Steyer 2008: 189ff).

¹² Das Log-Likelihood (LL) gilt als das zuverlässigste Maß zur Identifikation von lexikalischen Kollokationen (McEnergy/Tono/Xiao 2006: 55f., 217).

¹³ Diese 520 *Kookkurrenzen* (zur terminologischen Abgrenzung cf. Steyer 2008: 187ff) bilden die quantitative Basis der Untersuchung. Sie sind dem Artikel als Anhang beigelegt. Da es sich beim DWDS Kernkorpus um ein statisches Korpus handelt, sind sie auch jederzeit online nachprüfbar.

überprüft und selektiert werden musste. Die Überprüfung der Kombinationen hinsichtlich der lexikalischen Kriterien von Kollokationen ergab eine relativ geringe Menge typischer Kollokationen, was Reder (2006: 62) zufolge auf Grund der weitgehenden semantischen Verträglichkeit und Kombinationsfähigkeit im Grundwortschatz zu erwarten war. Zu einigen Ausgangswörtern konnte innerhalb der Untersuchungsmenge gar keine lexikalische Kollokation gefunden werden. So sind die häufigsten verbalen Partner¹⁴ von *Frau*: *sagen, sehen, kommen, lieben, sitzen, gehen, fragen, müssen, stehen, geben, sollen, lassen, wissen, sprechen, rufen, erzählen, nehmen, denken, kennen, dürfen, bitten*. Die einzige offenkundige Besonderheit bei diesen Wortpaaren tritt in der Struktur (*zur*) *Frau nehmen* auf, wobei dieses Wortpaar eindeutig zu den Idiomen gezählt werden muss, da die Gesamtbedeutung *heiraten* sich nicht aus den einzelnen Komponenten erschließt.

Insgesamt konnten 59 häufige¹⁵ lexikalische Kollokationen identifiziert werden. In zwei Fällen fand sich die gleiche hochfrequente und als Kollokation eingeordnete Wortkombination sowohl bei der baseninduzierten als auch bei der kollokatoreninduzierten Untersuchung. Es handelt sich dabei um: *Beispiel geben* und *Grund geben*. Für die Analyse der Wörterbücher war eine doppelte Verzeichnung nicht notwendig, weshalb die beiden Kollokationen in diesem Teil der Auswertung nur jeweils einmal geführt werden. Ihre Zugehörigkeit zum hochfrequenten Wortschatzbereich ist durch die Dopplung unterstrichen. Eine weitere Auffälligkeit der baseninduzierten Analyse war das Auftreten von *laden* und *einladen* in Kombination mit *Essen*. Die Sichtung der Belege ergab, dass *laden* in der Struktur *einladen* genutzt wird. Die Korpusanalyse hat die Trennung des Verbs in *ein-* und *laden* nicht erfassen können. Für die Analyse der Wörterbücher wird lediglich die Struktur *Essen einladen* untersucht. Entsprechend wird *Essen laden* als Kollokation nicht berücksichtigt. Analog wird mit den Kollokatoren *treten* und *eintreten* bei der Basis *Haus* verfahren. Es bleiben 56 hochfrequente lexikalische Kollokationen, welche im Folgenden, alphabetisch aufsteigend nach ihrer Basis sortiert, aufgelistet sind¹⁶:

Anlass geben, Arbeit gehen, Arbeit geben, Aufschluss geben, Auskunft geben, Beispiel geben, Beispiel folgen, Beispiel anführen, Beispiel liefern, Ende setzen, Ende kommen, Ende finden, Entwicklung sehen, Erinnerung bleiben, Essen einladen, Essen gehen, Essen servieren, Fall eintreten, Gedächtnis bleiben, Gefallen tun, Geld verdienen, Geld kosten, Geld ausgeben, Geld sparen, Geld stecken, Geld lassen, Gelegenheit geben, Genüge tun, Grund geben, Grund sehen, Hand halten, Hand drücken, Hand heben, Hand geben, Hand schütteln, Hände falten, Haus schicken, Haus treten/eintreten, Haus gehen, Interesse liegen, Leben bleiben, Leben lassen, Mühe geben, Pflicht tun, Ruhe lassen, Schnee liegen, Schritt tun, Schwerpunkt liegen, Sinn geben, Sinn kommen, Stadt ziehen, Unrecht tun, Verhältnis liegen, Wahrheit sagen, Wasser lassen, Zeit lassen.

Da die Abfrage der Kollokationsdatenbank des DWDS sich auf binäre Einheiten beschränkt, kann an dieser Stelle nicht differenziert werden, ob die Kollokationen präferiert mit einem oder mehreren weiteren Lexem/en, bspw. Präpositionen, auftreten.

¹⁴ Vor der Reduktion um Modal- und Funktionsverben.

¹⁵ *Häufig* sei verstanden als gemäß LL-Wert und somit neben anderen Parametern auch gemäß der Häufigkeit signifikant.

¹⁶ Die unscharfe Grenzziehung in der Kategorisierung von Mehrwortverbindungen und in der Auslegung einzelner Kriterien lässt an dieser Stelle zweifellos eine Diskussion über die Zuordnung der einen oder anderen *lexikalischen* Kollokation zu. So könnte man bspw. der Wortverbindung *Genüge tun* einen gewissen Grad an Unikalität/Idiomatizität unterstellen oder den Grad der Kompostionalität der Verbindungen *Wasser lassen* oder *Mühe geben* diskutieren. Dennoch sind all diese Wortverbindungen statistisch sehr gebräuchlich und auch die identifizierten Kollokationspartner gehören bis auf die Ausnahmen *Aufschluss* und *falten* zu den häufigsten 4000 Wörtern des Deutschen.

4 Analyse der Wörterbücher

Nach der Identifikation der 56 hochfrequenten Kollokationen zu den 26 Ausgangswörtern wurde das Vorkommen dieser Kollokationen in den vier großen einsprachigen LWB des Deutschen als Fremdsprache analysiert. Es handelt sich dabei um das *de Gruyter Wörterbuch Deutsch als Fremdsprache* (GWB), das *Langenscheidt Großwörterbuch Deutsch als Fremdsprache* (LDaF), das *Pons Großwörterbuch Deutsch als Fremdsprache* (PDaF) und das *Wahrig Großwörterbuch Deutsch als Fremdsprache* (WAR). Die Wörterbücher wurden jeweils auf einen Eintrag der Kollokation unter der nominalen Basis und unter dem verbalen Kollokator der Kollokation überprüft. Somit ergeben sich 112 mögliche Einträge der 56 untersuchten Kollokationen.

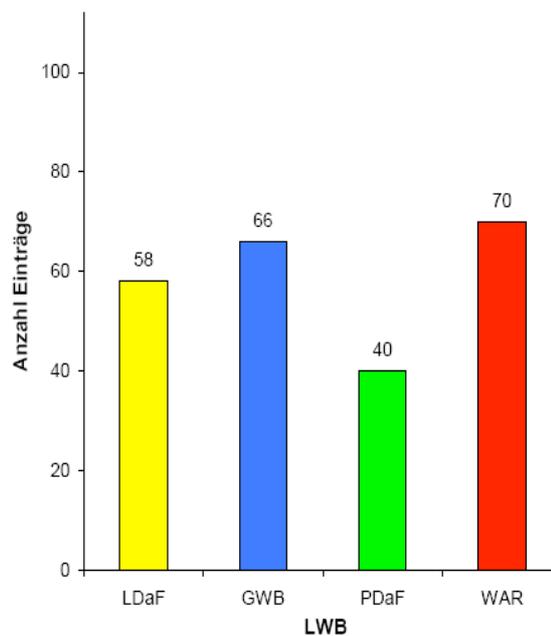


Abb. 1: Anzahl der Kollokationseinträge von 112 möglichen Eintragungen im LWB

In diesem Beitrag wird dabei nicht differenziert, in welchem Teil der Wörterbuchartikel¹⁷ die Kollokationen auftreten. Als Eintrag gewertet wurden zudem alle möglichen Formen, in denen die Kollokationen in den LWB verzeichnet sind, da die Korpusanalyse keine präzisere Aussage zu über die Zweiworteinheit hinausgehenden Kombinationskonventionen der Kollokationen (Numerus, Präpositionen, Konjunktionen, Pronomen) (siehe oben) treffen konnte¹⁸. Das quantitative Darstellungsoptimum von 112 Einträgen erreicht keins der untersuchten LWB. Den höchsten Wert bietet das WAR mit insgesamt 70 Einträgen, etwa ein Drittel der möglichen Eintragungen liefert das PDaF (Abb. 1). Die Differenz von 30 Einträgen zwischen dem LWB mit den meisten und den wenigsten verzeichneten Kollokationen verdeutlicht, wie uneinheitlich die Aufnahme von Kollokationen in die LWB schon bei einem kleinen Ausschnitt des Grundwortschatzes erfolgt¹⁹. Im Folgenden wird

¹⁷ Natürlich ist für eine Aussage über die Qualität der Kollokationsdarstellung der LWB auch die Mikrostruktur der Wörterbuchartikel zu berücksichtigen (cf. Steyer 2008: 193f.), sprich ob die Kollokationen im Beispielpart oder Definitionsteil des Artikels erscheint, ob sie als Idiom oder Kollokation gekennzeichnet sind etc. Dieser Artikel beschränkt sich jedoch auf quantitative Aspekte und somit das WAS, für Vorschläge zum WIE cf. u. a. Bahns 1996: 39ff, Steinbügl 2005: 46ff und Steyer 2008: 200f.

¹⁸ Zum Desiderat der an dieser Stelle vernachlässigten Berücksichtigung von Flexionsmerkmalen und Erweiterungskonventionen von Kollokationen cf. Steyer 2008: 197ff.

¹⁹ Wobei die Anzahl der eingetragenen Stichworte zwischen WAR (70 000) und PDaF (77 000) vergleichbar ist.

genauer analysiert, ob die Kollokationen unter dem nominalen oder/ und dem verbalen Stichwort verzeichnet sind.

Für einen Eintrag der Kollokation beim verbalen Kollokator spricht, dass dieser ohne die Angabe der kollokierenden Basen gar nicht in all seinen Lesarten erfasst werden könnte (Bahns 1996: 39). Zudem kann ein solcher Eintrag helfen, "synonyme Lexeme in ihrer Bedeutung und Verwendungsweise zu differenzieren" (Bahns 1996: 40). Häufig erschweren gerade die Kollokatoren das Verständnis von Kollokationen (Jehle 1990: 262), so dass aus rezeptiver Sicht ein Eintrag unter dem Kollokator wünschenswert wäre (ebd.). Wörterbücher, in denen unter Kollokatoren "Informationen über mögliche Basen und Basisklassen" (Hausmann 2007: 219) eingetragen sind, sind kognitionsorientiert (ebd.).

Für einen Eintrag von Kollokatoren unter der jeweiligen Basis spricht, dass der L2-Wörterbuchbenutzer bei der Textproduktion weit häufiger nach einem passenden Kollokator zu einer ihm bereits bekannten Basis sucht, als umgekehrt (cf. Jehle 1990: 116). Demnach werden WB dieser Struktur auch als produktionsorientiert bezeichnet. Gesuchte Kollokatoren findet der Nutzer beim Nachschlagen unter dem ihm bekannten Autosemantikon (Hausmann 2007: 218). Zudem ist natürlich nicht auszuschließen, dass auch die Basis, obgleich autosemantisch und somit auch ohne Kollokator definierbar, gerade im Falle einer Kollokation Verständnisprobleme bereitet.

Es ist also eine Vielzahl produktiver und rezeptiver Nutzungssituationen denkbar, in denen der Eintrag einer Kollokation unter ihren beiden Einzelexemen nützlich wäre (cf. Bahns 1996, 110; Reder 2006: 183).

Folgt man der Annahme, dass lexikalische Kollokationen rezeptiv unproblematisch sind und der Lerner für produktive Bedürfnisse nach passenden Kollokatoren im LWB suchen wird, sollten die 56 Kollokationen zumindest unter den nominalen Basen zu finden sein (cf. Jehle 1990: 263). Tendenziell scheinen die LWB diesem Prinzip durchaus zu folgen (cf. Abb. 2), denn von allen Kollokationen, die **nur** unter der Basis **oder** dem Kollokator in den LWB eingetragen sind, finden sich durchweg mehr ausschließlich *unter der Basis*. In allen LWB finden sich jedoch auch ausschließlich *unter dem Kollokator* eingetragene Kollokationen und sowohl unter Basis als auch unter dem Kollokator eingetragene Mehrwortverbindungen. Diese Uneinheitlichkeit ist für den Benutzer eines LWB sehr nachteilig, denn er kann sich nicht darauf verlassen, die Kollokation entweder unter dem Nomen, unter dem Verb oder gar unter beiden finden zu können. Die Werte zeigen, dass ein eindeutiges Konzept hinsichtlich des Eintragungsorts nicht vorhanden ist.

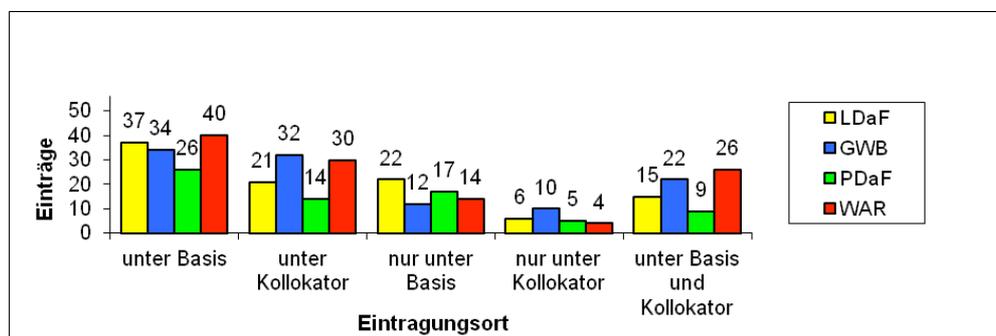


Abb. 2: Darstellung der Kollokationen nach Eintragungsort

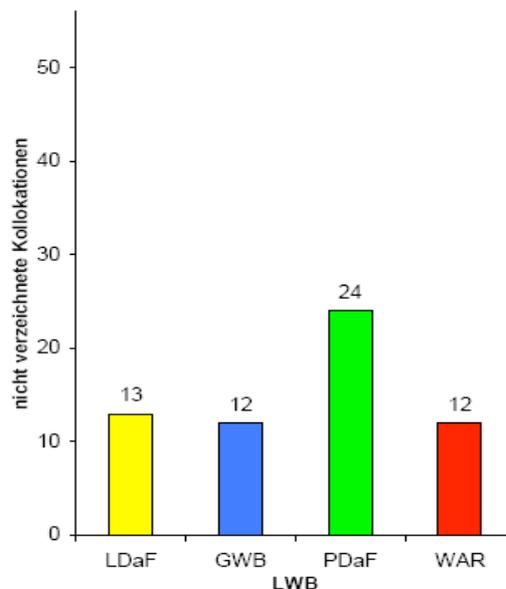


Abb. 3: Anzahl der nicht in den LWB eingetragenen Kollokationen

In allen LWB fehlen einige der 56 Kollokation (cf. Abb.3), wobei lediglich vier der hochfrequenten Kollokationen (*Entwicklung sehen, Geld stecken, Grund sehen, Verhältnis liegen*) in keinem der LWB verzeichnet sind, alle anderen sind in mindestens einem der LWB eingetragen. Dieser Wert zeigt zum einen, dass die durch die aufwendige Kollokationsidentifikation gefundenen Mehrwortverbindungen von den Wörterbuchautoren ebenfalls als relevant erachtet werden, wenn auch auf nicht nachvollziehbar unterschiedliche Art und Weise. Zum anderen wird klar, dass die insgesamt überhaupt nicht im LWB verzeichneten Kollokationen (cf. Abb. 3), wovon es bis zu 24 von 56 gibt, ein tatsächlich großes Desiderat bei der Entwicklung aktueller LWB darstellen. Fünf Kollokationen sind in jedem der untersuchten LWB sowohl unter der Basis als auch unter dem Kollokator eingetragen (*Geld sparen, Geld verdienen, Hand halten, Hand schütteln, Hause gehen*). Diese Schnittmenge sollte wesentlich größer ausfallen, wenn alle LWB auf der Basis großer, ausgewogener und repräsentativer Korpora erstellt worden wären.

5 Fazit

Es wurde für einen repräsentativen Lexikonausschnitt gezeigt, wie hochfrequente und somit zum Grundwortschatz gehörige Kollokationen im Korpus datengeleitet identifiziert werden können. Diese Auswahlmethode ermöglicht es, systematisch für den Fremdsprachenlerner tatsächlich relevante Mehrwortverbindungen zu benennen. Im Anschluss wurden die vier großen LWB des Bereichs Deutsch als Fremdsprache auf die Erfassung dieser Kollokationen hin untersucht und verglichen.

In der Wörterbuchanalyse beschränkt sich der geleistete Beitrag v.a. auf quantitative Aspekte der lexikographischen Erfassung. Dies lässt eine abschließende Bewertung, welches der LWB eine besonders/weniger gelungene Kollokationsdarstellung aufweist, nicht zu. Die durchweg unvollständige und uneinheitliche Verzeichnung der betrachteten Kollokationen impliziert jedoch für die weitere lexikographische Arbeit Forderungen, wie sie bereits Heine (2005: 353) für den verwandten Bereich der Funktionsverbgefüge gestellt hat:

Für weitere Wörterbücher bzw. Neuauflagen ist es wünschenswert, die mittlerweile vielfältigen Möglichkeiten, die die Arbeit mit elektronischen Korpora bietet, zu nutzen. Nur so können die bislang auf intuitiver und damit subjektiver Muttersprachen-Kompetenz von Autoren und Informanten beruhenden Angaben im Wörterbuch [...] durch objektivere, dem tatsächlichen Sprachgebrauch entsprechende, ergänzt oder ersetzt werden.

Auch Steyer (2008: 192) fordert, dass "[j]edem lexikografischen Arbeitsprozess [...] heute Korpusanalysen und im Idealfall auch Kookkurrenzanalysen vorgelagert sein" sollten. Insbesondere für den in diesem Beitrag untersuchten Bereich des Grundwortschatzes, welcher für den Lerner auf einem Sprachniveau erwerbsrelevant ist, auf dem er das LWB seines Vertrauens häufig nutzt, zeigen die Resultate, dass lexikographisch nachgebessert werden muss.

Dabei gilt es, neben der Auswahl und Aufnahme der relevanten Kollokationen auch eine konsequente, transparente und für den Lerner nützliche Lösung bezüglich des Eintragungsortes zu finden. Dem unter 4. geäußerten Vorschlag, Kollokationen sowohl unter der Basis als auch unter dem Kollokator zu verzeichnen, setzt der beschränkte Umfang von Printwörterbüchern eine natürliche Grenze. Eine Kollokation benötigt im Wörterbuch doppelt so viel Raum, wenn sie unter beiden Komponenten eingetragen ist. Gebräuchliche Komprimierungsverfahren wie bspw. die lexikographische Beschränkung auf das beteiligte Substantiv (cf. Steyer 2000, 106f.) sind jedoch kaum geeignet, bestehende Lernerbedürfnisse zu befriedigen. Konsequente Verweise zwischen den einzelnen Einträgen böten an dieser Stelle einen, wenn auch nicht unproblematischen, Kompromiss zwischen Platzökonomie und umfassender Registrierung an (ebd.), doch zumindest für den Bereich des Grundwortschatzes sollte eine umfassende Verzeichnung der Kollokationen angestrebt werden. Die Rechtfertigung unvollständiger Verzeichnung wichtiger sprachlicher Phänomene durch räumliche Begrenzung ist in Zeiten aufkommender elektronischer Wörterbücher jedoch ohnehin ein scheidendes Argument (cf. Haß-Zumkehr 2001: 378f.).

Die angesprochenen lexikographischen Fragestellungen *Welche Kollokationen sollten im LWB eingetragen sein?* und *Wo sollten diese Kollokationen verzeichnet sein?* müssen dem *Wie?* (cf. Kap.4) voran gestellt sein und jeder dieser Aspekte bedarf zweifellos weiterer wissenschaftlicher Betrachtung. Es ist wünschenswert, dass all diese Punkte in der weiteren lernerlexikographischen Arbeit stärkere Beachtung finden um den rezeptiven und produktiven Bedürfnissen der Lerner gerecht zu werden.

Literatur

Wissenschaftliche Literatur

- Bahns, Jens (1993): "Kollokation kontra Kontext. Wider ein zu weites Verständnis des Kollokationsbegriffs". *Praxis des neusprachlichen Unterrichts* 40: 30–37.
- Bahns, Jens (1996): *Kollokationen als lexikographisches Problem. Eine Analyse allgemeiner und spezieller Lernerwörterbücher des Englischen*. Tübingen: Niemeyer. (=Lexicographica Series maior 74).
- Bubenhof, Noah (2009): *Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse*. Berlin: de Gruyter. (=Sprache und Wissen).
- Engelberg, Stefan/Lemnitzer, Lothar (2001): *Lexikographie und Wörterbuchbenutzung*. Tübingen: Stauffenburg-Verlag. (=Stauffenburg Einführungen 14).
- Geyken, Alexander (2004): "Korpora als Korrektiv für einsprachige Wörterbücher". *Zeitschrift für Literaturwissenschaft und Linguistik* 136: 72–100.
- Glaboniat, Manuela (2010): *Profile deutsch. Gemeinsamer europäischer Referenzrahmen; Lernzielbestimmungen, Kannbeschreibungen, kommunikative Mittel; Niveau A1 - A2, B1 - B2, C1 - C2*. Berlin: Langenscheidt.
- Haß-Zumkehr, Ulrike (2001): *Deutsche Wörterbücher - Brennpunkt von Sprach- und Kulturgeschichte*. Berlin/New York: de Gruyter.
- Hausmann, Franz Josef (1984): "Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen". *Praxis des neusprachlichen Unterrichts* 4/31: 395–406.

- Hausmann, Franz Josef (2004): "Was sind eigentlich Kollokationen?". In: Steyer, Kathrin (ed.): *Wortverbindungen - mehr oder weniger fest*. Berlin, de Gruyter: 309–334.
- Hausmann, Franz Josef (2007): "Die Kollokationen im Rahmen der Phraseologie. Systematische und historische Darstellung". *Zeitschrift für Anglistik und Amerikanistik* 3/55: 217–234.
- Heine, Antje (2004): "*Funktionsverbgefüge in System, Text und korpusbasierter (Lerner-)Lexikografie*." *Dissertationsschrift zur Erlangung des Doktorgrades Doctor philosophiae (Dr. phil.)*. Leipzig: Universität Leipzig.
- Heine, Antje (2005): "Funktionsverbgefüge im Lernerwörterbuch. Ein Vergleich von "Langenscheidt Großwörterbuch Deutsch als Fremdsprache" (2003) und "De Gruyter Wörterbuch Deutsch als Fremdsprache." In: Barz, Irmhild/Bergenholtz, Henning/Korhonen, Jarmo (eds.): *Schreiben, Verstehen, Übersetzen, Lernen. Zu ein- und zweisprachigen Wörterbüchern mit Deutsch*. Frankfurt am Main, P. Lang: 345–356.
- Herbst, Thomas (1996): "What are Collocations: Sandy Beaches or False Teeth?". *English Studies* 77: 379–393.
- Jehle, Günter (1990): *Das englische und französische Lernerwörterbuch in der Rezension. Theorie und Praxis der Wörterbuchkritik*. Tübingen: Niemeyer. (=LexicographicaSeries maior 30).
- Karacic, Geriena (2004): "Verträglichkeit und Unverträglichkeit der Kollokate". In: Bracic, Stojan (ed.): *Linguistische Studien im Europäischen Jahr der Sprachen. Akten des 36. Linguistischen Kolloquiums in Ljubljana 2001 = Linguistic studies in the European year of languages*. Frankfurt am Main, Lang: 291–300.
- Konecny, Christine (2010): "Lexikalische Kollokationen und der Beitrag der Internet-Suchmaschine Google zu ihrer Erschließung und Beschreibung". In: Ptashnyk, Stefaniya/Hallsteinsdóttir, Erla/Bubenhofer, Noah (eds.): *Korpora, Web und Datenbanken. Computergestützte Methoden in der modernen Phraseologie und Lexikographie = Corpora, Web and Databases: computer-based methods in modern phraseology and lexicography*. Baltmannsweiler, Schneider Verlag Hohengehren: 77–94.
- Köster, Lutz/Neubauer, Fritz (2002): "Kollokationen und Kompetenzbeispiele im De Gruyter Wörterbuch Deutsch als Fremdsprache". In: Wiegand, Herbert Ernst (ed.): *Perspektiven der pädagogischen Lexikographie des Deutschen II. Untersuchungen anhand des "de Gruyter Wörterbuchs Deutsch als Fremdsprache"*. Tübingen, Niemeyer: 283–310.
- Kromann, Hans-Peder (1989): "Zur Funktionalen Beschreibung von Kollokationen in Übersetzungswörterbüchern". In: Gréciano, Gertrud (ed.): *Europhras 88. Phraséologie contrastive ; actes du Colloque international, Klingenthal-Strasbourg, 12 - 16 mai 1988*. Strasbourg, Univ. des Sciences Humaines Dép. d'Etudes Allemandes: 265–271.
- Lehr, Andrea (1998): "Kollokationen in Langenscheidts Großwörterbuch Deutsch als Fremdsprache". In: Wiegand, Herbert Ernst (ed.): *Perspektiven der pädagogischen Lexikographie des Deutschen. Untersuchungen anhand von "Langenscheidts Grosswörterbuch Deutsch als Fremdsprache"*. Tübingen, Niemeyer: 256–281.
- Markus, Tuulikki/Korhonen, Jarmo (2005): "Kollokationen in der deutschen Lernerlexikographie und in deutsch-finnischen Wörterbüchern". In: Barz, Irmhild/Bergenholtz, Henning/Korhonen, Jarmo (eds.): *Schreiben, Verstehen, Übersetzen, Lernen. Zu ein- und zweisprachigen Wörterbüchern mit Deutsch*. Frankfurt am Main, P. Lang: 327–344.
- McEnery, Tony/Tono, Yukio/Xiao, Richard (2006): *Corpus-based language studies. An advanced resource book*. London: Routledge. (=Routledge applied linguistics).
- Nesselhauf, Nadja (2005): *Collocations in a learner corpus*. Amsterdam etc.: Benjamins. (=Studies in corpus linguistics 14).
- Neubert, Albrecht (1966): "Analogien zwischen Phonologie und Semantik". In: Hintze, Fritz/Meier, Georg Friedrich/Seidel, Eugen (eds.): *Zeichen und System der Sprache*.

- Veröffentlichung des 2. Internationalen Symposions "Zeichen und System der Sprache " vom 8. 9. bis 15.9.1964 in Magdeburg. Berlin, Akademie Verlag: 106–116.
- Reder, Anna (2001): "Ein heißer Tipp - oder Kollokationslernen durch eine Ganzschrift". In: Häcki Buhofer, Annelies/Durco, Peter (eds.): *Wortschatz: Aneignung und Unterricht. IDT Publikation 2001*. <http://www.sprachwissenschaft.ch/IDT2001/>, Stand 30.05.2010
- Reder, Anna (2004): "Eine Frage aufwerfen oder: Kollokation als lexikographisches Problem in Wörterbüchern". In: Bracic, Stojan (ed.): *Linguistische Studien im Europäischen Jahr der Sprachen. Akten des 36. Linguistischen Kolloquiums in Ljubljana 2001 = Linguistic studies in the European year of languages*. Frankfurt am Main, Lang: 531–542.
- Reder, Anna (2006): *Kollokationen in der Wortschatzarbeit*. Wien: Praesens-Verlag.
- Scherfer, Peter (2001): "Zu einigen wesentlichen Merkmalen lexikalischer Kollokationen". In: Lorenz-Bourjot, Martine/Lüger, Heinz-Helmut (eds.): *Phraseologie und Phraseodidaktik*. Wien, Ed. Praesens: 4–19.
- Steinbügl, Birgit (2005): *Deutsch-englische Kollokationen. Erfassung in zweisprachigen Wörterbüchern und Grenzen der korpusbasierten Analyse*. Tübingen: Niemeyer. (=LexicographicaSeries maior 126).
- Steyer, Kathrin (2000): "Usuelle Wortverbindungen des Deutschen. Linguistisches Konzept und lexikografische Möglichkeiten." *Deutsche Sprache* 2/28: 101–125.
- Steyer, Kathrin (2008): "Kollokationen in deutschen Wörterbüchern und in der deutschen Wörterbuchforschung". In: Heid, Ulrich et al. (eds.): *Lexicographica 24. International Annual for Lexicography - Revue Internationale de Lexicographie - Internationales Jahrbuch für Lexikographie*. Berlin, New York: de Gruyter: 185–208.
- Tschirner, Erwin (2005): "Korpora, Häufigkeitslisten, Wortschatzerwerb". In: Heine, Antje/Hennig, Mathilde/Tschirner, Erwin (eds.): *Deutsch als Fremdsprache - Konturen und Perspektiven eines Faches. Festschrift für Barbara Wotjak zum 65. Geburtstag*. München, Iudicium: 133–152.
- Tschirner, Erwin (2008a): "Das professionelle Wortschatzminimum im Deutschen als Fremdsprache". *Deutsch als Fremdsprache* 4/45: 195–208.
- Viehweger, Dieter (1989): "Probleme der Beschreibung semantischer Vereinbarkeitsrelationen im allgemeinen einsprachigen Wörterbuch". In: Hausmann, Franz Josef et al. (eds.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Berlin; New York: de Gruyter: 888–893.
- Wallner, Franziska (2010): "Kollokationen in Wissenschaftssprachen: Zur lernerlexikographischen Relevanz der Textarten- und Diskursspezifik von Kollokationen". In: Ptashnyk, Stefaniya/ Hallsteinsdóttir, Erla /Bubenhof, Noah (eds.): *Korpora, Web und Datenbanken. Computergestützte Methoden in der modernen Phraseologie und Lexikographie = Corpora, Web and Databases: computer-based methods in modern phraseology and lexicography*. Baltmannsweiler, Schneider Verlag Hohengehren: 197–214.

Wörterbücher

- Balhar, Susanne/Cyffka, Andreas (2004): PONS Großwörterbuch Deutsch als Fremdsprache. Das neue Wörterbuch für Alltag, Unterricht, Studium und Beruf. Barcelona, etc.: Klett Sprachen.
- Götz, Dieter/Haensch, Günther/Wellmann, Hans (2007): Langenscheidt Großwörterbuch Deutsch als Fremdsprache. Das einsprachige Wörterbuch für alle die Deutsch lernen. Berlin, etc.: Langenscheidt.
- Jones, Randall L./Tschirner, Erwin (eds.) (2006): *A frequency dictionary of German. Core vocabulary for learners*. London: Routledge. (=Routledge frequency dictionaries).
- Kempcke, Günter (2000): *Wörterbuch Deutsch als Fremdsprache*. Berlin [u. a.]: de Gruyter.

Tschirner, Erwin (2008b): Grund- und Aufbauwortschatz Deutsch als Fremdsprache nach Themen. Berlin: Cornelsen. (=Lextra).

Wahrig-Burfeind, Renate (2008): Großwörterbuch Deutsch als Fremdsprache. Gütersloh: Wissen-Media-Verlag.

Anhang 1

Sachgruppe	Rank	Basis	Kookkurrenz	log-Likelihood	Frequenz des Bigramms
1. Personalien, Informationen zur Person	103	Frau	sagen	134'165'791	3000
			sehen	37'852'908	1203
			kommen	35'389'946	1162
			lieben	31'340'857	493
			sitzen	28'414'541	586
			gehen	26'597'412	881
			fragen	25'433'796	586
			müssen	25'003'271	1069
			stehen	22'724'778	752
			geben	22'320'083	870
			sollen	<u>21'944'878</u>	876
			lassen	20'773'276	780
			wissen	<u>19'928'604</u>	653
			sprechen	17'557'015	504
			rufen	<u>17'267'014</u>	377
			erzählen	15'953'965	338
			nehmen	15'901'682	545
			denken	<u>14'964'795</u>	440
			kennen	11'388'860	306
			dürfen	<u>11'246'830</u>	393
bitten	11'079'199	256			
2. Wohnen	159	Haus	kommen	119'387'412	2307
			gehen	<u>105'769'375</u>	1968
			verlassen	<u>44'602'192</u>	612
			stehen	36'243'286	882
			wohnen	<u>34'643'120</u>	429
			bringen	<u>33'821'387</u>	715
			bleiben	29'933'281	700
			fahren	<u>29'816'492</u>	541
			bauen	<u>28'145'964</u>	380
			müssen	26'629'224	935
			schicken	21'469'468	310
			lassen	18'488'768	617
			sehen	17'864'077	645
			sitzen	<u>15'179'840</u>	333
			liegen	14'198'020	421
			tragen	13'922'524	325
			sollen	13'610'367	557
			treten	13'036'831	323
			nehmen	12'814'747	407
			laufen	<u>12'721'895</u>	268
3. Umwelt	186	Stadt	liegen	19'058'726	472
			fahren	16'097'847	324
			kommen	13'831'984	507
			gehen	13'305'709	447
			verlassen	13'100'391	227
			geben	10'235'037	420
			müssen	8'249'442	432

			sollen	7'628'871	362
			sehen	<u>7'578'911</u>	355
			wohnen	<u>7'494'920</u>	124
			zerstören	7'164'856	112
			besetzen	6'672'069	102
			ziehen	6'136'037	159
			nehmen	5'680'988	226
			bauen	5'213'734	100
			entstehen	4'507'893	120
			bleiben	4'376'837	190
			führen	<u>4'217'431</u>	160
			bringen	4'178'912	169
			stehen	4'122'672	212
4. Reisen und Verkehr	104	Mensch	geben	59'817'734	1692
			müssen	54'070'781	1791
			leben	<u>43'381'348</u>	786
			sehen	<u>38'018'347</u>	1271
			sollen	<u>31'918'789</u>	1158
			wissen	27'491'643	842
			kommen	<u>24'437'007</u>	982
			sagen	<u>24'208'994</u>	1065
			stehen	<u>22'248'857</u>	783
			töten	19'805'970	278
			kennen	17'353'912	425
			glauben	16'761'202	454
			gehen	16'485'286	700
			tuen	16'214'948	515
			denken	16'170'908	486
			lassen	15'385'256	695
			lieben	<u>14'581'063</u>	290
			finden	14'486'459	469
			brauchen	<u>14'159'703</u>	375
			sprechen	14'055'984	460
5. Verpflegung	970	Essen	kommen	7'132'043	175
			bringen	6'440'157	123
			einladen	6'292'047	65
			geben	6'204'988	158
			schlafen	<u>5'067'440</u>	41
			schmecken	4'491'352	46
			gehen	4'338'075	115
			kochen	<u>4'229'942</u>	45
			müssen	<u>3'962'158</u>	133
			sollen	3'789'364	116
			bekommen	<u>3'728'006</u>	62
			laden	3'614'370	42
			holen	3'234'165	47
			lassen	3'110'902	95
			servieren	3'105'264	30
			sitzen	2'879'252	57
			sagen	2'175'385	89
			warten	1'970'552	35
			stehen	1'809'865	62
			trinken	<u>1'774'439</u>	29
6. Einkaufen	251	Geld	verdienen	73'555'996	679
			kosten	46'093'457	450
			geben	27'018'103	623

			brauchen	26'771'387	401
			ausgeben	<u>23'819'685</u>	235
			bekommen	21'519'993	311
			müssen	<u>20'283'032</u>	573
			nehmen	<u>16'315'182</u>	359
			sparen	<u>15'549'166</u>	153
			bezahlen	12'999'653	160
			sollen	12'756'648	383
			leihen	12'517'578	116
			kommen	12'037'943	367
			kaufen	10'191'411	142
			bringen	<u>10'150'306</u>	238
			erhalten	9'463'481	194
			kriegen	<u>9'000'017</u>	113
			schicken	8'872'731	129
			stecken	<u>7'909'892</u>	120
			lassen	7'874'413	262
7. Öffentliche und private Dienstleistungen	350	System	müssen	11'699'869	401
			bestehen	6'639'162	153
			entwickeln	6'624'395	118
			geben	6'574'460	248
			bilden	6'545'580	137
			sollen	5'537'791	228
			schaffen	4'279'063	99
			aufbauen	3'482'540	54
			bringen	3'336'839	114
			liegen	3'232'322	120
			ermöglichen	3'187'720	58
			einführen	<u>3'121'383</u>	52
			stellen	3'117'221	106
			darstellen	2'899'562	63
			bezeichnen	2'791'176	66
			führen	2'765'824	96
			entstehen	2'539'979	66
			lassen	2'538'782	135
			unterscheiden	2'359'654	51
			funktionieren	2'358'759	33
8. Körper und Gesundheit	179	Hand	legen	212'309'551	2573
			halten	174'211'641	2614
			nehmen	<u>141'743'057</u>	2336
			drücken	96'994'385	1081
			liegen	<u>93'005'137</u>	1671
			reichen	83'769'580	968
			heben	<u>78'463'965</u>	976
			strecken	71'667'935	695
			geben	64'578'672	1605
			küssen	60'909'150	628
			gehen	60'688'257	1427
			falten	53'226'646	439
			fallen	49'957'769	865
			greifen	<u>47'884'932</u>	634
			schütteln	45'103'130	527
			lassen	44'542'378	1186
			stützen	42'565'957	510
			schlagen	41'031'143	659
			fassen	<u>38'222'856</u>	560

9. Wahrnehmung und Bewegung	323	Sinn	pressen	30'889'907	335
			verstehen	<u>28'070'081</u>	487
			geben	15'307'040	526
			kommen	<u>14'321'060</u>	510
			müssen	10'260'234	478
			sprechen	10'086'124	276
			verlieren	<u>9'215'899</u>	188
			sollen	9'060'041	392
			liegen	8'558'906	276
			gelten	<u>8'303'758</u>	202
			wirken	7'182'165	153
			handeln	<u>6'298'882</u>	154
			brauchen	6'263'228	165
			lassen	5'834'470	280
			beeinflussen	<u>5'296'780</u>	88
			interpretieren	4'960'366	62
			erfassen	<u>4'882'400</u>	84
			heissen	4'793'675	145
			auffassen	4'752'884	70
			bezeichnen	4'747'436	115
nehmen	4'649'397	198			
10. Ausbildung	89	Beispiel	zeigen	51'543'569	769
			geben	44'418'599	929
			folgen	23'091'582	352
			anführen	20'207'047	197
			mögen	15'242'430	312
			nennen	13'876'388	257
			bieten	13'484'252	216
			erläutern	13'336'350	141
			sollen	13'258'467	418
			nehmen	<u>13'243'727</u>	328
			lassen	9'679'660	318
			finden	7'932'546	205
			verdeutlichen	<u>6'810'306</u>	67
			liefern	<u>6'741'586</u>	109
			sehen	6'389'277	265
			vorangehen	6'101'906	62
			dienen	5'634'114	114
			belegen	4'941'048	68
			wählen	4'712'544	92
			gelten	4'569'351	118
illustrieren	4'532'859	46			
11. Arbeitswelt	167	Arbeit	leisten	61'605'713	757
			verrichten	33'035'049	285
			müssen	27'320'320	934
			gehen	24'777'720	712
			beginnen	21'729'326	439
			sollen	19'253'251	673
			aufnehmen	18'145'448	294
			tuen	17'865'074	446
			geben	16'087'181	594
			kommen	11'825'874	501
			erleichtern	11'346'278	155
			lassen	<u>11'014'352</u>	444
			nehmen	11'002'087	364
			finden	10'537'504	308

			fortsetzen	9'787'074	152
			liegen	7'570'822	280
			beschäftigen	7'554'254	144
			schaffen	7'235'768	185
			beenden	7'139'496	113
			erledigen	6'519'020	101
12. Sprache	231	Grund	liegen	48'212'031	945
			geben	38'550'310	1016
			müssen	25'396'072	877
			nehmen	21'472'993	552
			legen	15'958'638	333
			bestehen	15'228'838	347
			angeben	14'360'188	194
			sollen	13'159'713	526
			vorliegen	12'812'303	171
			sehen	11'575'248	481
			erfolgen	11'063'563	214
			gehen	11'038'621	423
			dürfen	10'941'190	320
			annehmen	8'883'446	187
			kommen	8'587'781	410
			mögen	8'560'251	255
			wissen	7'430'818	294
			suchen	6'672'850	176
			lassen	6'568'599	325
			finden	6'446'916	222
13. Freizeit und Unterhaltung	204	Ende	gehen	51'564'336	1071
			bereiten	41'552'568	445
			führen	37'664'382	678
			nehmen	32'306'943	680
			setzen	24'039'822	461
			kommen	22'727'053	668
			müssen	22'446'406	742
			stehen	18'920'872	512
			finden	17'903'639	401
			sollen	17'791'017	575
			bringen	13'002'605	334
			betragen	12'724'678	199
			erreichen	11'312'240	229
			denken	10'027'390	262
			beginnen	9'216'034	221
			erfinden	8'483'091	193
			bedeuten	7'883'894	164
			sehen	7'691'650	344
			bleiben	7'385'524	250
			geben	5'957'199	298
14. Persönliche Beziehungen und Kontakte	305	Entwicklung	fördern	20'068'729	256
			laufen	11'804'685	225
			zeigen	11'346'034	273
			verfolgen	9'039'174	140
			beitragen	8'662'697	126
			hemmen	8'523'943	93
			vollziehen	8'439'918	124
			führen	8'254'703	227

			geben	7'976'995	335
			verlaufen	<u>7'833'764</u>	113
			nehmen	7'325'475	244
			müssen	6'554'686	349
			sollen	5'976'641	290
			stehen	5'775'084	235
			beeinflussen	5'484'724	87
			gehen	5'196'709	237
			beschleunigen	5'151'459	65
			bringen	<u>4'905'748</u>	172
			stellen	4'633'520	161
			sehen	4'466'401	244
15. Politik und Gesellschaft	160	Fall	müssen	38'283'984	1077
			sollen	23'463'789	708
			geben	20'389'614	638
			dürfen	19'378'071	447
			liegen	16'427'614	427
			handeln	14'062'500	274
			kommen	12'669'401	481
			bringen	12'322'576	332
			entscheiden	<u>11'437'108</u>	190
			lassen	10'523'612	399
			eintreten	<u>8'739'864</u>	151
			gelten	8'696'824	211
			bleiben	7'866'088	269
			zeigen	7'648'590	224
			mögen	7'074'334	214
			erledigen	6'207'583	93
			behandeln	6'207'583	125
			scheinen	6'207'583	167
			setzen	5'380'670	171
			gehen	5'101'504	256
			vorkommen	4'913'586	84
16. Allgemeine Begriffe	51	Jahr	laufen	84'814'961	1388
			betragen	60'630'913	886
			sollen	49'118'428	1768
			verurteilen	45'922'920	638
			sterben	<u>44'915'713</u>	741
			dauern	39'614'207	609
			kommen	33'806'130	1409
			beginnen	33'280'393	800
			vergehen	<u>30'663'262</u>	414
			leben	30'329'170	683
			bestrafen	28'117'969	360
			müssen	27'957'549	1455
			gehen	22'872'634	1007
			verlängern	22'614'666	283
			sehen	21'870'662	1095
			erreichen	20'712'864	516
			geben	20'068'356	1043
			arbeiten	19'296'141	454
			erhalten	19'107'725	560
			bleiben	<u>17'494'183</u>	704
			erscheinen	17'469'991	517

Anhang 2

Rank	Basis	Kookkurrenz	log-Likelihood	Frequenz des Bigramms
46	sagen	Herr	220570840	4354
		Frau	134165791	3000
		Mutter	103583018	1950
		Wort	85631719	2001
		Vater	83451514	1683
		Wahrheit	63520132	1088
		Stimme	56004976	1204
		Fräulein	43393301	705
		Gott	35065608	957
		Doktor	29545967	516
		Kind	29481714	1025
		Hand	28329409	1027
		Weile	26333728	453
		Mensch	24208994	1065
		Mama	22827866	358
		Papa	22806660	345
		Onkel	22645410	403
		Junge	22132676	478
		Tante	22039871	381
		Freund	22025627	584
57	geben	Antwort	153416191	2036
		Möglichkeit	127951172	1998
		Ausdruck	153416191	1796
		Wehrmacht	102610625	1170
		Anlass	97643662	1228
		Auskunft	95389697	1008
		Mühe	79716714	988
		Oberkommando	79086260	789
		Gelegenheit	76012012	1119
		Hand	64578672	1605
		Mensch	59817734	1692
		Hauptquartier	47651440	512
		Erklärung	45337837	847
		Befehl	44987944	698
		Beispiel	44418599	929
		Leute	43939663	963
		Zeit	40700010	1419
		Grund	38550310	1016
		Aufschluss	38427673	399
		Regierung	37124590	1089
81	sehen	Auge	178515391	3042
		Abschnitt	110688955	1414
		Gesicht	104936143	1806
		Seite/S.	66983037	1507
		Fenster	60879512	1006
		Mann	46027158	1358
		Wortlaut	44085264	548
		Mensch	38018347	1271
		Frau	37852908	1203
		Bild	34198662	840
		Blick	28878391	652
		Mal	27797781	574
		Herr	27479351	1017
		Übersicht	24956365	334

		Hand	24062534	852
		Abkommen	23768650	484
		Jahr	21870662	1095
		Kind	21355398	772
		Mutter	21140647	609
		Leute	21122942	587
82	lassen	Stich	71145850	715
		Ruhe	49941216	753
		Zeit	45984380	1439
		Hand	44542378	1186
		Zweifel	31950046	516
		Auge	27242197	776
		Wort	26701865	810
		Kind	23541633	768
		Herr	21737927	826
		Frau	20773276	780
		Wasser	19643469	495
		Kopf	18814764	538
		Tag	18756415	736
		Haus	18488768	617
		Mann	18387510	723
		Arm	17441212	422
		Blick	16746040	426
		Leben	16372444	629
		Mensch	15385256	695
		Minute	15277977	335
112	bleiben	Zeit	38722732	1059
		Haus	29933281	700
		Tag	24181233	698
		Leben	22336660	621
		Frage	20826824	601
		Jahr	17494183	74
		Weile	16874061	258
		Kraft	16026349	398
		Tür	13734359	304
		Erinnerung	13390293	242
		Nacht	12358243	288
		Gedächtnis	12303071	181
		Mensch	11868616	473
		Kind	11343202	396
		Erfolg	11288390	255
		Land	9672659	382
		Antwort	9656190	218
		Stunde	9292119	264
		Platz	9047001	228
		Rest	8957357	166
118	liegen	Bett	101597676	1229
		Hand	93005137	1671
		Interesse	53970610	903
		Grund	48212031	945
		Boden	42970723	703
		Ding	33147336	621
		Herz	27634282	505
		Tisch	27166692	476
		Schnee	23805476	308
		Verhältnis	22116333	460
		Schwerpunkt	21691941	233

		Auge	19780115	515
		Stadt	19058726	472
		Seite	18324346	491
		Tag	18145867	574
		Rücken	17339091	279
		Fall	16427614	427
		Sache	16383866	379
		Natur	16346592	344
		Straße	16258680	345
123	heißen	Name	2035987	40
		Gott	1463462	33
		Vorname	1194234	12
		Vorsitzende	1052403	20
		Herr	966955	33
		Menge	959753	18
		Wort	932702	28
		Mensch	882584	32
		Mann	872682	30
		Wahrheit	855157	17
		Leute	842901	21
		Kind	700008	24
		Sohn	691752	16
		Ding/Dinger	643258	17
		Dame	634097	14
		Vater	625088	18
		S./Seite	619011	22
		Sprache	605211	15
		Prophet	582440	8
		Philosoph	581825	9
124	denken	Mensch	16170908	486
		Frau	14964795	440
		Gott	14337119	336
		Augenblick	12524241	255
		Mann	11160087	364
		Zeit	10461278	398
		Herr	10226084	356
		Ende	10027390	262
		Vater	9380975	244
		Wort	9071987	286
		Leute	8408050	217
		Tag	8197442	303
		Mutter	8036255	218
		Kind	7453533	260
		Gedanke	7035423	180
		Ding	6421656	176
		Junge	5237881	116
		Mädchen	4848050	130
		Leben	4685704	202
		Fräulein	4262822	90
140	tun	Schritt	70995879	959
		Pflicht	45551025	589
		Gefallen	42794463	362
		Dienst	28546565	427
		Abbruch	25441919	244
		Arbeit	17865074	446
		Unrecht	16878691	211
		Schuldigkeit	16518082	130

		Mensch	16214948	515
		Genüge	15970925	150
		Gut(es)	14418995	233
		Sache	14104873	318
		Herr	13292584	444
		Mann	10783943	378
		Ding	10712009	261
		Einhalt	9553929	92
		Frau	8996276	340
		Gott	8271319	245
		Kind	7847465	285
		Leute	7612448	214
143	glauben	Gott	19471842	386
		Herr	16914363	447
		Mensch	16761202	454
		Leute	11755490	252
		Irrtum	7340773	102
		Frau	6646188	240
		Grund	6215082	190
		Vorsitzende	6083627	129
		Zeit	5678364	250
		Augenblick	4991360	124
		Wunder	4822929	81
		Regierung	4680916	177
		Volk	4072032	146
		Welt	3712835	151
		Wahrheit	3696182	90
		Leben	3670696	157
		Recht	3470337	124
		Sache	3384304	107
		Sieg	3326099	70