

Korpusmethoden zur Erarbeitung eines Wörterbuches leicht verwechselbarer Ausdrücke

Petra Storzjohann (Mannheim)

1 Einleitung

Beim Erlernen einer Fremdsprache kann es zu Situationen kommen, in denen beim Gebrauch formal, lautlich oder semantisch ähnlicher Wörter Unsicherheiten entstehen. Neben Homophonen, Homographen, falschen Freunden, Ausdrücken derselben Wortfamilie oder Dubletten gibt es auch Ausdrücke wie *musisch/musikalisch*, *nötig/notwendig*, *kindisch/kindlich/kindehaft*, die zu sprachlichen Zweifelsfällen gezählt werden. Bei diesen Ausdrücken, sogenannten Paronymen, handelt es sich um Wortpaare (oder lexikalische Gruppen), die neben Ähnlichkeiten der Form oft auch über semantische Gemeinsamkeiten sowie über referentielle Überschneidungsbereiche verfügen. Sie bezeichnen häufig ähnliche Konzepte, deren exakte kontextuelle Abgrenzung in bestimmten sprachlichen Situationen nicht immer erfolgreich gelingt. Das kann zu kommunikativen Verwirrungen bis hin zu Fehlinterpretationen führen. Von solchen Zweifeln sind Muttersprachler*innen wie Fremdsprachenlernende gleichermaßen betroffen, auch wenn es Fälle gibt, die typischerweise der einen oder anderen Gruppe zugeordnet werden können. Um Paronyme kontextuell angemessen zu verwenden bzw. deren Bedeutungsnuancen zu erkennen bzw. deren Bedeutungsnuancen zu kennen, wird linguistisches, aber auch außersprachliches Wissen benötigt. In gängigen ein- oder auch zweisprachigen Bedeutungswörterbüchern ist Beides gleichzeitig meist nicht zu finden.

Am Leibniz-Institut für Deutsche Sprache (IDS) wurde im Programmbereich „Lexikografie und Sprachdokumentation“ ein neuartiges Wörterbuch entwickelt, das leicht verwechselbare Ausdrücke in ihrem aktuellen öffentlichen Sprachge-

brauch deskriptiv beschreibt. Im Jahr 2018 erschien das elektronische Nachschlagewerk „Paronyme – Dynamisch im Kontrast“, das sich durch folgende drei Aspekte auszeichnet:

- 1) Erstens liegen mehrstufige kontrastive Beschreibungsebenen und flexible Darstellungsformen vor;
- 2) zweitens sind die Bedeutungserläuterungen kognitiv-konzeptuell angelegt, um einer langen Forderung nach einer stärker kognitiv ausgerichteten Lexikografie Rechnung zu tragen;
- 3) drittens werden Datengrundlagen und Analysemethoden genutzt, mit denen umfassend Paronyme ermittelt und diese anschließend erstmals empirisch ausgewertet werden konnten.

Nachfolgend soll zunächst das Nachschlagewerk kurz vorgestellt werden. Hauptaugenmerk des Beitrags liegt auf den Erläuterungen der eingesetzten Analysemethoden und der ermittelten Daten, auf deren Grundlage das Wörterbuch entstand und in den kommenden Jahren weiter ausgebaut wird. Das schließt zum einen die Beschreibung der Ermittlung der gesamten Stichwortliste ein, zum anderen aber auch die Darstellung innovativer kontrastiver Korpusverfahren zur Analyse von Wortpaaren.

2 Paronymie im Sprachgebrauch

Paronymie ist ein weit verbreitetes Phänomen, das bei genauerer Betrachtung größer ist als vorab erwartet wurde (vgl. Schnörch 2015). Rund 900 Paare bzw. lexikalische Gruppen (*nötig/notwendig/notwendigerweise, patriarchalisch/patriarchal/patriarchisch*) wurden in einem systematischen Prozess korpusgestützter und redaktioneller Arbeitsschritte ermittelt. Zu Verwechslungen und sprachlichen Zweifelsfällen, die bei Paronymen im Sprachgebrauch zu beobachten sind, kommt es aus unterschiedlichen Gründen. So müssen unterschiedliche linguistische, pragmatische, diskursive aber auch kognitiv-orientierte Parameter herangezogen werden, um das Zustandekommen der vielfältigen Verwechslungsfälle zu erklären und so in einem nächsten Schritt zu einem Typologierungsmodell zu gelangen.¹ Neben rein linguistischen Bedingungen unterliegt der Gebrauch von Paronymen in der Kommunikation auch kognitiven und wissenskonstituierenden Faktoren. In ihren jeweiligen Verwendungsweisen zeigen sie spezifische kommunikative Funktionen, individuelle konzeptuelle Präferenzen oder Diskurszugehörigkeiten. Sie unterliegen aber auch thematisch-sachorientiertem Gebrauchseinschränkungen oder ihre Verwechslungen lassen sich auf einen semantischen Wandel zurückzuführen, der erst in jünge-

¹ Auch Annäherungen an mögliche Definitionen sollten jene unterschiedlichen Faktoren berücksichtigen, die im kommunikativ-situativen Gebrauch eine Rolle spielen, und die über bisherige sprachsystematische Modelle hinausgehen (vgl. dazu die Klassifikationen von Lăzărescu 1999).

rer Vergangenheit stattgefundenen hat.² Darüber hinaus spielen auch Sprechereinstellungen, enzyklopädisches Wissen und kulturelle Erfahrungen bei der Verwendung oder Interpretation von Paronymen eine Rolle.

„Die richtige Zuordnung zu speziellen Wissensdomänen ebenso wie Langzeitwissen über Personen oder Sachverhalte tragen folglich dazu bei, kontextuell angemessen zu kommunizieren. Andernfalls kommt es zu Verwechslungen bis hin zu Fehlern bzw. Missverständnissen: Sachwissen oder der Mangel daran lässt sich besonders gut in metasprachlichen semantischen Aushandlungsprozessen beobachten. Solche Umstände können auch ein Nährboden für Bedeutungswandel sein.“ (Mell; Schnörch; Storjohann 2019: 66)

Fremdsprachenlernende haben ein sehr gutes Gefühl dafür, welche Informationen sie für die Verwendung von verwechselbaren Ausdrücken benötigen, um angemessen zu kommunizieren. Es zeigt sich, dass ihnen durchaus bewusst ist, dass sie neben rein lexikalischem auch außersprachliches und enzyklopädisches Wissen benötigen. Das belegen Auswertungen von Internetforen, deren Teilnehmer*innen sich über sprachliche Unsicherheiten austauschen: Es kommt klar zum Ausdruck, dass, z.B. enzyklopädisch-konzeptuelle Informationen oder Angaben zu referentiellen oder sachgebietsbezogenen kontextuellen Bezugsrahmen gezielt erfragt und diskutiert werden. In solchen Foren werden mit Blick auf Übersetzungsäquivalente komplexe kontextuelle Situationen erläutert. Das Ziel ist, verschiedene Faktoren sprachlicher und außersprachlicher Natur gleichermaßen in Fragen und Antworten zu thematisieren. So zeigt sich das gesamte Spektrum an Einflüssen zur Herstellung angemessener kommunikativer Akte. Auf diese Weise erfährt man, wann sprachliche Schwierigkeiten auftreten, auf welche konkreten Kontexte sie sich beziehen und was in Erfahrung gebracht werden soll. Aber man erfährt u.U. auch, wo bereits nachgeschlagen wurde und ob man mit den Angaben zufrieden ist.³

In Übereinstimmung mit Storrer (2013) zeigte sich, dass die Untersuchung solcher Foren für Lexikograf*innen wertvolle Einsichten über Nachschlagehandlungen und -bedürfnisse liefern können. Aus diesem Grund wurden im Projekt „Paronymwörterbuch“ internetbasierte Nutzerstudien durchgeführt, die die Konzeption maßgeblich beeinflussten. Um den Bedürfnissen an bestimmten Inhalten, Beschreibungs- bzw. Darstellungsformen gerecht werden zu können, verpflichtete sich das Projekt eines korpusgesteuerten Vorgehens, denn nur in authentischer Kommunikation, wie sie Korpusdaten widerspiegeln, können sprachliche wie auch außersprachliche Faktoren sowie sämtliche Facetten des Phänomens im Sprachgebrauch aufgedeckt werden.

² Das betrifft z.B. die Annäherung des Gebrauchs von *sensitiv* an *sensibel* (vgl. Storjohann 2015).

³ Siehe z.B. German Stackexchange (<https://german.stackexchange.com/questions/44541/the-difference-between-eigen-and-eigenartig> [18.06.2020]).

3 Paronymie in der Lexikografie

3.1 Printwörterbücher

Für das Deutsche gab es bis 2018 zwei Printwörterbücher, die Wortpaare mit Verwechslungspotenzial zum Gegenstand ihrer Beschreibung machten. Diese teilweise traditionell angelegten, z.T. veralteten Nachschlagewerke waren Pollmann; Wolk (2010) und Müller (1973). Neben typischen Paronymen enthalten sie Einträge, die verschiedenartige Fälle von Interferenzen darstellen, z.B. Wortgruppen wie *Erlebnis/Ergebnis/Ereignis*, Synonyme wie *Vers/Strophe* oder Paare, die nur lautliche Gemeinsamkeiten haben (*Korsett/Klosett*).

Für das Deutsche gab es bisher kein reines Paronymwörterbuch, welches Ausdrücke mit formalen und semantischen Ähnlichkeiten wie *formal/formell* oder *Wirkung/Wirksamkeit* kontrastiv beschreibt. Darüber hinaus gab es bis 2018 kein Wörterbuch, das paronyme Ausdrücke wie *musisch/musikalisch*, oder auch *konzeptionell/konzeptuell*, *methodisch/methodologisch* konsequent korpusgestützt ermittelt sowie die gegenwartssprachlichen Daten auswertet, interpretiert und dokumentiert, um alle nötigen Informationen für einen Wörterbucheintrag zu erhalten, die Nachschlagenden ausreichend Hilfestellung zum aktuellen und angemessenen Gebrauch solcher Wortpaare anbieten kann. Daher war es das Ziel des Projektes „Paronymwörterbuch“, genau diese Lücke(n) zu füllen, also sich dem Phänomen der sprachlichen Verwechslung aus korpuslinguistischer Sicht zu widmen und ein neues und vor allem neuartiges, dynamisches, benutzerdefiniertes Online-Nachschlagewerk zu entwickeln, das in der Lage ist, flexibel auf verschiedene Bedürfnisse eingehen zu können.

3.2 „Paronyme – Dynamisch im Kontrast“

„Paronyme – Dynamisch im Kontrast“ ist ein kostenloses Internetwörterbuch, das Verwendungen leicht verwechselbarer Ausdrücke des Deutschen in der öffentlichen Gebrauchssprache der Gegenwart umfassend beschreibt. Es ist über die Plattform OWID^{plus} zugänglich. In kontrastiven, dynamischen Einträgen informiert es auf unterschiedlichen Beschreibungstiefen über manifestierte sprachliche Konventionen. Die Einträge werden auf der Basis eines sehr großen Korpus und mithilfe komplementärer Analyseverfahren erarbeitet. Die extrahierten Daten werden anschließend lexikografisch interpretiert, um zuverlässige Einträge zu gewährleisten. Explorative, korpusanalytische Auswertungen sind generell deskriptiv angelegt, sie spiegeln typische und gegenwartssprachlich gut etablierte Verwendungskontexte wider.

Beim Öffnen des Wörterbuchs bekommen Nachschlagende zunächst eine kompakte Überblicksdarstellung, bei der alle Verwendungskontexte der Paronyme einerseits stichwortbezogen horizontal angeordnet sind; vertikal erfolgt zusätzlich eine Zuordnung zu den Verwendungskontexten des jeweiligen Partnerwortes, was

eine kontrastive Perspektive ermöglicht. Um zu erkennen, welche Art der Kontexte vorliegen, bedeutungsgleiche, ähnliche oder unterschiedliche, werden alle Verwendungen nach semantischen Merkmalen entsprechend gruppiert und farblich markiert. Auf diese Weise erscheinen zunächst diejenigen Verwendungen, die sehr starke semantische Überlappungen zeigen, gefolgt von Kontexten, die sich ähnlich sind, aber je eigene Nuancen aufweisen, gefolgt von Verwendungskontexten, die unterschiedlich sind (Abbildung 1).

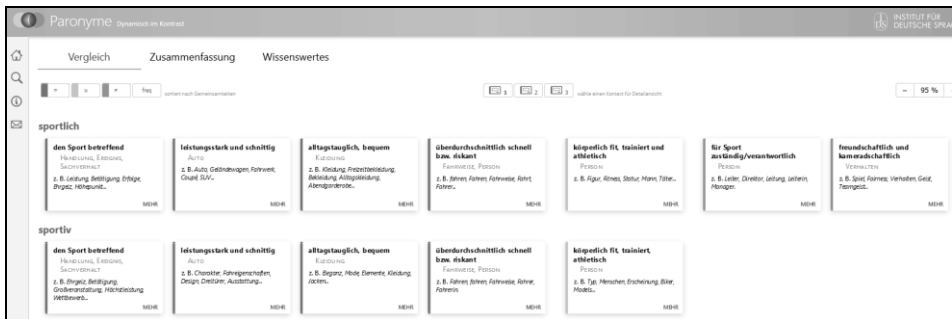


Abbildung 1: Kontrastive Überblicksdarstellung.

Die Reihenfolge der Gruppen kann individuell gewählt werden. Dynamisch lassen sich so unterschiedliche Übersichten nach diversen linguistischen Parametern generieren. Die Darstellung bietet einen schnellen Überblick darüber, wie viele Verwendungskontexte vorliegen und ob bzw. welche Gemeinsamkeiten und Unterschiede zwischen den Ausdrücken existieren. So geht Abbildung 1 hervor, dass *sportlich/sportiv* über vier Kontexte verfügen, in denen starke Überlappungen vorliegen. Bei einer Verwendung existieren Ähnlichkeiten und zusätzlich gibt es zwei Verwendungen nur für den Ausdruck *sportlich*. In der Überblicksdarstellung werden den Nutzer*innen je Verwendung drei Angaben in folgender Reihenfolge angeboten:

- eine Kurzparaphrase meist in Form von Synonymen (z.B. *sportlich* im Sinne von **leistungsstark und schnittig**);
- eine Angabe zur Referenz bzw. den Sachbezug (z.B. in Bezug auf ein **AUTO**);
- maximal fünf Kollokatoren, die beispielhaft die Referenz illustrieren und gleichzeitig typische lexikalische Kombinationen darstellen (z.B. *Auto, Geländewagen, Fahrwerk, Coupé, SUV*).

Jede Kurzparaphrase wird durch ein Bezugswort ergänzt, bei dem es sich um eine Kategorienbezeichnung handelt, die den referentiellen Kontext in einen Bezugsrahmen einbettet. Kurzerläuterungen zusammen mit den entsprechenden Bezugsrahmen oder Sachgebieten und gepaart mit beispielhaften Kontextpartnern verorten jedes Paronym für jede dokumentierte Verwendung.

Die Überblicksdarstellung dient der schnellen Erfassung und Orientierung. Mit ihrer stark durch visuelle Elemente unterstützten Struktur und der Kombination von lexikalischen und enzyklopädischen Informationen können Nachschlagende schnell kontextuelle Gegebenheiten erfassen und unterscheiden. Für bestimmte Anfragen können die hier angebotenen Informationen bereits ausreichende Antworten liefern. Darüber hinaus dient die Art der Präsentation als Navigationsstruktur zu weiteren, detaillierteren Beschreibungen aller Verwendungskontexte. Über einen Link in jedem Kontextfeld gelangt man so zu einer Detailansicht (Abbildung 2). Die Detailansicht umfasst folgende Angaben: die Kurzparaphrase, eine explizite Bedeutungserläuterung (Langparaphrase), ggf. eine Einordnung in eine Domäne/Diskursbereichsangabe (fehlt in Abbildung 2), max. zehn illustrative Kollokationen, ggf. typische Verwendungsmuster (fehlt in Abbildung 2), max. drei Korpusbelege, die den Ausdruck zusammen mit einigen oben erwähnten Kollokationen im Kontext zeigen, Synonyme und Antonyme mit Belegen.

sportlich, leistungsstark und schnittig ⁽⁴⁾

charakterisiert meist ein Auto bzw. dessen Erscheinungsbild dahingehend, dass es z. B. ein tiefergelegtes Fahrwerk sowie stärkere Motorleistung aufweist und optisch schnittig bzw. dynamisch wirkt

z. B. *Auto, Geländewagen, Fahrwerk, Coupé, SUV, Limousine, Dreitürer, Flitzer, Optik, Aussehen*

Verwendungsbeispiele ▼

Belege

Die Bayerischen Motorenwerke, groß geworden mit **sportlichen** Autos und Premium-Limousinen, haben nicht nur irgendwann auf **sportliche** Geländewagen (SUV) gesetzt, um damit viel Geld zu verdienen. Die Münchner haben schon vor Jahren eine große Elektroauto-Kampagne angeschoben. (Süddeutsche Zeitung, 15.06.2015, S. 19, Elektrofahrzeuge.)

Der neue Opel Corsa wird im Juli auf der British International Motor Show in London vorgestellt. [...] Der Dreitürer wird als **sportliches** Coupé positioniert und dient auch als Basis für ein späteres OPC-Topmodell; [...]. (Leipziger-Volkszeitung, 27.03.2006, S. 8, Neue C-Klasse steht in Genf.)

Sinnverwandte Wörter ▼

Synonyme:

athletisch ▷, leistungsstark ▷, schnittig ▷, sportiv ▷

Abbildung 2: Detailansicht eines Verwendungskontextes.

Neben rein linguistischen Details vermitteln die Angaben essentielle enzyklopädische Informationen, deren Kenntnis für angemessene sprachliche Interaktion unerlässlich ist. Die Langparaphrase zu *sportlich* im Sinne von ‚leistungsstark und schnittig‘ etwa enthält das Bezugswort **AUTO**, Kollokatoren wie *Auto, Geländewagen, Fahrwerk, Coupé, SUV, Limousine* etc. spiegeln usuelle Strukturen und die außersprachliche Realität wider. Die Gegenstandsbereiche, auf die die Ausdrücke referieren, sind generell explizit in die Paraphrase eingearbeitet. Sie dienen zudem der Navigation in der Übersichtsebene und unterstützen die Abgrenzung der Kontexte im Zuge der lexikografischen Disambiguierung. Als solcherart verstandene Bezugswörter (z.B. **AUTO, KLEIDUNG, FAHRWEISE, PERSON**) stellen sie entweder eine aus Kollokatoren abstrahierte kognitive Entität dar oder eine assoziierte Kategorie, die mithilfe der CNS-Methode gewonnen wird (siehe dazu 4.3).

Bei allen Nachschlagehandlungen können zusammengehörige Informationen gemeinsam abgerufen und gemeinsam abgespeichert werden. Um neue Wortschatzstrukturen oder Wortschatzelemente gesichert abzuspeichern, ist ein effektiver Semantisierungsprozess nötig. Ein solcher Prozess kann durch verschiedene Techniken des Assoziierens gefördert werden. Nach Haß (2006) zählen dazu u.a. Lernen im Kontext durch das Verknüpfen mit Weltwissen, Definitionen oder Erläuterungen mithilfe von Unter- und Oberbegriffen, Bilden von Paraphrasen, Analogien, Synonymen und Antonymen. Die konzeptuell ausgerichteten Angaben und die lexikalisch-kontextuellen Details im Paronymwörterbuch verstehen sich als Zusammenspiel verschiedenartiger Informationen, die einen solchen Semantisierungsprozess begünstigen sollen.

Ein weiteres Alleinstellungsmerkmal des Wörterbuchs ist die benutzerdefinierte Auswahlmöglichkeit aus unterschiedlichen Kontrastierungsmöglichkeiten. Nachschlagende entscheiden selbst, welche Verwendungen von Interesse sind, um sie individuell einem vergleichenden Studium zu unterziehen. Die Detailansichten von max. drei Verwendungskontexten eines einzelnen Ausdrucks oder von zwei (falls vorhanden drei) Paronymen können entsprechend selbstgewählt aufgerufen werden. Auf diese Weise lassen sich gleichzeitig bestimmte Kontexte im Detail parallel betrachten (Abbildung 3).

sportlich, leistungsstark und schnittig

(4)

charakterisiert meist ein Auto bzw. dessen Erscheinungsbild dahingehend, dass es z. B. ein tiefergelegtes Fahrwerk sowie stärkere Motorleistung aufweist und optisch schnittig bzw. dynamisch wirkt

z. B. *Auto, Geländewagen, Fahrwerk, Coupé, SUV, Limousine, Dreitürer, Flitzer, Optik, Aussehen*

sportiv, leistungsstark und schnittig

(2)

charakterisiert meist ein Auto bzw. dessen Erscheinungsbild dahingehend, dass es z. B. ein tiefergelegtes Fahrwerk sowie stärkere Motorleistung aufweist und optisch schnittig bzw. dynamisch wirkt

z. B. *Auto, Charakter, Fahreigenschaften, Design, Dreitürer, Ausstattung, Limousine*

Verwendungsbeispiele**Belege**

Die Bayerischen Motorenwerke, groß geworden mit **sportlichen** Autos und Premium-Limousinen, haben nicht nur irgendwann auf **sportliche** Geländewagen (SUV) gesetzt, um damit viel Geld zu verdienen. Die Münchner haben schon vor Jahren eine große Elektroauto-Kampagne **angeschoben**. (Süddeutsche Zeitung, 15.06.2015, S. 19, Elektrofahrzeuge.)

Der neue Opel Corsa wird im Juli auf der British International Motor Show in London vorgestellt. [...] Der Dreitürer wird als **sportliches Coupé** positioniert und dient auch als Basis für ein späteres OPC-Topmodell; [...]. (Leipziger-Volkszeitung, 27.03.2006, S. 8, Neue C-Klasse steht in Genf.)

Belege

Zu den witzigeren Autos des Jahres gehört der Space Runner von Mitsubishi, der nach dem Rezept 'kurz, klein und hoch' gestrickt ist. Dies macht ihn zu einem behenden Stadtfitzer, mit dem man auch in der extrem engen Tiefgarage des Drogeriemarktes, den unsere bessere Hälfte gerne ansteuert, gut zurechtkommt. Dem Anspruch, **sportive** Autos zu bauen, wird BMW mit dem 3er-Coupé wieder mehr gerecht als noch vor einigen Jahren. (Süddeutsche Zeitung, 31.12.1992, Das Autojahr 1992.)

Der MiTo soll in der Kleinwagenklasse reüssieren und modebewusste Käufer gewinnen. [...] Damit die Fahreigenschaften mit dem **sportiven** Design mithalten, spendierten die Ingenieure dem MiTo Feinheiten wie Rebound-Federn in den Stoßdämpfern, die das Einknicken in schnellen Kurven minimieren, und das D.N.A.-System. (FOCUS, 30.06.2008, S. 110-110, Kleinwagen.)

Sinnverwandte Wörter**Synonyme:**

athletisch ▷, leistungsstark ▷, schnittig ▷, sportiv ▷

Synonyme:

schnittig ▷, sportlich ▷

Abbildung 3: Kontrastive Detailansicht zweier Verwendungskontexte.

Neben dem bisher vorgestellten Informationsangebot, das sich vor allem auf vergleichende Strukturen bezieht, verfügt das neue Wörterbuch auch über die Rubriken „Zusammenfassung“ und „Wissenswertes“, die überwiegend narrativ-kommentierender Art sind. In der „Zusammenfassung“ werden wesentliche Beobachtungen zu Bedeutung und Verwendung erläuternd festgehalten. Eine Quintessenz fasst die Beziehung zwischen den beiden Paronymen kurz zusammen und bringt sie auf den Punkt. Darunter werden auch Beobachtungen zu Unterschieden in der Vorkommenshäufigkeit, zu Präferenzen oder Einschränkungen regionaler Art werden festgehalten (Abbildung 4).

Vergleich

Zusammenfassung

Wissenswertes

sportlich – sportiv

Quintessenz: Die Adjektive *sportlich* und *sportiv* teilen viele Kontexte und können in diesen gemeinsamen Verwendungen bedeutungsgleich gebraucht werden. Es gibt aber auch zwei Verwendungen von *sportlich*, die nicht für *sportiv* belegt sind. *Sportlich* ist daher semantisch vielseitiger.

Beide Adjektive sind Ableitungen vom Nomen *Sport*. *Sportlich* ist wesentlich häufiger im Korpus belegt als *sportiv*. Beide Adjektive teilen viele Kontexte und können sich in diesen Verwendungen gleichermaßen auf Personen, Handlungen, Sachverhalte oder Autos und Fahrweisen beziehen. Zahlreiche Kollokatoren kommen bei Verwendungen mit beiden Ausdrücken vor. Auffällig ist aber auch, dass die Kontexte, in denen sich *sportlich* und *sportiv* auf Kleidung beziehen, unterschiedlich stark favorisiert werden: *Sportlich* weist sechs verschiedene Verwendungen auf, von denen Bezüge zu Kleidung wenig frequent sind. Im Allgemeinen stehen Charakterisierungen von Aussehen (z. B. von Kleidung) oder von Bauweisen (z. B. bei Autos) bei *sportiv* im Vordergrund. Umgekehrt spielen Kontexte, in denen man sich mit *sportiv* auf den Bereich des Sportes bezieht, nur eine untergeordnete Rolle. Diese Kontexte sind aber zentrale Verwendungen von *sportlich*.

In Belegen, in denen *sportlich* oder *sportiv* die körperliche Fitness von Personen näher beschreiben, liegen unterschiedliche diskursive Präferenzen vor, d. h., sie werden tendenziell auch in unterschiedlichen Themenbereichen verwendet. Auffällig ist dabei, dass *sportlich* häufig in Polizeiberichten vorkommt, in denen Äußerlichkeiten von Tätern geschildert werden. Diese Textsorte ist nicht in gleichem Maße für *sportiv* belegt. Hier werden thematische Bereiche wie Gesundheit und Fitness stärker hervorgehoben.

Daneben weist *sportlich* aber auch zwei Verwendungen auf, die nicht für *sportiv* belegt sind. Zum einen können keine Kontexte für *sportiv* belegt werden, in denen damit Personen charakterisiert werden, die für den Bereich Sport zuständig sind. Zum anderen kann *sportiv* nicht in Kontexten verwendet werden, in denen man kameradschaftliches und faires Verhalten von Menschen, auch außerhalb des Sports, beschreibt. *Sportiv* ist somit in semantischer Hinsicht eingeschränkter.

Abbildung 4: Rubrik Zusammenfassung.

In der Rubrik „Wissenswertes“ finden Nutzer*innen Angaben zur Grammatik, zum Sprachwandel, zu weiteren Paaren mit gleichem Muster, aber auch Informationen zum mündlichen oder fachsprachlichen Gebrauch (z.B. *Parodontose/Parodontitis*, *Synthese/Synthesis*). Die Erarbeitung dieser Angaben wird in den kommenden Jahren vermehrte Aufmerksamkeit geschenkt. Generell ist die empirische Auswertung des Sprachgebrauchs die zentrale Grundlage für die Erarbeitung dieses Nachschlagewerkes. Etablierte Korpusverfahren sowie innovative kontrastive Methoden gepaart mit lexikografischer Kompetenz prägen die redaktionelle Praxis und werden im folgenden Abschnitt näher erläutert.

4 Datenbasis, Analysewerkzeuge und -methoden

Die Voraussetzung für ein gegenwartssprachliches, sprachgebrauchsorientiertes Wörterbuch ist ein geeignetes Korpus, das auf der Basis umfangreicher schriftsprachlicher Korpusarchive des IDS für spezielle lexikografische Zwecke zusammengestellt wurde.⁴ Mithilfe des Paronymkorpus wurde erstmals eine konsequent korpusbasierte und systematisch redaktionell ausgewertete Stichwortliste von Paronymen erstellt (vgl. Schnörch 2015). Darüber hinaus ist das Korpus die Datengrundlage zur Bedeutungsanalyse der Stichwörter, Quelle für die Daten der verschiedenen Angaben und der Belege in den Wörterbuchartikeln. Das Besondere an diesem Korpus ist seine relative regionale Ausgewogenheit und seine öffentliche Zugänglichkeit über das Recherchetool COSMAS II. Damit werden die Inhalte der Einträge nachprüfbar bzw. nachvollziehbar und somit die Wörterbucharbeit transparent. Insgesamt besteht die Datenbasis aus 27 verschiedenen Quellen (insgesamt 2,3 Milliarden Textwörter) mit regionalen und überregionalen Zeitungstexten aus den Jahrgängen 1990-2015.

Zeitungstexte zur sprachlichen Auswertung heranzuziehen, hat Nachteile, sie stellen aber dennoch eine geeignete Datenbasis dar. Zeitungen und Zeitschriften sind durchaus vorteilhaft für die Untersuchung von Paronymen im öffentlichen Sprachgebrauch: Sie sind stark verbreitete und redigierte Texte, die von einer Großzahl der Bevölkerung rezipiert werden. Generell bedienen sich allgemeine journalistische Texte eines konventionalisierten Sprachgebrauchs und sie decken auch eine Vielzahl unterschiedlicher Themen und Genres ab. Man denke etwa an Ressorts und Rubriken wie Sport (*sportlich/sportiv*), Politik (*populär/populistisch*), Kultur (*Adaption/Adaptation*), Gesellschaft (*patriarchalisch/patriarchal/patriarchisch*), Forschung (*Informationstechnik/Informationstechnologie*) oder Textsorten wie Wetterberichte (*klimatisch/klimatologisch*), Leserbriefe, Anzeigen (*muskulär/muskulös*). Zeitungen werden von unterschiedlichen Autor*innen verfasst, sodass sie neben typisch journalistischen Berichterstattungen auch persönliche Sprachstile abbilden, aber auch regionale wie überregionale Eigenarten ein- bzw. auszublenzen vermögen. Ein weiterer Vorteil von Zeitungen ist, dass sie gegenwärtige kulturelle Einstellungen und soziale Prägungen widerspiegeln, die sprachlich ihren Niederschlag finden.⁵

Im Moment beschränkt sich die lexikografische Arbeit auf die Auswertung Paronympaare in schriftsprachlichen Texten, künftig werden auch gesprochensprachliche Daten untersucht, um Unterschiede standardnaher und spontansprachlicher Kommunikation zu ermitteln.

⁴ Zum Paronymkorpus siehe: <http://www1.ids-mannheim.de/lexik/paronymwoerterbuch/dasparonymkorpus.html> [07.06.2020].

⁵ Das ist besonders für die Untersuchung von diskursgebundenen Paronymen wichtig (z.B. *national/nationalistisch, unehelich/nichtehelich/auferehelich*; Tanacković; Krtalić; Lacović 2014: 1).

4.1 Die Stichwörter

Bisher konnte man sich kaum eine Vorstellung über das Gesamtinventar an deutschen Paronymen machen, da lange nur individuelle Listen an verwechselbaren Ausdrücken vorlagen, die etwa während langjähriger Lehrtätigkeit angelegt wurden (vgl. z.B. Lăzărescu 1999). Diese vermittelten den Eindruck, dass Paronyme eine überschaubare Menge an Ausdrücken bilden. Dass das Phänomen Paronymie viel umfangreicher ist als ursprünglich angenommen, zeigte dann die erste systematisch korpusgestützt ermittelte und redaktionell ausgewertete Stichwortkandidatenliste (vgl. Schnörch 2015). Auf der Grundlage einer computerbasierten, zeichen- bzw. musterorientierten Suche von Wortpaaren im Korpus wurden gezielt Paare ermittelt, deren formale Bestandteile sich nur in bestimmten Mustern unterscheiden. Tabelle 1 zeigt, dass diese Muster entweder morphologische Elemente wie Suffixe oder auch einfache musterhafte Stringsequenzen sind.

Tabelle 1: Beispiele für automatisierte Stichwortsuche.

Zeichenkette	Beispiele
-bar/-lich	<i>sichtbar/sichtlich, verantwortbar/verantwortlich, nachweisbar/nachweislich</i>
-tiv/-bel	<i>operativ/operabel, repräsentativ/repräsentabel, sensitiv/sensibel</i>
-al/-ell	<i>spezial/speziell, formal/formell, ideal/ideell, original/originell</i>
-v/-onal	<i>emotiv/emotional, fiktiv/fiktional, operativ/operational</i>
-al/-sch	<i>kategorial/kategorisch, diktatorisch/diktatorial, hegemonial/hegemonisch</i>

Auf der Grundlage von ca. 200 Mustern wurden 9246 Wortpaare automatisch extrahiert, anschließend „manuell“ geprüft und in Häufigkeitsklassen einsortiert. Das Ergebnis dieser Auswertung waren rund 2000 Wortpaare, die sich als Paronymkandidaten qualifizierten. Nur rund 710 Paare weisen überhaupt eine ausreichende Frequenz auf, um statistische Analysen durchführen zu können. Darunter befinden sich auch zahlreiche Zusammensetzungen oder Negationen, die als Sublemma angesetzt werden können (z.B. **parteiisch/parteilich**: *innerparteiisch/innerparteilich, überparteiisch/überparteilich, unparteiisch/unparteilich*). Das Ergebnis der umfangreichen Auswertung waren circa 250 Paare, die als ausreichend frequente Hauptstichwörter empirisch untersucht und im Wörterbuch dokumentiert werden.

4.2 Kollokationsanalyse

Die Bedeutungsanalyse eines lexikalischen Ausdrucks setzt eine gründliche Untersuchung seines Gebrauchs voraus. Auf der Basis eines umfangreichen Korpus und unter Einsatz verschiedener korpuslinguistischer Verfahren werden zu diesem Zwecke Kollokationsprofile der Paronyme ermittelt, analysiert, interpretiert und redaktionell eingearbeitet. Das heißt, dass semantische Umgebungen eines Such-

wortes nach auffälligen Mustern untersucht werden, um typische kontextuelle Mitspieler, sogenannte Kollokatoren oder Kookkurrenzen ermitteln zu können. Dieses computerbasierte Verfahren (Kollokationsanalyse, vgl. Belica 1995) extrahiert statistisch signifikante kontextuelle Partnerwörter und ist mittlerweile in der deutschsprachigen Lexikografie ein etabliertes korpusanalytisches Vorgehen (vgl. z.B. Klosa; Storjohann 2011). Der entscheidende Vorteil dieser Methode liegt darin, Lexikograf*innen einen systematischen Zugang zu usuellen sprachlichen Strukturen und Mustern des Sprachgebrauchs zu ermöglichen. Aber auch auf der Seite der Rezipienten ist die Angabe von Kollokatoren in Wörterbüchern hilfreich, um usuelle Muster und musterhafte Konstruktionen erlernen und Assoziationen verknüpfen zu können.

„Für das Fach DaF/DaZ sind insbesondere linguistische Ansätze und Theorien interessant, die versuchen, Sprache umfassend in den Blick zu nehmen, die funktional orientiert sind, die den Spracherwerb in den Blick nehmen und somit Regeln und Regelmäßigkeiten aus der Lernerperspektive beschreiben. Das in den letzten Jahren verstärkt gewachsene Interesse an der Rolle formelhafter Sequenzen (Kollokationen, kommunikative Routinen, idiomatische Prägnungen) hat, zusammen mit der in der Spracherwerbsforschung neu bewerteten Rolle von Chunks beim Sprachenlernen, dazu geführt, dass das Verhältnis von Grammatik, Lexikon und kommunikativem Kontext insgesamt neu bewertet wird.“ (Fandrych; Hufeisen; Krumm; Riemer 2010: 7)

Abbildung 5 zeigt, dass Kollokatoren zusammen mit ihren syntagmatischen Mustern in einer Ergebnisliste nach Signifikanzwerten geordnet angezeigt werden. Deren Gruppierung und Einordnung nach verschiedenen Kontexten obliegt anschließend der lexikografischen Interpretation und Auswertung (vgl. Schnörch 2019). So wird jeder Ausdruck in seine wesentlichen semantischen Kontexte zerlegt, um unterscheidbare Verwendungen eines Lemmas zu erhalten. Diese entsprechen nicht zwangsläufig traditionellen Lesarten. Vielmehr rechtfertigen divergierende Referenz bzw. konzeptuelle Kategorien eine Trennung. In kontrastiv angelegten Kategorien werden die Arten der Verwendungsaspekte eines jeden Paronyms zu denen des Partnerausdrucks in Beziehung gesetzt. Abbildung 5 ist ein exemplarischer Ausschnitt der Kollokationsliste für das Adjektiv *effektiv*. Der Ausdruck *effektiv* z.B. kommt besonders häufig zusammen mit *Jahreszins* oder *Hypothekarkredite* vor und weist auf eine Verwendung im Bereich Bank- und Finanzwesen hin. Es taucht auch gemeinsam mit Verben wie *arbeiten*, *gestalten*, *nutzen* auf, was auf Eigenschaften von Handlungen hinweist, die dahingehend charakterisiert werden, einen Prozess in Bezug auf wirtschaftliche Aspekte als optimal wirksam zu bezeichnen. Aber auch der Gebrauch von *effektiv* im Kontextbereich der Medizin (*Mittel*) und der Kriminalität (*bekämpfen*, *Kriminalität*, *Schutz*) können mithilfe statistisch signifikanter Mitspieler schnell erfasst werden.

LLR	kumul.	Häufig	Kookkurrenzen	syntagmatische Muster
5428	17	17	Jahreszins anfängliche	88% Der/der anfängliche effektive Jahreszins E berücksichtigt
	356	339	Jahreszins	46% den/einem effektiven [...] Jahreszins von ...
4223	1495	1139	arbeiten	46% effektiv [...] arbeiten
3014	2126	631	gestalten	75% effektiver (zu) gestalten
2490	3728	1602	sehr	67% sehr [...] effektiv
2002	4350	622	nutzen	54% effektiv (zu) nutzen
1682	4792	442	möglichst	64% möglichst [...] effektiv
1645	4815	23	Fibor Hypothekarkredite	100% 5,90% Hypothekarkredite auf 10 Jahre effektiv 6,55 6,95% Fibor
	4894	79	Fibor	100% 10 Jahre effektiv 5,75 - 5,85% Fibor
1617	6335	1441	können	39% effektiv [...] können
1556	6783	448	Mittel	23% ein effektives [...] Mittel gegen/um ...
1507	7044	261	Methode	42% eine/und effektive [...] Methode um ...
1448	7059	15	bekämpfen Kriminalität	60% Kriminalität [...] effektiver (zu) bekämpfen
	7072	13	bekämpfen Armut	53% die Armut [...] effektiv (zu) bekämpfen
	7254	182	bekämpfen	59% effektiv (zu) bekämpfen
1446	7577	323	wesentlich	83% wesentlich [...] effektiver als
1426	7905	328	äußerst	58% äußerst [...] effektiv
1350	8195	290	schneller	80% schneller (und) effektiver
1292	8210	15	Kontrolle parlamentarische	66% eine effektive parlamentarische [...] Kontrolle der ...
	8559	349	Kontrolle	54% eine effektive [...] Kontrolle der ...
1159	8586	27	Hypothekarkredite	92% 11,95% Hypothekarkredite auf 5 Jahre effektiv 5,50 - 5,90% Hypothekarkredite auf
1155	8919	333	genutzt	55% effektiv [...] genutzt werden
1121	9261	342	Schutz	41% einen effektiven [...] Schutz vor/der ...
1111	9327	66	11,25% Hypothekarkredite	100% - 11,25% Hypothekarkredite auf 5 Jahre effektiv 5,27 -
1080	9585	258	eingesetzt	44% effektiv [...] eingesetzt werden
1053	10046	461	schnell	80% schnell (und) effektiv
1029	10694	648	Arbeit	35% eine effektive [...] Arbeit
995	10704	10	Maßnahmen ergreifen	70% effektive Maßnahmen [...] zu ergreifen um
	10711	7	Maßnahmen ergriffen	71% effektive Maßnahmen [...] ergriffen werden ...
	10985	274	Maßnahmen	31% effektive [...] Maßnahmen
958	11062	77	kostengünstiger	68% effektiver (und) kostengünstiger
948	11065	3	Bekämpfung Kriminalität organisiert	66% effektive Bekämpfung der organisierten Kriminalität
	11076	11	Bekämpfung Kriminalität	63% die/eine effektive Bekämpfung der Kriminalität

Abbildung 5: Kollokationsanalyse von *effektiv*.

Etwas anders sieht es auf den ersten Blick bei *effizient* aus. Aspekte der Energiegewinnung und -nutzung, der Einsatz von Ressourcen oder neuer Technologien sowie Fragen des Umweltschutzes (z.B. *Energie, Energienutzung, Ressourcen, erneuerbar, Technologien*) werden häufig im öffentlichen Sprachgebrauch thematisiert. Daneben werden aber auch Gemeinsamkeiten deutlich, wie etwa die Verwendung des Ausdrucks in allgemeinen, politischen oder ökonomischen Kontexten (*arbeiten, gestalten, Verwaltung, Strukturen, kostengünstig*). Hier werden bereits semantische und diskursive Überschneidungen zu *effektiv* im Sinne ‚erfolgreicher Prozesse‘ und ‚ökonomisch wirksamer Strukturen‘ sichtbar.

Die sorgfältige Auswertung der Kollokatoren und das Prüfen der Kontexte mithilfe größerer Textausschnitte ermöglicht nach und nach die Erschließung sämtlicher Bedeutungsaspekte.⁶ Gemeinsamkeiten und Unterschiede werden sukzessive herausgearbeitet. Der Nachteil dieser Methode liegt darin, dass sie für jeden einzelnen Ausdruck diese Analyse separat durchgeführt werden muss und eine

⁶ Das Bedeutungsspektrum von *effektiv* und *effizient* geht noch über die hier erwähnten Verwendungen hinaus; siehe http://www.owid.de/parowb/artikel/effektiv_effizient_effektivoll/diff [18.06.2020].

vergleichende Untersuchung nur „manuell“ durch Abgleichen erfolgen kann. Wie aufwändig diese Praxis ist, zeigt u.a. Schnörch (2019). Aus diesem Grund rückte ein bisher kaum genutztes Verfahren in den Blickpunkt, das zwei Ausdrücke vergleichend darstellt.

4.3 Das CNS-Verfahren

Die Arbeiten von Vachková (2010) (z.B. *Abendland/Morgenland*) und Marková (2012) (z.B. *kalt/kühl*) lieferten erste Erkenntnisse im Umgang mit einer innovativen kontrastiven Methode, dem Contrasting-Near-Synonyms-Verfahren (CNS)⁷, bei der die Kollokationsprofile zweier Ausdrücke miteinander verglichen werden und das Ergebnis in Form merkmalsbasierter topografischer Karten bereitgestellt wird. Heute ist dieses Instrument ein unverzichtbarer Teil der täglichen Wörterbucharbeit im Paronymprojekt, das komplementär zur klassischen Kollokationsanalyse genutzt wird. Die Contrasting-Near-Synonyms-Methode fußt auf einer stark visualisierten Erfassung verschieden gruppierter Kontexte in Form einer Matrix. Diese ordnet Ausdrücke mit ähnlichen Kollokationsprofilen nach semantischer Nähe in Abstufungen an, die in selbstorganisierenden Merkmalskarten (Self Organising Maps: SOM) farblich nachempfunden werden. Die CNS-Methode liefert damit eine kombinierte Merkmalskarte, die auf der Ermittlung lexikalisch-semantischer Relationen basiert und überblicksartig schnell die gebrauchsorientierte Kontrastierung eines Wortpaares ermöglicht, das über semantische Ähnlichkeiten verfügt oder auch nicht (siehe Abbildung 6).

⁷ Das CNS-Verfahren wurde am IDS im Programmbereich Korpuslinguistik entwickelt (vgl. Belica 2006). Zum Einsatz in der Lexikografie siehe auch Vachková; Belica (2009), Belica (2011). SOMs arbeiten auf der Grundlage von ermittelten Kollokationsprofilen, die in der CCDB, einer experimentellen Kooperationsdatenbank gespeichert sind (vgl. Belica ab 2001).



Abbildung 6: CNS von *effektiv* und *effizient* mit lexikografischer Auswertung.

„Erstellt man eine kombinierte lexikalische Merkmalskarte für ein Paar von semantisch verwandten Lexemen, indem man alle Lexeme, die mindestens einem der beiden Lexeme ähnlich sind, dem Prozess der Selbstorganisation gemeinsam unterzieht, so treten die Gemeinsamkeiten und Unterschiede in der Verwendung einzelner Wörter des untersuchten Wortpaares deutlich hervor.“ (Belica 2011: 168)

Gruppenweise werden also diejenigen lexikalischen Ausdrücke eingeordnet, die eine Ähnlichkeit im Kollokationsprofil mit den Kollokationsprofilen der beiden Suchwörter (hier *effektiv/effizient*) haben. Gerade die semantischen Ähnlichkeiten manifestieren sich in unterschiedlichen, farblich nuancierten Gruppen. Deren Darstellung erfolgt in einer gitterartigen Matrix, die einem Kontinuum nachempfunden ist. Auf diese Weise werden *effektiv* und *effizient* mit all denjenigen Ausdrücken, mit denen sie Teile ihres Kollokationsprofils gemeinsam haben, direkt zueinander in Beziehung gesetzt. Eine kombinierte SOM sorgt dafür, dass unstrukturierte Muster

für zwei Ausdrücke gleichzeitig analysiert, strukturiert sowie einsortiert werden und in visualisierter Form zur linguistischen Interpretation angeboten werden. Die lexikografische Leistung liegt in der Interpretation dieser Karte. Sie stellt ein Instrument für spontanes intuitives Assoziieren dar, indem die einzelnen Gruppen (Kästchen) näher betrachtet werden und ein assoziiertes Konzept darübergerlegt wird (Abbildung 6).⁸ Merkmalsbasierte kartografische Verfahren eignen sich gut zur Herleitung anschaulicher Bedeutungserläuterungen. Das liegt daran, dass bei dieser eher kognitiv orientierten Methode „das assoziative Moment eine grundlegende Rolle spielt“ (Vachková 2018: 260). Die assoziierten Verknüpfungen, die bei längerer Betrachtung entstehen, geben die entscheidenden Hinweise auf Strukturen, Affinitäten, Themen- und Sachgebiete, Domänen sowie Gegenstandsbezüge.

Besonders gut gelingt die Erfassung von diskursiven Eigenarten eines Ausdrucks oder die Ermittlung von Themen- bzw. Sachgebieten, in denen die Suchwörter besonders häufig vorkommen. Über die Darstellung und Bündelung lexikalisch-semanticischer Relationen gelangt man zu Einsichten über verschiedene Gebrauchsweisen eines jeden einzelnen Paronyms und erfasst schnell, wo Unterschiede vorliegen. So sind in Abbildung 6 die evozierten mentalen Konzepte zur Veranschaulichung schwarz oder weiß über die SOMs gelegt. *Effektiv* und *effizient* verfügen über diskursive Überschneidungen (im Original⁹ orange markiert). Mit beiden Adjektiven werden gleichermaßen ökonomische Strukturen und Abläufe bezeichnet. Wirtschaftliche Kontexte dominieren und bestimmte Aspekte der Kosten, des Gewinns, wirtschaftlicher Systeme bzw. Marketingstrategien werden häufig zum Ausdruck gebracht.

Ihre Verwendungen unterscheiden sich aber auch: Im unteren linken Bereich wird der Ausdruck *effektiv* präferiert, dem oberen rechten Bereich dagegen wird *effizient* zugeschrieben. Hier tauchen typischerweise Kontexte wie *effektive Zinsen*, *effektiver Jahreszins* u.ä. auf. Auch der heller erscheinende Bereich rechts unten mit Assoziationen wie Medizin und Therapie ist dem Ausdruck *effektiv* zuzuschreiben. Der obere rechte Bereich um *effizient* betrifft Themenfelder der Energie (vor allem der Stromerzeugung) und Technologie. Hier werden technische bzw. technologische Ressourcen sowie deren Wirksamkeit oder Nachhaltigkeit näher charakterisiert.

Was wie ein „lexikografisches Spielzeug“ (Vachková 2018: 260) anmutet, entpuppte sich im Laufe der Wörterbucharbeit als wichtiges Werkzeug zur Ermittlung essentieller lexikografischer Inhalte (Domänen, Diskursbereiche oder Themenkategorien). Dies spielt gerade in Hinblick auf das gemeinsame Darstellen linguistischer und enzyklopädischer Inhalte für kognitiv ausgerichtete Wörterbuchartikel, also der Verknüpfung von Sprach- und Weltwissen, eine zentrale Rolle. Der Nachteil dieser Methode ist aber, dass nur Verwendungskontexte, die im Sprachge-

⁸ Eine ausführliche Anleitung zum Vorgehen des Assoziierens geben Vachková; Belica (2009: 260).

⁹ Das Verfahren arbeitet mit einem Farbkontinuum, das leider durch die Schwarz-Weiß-Abbildungen nicht widergespiegelt wird.

brauch dominieren, prominent in Erscheinung treten. Und das CNS-Prozedere lenkt die Betrachtenden zu den relevanten thematischen Bereichen, aber konkrete lexikalische Muster (also syntagmatische Umgebungen) vermag diese Methode nicht zu verdeutlichen. Auch syntaktische Besonderheiten sind nicht hinreichend ableitbar. Kleinere semantische Nuancen lassen sich nur mithilfe der jeweils präferierten Kookkurrenzen oder kleinerer Kontextausschnitte näher untersuchen, also mit der klassischen Kollokationsanalyse. Bestenfalls ergänzen sich beide Herangehensweisen bei der lexikografischen Arbeit. Der unterstützende Beitrag der Kollokationsanalyse besteht konkret darin, beispielhafte sprachliche Muster nach ihrer Signifikanz geordnet anbieten zu können. Im Idealfall konkretisieren sie die Abstrahierungen der CNS-Methode und spitzen die gedanklichen Verknüpfungen zu, wie etwa Tabelle 2 zeigt.

Mithilfe komplementärer korpusgesteuerter Methoden erhalten Lexikograf*innen einen detaillierten Einblick in die sprachlichen Strukturen des Paronymgebrauchs. Dadurch weichen die Angaben der Wörterbucheinträge des neuen Nachschlagewerkes u.U. von denen anderer Wörterbücher ab.

Tabelle 2: Zusammenfassung Zuordnung von Bezugsrahmen, Diskurs/Domänenangabe und Kollokatoren.

effektiv	effizient
,optimal wirksam‘ ÖKONOMISCHE/R STRUKTUR, PROZESS, BEREICH häufig in Wirtschaft, Politik z.B. <i>Kontrolle, Arbeit, Verwaltung, Bereich, Strukturen, Lösungen, arbeiten, wirtschaften</i>	,optimal wirksam‘ ÖKONOMISCHE/R STRUKTUR, PROZESS, BEREICH häufig in Wirtschaft, Politik z.B. <i>Verwaltung, Strukturen, Prozesse, Lösungen, System, arbeiten, wirtschaften</i>
,allgemein erfolgreich‘ PROZESS, STRATEGIE, SACHVERHALT z.B. <i>Lernen, Strategien, Maßnahme, Kommuni- kation, Methode, Instrument</i>	,allgemein erfolgreich‘ PROZESS, STRATEGIE, SACHVER- HALT z.B. <i>Lernen, Instrument, Strategie, Me- thoden, Kommunikation, Maßnahmen</i>
,technologisch ressourcenschonend‘ ANLAGE, GERÄT häufig in Technik, Technologie z.B. <i>Technologien, Motoren, Technik, Anlagen, Kraftwerke</i>	,technologisch ressourcenschonend‘ ANLAGE, GERÄT häufig in Technik, Technologie z.B. <i>Technologien, Motoren, Technik, Anlagen, Kraftwerke</i>

effektiv	effizient
,wirkungsvoll‘ KRISENBEWÄLTIGUNG, KRIMINALI- TÄTSBEKÄMPFUNG z.B. <i>Bekämpfung, Strafverfolgung, Polizeiarbeit, Kriminalitätsbekämpfung</i>	
,medizinisch wirksam‘ MEDIZIN häufig in Medizin, Biologie z.B. <i>Ganzkörpertraining, Training, Therapie, Behandlung, Medikamente</i>	
,ökologisch nachhaltig‘ MAßNAHME, RESSOURCE häufig in Umweltschutz z.B. <i>Wärmedämmung, Klimaschutz, Energieein- sparung, Energienutzung</i>	
,real, tatsächlich‘ GELDBETRAG, FINANZPRODUKT häufig in Finanzwesen z.B. <i>Jahreszins, Zins, Zinssatz, Steuerbelastung</i>	
	,verfahrenstechnisch ressourcen- schonend‘ ENERGIEGEWINNUNG oft in Umwelt, Energie z.B. <i>Nutzung, Energienutzung, Einsatz, Stromerzeugung</i>

5 Schlussbemerkungen

Das neue Paronymwörterbuch wird in den kommenden Jahren kontinuierlich ausgebaut. Seit Beginn versucht das Projekt, der Forderung nachzukommen, lexikografische Tradition mit kognitiver Theorie stärker zu verknüpfen (vgl. z.B. Kövecses; Csábi 2014, Ostermann 2015), indem das Zusammenspiel von Bedeutung, Gebrauch, Konzept, Diskurserfahrung und lexikalischer Repräsentation stärker betont wird. Das geht nur, wenn neben einem ausgefeilten lexikografischen Konzept und einer ausgewogenen Datengrundlage auch entsprechende Analysetools

und -verfahren genutzt werden (vgl. Fuertes-Olivera 2013). Im Fokus dieses Beitrags stand das Zusammenwirken von Korpusdaten, linguistischer Analyse und lexikografischer Beschreibung, wie es zur Untersuchung und Dokumentationen des Paronymwörterbuchs ‚Paronyme – Dynamisch im Kontrast‘ genutzt wird. Ein derartiges Zusammenwirken ermöglicht die Verknüpfung sprachlicher und außersprachlicher Details in Wörterbuchartikeln, sodass lexikalische Muster zusammen mit enzyklopädischen Angaben nachgeschlagen werden können. Solche Informationen können folglich gleichzeitig abgespeichert werden und einen erfolgreicherem Semantisierungsprozess anstoßen. Anhand konkreter Wörterbuchbeispiele wurde einerseits gezeigt, welche Korpusinformationen für einen Wörterbucheintrag relevant sind, und andererseits wurden die unterschiedlichen Methoden diskutiert, mit denen unterschiedliche Details, sprachliche wie außersprachliche, gewonnen werden.

Von Anfang an lag es im Interesse des Projekts, das neue elektronische Nachschlagewerk so zu konzipieren, dass verschiedene Aspekte der Wortschatzvermittlung berücksichtigt werden und dass ein reiches, vernetztes Informationsangebot für Lehrende und Lernende angeboten wird, um diese lernpsychologisch zu unterstützen. Das bedeutete, sich kritisch mit bestehenden Wörterbüchern zu beschäftigen, potenzielle Nutzer*innen und ihre Bedürfnisse vorab zu studieren und für getroffene inhaltliche Entscheidungen passende sprachtechnologische Ansätze zu nutzen. Dazu müssen lexikografische Praktiken und Traditionen aufgebrochen werden. Das heißt in der Konsequenz aber auch, dass das neue Wörterbuch Lernende beim Umgang mit dieser Ressource z.T. vor unbekannte Nachschlagehandlungen stellt. Innovative Präsentationen sollen einen großen Beitrag zur sachgemäßen Erfassung und zur richtigen Methodik des Sichtens und Verankerns von Wortschatz und seinen Strukturen leisten. Nutzer*innen sollen komplexe Strukturen und umfangreiche Informationen zu einem Stichwort, seinen Vernetzungen und seinen Beziehungen kognitiv verankern und dabei nicht den Überblick verlieren.

Literatur

- Belica, Cyril (1995): *Statistische Kollokationsanalyse und -clustering. Korpuslinguistische Analysemethoden*. Mannheim: Institut für Deutsche Sprache.
- Belica, Cyril (ab 2001): *Koolekurrenzdatenbank CCDB – V3.3. Eine korpuslinguistische Denk- und Experimentierplattform*. Mannheim: Institut für Deutsche Sprache. (<http://corpora.ids-mannheim.de/ccdb/> [15.09.2018])
- Belica, Cyril (2006): *Modellierung semantischer Nähe: Kontrastierung von nahen Synonymen. Korpusanalytische Methode*. Mannheim: Institut für Deutsche Sprache. (<http://corpora.ids-mannheim.de/ccdb/> [15.09.2018])

- Belica, Cyril (2011): Semantische Nähe als Ähnlichkeit von Kookkurrenzprofilen. In: Abel, Andrea; Zanin, Renata (Hrsg.): *Korpora in Lehre und Forschung*. Bozen: Freie Universität Bozen, 155-178.
- COSMAS II: *Corpus Search, Management and Analysis System*. Mannheim: Institut für Deutsche Sprache. (<http://www.ids-mannheim.de/cosmas2/web-app/> [15.05.2018])
- DGD: *Datenbank Gesprochenes Deutsch*. (https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome [15.09.2018])
- Fandrych, Christian; Hufeisen, Britta; Krumm, Hans-Jürgen; Riemer, Claudia (2010): Deutsch als Fremd- und Zweitsprache als spezifisches Lehr- und Forschungsgebiet. In: Krumm, Hans-Jürgen u.a. (Hrsg.): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch*, 1. Halbband. Berlin/Boston: de Gruyter, 1-18.
- Fuertes-Olivera, Pedro A. (2013): e-lexicography: The Continuing Challenge of Applying New Technology to Dictionary-Making. In: Jackson, Howard (Hrsg.): *The Bloomsbury Companion to Lexicography*. London: Bloomsbury, 323-340.
- Haß, Frank (2006): *Fachdidaktik Englisch*. Stuttgart: Klett.
- Klosa, Annette; Storjohann, Petra (2011): Neue Überlegungen und Erfahrungen zu den lexikalischen Mitspielern. In: Klosa, Annette (Hrsg.): *elexiko. Erfahrungsberichte aus der lexikographischen Praxis eines Internetwörterbuchs*. Tübingen: Narr Francke Attempto, 49-80.
- Kövecses, Zoltán; Csábi, Szilvia (2014): Lexicography and cognitive linguistics. In: *Revista Española de Lingüística Aplicada* 27/1, 118-139.
- Lăzărescu, Ioan (1999): *Die Paronymie als lexikalisches Phänomen und die Paronomasie als Stilmittel im Deutschen*. Bukarest: Editura Anima.
- Marková, Věra (2012): *Synonyme unter dem Mikroskop. Eine korpuslinguistische Studie*. Tübingen: Narr Francke Attempto.
- Mell, Ruth M.; Schnörch, Ulrich; Storjohann, Petra (2019): Korpussemantische Einflussfaktoren auf Eigenschaften und Funktionen von Paronymen. In: Storjohann, Petra (Hrsg.): *Paronymie im deutschen Sprachgebrauch*. Themenheft Deutsche Sprache 1/2019. Berlin: Erich Schmidt Verlag, 53-67.
- Müller, Wolfgang (1973): *Leicht verwechselbare Wörter. Duden Taschenwörterbücher Band 17*. Mannheim: Bibliographisches Institut.
- Ostermann, Carolin (2015): *Cognitive lexicography. A New Approach to Lexicography Making Use of Cognitive Semantics*. Berlin/Boston: de Gruyter.

- OWID^{plus}: Online-Wortschatz-Informationssystem Deutsch^{plus}.
(<http://www.owid.de/plus/> [15.05.2018])
- Paronyme – Dynamisch im Kontrast. (<http://www.owid.de/parowb/>
[15.05.2018])
- Paronymkorpus. (www1.ids-mannheim.de/lexik/paronymwoerterbuch/das-paronymkorpus.html [15.05.2018])
- Pollmann, Christoph; Wolk, Ulrike (2010): *Pons Wörterbuch der verwechselten Wörter. 1000 Zweifelsfälle verständlich erklärt*. Stuttgart: Klett.
- Schnörch, Ulrich (2015): Wie viele Paronympaare gibt es eigentlich? Das Zusammenspiel aus korpuslinguistischen und redaktionellen Verfahren zur Ermittlung einer Paronymstichwortliste. In: *Sprachreport* 2015/4, 16-26.
- Schnörch, Ulrich (2019): Kookkurrenzanalyse und Vergleich: Überlegungen zur Methodenanwendung bei der lexikografischen Beschreibung von Paronymen. In: Storjohann, Petra (Hrsg.): *Paronymie im deutschen Sprachgebrauch*. Themenheft Deutsche Sprache 1/2019. Berlin: Erich Schmidt Verlag, 12-24.
- Storjohann, Petra; Ulrich Schnörch (2017): Sprachlicher Wandel der Gegenwart und seine Dokumentation in einem Wörterbuch. In: Vachková, Marie; Šemelík, Martin; Kloudová, Věra (Hrsg.): Themenheft *Lexikographie*. Germanistica Pragensia, AUC Philologica, 4/2016. Prag: Charles University Karolinum Press, 133-172.
- Storjohann, Petra (2015): Was ist der Unterschied zwischen sensitiv und sensibel? In: *Zeitschrift für Angewandte Linguistik* 62/1, 99-122.
- Storjohann, Petra (2019): Paronyme – Dynamisch im Kontrast. Ein kognitiv ausgerichtetes, multifunktionales, dynamisches Nachschlagewerk. In: Storjohann, Petra (Hrsg.): *Paronymie im deutschen Sprachgebrauch*. Themenheft Deutsche Sprache 1/2019. Berlin: Erich Schmidt Verlag, 82-94.
- Storrer, Angelika (2013): Representing dictionaries in hypertextual form. In: Gouws, Rufus H.; Heid, Ulrich; Schweickard, Wolfgang; Wiegand, Herbert Ernst (Hrsg.): *Dictionaries. An International Encyclopedia of Lexicography*. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography. Berlin/New York: de Gruyter, 1244-1253.
- Tanacković, Sanjica Faletar; Krtalić, Maja; Lacović, Darko (2014): Newspapers as a Research Source: Information Needs and Information Seeking of Humanities Scholars. Digital Transformation and the Changing Role of News Media in the 21st Century. In: *IFLA 13-14 August 2014 in Genf*. (http://www.ifla.org/files/assets/newspapers/Geneva_2014/s6-lacovic-en.pdf [15.08.2018])

- Vachková, Marie; Cyril Belica (2009): Self-Organizing Lexical Feature Maps. Semiotic Interpretation and Possible Application in Lexicography. In: *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* (IJGLSA) 13/2, 223-260.
- Vachková, Marie (2010): Zur Erforschung und Erfassung der diskursgebundenen semantischen Kontraste auf der Grundlage des SOM-Modells. In: *Tématickéčíslo Germanistica Pragensia XX*. AUC Philologica, 2/2010. Prag: Charles University Karolinum Press, 193-208.
- Vachková, Marie (2018): Methoden zur korpuslinguistischen Erforschung von semantischen Relationen. In: Engelberg, Stefan; Kämper, Heidrun; Storjohann, Petra (Hrsg.): *Wortschatz: Theorie, Empirie, Dokumentation. Germanistische Sprachwissenschaft um 2020*. Band 3. Berlin/Boston: de Gruyter, 239-268.