

**BUCH, HÖRBUCH, E-BOOK & CO.  
LEXIKALISCHER WANDEL UND  
LEMMAUPDATE EINES DEUTSCHEN  
LERNERWÖRTERBUCHPORTALS**

*Zita Hollós*

Károli Gáspár Universität der Reformierten Kirche (Ungarn)

**„BUCH“, „HÖRBUCH“, „E-BOOK“ & CO.  
LEXICAL CHANGE AND LEMMA UPDATE OF A  
GERMAN LEARNER’S DICTIONARY PORTAL**

*Zita Hollós*

Károli Gáspár University of the Reformed Church (Hungary)

<https://doi.org/10.60055/GerSk.2025.izv.1.93-119>

*Abstract:* Im Zuge der Technisierung des Alltags und der Kommunikation treten allmählich traditionelle Medien wie Bücher, Briefe, Radio oder Fernsehen, um nur einige zu nennen, in den Hintergrund. Die Gegenstände unserer Kulturepoche, der des 20. Jahrhunderts, werden nach und nach durch immaterielle Phänomene verdrängt. Die Jugend benutzt diese und ebenfalls die sie bezeichnenden Wörter, wie z. B. *Hörbuch*, *Audiobook*, *E-Book* oder *E-Mail*, *Post*, *Like* sowie verschiedene Online- oder On-demand-Produkte tagtäglich und auch im Umgang mit Social Media. Was bedeutet das für die Pflege eines Lernerwörterbuchportals für Deutsch als Fremdsprache wie E-KOLLEX DAF? Welche Daten müssen als erstes aktualisiert werden, damit diese Ressource die Jugend weiterhin anspricht und *up to date* bleibt? Welche Wörterbuchportale und Tools bzw. Korpusstatistiken können den Lexikographen bei der Aktualisierung des Lemmabestandes, zunächst für den Themenkreis „Medien“, behilflich sein? Auf diese Fragen werden im folgenden Beitrag Antworten gesucht.

*Schlüsselwörter:* Neologismen, Medien, Lemmabestand, Wörterbuchportal, DaF

*Abstract:* In the course of the digitalisation of everyday life and communication, traditional media such as books, letters, radio and television, to name but a few, are gradually fading into the background. The objects of our cultural epoch, that of the 20th century, are gradually being replaced by immaterial phenomena. Young people use these and also the words that describe them, such as *Hörbuch*, *Audiobook*, *E-Book* or *E-Mail*, *Post*, *Like* and various online or on-demand products on a daily basis and also in dealing with social media. What does this mean for the maintenance of a learner dictionary portal for German as a foreign language such as E-KOLLEX DAF? Which data needs to be updated first so that this resource continues to appeal to young people and remains up to date? Which dictionary portals and tools or corpus statistics can help lexicographers to update the lemma inventory, initially for the “Media” topic area? Answers to these questions are sought in the following article.

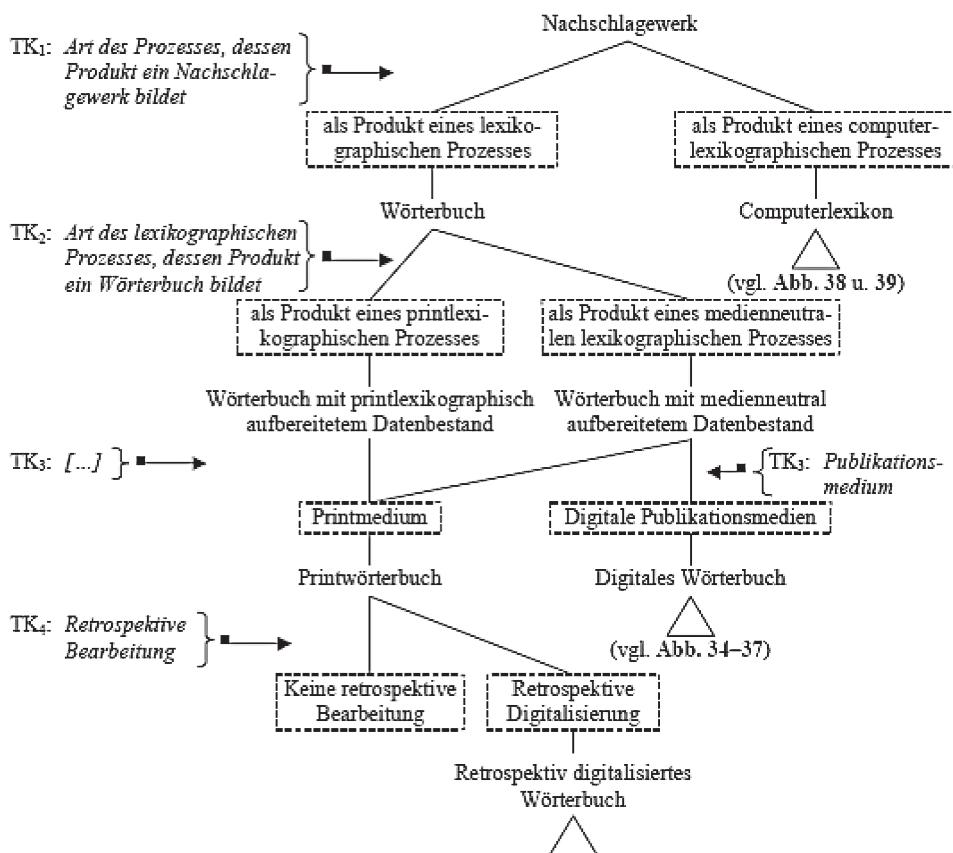
*Keywords:* neologisms, media, lemma collection, dictionary portal, GFL

## 1. Einleitung

Lexikographische Prozesse wurden bereits vor Jahrzehnten mit Computer-einsatz realisiert, dennoch ist die heutige Art und Weise des Einsatzes der Technologie kaum mit ihren Anfängen zu vergleichen. Während früher der Computer in erster Linie in der Phase der Datenaufbereitung und -auswertung, d.h. bei der Speicherung und Publikation der Daten zum Einsatz kam, erfährt heutzutage vor allem die Phase der Datenbeschaffung eine radikale Veränderung – eher eine Revolutionierung – dank der elektronischen Korpora und der statistischen Analysetools als Ergebnis der rasanten Entwicklung der Korpuslinguistik, der maschinellen Sprachverarbeitung sowie des Einsatzes von KI, nicht nur in der Lexikographie, sondern auch in allen benachbarten Disziplinen.

Der Prozess begann in der zweiten Hälfte des 20. Jahrhunderts mit dem Aufbau des elektronischen British National Corpus (BNC) im Dienste des Belegprinzips, gefolgt von der Analyse der Konkordanzen für die Gewinnung typischer Wortverbindungen und Kollokationen zu den Stichwörtern. Um die Jahrtausendwende erfolgte dann die Digitalisierung der Wörterbücher, zunächst als CD-ROM neben den Printpublikationen und bald auch als offline und online Wörterbücher. Am Anfang der Technisierung der Lexikographie entstanden viele retrospektive digitalisierte Wörterbücher, um printlexikographische Produkte für das neue Medium tauglich zu machen oder bereits im Interesse der Sprachdokumentation. Mithilfe des Aufbaus verschiedener Datenbanken mit z.T. nicht medienspezifisch aufbereiteten Datenbeständen konnten nun aber auch innovative zweisprachige, korpuslexikographische Projekte – wie das SZÓKAPTÁR/KOLLEX – gestartet werden, deren Endprodukte abhängig von dem Publikationsmedium entweder Printwörterbücher oder digitale Wörterbücher waren, ggf. beides, wie auch im Falle der KOLLEX-Datenbank.

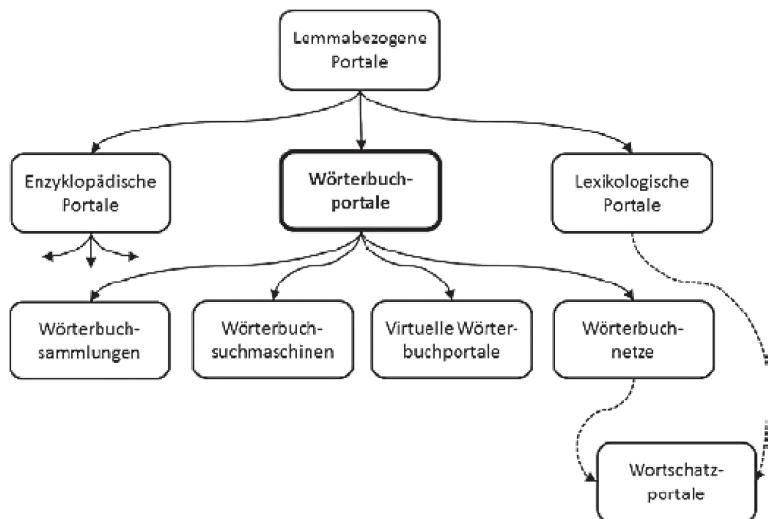
Heutzutage kann man – etwas vereinfacht betrachtet – zwei Zielsetzungen mit der Digitalisierung in der Lexikographie verfolgen: die Sprachdokumentation zur Bewahrung früherer, qualitativ hochwertiger printlexikographischer Produkte oder den Aufbau neuer lexikalischer Datenbanken im Rahmen eines medienneutral konzipierten lexikographischen Projekts (vgl. Abb. 1.), ggf. mit kreativen oder KI-gesteuerten technischen Lösungsvorschlägen (Einbindung verschiedener Statistiken und Korpora, Vernetzung der Ressourcen etc.) und mit innovativen Designelementen (multimodale Inhalte, neuartige Visualisierungen u.a.) für die Datenpräsentation.



**Abb. 1:** Typologieausschnitt für Nachschlagewerke (Wiegand et al. 2010, 89)

Mit der technischen Entwicklung der Vernetzungsmöglichkeiten der Ressourcen und mit dem Wunsch der Integration verschiedener lexikographischer

oder lexikologischer Produkte i.w.S. entstanden die ersten deutschen, lemmabbezogenen Portale (vgl. Abb. 2.), darunter das lexikologische Portal „Deutscher Wortschatz“ (DW) an der Universität Leipzig, das Wortschatzportal „Digitales Wörterbuch der Deutschen Sprache“ (DWDS) an der Berlin-Brandenburgischen Akademie der Wissenschaften oder das Wörterbuchnetz „Online Wortschatz-Informationssystem Deutsch“ (OWID) am IDS.



**Abb. 2:** Typologie von Wörterbuchportalen (Klosa/Müller-Spitzer 2016, 57)

Es ist leicht abzusehen, dass eine noch höhere Integration der lemmabbezogenen Portale, d.h. die Verbindung der Vorteile enzyklopädischer, lexikologischer Portale sowie die der Wörterbuchportale – evtl. mit der Einbindung anderer sprachdokumentarischen und/oder sprachlichen Ressourcen wie z. B. Konstrukta – immer mehr die Zukunft bestimmen werden.<sup>6</sup> Aus diesem Grund müssen auch für DaF-Lerner konzipierte online Nachschlagewerke wie das datenbankbasierte und korpusgestützte, deutsch-ungarische Lernerwörterbuchportal „Elektronisches KOLlokationsLEXikon Deutsch als Fremdsprache“ (E-KOLLEX DAF) stets nach innovativen Lösungen suchen.

Das Thema für den vorliegenden Beitrag, das Update des Lernerwörterbuchportals E-KOLLEX DAF, ist aus der Not heraus geboren. Die konkreten Daten und

<sup>6</sup> An ihrer Einbindung in sog. Schreib- und Übersetzungsassistenten wird z.Z. intensiv gearbeitet.

damit die Lemmakandidaten wurden zwischen 2005-2014 nach und nach in die Datenbank eingegeben, die dann 2014 in der ersten Printpublikation und dann ab 2016 als Wörterbuchportalartikel sukzessive online veröffentlicht wurden.<sup>7</sup> Bestimmte Daten sind somit fast 20 Jahre alt. Im Hinblick auf die äußere Selektion (Lemmabestand) kann dies heutzutage zu unangenehmen Lücken führen.<sup>8</sup> In der einschlägigen, neueren Fachliteratur zum Wortschatz z. B. am IDS (Engelberg/Lobin/Steyer/Wolfer (Hrsg.) 2018 u.a.) oder zur (digitalen) Lexikographie (Gouws et al. (Hrsg.) 2013 oder Jackson 2022 u.a.) findet sich kein Beitrag, der sich mit dem Updateen bereits existierender online Wörterbücher beschäftigt.<sup>9</sup> Zwar werden Internetressourcen für die Möglichkeit ihrer kontinuierlichen und zeitgleichen Aktualisierung angepriesen, es gibt aber außer dem DWDS- und OWID-Neo-Projekt m.W. kein Nachschlagewerk in der wissenschaftlichen Lexikographie, in dem diese Chance wirklich wahrgenommen würde. Diese Tat- sache ist kaum verwunderlich, da gerade Akademien, Sprachinstitute und vor allem Verlage wie Duden der Pflege ihrer Datenbestände der Nachschlagewerke in laufenden Projekten verpflichtet sind. Ein Projekt für eine zwei-, inzwischen z.T. einsprachige Onlineresource wie E-KOLLEX DAF, das auf der Basis der seit 2005 existierenden KOLLEX-Datenbank sukzessive in den letzten Jahren in einer universitären lexikographischen Werkstatt entstandenen ist und seit dem ersten Onlinerelease 2016 ständig weiterentwickelt wird, muss sich nun dieser Herausforderung ebenso stellen.

Bevor ich mich jedoch systematisch dem Thema der Lemmaselektion im Kapitel 3 und dem der Aktualisierung der Lemmaliste im Kapitel 4 widmen kann, müssen die Entstehungsvoraussetzungen der Datenbank sowie ihre jetzige Onlinepräsenz in Form eines frei zugänglichen Wörterbuchportals für Deutsch als Fremdsprache im nächsten Kapitel kurz vorgestellt werden.

## 2. Entstehung eines deutschen Lernerwörterbuchportals

Das erste Produkt des 2005 gestarteten, datenbankbasierten und korpusge- stützten Wörterbuchprojekts war ein 2014 in Ungarn publiziertes, zweisprachi-

<sup>7</sup> Die genauen Daten der Printpublikation sind auf Seite 12 der Open-access-Veröffentlichung auf dem IDS-Publikationsserver dokumentiert: [https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/12089/file/Gesamt\\_dt-ungar-WB\\_Hollos.pdf](https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/12089/file/Gesamt_dt-ungar-WB_Hollos.pdf)

<sup>8</sup> Für die Aktualisierung der einzelnen Sememe zu den Lexemen (als Stichwörter im Portal) ist diese Zeitspanne nicht so ausschlaggebend. Des Weiteren wurde die innere Selektion (Datensortiment der statistisch signifikanten deutschen Kollokationen) Jahre später gestartet. Diese Daten sind ebenfalls nicht so zeitsensitiv – wie die Sememe –, dennoch muss auch diese Selektion bald in einem nächsten Beitrag genauer unter die Lupe genommen werden.

<sup>9</sup> Mir ist lediglich ein Beitrag mit dem Thema zur Lemmaselektion für ein Neologismen- wörterbuch bekannt (Herberg 2002) und keiner zum Update des Lemmabestandes innerhalb eines zweisprachigen Wörterbuchportals.

ges, syntagmatisches Lernerwörterbuch für Deutschlerner der Mittel- und Oberstufe mit dem Titel „SZÓKAPTÁR: Német–magyar SZÓkapcsolatTÁR. KOLLEX. Deutsch-ungarisches Kollokationslexikon“, das insgesamt mehr als 61 000 Kollokationen, Wortverbindungen, Kombinationen mit Valenzangaben enthält, von denen mehr als 48 000 auch ins Ungarische übersetzt sind. Das im IDS-Verlag 2023 publizierte KOLLEX (Hollós 2023) verzeichnet somit ebenfalls ca. 61 000 Kollokationen/Wortverbindungen und mehr als 10 000 verschiedene Lexeme als Kollokatoren/Kotextpartner zu 2262 Lexemen als Lemmata. Nach der Umstellung des Redaktionssystems von einer Desktop-Applikation auf ein webbasiertes User-Interface im Jahre 2016, gefolgt von der Entwicklung eines Webdesigns für die Onlineversion, konnte ein Teil der Datenbank bereits in demselben Jahr unter der Internetadresse [www.kollex.hu/szotar](http://www.kollex.hu/szotar) der Öffentlichkeit kostenlos zugänglich gemacht werden.

Das Herzstück des „Deutschen Lernerwörterbuchportals“ bilden die 1274 substantivischen Wörterbuchartikel, in denen mehr als 30 000 typische deutsche Kollokationen/Wortkombinationen und ihre jeweiligen ungarischen Äquivalente enthalten sind. Die anhand des Leipziger Korpus *Deutscher Wortschatz*<sup>10</sup> ermittelten, statistisch signifikanten und manuell sememspezifisch ausgewählten Kollokationen und lernerrelevanten Wortkombinationen werden im Portal nach ihren didaktisch motivierten Strukturtypen unter sog. Strukturlinks geordnet. Dank des umfangreichen Datenangebots des Kollokationslexikons für Deutschlerner (KOLLEX 2023) können die verschiedenen Angabetypen aus der Datenbank Schritt für Schritt auch auf der Datenpräsentationsebene, also im Wörterbuchportal „Elektronisches Kollokationslexikon Deutsch als Fremdsprache“ (E-KOLLEX DAF) veröffentlicht werden.

Für das Lernerwörterbuchportal wurden mehrere Module geplant, drei davon sind bereits realisiert.<sup>11</sup> Die ersten zwei Module *KOLLEX* und *Schulwörterbuch (Suliszótár)* bilden das lexikographische Grundgerüst des Lernerwörterbuchportals, weil sie alle 1274 substantivischen Wörterbuchartikel aus dem Printwörterbuch SZÓKAPTÁR/KOLLEX (2014) und insgesamt 1280 printlayout-formatierte Wörterbuchartikel aus dem zweisprachigen Schulwörterbuch „Suliszótár“ (Hollós 2001) umfassen, das in den letzten Jahren sukzessive retrodigitalisiert wurde. Das *Schulwörterbuch-Modul*, das für die zweisprachigen Lernerwörterbuchartikel aus der Datenbank vorgesehen wurde, ist gleich in der Mitte des Portals unter dem Tab **Schulwörterbuch** zu finden (vgl. Abb. 3). Zu den Daten des *KOLLEX-Moduls* gehören der grün hinterlegte Formkommentar mit dem Lemma, die

<sup>10</sup> Das Korpus ist unter dem folgenden Link zu finden: <http://wortschatz.uni-leipzig.de/> [am 1.10.2023].

<sup>11</sup> Zu den geplanten und (teil)realisierten Modulen siehe Hollós 2022.

Sememe links neben dem Lernerwörterbuchartikel und die Kollokationen/Wortverbindungen unter den Strukturlinks **Lemma + SUBS**, **Lemma + ADJ**, **Lemma + VERB** und **KOMB** mit **Lemma**.

Das dritte Modul *Externe Ressourcen* impliziert die Integration lemmabegleitender Portale im Dienste neuer, fremdsprachendiffaktischer Angaben- bzw. Datentypen. Bei E-KOLLEX DAF sind diese das lexikologische Portal DW, darin der auf „Assoziogramm“ getaufte Datentyp und das Wortschatzportal DWDS, darunter die Daten mit den Namen „Belege“ und „Kookkurrenzen“. Diese spezifischen Datentypen werden im Portal als „Fenster im Fenster“ – in Form von unteren Tabs, von wo sie auch in einem neuen Fenster geöffnet werden können – organisch integriert. Den Eintrag zum Lemma **Buch**<sup>12</sup> im Portal gibt Abb. 3 wieder:

**Abb. 3:** Eintrag zum Lemma **Buch** im E-KOLLEX DAF, online unter: <http://kollex.hu/szotar/index.php/Welcome/search/Buch> (Zugriff am 13.06.2024)

E-KOLLEX DAF ist also kontinuierlich im Aufbau und enthält z.Z. (Mitte Juni 2024) folgende „Teilmengen“ an Daten:

- **Wörterbuchartikel mit Kollokationen/Wortverbindungen: 1274**
- zu allen Substantivlemmata aus KOLLEX

<sup>12</sup> Der dazugehörige Wörterbuchartikel aus dem Printwörterbuch ist als Open-Access-Publikation auf dem IDS-Publikationsserver unter [https://ids-pub.bsz-bw.de/frontdoor/de-liver/index/docId/12089/file/Gesamt\\_dt-ungar-WB\\_Hollos.pdf](https://ids-pub.bsz-bw.de/frontdoor/de-liver/index/docId/12089/file/Gesamt_dt-ungar-WB_Hollos.pdf) zu finden.

- **Wörterbuchartikel mit Äquivalenten: 988**  
zu allen Verb-, Adjektiv- und Adverblemmata aus KOLLEX
- **Korpusgestützt ermittelte, typische Kollokationen/ Wortverbindungen: ca. 30 000**  
zu allen Substantivlemmata aus KolleX
- **Bilder zu einzelnen Sememen: ca. 300**  
zu den Substantivlemmata in den ersten Lemmateilstrecken A-Z, außerdem vereinzelt in den Lemmastrecken A-Z zu anderen Lemmata
- **Dynamische Links zu Angaben im DWDS und DW: 6786**  
pro Lemma drei Links
- **Retrodigitalisierte Wörterbuchartikel aus Schulwörterbuch: 1280**  
605 Substantiv-, 366 Verb- und 239 Adjektivartikel
- **Links zum PONS-Bildwörterbuch, zu freedictionary u.a.**  
in der ersten Lemmateilstrecke A-Z in den Substantivartikeln

Das deutsche Lernerwörterbuchportal E-KOLLEX DAF kann man in gewissem Maße auch als eine Art Experimentierplattform betrachten, wo sich die klassischen lernerlexikographischen Methoden mit den Möglichkeiten der neuen Technologien treffen und die Synergieeffekte dieser Begegnung sukzessive auf der Datenpräsentationsebene sichtbar werden.

### 3. Lemmaselektion

Vor Beginn des Aufbaus der Datenbank musste nicht nur die Datenmodellierung und damit die Hierarchie der mehr als hundert verschiedenen Angabetypen bestimmt werden, sondern auch die sog. äußere Selektion, d.h. die Lemmaselektion für das syntagmatische Lernerwörterbuch erfolgen. Dies geschah parallel zur technischen Entwicklung der Datenbank und des dazugehörigen Wörterbucheditors mit Hilfe des Basiswörterbuchs von PONS, in dessen Lemmaliste der Grundwortschatz des sog. Zertifikat Deutsch<sup>13</sup> mit einem Symbol „●“ markiert war. Aus diesem Zertifikatswortschatz für die damalige Sprachprüfung für Deutsch als Fremdsprache wurden nur die Substantive, Verben, Adjektive und Adverbien, die als Basen einer Kollokation im hausmannschen Sinne fungieren konnten und sich für die Schüler als relevant erwiesen, ausgewählt. Die Gesamtlemmazahl betrug 2262, die – wenn man die vollständig erarbeiteten Wörterbuchartikel betrachtet –, auch für das heutige Wörterbuch-

<sup>13</sup> Dieses Zertifikat galt bis 2009 als international anerkannte Sprachprüfung, mit dem man sich für ein Studienkolleg in Deutschland bewerben konnte. Es orientierte sich an den Bedürfnissen der jungen erwachsenen Deutschlerner und an der Stufe B1 des *Gemeinsamen europäischen Referenzrahmens für Sprachen*.

portal zutrifft. Für Demonstrationszwecke gibt es zwar einige neue Lemmazeichen im Portal (z. B.: Schmetterling, Ostern oder E-Book), aber zu ihnen fehlt momentan die innere Selektion, d.h. die korpusgestützt ermittelten Kollokationen/Wortverbindungen, ganz zu schweigen von den dazugehörigen Übersetzungen, stilistisch-pragmatischen Markierungen und morphosyntaktischen Restriktionen, die zu allen Lemmata in der Datenbasis – als Grundlage für die Printpublikation – auch verzeichnet werden mussten. Inzwischen bedarf diese Lemmaauswahl wegen des lexikalischen Sprachwandels eines Updates in Bezug auf neue (evtl. veraltete) Lemmakandidaten. Aber wie? Eine mögliche Antwort liefert Androutsopoulos: „Die üblichen analytischen Zugänge zum Erneuerungsprozess sind die Untersuchungen der lexikographischen Praxis und die Untersuchung von ‘Wortkarrieren’ im öffentlichen Sprachgebrauch.“ (Androutsopoulos 2005, 175) Da dies für das Update des relativ großen KOLLEX-Lemmabestandes systematisch erfolgen müsste, werden nun verschiedene lexikographische, möglichst korpusgestützt erarbeitete, deutsche Print- und Onlineressourcen, bedingt auch vereinzelte „Wortkarrieren“ (v.a. jugendsprachlicher Ausdrücke in der öffentlichen Kommunikation) genauer untersucht.<sup>14</sup> Sie werden in den folgenden Kapiteln für die Gewinnung neuer Lemmakandidaten, die solche Neulexeme (selten Neusemmeme) darstellen, die zu ihrer Zeit als Neologismen<sup>15</sup> galten, heute jedoch u.U. keine mehr sind, vorgestellt, um einen Vorschlag für eine praxisnahe und leicht operationalisierbare Methode mit dem Ziel der Aktualisierung des Lemmabestandes von E-KOLLEX DAF im Kapitel 4 zu unterbreiten.

### 3.1. Quellen für Lemmakandidaten: Printressourcen

Damit die Lemmakandidatenliste eines Lernerwörterbuchs oder Lernerwörterbuchportals erstellt werden kann, müssen mögliche Sekundärquellen für die Selektion der möglichen Lemmata relativ zum Benutzerkreis und zum Wörterbuchgegenstand gefunden werden. Da es hier um eine bereits existierende Res-

<sup>14</sup> Androutsopoulos (2005) zitiert und beschreibt ein Stufenmodell der Verbreitung sprachlicher Innovationen als komplexe Diffusionsprozesse.

<sup>15</sup> Ich verwende den Terminus im Sinne des Neologismenwörterbuchs im OWID (OWID-Neo), da er in der Tradition der Neologismusforschung steht. Die dort praktizierte, korpuslinguistische Gewinnung und Beschreibung der Neologismen für das Print- und Onlinewörterbuch des IDS ist wegweisend. Demnach lautet die Definition eines Neologismus folgendermaßen: „Ein Neologismus ist eine lexikalische Einheit bzw. eine Bedeutung, die in einem bestimmten Abschnitt der Sprachentwicklung in einer Kommunikationsgemeinschaft aufkommt, sich ausbreitet, als sprachliche Norm allgemein akzeptiert und in diesem Entwicklungsabschnitt von der Mehrheit der Sprachbenutzer über eine gewisse Zeit hin als neu empfunden wird.“ (online unter: <https://www.owid.de/wb/neo/konzept/Konzept21.html>, Zugriff am 13.06.2024)

source geht, d.h. die Auswahl der Lemmakandidaten nicht von null angefangen wird, sind neue Korpustools oder -statistiken nicht zielführend und deshalb kann man zu diesem Zweck keine Primärquellen nutzen. Die Aufgabe ist also, angemessene gedruckte oder online Nachschlagewerke für die Aktualisierung der bereits existierenden Lemmaliste zu finden.

Es gibt alphabetische DaF-Wortschatzsammlungen wie die Wortlisten für das Goethe-Zertifikat A1, A2 oder B1, gemeinsam vom Goethe-Institut und vom ÖSD herausgegeben. Im Hinblick auf die äußere Selektion für den anvisierten Benutzerkreis des Lernerwörterbuchportals ist die Liste aus dem Goethe-Zertifikat B1 (2016) (im Weiteren: Wortliste B1) relevant.

Der dem Goethe-Zertifikat B1 zugrunde liegende Wortschatz umfasst circa 2400 lexikalische Einheiten, die Deutschlernende auf der Niveaustufe B1 kennen sollten. Er ist sowohl für die Jugendprüfung wie für die Erwachsenenprüfung anzuwenden. Welche Wörter und Ausdrücke in einem Prüfungssatz vorkommen, hängt u. a. von den Inhalten und Themen der jeweiligen Aufgaben ab. Es wird von den Prüfungsteilnehmenden erwartet, dass sie den hier vorliegenden Wortschatz zumindest rezeptiv beherrschen, d. h. dass diese beim Bearbeiten der geschriebenen Texte verstanden werden müssen. (Wortliste B1 2016, 5)

Diese Wortliste B1 ist auch im Internet zugänglich und enthält nicht nur eine alphabetische Auflistung der zertifikatsrelevanten Wörter, sondern auch nach formalen oder thematischen Prinzipien sortierte Wortlisten (z. B.: für *Abkürzungen* oder *Anglizismen*, bzw. für *Farben* oder *Himmelsrichtungen*), die in der alphabetischen Liste nicht enthalten sind und deshalb auch mitberücksichtigt werden sollten. Die Beschaffenheit dieses Zertifikatswortschatzes ist vor allem für die Überprüfung der in der Zwischenzeit durch die veränderte Wirklichkeit entstandenen Lücken des Grundwortschatzes im Portal geeignet. Vor allem die Liste der 73 Anglizismen (Wortliste B1 2016, 8f.), von denen 17 bereits im E-KOL-LEX DAF verzeichnet sind, ist eine wertvolle Quelle für die Behebung der wegen des Sprachwandels entstandenen Lemmalücken im Lernerwörterbuchportal und müsste deshalb so schnell wie möglich in den Lemmabestand des Portals integriert werden. Einige Beispiele für bereits im Portal verzeichnete Anglizismen vom Anfang des Wörterverzeichnisses sind *Baby*, *Bar*, *Cola*, *Computer*, *DVD*, *E-Book*, *Fax* oder *fit*. Die folgende Liste zeigt alle Anglizismen für Buchstabe C aus dem Goethe-Zertifikat B1<sup>16</sup>:

das Camp, -s  
campen, campt, campte, hat gecampt  
die Castingshow, -s  
der CD-Player, -

---

<sup>16</sup> Eine wichtige Lemmalücke stellt das Fehlen des Anglizismus *chillen* dar (vgl. Androutsopoulos 2005).

der Chat(room), -s  
chatten, chattet, chattete, hat gechattet  
checken, checkt, checkte, hat gecheckt  
der Chip, -s  
die City, -s  
der Club, -s  
die/das Cola, -s  
der/das Comic, -s  
der Computer, -  
cool

Bei der weiteren Überprüfung der Gesamtlemmaliste könnte eine moderne, thematische Wortschatzsammlung „Grund- und Aufbauwortschatz. Deutsch als Fremdsprache“ von Tschirner (2016) behilflich sein. Sie wurde nach Häufigkeit in Korpora und im Hinblick auf die DaF-Perspektive erstellt. Es ist eine nach Themen zusammengestellte und gleichzeitig in Grund- (A1-B1) und Aufbauwortschatz (B2) eingeteilte Sammlung, deren empirische Korpusgrundlage zwar bis 2006 zurückreicht, dennoch könnte sie bei der Überprüfung der Lemmaliste große Dienste leisten, gleichzeitig für die Ausweisung der Lemmazichen zu bestimmten Themen genutzt werden und damit zur thematisch orientierten Weiterentwicklung der E-KOLLEX-Ressource beitragen.

Abb. 4 zeigt die vier verschiedenen DaF-Listen, die bisher erwähnt wurden: die von Zertifikat Deutsch, Wortliste A2, Wortliste B1 und die von Tschirner A1-B2. Die Listen zeigen nur die Buchstabenstrecke C, aus denen die Substantive für diese Darstellung selektiert wurden. Der erste Grund hierfür liegt im Umfang des vorliegenden Beitrages, da im Kapitel 4, wo die Methode des Updatens vorgestellt und erprobt wird, die Selektionsmethode notwendigerweise auf die Substantive (und auf einen bestimmten Themenbereich) eingeschränkt werden muss. Den zweiten Grund liefert Eisenberg (2018, 28), der mit Blick auf die Hierarchie der Entlehnbarkeit<sup>17</sup> folgende Reihenfolge angibt: Substantiv > Adjektiv > Verb > Adverb > Präposition > andere Funktionswörter > Derivation > Flexion > Laut.<sup>18</sup> Damit korreliert auch die Behauptung von Eichinger (2007, 32), der basierend auf seinen früheren Forschungen betont, dass es bei Substantiven „um

<sup>17</sup> Bei der „Geburt“ der Neologismen ist neben der Wortbildung oft auch die Entlehnung in irgendeinem Maße beteiligt, oft gehen sie Hand in Hand, wie es bei Eichinger (2007, 32ff.) bei den Wortfamilienkernen an vielen Beispielen für Neologismen aus den 90er Jahren hinlänglich belegt wird.

<sup>18</sup> Diese Reihenfolge ist im Bereich der offenen Klassen umstritten und manchmal werden die Verben vor die Adjektive gestellt (vgl. ebenda). Im Falle der Neologismen – wenn bei ihnen (auch) die Entlehnung eine Rolle spielt – ist dieses Prinzip der Verben, aufgrund der Listen im OWID-Neo, absolut einleuchtend.

unvergleichlich höhere Fallzahlen geht, und um deutlich stärker ausgebauten morphologische Zusammenhänge“ (ebenda) als bei Verben oder Adjektiven.

In der Tabelle der Wortliste B1 (vgl. Abb. 4) wurden die Anglizismen in die alphabetische Lemmareihe mit hellgrünem Hintergrund integriert. Bei dieser Buchstabenstrecke ist sehr auffällig, dass viele Lemmata fremder (französischer oder englischer) Herkunft sind. Damit ist diese prozentuale Verteilung für andere Lemmastrecken nicht unbedingt ausschlaggebend.

**Zertifikat Deutsch  
(bis 2009)**

Café
Camion
Camping
CD
CD-ROM
Chance
Charakter
Chauffeur(in)
Chef(in)
Chemie
Club
Coiffeur, Coiffeuse
Cola
<b>Computer</b>
Couch
Couvert
Creme

**Goethe-Zertifikat A2, B1  
(2016)**

Café
<b>Cafeteria</b>
Chef(in)
Club
Comic
<b>Computer</b>
Creme
Chance
Chip
City
Club
Coiffeur, Coiffeuse
Cola
Comic
Computer
Couch
Creme
Cousin, Cousine
Couvert

**Tschirner A1-B2  
(2016)**

Café
CD
Celsius
Center
Champion
Chance
Chaos
Charakter
Chef
Chemie
Chip
Chor
Christ
Club
<b>Computer</b>

**Abb. 4:** Verschiedene DaF-Wortlisten für Buchstabe C (nur Substantive)

Wie aus den Tabellen leicht ersichtlich wird, gibt es gewisse Überlappungen bei den Listen, aber keine Liste (weder die von Goethe-Zertifikat noch die von Tschirner) kann allein für die Überprüfung oder Aktualisierung verwendet werden. In allen Listen konnten nur zwei gleiche Einträge gefunden werden, die fett hervorgehoben wurden, die Lemmata **Café** und **Computer**.

Die Abb. 5 zeigt die jetzige Lemmaliste für C im Lernerwörterbuchportal, die z.Z. (Mitte Juni 2024) zwölf Lemmata, ausschließlich Substantive, enthält.<sup>19</sup>

<sup>19</sup> Aus den siebzehn Wörtern der als Grundlage dienenden, damaligen ZD-Liste wurden in die Lemmastrecke C für das Printwörterbuch zwölf ausgewählt. Wegefallen sind Wörter wie z. B. *Camion* oder *Chauffeur*, weil sie für die Lebenswelt der Schüler wenig relevant sind.

Ergebnisse für "C" - 12
Café
Camping
CD
CD-ROM
Chance
Charakter
Chef
Chemie
Cola
Computer
Couch
Creme

**Abb. 5:** Lemmalisten für Buchstabe C im E-KOLLEX DAF; online unter: <http://kollex.hu/szotar/index.php>Welcome/search/C> (Zugriff am 13.06.2024)

Wenn man in der ersten Kontrollrunde des Updatens nur die grün hinterlegten Anglizismen der Wortliste B1 ins Portal aufnimmt – von den 10 nur 7, weil die anderen bereits vorhanden sind –, dann hat man schon wesentlich zur Aktualisierung des Lemmabestandes beigetragen, da solche Wörter wie *Casting*, *Chat*, *City* oder *Comic*<sup>20</sup> als neue Lemmata nun darin nachgeschlagen werden könnten. Dieser Vorgang könnte den ersten Schritt des Updatens der Lemmastrecken im E-KOLLEX DAF darstellen.

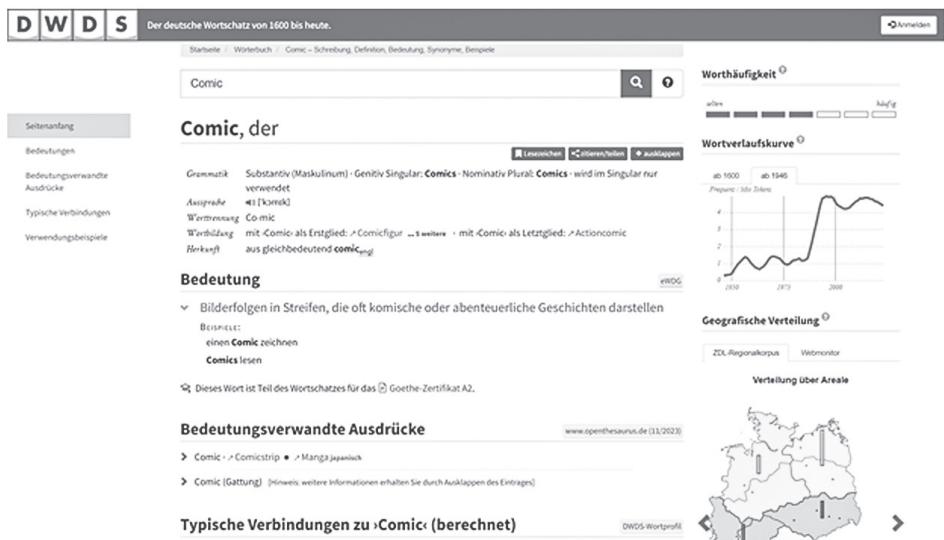
### 3.2. Quellen für Lemmakandidaten: Onlineressourcen

Generell kann man unter den Sekundärquellen für die Aktualisierung des Lernerwörterbuchportals zwischen deutschen Onlinewörterbüchern für DaF (wie PONS DaF) und deutschen Wörterbuchportalen (vgl. Abb. 2), wozu auch das Wortschatzportal DWDS zu rechnen ist, unterscheiden. Sie sind nur dann brauchbar, wenn sie die Markierung „neu“ oder „Neo“ für Neologismen vorsehen. Deutsche Lernerwörterbücher, egal ob print oder online, machen davon keinen Gebrauch.<sup>21</sup> DWDS und OWID können jedoch dafür hilfreich sein.

<sup>20</sup> In den Listen sind nur die Substantive verzeichnet, aber von den anderen Wortarten kämen noch solche Wörter wie *chatten*, *ckecken* oder *cool* hinzu.

<sup>21</sup> Der online Duden (duden.de) ist weder ein Lernerwörterbuch noch ist darin eine Markierung für Neologismen vorgesehen, deshalb ist diese Ressource für Zwecke des systematischen Updatens der Lemmaliste ungeeignet.

Bei Wörterbuchportalen fehlt meistens die didaktische Perspektive, wobei DWDS sukzessive die Markierung des Goethe-Zertifikatswortschatzes vorgenommen hat. Der Eintrag aus DWDS zum Lemma **Comic** (Abb. 6) zeigt, dass dieses Lemmazichen bereits als solches markiert wurde. Nach der Bedeutung, vor dem Teil „Bedeutungsverwandte Ausdrücke“, ist zu lesen: „Dieses Wort ist Teil des Wortschatzes für das Goethe-Zertifikat A2.“ (vgl. auch Abb. 4 mit den DaF-Wortlisten). Ein filterbasiertes Suchen nach Zertifikatswortschätzten ist jedoch z.Z. noch nicht gewährleistet.



**Abb. 6:** Anfang des Eintrages zum Lemma **Comic** im DWDS, online unter: <https://www.dwds.de/wb/Comic> (Zugriff am 13.06.2024)

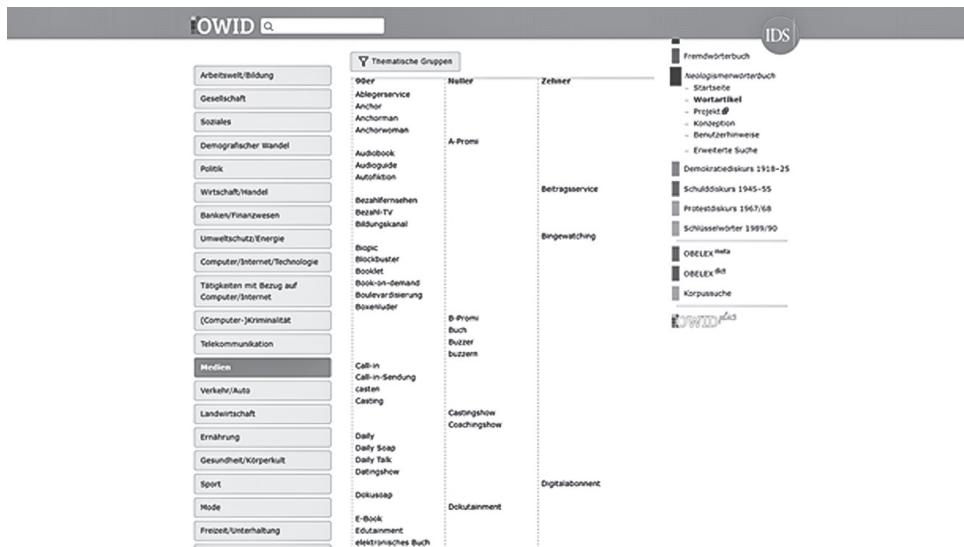
DWDS ermöglicht auch die gefilterte Suche im ganzen Wortschatzportal nach diasystematischen Angaben, darunter nach der Diamarkierung „Gruppensprache“ (z. B. Jugendsprache, z.Z. 59 Einträge), nach der Diamarkierung „Fachgebiet“ – z. B. Medien (z.Z. 118 Einträge, darunter: *E-Book, Streamer*), Internet (z.Z. 68 Einträge, darunter: *Like, Tweet, Follower, Surfer*) oder Informations- und Telekommunikationstechnik (z.Z. 552 Einträge, darunter *Account, Applikation*) – oder nach Diamarkierung „Gebrauchszeitraum“ (z. B. veraltend, z.Z. 2993 Einträge). Ein wichtiges Desiderat bildet bei der letzteren Filterfunktion die Angabe der Gebrauchszeit „Neo“, d.h. man kann im DWDS nicht nach Neologismen suchen. Der Grund ist einfach: während im WDG dieser Angabentyp in der Wörterbuchkonzeption vorgesehen war, ist er dies im DWDS nicht

mehr. Es ist im Portal umständlich herauszufinden, welche der gefilterten Wörter neu dazugekommen sind. Man kann zwar in der Lemmadatenbank des DWDS nach dem Artikeldatum der einzelnen Lemmata suchen, aber dieses Vorgehen ist ein Umweg und oft auch nicht ausschlaggebend.<sup>22</sup> Es wäre wünschenswert, wenn beide Daten gleichzeitig, ggf. mit einer kombinierten Filtersuche, recherchierbar wären. Dennoch kann man die Ergebnisse der gefilterten Suche nach Themengebieten wie z. B. „Medien“ oder „Internet“ in einer späteren Phase des Aktualisierungsprozesses zur Kontrolle heranziehen.

Unter den Onlineressourcen ist das Wörterbuchnetz „Online Wortschatz-Informationssystem Deutsch“ am IDS besonders wichtig. Es enthält mehrere wissenschaftliche Teilwörterbücher, darunter das Neologismenwörterbuch (im Weiteren: OWID-Neo). In dieser Ressource der Neologismen wird der gezielte Zugriff auf bestimmte Gruppen von Neologismen mit Hilfe der systematischen, zweifachen filterbasierten Suche ermöglicht: einerseits nach Themen (z. B. Computer/Internet/Technologie) und andererseits nach Neologismustypen (z. B. Neulexem mit Untertypen, Neubedeutung), nach Zeitspannen (z. B. Neologismen der Zehnerjahre), nach pragmatischen Merkmalen (z. B. Jugendsprache) oder nach der Herkunft (z. B. aus Englisch) u.a. Abb. 7 zeigt die Kombination der Filterung nach inhaltlich gruppierten Stichwörtern in 21 thematischen Gruppen<sup>23</sup> und nach Zeiträumen in drei Epochen (90er, Nuller- und Zehnerjahre):

<sup>22</sup> Ein jüngeres Datum muss allerdings nicht heißen, dass es sich um einen Neologismus handelt. Mit dem neuen Lemma kann durchaus eine frühere Lemmalücke (wie z. B. das Kompositum *Übungsheft* vom 17.11.2022) im Portal – hier als Minimalartikel realisiert – geschlossen werden.

<sup>23</sup> Eichinger (2007: 36) spricht in diesem Zusammenhang, noch vor dem Erscheinen von OWID-Neo, also kurz nach der Printpublikation des Neologismenwörterbuchs des IDS (Herberg/Kinne/Steffens 2004), über vier Bereiche, die „Modernisierungssphänomene“ (ebenda) beinhalten. Die ersten beiden enthalten Wortschatzelemente, die im Zuge der elektronischen Revolution in eher technischer und in eher gesellschaftlicher Sicht entstanden sind, während die anderen zwei Bereiche Erscheinungen der gesellschaftlichen Modernisierung auf globaler und nationaler Ebene umfassen.



**Abb. 7: Inhaltlich gruppierte Stichwörter aus dem Themenbereich „Medien“ im OWID-Neo (Zugriff am 13.06.2024)**

Sogar eine erweiterte, kombinierte Suche steht dem kundigen Benutzer im OWID-Neo zur Verfügung, womit man die Treffer noch spezifischer, z. B. nach Neologismustyp, Aufkommen (Epoche) und Herkunft (Sprache) usw. herausfiltern kann. Um die Nützlichkeit dieser erweiterten Suche für lernerlexikographische Zwecke zu demonstrieren, wurde eine Suche nach Neubedeutungen im OWID-Neo gestartet: sie ergab 201 Treffer gegenüber 2279 Treffern als Neulexeme. Wenn man die Neubedeutungen nach der Herkunft „Englisch“ weiter spezifiziert, bleiben nur noch fünf Sememe zu den Lexemen *Exit*, *Framing*, *ghosten*, *Pushback* und *Update*. Wenn man eine Suche nach Neulexemen startet, die pragmatisch als „Jugendsprache“ markiert sind, bekommt man nur zehn Treffer. Auf diese Weise kann man auch schwer auffindbare Neusemmeme oder Neologismen aus der Jugendsprache, die für die Benutzergruppe hochrelevant sind, kinderleicht entdecken und rasch in die Lemmaliste des Portals integrieren.

Leider ist die für die Aktualisierung nützlichste Kombination, die des Themas mit dem Aufkommen und/oder mit dem NeologismusTyp, im OWID-Neo nicht möglich, dennoch ist die eben vorgestellte Filtermethode (vgl. Abb. 7) der beste Ausgangspunkt für das systematische, themengesteuerte Updaten des Lemmabestandes im Portal E-KOLLEX DAF. Dies wird im nächsten Kapitel am Beispiel des Themenbereichs „Medien“ eingehend erläutert.

#### 4. Update des Lemmabestandes im Lernerwörterbuchportal

Bei der Aktualisierung der Lemmaliste von E-KOLLEX DAF müssen gemäß den früheren Ausführungen sowohl die B1-Lerner als Zielgruppe, als auch der Wörterbuchtyp „syntagmatisches Wörterbuch“ mit Fokus auf Kollokationen sowie die Wörterbuchfunktion „Sprachproduktion“ im Auge behalten werden. Daraus folgt, dass aus den im Kapitel 3 vorgestellten Print- und Onlineressourcen für die Selektion der neuen Lemmakandidaten eine solche Methodenkombination gewählt werden muss, die diesem Zweck am besten dient. Bestimmte Ressourcen (Wortlisten der Goethe-Zertifikate oder DWDS mit bestimmten Filtereinstellungen) wurden bereits aus guten Gründen als Kontrollinstanzen für die Selektion kategorisiert. Das OWID-Neo kann jedoch wegen seines Wörterbuchgegenstandes *Neologismen* und seiner komplexen Such- und Zugriffsmöglichkeiten sowie mit seinen erweiterten Filterfunktionen und deren Kombinierbarkeit als die beste Grundlage für die schwer bewältigbare Aktualisierung der Lemmaliste des Portals qualifiziert werden.

An dieser Stelle muss noch ein theoretisches Problem kurz angesprochen werden, nämlich der Sprachwandel bei Neologismen. Davon ist vor allem der Neologismenbestand der 90er Jahre, aber auch der der Zehnerjahre im OWID-Neo betroffen. Die für die 90er Jahre als Neologismen kategorisierten Lexeme sind heute entweder standardsprachliche Einheiten geworden (z. B. *anklicken*) oder aus dem Sprachgebrauch verschwunden (z. B. *Aldisierung*).<sup>24</sup> Man kann davon ausgehen, dass die bereits standardsprachlichen, auch für Lerner wichtigen Elemente des Wortschatzes bereits in den Listen der Deutschzertifikate enthalten sind, die als die wesentliche Grundlage für die erste Lemmaselektion gedient haben und auch als Kontrollinstanz für die Aktualisierung der Ressource dienen werden. Ebenfalls kann die Mitte und das Ende der Epoche der Zehnerjahre als problematisch betrachtet werden, da hier vermutlich noch viele Lexeme (z. B. *Faktenfinder*, *Abgassoftware*) vorzufinden sind, die mit großer Wahrscheinlichkeit nicht mehr lange im Sprachgebrauch erhalten bleiben. Aus diesem Grund sollte man abwarten, bis die konkreten relativen Häufigkeiten der letzten 20-30 Jahre für sie im OWID-Neo vorhanden sind. Daraus folgt, dass man sich beim Update des Lemmabestandes vor allem auf die zusammen 1008 Neologismen der Nullerjahre und auf die Neologismen in der Anfangsperiode der Zehnerjahre konzentrieren sollte.

<sup>24</sup> Ein Teil der Letzteren werden auch „historische Exotismen“ (Eichinger 2007: 39) genannt. Sie waren damals aktuell und werden heute nur dann gebraucht, wenn im Rückblick von diesen Zeiten und ihren Erscheinungen gesprochen wird.

#### 4.1. Thematisch gesteuerte Aktualisierung der Lemmaliste

Es spricht mehreres dafür, dass man im Falle eines Lernerwörterbuchportals in erster Linie thematisch vorgehen müsste. Erstens wird damit gewährleitet, dass die für die anvisierte Benutzergruppe relevantesten Themen zuerst herangezogen werden. Zweitens ist eine thematische Erweiterung sowohl logisch als auch methodisch besser begründbar als eine formale, nach dem Alphabet. Außerdem ist die thematische Steuerung des Wortschatz- und Kollokationserwerbs im Rahmen eines „Produktionswörterbuchs“ die einzige, didaktisch und psycholinguistisch vertretbare Methode.

Es ist unmöglich, im Rahmen eines kurzen Beitrages alle 21 Themenbereiche des OWID-Neo für das Testen einer praktikablen Methodenkombination zum Zweck der Aktualisierung des Lemmabestandes im E-KOLLEX DAF einzeln durchzuspielen, deshalb wurden als erster Schritt bei der Erarbeitung der Methodologie zwei Probeselektionen für die Themenbereiche „Computer/Internet/Technologie“ und „Medien“ durchgeführt.<sup>25</sup> Wörter aus dem ersten Bereich wie *Computer*, *E-Mail*, *Internet*, *Whatsapp*, *Post*, *Like* u.a. sowie Wörter aus dem zweiten wie *Audiobook*, *E-Book*, *Hörbuch*, *Influencer*, *Soap* u.a. sind in der heutigen durch und durch digitalisierten Welt – vor allem in der gesprochenen Sprache – unentbehrlich. Einige Beispiele für Neologismen der 90er Jahre wie *Computer*, *E-Mail* oder *Internet* wurden bereits im Lernerwörterbuchportal lemmatisiert, alle anderen – außer *E-Book* als Testlemma dieses Beitrags – noch nicht. Aus den beiden, für die obigen Themen erarbeiteten Lemmaselektions-tabellen mit insgesamt 435 Neulexemen wird die erstere, weniger umfangreiche Selektion für „Medien“ im vorliegenden Aufsatz zu Demonstrationszwecken einer möglichen thematisch orientierten Aktualisierung der Lemmaliste gewählt.

Die insgesamt 152 Neologismen-Lemmakandidaten (z. B. *Blockbuster*, *Pilot-*, *casten* oder *kultig*) sind Stichwörter aus der Sachgruppe „Medien“ von OWID-Neo. An diesem, relativ kleinen medienspezifischen Themenkreis kann man das gewählte Vorgehen testen, optimieren und letztendlich die Ergebnisse anschaulich präsentieren. Hier wird der Anfang der Neologismenstrecke – darunter auch **Audiobook** – aus der ersten Selektionstabelle wiedergegeben:<sup>26</sup>

<sup>25</sup> An dieser Stelle möchte ich mich bei den EMLex-Studentinnen Griselda Febrina Talitha und Aidana Koshkin für ihre Vorarbeiten bei der Lemmaselektion zum Zweck der Aktualisierung des Lernerwörterbuchportals im Rahmen ihres Praktikums an der KRE in Budapest bedanken.

<sup>26</sup> Vergleicht man Tab. 1 mit Tab. 2, also mit dem Ergebnis der Endselektion, sieht man, dass es kein einziger dieser ersten Lemmakandidaten, nicht einmal **Audiobook**, in die Endselektion geschafft hat. Der Grund für Letzteres liegt auf der Hand: *Hörbuch* hat sich auf Kosten von *Audiobook* in der Sprachverwendung durchgesetzt, dies bezeugen die Wortverlaufskurven im DWDS eindeutig.

Lemmakandidaten
Ablegerservice
Anchor
Anchorman
Anchorwoman
A-Promi
Audiobook
Audioguide
Autofiktion
Beitragsservice
Bezahlfernsehen
Bezahl-TV
Bildungskanal
Bingewatching
Biopic

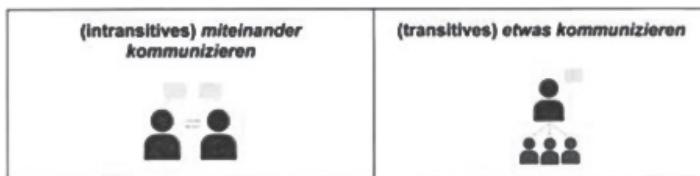
**Tab. 1:** Ausschnitt aus der potenziellen Lemmakandidatenliste anhand von OWID-Neo zu „Medien“

Auch für die Wahl dieser Themengruppe spricht, dass Social Media im Leben der Jugendlichen seit langem eine enorme Rolle spielt. Als Folge entstehen viele neuartige Ausdrucksmittel im Umgang mit den neuen Medien, aber auch neue Kommunikationspraktiken.

Neben den sprachlichen Besonderheiten der computervermittelten digitalen Kommunikation ist jedoch interessant, wie sich Kommunikationspraktiken selbst durch die spezifischen Kommunikationsbedingungen von Social Media und deren stetig steigender Bedeutung und Verankerung im Alltag verändert hat (sic). Grundannahme des vorliegenden Modells ist, dass sich hier ein Wandel von Kommunikation abzeichnet, der sich anhand verschiedener sprachlicher Neuschöpfungen und Neukonzeptualisierungen skizzieren lässt. Insbesondere auf dem unter Jugendlichen beliebtesten Social Media-Kanal, der Videoplattform YouTube, steht die one-to-many-Kommunikation im Vordergrund. Ziel ist es, mit dem Video möglichst viele Follower anzusprechen und Likes zu erzielen, um die Reichweite zu steigern und so das mediale Image zu profilieren. (Böhnert/Lemke 2018, 40)

Abb. 8 aus Böhnert/Lemke (2018, 45) zeigt am Beispiel der Valenz des Verbs *kommunizieren* das oben erwähnte Modell und den Unterschied zwischen zwei Formen der Kommunikation, die der traditionellen, zeitgleichen face-to-

face-Kommunikation und die der „one-to-many-Kommunikation“. Letztere ist im Zuge des medienbedingten Wandels auf Social Media-Plattformen zunehmend präsent.



**Abb. 8:** Zwei Formen der Kommunikation anhand der Sememe des Verbs *kommunizieren* (Böhnert/Lemke 2018, 45)

An dieser Stelle sollte ein letztes, variationslinguistisches Argument nicht unerwähnt bleiben. Androutsopoulos (2005) betont sehr früh die Rolle der Massenmedien im Sprachwandelprozess im Hinblick auf lexikalische Neuerungen. Medien sind einerseits „Quellen für Sprachmoden“ (Androutsopoulos 2005, 173) und andererseits „Indikatoren für die gesellschaftliche Verbreitung soziolektal markierten Wortschatzes“ (ebenda). Dies gilt in der Zeit von Social Media verstärkt für fast alle Altersgruppen.

#### 4.2. Vorschlag für eine Methodenkombination für das Update der Lemmaliste im E-KOLLEX DAF

Auf der Grundlage der bisherigen Ausführungen kann man also für die systematische Aktualisierung und Erweiterung der Lemmaliste aus einer bestimmten Teilmenge der thematischen Lemmabestände des OWID-Neo ausgehen. Die Gesamtstichwortliste im OWID-Neo ohne Phraseologismen, Strichlemmata oder versteckte neue Wörter – selbstverständlich ohne die Wörter unter Beobachtung – beträgt 2519 Einheiten. Davon sind 152 dem Themenkreis „Medien“ zugeordnet, das macht ca. 6% aus.

Beim ersten Schritt der Selektion musste geprüft werden, welche Einheiten als Neologismen der Zehnerjahre gelten, damit hier mehr auf den Anfang und die Mitte der Epoche geachtet werden kann.<sup>27</sup> Dann musste verzeichnet werden, welche Lexeme bereits im DWDS dokumentiert sind und wenn ja, wann sie aufgenommen wurden. Damit kann man recht „frische“ Neologismen besser entdecken und bei der Selektion gesondert behandeln. Von den 152 Neologismen sind zur Zeit des Selektionsprozesses (Januar 2024) 69 im DWDS als eigenständiges Stichwort noch nicht eingetragen. Bei ihnen ist Vorsicht geboten. Der Lernerperspektive wird

<sup>27</sup> Von insgesamt 606 Neologismen gehören 569 zum Anfang und zur Mitte der Epoche der Zehnerjahre.

dadurch Rechnung getragen, dass als nächstes die Wortlisten der Goethe-Zertifikate und die Stichwortliste des Lernerwörterbuchs von Langenscheidt (LGWDaF) in die Selektion mit einbezogen werden. Das Endergebnis der Selektion kann man Tabelle 2 entnehmen.

LEMMA-KANDIDATEN	OWID-NEO 10ER JAHRE	DWDS (JAHR)	GOETHE-ZERTIFIKAT (B2/C1)	LGWDaF
Casting	nein	ja (1999)	nein	ja
E-Book	nein	ja (2020)	ja	ja
Flyer	nein	ja (2021)	nein	nein
Hörbuch	nein	ja (2019)	nein	ja
Influencer	ja	ja (2018)	nein	nein
Realityshow	nein	ja (1999)	nein	nein
Sitcom	nein	ja (2015)	nein	nein
Soap	nein	ja (2015)	nein	nein
Spoiler	ja	ja (2021)	nein	nein
Vlog	nein	ja (2021)	nein	nein
Voting	nein	ja (2016)	nein	nein

Tab. 2: Endgültige Lemmakandidatenliste für „Medien“

Die OWID-Neo-Liste für „Medien“ wurde von 152 Einheiten zuerst auf 40 potenzielle Lemmakandidaten reduziert, aus denen zuletzt nur 11 Lemmata (darunter auch **Hörbuch** und **E-Book**) für das Update – gemäß der Befolgung der oben beschriebenen Selektionsprinzipien – ausgewählt wurden.<sup>28</sup>

Aus diesen Proportionen folgt, dass anhand der Stichwortliste des OWID-Neo um die 180 neu aufzunehmende Lemmata zu erwarten sind, ohne die Zunahme der bereits angesprochenen Kontrollgänge mit Hilfe anderer Ressourcen.<sup>29</sup> Somit ist insgesamt mit etwa 200 neuen Stichwörtern im Portal zu rechnen: dies ist fast um ein Zehntel mehr als der jetzige Lemmabestand.

<sup>28</sup> Von den 283 Neologismen zum Bereich „Computer/Internet/Technologie“ sind letztendlich 50 Kandidaten geblieben: davon sind 4 Wörter (*DVD*, *E-Mail*, *Internet*, *Mail*) bereits im Portal, 3 sind Neusememe (*Adresse 2*, *Seite 9*, *Wolke 2*) und 2 können modifiziert (als andere Wortart oder nur eine seiner Konstituenten) aufgenommen werden. Somit können 44 Wörter aus diesem Bereich als neue Lemmata beim Updaten der Stichwortliste ins Portal Eingang finden.

<sup>29</sup> Die früher erläuterten Kontrollen der upgedateten Lemmaliste des Portals tragen dazu bei, dass das Endergebnis der äußereren Gesamtselektion theoretisch und praktisch abgesichert wird. Diese Prüfprozesse können erst nach den thematisch gesteuerten Selektionsvorgängen erfolgen, da die Listen – mit einer Ausnahme (Tschirner 2016) – alphabetisch und nicht thematisch sortiert sind.

### 4.3. Zwischenbilanz

In den vorangehenden Kapiteln wurde dafür plädiert, dass man beim systematischen Update der Lemmaliste eines Lernerwörterbuchportals thematisch – mit Hilfe von OWID-Neo – vorgehen müsste. Damit wird nämlich gewährleistet, dass die für die spezifische Benutzergruppe jugendlicher Deutschlerner relevantesten Themen als erste herangezogen werden können. Man sollte sich beim Update des Lemmabestandes vor allem auf die Neologismen der Nullerjahre im OWID-Neo<sup>30</sup> konzentrieren. Die für die Aktualisierung nützlichste Filterkombination, die des Themas mit dem Aufkommen, ist leider im OWID-Neo nicht möglich, dennoch ist die vorgestellte Filtermethode nach inhaltlichen Gruppen (vgl. Abb. 7 und Tab. 1) und die im Kapitel 4.2 vorgeschlagene, komplexe, aber dennoch operationalisierbare Methodenkombination der wohl beste Weg für die systematische und themengesteuerte Aktualisierung der Lemmaliste im Portal E-KOLLEX DAF. Dies wurde am Beispiel des Themenbereichs „Medien“ und am Ergebnis der zugehörigen Selektion (vgl. Tab. 2) in diesem Beitrag demonstriert.

Wenn man in der ersten Runde des Updatens nur die Anglizismen der Wortliste B1 des Goethe-Zertifikats ins Portal aufnimmt, wofür vieles spricht, dann könnte man zur Aktualisierung des Lemmabestandes im Hinblick auf Fremdwörter englischer Herkunft auf die Schnelle wesentlich beitragen, da nun solche „neuen“ Wörter wie Chat, City, Comic oder chatten oder cool – wenn man nur Neuzugänge mit C erwähnt – darin auffindbar wären. Dieser Vorgang könnte also den ersten Schritt des formal und alphabetisch gesteuerten Updatens der Lemmastrecken im E-KOLLEX DAF darstellen. Die zweite und dritte Runde einer Aktualisierung des Lemmabestandes könnte dann die Aufnahme der bereits selektierten Lemmakandidaten beider Themenbereiche „Medien“ und „Computer/Internet/Technologie“ ins Portal darstellen. Diese Vorgänge wären die wichtigsten Schritte eines thematisch gesteuerten Updatens des Lemmabestandes. Mit diesen ersten drei Update-Runden kann man bereits augenfällige Ergebnisse für die Lemmastrecken von E-KOLLEX DAF erzielen.

<sup>30</sup> Auch die Neologismen aus den Zehnerjahren müssen unter Berücksichtigung von DWDS mit einbezogen werden, damit auch die neuesten Lexeme aus Kommunikationssituativen deutscher Jugendlicher in die Ressource Eingang finden können.

## 5. Integration eines neuen Lemmakandidaten ins Wörterbuchportal

Zum Abschluss muss noch die Integration frisch gewonnener Lemmakandidaten als neue Lemmazzeichen und somit als neuer Artikel im Wörterbuchportal E-KOLLEX DAF am Lemma **E-Book** exemplarisch vorgestellt werden. Sie erfolgt auf zwei Ebenen der Onlineressource: auf der Datenmodellierungsebene (vgl. Abb. 9) und auf der Datenpräsentationsebene (vgl. Abb. 10).

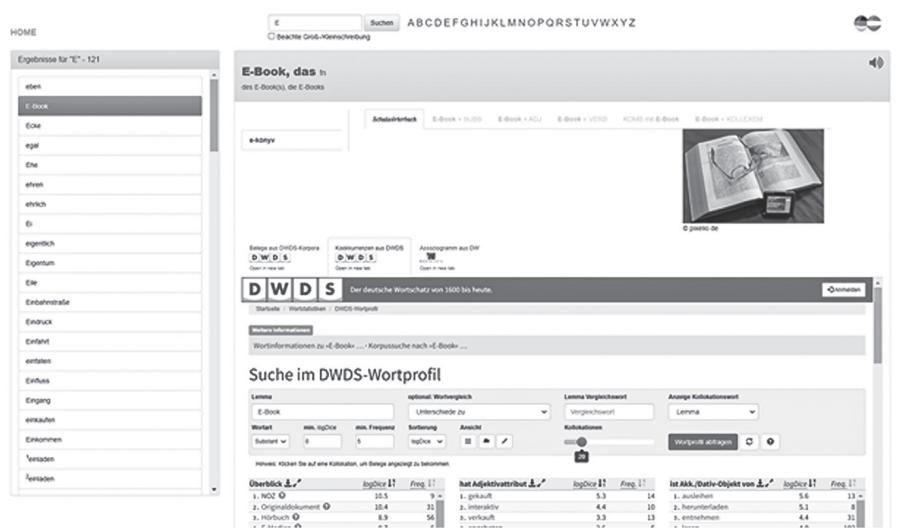
Das erste Printscreen-Bild (Abb. 9) zeigt den neu angelegten Wörterbuchartikel zu **E-Book** in der Datenbank, d.h. auf der Datenmodellierungsebene:

The screenshot shows the KOLLEX Editor interface with the following details:

- Header:** KOLLEX, Új szó, Szavak, keresés, Keresés.
- Breadcrumbs:** E-Book - HUNGARIAN [Lock WORD For Me]
- Navigation:** Structure, Uploads, Actions.
- Main Area:**
  - Campaign Skeleton:** Shows a tree structure under 'CAMPAGN - edit' with nodes like 'Cszövek - E-Book', 'Cmonon - 0', 'CLemma - E-Book', etc.
  - Position:** A dropdown menu with 'Value' selected.
  - Add elements:** A button to add new elements.
  - Előnézet:** A preview window showing the entry 'E-Book | E-Book' with the note 'des des E-Book(s) die E-Books | e-könyv Buch im digitalen Format (nukleus.nuvelingualisdictio\_4M\_E\_Book\_4)' and parts of speech: SURN, ADJ, VERB, KONJ, KOLLEREM.

**Abb. 9:** Neuer Eintrag in der Datenbank zum Lemma **E-Book** im nicht öffentlichen KOLLEX-Editor (Zugriff am 13.02.2024)

Das zweite Printscreen-Bild (Abb. 10) zeigt den neu angelegten Wörterbuchartikel zu **E-Book** im Portal, d.h. auf der Datenpräsentationsebene:



**Abb. 10:** Neuer Eintrag im Portal E-KOLLEX DAF zum Lemma **E-Book**, online unter:  
<http://kollex.hu/szotar/index.php/Welcome/search/E-Book> (Zugriff am 13.02.2024)

Die oberen Tabs, die sog. Inhaltslinks mit Angabe des Lemmas und die der Wortart des Kollokators, z. B. **E-Book** + ADJ (vgl. Abb. 10) sind inaktiv, deshalb grau auf dem Printscreenshot-Bild, da das Ergebnis der inneren Selektion, d.h. die Kollokationen, noch nicht in die Datenbank eingegeben wurden.

Der obige Wörterbuchartikel zum neuen Lemmazichen **E-Book** oder Einträge zu „alten“ Lemmata aus dem Themenbereich „Computer/Internet/Technologie“ wie zu **Computer**, **E-Mail** und **Internet** sowie alle 1274 substantivischen Wörterbuchartikel sind auf der Internetseite des Lernerwörterbuchportals unter [www.kollex.hu/szotar](http://www.kollex.hu/szotar) frei recherchierbar.

## 6. Resümee

Als Resümee soll hier ein Zitat aus dem Aufsatz von Eichinger (2007) mit dem Titel „So viele Wörter. Was neue Wörter über die deutsche Sprache aussagen“, den er kurz nach dem Erscheinen des Neologismenwörterbuchs am IDS (Herberg/Kinne/Steffens 2004) publiziert hat, wiedergegeben werden:

In der Moderne haben sich die Bedingungen von Alltag und Fachlichkeit in spezifischer Weise überlagert, die Sprache des öffentlichen Verkehrs, also der Verwaltung und ähnlicher Bereiche nimmt zu, greift aber ebenfalls weiter in den Alltag aus und die Medien in ihrer neuen Technik und Vielfalt bedienen nun jeweils eine unterschiedliche Kundschaft. Das Alles (sic) mündet in einem Bild erhöhter Va-

riation, das nur deshalb nicht als Wirrnis erscheinen sollte, weil es den Tatbestand reflektiert, dass sich unsere Gesellschaft nicht mehr nur in den traditionellen Kategorien von gesellschaftlicher Schichtung, regionaler Verteilung usw. beschreiben lässt, sondern darüber hinaus durch Gruppierungen, die ihr Einverständnis auch in Öffentlichkeit unter anderem auch dadurch symbolisieren, dass sie ihre eigenen sprachlichen Konventionen als Sozialsymbol in die Öffentlichkeit bringen. (Eichinger 2007, 42)

Vor diesem Hintergrund wird die Schwierigkeit des Aufbaus und des fortwährenden Updatens von E-KOLLEX DAF besonders gut ersichtlich. Ziel dieses Beitrages war gemäß dem obigen Zitat, diesem veränderten gesellschaftlichen und sprachlichen Umstand in einem deutschen Lernerwörterbuchportal – zunächst im Bezug auf die äußere Selektion – zumindest in Teilen Rechnung zu tragen, indem für die Aktualisierung der Lemmaliste eine operationalisierbare Methodenkombination vorgeschlagen und erprobt wurde. Sie ermöglicht, den Lemmabestand des Portals sukzessive so upzudaten, dass die sprachlichen Konventionen der deutschen Jugendlichen und jungen Erwachsenen als Sozialsymbol in einer solchen Form zugänglich gemacht werden, dass sich die Deutschlerner diese mithilfe der Ressource E-KOLLEX DAF benutzerfreundlich aneignen können.

#### LITERATURVERZEICHNIS/REFERENCES

- Androutsopoulos, Jannis. 2005. „... und jetzt gehe ich chillen: Jugend- und Szenesprachen als lexikalische Erneuerungsquellen des Standards.“ In *Standardvariation. Wie viel Variation verträgt die deutsche Sprache?* (171–206), hrsg. von L. M. Eichinger und W. Kallmeyer. Berlin/New York: de Gruyter.
- Böhnert, Katharina / Lemke, Ilka. 2018. „Haten, hashtaggen, posten, liken – Kommunikationspraktiken in Social Media und ihre sprachliche Realisierung als Gegenstand im Deutschunterricht.“ *Praxis Deutsch* 267: 40–45.
- Engelberg, Stefan / Lobin, Henning / Steyer, Kathrin / Wolfer, Sascha (Hrsg.). 2018. *Wortschätz. Dynamik, Muster, Komplexität*. Berlin/Boston: de Gruyter.
- Eichinger, Ludwig. 2007. „So viele Wörter. Was neue Wörter über die deutsche Sprache aussagen.“ *Studia Germanica Universitatis Vesprimiensis*, hrsg. von C. Földes 11 (1): 27–43.
- Eisenberg, Peter 2018. *Das Fremdwort im Deutschen*. Berlin/Boston: de Gruyter.
- Gouws, Rufus H. / Heid, Ulrich / Schweickard, Wolfgang / Wiegand, Herbert Ernst (eds.). 2013. *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin/Boston: de Gruyter.
- Herberg, Dieter. 2002. „Der lange Weg zur Stichwortliste. Aspekte der Stichwortselektion für ein allgemeinsprachliches Neologismenwörterbuch“. In *Ansichten der deutschen Sprache* (237–250), hrsg. von U. Haß-Zumkehr et al. Tübingen: Niemeyer.
- Herberg, Dieter / Kinne, Michael / Steffens, Doris. 2004. *Neuer Wortschatz. Neologismen der 90er Jahre im Deutschen*. Berlin/New York: de Gruyter.

- Hollós, Zita. 2001. *Német-magyar Suliszótár*. [Deutsch-ungarisches Schulwörterbuch.] Második, javított és bővített kiadás. Szeged: Grimm Kiadó.
- Hollós, Zita. 2022. „Cross-Media-Publishing in der korpusgestützten Lernerlexikographie. Entstehung eines Lernerwörterbuchportals DaF.“ In *Dictionaries and Society. Proceedings of the XX EURALEX International Congress* (436–447), hrsg. von A. Klosa-Kückelhaus et al. Mannheim: IDS-Verlag. [https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/11220/file/Dictionaries\\_and\\_Society\\_Proceedings\\_of\\_the\\_XXEuralex\\_2022.pdf](https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/11220/file/Dictionaries_and_Society_Proceedings_of_the_XXEuralex_2022.pdf) [Zugriff am 02.02.2024].
- Hollós, Zita. 2023. *KolleX: deutsch-ungarisches KOLLokationsLEXikon. Korpusbasier tes Kollokationswörterbuch Deutsch als Fremdsprache*. SZÓkapTÁR: Német–magyar SZÓkapcsolatTÁR. Korpuszalapú kollokációs tanulószótár. Mannheim: IDS-Verlag. [https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/12089/file/Gesamt\\_dt-ungar-WB\\_Hollos.pdf](https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/12089/file/Gesamt_dt-ungar-WB_Hollos.pdf) [Zugriff am 02.02.2024].
- Jackson, Howard 2022. *The Bloomsbury Handbook of Lexicography*. Second Edition. London: Bloomsbury Academic.
- Klosa, Annette / Müller-Spitzer, Carolin (Hrsg.). 2016. *Internetlexikografie. Ein Kompendium*. Berlin/Boston: de Gruyter.
- Tschirner, Erwin. 2016. *Grund- und Aufbauwortschatz. Deutsch als Fremdsprache nach Themen*. Berlin: Cornelsen.
- Wanzeck, Christiane. 2010. *Lexikologie. Beschreibung von Wort und Wortschatz im Deutschen*. Göttingen: Vandenhoeck & Ruprecht.
- Wiegand, Herbert Ernst / Beßwenger, Michael / Gouws, Rufus H. / Kammerer, Matthias / Storrer, Angelika / Wolski, Werner (Hrsg.). 2010. *Wörterbuch zur Lexikographie und Wörterbuchforschung. Dictionary of Lexicography and Dictionary Research*. Bd. 1. Berlin/New York: de Gruyter.

#### QUELLEN DER BEISPIELE/SOURCES OF EXAMPLES

- DW = *Deutscher Wortschatz*. Projekt der Universität Leipzig, der Sächsischen Akademie der Wissenschaften zu Leipzig und des Instituts für Angewandte Informatik. <https://wortschatz.uni-leipzig.de/de> [Zugriff am 02.02.2024].
- DWDS = *Digitales Wörterbuch der deutschen Sprache. Das Wortauskunftsyste m zur deutschen Sprache in Geschichte und Gegenwart*, hrsg. v. der Berlin-Brandenburgischen Akademie der Wissenschaften. <https://www.dwds.de> [Zugriff am 02.02.2024].
- Wortliste B1 = *Goethe-Zertifikat B1. Deutschprüfung für Jugendliche und Erwachsene*. 2016. [https://www.goethe.de/pro/relaunch/prf/de/Goethe-Zertifikat\\_B1\\_Wortliste.pdf](https://www.goethe.de/pro/relaunch/prf/de/Goethe-Zertifikat_B1_Wortliste.pdf) [Zugriff am 02.02.2024].
- E-KolleX DaF = *Elektronisches Kollokationslexikon Deutsch als Fremdsprache. Deutsches Lernerwörterbuchportal für die Sprachproduktion* 2016ff., hrsg. v. Zita Hollós. <http://kollex.hu/szotar> [Zugriff am 02.02.2024].
- OWID-Neo = Neologismenwörterbuch 2006ff. In *OWID – Online Wortschatz-Informati onssystem Deutsch*, hrsg. v. Leibniz-Institut für Deutsche Sprache, Mannheim. <https://www.owid.de/docs/neo/start.jsp> [Zugriff am 02.02.2024].

✉ Assoz. Prof. Dr. Zita Hollós, PhD, habil.  
ORCID iD: 0009-0002-4151-6858  
Institute of German and Netherlandic Studies  
Faculty of Humanities und Social Sciences  
Károli Gáspár University of the Reformed Church in Hungary  
Reviczky u. 4-6.  
1088 Budapest, HUNGARY  
E-mail: hollos.zita@kre.hu