

## Ressourcen

Lea Hoffmann\*

# Online-Ressourcen zur Phraseologie des Deutschen

## Online Resources on Phraseology of German

<https://doi.org/10.1515/zgl-2018-0019>

Dieser Beitrag gibt einen Überblick über ausgewählte Online-Ressourcen für die Recherche und Analyse von deutschen Phrasemen bzw. Phraseologismen. Von den technischen Entwicklungen im World Wide Web konnte auch die Lexikografie profitieren: Digitale Nachschlagewerke und Online-Wörterbücher sind seit einiger Zeit in großer Zahl als frei zugängliche Angebote verfügbar und dabei fortlaufend erweiterbar – in Bezug auf ihren Umfang sind ihnen keine Grenzen gesetzt. Auch automatische Verfahren erleichtern es, große Sprachdatenmengen aus Korpora auf ihre Lexik hin zu untersuchen und für die Nutzerin oder den Nutzer sichtbar zu machen.

Die in diesem Beitrag dargestellten Online-Ressourcen nutzen diese Chancen mehr oder weniger umfangreich. Manche weiten die Idee des phraseologischen Wörterbuchs schlicht auf das Netz aus, andere wiederum lassen die Nutzerin oder den Nutzer selbst zum Lexikografen werden und lassen ihn die Daten, die sonst nur für die Wissenschaftlerinnen und Wissenschaftler nutzbar und auswertbar waren, selbst sehen und damit experimentieren. Besonders zu Sprichwörtern finden sich viele Angebote, die von Privatpersonen ohne institutionell-akademische Anbindung betrieben werden.<sup>1</sup> Für diesen Artikel wurden solche Projekte ausgewählt, deren Datenbankaufbau transparent dokumentiert wurde.

---

<sup>1</sup> Auf ein besonders interessantes Angebot, das eine Privatperson verwaltet, soll an dieser Stelle aufgrund des Detailreichtums hingewiesen werden: *schwäbisch schwätza* <<https://www.schwaebisch-schwaetza.de/>> ist eine vielseitige Online-Ressource zur schwäbischen Sprache und Kultur von Peter-Michael Mangold, in die ein über 11.000 Einträge umfassendes Wörterbuch zu schwäbisch-spezifischen Phrasemen inklusive authentischen Audio-Beispielen integriert ist.

---

\*Kontaktperson: **Lea Hoffmann**: Institut für Germanistik/Linguistik, Universität Duisburg-Essen, Berliner Platz 6–8, D-45127 Essen, E-Mail: [lea.hoffmann@uni-due.de](mailto:lea.hoffmann@uni-due.de)

# 1 Zu Sprichwörtern und Idiomen

Innerhalb des Portals für lexikografische Online-Ressourcen OWID des Instituts für Deutsche Sprache (IDS) in Mannheim wurde das *Sprichwörterbuch* <<https://www.owid.de/wb/sprw/start.html>> veröffentlicht, das zu der Projektgruppe *Usuelle Wortverbindungen (UWV)* gehört und von Kathrin Steyer u. a. entwickelt wurde. Das Wörterbuch stellt aktuell gängige Sprichwörter des Deutschen dar und besteht aus drei Modulen. Neben Gebrauchsbesonderheiten, Bedeutungserläuterungen, typischen Verwendungsweisen und anderen Angaben sind insbesondere Hyperlinks zu anderen Sprichwort-Einträgen vorhanden, die sich aus den Basiskomponenten des angezeigten Sprichworts ableiten. Das erste Modul *EU-Projekt SprichWort* umfasst 300 Artikel zu denjenigen Sprichwörtern, die innerhalb des multilingualen EU-Projekts *SprichWort. Eine Internetplattform für das Sprachenlernen* erarbeitet wurden. Diese wurden außerdem auf der *SprichWort-Plattform* <<http://www.sprichwort-plattform.org>> veröffentlicht, im Gegensatz zum *Sprichwörterbuch* jedoch mit anderem Fokus (Vergleichen der Sprichwort-Äquivalente im Slowenischen, Slowakischen, Tschechischen und Ungarischen). Über diese Sprichwörter hinaus wurden im Modul *Häufige Sprichwörter in DeReKo* weitere Sprichwörter aus dem *Deutschen Referenzkorpus (DeReKo)*<sup>2</sup> aufgenommen. Aufgrund ihrer ähnlichen Funktion wurden auch Werbeslogans im gleichnamigen Modul in das *Sprichwörterbuch* mit aufgenommen.

Die Seite <<http://www.ettinger-phraseologie.de/>> bietet eine Online-Version des mittlerweile vergriffenen Wörterbuchs zu Deutschen Redewendungen<sup>3</sup> von Regina Hessky und Stefan Ettinger, das zu einer Phrasendatenbank erweitert wurde. Dieses Angebot hat sich vor allem dem Lernen und Lehren von Phrasemen und Idiomen verschrieben. Die thematisch angeordneten Einträge der Redewendungen bestehen aus Bedeutungserklärungen und Anwendungsbeispielen. Besonders an dieser Ressource ist die Ausführlichkeit letzterer, die nicht nur durch einen umfangreichen Textausschnitt präsentiert werden, sondern auch deren Quelle mitsamt Hyperlink aufführen.

Das *Online Lexikon zur diachronen Phraseologie (OLdPhras)*, <<https://www.oldphras.net/web/>> erweitert das Spektrum an digitalen Sprichwörterbüchern um eine Ressource, die den Sprachwandel in den Blick nimmt. Entwickelt wurde eine Datenbank mit neuhochdeutschen Sprichwörtern und Redewendungen einschließlich diachroner Informationen unter der Leitung von Annelies Häcki

<sup>2</sup> Weiterführend zum *DeReKo* siehe auch Lüngen, Harald (2017): *DeReKo – Das Deutsche Referenzkorpus*. In: *ZGL* 2017; 45(1): 161–170.

<sup>3</sup> Hessky, Regina/Ettinger, Stefan (1997): *Deutsche Redewendungen. Ein Wörter- und Übungsbuch für Fortgeschrittene*, Tübingen.

Buhofer u. a. Die Artikel umfassen unter anderem Informationen zu Transformationen, Vorkommen in unterschiedlichen Textsorten und Bedeutungserklärungen der einzelnen Sprichwörter. Außerdem werden Belegzahlen aus verschiedenen Jahrhunderten (1650 bis 1900) geliefert sowie graphisch dargestellt und die kulturhistorische Entwicklung der Sprichwörter kurz beschrieben.

## 2 Zu Kollokationen

Das *Kollokationenwörterbuch*<sup>4</sup> von Buhofer u. a. findet man als Betaversion online unter <<https://kollokationenwoerterbuch.ch/web/>>. Bei dieser Online-Ressource handelt es sich, ähnlich zur *ettinger-Phraseologie*, um eine elektronische Erweiterung des Printangebots – auf der Seite kann das Wörterbuch durchsucht werden, sodass zu einem eingegebenen Stichwort eine Auswahl an Kollokationen angezeigt wird. Der Basiswortschatz umfasst 2.000 Lemmata, die durch die Ermittlung der Schnittmenge von verschiedenen Lern- bzw. Grundwortschätzen und Frequenzlisten aus Korpora entstanden sind. Die Kollokationen wurden empirisch ermittelt, indem Kookkurrenzanalysen in drei großen bereits bestehenden Korpora<sup>5</sup> durchgeführt und deren Ergebnisse aufbereitet wurden. Zielgruppe des Wörterbuchs sind Sprachlernende, wodurch sich die Auswahl und die Begrenzung des Wortschatzes erklären. Über die Angabe der Kollokationen geht das Wörterbuch jedoch nicht hinaus.

Eine im Gegensatz zur letztgenannten Ressource eher interaktiv angelegte Datenbank stellt die *Kookkurrenz-Datenbank ccbd* <<http://corpora.ids-mannheim.de/ccdb/>> von Cyril Belica dar. Die als „Experimentier- und Denkplattform“ ausgeschriebene Datenbank beruht auf einem Ausschnitt des *DeReKo* (davon ungefähr 2.2 Milliarden Textwörter) und enthält Kookkurrenzprofile zu ungefähr 220.000 Bezugswörtern. Zu jedem Bezugswort werden bis zu fünf Kookkurrenzprofile mit Verwendungsbeispielen in Form von KWIC-Kontexten dargestellt. Da die Daten vollautomatisch erzeugt wurden, müssen sie erst einer linguistischen Analyse unterzogen werden, um einschätzen zu können, ob es sich bei den Kookkurrenzprofilen tatsächlich um Phraseme handelt. Hierin besteht zugleich die besondere Stärke der *ccdb*, indem der Nutzer oder die Nutzerin so dazu angeleitet

---

<sup>4</sup> Buhofer, Annelies Häcki (2014): Feste Wortverbindungen des Deutschen: Kollokationen-Wörterbuch für den Alltag, Tübingen.

<sup>5</sup> Darunter fällt auch das *DWDS Kernkorpus*. Weiterführend zum *Digitalen Wörterbuch der Deutschen Sprache* siehe auch Geyken, Alexander et. al (2017): Die Korpusplattform des „Digitalen Wörterbuchs der deutschen Sprache“ (DWDS). In: *ZGL* 2017; 45(2): 327–344.

wird, mit der Datenbank zu experimentieren und die Daten mit eigenen Strategien und Kriterien zu interpretieren.

Mit der Ressource *DiaCollo* <<https://clarin-d.de/de/kollokationsanalyse-in-diachroner-perspektive>> kann man den Wandel typischer Wortverbindungen, genauer Kollokationen im Verlauf eines bestimmten Zeitabschnitts, untersuchen. Man kann hier selbst entscheiden, auf welcher Korpusgrundlage man sich die Daten generieren lassen möchte. Zur Wahl steht unter anderem *Das Deutsche Textarchiv (DTA)*<sup>6</sup> auf <<http://kaskade.dwds.de/dstar/dta/diacollo/>>, das aus Texten im Zeitraum von 1600 bis 1900 besteht.<sup>7</sup> Ein besonderes Werkzeug, das sehr schön illustriert, welche Möglichkeiten *DiaCollo* für die Darstellung des Wandels von typischen Wortverbindungen bietet, ist die Wortwolke (*cloud*). Mit der *cloud* werden die zehn stärksten (lemmatisierten) Kollokatoren pro Dekade für ein eingegebenes Stichwort visualisiert. Farben markieren dabei die Stärke des gemeinsamen Aufkommens. Mit der Abspieltaste kann man eine Animation starten, die die Kollokatoren der jeweiligen Zeitabschnitte ineinander übergehen lässt und so visuell die sprachliche Veränderung offenbart.

### 3 Weitere Online-Ressourcen

Das *DTWW (Deutsch-türkische Wortverbindungen Wirtschaft)*, <<http://wvonline.ids-mannheim.de/dtww/index.htm>>, ebenfalls dem Projekt *UWV* zugehörig und von Ayfer Aktaş erarbeitet, ist eine ca. 900 Einträge umfassende Sammlung von Phrasemen speziell aus der Domäne Wirtschaft. Alphabetisch sortiert lassen sich die deutschen Wortverbindungen mit Korpusbelegen aus *DeReKo* anzeigen und werden direkt mit der türkischen Entsprechung sowie deren wörtlicher Übersetzung verglichen.

Eine Ressource, die auch von der *UWV*-Projektgruppe erarbeitet wird und sich noch in den Startlöchern befindet, ist das *PrepCon* <<http://uwv.ids-mannheim.de/prepcon/modul1/tables.html>>, dessen Ergebnisse in *OWIDplus* veröffentlicht werden. Das erste Modul *PrepCon explorativ* ist bereits einsehbar und enthält eine Datenbank von Präposition-Nomen-Verbindungen. Tabellarisch werden die lexikografischen Einträge unter anderem mit Hinzunahme der KWICs und Frequenzangaben präsentiert. Diese Ansicht kann dann auf einzelne Präpositionen und

<sup>6</sup> Weiterführend zum *DTA* siehe auch Wiegand, Frank et. al (2018): Recherchieren, Arbeiten und Publizieren im Deutschen Textarchiv: ein Praxisbericht. In: *ZGL* 2018; 46(1): 147–161.

<sup>7</sup> Ebenfalls auswählbare Korpora sind *DTA + DWDS*, *DWDS Kernkorpus* und die Wochenzeitung *Die ZEIT*.

die zugehörigen Nomina oder umgekehrt beschränkt werden. Die Nutzerin oder der Nutzer kann angeleitet durch die Hinweise der Datenbank auf diese Weise selbst den usuellen Gebrauch ableiten. Es folgen zwei Module, die verschiedene Beschreibungstiefen der Präposition-Nomen-Verbindungen behandeln sollen. Das Projekt soll Ende 2019 abgeschlossen sein.

Wer über diese Ressourcen hinaus mit eigenen Sprachdaten Phraseme erforschen möchte, kann mit frei und kostenlos verfügbarer korpuslinguistischer Software experimentieren, die in den meisten Fällen auch phraseologische Tools enthält. Eine Auswahl dieser Programme ist auf <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/links/software> aufgelistet.

# DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT  
DUISBURG  
ESSEN

*Offen im Denken*

ub | universitäts  
bibliothek

Dieser Text wird über DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt. Die hier veröffentlichte Version der E-Publikation kann von einer eventuell ebenfalls veröffentlichten Verlagsversion abweichen.

**DOI:** 10.1515/zgl-2018-0019

**URN:** urn:nbn:de:hbz:464-20201022-140948-9

Alle Rechte vorbehalten.

Dieser Beitrag ist mit Zustimmung des Rechteinhabers aufgrund einer (DFG-geförderten) Allianz- bzw. Nationallizenz frei zugänglich.