
Nutzung des DWDS-Wortprofils beim Aufbau eines lexikalischen Informationssystems zu deutschen Stützverbgefügen

Jörg Didakowski, Nadja Radtke
Berlin-Brandenburgische Akademie der Wissenschaften,
Technische Universität Dortmund
didakowski@bbaw.de, nadja.radtke@tu-dortmund.de

Abstract

Im Hinblick auf eine Zuarbeit für die lexikographische Arbeit an einem lexikalischen Informationssystem zu deutschen Stützverbgefügen wird das DWDS-Wortprofil vorgeschlagen. Mithilfe dieses Werkzeugs kann eine zeitintensive und mühsame Rechercharbeit über eine Textsuchmaschine vermieden werden, indem auf Basis eines ausgewogenen Korpus potenzielle Stützverbgefüge bereitgestellt werden. Des Weiteren wird die Einbeziehung von Assoziationsmaßen vorgeschlagen, um die Menge des zu sichtenden Materials für die lexikographische Arbeit weiter reduzieren zu können.

Keywords: Stützverbgefüge; Assoziationsmaße; computerlinguistische Verfahren

1 Einleitung

Neben den traditionellen Textkorpora stehen den Lexikographen heutzutage auch digitale Textkorpora zur Verfügung, die über Abfrage- und Analysewerkzeuge die Möglichkeiten der lexikographischen Arbeit stark erweitern (Engelberg & Lemnitzer 2009). Eines dieser Werkzeuge stellt das DWDS-Wortprofil dar, welches beim Aufbau eines lexikalischen Informationssystems zu deutschen Stützverbgefügen (das SVG-Wiki) hilft bzw. den Aufbau von diesem gar erst ermöglicht. Dieses Werkzeug ist Teil des Digitalen Wörterbuchs der deutschen Sprache (DWDS), eines Projekts der Berlin-Brandenburgischen Akademie der Wissenschaften. Ziel des vorliegenden Beitrags ist es, die Potenziale der Nutzung des DWDS-Wortprofils beim Aufbau des SVG-Wikis aufzuzeigen.

Im 2. Kapitel des Beitrags stellen wir zunächst das SVG-Wiki vor, dabei legen wir den Gegenstand fest, bestimmen die Zielgruppe, führen die Komponenten des SVG-Wikis ein und heben abschließend die korpusbasierte Erarbeitung des SVG-Wikis hervor. Im 3. Kapitel beschreiben wir das DWDS-Wortprofil, das für die Datenerhebung des SVG-Wikis grundlegend ist. Anschließend gehen wir im 4. Kapitel des Beitrags genauer auf die Nutzung des DWDS-Wortprofils beim Aufbau des SVG-Wikis ein.

2 Das SVG-Wiki: Ein lexikalisches Informationssystem zu deutschen Stützverbgefügen

Den Gegenstand des SVG-Wikis bilden Stützverbgefüge (SVG, engl. *support verb constructions*) des Deutschen wie z.B. *Anwendung finden*, *zur Anwendung kommen* und *Kritik üben*, die aus einem prädikativen Nomen (z.B. *Anwendung*) und einem semantisch blassen Stützverb (z.B. *finden* oder *kommen*) konstruiert werden¹. Abgegrenzt werden SVG von freien Konstruktionen (z.B. *zur Party kommen*) sowie von Idiomen (z.B. *zu Potte kommen*); beschrieben werden sie in Bezug auf ihre Systematik und hinsichtlich ihrer Leistungen (z.B. Pottelberge van 2001; Seifert 2004; Heine & Wotjak 2005; Storrer 2007).

SVG haben bereits Eingang in die Grammatiken, Wörterbücher und Lehrwerke des DaF-Unterrichts gefunden. Wie sie am besten eingeführt und behandelt werden, wird in der Forschungsliteratur intensiv diskutiert. Vermittelt werden sie bei der Arbeit mit dem Wortschatz sowie im Grammatikunterricht, indem einerseits ausgewählte SVG aufgeführt und andererseits ausgewählte grammatische und stilistische Eigenschaften der SVG beschrieben werden. Wie man dabei zu einer adäquaten Auswahl der SVG gelangt und wie Vollständigkeit und Systematik hinsichtlich der grammatischen und stilistischen Eigenschaften der SVG erreicht werden können, bleibt jedoch offen. Die zu vermitteln den SVG auszuwählen, um dann diese in Hinblick auf ihre Systematik behandeln zu können, fällt hier nach wie vor in das Aufgabenfeld der DaF-Lehrenden. Eine vor diesem Hintergrund hilfreiche Ressource zu deutschen SVG für die DaF-Lehrenden gibt es noch nicht. Das SVG-Wiki schließt diese Lücke, indem es die DaF-Lehrenden als zukünftige Nutzer vorsieht und aufbauend auf den bereits ausgereiften Vorschlägen für die Printlexikographie (Heine 2006; Heine 2008) eine Wörterbuchkomponente (Spezialwörterbuch) mit einer Grammatikkomponente verbindet. Das SVG-Wiki ist als ein digitales wikibasiertes Informationssystem realisiert², das den Nutzern im Internet zur freien Verfügung stehen wird und kontinuierlich erweitert werden soll (siehe Abbildung 1).

1 In der deutschen linguistischen Fachliteratur findet man unterschiedliche Termini wie *Nominalisierungsverbgefüge*, *Funktionsverbgefüge* und *Streckverbgefüge*, die ebenfalls unterschiedlich begrifflich gefasst sind. Bei der Bestimmung des Gegenstandes des SVG-Wikis bedienen wir uns einem möglichst weiten Begriff der SVG: Das prädikative Nomen steht dabei im Akkusativ oder kommt als Präpositionalphrase vor, es ist abstrakt, wird deverbale oder deadjektivisch gebildet, kann ein Fremdexem sein und idiomatisch verwendet werden.

2 Das Wikisystem (hier: MediaWiki) ermöglicht u.a., die jeweiligen Komponenten des SVG-Wikis kollaborativ auszubauen. So können z.B. die Nutzer die noch nicht im SVG-Wiki berücksichtigten Stützverben zur Erweiterung des SVG-Wikis auf den dafür vorgesehenen Seiten des SVG-Wikis vorschlagen.



Abbildung 1: Hauptseite des SVG-Wikis.

Ziel des SVG-Wikis ist es, die SVG in ihrem Grundbestand festzuhalten, die DaF-Lehrenden bei der Auswahl der SVG zu unterstützen und die SVG in Bezug auf ihre grammatischen und stilistischen Eigenschaften zu beschreiben.

Das Besondere an dem SVG-Wiki ist, dass die jeweiligen Komponenten korpusbasiert erarbeitet werden. Ausgegangen wird hierbei von einer Lemmaliste der Stützverben, die aus 23 ausgewählten Grammatiken und unter Berücksichtigung der Forschungsliteratur erstellt wurde; sie enthält ca. hundert Stützverben mit den für sie charakteristischen und für die Datenerhebung relevanten Merkmalen. So wird z.B. bei dem Stützverb *finden* eingetragen, dass dieses mit einem prädikativen Nomen im Akkusativ vorkommt. Anhand dieser Lemmaliste werden zunächst die prädikativen Nomina der jeweiligen Stützverben³ mithilfe des DWDS-Wortprofils ermittelt; von den ermittelten prädikativen Nomina ausgehend, besteht im weiteren Schritt der Erarbeitung die Möglichkeit, die im SVG-Wiki noch nicht berücksichtigten Stützverben ebenfalls mithilfe des DWDS-Wortprofils zu entdecken. Anschließend findet die lexikographische Beschreibung der Stützverben, der prädikativen Nomina und der SVG anhand des DWDS-Kernkorpus statt. Parallel dazu wird die Grammatikkomponente des SVG-Wikis erarbeitet. Das DWDS-Kernkorpus stellt somit die primäre Quelle für die Wörterbuchbasis der Wörterbuchkomponente dar und bildet gleichzeitig die Datengrundlage für die Grammatikkomponente des SVG-Wikis.

3 Kamber (2008) geht ebenfalls in seiner korpusbasierten Untersuchung zu den nominalen Prädikaten des Deutschen von den jeweiligen Verben aus.

3 Das DWDS-Wortprofil

Das DWDS-Wortprofil (Didakowski & Geyken 2013) stellt Kookkurrenzpaare für verschiedene grammatische Relationen wie z.B. Akkusativ-/Dativobjekt, Genitivattribut, Adjektivattribut und präpositionales Komplement/Modifizierer bereit. Die Kookkurrenzpaare werden mithilfe von computerlinguistischen Verfahren automatisch extrahiert. In Kilgarrieff et al. (2004) wird für die automatische Extraktion grammatischer Kookkurrenzpaare die flache *Sketch-Grammar* vorgeschlagen, mit der über reguläre Ausdrücke Kookkurrenzpaare für bestimmte grammatische Relationen extrahiert werden können. Ivanova et al. (2008) zeigen jedoch, dass es für das Deutsche sinnvoll ist, um gute Ergebnisse zu erzielen, auf eine reichhaltigere linguistische Analyse zurückzugreifen. Beim DWDS-Wortprofil wird für eine reichhaltigere Analyse die TAGH-Morphologie (Geyken & Hanneforth 2006) und der regelbasierte Parser SynCoP (Syntactic Constraint Parser, Didakowski 2008) verwendet. So kann die relativ reichhaltige Morphologie und freie Wortstellung im Deutschen angemessen behandelt und die Kookkurrenzen mit gewünschter Qualität extrahiert werden.

Im DWDS-Wortprofil sind die Kookkurrenzpaare mit Werten verschiedener Assoziationsmaße versehen. Derzeit werden drei verschiedene Assoziationsmaße berechnet: 1) die reine Frequenz, 2) das auf dem Dice-Koeffizienten basierende logDice-Maß (Rychlý 2008) und 3) das auf Mutual-Information basierende MI-log-Freq-Maß (Kilgarrieff & Tugwell 2002). Mithilfe dieser Maße können Kookkurrenzpaare nach Verbindungsstärke bzw. Anziehungskraft sortiert werden. Hierbei wird das Assoziationsmaß in der Regel so gewählt, dass die entsprechende Sortierung für eine bestimmte Aufgabe am geeignetsten ist (Evert 2008).

Die Korpusgrundlage für das DWDS-Wortprofil bilden das DWDS-Kernkorpus und verschiedene verbreitete Zeitungen (*Süddeutsche Zeitung, DIE ZEIT, Berliner Zeitung, DIE WELT, Der Tagesspiegel, Bild*). Das DWDS-Kernkorpus ist ein Referenzkorpus der *deutschen* Sprache des 20. Jahrhunderts und ist ausgeglichen bezüglich verschiedener Textsorten, die zudem gleichmäßig über das 20. Jahrhundert verteilt sind. Es umfasst über 100 Millionen laufende Wortformen (Tokens) und hat damit eine vergleichbare Größe wie das British National Corpus (Geyken 2007). Das DWDS-Kernkorpus stellt somit als ausgeglichenes Referenzkorpus das Herzstück der Korpusbasis des DWDS-Wortprofils dar und nimmt damit eine besondere Stellung ein. Die gesamte Korpusgrundlage des DWDS-Wortprofils umfasst ca. 1,7 Milliarden laufende Wortformen (Tokens) und reicht zeitlich vom Anfang des 20. Jahrhunderts bis heute. Die Kookkurrenzpaare sind hierbei für die gesamte Korpusbasis und auch für die einzelnen Subkorpora berechnet. So können Kookkurrenzpaare z.B. ausschließlich auf Basis des DWDS-Kernkorpus abgefragt werden.

Das DWDS-Wortprofil ist einerseits über die DWDS-Webseite und andererseits innerhalb der CLARIN-Infrastruktur über WebLicht (Hinrichs et al. 2010) zugänglich, wo es in Verarbeitungsketten integriert werden kann.

Ein Beispiel für eine DWDS-Wortprofil-Abfrage auf der DWDS-Webseite ist in Abbildung 2 zu sehen. Hier wurde das Verb *finden* für die Akkusativ-/Dativobjekt-Relation unter Verwendung des MI-log-

Freq-Maßes auf der Basis des DWDS-Kernkorpus abgefragt. Die relevanten Kookkurrenzpartner zu dem Verb *finden* werden als Wortwolke dargestellt. Je größer der Wert des Assoziationsmaßes eines Kookkurrenzpaares ist, desto größer wird der Kookkurrenzpartner in der Wolke dargestellt. Eine alternative Darstellungsform zu dieser Wortwolke ist die Tabellenansicht, in der die Kookkurrenzpartner nach dem Assoziationsmaß sortiert aufgelistet und genauere Informationen zu Wortkategorien und Assoziationswerten aufgeführt sind.

Hervorzuheben ist dabei, dass im DWDS-Wortprofil über die einzelnen Kookkurrenzpartner direkt auf die entsprechenden Korpusbelege zugegriffen werden kann. Erst dadurch wird eine sinnvolle lexikographische Arbeit möglich. In Abbildung 3 sind die Belege für das Kookkurrenzpaar *Anwendung finden* aufgeführt.

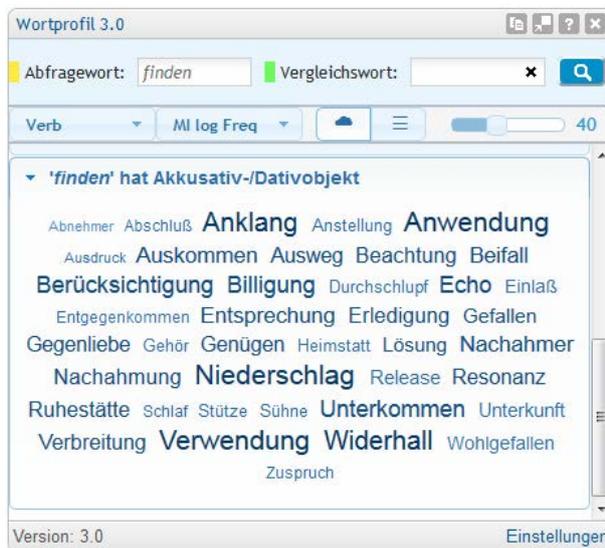


Abbildung 2: DWDS-Wortprofil-Wortwolke.

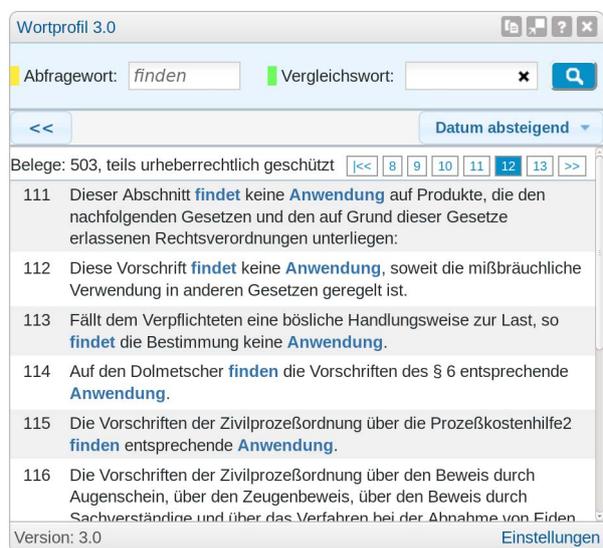


Abbildung 3: DWDS-Wortprofil-Belege.

Mit dem DWDS-Wortprofil ist es also möglich, auf strukturierte Weise Kookkurrenzpaare mit den dazugehörigen Korpusbelegen zu ermitteln. Verschiedene Assoziationsmaße können dazu verwendet werden, bestimmte Kookkurrenzpaare aus der Menge der Kookkurrenzpaare hervorzuheben. Die große Korpusbasis zusammen mit dem DWDS-Kernkorpus als ausgeglichenem Bestandteil gewährleistet hierbei ein breites Spektrum an Kookkurrenzen und repräsentative Ergebnisse.

4 Nutzung des DWDS-Wortprofils beim Aufbau des SVG-Wikis

Im Folgenden begründen wir die Wahl der Korpusbasis für die Ermittlung der prädikativen Nomina beim Aufbau des SVG-Wikis und gehen kurz auf verschiedene Möglichkeiten zur Ermittlung der prädikativen Nomina ein. Hierbei heben wir die Potenziale des DWDS-Wortprofils hervor. Im Weiteren

erläutern wir, ob die Assoziationsmaße dabei helfen können, die Menge der Kookkurrenzpaare für mögliche prädikative Nomina zu verkleinern, sodass weniger Kookkurrenzpaare gesichtet werden müssen und dabei trotzdem der Grundbestand der SVG festgehalten werden kann.

4.1 Ermittlung der prädikativen Nomina

Bei der Ermittlung der prädikativen Nomina wird die Korpusbasis auf das DWDS-Kernkorpus eingeschränkt, damit mit Blick auf die Zielsetzung des SVG-Wikis die Dekaden des 20. Jahrhunderts sowie verschiedene Textsorten gleichermaßen vertreten sind.

Zur Ermittlung der prädikativen Nomina kann einerseits die DWDS-Suchmaschine und andererseits das DWDS-Wortprofil genutzt werden. Bei der Nutzung der DWDS-Suchmaschine erhält man z.B. für die Abfrage zu dem Verb *finden* eine Liste mit 82.864 Treffern, in der nach den prädikativen Nomina manuell gesucht werden muss. Bei der Nutzung des DWDS-Wortprofils reduziert man bereits durch die Wahl einer grammatischen Relation die Menge der Treffer. So erhält man bei dem Verb *finden* durch die Wahl der Akkusativ-/Dativobjekt-Relation eine Liste mit 779 Kookkurrenzpaaren, die dann durch das Zugreifen auf einzelne Korpusbelege manuell nach prädikativen Nomina klassifiziert werden können. Hierbei beträgt die durchschnittliche Anzahl an Kookkurrenzpaaren für die 31 in den Grammatiken am häufigsten genannten Stützverben pro Verb mit dem Nomen im Akkusativ ca. 437 und pro Verb mit dem Nomen als Präpositionalphrase ca. 843. Hervorzuheben ist im Weiteren, dass dem DWDS-Wortprofil eine reichhaltigere linguistische Analyse (siehe Kapitel 3) zugrunde liegt und somit bestimmte Fälle, die die Suche mit der DWDS-Suchmaschine zusätzlich erschweren, vermieden werden. Zu solchen Fällen gehören Verben mit einem abtrennbaren Präfix (z.B. *ausüben*) oder Verben, die bezüglich einer Wortform homograph zu einem anderen Verb sind (z.B. *geraten* und *raten*). Somit ermöglicht das DWDS-Wortprofil, die Ermittlung der prädikativen Nomina überhaupt in einem realistischen Zeitrahmen bewältigen zu können.

4.2 Verwendung von Assoziationsmaßen bei der Ermittlung der prädikativen Nomina

Über das DWDS-Wortprofil ist es möglich, die abgefragten Kookkurrenzpaarlisten nach verschiedenen Assoziationsmaßen zu sortieren (siehe Kapitel 3). Hierbei ist die Frage, ob die Assoziationsmaße in der Lage sind, die Kookkurrenzpaarlisten so zu sortieren, dass am Anfang der Listen die Dichte der prädikativen Nomina sehr hoch ist und am Ende nur wenige prädikative Nomina vorkommen. So könnten die Kookkurrenzpaarlisten verkleinert werden, ohne dass zu viele prädikative Nomina verloren gehen. Auf diese Weise kann Recherchearbeit eingespart werden.⁴

4 Langer (2009) versucht, Funktionsverbgefüge vollautomatisch aus Korpora zu gewinnen. Dies will er hauptsächlich über Assoziationsmaße realisieren. Er zeigt, dass die Maße für so eine Aufgabenstellung

Um in Bezug zu der oben genannten Fragestellung eine Bewertung der Assoziationsmaße durchzuführen, folgen wir der Methodik in Evert et al. (2000) und beurteilen die Maße anhand von Precision und Recall. Hierzu wurden vier Verben herangezogen: *bringen*, *finden*, *kommen* und *üben*. Zu diesen Verben wurden mithilfe des DWDS-Wortprofils unter Berücksichtigung der jeweiligen grammatischen Relationen und durch die Wahl des DWDS-Kernkorpus als Korpusbasis Kookkurrenzpaarlisten ermittelt, wobei eine Minimalfrequenz für die Kookkurrenzpaare auf fünf festgelegt wurde. Die Listen wurden dann manuell vollständig gesichtet und nach prädikativen Nomina klassifiziert. Insgesamt wurden 859 prädikative Nomina aus 9.166 Kookkurrenzpaaren identifiziert. Die prädikativen Nomina machen demnach 27% der Gesamtmenge aus. Über die so erstellte Referenzmenge können nun Precision und Recall ermittelt werden. Zur Bewertung des Nutzens der Assoziationsmaße wird hier zusätzlich eine Zufallsortierung der Kookkurrenzlisten hinzugezogen. Die Zufallsliste deckt den Fall ab, dass kein Assoziationsmaß zur Sortierung verwendet wird.

Der Verlauf von Precision und Recall zu den einzelnen Assoziationsmaßen und der Zufallsliste ist in den Diagrammen in Abbildung 4 und 5 zu sehen. Auf den X-Achsen ist der Anteil an Kookkurrenzpaaren, der durch das Verkürzen der Kookkurrenzpaarlisten entsteht, in Prozent aufgetragen. Precision und Recall zu den einzelnen Anteilen sind jeweils auf den Y-Achsen in Prozent ablesbar.

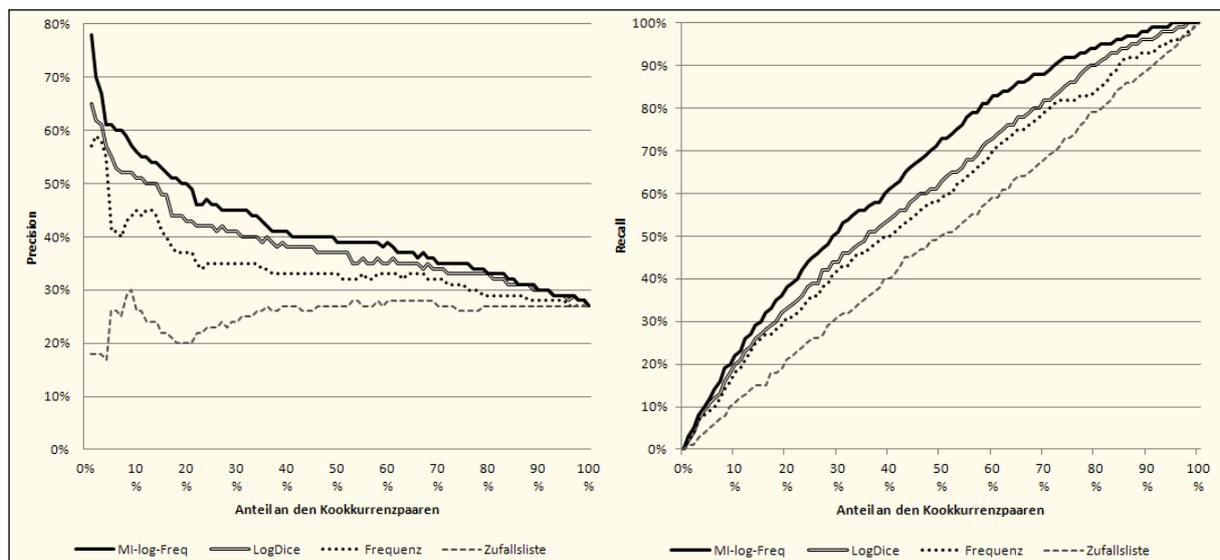


Abbildung 4: Precision-Kurven.

Abbildung 5: Recall-Kurven.

Der Verlauf der Precision-Kurven in Abbildung 4 zeigt, dass über die Sortierung nach dem MI-log-Freq-Maß die besten Precision-Werte erreicht werden. Hier liegt der Anteil der prädikativen Nomina sogar bei 78%, wenn 1% der Kookkurrenzpaare herangezogen wird. Mit anwachsendem Anteil an Kookkurrenzpaaren flachen die Precision-Kurven der Assoziationsmaße anfangs ab und fallen stetig auf das Grundniveau von 27%, welches durch den Anteil an prädikativen Nomina gesetzt ist. Bei der Zu-

nicht ausreichend sind. Die vollautomatische Extraktion, die Langer (2009) im Sinn hat, wird in unserer Vorgehensweise jedoch nicht verfolgt.

fallssortierung hingegen bewegt sich die Precision lediglich nahe am Grundniveau und sogar darunter. Der Verlauf der Recall-Kurven in Abbildung 5 zeigt ergänzend dazu, dass unter Verwendung des MI-log-Freq-Maßes nur die Hälfte der Kookkurrenzpaare gesichtet werden muss, um bereits 72% der prädikativen Nomina zu ermitteln. Sind die Kookkurrenzpaare nach dem Zufall sortiert, bekommt man hingegen lediglich ca. 50% der prädikativen Nomina.

Hier wird deutlich, dass die Assoziationsmaße hilfreich sind, wenn man bei der Ermittlung der prädikativen Nomina den Umfang und den damit verbundenen zeitlichen Aufwand reduzieren möchte und gleichzeitig möglichst viele prädikative Nomina als Grundbestand ermitteln will. Hierbei hat sich das MI-log-Freq-Maß als am geeignetsten herausgestellt.

5 Zusammenfassung

Das DWDS-Wortprofil hat das Potential, den Aufbau eines lexikalischen Informationssystems zu deutschen Stützverbgefügen entscheidend zu erleichtern. Die Aufgabe des DWDS-Wortprofils liegt bei dem Aufbau des SVG-Wikis darin, mögliche prädikative Nomina für Stützverben bereitzustellen. Dadurch wird eine zeitintensive und mühsame Recherche über eine Textsuchmaschine vermieden. Unter dem Aspekt zeitlicher Restriktionen ist das DWDS-Wortprofil sogar unabdingbar. Über eine zusätzliche Bewertung nach Assoziationsmaßen kann zudem weitere Recherchezeit eingespart werden.

6 Literaturhinweise

- Das digitale Wörterbuch der deutschen Sprache (DWDS). Accessed at: www.dwds.de [11/04/2014].
- Didakowski, J. (2008). Local Syntactic Tagging of Large Corpora Using Weighted Finite State Transducers. In A. Storrer, A. Geyken et al. (eds.) *Text Resources and Lexical Knowledge. Selected Papers from the 9th Conference on Natural Language Processing, KONVENS 2008*. Berlin et al.: Mouton de Gruyter, pp. 65-78.
- Didakowski, J., Geyken, A. (2013). From DWDS corpora to a German Word Profile – methodological problems and solutions. In *Network Strategies, Access Structures and Automatic Extraction of Lexicographical Information, 2nd Work Report of the Academic Network “Internet Lexicography” (OPAL - Online publizierte Arbeiten zur Linguistik X/2012)*. Mannheim: Institut für Deutsche Sprache, pp. 43-52.
- Engelberg, S., Lemnitzer, L. (2009). *Lexikographie und Wörterbuchbenutzung*. Tübingen: Stauffenburg.
- Evert, S., Heid, U. et al. (2000). Methoden zum qualitativen Vergleich von Signifikanzmaßen zur Kollokationsidentifikation. In W. Zühlke, Ernst G. Schukat-Talamazzini (eds.) *Sprachkommunikation, KONVENS 2000*. Berlin et al.: VDE, pp. 215-220.
- Evert, S. (2008). Corpora and collocations. In A. Lüdeling, M. Kytö (eds.) *Corpus Linguistics. An International Handbook of the Science of Language and Society*. Berlin: Mouton de Gruyter, pp. 1212-1248.
- Geyken, A., Hanneforth, T. (2006). TAGH: A Complete Morphology for German based on Weighted Finite State Automata. In A. Yli-Jyrä, L. Karttunen et al. (eds.) *Finite State Methods and Natural Language Processing*. Berlin et al.: Springer, pp. 55-66.

- Geyken, A. (2007). The DWDS corpus: A reference corpus for the German language of the 20th century. In Ch. Fellbaum (eds.) *Idioms and Collocations. Corpus-based Linguistic and Lexicographic Studies*. London et al.: Continuum, pp. 23-41.
- Heine, A., Wotjak, B. (2005). Zur Abgrenzung und Beschreibung verbonominaler Wortverbindungen (Wortidiome, Funktionsverbgefüge, Kollokationen). In *Deutsch als Fremdsprache. Zeitschrift für Theorie und Praxis des Deutschunterrichts für Ausländer*, 42(3), pp. 143-153.
- Heine, A. (2006). Funktionsverbgefüge in System, Text und korpusbasierter (Lerner-) Lexikographie. Frankfurt a. M. et al.: Peter Lang.
- Heine, A. (2008). Funktionsverbgefüge richtig verstehen und verwenden. Ein korpusbasierter Leitfaden mit finnischen Äquivalenten. Frankfurt a. M. et al.: Peter Lang.
- Hinrichs, E., Hinrichs, M. et al. (2010). WebLicht: Web-based LRT services for German. In *Proceedings of the ACL 2010, System Demonstrations (ACLDemos '10), Association for Computational Linguistics*. Stroudsburg, PA (USA), pp. 25-29.
- Ivanova, K., Heid, U. et al. (2008). Evaluating a German Sketch Grammar: A Case Study on Noun Phrase Case. In *Proceedings of the 6th International Conference on Language Resources and Evaluation, 2008*. Marrakech (Morocco), pp. 2101-2107.
- Kilgarriff, A., Tugwell, D. (2002). Sketching Words. In M.-H. Corréard (eds.) *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins, EURALEX*, pp. 125-137.
- Kilgarriff, A., Rychlý, P. et al. (2004). The Sketch Engine. In *Proceedings of EURALEX 2004*. Lorient (France), pp. 105-116.
- Kamber, A. (2008). Funktionsverbgefüge – empirisch. Eine korpusbasierte Untersuchung zu den nominalen Prädikaten des Deutschen. Tübingen: Max Niemeyer.
- Langer, S. (2009). Funktionsverbgefüge und automatische Sprachverarbeitung. München: LINCOM. *MediaWiki*. Accessed at: www.wikipedia.org [11/04/2014].
- Pottelberge van, J. (2001). Verbonominale Konstruktionen, Funktionsverbgefüge. Vom Sinn und Unsinn eines Untersuchungsgegenstandes. Heidelberg: Universitätsverlag C. Winter.
- Rychlý, P. (2008). A lexicographer-friendly association score. In P. Sojka, A. Horák (eds.): *Proceedings of Second Workshop on Recent Advances in Slavonic Natural Languages Processing, RASLAN 2008*. Brno: Masaryk University, pp. 6-9.
- Seifert, J. (2004). Funktionsverbgefüge in der deutschen Gesetzessprache (18. – 20. Jahrhundert). Hildesheim et al.: Georg Olms.
- Storrer, A. (2007). Corpus-based investigations on German support verb constructions. In Ch. Fellbaum (eds.) *Idioms and Collocations. Corpus-based Linguistic and Lexicographic Studies*. London et al.: Continuum, pp. 164-187.
- WebLicht*. Accessed at: <https://weblicht.sfs.uni-tuebingen.de> [11/04/2014].

